HIV Sequence Compendium 2009

Editors

Carla Kuiken
Los Alamos National Laboratory

Thomas Leitner
Los Alamos National Laboratory

Brian Foley
Los Alamos National Laboratory

Beatrice Hahn
University of Alabama

Preston Marx
Tulane National Primate Research Center

Francince McCutchan
Henry M. Jackson Foundation

Steven Wolinsky
Northwestern University

Bette Korber
Los Alamos National Laboratory

Project Officer
Geetha Bansal
Division of AIDS
National Institute of Allergy and Infectious Diseases

Los Alamos HIV Sequence Database and Analysis Staff
Werner Abfalterer, Gayathri Athreya, Will Fischer, Bob Funkhouser, Chien-Chi Lo, Jennifer Macke, James J. Szinger, James Thurmond, Hyejin Yoon, Ming Zhang

This publication is funded by the Division of AIDS, National Institute of Allergy and Infectious Diseases, through interagency agreement IAA Y1-AI-8309-1 "HIV/SIV Database and Analysis Unit" with the U.S. Department of Energy.

Published by
Theoretical Biology and Biophysics
Group T-6, Mail Stop K710
Los Alamos National Laboratory
Los Alamos, New Mexico 87545 U.S.A.

LA-UR 09-03280

http://www.hiv.lanl.gov/

HIV Sequence Compendium 2009

Published by
Theoretical Biology and Biophysics
Group T-6, Mail Stop K710
Los Alamos National Laboratory
Los Alamos, New Mexico 87545 U.S.A

LA-UR 09-03280
Approved for public release; distribution is unlimited.

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by Los Alamos National Security, LLC, for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396.

This report was prepared as an account of work sponsored by an agency of the U.S. Government. Neither Los Alamos National Security, LLC, the U.S. Government nor any agency thereof, nor any of their employees make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by Los Alamos National Security, LLC, the U.S. Government, or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of Los Alamos National Security, LLC, the U.S. Government, or any agency thereof.

Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

This report was prepared as an account of work sponsored by NIH/NIAID/DAIDS under contract number IAA Y1-AI-8309-1 "HIV/SIV Database and Analysis Unit".

Contents

Contents **iii**

I Preface **1**
- I-1 Introduction 1
- I-2 Acknowledgements 1
- I-3 Citing the compendium or database 1
- I-4 About the PDF 1
- I-5 About the cover 2
- I-6 Genome maps 4
- I-7 HIV/SIV proteins 5
- I-8 Landmarks of the genome 6
- I-9 Amino acid codes 8
- I-10 Nucleic acid codes 8

II HIV-1/SIVcpz Complete Genomes **9**
- II-1 Introduction 9
- II-2 Annotated features 10
- II-3 Sequences 12
- II-4 Alignments 16

III HIV-2/SIV Complete Genomes **149**
- III-1 Introduction 149
- III-2 Annotated features 150
- III-3 Sequences 152
- III-4 Alignments 154

IV PLV Complete Genomes **219**
- IV-1 Introduction 219
- IV-2 Sequences 220
- IV-3 Alignments 222

V HIV-1/SIVcpz Proteins **301**
- V-1 Introduction 301
- V-2 Annotated features 302
- V-3 Sequences 304
- V-4 Alignments 310

VI HIV-2/SIV Proteins **365**
- VI-1 Introduction 365
- VI-2 Annotated features 366
- VI-3 Sequences 367
- VI-4 Alignments 369

VII PLV Proteins **395**
- VII-1 Introduction 395
- VII-2 Sequences 396
- VII-3 Alignments 403

I

Preface

I-1 Introduction

This compendium is an annual printed summary of the data contained in the HIV sequence database. In these compendia we try to present a judicious selection of the data in such a way that it is of maximum utility to HIV researchers. Each of the alignments attempts to display the genetic variability within the different species, groups and subtypes of the virus.

This compendium contains sequences published before January 1, 2009. Hence, though it is called the 2009 Compendium, its contents correspond to the 2008 curated alignments on our website.

The number of sequences in the HIV database is still increasing exponentially. In total, at the time of printing, there were 229,451 sequences in the HIV Sequence Database, an increase of 17% since last year.

The number of near complete genomes (>7000 nucleotides) increased to 2099 by end of 2008, reflecting a smaller increase than in previous years. However, as in previous years, the compendium alignments contain only a small fraction of these. Included in the alignments are a small number of sequences representing each of the subtypes and the more prevalent circulating recombinant forms (CRFs) such as 01 and 02, as well as a few outgroup sequences (group O and N and SIV-CPZ). Of the rarer CRFs we included one representative each. A more complete version of all alignments is available on our website, `http://www.hiv.lanl.gov/content/sequence/NEWALIGN/align.html`

Reprints are available from our website in the form of both HTML and PDF files. As always, we are open to complaints and suggestions for improvement. Inquiries and comments regarding the compendium should be addressed to seq-info@lanl.gov

I-2 Acknowledgements

The HIV Sequence Database and Analysis Project is funded by the Vaccine and Prevention Research Program of the AIDS Division of the National Institute of Allergy and Infectious Diseases (Dr. Geetha Bansal, Project Officer) through interagency agreement IAA Y1-AI-8309-1 “HIV/SIV Database and Analysis Unit” with the U.S. Department of Energy.

I-3 Citing the compendium or database

The LANL HIV Sequence Database may be cited in the same manner as this compendium:

HIV Sequence Compendium 2009. Carla Kuiken, Thomas Leitner, Brian Foley, Beatrice Hahn, Preston Marx, Francince McCutchan, Steven Wolinsky, and Bette Korber editors. 2009. Publisher: Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, New Mexico. LA-UR 09-03280.

I-4 About the PDF

The complete *HIV Sequence Compendium 2009* is available in Adobe Portable Document Format (PDF) from our website, `http://www.hiv.lanl.gov/`. The PDF version is hypertext enabled and features ‘clickable’ table-of-contents, indexes, references and links to external web sites.

This volume is typeset using LaTeX.

I-5 About the cover

The cover of this compendium depicts a phylogenetic analysis method to classify sequences into known subtype clades with the Branching Index (BI) (Figure I.1). The purpose of the experiment was to examine variation in BI classification accuracy across known subtypes and over the HIV-1 genome, and thereby to establish threshold BI values for inferring whether a sequence belongs to a known subtype clade. The *x*-axis of each panel depicts position along the HIV-1 genome. The *y*-axis is the BI, which varies from 0 (complete divergence from other members of the subtype clade) to 1 (no divergence from a known subtype clade, relative to other clades). Each panel depicts a known subtype and results from computing the BI for many sequences of that subtype. Thus, the figure depicts a quantification of BI classification accuracy for HIV-1 subtypes. This analysis complements results from the LANL HIV Database PhyloPlace tool by establishing BI error rates and confidence levels for subtype inference.

We sampled 10,000 random-length fragments over the 2005 HIV-1 M-group subtype reference alignment, and randomly chose a query sequence for each. We computed the BI twice for each fragment sampled, once with all sequences included (thus generating the data points used to draw the upper curve, colored by subtype and represented by + symbols) and once with a subset of sequences that excluded the subtype of the query sequence (generating the data points used to draw the lower curve, colored by grayscale and represented by ∘ symbols). This second scenario emulates conditions where the query sequence is from an unknown subtype. The figure depicts the BI from each fragment as a pair of points. Both the midpoint (+ or ∘) and the extent of each fragment are shown. Vertical lines connect paired observations when one is misclassified. We fit smooth curves to the resulting data with loess locally weighted regression. Horizontal lines depict the optimal thresholds for inferring that a sequence is of known subtype with greatest accuracy, either among all subtypes (black) or per subtype (gray). Sequences yielding BI values above 0.66 are significantly associated with the subtype, with 93.5% confidence. For more information about BI, see Wilbe *et al.* 2003 (*Virology* **316**:116–125). For additional information about the generation of these curves and their use in classifying HIV-1 sequences, see Hraber *et al.* 2008 (*J. Gen. Virol.* **89**:2098–2107).

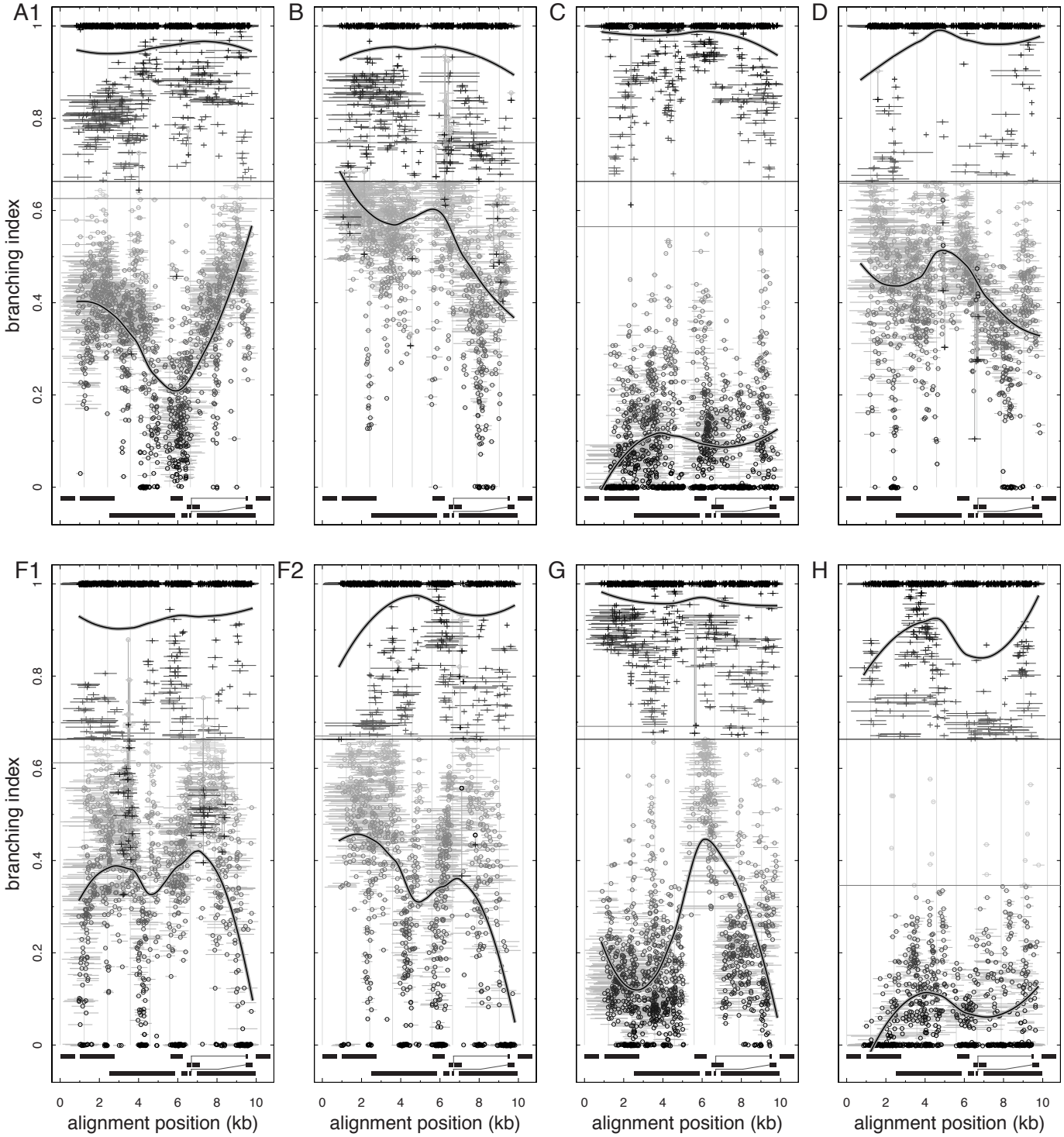
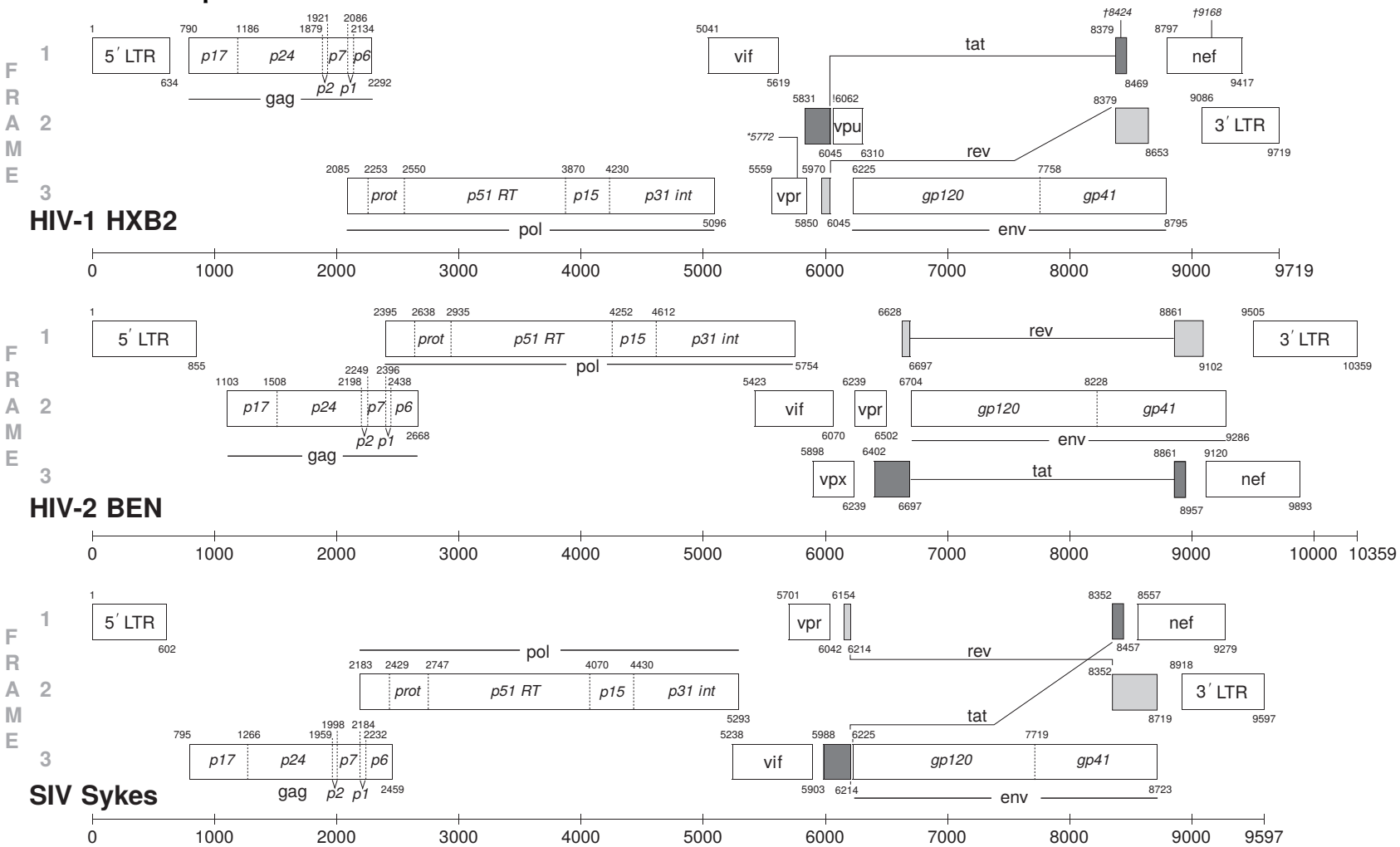


Figure I.1: The branching index classifies HIV-1 subtypes.

I-6 Genome maps



Landmarks of the HIV-1, HIV-2, and SIV genomes. Open reading frames are shown as rectangles. The gene start, indicated by the small number in the upper left corner of each rectangle normally records the position of the `a` in the `atg` start codon for that gene, while the number in the lower right records the last position of the stop codon. For *pol*, the start is taken to be the first `t` in the sequence `ttttttag`, which forms part of the stem loop that potentiates ribosomal slippage on the RNA and a resulting −1 frameshift and the translation of the Gag-Pol polyprotein. The *tat* and *rev* spliced exons are shown as shaded rectangles. In HXB2, *5772 marks the position of a frameshift in the *vpr* gene caused by an "extra" `t` relative to most other subtype B viruses; !6062 indicates a defective `acg` start codon in *vpu*; †8424 and †9168 mark premature stop codons in *tat* and *nef*. See Korber *et al.*, Numbering Positions in HIV Relative to HXB2CG, in *Human Retroviruses and AIDS*, 1998, p. 102. Available from `http://www.hiv.lanl.gov/content/sequence/HIV/REVIEWS/HXB2.html`

I-7 HIV/SIV proteins

Name	Size	Function	Localization
Gag			
MA	p17	membrane anchoring; env interaction; nuclear transport of viral core (myristylated protein)	virion
CA	p24	core capsid	virion
NC	p7	nucleocapsid, binds RNA	virion
	p6	binds Vpr	virion
Pol			
Protease (PR)	p15	Gag/Pol cleavage and maturation	virion
Reverse Transcriptase (RT)	p66, p51	reverse transcription, RNAse H activity	virion
RNase H	p15		virion
Integrase (IN)	p31	DNA provirus integration	virion
Env	gp120/gp41	external viral glycoproteins bind to CD4 and secondary receptors	plasma membrane, virion envelope
Tat	p16/p14	viral transcriptional transactivator	primarily in nucleolus/nucleus
Rev	p19	RNA transport, stability and utilization factor (phosphoprotein)	primarily in nuleolus/nucleus shuttling between nucleolus and cytoplasm
Vif	p23	promotes virion maturation and infectivity	cytoplasm (cytosol, membranes), virion
Vpr	p10-15	promotes nuclear localization of preintegration complex, inhibits cell division, arrests infected cells at G2/M	virion nucleus (nuclear membrane?)
Vpu	p16	promotes extracellular release of viral particles; degrades CD4 in the ER; (phosphoprotein only in HIV-1 and SIVcpz)	integral membrane protein
Nef	p27-p25	CD4 and class I downregulation (myristylated protein)	plasma membrane, cytoplasm, (virion?)
Vpx	p12-16	Vpr homolog present in HIV-2 and some SIVs, absent in HIV-1	virion (nucleus?)
Tev	p28	tripartite tat-env-rev protein (also named Tnv)	primarily in nucleolus/nucleus

I-8 Landmarks of the genome

HIV genomic structural elements

LTR Long terminal repeat, the DNA sequence flanking the genome of integrated proviruses. It contains important regulatory regions, especially those for transcription initiation and polyadenylation.

TAR Target sequence for viral transactivation, the binding site for Tat protein and for cellular proteins; consists of approximately the first 45 nucleotides of the viral mRNAs in HIV-1 (or the first 100 nucleotides in HIV-2 and SIV.) TAR RNA forms a hairpin stem-loop structure with a side bulge; the bulge is necessary for Tat binding and function.

RRE Rev responsive element, an RNA element encoded within the env region of HIV-1. It consists of approximately 200 nucleotides (positions 7327 to 7530 from the start of transcription in HIV-1, spanning the border of gp120 and gp41). The RRE is necessary for Rev function; it contains a high affinity site for Rev; in all, approximately seven binding sites for Rev exist within the RRE RNA. Other lentiviruses (HIV-2, SIV, visna, CAEV) have similar RRE elements in similar locations within env, while HTLVs have an analogous RNA element (RXRE) serving the same purpose within their LTR; RRE is the binding site for Rev protein, while RXRE is the binding site for Rex protein. RRE (and RXRE) form complex secondary structures, necessary for specific protein binding.

PE Psi elements, a set of 4 stem-loop structures preceding and overlapping the Gag start codon which are the sites recognized by the cysteine histidine box, a conserved motif with the canonical sequence CysX2CysX4HisX4Cys, present in the Gag p7 MC protein. The Psi Elements are present in unspliced genomic transcripts but absent from spliced viral mRNAs.

SLIP An TTTTTT slippery site, followed by a stem-loop structure, is responsible for regulating the -1 ribosomal frameshift out of the Gag reading frame into the Pol reading frame.

CRS Cis-acting repressive sequences postulated to inhibit structural protein expression in the absence of Rev. One such site was mapped within the pol region of HIV-1. The exact function has not been defined; splice sites have been postulated to act as CRS sequences.

INS Inhibitory/Instability RNA sequences found within the structural genes of HIV-1 and of other complex retroviruses. Multiple INS elements exist within the genome and can act independently; one of the best characterized elements spans nucleotides 414 to 631 in the gag region of HIV-1. The INS elements have been defined by functional assays as elements that inhibit expression posttranscriptionally. Mutation of the RNA elements was shown to lead to INS inactivation and up regulation of gene expression.

Genes and gene products

GAG The genomic region encoding the capsid proteins (group specific antigens). The precursor is the p55 myristylated protein, which is processed to p17 (MAtrix), p24 (CApsid), p7 (NucleoCapsid), and p6 proteins, by the viral protease. Gag associates with the plasma membrane where the virus assembly takes place. The 55 kDa Gag precursor is called assemblin to indicate its role in viral assembly.

POL The genomic region encoding the viral enzymes protease, reverse transcriptase, RNAse, and integrase. These enzymes are produced as a Gag-Pol precursor polyprotein, which is processed by the viral protease; the Gag-Pol precursor is produced by ribosome frameshifting near the 3′end of gag.

ENV Viral glycoproteins produced as a precursor (gp160) which is processed to give a noncovalent complex of the external glycoprotein gp120 and the transmembrane glycoprotein gp41. The mature gp120-gp41 proteins are bound by non-covalent interactions and are associated as a trimer on the cell surface. A substantial amount of gp120 can be found released in the medium. gp120 contains the binding site for the CD4 receptor, and the seven transmembrane domain chemokine receptors that serve as co-receptors for HIV-1.

TAT Transactivator of HIV gene expression. One of two essential viral regulatory factors (Tat and Rev) for HIV gene expression. Two forms are known, Tat-1 exon (minor form) of 72 amino acids and Tat-2 exon (major form) of 86 amino acids. Low levels of both proteins are found in persistently infected cells. Tat has been localized primarily in the nucleolus/nucleus by immunofluorescence. It acts by binding to the TAR RNA element and activating transcription initiation and elongation from the LTR promoter, preventing the 5′LTR AATAAA polyadenylation signal from causing premature termination of transcription and polyadenylation. It is the first eukaryotic transcription factor known to interact with RNA rather than DNA and may have similarities with prokaryotic anti-termination factors. Extracellular Tat can be found and can be taken up by cells in culture.

REV The second necessary regulatory factor for HIV expression. A 19 kDa phosphoprotein, localized primarily in the nucleolus/nucleus, Rev acts by binding to RRE and promoting the nuclear export, stabilization and utilization of the unspliced viral mRNAs containing RRE. Rev is considered the most functionally conserved regulatory protein of lentiviruses. Rev cycles rapidly between the nucleus and the cytoplasm.

VIF Viral infectivity factor, a basic protein of typically 23 kDa. Promotes the infectivity but not the production of viral particles. In the absence of Vif the produced viral particles are defective, while the cell-to-cell transmission of virus is not affected significantly. Found in almost all lentiviruses, Vif is a cytoplasmic protein, existing in both a soluble cytosolic form and a membrane-associated form. The latter form of Vif is a peripheral membrane protein that is tightly associated with the cytoplasmic side of cellular membranes. In 2003, it was discovered that Vif prevents the action of the cellular APOBEC-3G protein which deaminates DNA:RNA heteroduplexes in the cytoplasm.

VPR Vpr (viral protein R) is a 96-amino acid (14 kDa) protein, which is incorporated into the virion. It interacts with the p6 Gag part of the Pr55 Gag precursor. Vpr detected in the cell is localized to the nucleus. Proposed functions for Vpr include the targeting the nuclear import of preintegration complexes, cell growth arrest, transactivation of cellular genes, and induction of cellular differentiation. In HIV-2, SIV-SMM, SIV-RCM, SIV-MND-2 and SIV-DRL the Vpx gene is apparently the result of a Vpr gene duplication event, possibly by recombination.

VPU Vpu (viral protein U) is unique to HIV-1, SIVcpz (the closest SIV relative of HIV-1), SIV-GSN, SIV-MUS, SIV-MON and SIV-DEN. There is no similar gene in HIV-2, SIV-SMM or other SIVs. Vpu is a 16 kDa (81-amino acid) type I integral membrane protein with at least two different biological functions: (a) degradation of CD4 in the endoplasmic reticulum, and (b) enhancement of virion release from the plasma membrane of HIV-1-infected cells. Env and Vpu are expressed from a bicistronic mRNA. Vpu probably possesses an N-terminal hydrophobic membrane anchor and a hydrophilic moiety. It is phosphorylated by casein kinase II at positions Ser52 and Ser56. Vpu is involved in Env maturation and is not found in the virion. Vpu has been found to increase susceptibility of HIV-1 infected cells to Fas killing.

NEF A multifunctional 27-kDa myristylated protein produced by an ORF located at the 3′end of the primate lentiviruses. Other forms of Nef are known, including nonmyristylated variants. Nef is predominantly cytoplasmic and associated with the plasma membrane via the myristyl residue linked to the conserved second amino acid (Gly). Nef has also been identified in the nucleus and found associated with the cytoskeleton in some experiments. One of the first HIV proteins to be produced in infected cells, it is the most immunogenic of the accessory proteins. The nef genes of HIV and SIV are dispensable *in vitro*, but are essential for efficient viral spread and disease progression *in vivo*. Nef is necessary for the maintenance of high virus loads and for the development of AIDS in macaques, and viruses with defective Nef have been detected in some HIV-1 infected long term survivors. Nef downregulates CD4, the primary viral receptor, and MHC class I molecules, and these functions map to different parts of the protein. Nef interacts with components of host cell signal transduction and clathrin-dependent protein sorting pathways. It increases viral infectivity. Nef contains PxxP motifs that bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of HIV but not for the downregulation of CD4.

VPX A virion protein of 12 kDa found in HIV-2, SIV-SMM, SIV-RCM, SIV-MND-2 and SIV-DRL and not in HIV-1 or other SIVs. This accessory gene is a homolog of HIV-1 vpr, and viruses with Vpx carry both vpr and vpx. Vpx function in relation to Vpr is not fully elucidated; both are incorporated into virions at levels comparable to Gag proteins through interactions with Gag p6. Vpx is necessary for efficient replication of SIV-SMM in PBMCs. Progression to AIDS and death in SIV-infected animals can occur in the absence of Vpr or Vpx. Double mutant virus lacking both vpr and vpx was attenuated, whereas the single mutants were not, suggesting a redundancy in the function of Vpr and Vpx related to virus pathogenicity.

Structural proteins/viral enzymes The products of *gag*, *pol*, and *env* genes, which are essential components of the retroviral particle.

Regulatory proteins Tat and Rev proteins of HIV/SIV and Tax and Rex proteins of HTLVs. They modulate transcriptional and posttranscriptional steps of virus gene expression and are essential for virus propagation.

Accessory or auxiliary proteins Additional virion and non-virion-associated proteins produced by HIV/SIV retroviruses: Vif, Vpr, Vpu, Vpx, Nef. Although the accessory proteins are in general not necessary for viral propagation in tissue culture, they have been conserved in the different isolates; this conservation and experimental observations suggest that their role *in vivo* is very important. Their functional importance continues to be elucidated.

Complex retroviruses Retroviruses regulating their expression via viral factors and expressing additional proteins (regulatory and accessory) essential for their life cycle.

I-9 Amino acid codes

A	Alanine
B	Aspartic Acid or Asparagine
C	Cysteine
D	Aspartic Acid
E	Glutamic Acid
F	Phenylalanine
G	Glycine
H	Histidine
I	Isoleucine
K	Lysine
L	Leucine
M	Methionine
N	Asparagine
P	Proline
Q	Glutamine
R	Arginine
S	Serine
T	Threonine
V	Valine
W	Tryptophan
X	unknown or “other” amino acid
Y	Tyrosine
Z	Glutamic Acid or Glutamine
.	gap
-	identity
*	stop codon
#	incomplete codon

I-10 Nucleic acid codes

A	Adenine
C	Cytosine
G	Guanine
T	Thymine
U	Uracil
M	A or C
R	A or G
W	A or T
S	C or G
Y	C or T
K	G or T
V	A or C or G
H	A or C or T
D	A or G or T
B	C or G or T
N	unknown
.	gap
-	identity

II

HIV-1/SIVcpz Complete Genomes

Contents

II-1	Introduction	**9**
II-2	Annotated features	**10**
II-3	Sequences	**12**
II-4	Alignments	**16**

II-1 Introduction

The goal in selecting the sequences to present in the published compendium alignment is to display a set of sequences that represent the genetic variation of HIV-1 worldwide. Selections are made based on the list of identified subtypes and CRFs together with information from phylogenetic trees. Some effort was also put into representing a diversity of geographic locations and sampling times for each subtype, with an emphasis on more recent samples. We include more sequences from the more prevalent subtypes, only one representative for each CRF (except for the widespread CRFs 01 and 02), and group N, group O, and CPZ sequences. Within the limited space of the compendium, this selection is intended to show the worldwide variation of HIV-1 subtypes and CRFs. For URFs, of which there is a growing number, this alignment contains none. The web-based version of this alignment contains all URFs and many additional representatives of most subtypes and CRFs.

As always, the HXB2 sequence is the master sequence in this alignment. This is also the HIV Database genome coordinate standard sequence. The alignment was generated by an iterative process between automated alignment using HMMER and manual editing using BioEdit and Se-Al. As in previous years, the alignment presented is not suggested to be optimal, with the absolute minimum number of gaps and mismatches. It is a compromise between optimal alignment, readability, and an attempt to keep insertions and deletions from altering the protein reading frame presentation. Most gaps have been introduced in multiples of 3 bases to maintain open reading frames when translated directly from the alignment.

At the bottom of the alignment, the translated HXB2 protein sequence is shown. The HIV genome has many overlapping coding regions, and all are shown. For more complete annotation of functional domains, see the protein sequence alignments in Chapter V.

II-2 Annotated features

Features of HIV-1 annotated in the alignment that follows.

Feature	Location	Page
5' LTR U3 start	1	16
TCF-1 alpha	315-329	20
NF-κ-B-II	350-359	20
NF-κ-B-I	364-373	20
Sp1-III	375-386	20
Sp1-II	388-397	20
Sp1-I	398-408	20
TATA Box	427-431	22
TAR element start	453	22
5' LTR U3 end	455	22
+1 mRNA start site	456	22
5' LTR R repeat begin	456	22
TAR element end	513	22
Poly-A signal	527-532	22
5' LTR R repeat end	551	22
5' LTR U5 start	552	22
Extensive secondary structure	568-605	22
5' LTR U5 end	633	24
Lys tRNA primer binding site	634-653	24
Packaging loops begin	681	24
Packaging loops end	789	26
Gag and Gag-Pol start	790	26
Gag p17 Matrix end	1185	32
Gag p24 Capsid start	1186	32
Gag p24 Capsid end	1881	40
Gag p2 start	1882	40
Gag p2 end	1920	42
Gag-Pol fusion TF protein start	1921	42
Gag p7 nucleocapsid start	1921	42
Gag p7 nucleocapsid end	2085	44
Gag-Pol -1 ribosomal slip site	2085	44
Gag p1 start	2086	44
Gag p1 end	2133	44
Gag p6 start	2134	44
Gag-Pol TF end	2252	46
Pol protease start	2253	46
Gag p6 end	2292	46
Gag end	2292	46
Pol Protease end	2549	50
Pol p66 and p51 RT start	2550	50
p51 end and p66 RT continue	3869	66
Pol p15 RNAse H start	3870	66
Pol p66 RT, Pol p15 Rnase H end	4229	70
Pol p31 Integrase start	4230	70
Vif start	5041	80
Pol p31 integrase end	5096	80
Vpr start	5559	86
Vif end	5619	86

Feature	Location	Page
frameshift insert in HXB2	5772	88
Vpr premature end (HXB2 only)	5795	88
Tat exon 1 start	5831	90
Vpr end	5850	90
Rev exon 1 start	5970	92
Tat Rev exon 1 end	6045	92
intron start	6046	92
Vpu start (ACG in HXB2)	6062	92
Env start	6225	94
Vpu end	6310	96
Env signal peptide end	6314	96
Env gp120 start	6315	96
V1 loop start	6615	100
V1 loop end	6691	102
V2 loop start	6696	102
V2 loop end	6811	104
V3 loop start	7110	108
V3 loop end	7217	110
V4 loop start	7377	112
V4 loop end	7477	114
V5 loop start	7602	116
V5 loop end	7636	116
Rev Responsive Element (RRE) region	7710	116
Env gp120 end	7757	118
Env gp41 start	7758	118
RRE end	8061	122
Tat Rev intron end	8378	126
Tat Rev exon 2 start	8379	126
Tat premature stop in HXB2	8424	126
Tat end	8469	126
Rev end (TAA) in some lineages	8605	128
Rev end	8653	130
Env gp41, gp160 end	8795	132
Nef start	8797	132
3' LTR U3 start	9086	136
Nef premature end in HXB2	9168	136
TCF-1 alpha binding	9400-9414	140
Nef end	9417	140
NF-κ-B-II	9435-9444	140
NF-κ-B-I	9449-9458	142
Sp1-III	9462-9471	142
Sp1-II	9473-9482	142
Sp1-I	9483-9493	142
TATA box	9512-9516	142
TAR element start	9538	142
3' LTR U3 end	9540	142
3' LTR repeat start	9541	142
TAR element end	9599	144
Poly-A signal	9612-9617	144
3' LTR R repeat end	9636	144
3' LTR U5 start	9637	144
Extensive secondary structure in this region?	9646	144
3' LTR U5 end	9719	146

II-3 Sequences

Sequences included in the HIV-1/SIVcpz complete genome alignment.

Name	Accession	Country	Author	Reference
B.FR.83.HXB2	K03455	France	van Beveren, C.P.	*(in) Weiss,RL, Teich,N, Varmus,H and Coffin,J(Eds); RNA TUMOR VIRUSES, SECOND EDITION,* **2**, Vol 2: 1102-1123; Cold Spring Harbor Laboratory, Cold Spring Harbor (1985)
A1.GE.99.99GEMZ011	DQ207944	Georgia	Zarandia, M.	*ARHR* **22**(5):470-476 (2006)
A1.KE.00.KER2008	AF457052	Kenya	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
A1.KE.00.KNH1144	AF457066	Kenya	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
A1.KE.00.KSM4024	AF457077	Kenya	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
A1.KE.00.MSA4069	AF457080	Kenya	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
A1.KE.00.NKU3005	AF457089	Kenya	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
A1.RU.00.RU00051	EF545108	Russia	Thomson, M.M.	*ARHR* **23**(12):1599-1604 (2007)
A1.RW.93.93RW_024	AY713406	Rwanda	Brown, B.K.	*J Virol* **79**(10):6089-6101 (2005)
A1.SE.95.SE8891	AF069673	Sweden	Carr, J.K.	*AIDS* **13**(14):1819-1826 (1999)
A1.SE.95.UGSE8131	AF107771	Sweden	Carr, J.K.	*AIDS* **13**(14):1819-1826 (1999)
A1.TZ.01.A173	AY253305	Tanzania	Arroyo, M.A.	*ARHR* **20**(8):895-901 (2004)
A1.UA.01.01UADN139	DQ823357	Ukraine	Saad, M.D.	*ARHR* **22**(8):709-714 (2006)
A1.UG.92.92UG037	AB253429	Uganda	Sakamoto, Y.	Unpublished
A1.UG.99.99UGA07072	AF484478	Uganda	Harris, M.E.	*ARHR* **18**(17):1281-1290 (2002)
A2.CD.97.97CDKTB48	AF286238	D.R.C.	Gao, F.	*ARHR* **17**(8):675-688 (2001)
A2.CY.94.94CY017_41	AF286237	Cyprus	Gao, F.	*ARHR* **17**(8):675-688 (2001)
A.SN.01.DDI579	AY521629	Senegal	Meloni, S.T.	*J Virol* **78**(22):12438-12445 (2004)
A.SN.01.DDJ369	AY521631	Senegal	Meloni, S.T.	*J Virol* **78**(22):12438-12445 (2004)
A.SN.96.DDJ360	AY521630	Senegal	Meloni, S.T.	*J Virol* **78**(22):12438-12445 (2004)
A.CD.02.02CD_KTB035	AM000055	D.R.C.	Vidal, N.	*ARHR* **22**(2):182-187 (2006)
A.CD.97.97CD_KCC2	AM000053	D.R.C.	Vidal, N.	*ARHR* **22**(2):182-187 (2006)
A.CD.97.97CD_KTB13	AM000054	D.R.C.	Vidal, N.	*ARHR* **22**(2):182-187 (2006)
B.AR.04.04AR151516	DQ383752	Argentina	Pando, M.A.	*Retrovirology* **3** , 59 (2006)
B.AU.87.MBC925	AF042101	Australia	Oelrichs, R.B.	*J Biomed Sci* **7**(2):128-135 (2000)
B.BO.99.BOL0122	AY037270	Bolivia	Carr, J.K.	*AIDS* **15**(15):F41-F47 (2001)
B.BR.03.BREPM2012	EF637046	Brazil	Sa-Filho, D.	*ARHR* **23**(9):1087-1094 (2007)
B.CA.97.CANB3FULL	AY779553	Canada	Cali, L.	*ARHR* **21**(8):728-733 (2005)
B.CN.05.05CNHB_hp3	DQ990880	China	Tan, J.-x.	*Chin Med J(Engl)* **120**(9); 831-3 (2007)
B.CO.01.PCM001	AY561236	Colombia	Sanchez, G.I.	*Am J Trop Med Hyg* **74**(4); 674-7 (2006)
B.GB.83.CAM1	D10112	United Kingdom	Saurya, S.	*ARHR* **19**(1):73-76 (2003)
B.GB.86.GB8	AJ271445	United Kingdom	Farrar, G.H.	*J Med Virol* **34**(2):104-113 (1991)
B.GE.03.03GEMZ010	DQ207942	Georgia	Zarandia, M.	*ARHR* **22**(5):470-476 (2006)
B.IT.05.SG1	DQ672623	Italy	Calugi, G.	*J Virol* **80**(23):11892-11896 (2006)
B.JP.05.DR6538	AB287363	Japan	Sakamoto, Y.	Unpublished
B.KR.05.05CSR3	DQ837381	S. Korea	Cho, Y.	Unpublished
B.NL.00.671_00T36	AY423387	Netherlands	Geels, M.J.	*J Virol* **77**(23):12430-12440 (2003)
B.RU.04.04RU128005	AY682547	Russia	Galkin, A.N.	*ARHR* **22**(11):1192-1197 (2006)
B.TH.00.00TH_C3198	AY945710	Thailand	Watanaveeradej, V.	*ARHR* **22**(8):801-7 (2006)

Name	Accession	Country	Author	Reference
B.UA.01.01UAKV167	DQ823362	Ukraine	Saad, M.D.	*ARHR* **22**(8):709-714 (2006)
B.US.04.ES10_53	EF363127	United States	Blankson, J.N.	*J Virol* **81**(5):2508-2518 (2007)
B.US.99.PRB959_03	AY331296	United States	Bernardin, F.	*J Virol* **79**(17):11523-11528 (2005)
C.AR.01.ARG4006	AY563170	Argentina	Carrion, G.	*ARHR* **20**(9):1022-1025 (2004)
C.BR.04.04BR013	AY727522	Brazil	Sanabani, S.	*ARHR* **22**(2):171-176 (2006)
C.BW.00.00BW07621	AF443088	Botswana	Novitsky, V.	*J Virol* **76**(11):5435-5451 (2002)
C.CN.98.YNRL9840	AY967806	China	Qiu, Z.	*ARHR* **21**(12):1051-1056 (2005)
C.ET.02.02ET_288	AY713417	Ethiopia	Brown, B.K.	*J Virol* **79**(10):6089-6101 (2005)
C.GE.03.03GEMZ033	DQ207941	Georgia	Zarandia, M.	*ARHR* **22**(5):470-476 (2006)
C.IL.99.99ET7	AY255824	Israel	Harris, M.E.	*ARHR* **19**(12):1125-33 (2003)
C.IN.99.01IN565_10	AY049708	India	Khurana, S.	Unpublished
C.KE.00.KER2010	AF457054	Kenya	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
C.MM.99.mIDU101_3	AB097871	Myanmar	Takebe, Y.	*AIDS* **17**(14):2077-87 (2003)
C.MW.93.93MW_965	AY713413	Malawi	Brown, B.K.	*J Virol* **79**(10):6089-6101 (2005)
C.SN.90.90SE_364	AY713416	Senegal	Brown, B.K.	*J Virol* **79**(10):6089-6101 (2005)
C.SO.89.89SM_145	AY713415	Somalia	Brown, B.K.	*J Virol* **79**(10):6089-6101 (2005)
C.TZ.02.CO178	AY734556	Tanzania	Arroyo, M.A.	*AIDS* **19**(14):1517-1524 (2005)
C.UY.01.TRA3011	AY563169	Uruguay	Carrion, G.	*ARHR* **20**(9):1022-1025 (2004)
C.YE.02.02YE511	AY795906	Yemen	Saad, M.D.	*ARHR* **21**(7):644-648 (2005)
C.ZA.04.04ZASK164B1	DQ056405	S. Africa	Rousseau, C.M.	*J Virol Methods* **136**(1-2):118-125 (2006)
C.ZA.05.05ZAPSK240B1	DQ369991	S. Africa	Rousseau, C.M.	*J Virol Methods* **136**(1-2):118-125 (2006)
C.ZM.02.02ZM108	AB254141	Zambia	Tatsumi, M.	Unpublished
D.CD.83.ELI	K03454	D.R.C.	Alizon, M.	*Cell* **46**(1):63-74 (1986)
D.CM.01.01CM_0009BBY	AY371155	Cameroon	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
D.KE.01.NKU3006	AF457090	Kenya	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
D.KR.04.04KBH8	DQ054367	S. Korea	Cho, Y.	Unpublished
D.TD.99.MN011	AJ488926	Chad	Vidal, N.	*JAIDS* **33**(2):239-246 (2003)
D.TZ.01.A280	AY253311	Tanzania	Arroyo, M.A.	*ARHR* **20**(8):895-901 (2004)
D.UG.99.99UGK09259	AF484498	Uganda	Harris, M.E.	*ARHR* **18**(17):1281-1290 (2002)
D.YE.01.01YE386	AY795903	Yemen	Saad, M.D.	*ARHR* **21**(7):644-648 (2005)
D.YE.02.02YE516	AY795907	Yemen	Saad, M.D.	*ARHR* **21**(7):644-648 (2005)
D.ZA.90.R1	EF633445	S. Africa	Jacobs, G.B.	*ARHR* **23**(12):1575-8 (2007)
F1.AR.02.ARE933	DQ189088	Argentina	Aulicino, P.C.	*ARHR* **21**(2):158-164 (2005)
F1.BE.93.VI850	AF077336	Belgium	Laukkanen, T.	*Virology* **269**(1):95-104 (2000)
F1.BR.01.01BR125	DQ358802	Brazil	Sanabani, S.	*Infect Genet Evol* **6**(5):368-377 (2006)
F1.ES.x.P1146	DQ979023	Spain	Sierra, M.	Unpublished
F1.FI.93.FIN9363	AF075703	Finland	Laukkanen, T.	*Virology* **269**(1):95-104 (2000)
F2.CM.02.02CM_0016BBY	AY371158	Cameroon	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
F2.CM.95.MP257	AJ249237	Cameroon	Triques, K.	*ARHR* **16**(2):139-151 (2000)
F2.CM.97.CM53657	AF377956	Cameroon	Carr, J.K.	*Virology* **286**(1):168-181 (2001)
G.BE.96.DRCBL	AF084936	Belgium	Oelrichs, R.B.	*ARHR* **15**(6):585-589 (1999)
G.CM.01.01CM_4049HAN	AY371121	Cameroon	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
G.CU.99.Cu74	AY586547	Cuba	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
G.ES.99.X138	AF450098	Spain	Delgado, E.	*JAIDS* **29**(5):536-543 (2002)
G.GH.03.03GH175G	AB287004	Ghana	Takekawa, N.	Unpublished
G.KE.93.HH8793_12_1	AF061641	Kenya	Salminen, M.	*ARHR* **8**(9):1733-1742 (1992)
G.NG.01.01NGPL0669	DQ168576	Nigeria	Carr, J.K.	Unpublished
G.PT.x.PT2695	AY612637	Portugal	Esteves, A.M.	Unpublished
G.SE.93.SE6165	AF061642	Sweden	Carr, J.K.	*Virology* **247**(1):22-31 (1998)
H.BE.93.VI991	AF190127	Belgium	Janssens, W.	*AIDS* **14**(11):1533-1543 (2000)

Name	Accession	Country	Author	Reference
H.BE.93.VI997	AF190128	Belgium	Janssens, W.	*AIDS* **14**(11):1533-1543 (2000)
H.CF.90.056	AF005496	C.A.R.	Gao, F.	*J Virol* **72**(7):5680-5698 (1998)
J.CD.97.J_97DC_KTB147	EF614151	D.R.C.	Abecasis, A.B.	*J Virol* **81**(16):8543-8551 (2007)
J.SE.93.SE7887	AF082394	Sweden	Laukkanen, T.	*ARHR* **15**(3):293-297 (1999)
K.CD.97.EQTB11C	AJ249235	D.R.C.	Triques, K.	*ARHR* **16**(2):139-151 (2000)
K.CM.96.MP535	AJ249239	Cameroon	Triques, K.	*ARHR* **16**(2):139-151 (2000)
01_AE.CF.90.90CF11697	AF197340	C.A.R.	Anderson, J.P.	*J Virol* **74**(22):10752-10765 (2000)
01_AE.CN.05.FJ051	DQ859178	China	Huang, H.L.	*Chin Med J* **119**(19):1622-1628 (2006)
01_AE.CN.06.FJ054	DQ859180	China	Huang, H.L.	*Chin Med J* **119**(19):1622-1628 (2006)
01_AE.HK.x.HK001	DQ234790	Hong Kong	Tsui, S.K.W.	Unpublished
01_AE.TH.01.01TH_R2184	AY945730	Thailand	Watanaveeradej, V.	*ARHR* **22**(8):801-7 (2006)
01_AE.TH.02.OUR769I	AY358062	Thailand	Tovanabutra, S.	*ARHR* **20**(5):465-75 (2004)
01_AE.TH.90.CM240	U54771	Thailand	Carr, J.K.	*J Virol* **70**(9):5935-5943 (1996)
01_AE.US.00.00US_MSC1164	AY444804	United States	Tovanabutra, S.	*ARHR* **21**(5):424-429 (2005)
02_AG.CM.02.02CM_4082STN	AY371141	Cameroon	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
02_AG.EC.x.ECU41	AY151001	Ecuador	Carrion, G.	*ARHR* **19**(4):329-32 (2003)
02_AG.GH.03.GHNJ196	AB231898	Ghana	Tatsumi, M.	Unpublished
02_AG.NG.01.PL0710	DQ168577	Nigeria	Carr, J.K.	Unpublished
02_AG.NG.x.IBNG	L39106	Nigeria	Howard, T.M.	*ARHR* **10**(12):1755-1757 (1994)
02_AG.SE.94.SE7812	AF107770	Sweden	Laukkanen, T.	Unpublished
02_AG.SN.98.MP1211	AJ251056	Senegal	Toure-Kane, C.	*ARHR* **16**(6):603-609 (2000)
03_AB.RU.97.KAL153_2	AF193276	Russia	Liitsola, K.	*ARHR* **16**(11):1047-1053 (2000)
04_cpx.CY.94.CY032	AF049337	Cyprus	Gao, F.	*J Virol* **72**(12):10234-10241 (1998)
05_DF.BE.x.VI1310	AF193253	Belgium	Laukkanen, T.	*Virology* **269**(1):95-104 (2000)
06_cpx.AU.96.BFP90	AF064699	Australia	Oelrichs, R.B.	*ARHR* **14**(16):1495-1500 (1998)
08_BC.CN.97.97CNGX_6F	AY008715	China	Piyasirisilp, S.	*J Virol* **74**(23):11286-11295 (2000)
09_cpx.GH.96.96GH2911	AY093605	Ghana	McCutchan, F.E.	*ARHR* **20**(8):819-826 (2004)
10_CD.TZ.96.96TZ_BF061	AF289548	Tanzania	Koulinska, I.N.	*ARHR* **17**(5):423-431 (2001)
11_cpx.GR.x.GR17	AF179368	Greece	Paraskevis, D.	*ARHR* **16**(9):845-855 (2000)
12_BF.AR.99.ARMA159	AF385936	Argentina	Carr, J.K.	*AIDS* **15**(15):F41-F47 (2001)
13_cpx.CM.96.1849	AF460972	Cameroon	Wilbe, K.	*ARHR* **18**(12):849-56 (2002)
14_BG.DE.01.9196_01	AY882421	Germany	Harris, B.	*ARHR* **21**(7):654-660 (2005)
15_01B.TH.99.99TH_MU2079	AF516184	Thailand	Viputtijul, K.	*ARHR* **18**(16):1235-1237 (2002)
16_A2D.KR.97.97KR004	AF286239	S. Korea	Gao, F.	*ARHR* **17**(8):675-688 (2001)
17_BF.PE.02.PE02_PCR0155	EU581828	Peru	Carr, J.K.	Unpublished
18_cpx.CU.99.CU76	AY586540	Cuba	Thomson, M.M.	*AIDS* **19**(11):1155-63 (2005)
19_cpx.CU.99.CU7	AY894994	Cuba	Casado, G.	*JAIDS* **40**(5):532-537 (2005)
20_BG.CU.99.Cu103	AY586545	Cuba	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
21_A2D.KE.91.KNH1254	AY945737	Kenya	Visawapoka, U.	*ARHR* **22**(7):695-702 (2006)
22_01A1.CM.01.01CM_0001BBY	AY371159	Cameroon	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
23_BG.CU.03.CB118	AY900571	Cuba	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
24_BG.CU.03.CB378	AY900574	Cuba	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
24_BG.CU.03.CB471	AY900575	Cuba	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
25_cpx.CM.01.101BA	DQ826726	Cameroon	Carr, J.K.	Unpublished
27_cpx.CD.97.97CDKTB49	AJ404325	D.R.C.	Vidal, N.	*ARHR* **16**(18):2059-2064 (2000)
28_BF.BR.99.BREPM12609	DQ085873	Brazil	Sa Filho, D.J.	*ARHR* **22**(1):1-13 (2006)
29_BF.BR.02.BREPM119	AY771590	Brazil	Sa Filho, D.J.	*ARHR* **21**(2):145-151 (2005)
31_BC.BR.02.110PA	EF091932	Brazil	Santos, A.F.	*AIDS* **20**(16):2011-2019 (2006)
32_06A1.EE.01.EE0369	AY535660	Estonia	Adojaan, M.	*JAIDS* **39**(5):598-605 (2005)
33_01B.MY.05.05MYKL007_1	DQ366659	Malaysia	Tee, K.K.	*JAIDS* **43**(5):523-529 (2006)
34_01B.TH.99.OUR2478P	EF165541	Thailand	Tovanabutra, S.	*ARHR* **23**(6):829-833 (2007)

Name	Accession	Country	Author	Reference
35_AD.AF.05.05AF095	EF158041	Afghanistan	Sanders-Buell, E.	*ARHR* **23**(6):834-839 (2007)
36_cpx.CM.00.00CMNYU830	EF087994	Cameroon	Powell, R.L.	*ARHR* **23**(8):1008-1019 (2007)
37_cpx.CM.00.00CMNYU926	EF116594	Cameroon	Powell, R.L.	*ARHR* **23**(7):923-933 (2007)
38_BF1.UY.05.UY05_4752	FJ213780	Uruguay	Ruchansky, D.	Unpublished
39_BF.BR.03.03BRRJ103	EU735534	Brazil	Guimaraes, M.L.	*AIDS* **22**(3):433-435 (2008)
40_BF.BR.04.04BRRJ115	EU735538	Brazil	Guimaraes, M.L.	*AIDS* **22**(3):433-435 (2008)
42_BF.LU.03.luBF_05_03	EU170155	Luxembourg	Struck, D.	Unpublished
43_02G.SA.03.J11223	EU697904	Saudi Arabia	Badreddine, S.	*ARHR* **23**(5):667-674 (2007)
U.CD.83.83CD003	AF286236	D.R.C.	Gao, F.	*ARHR* **17**(12):1217-1222 (2001)
U.CD.90.90CD121E12	AF457101	D.R.C.	Mokili, J.L.	*ARHR* **18**(11):817-823 (2002)
U.GR.99.GR303	AY046058	Greece	Paraskevis, D.	*J Gen Virol* **82**(Pt 10):2509-2514 (2001)
U.NL.95.U_NL_95_H10986_D1	EF029066	Netherlands	van der Hoek, L.	*ARHR* **23**(3):466-470 (2007)
CPZ.CD.90.ANT	U42720	D.R.C.	Vanden Haesevelde, M.M.	*Virology* **221**(2):346-350 (1996)
CPZ.CM.01.SIVcpzCAM13	AY169968	Cameroon	Nerrienet, E.	*J Virol* **79**(2):1312-9 (2005)
CPZ.CM.05.SIVcpzEK505	DQ373065	Cameroon	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.CM.05.SIVcpzLB7	DQ373064	Cameroon	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.CM.05.SIVcpzMB66	DQ373063	Cameroon	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.CM.05.SIVcpzMT145	DQ373066	Cameroon	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.GA.88.GAB1	X52154	Gabon	Huet, T.	*Nature* **345**(6273):356-359 (1990)
CPZ.TZ.00.TAN1	AF447763	Tanzania	Santiago, M.L.	*J Virol* **77**(3):2233-2242 (2003)
CPZ.US.85.CPZUS	AF103818	United States	Gao, F.	*Nature* **397**(6718):436-441 (1999)
CPZ.CM.x.SIVcpzDP943	EF535993	Cameroon	Heuverswyn, F.V.	*Virology* **368**(1):155-171 (2007)
CPZ.CM.x.SIVcpzMB897	EF535994	Cameroon	Heuverswyn, F.V.	*Virology* **368**(1):155-171 (2007)
N.CM.02.DJO0131	AY532635	Cameroon	Bodelle, P.	*ARHR* **20**(8):902-908 (2004)
N.CM.04.04CM_1015_04	DQ017382	Cameroon	Yamaguchi, J.	*ARHR* **22**(1):83-92 (2006)
N.CM.04.04CM_1131_03	DQ017383	Cameroon	Yamaguchi, J.	*ARHR* **22**(1):83-92 (2006)
N.CM.95.YBF30	AJ006022	Cameroon	Simon, F.	*Nat Med* **4**(9):1032-1037 (1998)
N.CM.97.YBF106	AJ271370	Cameroon	Roques, P.	*AIDS* **18**(10):1371-1381 (2004)
O.BE.87.ANT70	L20587	Belgium	Vanden Haesevelde, M.	*J Virol* **68**(3):1586-1596 (1994)
O.CM.91.MVP5180	L20571	Cameroon	Gurtler, L.G.	*J Virol* **68**(3):1581-1585 (1994)
O.CM.96.96CMABB637	AY169810	Cameroon	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.98.98CMA104	AY169802	Cameroon	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.99.99CMU4122	AY169815	Cameroon	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.FR.92.VAU	AF407418	France	Vartanian, J.P.	*J Gen Virol* **83**(Pt 4):801-805 (2002)
O.SN.99.SEMP1299	AJ302646	Senegal	Toure-Kane, C.	*ARHR* **17**(12):1211-1216 (2001)
O.US.99.99USTWLA	AY169814	United States	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)

```
                          5' LTR U3 start
B.FR.83.HXB2              .TGGAAGGGCTAATTCAC...TCCCAACGA...AGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTC...CCTGATTAGCAGAACTACACACCAGGGCCAGGG...ATCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGA  157
A1.GE.99.99GEMZ011        ..........................................................................................................................................................................    0
A1.KE.00.KER2008          ..........................................................................................................................................................................    0
A1.KE.00.KNH1144          ..........................................................................................................................................................................    0
A1.KE.00.KSM4024          ..........................................................................................................................................................................    0
A1.KE.00.MSA4069          ..........................................................................................................................................................................    0
A1.KE.00.NKU3005          ..........................................................................................................................................................................    0
A1.RU.00.RU00051          ..........................................................................................................................................................................    0
A1.RW.93.93RW_024         ..........................................................................................................................................................................    0
A1.SE.95.SE8891           ..........................................................................................................................................................................    0
A1.SE.95.UGSE8131         ..........................................................................................................................................................................    0
A1.TZ.01.A173             ..........................................................................................................................................................................    0
A1.UA.01.01UADN139        ..........................................................................................................................................................................    0
A1.UG.92.92UG037          .----T---T-----T--...---A-GAA-...--G-----A---------------G-----------------------...-------G------T---------------A--...--------C-----A--A-------------T---------------G--  157
A1.UG.99.99UGA07072       ..........................................................................................................................................................................    0
A2.CD.97.97CDKTB48        ..........................................................................................................................................................................    0
A2.CY.94.94CY017_41       ..........................................................................................................................................................................    0
A.SN.01.DDI579            ..........................................................................................................................................................................    0
A.SN.01.DDJ369            ..........................................................................................................................................................................    0
A.SN.96.DDJ360            ..........................................................................................................................................................................    0
A.CD.02.02CD_KTB035       .----TA--T-----TGG...-----GAA-...--------C---------T-----G-----------------------...-------G------T------------------...G-------------A--A-------------T------------------  157
A.CD.97.97CD_KCC2         .----T---T-----TGG...-----GAA-...--------C---------------G----T-----G------------...-------G------T------------------...--T----TC--T--A--A-------------T----T-------------  157
A.CD.97.97CD_KTB13        ..........................................................................................................................................................................    0
B.AR.04.04AR151516        ..........................................................................................................................................................................    0
B.AU.87.MBC925            .--------------T--...------AAG...------------------------------------------------...-------G-------------------------...------------A-A----------------T------------------  157
B.BO.99.BOL0122           ..........................................................................................................................................................................    0
B.BR.03.BREPM2012         ..........................................................................................................................................................................    0
B.CA.97.CANB3FULL         ..........................................................................................................................................................................    0
B.CN.05.05CNHB_hp3        .--------------T--...------AAG...--------G---------------G-------T----------T----...-------G-G-C-----T--------------A...--------C--T-------------------T------------------  157
B.CO.01.PCM001            ..........................................................................................................................................................................    0
B.GB.83.CAM1              .--------A-----T--...------A--...-----G------------------------------------------...-------G-------------------------...-----------------A-------------T------------------  157
B.GB.86.GB8               ..-------------TGG...------AA-...--------G---------------G-----------------------...-------G-------------------------...-C-----TC----------------------TT--A--------------  156
B.GE.03.03GEMZ010         ..........................................................................................................................................................................    0
B.IT.05.SG1               ..........................................................................................................................................................................    0
B.JP.05.DR6538            .----T---------T--...------AA-..A--------A---------------G-----------------AT----...-------G---T--T------------------...-----G-TC--------------G-------T-------------A-G--  158
B.KR.05.05CSR3            .--------T--------...------AAG...-----------------CT-----------------------------...-------G---AGG-------------------...G------------------------------T----T-------------  157
B.NL.00.671_00T36         ..........................................................................................................................................................................    0
B.RU.04.04RU128005        ..........................................................................................................................................................................    0
B.TH.00.00TH_C3198        ..........................................................................................................................................................................    0
B.UA.01.01UAKV167         ..........................................................................................................................................................................    0
B.US.04.ES10_53           .--------------T--...-------AG...--------------------T---G-T---A-----------------...-------G---A-----T---------------...-C---------C-----------C-------T---------------G--  157
B.US.99.PRB959_03         ..........................................................................................................................................................................    0
C.AR.01.ARG4006           ..........................................................................................................................................................................    0
C.BR.04.04BR013           ..........................................................................................................................................................................    0
C.BW.00.00BW07621         ..........................................................................................................................................................................    0
C.CN.98.YNRL9840          ..........................................................................................................................................................................    0
C.ET.02.02ET_288          ..........................................................................................................................................................................    0
C.GE.03.03GEMZ033         ..........................................................................................................................................................................    0
C.IL.99.99ET7             ..........................................................................................................................................................................    0
C.IN.99.01IN565_10        ..........................................................................................................................................................................    0
C.KE.00.KER2010           ..........................................................................................................................................................................    0
C.MM.99.mIDU101_3         ..........................................................................................................................................................................    0
C.MW.93.93MW_965          ..........................................................................................................................................................................    0
C.SN.90.90SE_364          ..........................................................................................................................................................................    0
C.SO.89.89SM_145          ..........................................................................................................................................................................    0
C.TZ.02.CO178             ..........................................................................................................................................................................    0
C.UY.01.TRA3011           ..........................................................................................................................................................................    0
C.YE.02.02YE511           ..........................................................................................................................................................................    0
C.ZA.04.04ZASK164B1       ..........................................................................................................................................................................    0
C.ZA.05.05ZAPSK240B1      ..........................................................................................................................................................................    0
C.ZM.02.02ZM108           .--------T-----T--...--TA-GAA-...--------G---------T-----G----T-------------T----...-------G---A---------T-G--A------...G-------C----------------------T---------G--------  157
D.CD.83.ELI               ..........................................................................................................................................................................    0
D.CM.01.01CM_0009BBY      ..........................................................................................................................................................................    0
D.KE.01.NKU3006           ..........................................................................................................................................................................    0
D.KR.04.04KBH8            ..................................................-------------------------------...-----C-G---T---------------------...-----------------------T-------T-G----------------  114
D.TD.99.MN011             ..........................................................................................................................................................................    0
D.TZ.01.A280              ..........................................................................................................................................................................    0
D.UG.99.99UGK09259        ..........................................................................................................................................................................    0
D.YE.01.01YE386           ..........................................................................................................................................................................    0
D.YE.02.02YE516           ..........................................................................................................................................................................    0
D.ZA.90.R1                ..........................................................................................................................................................................    0
F1.AR.02.ARE933           ..........................................................................................................................................................................    0
F1.BE.93.VI850            ..........................................................................................................................................................................    0
F1.BR.01.01BR125          ..........................................................................................................................................................................    0
F1.ES.x.P1146             ..........................................................................................................................................................................    0
F1.FI.93.FIN9363          ..........................................................................................................................................................................    0
F2.CM.02.02CM_0016BBY     ..........................................................................................................................................................................    0
F2.CM.95.MP257            ..........................................................................................................................................................................    0
F2.CM.97.CM53657          ..........................................................................................................................................................................    0
G.BE.96.DRCBL             .............................................AC-TGA-CTAT-GG---TA-T--CA----A-T----...--A----G-------------------------...-CT----TC---------C----G-------T---A--G------A-G--  119
G.CM.01.01CM_4049HAN      ..........................................................................................................................................................................    0
G.CU.99.Cu74              ..........................................................................................................................................................................    0
G.ES.99.X138              ..........................................................................................................................................................................    0
G.GH.03.03GH175G          .----T---T-----TGG...---A-GAA-...--------G---------------G----TA----------A-T----...--A----G-------------------------...-CT---CTC--------------G-------T---A---------A-A--  157
G.KE.93.HH8793_12_1       ..........................................................................................................................................................................    0
G.NG.01.01NGPL0669        ..........................................................................................................................................................................    0
G.PT.x.PT2695             ..............................................................................................-------T--------T------...-CT----TC--------------G-------T---A---------A-G--   73
G.SE.93.SE6165            ..........................................................................................................................................................................    0
H.BE.93.VI991             ..........................................................................................................................................................................    0
H.BE.93.VI997             ..........................................................................................................................................................................    0
H.CF.90.056               ..........................................................................................................................................................................    0
J.CD.97.J_97DC_KTB147     ..........................................................................................................................................................................    0
J.SE.93.SE7887            ..........................................................................................................................................................................    0
K.CD.97.EQTB11C           ..........................................................................................................................................................................    0
K.CM.96.MP535             ..........................................................................................................................................................................    0
```

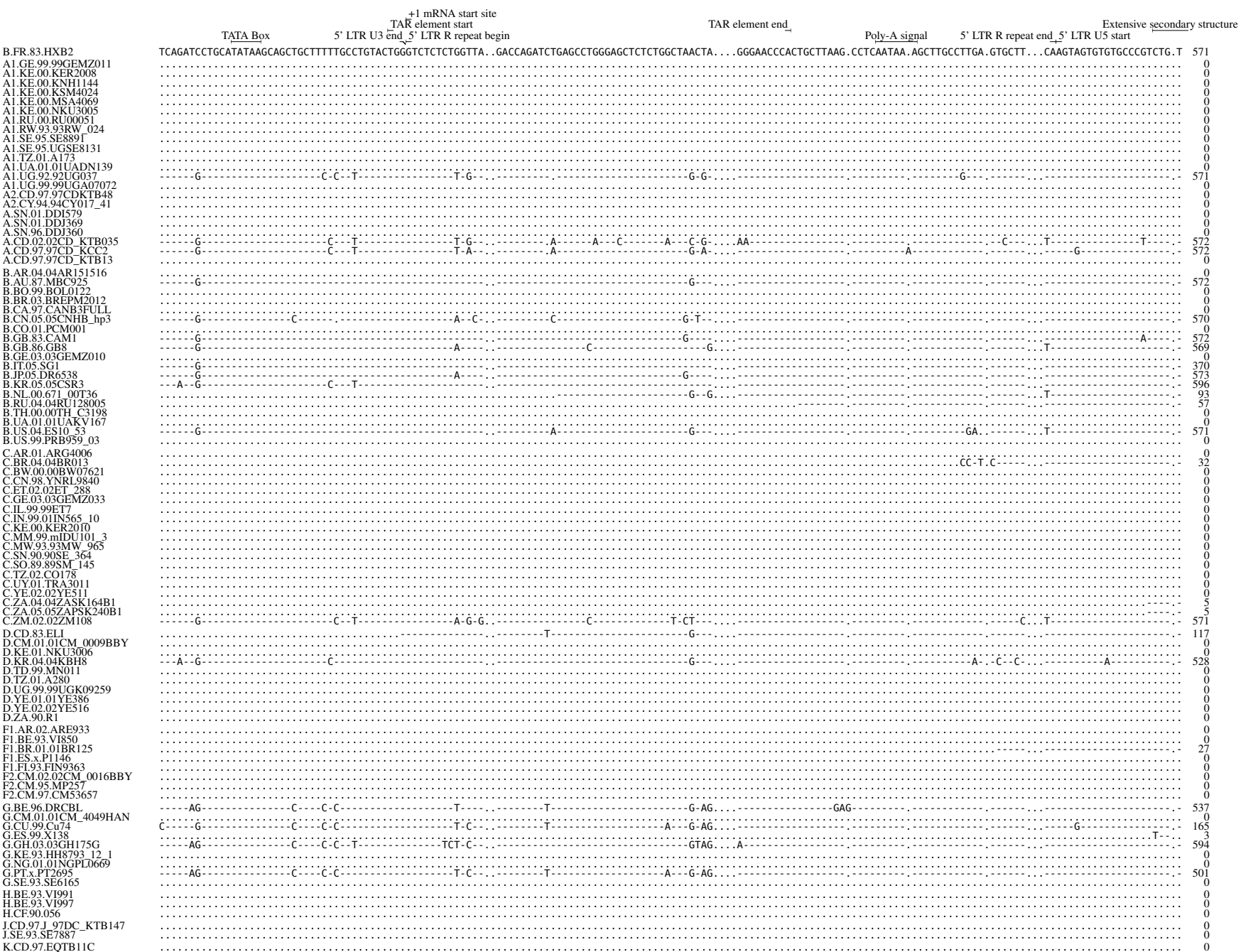
```
                               5' LTR U3 start
B.FR.83.HXB2                   .TGGAAGGGCTAATTCAC...TCCCAACGA...AGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTC...CCTGATTAGCAGAACTACACACCAGGGCCAGGG...ATCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGA  157
01_AE.CF.90.90CF11697          ..................................................-----------...----------------------------------...--------C------TGT-------------------------------   94
01_AE.CN.05.FJ051              .----T------G--T--...---A-GAA-...--------A---------T-A---G----TA-T----------T----...-------G---A-----------------------...--------C------TGT------------T-------A----------  157
01_AE.CN.06.FJ054              .----T------G--T--.TTC-A-G-TA-..AGACA-C--A---A-----G-A---G-----A-T-----T----T----..C-G----CGC--C----T---------------..ATC------CG-GT-ACGT----A--------T----T-----A----C--  162
01_AE.HK.x.HK001               ...........................................................................................................................................    0
01_AE.TH.01.01TH_R2184         ...........................................................................................................................................    0
01_AE.TH.02.OUR769I            ...........................................................................................................................................    0
01_AE.TH.90.CM240              ...........................................................................................................................................    0
01_AE.US.00.00US_MSC1164       ...........................................................................................................................................    0
02_AG.CM.02.02CM_4082STN       ...........................................................................................................................................    0
02_AG.EC.x.ECU41               ...........................................................................................................................................    0
02_AG.GH.03.GHNJ196            .----T---------T--...---A-GAA-...-----G--G--------------G----T----------A-T----...--A----G------------------------...--------C-------------------T---A---------A-A--  157
02_AG.NG.01.PL0710             ...........................................................................................................................................    0
02_AG.NG.x.IBNG                ...........................................................................................................................................    0
02_AG.SE.94.SE7812             ...........................................................................................................................................    0
02_AG.SN.98.MP1211             ...........................................................................................................................................    0
03_AB.RU.97.KAL153_2           ...........................................................................................................................................    0
04_cpx.CY.94.CY032             ...........................................................................................................................................    0
05_DF.BE.x.VI1310              ...........................................................................................................................................    0
06_cpx.AU.96.BFP90             .----T---------T--...---A-GAA-...-----G--G--------------G----T----------T-T----...-------G---A--------------------...CC------C---T----T-------------------------------  157
08_BC.CN.97.97CNGX_6F          ...........................................................................................................................................    0
09_cpx.GH.96.96GH2911          ...........................................................................................................................................    0
10_CD.TZ.96.96TZ_BF061         ...........................................................................................................................................    0
11_cpx.GR.x.GR17               ...........................................................................................................................................    0
12_BF.AR.99.ARMA159            .--------T-----T-T...---A-GAA-...--------G-------------------------T-----...------G-----------------------...G------TG-------------G------T-----------------  157
13_cpx.CM.96.1849              ...........................................................................................................................................    0
14_BG.DE.01.9196_01            ...........................................................................................................................................    0
15_01B.TH.99.99TH_MU2079       ...........................................................................................................................................    0
16_A2D.KR.97.97KR004           ...........................................................................................................................................    0
17_BF.PE.02.PE02_PCR0155       ...........................................................................................................................................    0
18_cpx.CU.99.CU76              ...........................................................................................................................................    0
19_cpx.CU.99.CU7               ...........................................................................................................................................    0
20_BG.CU.99.Cu103              ...........................................................................................................................................    0
21_A2D.KE.91.KNH1254           ...........................................................................................................................................    0
22_01A1.CM.01.01CM_0001BBY     ...........................................................................................................................................    0
23_BG.CU.03.CB118              ...........................................................................................................................................    0
24_BG.CU.03.CB378              ...........................................................................................................................................    0
24_BG.CU.03.CB471              ...........................................................................................................................................    0
25_cpx.CM.01.101BA             ...........................................................................................................................................    0
27_cpx.CD.97.97CDKTB49         ...........................................................................................................................................    0
28_BF.BR.99.BREPM12609         ...........................................................................................................................................    0
29_BF.BR.02.BREPM119           ...........................................................................................................................................    0
31_BC.BR.02.110PA              ...........................................................................................................................................    0
32_06A1.EE.01.EE0369           ...........................................................................................................................................    0
33_01B.MY.05.05MYKL007_1       ...........................................................................................................................................    0
34_01B.TH.99.OUR2478P          ...........................................................................................................................................    0
35_AD.AF.05.05AF095            ...........................................................................................................................................    0
36_cpx.CM.00.00CMNYU830        ...........................................................................................................................................    0
37_cpx.CM.00.00CMNYU926        ...........................................................................................................................................    0
38_BF1.UY.05.UY05_4752         ...........................................................................................................................................    0
39_BF.BR.03.03BRRJ103          ...........................................................................................................................................    0
40_BF.BR.04.04BRRJ115          ...........................................................................................................................................    0
42_BF.LU.03.luBF_05_03         ...........................................................................................................................................    0
43_02G.SA.03.J11223            ...........................................................................................................................................    0
U.CD.83.83CD003                ...........................................................................................................................................    0
U.CD.90.90CD121E12             ...........................................................................................................................................    0
U.GR.99.GR303                  ...........................................................................................................................................    0
U.NL.95.U_NL_95_H10986_D1      ............................................................................................--...-----------------------------------------------   52
CPZ.CD.90.ANT                  ...........................................................................................................................................    0
CPZ.CM.01.SIVcpzCAM13          ...........................................................................................................................................    0
CPZ.CM.05.SIVcpzEK505          ...........................................................................................................................................    0
CPZ.CM.05.SIVcpzLB7            ...........................................................................................................................................    0
CPZ.CM.05.SIVcpzMB66           ...........................................................................................................................................    0
CPZ.CM.05.SIVcpzMT145          ...........................................................................................................................................    0
CPZ.GA.88.GAB1                 .--------T--G--T--...---AGGA--...--------G-------C--C---G----T-----------T----...-----C-G-----------A---A-----A...-CA----TC-----CTGT------------T---A-----G--CC-GAC  157
CPZ.TZ.00.TAN1                 ...........................................................................................................................................    0
CPZ.US.85.CPZUS                .--------T--G--T-T...---AG-A--...--------G-----A-----C---G-T-----T--------AAT----...-----C-G-----T-----C--------T--A...G-A--------C--A--A-A---GC------TT---T----C--TC-CAC  157
CPZ.CM.x.SIVcpzDP943           ...........................................................................................................................................    0
CPZ.CM.x.SIVcpzMB897           ...........................................................................................................................................    0
N.CM.02.DJO0131                ...........................................................................................................................................    0
N.CM.04.04CM_1015_04           ...........................................................................................................................................    0
N.CM.04.04CM_1131_03           ...........................................................................................................................................    0
N.CM.95.YBF30                  ...........................................................................................................................................    0
N.CM.97.YBF106                 ...........................................................................................................................................    0
O.BE.87.ANT70                  .--------T-----T--...-----TAA-...---GC---A-----G-----T---G-G--TA----T--G--A-T----...-------G-----------------A-----A...-C---G-TC--------A---------TG-TT--A----------GTC  157
O.CM.91.MVP5180                .----T---T-----T--...-----TAAG...---GC---A-----G-----C-----A--T-----T--G--A-T----...-------G---TGT--------G--A-----A...CCT----TC--------A----------TG-TT--A--G--------GTC  157
O.CM.96.96CMABB637             ...........................................................................................................................................    0
O.CM.98.98CMA104               ...........................................................................................................................................    0
O.CM.99.99CMU4122              ...........................................................................................................................................    0
O.FR.92.VAU                    ...........................................................................................................................................    0
O.SN.99.SEMP1299               .----T---T-----T--...-----TAA-...---GC---G-----G-----T---G-A--T-----T--G--A-T----...-------G---GG-----------A------...CCA----TC--------A-----G----TA-TT---T----------GTC  157
O.US.99.99USTWLA               ...........................................................................................................................................    0
```

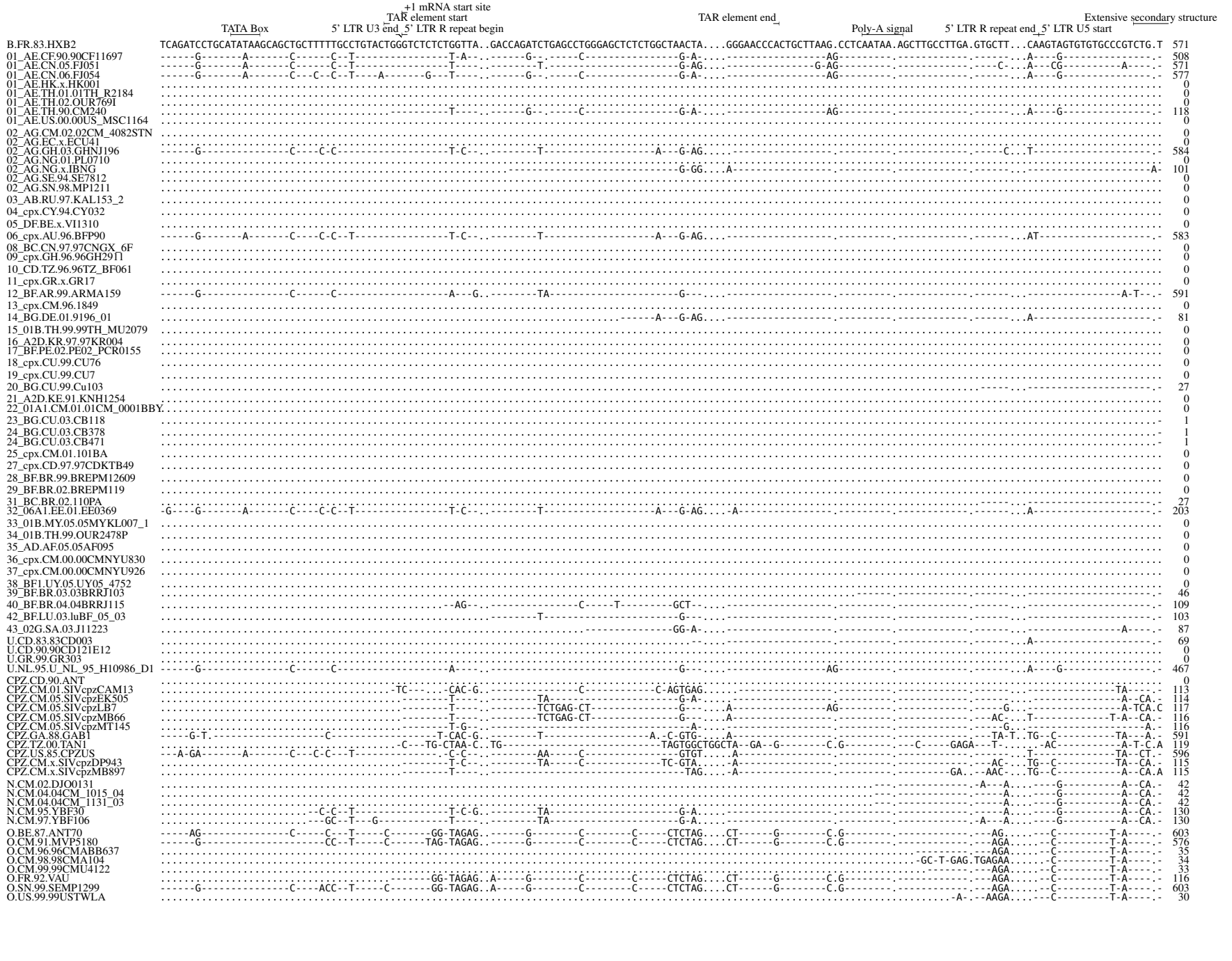
```
B.FR.83.HXB2            G...CCAGAGAAGTTAGAAGAAGCC...AACAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGAATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCAC...ATGGCCCGAGAG......CTGCATCCGG 312
A1.GE.99.99GEMZ011      .......................................................................................................................................................................... 0
A1.KE.00.KER2008        .......................................................................................................................................................................... 0
A1.KE.00.KNH1144        .......................................................................................................................................................................... 0
A1.KE.00.KSM4024        .......................................................................................................................................................................... 0
A1.KE.00.MSA4069        .......................................................................................................................................................................... 0
A1.KE.00.NKU3005        .......................................................................................................................................................................... 0
A1.RU.00.RU00051        .......................................................................................................................................................................... 0
A1.RW.93.93RW_024       .......................................................................................................................................................................... 0
A1.SE.95.SE8891         .......................................................................................................................................................................... 0
A1.SE.95.UGSE8131       .......................................................................................................................................................................... 0
A1.TZ.01.A173           .......................................................................................................................................................................... 0
A1.UA.01.01UADN139      .......................................................................................................................................................................... 0
A1.UG.92.92UG037        T...GA---TG-AG----G-----T...-CTGG-----------AT---C-A---------A-AT---AA-------------TGA-----A----ACA---AG-----A----------A----G---CGAGTA---...-AA--AA----......--A------- 312
A1.UG.99.99UGA07072     .......................................................................................................................................................................... 0
A2.CD.97.97CDKTB48      .......................................................................................................................................................................... 0
A2.CY.94.94CY017_41     .......................................................................................................................................................................... 0
A.SN.01.DDI579          .......................................................................................................................................................................... 0
A.SN.01.DDJ369          .......................................................................................................................................................................... 0
A.SN.96.DDJ360          .......................................................................................................................................................................... 0
A.CD.02.02CD_KTB035     T...T----TGCAG----G--G--T...-CTG-----------A----C-CC------GA-A-A--AA-------------T--T--A-AG---------AT----CA---------T-----G---C-GA-A---...C-A----------......-A-------- 312
A.CD.97.97CD_KCC2       -...-----AGCAG-------G--T...-CTGGG---------AT---C-----------A-AT---AA-------------T--T--------------AT-----A---------T-----G-----AA-A---...-GA---AA----.....---------- 312
A.CD.97.97CD_KTB13      .......................................................................................................................................................................... 0
B.AR.04.04AR151516      .......................................................................................................................................................................... 0
B.AU.87.MBC925          -...----------A-------G---...--TG-----------A--------------A-----------G-----G-----------------------------------------C--C------...------------.....---------- 312
B.BO.99.BOL0122         .......................................................................................................................................................................... 0
B.BR.03.BREPM2012       .......................................................................................................................................................................... 0
B.CA.97.CANB3FULL       .......................................................................................................................................................................... 0
B.CN.05.05CNHB_hp3      T...------C--G-----------...-------------G-T-----C----T---A-A----AA-------CA--------A------------C-GAT-----A------------A------C--C------...-------------.....AAA-----T- 312
B.CO.01.PCM001          .......................................................................................................................................................................... 0
B.GB.83.CAM1            -...------C--G-------G---...--T---A--------G-----------------A------A-----G------------A----A------AC--AT----A-------------------------------...-------------.....AA-------- 312
B.GB.86.GB8             -...------G--G-----A-G---...--TG-----------A-T-T-----------A------A-----G--A--G-----A----A----------T-----A--------A-----------------...--------C--A......-C-------- 311
B.GE.03.03GEMZ010       .......................................................................................................................................................................... 0
B.IT.05.SG1             ........................................TGG-AGGG--AGTTTAAA-G-GA---CAAC-.----A---G-------A-G--------AC----G--A---------AGCATC-GG..AG-ACTG-A---A..CTGCTGA-----A- 114
B.JP.05.DR6538          -...------C--G----G--G---...--T------------A-T--C-----------A-----------G-----A-----A---------------T-----A----------A-------C---------...-A-------------.....-AA------- 313
B.KR.05.05CSR3          -...AA-------A-------G-T-...-CTG--A--------------C--C--------A-A----A------G--A--G--T----------------AT-----A----------------------C--T...G-A--------C.....---------- 312
B.NL.00.671_00T36       .......................................................................................................................................................................... 0
B.RU.04.04RU128005      .......................................................................................................................................................................... 0
B.TH.00.00TH_C3198      .......................................................................................................................................................................... 0
B.UA.01.01UAKV167       .......................................................................................................................................................................... 0
B.US.04.ES10_53         -...A--------G-------G---...---G-----------A---------------A-A----------G----------A----A----------C------A--------------C--C-GC---...G-----------.....---------- 312
B.US.99.PRB959_03       .......................................................................................................................................................................... 0
C.AR.01.ARG4006         .......................................................................................................................................................................... 0
C.BR.04.04BR013         .......................................................................................................................................................................... 0
C.BW.00.00BW07621       .......................................................................................................................................................................... 0
C.CN.98.YNRL9840        .......................................................................................................................................................................... 0
C.ET.02.02ET_288        .......................................................................................................................................................................... 0
C.GE.03.03GEMZ033       .......................................................................................................................................................................... 0
C.IL.99.99ET7           .......................................................................................................................................................................... 0
C.IN.99.01IN565_10      .......................................................................................................................................................................... 0
C.KE.00.KER2010         .......................................................................................................................................................................... 0
C.MM.99.mIDU101_3       .......................................................................................................................................................................... 0
C.MW.93.93MW_965        .......................................................................................................................................................................... 0
C.SN.90.90SE_364        .......................................................................................................................................................................... 0
C.SO.89.89SM_145        .......................................................................................................................................................................... 0
C.TZ.02.CO178           .......................................................................................................................................................................... 0
C.UY.01.TRA3011         .......................................................................................................................................................................... 0
C.YE.02.02YE511         .......................................................................................................................................................................... 0
C.ZA.04.04ZASK164B1     .......................................................................................................................................................................... 0
C.ZA.05.05ZAPSK240B1    .......................................................................................................................................................................... 0
C.ZM.02.02ZM108         C...----G-G-AG--------A--...-GTG--------G---A-T-T---C---------C--A--A-------------TGAAC-C--------A---A------A-----------T-T-----CGCAGA---...--------C---......--A------- 312
D.CD.83.ELI             .......................................................................................................................................................................... 0
D.CM.01.01CM_0009BBY    .......................................................................................................................................................................... 0
D.KE.01.NKU3006         .......................................................................................................................................................................... 0
D.KR.04.04KBH8          A...G-----G--G-------G---...--TG-----------A-T--C-----------A------A---------A--G------------C-----C--A-------A---A-----AAA---------G-A--T...-A-------AT-.....---------- 269
D.TD.99.MN011           .......................................................................................................................................................................... 0
D.TZ.01.A280            .......................................................................................................................................................................... 0
D.UG.99.99UGK09259      .......................................................................................................................................................................... 0
D.YE.01.01YE386         .......................................................................................................................................................................... 0
D.YE.02.02YE516         .......................................................................................................................................................................... 0
D.ZA.90.R1              .......................................................................................................................................................................... 0
F1.AR.02.ARE933         .......................................................................................................................................................................... 0
F1.BE.93.VI850          .......................................................................................................................................................................... 0
F1.BR.01.01BR125        .......................................................................................................................................................................... 0
F1.ES.x.P1146           .......................................................................................................................................................................... 0
F1.FI.93.FIN9363        .......................................................................................................................................................................... 0
F2.CM.02.02CM_0016BBY   .......................................................................................................................................................................... 0
F2.CM.95.MP257          .......................................................................................................................................................................... 0
F2.CM.97.CM53657        .......................................................................................................................................................................... 0
G.BE.96.DRCBL           T...---TCAG--G----G------...--T------------A---TC-A--------CA-CT---A------------G---GAA--C---------C-G-T-----AA--------TA--------CGGAGA---...--A---------.....---------- 274
G.CM.01.01CM_4049HAN    .......................................................................................................................................................................... 0
G.CU.99.Cu74            .......................................................................................................................................................................... 0
G.ES.99.X138            .......................................................................................................................................................................... 0
G.GH.03.03GH175G        T...----CAG--G----G--G---...--T----------C--A---TC-C--------CA-CT---A---------A--G--TGAA--C-A-------C-G-TA----A---------TA--------CGGAGA--T...C-T---------.....---------- 312
G.KE.93.HH8793_12_1     .......................................................................................................................................................................... 0
G.NG.01.01NGPL0669      .......................................................................................................................................................................... 0
G.PT.x.PT2695           T...----CAG--G----G------...--T------------A---T--A--------CA-CT---A------------G--TG-A--C-AG--G---C-G-TA-----A--------TA--------CGGAGA---...--A---------.....---------A 228
G.SE.93.SE6165          .......................................................................................................................................................................... 0
H.BE.93.VI991           .......................................................................................................................................................................... 0
H.BE.93.VI997           .......................................................................................................................................................................... 0
H.CF.90.056             .......................................................................................................................................................................... 0
J.CD.97.J_97DC_KTB147   .......................................................................................................................................................................... 0
J.SE.93.SE7887          .......................................................................................................................................................................... 0
K.CD.97.EQTB11C         .......................................................................................................................................................................... 0
K.CM.96.MP535           .......................................................................................................................................................................... 0
```

```
B.FR.83.HXB2                G...CCAGAGAAGTTAGAAGAAGCC...AACAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGAATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCAC...ATGGCCCGAGAG......CTGCATCCGG  312
01_AE.CF.90.90CF11697       T......A-AG-AG----G----A...--------AT--A-G--A-T--C----------CA------AA------G-A-----TGAA--A---------C-GAT-----A---------TTC-------CGAAGA--------A----------------G---A-  249
01_AE.CN.05.FJ051           C......A-AG-AGC------G-A....-C---------A----G-T--C----------CA------AA----------GG--TGAAC-CG---------C-GAT-----A---------TTC-------CGAA-A--------A----------A......AA-T-C--AC  312
01_AE.CN.06.FJ054           C..C---AGAG-CG----G--GAA....-C---------A-----TT--C------R---CA-A----A------CC-A--G--GAA--AG--------C-GAT-G---A---G-C---TGC-------CGAAGA-----.G-A----------.----G---A-  318
01_AE.HK.x.HK001            ..........................................................................................................................  0
01_AE.TH.01.01TH_R2184      ..........................................................................................................................  0
01_AE.TH.02.OUR769I         ..........................................................................................................................  0
01_AE.TH.90.CM240           ..........................................................................................................................  0
01_AE.US.00.00US_MSC1164    ..........................................................................................................................  0
02_AG.CM.02.02CM_4082STN    ..........................................................................................................................  0
02_AG.EC.x.ECU41            ..........................................................................................................................  0
02_AG.GH.03.GHNJ196         T.......CAG--G----G---------..--TG------------A-GT---A--------CA-CT---A---------------TGAA--C--------AC-G-TC----A---------------G-----CACA------.CA--AA----------.----------  312
02_AG.NG.01.PL0710          ..........................................................................................................................  0
02_AG.NG.x.IBNG             ..........................................................................................................................  0
02_AG.SE.94.SE7812          ..........................................................................................................................  0
02_AG.SN.98.MP1211          ..........................................................................................................................  0
03_AB.RU.97.KAL153_2        ..........................................................................................................................  0
04_cpx.CY.94.CY032          ..........................................................................................................................  0
05_DF.BE.x.VI1310           ..........................................................................................................................  0
06_cpx.AU.96.BFP90          T..-------G--G--------CTT...-CT------------A-T--C----------CA-AT---A------GCA--G--TGAA--A-----------A------A----------TC-------AGGAGA-------..--A-------------.AAA-------  312
08_BC.CN.97.97CNGX_6F       ..........................................................................................................................  0
09_cpx.GH.96.96GH29I1       ..........................................................................................................................  0
10_CD.TZ.96.96TZ_BF061      ..........................................................................................................................  0
11_cpx.GR.x.GR17            ..........................................................................................................................  0
12_BF.AR.99.ARMA159         C.......T-G--G-----A-G-------..--TG------------A-T-----C-------CA------AA----------A--TGAA--C--------AC-GATA----A---------------G--GC-AAGA------.T-A---A----------.AAA-----A-  312
13_cpx.CM.96.1849           ..........................................................................................................................  0
14_BG.DE.01.9196_01         ..........................................................................................................................  0
15_01B.TH.99.99TH_MU2079    ..........................................................................................................................  0
16_A2D.KR.97.97KR004        ..........................................................................................................................  0
17_BF.PE.02.PE02_PCR0155    ..........................................................................................................................  0
18_cpx.CU.99.CU76           ..........................................................................................................................  0
19_cpx.CU.99.CU7            ..........................................................................................................................  0
20_BG.CU.99.Cu103           ..........................................................................................................................  0
21_A2D.KE.91.KNH1254        ..........................................................................................................................  0
22_01A1.CM.01.01CM_0001BBY. ..........................................................................................................................  0
23_BG.CU.03.CB118           ..........................................................................................................................  0
24_BG.CU.03.CB378           ..........................................................................................................................  0
24_BG.CU.03.CB471           ..........................................................................................................................  0
25_cpx.CM.01.101BA          ..........................................................................................................................  0
27_cpx.CD.97.97CDKTB49      ..........................................................................................................................  0
28_BF.BR.99.BREPM12609      ..........................................................................................................................  0
29_BF.BR.02.BREPM119        ..........................................................................................................................  0
31_BC.BR.02.110PA           ..........................................................................................................................  0
32_06A1.EE.01.EE0369        ..........................................................................................................................  0
33_01B.MY.05.05MYKL007_1    ..........................................................................................................................  0
34_01B.TH.99.OUR2478P       ..........................................................................................................................  0
35_AD.AF.05.05AF095         ..........................................................................................................................  0
36_cpx.CM.00.00CMNYU830     ..........................................................................................................................  0
37_cpx.CM.00.00CMNYU926     ..........................................................................................................................  0
38_BF1.UY.05.UY05_4752      ..........................................................................................................................  0
39_BF.BR.03.03BRRJ103       ..........................................................................................................................  0
40_BF.BR.04.04BRRJ115       ..........................................................................................................................  0
42_BF.LU.03.luBF_05_03      ..........................................................................................................................  0
43_02G.SA.03.J11223         ..........................................................................................................................  0
U.CD.83.83CD003             ..........................................................................................................................  0
U.CD.90.90CD121E12          ..........................................................................................................................  0
U.GR.99.GR303               ..........................................................................................................................  0
U.NL.95.U_NL_95_H10986_D1   T......AGAG-AG-------G-------..--TG------G-----A-T--C----------CA------A-------------TG-A--C-----------AGA---GA----------TC-------CGAGTA------.CAA-------T........--A-------  207
CPZ.CD.90.ANT               ..........................................................................................................................  0
CPZ.CM.01.SIVcpzCAM13       ..........................................................................................................................  0
CPZ.CM.05.SIVcpzEK505       ..........................................................................................................................  0
CPZ.CM.05.SIVcpzLB7         ..........................................................................................................................  0
CPZ.CM.05.SIVcpzMB66        ..........................................................................................................................  0
CPZ.CM.05.SIVcpzMT145       ..........................................................................................................................  0
CPZ.GA.88.GAB1              A...GAG---C--G-----C---------..--TG-------T----A-T--C----G--T--CA-AT-T-A------G-----A--TGA---C-A---G--C--G-TC---C-C---------A-G--G--C--AAGA--T...--T---A----A......-AA-------  312
CPZ.TZ.00.TAN1              ..........................................................................................................................  0
CPZ.US.85.CPZUS             A...GA----G-AG----GCGG-------..T----------C--A--TAC--C-G-----CA--T---A-----------A--TGAAC-TG-C-----C---ATC---CA--------CTGAA--G--TCGGAGA------.GA--TAA----------.----------A-  312
CPZ.CM.x.SIVcpzDP943        ..........................................................................................................................  0
CPZ.CM.x.SIVcpzMB897        ..........................................................................................................................  0
N.CM.02.DJO0131             ..........................................................................................................................  0
N.CM.04.04CM_1015_04        ..........................................................................................................................  0
N.CM.04.04CM_1131_03        ..........................................................................................................................  0
N.CM.95.YBF30               ..........................................................................................................................  0
N.CM.97.YBF106              ..........................................................................................................................  0
O.BE.87.ANT70               A...GA---AG--GC----AG-CTAGGAG-T-C-T-T----GGG-T-ATC-CC-G--T--A-CAT-TGCC------T-T--A--TA-AC-T-A----A-AC-GAT-----A------T--ATCT----GCAACACC--T...G-T--TATGATA......ACT--C--A-  315
O.CM.91.MVP5180             A...G----AG--GC---GAG-CTGGGT--T-C-AAT--AG-TG-T--TC-TC----T--A-CTT-TAAT------GCT--G--TG-AC-CG-G--GA-AC--A-A---CA------T--ATCAT---GC--AACA--T....-A----TGC-A......AA---C--A-  315
O.CM.96.96CMABB637          ..........................................................................................................................  0
O.CM.98.98CMA104            ..........................................................................................................................  0
O.CM.99.99CMU4122           ..........................................................................................................................  0
O.FR.92.VAU                 ..........................................................................................................................  0
O.SN.99.SEMP1299            A...GA--CTG--GC---G---CTAGGA-----GT-T----GGG-T---C-CC-G--T--A-CTT--AAC-----CT-T--A---AACC-CG-GC--A-AC-GA-A---CA------T--ATCA----GCAGCACC--T...G-T--TATG-TA......ACCA-C--A-  315
O.US.99.99USTWLA            ..........................................................................................................................  0
```

```
                                 TCF-1 alpha                                                
```

```
                                     TCF-1 alpha                                       NF-κ-B-II                      NF-κ-B-I      Sp1-III      Sp1-II    Sp1-I
B.FR.83.HXB2                  AGTACTTCAAG............AACTGCTGACATC.........................GAGCTTGCTACAAGGG......ACTTTCCG.................CTGGGGACTTTCCAG...GGAGGCGTGGCCTGG.GCGGGACTG.GGGAGTGGCGAGCCC.  414
01_AE.CF.90.90CF11697         ------A---A............G----------GA.........................--AG--T---ACCA--......----.---.................---------------..G-----T----TT---.----AGT--.---------T-A---.  351
01_AE.CN.05.FJ051             ---T--AT--A............G----------AA.........................--AG--T---ACTA-A......----.---.................---------------..G-----T------G--.----AGT--.---------TT----.  414
01_AE.CN.06.FJ054             ------AT--A............G----------AT.........................--AGC-T---ACT---......----.---.................------T------C-..G-----T------GC-.----AGT--.---------T-A---.  420
01_AE.HK.x.HK001              ........................................................................................................................................................................  0
01_AE.TH.01.01TH_R2184        ........................................................................................................................................................................  0
01_AE.TH.02.OUR769I           ........................................................................................................................................................................  0
01_AE.TH.90.CM240             ........................................................................................................................................................................  0
01_AE.US.00.00US_MSC1164      ........................................................................................................................................................................  0
02_AG.CM.02.02CM_4082STN      ........................................................................................................................................................................  0
02_AG.EC.x.ECU41              ........................................................................................................................................................................  0
02_AG.GH.03.GHNJ196           ---TT-A---A............G----T-----CAGAAGT.........TGCTGACAG..-GA---T--G-CT---......--------.................---------------...-----T----TG---.-A--AGT--.---------T-A---.  428
02_AG.NG.01.PL0710            ........................................................................................................................................................................  0
02_AG.NG.x.IBNG               ........................................................................................................................................................................  0
02_AG.SE.94.SE7812            ........................................................................................................................................................................  0
02_AG.SN.98.MP1211            ........................................................................................................................................................................  0
03_AB.RU.97.KAL153_2          ........................................................................................................................................................................  0
04_cpx.CY.94.CY032            ........................................................................................................................................................................  0
05_DF.BE.x.VI1310             ........................................................................................................................................................................  0
06_cpx.AU.96.BFP90            ---TT-A---A............G----------AGGC............TGCTGACAAA.--AG--T---ATG---......--------.................---------------...-----T---------.-----G---.---------T-A---.  426
08_BC.CN.97.97CNGX_6F         ........................................................................................................................................................................  0
09_cpx.GH.96.96GH2911         ........................................................................................................................................................................  0
10_CD.TZ.96.96TZ_BF061        ........................................................................................................................................................................  0
11_cpx.GR.x.GR17              ........................................................................................................................................................................  0
12_BF.AR.99.ARMA159           ---GG-AT--A............G----AGACTGCTGACACAGAGAC....TGCTGACACA.--AGAAT---A-G--......--------.................---------------...A-G-CG-GCCAGA--.---T-----.---------TCA---.  434
13_cpx.CM.96.1849             ........................................................................................................................................................................  0
14_BG.DE.01.9196_01           ........................................................................................................................................................................  0
15_01B.TH.99.99TH_MU2079      ........................................................................................................................................................................  0
16_A2D.KR.97.97KR004          ........................................................................................................................................................................  0
17_BF.PE.02.PE02_PCR0155      ........................................................................................................................................................................  0
18_cpx.CU.99.CU76             ........................................................................................................................................................................  0
19_cpx.CU.99.CU7              ........................................................................................................................................................................  0
20_BG.CU.99.Cu103             ........................................................................................................................................................................  0
21_A2D.KE.91.KNH1254          ........................................................................................................................................................................  0
22_01A1.CM.01.01CM_0001BBY    ........................................................................................................................................................................  0
23_BG.CU.03.CB118             ........................................................................................................................................................................  0
24_BG.CU.03.CB378             ........................................................................................................................................................................  0
24_BG.CU.03.CB471             ........................................................................................................................................................................  0
25_cpx.CM.01.101BA            ........................................................................................................................................................................  0
27_cpx.CD.97.97CDKTB49        ........................................................................................................................................................................  0
28_BF.BR.99.BREPM12609        ........................................................................................................................................................................  0
29_BF.BR.02.BREPM119          ........................................................................................................................................................................  0
31_BC.BR.02.110PA             ........................................................................................................................................................................  0
32_06A1.EE.01.EE0369          .....................................................................................................................-------..-----T---------.-----G---.-----------A---.  46
33_01B.MY.05.05MYKL007_1      ........................................................................................................................................................................  0
34_01B.TH.99.OUR2478P         ........................................................................................................................................................................  0
35_AD.AF.05.05AF095           ........................................................................................................................................................................  0
36_cpx.CM.00.00CMNYU830       ........................................................................................................................................................................  0
37_cpx.CM.00.00CMNYU926       ........................................................................................................................................................................  0
38_BF1.UY.05.UY05_4752        ........................................................................................................................................................................  0
39_BF.BR.03.03BRRJ103         ........................................................................................................................................................................  0
40_BF.BR.04.04BRRJ115         ........................................................................................................................................................................  0
42_BF.LU.03.luBF_05_03        ........................................................................................................................................................................  0
43_02G.SA.03.J11223           ........................................................................................................................................................................  0
U.CD.83.83CD003               ........................................................................................................................................................................  0
U.CD.90.90CD121E12            ........................................................................................................................................................................  0
U.GR.99.GR303                 ........................................................................................................................................................................  0
U.NL.95.U_NL_95_H10986_D1     ------A---A............G--------T-AA.........................--AG--T---ACT---......--------.................---------------..-----T--TA-----.-----GT--.---------T-A---.  309
CPZ.CD.90.ANT                 ........................................................................................................................................................................  0
CPZ.CM.01.SIVcpzCAM13         ........................................................................................................................................................................  0
CPZ.CM.05.SIVcpzEK505         ........................................................................................................................................................................  0
CPZ.CM.05.SIVcpzLB7           ........................................................................................................................................................................  0
CPZ.CM.05.SIVcpzMB66          ........................................................................................................................................................................  0
CPZ.CM.05.SIVcpzMT145         ........................................................................................................................................................................  0
CPZ.GA.88.GAB1                ------A---A.GACTGACTTTAG----GC-CATG-GCACAAGAAC....TGCTGACT..........--G-GG-.......------A...................A---------G-.GGA----T-GTCGGGA-.-----GT--.---T-----TTTG--C  435
CPZ.TZ.00.TAN1                ........................................................................................................................................................................  0
CPZ.US.85.CPZUS               --------CG-.....AACTGATGG---TAAC-GCGCAGGCGCAATAAAACTGCTGACTGA.-GA---T---AGG-.......------A...................A-----G-----A....-G--GTG--T-A--.----A--A-.--CGTG-TTTTA---.  441
CPZ.CM.x.SIVcpzDP943          ........................................................................................................................................................................  0
CPZ.CM.x.SIVcpzMB897          ........................................................................................................................................................................  0
N.CM.02.DJO0131               ........................................................................................................................................................................  0
N.CM.04.04CM_1015_04          ........................................................................................................................................................................  0
N.CM.04.04CM_1131_03          ........................................................................................................................................................................  0
N.CM.95.YBF30                 ........................................................................................................................................................................  0
N.CM.97.YBF106                ........................................................................................................................................................................  0
O.BE.87.ANT70                 --CT----C--............--GGA--A-A-A-..............TGCTGACCTG.A--A-----GAC-CT-TGGA..-------AGCAAAGACTGCTGACACTGC-------------...T-G-AG-GACAGG--.----TT-G..---------T-A---.  448
O.CM.91.MVP5180               --CT----CCC...........A-GTAA------C...............TGCGGG...........................-------A....GACTGCTGACACTGC-------------.CGT-G-AG-GATAAG--.----TT-G..---------T-A---.  421
O.CM.96.96CMABB637            ........................................................................................................................................................................  0
O.CM.98.98CMA104              ........................................................................................................................................................................  0
O.CM.99.99CMU4122             ........................................................................................................................................................................  0
O.FR.92.VAU                   ........................................................................................................................................................................  0
O.SN.99.SEMP1299              --CT---T--C.............--GGA--A-A-C..............TGCTGACCTG.A--A-----GAC-CT-CGGG..-------AGCAGAGACTGCTGACACGGC-------------..T-TG--A-G-A-AG--.----TT-G..---------T-A---.  448
O.US.99.99USTWLA              ........................................................................................................................................................................  0
```


+1 mRNA start site
TAR element start
TAR element end
TATA Box
5' LTR U3 end
5' LTR R repeat begin
Poly-A signal
5' LTR R repeat end
5' LTR U5 start
Extensive secondary structure
B.FR.83.HXB2
TCAGATCCTGCATATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTA..GACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTA....GGGAACCCACTGCTTAAG.CCTCAATAA.AGCTTGCCTTGA.GTGCTT...CAAGTAGTGTGTGCCCGTCTG.T 571
A1.GE.99.99GEMZ011 0
A1.KE.00.KER2008 0
A1.KE.00.KNH1144 0
A1.KE.00.KSM4024 0
A1.KE.00.MSA4069 0
A1.KE.00.NKU3005 0
A1.RU.00.RU00051 0
A1.RW.93.93RW_024 0
A1.SE.95.SE889T 0
A1.SE.95.UGSE8131 0
A1.TZ.01.A173 0
A1.UA.01.01UADN139 0
A1.UG.92.92UG037 571
A1.UG.99.99UGA07072 0
A2.CD.97.97CDKTB48 0
A2.CY.94.94CY017_41 0
A.SN.01.DDI579 0
A.SN.01.DDJ369 0
A.SN.96.DDJ360 0
A.CD.02.02CD_KTB035 572
A.CD.97.97CD_KCC2 572
A.CD.97.97CD_KTB13 0
B.AR.04.04AR151516 0
B.AU.87.MBC925 572
B.BO.99.BOL0122 0
B.BR.03.BREPM2012 0
B.CA.97.CANB3FULL 0
B.CN.05.05CNHB_hp3 570
B.CO.01.PCM001 0
B.GB.83.CAM1 572
B.GB.86.GB8 569
B.GE.03.03GEMZ010 0
B.IT.05.SG1 370
B.JP.05.DR6538 573
B.KR.05.05CSR3 596
B.NL.00.671_00T36 93
B.RU.04.04RU128005 57
B.TH.00.00TH_C3198 0
B.UA.01.01UAKV167 0
B.US.04.ES10_53 571
B.US.99.PRB959_03 0
C.AR.01.ARG4006 0
C.BR.04.04BR013 32
C.BW.00.00BW07621 0
C.CN.98.YNRL9840 0
C.ET.02.02ET_288 0
C.GE.03.03GEMZ033 0
C.IL.99.99ET7 0
C.IN.99.01IN565_10 0
C.KE.00.KER2010 0
C.MM.99.mIDU101_3 0
C.MW.93.93MW_965 0
C.SN.90.90SE_364 0
C.SO.89.89SM_145 0
C.TZ.02.CO178 0
C.UY.01.TRA3011 0
C.YE.02.02YE511 0
C.ZA.04.04ZASK164B1 5
C.ZA.05.05ZAPSK240B1 5
C.ZM.02.02ZM108 571
D.CD.83.ELI 117
D.CM.01.01CM_0009BBY 0
D.KE.01.NKU3006 0
D.KR.04.04KBH8 528
D.TD.99.MN011 0
D.TZ.01.A280 0
D.UG.99.99UGK09259 0
D.YE.01.01YE386 0
D.YE.02.02YE516 0
D.ZA.90.R1 0
F1.AR.02.ARE933 0
F1.BE.93.VI850 0
F1.BR.01.01BR125 27
F1.ES.x.P1146 0
F1.FI.93.FIN9363 0
F2.CM.02.02CM_0016BBY 0
F2.CM.95.MP257 0
F2.CM.97.CM53657 0
G.BE.96.DRCBL 537
G.CM.01.01CM_4049HAN 0
G.CU.99.Cu74 165
G.ES.99.X138 3
G.GH.03.03GH175G 594
G.KE.93.HH8793_12_1 0
G.NG.01.01NGPL0669 0
G.PT.x.PT2695 501
G.SE.93.SE6165 0
H.BE.93.VI991 0
H.BE.93.VI997 0
H.CF.90.056 0
J.CD.97.J_97DC_KTB147 0
J.SE.93.SE7887 0
K.CD.97.EQTB11C 0
K.CM.96.MP535 0



HIV-1/SIVcpz
Complete Genomes

```
                            Extensive secondary structure                    Lys tRNA primer binding site
                                                           5' LTR U5 end                                                                    Packaging loops begin
B.FR.83.HXB2                TGTGTG..ACTCTGGTAAC.TAGAGATCCCTCAGACCC..TTTTAGTCAGTGTGGAAAA..TCTCTAGC.....AGTGGCGCCCGAACAGGGACCTGAAAG..........................CGAAAGGGAAACCAGAGGAGC...TCTCTCGAC...GCAGGA  696
A1.GE.99.99GEMZ011          .........................................................................................................................................................................    0
A1.KE.00.KER2008            .........................................................................................................................................................................    0
A1.KE.00.KNH1144            .........................................................................................................................................................................    0
A1.KE.00.KSM4024            .........................................................................................................................................................................    0
A1.KE.00.MSA4069            .........................................................................................................................................................................    0
A1.KE.00.NKU3005            .........................................................................................................................................................................    0
A1.RU.00.RU00051            .........................................................................................................................................................................    0
A1.RW.93.93RW_024           .........................................................................................................................................................................    0
A1.SE.95.SE8891             .........................................................................................................................................................................    0
A1.SE.95.UGSE8131           ........................................................AAA.....--------...--------------------C----AGTGAAA...GTAATAGGGACTCGAAAG------T...T------A-T...---------..------   94
A1.TZ.01.A173               .........................................................................................................................................................................    0
A1.UA.01.01UADN139          .........................................................................................................................................................................    0
A1.UG.92.92UG037            -T----..-----------.-----------------A..C-C---A-T-A--AA---..--------...--------------------T------.CGAAA...GTAATAGGGACTCGAAAA------T...T------A--AAGC-----T---..------  718
A1.UG.99.99UGA07072         .........................................................................................................................................................................    0
A2.CD.97.97CDKTB48          ....................................................................................................................................TTTTTCGAAG---.--TA...----G-A--T..-------------..---A--   42
A2.CY.94.94CY017_41         ..........................................................................TTGA-AA-CGAAAGT-------...........................ACTCGAAAG------T...T---------T...---------..------   64
A.SN.01.DDI579              .........................................................................................................................................................................    0
A.SN.01.DDJ369              .........................................................................................................................................................................    0
A.SN.96.DDJ360              .........................................................................................................................................................................    0
A.CD.02.02CD_KTB035         ------..C--T-------.--AG---------C---.C----AAAGAA---A---..------C-.....-T--------A-C-------TC-----......................------T...TA-----A--T...---------..------  693
A.CD.97.97CD_KCC2           ------..-----------.-----------------A...CC-TAG---A---A----..--------...--------------------TC-----CGAAAG..................TA----T...T------A--T...---------..------  699
A.CD.97.97CD_KTB13          .........................................................................................................................................................................    0
B.AR.04.04AR151516          .........................................................................................................................................................................    0
B.AU.87.MBC925              ------..-----------.------------------..-------.------------------..--------...--------------------T--T---.........................------A-------------..---------..------  697
B.BO.99.BOL0122             .........................................................................................................................................................................    0
B.BR.03.BREPM2012           ..............................................------------------..--------...------------------..........................ACTTGAAAG------TA--G---------A...---------..------   90
B.CA.97.CANB3FULL           .........................................................................................................................................................................    0
B.CN.05.05CNHB_hp3          ------..-----------.------------------...------TT----------..--------...--------------------GC-----.........................------T--G---------A...---------..------  694
B.CO.01.PCM001              .........................................................................................................................................................................    0
B.GB.83.CAM1                ------..-----------.------------------..-------.------------------..--------...--------------------T-----A.......................------TAG----G--------..---------..------  697
B.GB.86.GB8                 ------..---------G-.------------------..-------------A-----..--------...--------------------T------.....................T-----TTG------------..-------------..---A-G  694
B.GE.03.03GEMZ010           .........................................................................................................................................................................    0
B.IT.05.SG1                 ------..-----------.------------------..-------.--G----------------..--------...---------------------------.........................------A-------------..---------..------  495
B.JP.05.DR6538              ------..-----------.---------C--------..-------.------------------..--------...--------------------T------.....................T-----TAG------------..---------..------  698
B.KR.05.05CSR3              ------..-----------.----------------AA..------------------..--------...--------------------T------.........................-------A------------..---------..------  721
B.NL.00.671_00T36           -A----..-----------.------------------..-------.------------------..--------...------------------..........................ACTTGAAGA...---TA-----------A...---------..------  215
B.RU.04.04RU128005          ------..-----------.------------------..-------.-------------R-----..--------...------------------..........................ACYYGAAAG------A-------------..---------..------  182
B.TH.00.00TH_C3198          .........................................................................................................................................................................    0
B.UA.01.01UAKV167           .........................................................................................................................................................................    0
B.US.04.ES10_53             ------..-----------.------------------..-------.------------------..--------...--------------------AC-----.........................------A---------A--A...---------..------  696
B.US.99.PRB959_03           .........................................................................................................................................................................    0
C.AR.01.ARG4006             .........................................................................................................................................................................    0
C.BR.04.04BR013             ------..---------G-.--------------TA-..C---T-------------..--------...------------------..........................ACTTGAAAG------TAG----------A...---------..------  157
C.BW.00.00BW07621           ......................................................................--------...------------------..........................ACTTGAAAGT-----TA-G---------A...---------..------   71
C.CN.98.YNRL9840            .........................................................................................................................................................................    0
C.ET.02.02ET_288            .........................................................................................................................................................................    0
C.GE.03.03GEMZ033           .........................................................................................................................................................................    0
C.IL.99.99ET7               .........................................................................................................................................................................    0
C.IN.99.01IN565_10          .................................................AAAAA.....--------...------------------..........................ACCCGAAAG------TA-G------A--A...---------..------   76
C.KE.00.KER2010             .........................................................................................................................................................................    0
C.MM.99.mIDU101_3           ...................................................AAA.....--------...------------------..........................ACCTGAAAG------TA-G---------A...---------..------   74
C.MW.93.93MW_965            .........................................................................................................................................................................    0
C.SN.90.90SE_364            .........................................................................................................................................................................    0
C.SO.89.89SM_145            .........................................................................................................................................................................    0
C.TZ.02.CO178               .........................................................................................................................................................................    0
C.UY.01.TRA3011             .........................................................................................................................................................................    0
C.YE.02.02YE511             .........................................................................................................................................................................    0
C.ZA.04.04ZASK164B1         ------..-----------.------------------..--G-T-GT----------..--------...------------------..........................ACGCGAAAG------T--G------A--A...---------..------  130
C.ZA.05.05ZAPSK240B1        ------..G----------.------------------....---T-GT----------..--------...------------------..........................ACCCGAAAG------TA-G---------A...---------..---A--  129
C.ZM.02.02ZM108             ------..-----------.-----------------T..----T-GT----------..--------...--------------------GC-----.........................------TAG----------A...A--------..------  696
D.CD.83.ELI                 ------..-----------.------------------..C--------A--------..--------...------------------..........................ACCTGAAAG------TAG------------...---------..------  242
D.CM.01.01CM_0009BBY        .........................................................................................................................................................................    0
D.KE.01.NKU3006             .........................................................................................................................................................................    0
D.KR.04.04KBH8              ----A..-----------.-----------------T..CG--T--------AA----..-----C--...--------------------T------.........................------TAG------------...----C----..------  653
D.TD.99.MN011               .........................................................................................................................................................................    0
D.TZ.01.A280                .........................................................................................................................................................................    0
D.UG.99.99UGK09259          .........................................................................................................................................................................    0
D.YE.01.01YE386             .........................................................................................................................................................................    0
D.YE.02.02YE516             .........................................................................................................................................................................    0
D.ZA.90.R1                  .....................................................................................................................TCGGCTGGAA---.--T...TAAGTGA-GAG...-------C.....------   40
F1.AR.02.ARE933             ................................................................................................................................................-T.......------    8
F1.BE.93.VI850              ....................................................................................................................GGACTTG.AAG------TAG--------A--A...-----------..--G--G   46
F1.BR.01.01BR125            ---A--..-----------.--A-----------A-A..--------T---------A.--------...------------------..........................ACCTGAAAG------TAG--------A--A...---------..------  153
F1.ES.x.P1146               .........................................................................................................................................................................    0
F1.FI.93.FIN9363            ..................................................................................................................................AAAG------TAG--------A--A...---------..------   40
F2.CM.02.02CM_0016BBY       .........................................................................................................................................................................    0
F2.CM.95.MP257              .........................................................................................................................................................................    0
F2.CM.97.CM53657            .........................................................................................................................................................................    0
G.BE.96.DRCBL               ------..-----------.-----------------A..C-C---AT-----AA---..--------...--------------------T-----ACGAAAG...TTAATAGGGACTCGAAAG------T...T------A--T...---------..------  682
G.CM.01.01CM_4049HAN        .........................................................................................................................................................................    0
G.CU.99.Cu74                ------..-----------.-----------------A..C-C---A-G----AA---..--------...--------------------T------CGAAAG...TTAACAGGGACTCGAAAG------T...T------A--T...---------..------  310
G.ES.99.X138                ..----..-----------.-----------------A..C-C---AT-----AA---..--------...--------------------T------CGAAAG.....TTMACAGGSACTCGAAAGCGAAAGTT------A--T...---------..------  146
G.GH.03.03GH175G            ------..-----------.-----------------A..C-C---A------AA---..--------...---------------------C-----TGAAAG...TTAATAGGGACTCGAAAG------T...T------A--T...---------..------  739
G.KE.93.HH8793_12_1         ................................................G-CTT---..--------...------------------.A--TG--AGTGAAG...TTAACAGGGACTCGAAAG------T...T------A--A...---------..------   99
G.NG.01.01NGPL0669          .........................................................................................................................................................................    0
G.PT.x.PT2695               ------..-----------.-----------------A..C-C---AT----AA---..-------A....--------------------T------CGAAAG...TTAAACGGGACTCGAAAG------T...T------A--T...---------..------  646
G.SE.93.SE6165              ................................................G-CTT---..--------...--------------------T------CGAAAG...TTAACAGGGACTCGAAAG------T...T------A--T...---------..------  100
H.BE.93.VI991               ...................................................AAA.....--------...------------------..........................ACTTGAAAG------CA---------A---..---------..------   74
H.BE.93.VI997               ...............................................................................................................................................---------..------   15
H.CF.90.056                 ..............................................................................................................................TTGATAGT-----TA---------A--A...---------..------   43
J.CD.97.J_97DC_KTB147       .........................................................................................................................................................................    0
J.SE.93.SE7887              ...........................................................................................................................................T---------..------   16
K.CD.97.EQTB11C             .........................................................................................................................................................................    0
K.CM.96.MP535               .........................................................................................................................................................................    0
```

```
                          Extensive secondary structure                                         Lys tRNA primer binding site
                                                                            5' LTR U5 end                                                                                  Packaging loops begin
B.FR.83.HXB2              TGTGTG..ACTCTGGTAAC.TAGAGATCCCTCAGACCC..TTTTAGTCAGTGTGGAAAA..TCTCTAGC.....AGTGGCGCCCGAACAGGGACCTGAAAG..........................CGAAAGGGAAACCAGAGGAGC...TCTCTCGAC...GCAGGA  696
01_AE.CF.90.90CF11697     GA-AG-..-----------.----------------A..C-C---A-T-A--AA---..---------....-------------------T------CGAAAG...TTAATAGGGACTCGAAAG------T...T------A--T...---------..------  653
01_AE.CN.05.FJ051         GT-AG-..-----------.----------------A..C-C---A-T-A--AA---..---------....-------------------T------TGAAAG...TTAATAGGGACTCGAAAG--G---T...T------A--A...---------..------  716
01_AE.CN.06.FJ054         GT-AG-..-----------.---------------A...C-GC--M-T-A--AA---..---------....-------------------T------CGAAAG...TTAATAGGGACTCGAAAG--G---T...T------A--T...---------..------  721
01_AE.HK.x.HK001          .....................................................................-----------T------CGAAAG...TTAATAGGGACTCGAAAG------T...T------A--T...--------GC..------   76
01_AE.TH.01.01TH_R2184    ...............................................................................................................................................................    0
01_AE.TH.02.OUR769I       ...............................................................................................................................................................    0
01_AE.TH.90.CM240         GT-AG-..-----------.----------------A..C-C---A-T-A--AA---..------C-....------------------CA--CG--AGCGAAA..GTTAATAGGGACTCGAAAG------T...T------A--T...---------..------  264
01_AE.US.00.00US_MSC1164  ...............................................................................................................................................................    0
02_AG.CM.02.02CM_4082STN  ...............................................................................................................................................................    0
02_AG.EC.x.ECU41          ...............................................................................................................................................................    0
02_AG.GH.03.GHNJ196       ------..-----------.----------------A..A-------TTG--AA---..---------....--------------------G---G.........TTAATAGGGACGCGAAAG------T...T------A--.....-A-----------.----C-G  721
02_AG.NG.01.PL0710        ...............................................................................................................................................................    0
02_AG.NG.x.IBNG           ------..-----------.----------------A..C-C---A-T----AA---..---------....-----------------.--T---CG.........GTAATAGGGACTCGAAAG------T...T------A--A...----------..---A-G  238
02_AG.SE.94.SE7812        .........................................---------....-------------------T----G.........TTAATAGGGACTCGAAAG------T...T------A--A...---------..-----G   85
02_AG.SN.98.MP1211        ...............................................................................................................................................................    0
03_AB.RU.97.KAL153_2      ...............................................................................................................................................................    0
04_cpx.CY.94.CY032        ..........................................................................................T------TGAAAG....TTAATAGGACTCGAAAG------T...T------A--T...---------..------   63
05_DF.BE.x.VI1310         ......................A.....---------....------------------..........................ACTTGAAAA------TAG--------A--A...A--------..------   72
06_cpx.AU.96.BFP90        ------..-----------.----------------A..C-C---A------AA---..---------....-------------------T------CGAAAG...TTAATAGGGACTCGAAAG------T...T------A--T...---------..------  728
08_BC.CN.97.97CNGX_6F     ...............................................................................................................................................................    0
09_cpx.GH.96.96GH2911     ...............................................................................................................................................................    0
10_CD.TZ.96.96TZ_BF061    .....................CC.....---------....------------------T.........................ACCTGAAAG------TAG-----------A...---------..------   74
11_cpx.GR.x.GR17          .....................................................................................................................................--T---..------   12
12_BF.AR.99.ARMA159       ------..-----------.----------------A..--C---------------..---------....-------------------T------.......................------TAG-----------A...---------..------  716
13_cpx.CM.96.1849         ....................AAA.....---------....------------------..........................GACGTGAAAGCG--A-CT--------A--T...---------..------   73
14_BG.DE.01.9196_01       ------..-----------.----------------A..C-C--AAT-----AA--...---------....------------------.GA--TG--AGCGAAA...GTAACAGGGACTCGAAAG------T...T------A--T...---------..------  224
15_01B.TH.99.99TH_MU2079  .............TAATTCGCCCTT---..---------....---------------------T.--AGTGAAAG..TTAATAGGGACTCGAAAG------T...T------A--T...---------..------  107
16_A2D.KR.97.97KR004      ....................................................................C-T--AGTGAAA...GTAATAGGGACTTGAAAG------T...T------A--A...---------..------   63
17_BF.PE.02.PE02_PCR0155  ...............................................................................................................................................................    0
18_cpx.CU.99.CU76         ..............................................................................................................................---------..------   14
19_cpx.CU.99.CU7          ...............................................................................................................................................................    0
20_BG.CU.99.Cu103         ------..-----------.----------------A..C-C---A------AA---..---------....-------------------T------TGAAAG...TTAAAAGGGACTCGAAAG------T...T------A--T...---------..------  172
21_A2D.KE.91.KNH1254      ...............................................................................................................................................................    0
22_01A1.CM.01.01CM_0001BBY...............................................................................................................................................................    0
23_BG.CU.03.CB118         ------..-----------.----------------A..C-C---A------AA---..---------....-------------------T------CGAAAG...TTAAAAGGGACTCGAAAG--G---T...T------A--T...---------..------  146
24_BG.CU.03.CB378         ------..-----------.----------------A..C-C---A-R----AA---..---------....-------------------T------TGAAAG...TTAATAGGGACTCGAAAG------T...T------A--T...---------..------  146
24_BG.CU.03.CB471         ------..-----------.---------------AA..C-C---AT-----AA---..---------....--------------------------TGAAAG...TTAATAGGGACTCGAAAG------T...T------A--T...---------..------  146
25_cpx.CM.01.101BA        ...............................................................................................................................................................    0
27_cpx.CD.97.97CDKTB49    ...............................................................................................................................................................    0
28_BF.BR.99.BREPM12609    ..................................-C--....------------------..........................ACTTGAAAG------TAG------------..---------..------   67
29_BF.BR.02.BREPM119      ...............................................................................................................................................................    0
31_BC.BR.02.110PA         ------..-----------.---------------A-..C---W-------------..---------....------------------..........................ACTCGAAAG------TA-G----------A...---------..------  152
32_06A1.EE.01.EE0369      ------..-----------.--------------A-A..C-----A------AA---..---------....-------------------------ACGAAAG...TTAATAGGGACTCGAAAG------T...T------A--T...---------..------  348
33_01B.MY.05.05MYKL007_1  .........................................-------------------T-----ACGAAAG...TTAATAGGGACTCGAAAG------T...T------A--T...---------..------   84
34_01B.TH.99.OUR2478P     ...............................................................................................................................................................    0
35_AD.AF.05.05AF095       ...............................................................................................................................................................    0
36_cpx.CM.00.00CMNYU830   ...............................................................................................................................................................    0
37_cpx.CM.00.00CMNYU926   ...............................................................................................................................................................    0
38_BF1.UY.05.UY05_4752    ------..-----------.----------------..-----------------..---------....-------------------T------.......................------TAG--------A--A...---------..------  125
39_BF.BR.03.03BRRJ103     ------..-----------.----------------T...----------------..---------....--------------------C-----.......................------TAG-----------A...---------..------  170
40_BF.BR.04.04BRRJ115     ------..-----------.----------------A..------------------..---------....------------------..........................ACGCGAAAG------TAG-----------A...---------..------  234
42_BF.LU.03.luBF_05_03    ------..-----------.----------------A..Y-------C----------..---------....------------------..........................ACYTGAAAG------TAG------------..---------..------  228
43_02G.SA.03.J11223       ------..-----------.----------------A..C-C---A-G----AA---..---------....-------------------T------TGAAAG...TTAATAGGGACTCGAAAG--R---T...T------A--T...---------..------  232
U.CD.83.83CD003           ------..---------T-.----------------A..C-----GT-----------..---------....------------------..........................ACCGGAAAG------TA---------A--T...---------..------  194
U.CD.90.90CD121E12        ...............................................................................................................................................................    0
U.GR.99.GR303             ...............................................................................................................................................................    0
U.NL.95.U_NL_95_H10986_D1 ------..-----------.----------------A..------------------..---------....-------------------T------.......................------TAGG-------A--T...---------..------  591
CPZ.CD.90.ANT             ..........................................................---..........................ACTTGAGGAG.CT-CA------...............TGT---..GC-----   37
CPZ.CM.01.SIVcpzCAM13     -CA--T..---------T-.----------------TA..AAA--TAGTCCAG-A---..---------....------------------..........................ACCTGAAAG------TA-G---G----CTGA.A----C-----..---C-G  239
CPZ.CM.05.SIVcpzEK505     -CG--A..---------T-.----------------TA..A--CTACT.-A--AA---..---------....------------------..........................ACTTGAAAGT-----TA-..--G----CTGA.A--C----G-..-------  237
CPZ.CM.05.SIVcpzLB7       -T----..---------T-.---------------A-T...-G-T---T-C-C-------..---------....------------------..........................ACTTGAAAA------TA--.-TG---C--T...------------..--G--G  240
CPZ.CM.05.SIVcpzMB66      ------..---------T-.---------------TAA..C-G-----T---C--CC--AA--------....------------------..........................ACTTGAAAA------TA-G-------A--T...---------..------  243
CPZ.CM.05.SIVcpzMT145     -CA---..---------T-.---------------A-A..-C-ATA-TGA-AGAA---..---------....------------------..........................ACTTGAAAGT-----TA-----G----CTGA.A-----------..---C-G  242
CPZ.GA.88.GAB1            -CA-AC..-----------.---------------TTA..AA---TAGT.CAA------..---------....----------------------GT.......................------T..-----G----CTGA.A----C-----..---C-G  714
CPZ.TZ.00.TAN1            -ACCGCG.T--GCCC-GGGG---------------TTT..G-AGT-G-TAA--AA---..------C-....-------------------T---G-A.......................GC-GG-AACGCGGCCCCTG-A..........-.------  237
CPZ.US.85.CPZUS           C.A---..-----------.---------------A-T..C--A-AAGTAG---A---..---------....-------------------T-----A.......................------TAC---------CTGA.AA--TC---------..--G  720
CPZ.CM.x.SIVcpzDP943      -CA---..-----------.---------------A-A..AA---AGT.AG---A---..---------....-------------------T-----.......................T-----T-G---------CTGA.C---TC-------..---C-G  240
CPZ.CM.x.SIVcpzMB897      -T----..---------T-.---------------TAT..-AAGT---T---C------..---------....-------------------T-----A.......................------T-T----G---A--T...---------..------  240
N.CM.02.DJO0131           -CG--A..---------C-.----------------A...CC-G-A-T-A---A----..---------....------------------..........................ACTTGAAAA------TAG----G----CTGA.A-----------..------  168
N.CM.04.04CM_1015_04      -CG--A..---------C-.----------------A...C--G-G-T-A--A----..---------....------------------..........................ACTTGAAAA------TAG----G----CTGAAA---------..------  169
N.CM.04.04CM_1131_03      -CG--A..---------C-.----------------A...-C-G-A-T-A--A----..---------....------------------..........................ACTTGAAAA------TAG----G----CTGA.A-----------..---A-G  168
N.CM.95.YBF30             -CG--A..---------C-.----------------A...-C---A-T-A---A----..---------....------------------..........................ACTTGAAAA------TAG----G----CTGA.A-----------..------  256
N.CM.97.YBF106            -CG--A..---------C-.---------------TTA..CCA---A-T-A--AA---..------------....-S----------------..........................ACTTGAAAA------TAG----G----CTGA.A-----------..---A-G  256
O.BE.87.ANT70             -C.....A--C-----GT-.---------------T-A..C--AGACTGAA-CA-----.---------....-------------------T----G.......................T-----T--------G-A--AA.AA-CTC-----..---ACG  726
O.CM.91.MVP5180           -C.....A--C-----GT-.---------------T-A..C--AGACTGAA-CA-----.---------....-------------------GC-----.......................T-----T-G------G-A--AA.AA-CTC-----..---ACG  700
O.CM.96.96CMABB637        -C.....A--C------T-.---------------T-A..CG-AGACTGAA-CA-----.---------....------------------..............................A.CTTGA.---T--------G-A--AA.AA-CTC-----..---ACG  153
O.CM.98.98CMA104          -C.....A--C------G-.---------------T-A..CG-AGACTGAA-CA-----.---------....------------------..........................ACCGGAAAGA-----T--------G-A--AA.AA-CTC-----..---ACG  158
O.CM.99.99CMU4122         -C.....A--C------T-.---------------T-A..CG-AGAYTGAG-CA-----.---------....------------------..........................ACCTGAAAAG-----T--------G-A--AA.AA-CTC-----..---ACG  157
O.FR.92.VAU               -C.....A--C-----GT-.---------------T-A..C--AGACTGAA-CA-G---A.--------....------------------...........................AC------T--------G-A--AA.AA-CTC-----..---ACG  234
O.SN.99.SEMP1299          -C.....G--C------T-.---------------T-A..CG-AGACTGAG-CA-----.---------....--------------------G-----.......................A-----T--------G-A--AA.AA-CTC-----..---ACG  727
O.US.99.99USTWLA          -C.....A--C-----GT-.----A----------G-A..CG-AGCCTGAA-CA-----.---------....------------------.............................ACTT-----T-------AG-A--AA.AA-CTC-----..---AC-  148
```

Packaging loops end ⌄ Gag and Gag-Pol start

```
B.FR.83.HXB2               ..CTCGGCTTG.CTGAAG......CGCGCACGGCAAGAGGCGAG...GGGCGGCGA..CTGGTGAGTACGCCAAAAA.............TTTTGACTAGC......GGAGGCTAGAAGG..........................AGAGAGATGGGTGCGAGAGCG 804
Gag                                                                                                                                                                   _M__G__A__R__A_
A1.GE.99.99GEMZ011         ..............................................................................................................................................................--------- 9
A1.KE.00.KER2008           ..............................................................................................................................................................--------- 9
A1.KE.00.KNH1144           ..............................................................................................................................................................--------- 9
A1.KE.00.KSM4024           ..............................................................................................................................................................--------- 9
A1.KE.00.MSA4069           ..............................................................................................................................................................--------- 9
A1.KE.00.NKU3005           ..............................................................................................................................................................--------- 9
A1.RU.00.RU00051           ..---------.----G-......T--.---A------------...A--------..-Y------------T---GA...........T-----------......-------------..........................--------------------- 108
A1.RW.93.93RW_024          ..............................................................................................................................................................--------- 9
A1.SE.95.SE8891            ..............................................................................................................................................................--------- 9
A1.SE.95.UGSE8131          ..---------.------.....GT--A---A------------...A--------..-------------------...........TT-----------......-------------..........................--------------------- 200
A1.TZ.01.A173              ..............................................................................................................................................................--------- 9
A1.UA.01.01UADN139         ..............................................................................................................................................................--------- 9
A1.UG.92.92UG037           ..---------.----G-......T--A---A------------...A--------..-------------------...........TT-----------......-------------..........................--------------------- 823
A1.UG.99.99UGA07072        ..............................................................................................................................................................--------- 9
A2.CD.97.97CDKTB48         ..---------.----G-......T--A----------------...AA---A--G..--------------T----............T-----------......-------------..........................--------------------- 151
A2.CY.94.94CY017_41        ..---------.----G-......T--A------------------..G---------..--------------T--T-............T-----------......A------------..........................--------------------- 173
A.SN.01.DDI579             ............................................................................................................................................................----------- 11
A.SN.01.DDJ369             ............................................................................................................................................................----------- 11
A.SN.96.DDJ360             ............................................................................................................................................................----------- 11
A.CD.02.02CD_KTB035        ..---------.----G-......T--A---A------------...A--------..--------------T----............T-----------......A------------..........................--------------------- 801
A.CD.97.97CD_KCC2          ..---------.----G-......T--A---A------------...A--------..--------------T----.............--C--------......-------------..........................--------------------- 806
A.CD.97.97CD_KTB13         ...................................................................................................................................................-------------------- 20
B.AR.04.04AR151516         ............................................................................................................................................................----------- 11
B.AU.87.MBC925             ..---------.---T--......--------------------...---------..-------------------A............-----------......-------------..........................--------------------- 806
B.BO.99.BOL0122            ..............................................................................................................................................................--------- 9
B.BR.03.BREPM2012          ..---------.------......-------A------------...---------..---------------T---A............-----------......-------------..........................--------------------- 199
B.CA.97.CANB3FULL          .................................-----------...---------..-------------------.............-----------......-------------..........................--------------------- 84
B.CN.05.05CNHB_hp3         ..---------.------......-------A------------...---------..-------------------...........TT-----------......-------------..........................--------------------- 800
B.CO.01.PCM001             ..............................................................................................................................................................--------- 9
B.GB.83.CAM1               ..---------.------......-------A------------...---------..-------------------.........TTTA-----------......-------------..........................--------------------- 805
B.GB.86.GB8                A.---------.------......-----G--------------...---------..-------------------.............-----------......-------------..........................--------------------- 803
B.GE.03.03GEMZ010          ..............................................................................................................................................................--------- 9
B.IT.05.SG1                ..---------.------......-------A------------...---------..-------------------...........TT-----------......-------------..........................--------------------- 603
B.JP.05.DR6538             ..---------.----G.......--------------------...---------..-------------------A..........TT-----------......-------------..........................--------------------- 808
B.KR.05.05CSR3             ..---------.------......--------------------...---------..-C-----------------............T-----------......-------------..........................--------------------- 829
B.NL.00.671_00T36          ..---------.------AAGAAG-----G--------------...---------..--------------T----A............-----------......-------------..........................--------------------- 330
B.RU.04.04RU128005         ..---------.------......T------A------------...---------..-------------------A........TTTT-----------......-------------..........................--------R------------ 295
B.TH.00.00TH_C3198         ..............................................................................................................................................................--------- 9
B.UA.01.01UAKV167          ..............................................................................................................................................................--------- 9
B.US.04.ES10_53            ..---------.------......--------------------...---------..-C-----------------............T-----------......-------------..........................--------------------- 804
B.US.99.PRB959_03          ....................................................................................................................................................................--------------- 15
C.AR.01.ARG4006            ..............................................................................................................................................................--------- 9
C.BR.04.04BR013            ..---------.------......T--A-T--------------...A-------G..-------------------........TTTTA-----------......-------------..........................--------------------- 265
C.BW.00.00BW07621          ..---------.------......T--A-T--------------...A--------..-------------------.........TTTTA----------......-------------..........................--------------------- 181
C.CN.98.YNRL9840           ....................................................................................................................................................................--------------- 15
C.ET.02.02ET_288           ..............................................................................................................................................................--------- 9
C.GE.03.03GEMZ033          ..............................................................................................................................................................--------- 9
C.IL.99.99ET7              ..............................................................................................................................................................--------- 9
C.IN.99.01IN565_10         ..---------.------......T--A-T--------------...A-G------..-------------------........TTTTTA----------......-------------..........................--------------------- 185
C.KE.00.KER2010            ..............................................................................................................................................................--------- 9
C.MM.99.mIDU101_3          ..---------.------......T--A-T--------------...A-G------..-------------------........TTTTTA----------......-------------..........................--------------------- 183
C.MW.93.93MW_965           ..............................................................................................................................................................--------- 9
C.SN.90.90SE_364           ..............................................................................................................................................................--------- 9
C.SO.89.89SM_145           ..............................................................................................................................................................--------- 9
C.TZ.02.CO178              ..............................................................................................................................................................--------- 9
C.UY.01.TRA3011            ..............................................................................................................................................................--------- 9
C.YE.02.02YE511            ..............................................................................................................................................................--------- 9
C.ZA.04.04ZASK164B1        ..---------.------......T--A-T--------------...A-------G..-------------------........TTTTTA----------......-------------..........................--------------------- 239
C.ZA.05.05ZAPSK240B1       ..---------.------......T--A-T--------------...A--------..-------------------........TTTTTA----------......-------------..........................--------------------- 238
C.ZM.02.02ZM108            ..---------.------......T--A-T--------------...A-------G..-------------------.........TTTTA----------......-------------..........................----T---------------- 803
D.CD.83.ELI                ..---------.------......--------------------...----A----..-------------T-----............T-----------......-------------..........................--------------------- 350
D.CM.01.01CM_0009BBY       ..............................................................................................................................................................--------- 9
D.KE.01.NKU3006            ..............................................................................................................................................................--------- 9
D.KR.04.04KBH8             ..--T------.------......-------A------------...---TA----..-------------T-----...........TT-----------......A------------..........................G--A----------A------ 763
D.TD.99.MN011              .....................................................................................................................-............................--------------------- 22
D.TZ.01.A280               ..............................................................................................................................................................--------- 9
D.UG.99.99UGK09259         ..............................................................................................................................................................--------- 9
D.YE.01.01YE386            ..............................................................................................................................................................--------- 9
D.YE.02.02YE516            ..............................................................................................................................................................--------- 9
D.ZA.90.R1                 ..---------.------......-----G--------------...----A----..-------------T---TAA............-----------......--------A----..........................--------------------- 150
F1.AR.02.ARE933            ..---------.------......T--A----------------...A--------..-------------------.............-----------......A------------..........................--------------------- 113
F1.BE.93.VI850             A.---------.------T.....GCAA----------------...A--------..--------------G----.......TTTTTT-----------......-------------..........................--------------------- 159
F1.BR.01.01BR125           ..---------.------......T--A----------------...A-------G..-------------------.............-----------......A------------..........................--------------------- 258
F1.ES.x.P1146              ..---------.----G-......T--.-------.--------...A--------..-------------------............T-----------......-------------..........................--------------------- 102
F1.FI.93.FIN9363           ..---------.------......T--A----------------...A--------..-------------------............T-----------......-------------..........................--------------------- 145
F2.CM.02.02CM_0016BBY      ..............................................................................................................................................................--------- 9
F2.CM.95.MP257             .......................................................................................................................................................A--------------- 16
F2.CM.97.CM53657           ..............................................................................................................................................................--A------ 9
G.BE.96.DRCBL              ..---------.----G-......T--A---A------------...A--------..-------------------............T-----------......-------------..........................--------------------- 789
G.CM.01.01CM_4049HAN       ..............................................................................................................................................................--------- 9
G.CU.99.Cu74               ..---------.----G-......T--A---A------------...A--------..-------------------...........CT-----------......-------------..........................--------------------- 417
G.ES.99.X138               ..---------.------.....GT--R---A------------...A--------..---------------T---.............-----------......-------------..........................--------------------- 254
G.GH.03.03GH175G           ..---------.------......T--A---A------------...A--------..-------------------..........TCC-----------......-------------..........................--------------------- 847
G.KE.93.HH8793_12_1        ..---------.----G-......T--A-G-A------------...A--------..-----------------C-............T-----------......-------------..........................--------------------S 206
G.NG.01.01NGPL0669         ..............................................................................................................................................................--------- 9
G.PT.x.PT2695              ..---------.------.....GT--A---A------------...A-------G..---------------T---A............-----------......-------------..........................--------------------- 755
G.SE.93.SE6165             ..---------.----G-......T--A---A------------...A--------..-------------------............T-----------......-------------..........................G-------------------- 207
H.BE.93.VI991              ..---------.------......T--A---A------------...--------TGA-------------G-----.......TTTTTA-----------......-------------..........................--------------------- 190
H.BE.93.VI997              ..---------.------......T--A-----------------.GG-C-GC----..--------------G----.......TTTTTA-----------......-------------..........................--------------------- 131
H.CF.90.056                ..---------.------......T--A---A------------...A--------..-------------------........TTTTG-----------......-------------..........................--------------------- 151
J.CD.97.J_97DC_KTB147      .....................................................................................................................................................................-- 2
J.SE.93.SE7887             ..---------.------......T--A----------------...---------..------------------TA...........T-----------......-------------..........................--------------------- 127
K.CD.97.EQTB11C            ........................................................................................................................................................--------------T 15
K.CM.96.MP535              ........................................................................................................................................................--------------- 15
```

```
                                                                                                                                     Packaging loops end Gag and Gag-Pol start
B.FR.83.HXB2                 ..CTCGGCTTG.CTGAAG......CGCGCACGGCAAGAGGCGAG...GGGCGGCGA..CTGGTGAGTACGCCAAAAA.............TTTTGACTAGC......GGAGGCTAGAAGG...........................AGAGAGATGGGTGCGAGAGCG  804
Gag                                                                                                                                                                                    _M__G__A__R__A_
01_AE.CF.90.90CF11697        ..---------.----G-......T--T---A------------...A--------..-------------------.............TT--C------......-------------...........................---------------------  759
01_AE.CN.05.FJ051            ..---------.----G-......T--A---A------------...A--------..-------------------.............-----------......-------------...........................---------------------  820
01_AE.CN.06.FJ054            ..---------.----G-......T--A---A------------...A--------..-------------------.............-----------......-------------...........................---------------------  826
01_AE.HK.x.HK001             ..---------.----G-......T--A---A------------...A--------..-------------------.............-----------......-------------...........................---------------------  181
01_AE.TH.01.01TH_R2184       ...............................................................................................................................................................---------    9
01_AE.TH.02.OUR769I          ...............................................................................................................................................................---------    9
01_AE.TH.90.CM240            ..---------.----G-......T--A---A------------...A--------..-------------------.............-----------......-------------...........................---------------------  369
01_AE.US.00.00US_MSC1164     ...............................................................................................................................................................---------    9
02_AG.CM.02.02CM_4082STN     ...............................................................................................................................................................---------    9
02_AG.EC.x.ECU41             ........................................................................................................................................................................    0
02_AG.GH.03.GHNJ196          A.---------.----G-......T--A---A------------...A--------..-------------------.............T----------......-------------...........................---------------------  827
02_AG.NG.01.PL0710           ...............................................................................................................................................................---------    9
02_AG.NG.x.IBNG              A.---------.----G-......T-----A-------------...A--------..-------------------.............-----------......-------------...........................---------------------  344
02_AG.SE.94.SE7812           A.---------.----G-......T--A---A------------...A--------..-------------------.............T----------......-------------...........................---------------------  191
02_AG.SN.98.MP1211           ...................................................................................................................................................---.------------   16
03_AB.RU.97.KAL153_2         ..........................................................................................................--......-------------...........................---------------------   36
04_cpx.CY.94.CY032           ..---------.----G-......T--A---A------------...A--------..-------------------.............TT---------......-------------...........................---------------------  170
05_DF.BE.x.VI1310            ..---------.------......------A-------------..---TA----.-------------T-----.............T----------......-------------...........................---------------------  180
06_cpx.AU.96.BFP90           ..---------.----G-......T--A---A------------...A--------..-------------------.............T----------......A------------...........................---------------------  832
08_BC.CN.97.97CNGX_6F        .............................................................................................................................................................-----------   11
09_cpx.GH.96.96GH2911        .........................................................................................................................................................---------------   15
10_CD.TZ.96.96TZ_BF061       ..---------.------......-------A------------..----A----..AC-----------T----.............TT---------......-------------...........................---------------------  183
11_cpx.GR.x.GR17             ..---------.----G-......T--A-G-A------------...A-----G..------------------A...........T----------......-------------...........................---------------------  121
12_BF.AR.99.ARMA159          ..---------.------......--------------------..A--------..-------------------.............-----------......-------------...........................---------------------  824
13_cpx.CM.96.1849            ..---------.----GC......------A-------------..-------G..-------------------.............TT--------T......-------------...........................---------------------  182
14_BG.DE.01.9196_01          ..---------.---GGT......--A---A-------------...A--------..--------------T----.............-----------......-------------...........................---------------------  330
15_01B.TH.99.99TH_MU2079     ..---------.----G-......T--A---A------------...A--------..-------------------.............T----------......-------------...........................---------------------  212
16_A2D.KR.97.97KR004         ..---------.------......---A-GG..-----------..---------..-------------TG----.............TT---------......--------A----...........................---------------------  171
17_BF.PE.02.PE02_PCR0155     ...............................................................................................................................................................---------    9
18_cpx.CU.99.CU76            ..---------.----G-......T--A---A------------...A--------..-------------------.............T----------......-------------...........................---------------------  121
19_cpx.CU.99.CU7             .......................................................................................................................---------...........................----------G-------A   30
20_BG.CU.99.Cu103            ..---------.----G-......T--A---A------------...A--------..------------------C.............T----------......-------------...........................---------------------  279
21_A2D.KE.91.KNH1254         ...............................................................................................................................................................---------    9
22_01A1.CM.01.01CM_0001BBY.  ...............................................................................................................................................................---------    9
23_BG.CU.03.CB118            ..---------.----G-......T--A---A------------...A--------..------------------M.............T----------......-------------...........................---------------------  253
24_BG.CU.03.CB378            ..---------.----G-......T--A---A------------...A--------..------------------C.............T----------......-------------...........................---------------------  253
24_BG.CU.03.CB471            ..---------.----G-......T--A---A------------...A--------..------------------C.............T----------......-------------...........................---------------------  253
25_cpx.CM.01.101BA           ...............................................................................................................................................................---------    9
27_cpx.CD.97.97CDKTB49       .........................................................................................................................................................---------------   15
28_BF.BR.99.BREPM12609       ..---------.------......--------------------..---------..-------------------.............-----------......-------------...........................---------------------  173
29_BF.BR.02.BREPM119         ........................................................................................................................................................................    0
31_BC.BR.02.110PA            ..---------.------......T--A-T--------------...A-----G..-----------------...........TTTTA------......-------------...........................---------------------  259
32_06A1.EE.01.EE0369         ..---------.------......T--A----------------...A--------..-------------------.............T----------......A------------...........................---------------------  454
33_01B.MY.05.05MYKL007_1     ..---------.------......T--A---A------------...A--------..-------------------.............-----------......-------------...........................---------------------  189
34_01B.TH.99.OUR2478P        ...............................................................................................................................................................---------    9
35_AD.AF.05.05AF095          ...............................................................................................................................................................---------    9
36_cpx.CM.00.00CMNYU830      ...............................................................................................................................................................---------    9
37_cpx.CM.00.00CMNYU926      ...............................................................................................................................................................---------    9
38_BF1.UY.05.UY05_4752       ..---------.------......-----G-A------------...A--------..-------------------.............T----------......-------------...........................---------------------  230
39_BF.BR.03.03BRRJ103        ..---------.------......T-------------------..---------..-----------------G---.............-----------......-------------...........................---------------------  278
40_BF.BR.04.04BRRJ115        ..---------.-----A....G-----T--------------..G--A------..-----------------T-.............T----------......-------------...........................---------------------  344
42_BF.LU.03.luBF_05_03       ..---------.------......T-------------------..---------..-------------------.............T----------......-------------.........................-R---...-------------  333
43_02G.SA.03.J11223          ..---------.----G-......T--A---A------------...A--------..-------------------.............TT---------......-------------...........................---------------------  338
U.CD.83.83CD003              ..---------.------......T--A----------------..---------..-----------------G---.............T----------......-------------...........................---------------------  302
U.CD.90.90CD121E12           ...........................................................................................................................................................................-A----    6
U.GR.99.GR303                ..........................-A-----------ACG--T------..------------------A...........TT---------......-------------...........................---------------------   93
U.NL.95.U_NL_95_H10986_D1    ..---------.------......T--A-T--------------...A--------..-----------------G---AA..........T----------......-------------...........................---------------------  701
CPZ.CD.90.ANT                ..-----TC--.---C--.........-----GTA---GA-----------.....----A....AG--------CTTC-..........TTTTGTT--G-TG----T.........-.C-ATCCTAGGGGAAGGTCGAAGTCTCTAGGAAC...--GA-------A---G-G---  168
CPZ.CM.01.SIVcpzCAM13        A.---------.T---G.......T--A---AA-----------....---------.------------A-..................----TG----.......--T------CCTAGAGGAAGGGCGAAGTCTCTAGGGG...GAGA----------------  366
CPZ.CM.05.SIVcpzEK505        ..-------C-.T---G.......T--A---A--G---------....------GA.AA----------A-..................--------G--.......--T---C--G--AAGGA.......................G-GAG---------------  343
CPZ.CM.05.SIVcpzLB7          A.---------.----------.....T--A------------......---A-----.-------------T--...............C------------....G-T-----------A.......................GAGAGA--------------  350
CPZ.CM.05.SIVcpzMB66         ..---------.----------.....T--A------------......---A-----.-------------T--...............T------------....GC--T----------A.......................GAGAG---------------  353
CPZ.CM.05.SIVcpzMT145        A.---------.----G.......T--A-G-A-----------..GC----A-TCAAG-----------A-..................----TG-----.......--T------CCTAGAGGAAGGGCGAAGTCTCTAGGGG...GAGA----------------  373
CPZ.GA.88.GAB1               A.---------.----G.......T--A---A-----------....---------.------------A-..................----TG-----.......--T------CCTAGAGGAAGGGCGAAGTCTCTAGGGG...GAGA----------------  841
CPZ.TZ.00.TAN1               ..----------.---CA......GCGCA-TCA-------------....-CG-A-TCCGG----------AAA......................---TGT-GT-.......--T--GC-ACCCTAGAGGAAGGGCGAAGTCTCTAGGTAACAG-G--A--------------  367
CPZ.US.85.CPZUS              A.-------C-.----G.......T--A---A--G----------...-C-AACTCC.AGA----------AT.........................--TG-C----......A-T----G--CCTAGGGGAAGGGCGAAGTCTCTAGGGG...-AGAG---------------  846
CPZ.CM.x.SIVcpzDP943         A.---------.----G.......T--A---A-T-----------....-CG-A-TCCAGA-------G..............TTTTTTC----TG-C---........--T----G--CCTAGAGGAAGGGCGAAGTCTCTAGGGG...-AGAG---------------  372
CPZ.CM.x.SIVcpzMB897         ..-----A----.----GT......GCGCAC-A-----G-------..-----C-GCGA.------------T--...............T-------------.......-A-A---------A.......................GAGAG---------------  349
N.CM.02.DJO0131              ..-------C-.T--GT-........-A---A--G----------.....------GA.AGT-----------A-.....................--------G--.......--T---C----A-CAGGA.....................GAGAG---------------  274
N.CM.04.04CM_1015_04         ..-------C-.T--GGT.......--A---A--G----------.....------GA.AGT-----------A-.....................--------G--.......--T---C----A-CAGGA.....................GAGAG---------------  276
N.CM.04.04CM_1131_03         A.-------C-.T--GT-........-A---A--G----------.....------GA.AGT-----------A-.....................--------G--.......--T---C----A-TAGGA.....................GAGAG---------------  275
N.CM.95.YBF30                ..-------C-.T--G.......T--A---A--G----------.....------GA.AGT-----------A-.....................--------G--.......--T---C----A-TAGGA.....................GAGAG---------------  362
N.CM.97.YBF106               A.-------C-.T--G.......T--A---A--G----------.....------GA.AGT-----------A-.....................--------G--.......--T---C----A-TAGGA.....................GAGAG---------------  363
O.BE.87.ANT70                GG--------A.GC-G--......T--A-C--CT-----------....A-GAA-TCA.-A-A.-G--GA-TA-.....................---TG--G--.......--T---C---CCTAGGGGAAGGGCGAAGTCTCTAGGGG...--GA------------T---  855
O.CM.91.MVP5180              GG--------A.GC-G--......T--A-CT-CT-----------....A-GAA-TCA.-AA-A-G--GA-TA--....................--TG--G--.......--T---C---CCTAGGGGAAGGGCGAAGTCCCTAGGGG...--GA----------------  830
O.CM.96.96CMABB637           GG--------A.GC-G--......T--A-C--CT-----------....A-GAA-TCA.-AA-A-G--GA-TA--....................--TG--G--.......--T---C---CCTAGGGGAAGGACGAAGTCTCTAGGGG...--GA----------------  282
O.CM.98.98CMA104             GG--------A.GC-G--......T--A-C--CT-----------....A-GAA-TCA.-A-A.-G--GA-TA--....................--TG--G--.......--T---C---CCTAGGGGAAGGGCGAAGTCCCTAGGGG...--GA----------------  287
O.CM.99.99CMU4122            GG--------A.GC-G--......T--A-CT-CT-----------.....AG-A-TCA.-A-AA-G--GA-TA--....................--TG--G--.......--T---C---CCTAGGGGAAGGGCGAAGTCTCTAGGGG...--GA------------C---  286
O.FR.92.VAU                  GG--------A.GC-G--......T--A-C--CT-----------....A-GAA-TCA.-AA-A-G--GA-TA-.....................---TG--G--.......--T---C---CCTAGGGGAAGGGCGAAGTCTCTAGGGG...--GA----------------  364
O.SN.99.SEMP1299             GG--------A.GC-G--......T--A-C--CT-----------....A-GAA-TCA.-A-AG-G--GA-TA--....................--TG--G--.......--T---C---CCTAGGGGAAGGGCGAAGTCTCTAGGGG...--GAG---------------  857
O.US.99.99USTWLA             GG--------A.GC-G--......T--A-C--CT-----------....A-GAA-TCA.-GA-AAG--GA-TA--....................--TG--G--.......--T---C---CCTAGGGGAAGGGCGAAGTCCCTAGAGG...-AGA----A-----------  278
```

```
B.FR.83.HXB2            TCAGTATTAAGCGGGGGAGAATTAGATCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGG  974
Gag                     _S__V__L__S__G__G__E__L__D__R__W__E__K__I__R__L__R__P__G__G__K__K__K__Y__K__L__K__H__I__V__W__A__S__R__E__L__E__R__F__A__V__N__P__G__L__L__E__T__S__E__G__C__R__Q__I__L__G
A1.GE.99.99GEMZ011      ------------------C----G---AA------G-GT-------------------------------------G-C--------T---------------------------G---A--------C----C---A----T---------A-------A---CA-------A---A  179
A1.KE.00.KER2008        -----------T------A--------GC------G----------------------------------------G-C-G-----CT-----------------------G---AA-------C----C--------T--------T--------A---CAG------A--AA  179
A1.KE.00.KNH1144        -----------T------A--------A-------G----------------------------------------G-A-G------C------------------------G---A--------C----C---A----T---------G-------A---CAG------A-A-A  179
A1.KE.00.KSM4024        --------G---------A--------GC------G--------------------------------------------G------C----C-------------------G--TA----T---C----C---A----T--G-G---A-------A---CA-------A---A  179
A1.KE.00.MSA4069        ---A-------T------A--------GC------G--G-------------------------------------G---G------T--------------------A-G---A----T---C----C--------T---------A-------A---CA-------A--AA  179
A1.KE.00.NKU3005        --------------A---A--------GC------------------------------G----------------G-------------------------------G---A----T---C----C--------T---------A-------A---CA-------T-A-A  179
A1.RU.00.RU00051        ---A-------T------A--------GC------G-------------------------G----C---------G-C------CC----------------------G---A-------GC----C--GA----T-------------------A---CA---------A-A  278
A1.RW.93.93RW_024       ------------------A--------GC------G----------------------------------------G-A-G------C----G-----------------G---A----T---C----C---A----T------T--A-------A---CA-------A-A-A  179
A1.SE.95.SE8891         ---AA------T----A-A--AA----GCT-----G-----G---------------------------------------C-G--------G-----------A--------G---AA-C-----C----C---A----T---------A-------A---CA-------A---A  179
A1.SE.95.UGSE8131       ------------------A--------GCT-----G-----------G--------------C----------G-C--------C-----------------------G---AA---T---C----C--------T---------G-------A---CA---G---A-A-A  370
A1.TZ.01.A173           ---A-------T------A----G---GC------G----------------------------------------G-C-------CC----------------------G--CA--------C----C---A----T---------A-------A---CA---G---A--AA  179
A1.UA.01.01UADN139      ------------------A--------GC-----------------------------------------------G-A-------CC----------------------G---A-------GC----C---A----T---------G-------A---CA----------A  179
A1.UG.92.92UG037        -----------T------A--------GC------G----------------------------------------G-----------C-----------------------G---A----T---C----C---A----T---------A-------A---CA-------A---A  993
A1.UG.99.99UGA07072     -----------T------A--------GC------G----------------------------------------G-C--------C-----------------------G--CA----T---C----C--------T---------G-------A---CA-------A---C  179
A2.CD.97.97CDKTB48      --------G-----C---A-------AGCT-----G----------------------------------------G-C-G------T------------------------G---AA----T--A-C--C--CA----T---------GA-AC---A-------G----T-T--  321
A2.CY.94.94CY017_41     ---A--------------A--------GCT-----G----------------------------------------G-C-G------T-G----------------------G--GAA----T--A----C--------T---------C----G--A-------------A-AA-  343
A.SN.01.DDI579          ------------------A-----A--AAG-----G------------------------------C---------G-A--------T-----------------------G---A--------C----C---A----T---------A-------A---CA---G---A--CA  181
A.SN.01.DDJ369          ------------------A--------GC------G----------------------------------------G-A-G--G---C--A--------------------G---A----T---C----C---A----T---------G-------A---CA---G---A-A-A  181
A.SN.96.DDJ360          ------------------A--------AAG-----G------------------------------C---------G-A-G------T--A--------------------G---A--------C----C---A----T---------A-------A---CA---G---A---A  181
A.CD.02.02CD_KTB035     -----G------------A--------GC------G-----------------------------------CG-C-G------T----------------A-----G--CA----T---C----C---A----T-------------------A---CA-------A-A-A  971
A.CD.97.97CD_KCC2       ------------------A--------TC------G----------------------------------------G-A-G------T------------------------G---AA-------C----C--------T---------G-----------CA-------A--A-  976
A.CD.97.97CD_KTB13      --G---------------A--------TC---------------------------------------------G-C-G------T------------------------G--GA----T---C----C---A----T---------GG------A---CA-------A----  190
B.AR.04.04AR151516      ------A-----------A--------A--------------------C-----------C-T----------G------------------------------------------T-----------------T-----G-----------------------------A-A  181
B.AU.87.MBC925          ---------------------------A---------G----------------A-----------------C-----------------------------------G----------T--------------------------------------------------A-  976
B.BO.99.BOL0122         --G--------T-C---G---------AA------G------------C----------------------C-G---------------------------A-------------T-------C-----------T-------------GG--------------------  179
B.BR.03.BREPM2012       ------------------C--------AA---------G-------C----------A----------------G---------------------------------G---------------------------T-----GT-------T---------------T---A  369
B.CA.97.CANB3FULL       --------------------T------A----------G---------------G-------------------G------------------------------A-----G-----T--------C--------A-------------G----------------T---A  254
B.CN.05.05CNHB_hp3      -----------T--C---C----G---A----------G---------------------------------------------------------------------G--------------------------C---------G-------T--------------CA  970
B.CO.01.PCM001          ---------------------------A----TT----------T-----------A-----------G-----------------------------------A------A----T-----------------TC-G---T----------T-----------------  179
B.GB.83.CAM1            ---------------------------AA---------------------------------------------------------------------------------------------------------T-----G-----------------------------  975
B.GB.86.GB8             ---------------------------A----------------------------------------------G----------G---------------------------------------------------------------------------------A-A  973
B.GE.03.03GEMZ010       -----------------GA---------A-------------------C-----------------C-G---------------------------G----A-----------C-A------------C-----T--------------G-------A-------A-A--  179
B.IT.05.SG1             ---T---N-T--------A--------A-G------------------------------------------C----------------------------------------------------C--------------------G-G------------G--------  773
B.JP.05.DR6538          -----------------------G--CAA------G-----------G--------A------------------C-C----------------------------------------T-----C-----------------------G----------C-A------------  978
B.KR.05.05CSR3          ------------------C--------A-------------------------------------------------------C-------------------A-----------T--------C---A--------------G-------T------------A-T--  999
B.NL.00.671_00T36       ---------------------------A---------------------------------------G---------------------------------A-------------------------------T-----G-----------------------------  500
B.RU.04.04RU128005      ---A-------T-C---G--------CAAG------------------C-------A-----------------G-----------------R------------A-----------------C--------------R-----------G-----------G-------A  465
B.TH.00.00TH_C3198      ------------------C--------A----------G-----------------A-----------------G----------C-------------------------G-----------------------------------------------------------  179
B.UA.01.01UAKV167       --------------------C------AA-----------------------------------------------------------------------------------------------------------C-----G--------------C--------------  179
B.US.04.ES10_53         ------------------A----G---AA-------------------------------------C------G---------------------------A-----------------------C-----A---------G----G--------------A----  974
B.US.99.PRB959_03       --------------C---A---------A---------G--------G--------------C----G------------G-----------------------A--------------T--------------A---------G-------A-----------------  185
C.AR.01.ARG4006         ---A-------A--C---A--------ACT--------G---AA-----A-----------------CC----TGA-G-----CC----C-----------------G---A----T-----G-C--------T---------A---------------------A-AA-  179
C.BR.04.04BR013         ---------C-A--C-A-A--------ACT--------G---A-----C------------------------TGA-G-----CC----C-----------------G---A----T---C--GCC--------T-----G---G---------------------A-AAA  435
C.BW.00.00BW07621       ---A-------A----A-A--------AA-------------A----------------G-GA--GC-T----TGA-------C-----------------------G---A----T---C----C--------T-----G----------------A-------A-AAA  351
C.CN.98.YNRL9840        ---A-------A------A--------AA-------------A-------A---------------C-C----TGC------CC----------------------G---A----T---C----C--------T-----G--------C-------A-------A-AAA  185
C.ET.02.02ET_288        ---------C-A--C-A-A--------AA----------G--A-A-----------A---------C-T----TGC------CC----------------------G---A----T---C----C--------T-----G------AC---A----A-------A-AAA  179
C.GE.03.03GEMZ033       ---A-------A------A-------CAA-------------A-----------------------C-C----TGC------CC----------------------G---A-------C----C--------T-----G---G----C-------A-------A-AAA  179
C.IL.99.99ET7           ---A-------A------C----G---A-G------------A-----------------------C-T----TGC-------T----------------------G---A----T-T-C----C--------T-----G---G-------------A-------A--A-  179
C.IN.99.01IN565_10      -----------A--C---A--------AA-------------A-----------------------CGC----TGA------CT----------------------G---A----T---C----C--------T-----G---G-------------A-------A-AAA  355
C.KE.00.KER2010         ---A-------A------A--------GC-------------A-A---------------------C-T----GGC------CC----------------------G---A----T---C----C--------T-----G-----------A----A-------A-AAA  179
C.MM.99.mIDU101_3       ---A-------A------A--------AA-------------A-----------------------C-C----TGC------CC----------------------G---A----T---C----C--------T-----G-----------------A-------A-AAA  353
C.MW.93.93MW_965        --CA-------A--A---A--------AC-------------AA-C--------------------C-T----TGC------CC----------------------G---A----T---C----C----C---T---------G-------------A-------A-AAA  179
C.SN.90.90SE_364        ---A-------A------A--------TC---------G---A-----------------------C-C----TGC------CC---------------A------G---A----T---C----C--------T-----G-----------------G------A--AA  179
C.SO.89.89SM_145        ---A-------A--C---A--------GC---------G---A-----------------------C-C----TGA------CC----------------------G---A----T---C----C--------T-----G-----------------A-------A--AA  179
C.TZ.02.CO178           ---A--C----A--C---A--------AA---------G---A-----------------------C-C-----GA------CC----------------------G---A----T---C----C--------T-----G-----------A----A-------A-AAA  179
C.UY.01.TRA3011         ---A-------A--C---A--------AC---------G---A------A----------------C-T----TGA-G-----CC----C-----------------G---A----T---C--G-C--------T-------------------T----A-------A-AAA  179
C.YE.02.02YE511         ---A-------A--C---A--------ACG------------A--------------------------------TGC------CC----------------------G---A----T--------C--------T-------------------------A-------A-AA-  179
C.ZA.04.04ZASK164B1     ---A-------A------A--------AC-------------AA----------------------C-T----TGC-------T----------------------G---A----T---C----C---A----T---------GAG-----A----A-------G-AAA  409
C.ZA.05.05ZAPSK240B1    -----------A--C-A-A--------AA-------------A-----------------------C-T----TG-------C-----------------------G---A----T---C----C--------T--G---T--G-------------A-------A--CA  408
C.ZM.02.02ZM108         ---A-------A----A-A-G------GC-------------A-----------------------C----C-TGC------CC----------------------G---A----T---C----C--------T-----------------------A-------A--AA  973
D.CD.83.ELI             ------------------A--------AA-------------------C-------A-----------------G-C---------------------------------------AT---C-------------T----------------------A-------A-A--  520
D.CM.01.01CM_0009BBY    --------------A---A----G---GC------G------------C-----------CGC----GG----GGC-----------------------------------------T---C-------------T--G-----------G-------A-------A-AA-  179
D.KE.01.NKU3006         -----------T------A--------GA-------------------C---------T---C----G-----GC------------------------------------------T---A----------T--T---------C-------------------------A  179
D.KR.04.04KBH8          --T---------------A--------AC-----A-------------C-------AA-----A----TG--C---C-G-------------------------------A--------T---C-------------T-----G------A-----------------A-A--  933
D.TD.99.MN011           ------------------A----G---GAC--------G----A----C-------------C-----G----GC-------------------------------------------T---C-------------T-----G--------T-------A---------A--  192
D.TZ.01.A280            ------------------C--------GC-----------------------------------------C--C-----------------------------T-----------T---C-------------T-----G-----C----------A------......  173
D.UG.99.99UGK09259      ---------C------A-A--------GAC-------GG---------C------------------------GGC--------C-------------------A------------T---C-------------T---------G-------------A-----------A  179
D.YE.01.01YE386         ------------------A--------GAT------------------C-------------C----G-----GC-------------------------------------------T---C-------------T-----------------------A-------A----  179
D.YE.02.02YE516         ------------------A--------GA-------------------C------------GC---------CG-C------------------------------------------T---C-------------T--------------T-------AG--------A--  179
D.ZA.90.R1              -----T------------AG-------GC---------G-----------------A---------C------G-C-----------------------------------A----T---C-------------T-----------------------A---G---A---C  320
F1.AR.02.ARE933         ------------------A--------AC----------G-------------G-------------------G-C-G------T----G------------------------------A--G----------TC--------G-----------C--A-----A-A--  283
F1.BE.93.VI850          ---A--------------A--------GA--------------A---------G-------------G--------A-G------C--A-----------------------------T---C--G----------TC--------------------CA-A-----A-AA-  329
F1.BR.01.01BR125        ------------------A--------GC------------------------G-------------------G-C-G------T-----------------------------T---C-------------TC--------------------C--A-----A-A--  428
F1.ES.x.P1146           ------------------A--C-----GC----------------C----A----------------------G-C-G--G---C--------------------T-----------T------G----------TC--------C------------ARA-----A-A-C  272
F1.FI.93.FIN9363        ------------------A--C-----GC------------------------G------------C------G-A--------C------------------------------T---A-AG----------TC--------------------CA-A-----A-A-C  315
F2.CM.02.02CM_0016BBY   ------------------A--------GAC-----G-----------------G-A-G---------------GGC------------------------A---------------T---C-------------T-----G---A-------------A-------A-A--  179
F2.CM.95.MP257          ------------------A--------GC------------------------G-----G-------------GGC-G----------------------------A--------T---C-------------T-----G---A-------------AGA-----A-A--  186
F2.CM.97.CM53657        ---C--------------A--------GAT-T---------------------------G-------------GGC-G------------------------------A----T---C-------------T-----G---AAG-----------A-------A-A--  179
G.BE.96.DRCBL           ------------------A--------GCT-----G-----------G-------------------G-----G-A-G------T----------------------G--CA----T---C----C--------T---------G-------T---CA-A-----A---C  959
G.CM.01.01CM_4049HAN    ------------------A--------TCT------------------------------------------C--C-G---------------------------A-G--GA----T---C----C---A----T---------G-------A---CAG--GC--A----  179
G.CU.99.Cu74            -----G-----------CG--------GC------G---------C-G--------------------------G-A-G------C-C--------------------G---AA---T---C-C--C--------T-------------G---T---CA-------A--AA  587
G.ES.99.X138            ------------------A--------GC------------------G--------------------------G-W-G------T----------------------G---A----T---C-C--C---A----T---------G-------T---CAG-----M----  424
G.GH.03.03GH175G        ------------------A--------GCT-----------------G--------------------------G-A-G------C-------------------A-G--GA----T---C----CT-------T---------G-------T---CAG-----A--AT 1017
G.KE.93.HH8793_12_1     ---S--------------A--------GC---------------SC-G-C---------------------C-G-C-G------C----------------A---A-G--GA----T---C----C--------T---------G-------T---CA-------A--A-  376
G.NG.01.01NGPL0669      ------------------A--------GC------------------G--------A-----------------G-A--------T----------------------G---A----T---C----C--------C---------G-------A---CAG-----A--AA  179
G.PT.x.PT2695           ------------------A--------GC------------------G--------------------------G-A-G------T----------------------G---AA---T---C-C--C---A----T---------G-------T---CA-------A--A-  925
G.SE.93.SE6165          ----------C-------A--------GCT-----------------G-------------G----TC-------A--------C-----------------------G--GA----T---C----C---A----T---------G-------T---CA-------A--A-  377
H.BE.93.VI991           ------------------A--------GCT-----G-------------------------G-----------GGC--------C-G--------------------G---A----T---C----C--C-A---T---------G----T-----CCA---------A--  360
H.BE.93.VI997           -----------------CG--------ACT-T---G-------------------------------------GGC-------------------------A-----G---A----T---C----C--C-----T------T--G-----------CT-------A-A-A  301
H.CF.90.056             --------------C---A--------GCT-----G---------C---------------------------GGC--------C-----------------------G---A----T---C----C--C-----T---------C-----------CT---G---A-A-A  321
J.CD.97.J_97DC_KTB147   --G-A-AGC-T---------G-CC--AGC------G------T--C-G----G-C-----G----------C-GGCG------------------------------G--CA---AT---C----C--------TC--T-GTG-G--------G--GA---G-----A-T  172
J.SE.93.SE7887          ---A-------T------A--------GAT-----------------G-----------G-------------GGA-----G---C-----------------------G--CA----T---C----C--------TC----GT--G--A--------CA---------A-T  297
K.CD.97.EQTB11C         ------------------A-------CAA--------------A----C----------------------C-GGC--------C--------------------------------T---C----C--AA----T-----G---GT----------C-G------A-AA-  185
K.CM.96.MP535           ------------------A--------GC----------G--------C--------------------------C-G------C--------------------------------T---C-C--C--------T-----G---A-----------C-G------A-AAC  185
```

B.FR.83.HXB2	TCAGTATTAAGCGGGGGAGAATTAGATCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGG	974
Gag	_S__V__L__S__G__G__E__L__D__R__W__E__K__I__R__L__R__P__G__G__K__K__K__Y__K__L__K__H__I__V__W__A__S__R__E__L__E__R__F__A__V__N__P__G__L__L__E__T__S__E__G__C__R__Q__I__L__G	
01_AE.CF.90.90CF11697	-----------T------A--------GC------G---------C---------------------------G-A-G------T----------------A---T-G---A--------C----C--------T---------G-------A---CA----T--A-A-A	929
01_AE.CN.05.FJ051	-----------T------A--------GC-------------------C----------------------C-G-A-G------T----G-----------A---T-G---A----T---A----C--------T---------G-------A---CA-------A-A-A	990
01_AE.CN.06.FJ054	----------CT------A--------GC-------------------C----G-------------------GGA-G------T----------------A---T-----A--------C----C--------T---------G-------A---CA----T--A-A-A	996
01_AE.HK.x.HK001	---A------CA----A-A--------GC-------------------C------------------------TGA--------C----------------A---T-G---A--------C----C--------T---------G-G-----A---CA---G---A-A-A	351
01_AE.TH.01.01TH_R2184	-----------T------A--C-----GC-------------------C------------------------GGC-G------C----------------A---T-G---A--------C----C--------T---------G-------A---CA---G---A-A-A	179
01_AE.TH.02.OUR769I	-----------T------A--------GC-------------------C------------G------------GA-G------T--A-------------A---T-G---A--------C----C--------T---------G-------A---CA-------A---A	179
01_AE.TH.90.CM240	-----------T------A--------GC------------------GC------------GA----------GGC-G------T----------------A---T-----A--------C----C---A--T-T---------G-------A---CA-------A-A-A	539
01_AE.US.00.00US_MSC1164	----------CT------A--------AA-------------------C------------G------G----GC---------T----------------A---T-G---A--------C----C--------T---------G-------A---CA----C--A-A-A	179
02_AG.CM.02.02CM_4082STN	------------------A--------GC------G--G--------G--------A----------------G-C----G---C----------------------G---A--------C----C--------T---------A-------A--CCA----C--A--AA	179
02_AG.EC.x.ECU41	....----------------------G---A--------C----C-----T--T---------G-------A---CA---CC----A-A	85
02_AG.GH.03.GHNJ196	-----------T------A--------GC------G-----------GC-----------------C-G------C-------------------------------G---A-------TC----C--------T---------G-------A---CA---GC--T---A	997
02_AG.NG.01.PL0710	-----------T------A--------TC------G-------------------------------G-------A-------------------------------G---A----T---C----C----A-T-A---------G-------A---CA----C--A---A	179
02_AG.NG.x.IBNG	-----------T------A--------GC------G-------------------------------------G-C--------C----------------------G---A--------C----C--------T---------G----G--A---CA----C--A---A	514
02_AG.SE.94.SE7812	-----------T------A--------GC------G-------------------------------------G-C--------C----------------A---T-G---A--------C----C------------------G-------A---CA----C--A---A	361
02_AG.SN.98.MP1211	-----G-----T------A--------GC------G-G-----------------------------------G-C--------C----------------A---T-G---A--------C----C--------T---------G-------A---CA-------A-A-A	186
03_AB.RU.97.KAL153_2	------------------A--------GC----------------------------------G---------G-A-------CC----------------------G---A-------GC----C---A----T-----------------A---CA-----------A	206
04_cpx.CY.94.CY032	-----------T------A--------GC------G-GG----------------------------------G-C-G------C----------------A---T-G---A--------C----C--------T---------G-------A---CA----T--A---A	340
05_DF.BE.x.VI1310	------------------A--------GC-------------------C-------A-----------G----G-C--------T-------------------------C-----T---C-------------T-----G----------------A-------A-AA-	350
06_cpx.AU.96.BFP90	-----------T------A--------GA-------------------C--------------------------A-G-----CC----------------------G---A--------A----C--------T---------G-------A---CA---G---A-A-A	1002
08_BC.CN.97.97CNGX_6F	---A-------A------A--------AA-------------A-------------------A---C-C----TGC-------CC----------------------G---A----T---C----C--------T-----G---------------AG-------A-AAA	181
09_cpx.GH.96.96GH2911	------C----T------A--------GC------G-------------A-----------------------G-A-G--G---C--------------------T-G---A-----------G----A-----C---------A---G---G---CA-------T---A	185
10_CD.TZ.96.96TZ_BF061	------------------A--------GA----------------C--C-----------------C--------C--------T-------------------------C-----T---A-------------T----------------------A-------A-A--	353
11_cpx.GR.x.GR17	------------------A--------GC---------------------------------------G----GGC--------C----------------------G---A--------C----C---A----T---------GG------A----A-------A----	291
12_BF.AR.99.ARMA159	---------------------------A-------------------G-------------------------G--A---------------------C--A----TA-A--	994
13_cpx.CM.96.1849	------------------A--------AA------------------G-------------------G---C-G-A-G------C-------------------A--G--GA----T---C----C--------T---------A-------T---CA----------A-	352
14_BG.DE.01.9196_01	------------------A--------GC------------------G-------------------------GGA-G------T----------------------G---A----T---C-C--C----A---T---------G-------T---CA-------A----	500
15_01B.TH.99.99TH_MU2079	-----------T------A--------GC-------------------C-------A-----------G------C--------C----------------A---T-G---A--------A----C---A----T------T--G-------A---CA---------A-A	382
16_A2D.KR.97.97KR004	------------------A--------GC-------------------C-------A----------------G-C--------C----------------------------C--T---A-------------T-----G----------------A-------A-A--	341
17_BF.PE.02.PE02_PCR0155	--G------------------------A---C--C-----------A---------------------C--A-----A-A-C	179
18_cpx.CU.99.CU76	------------------A--------AA---------------------------A----------------G-C--------T-G--------------------G--GA----T---C-C--C------------------G-------A---CAG------A-A-A	291
19_cpx.CU.99.CU7	---T-------T------A--------GC---------------------T--C------------C------G-A-G------T--------------------T-G---A--------A----C---A----T------T--G-------A---CAG------A-A-A	200
20_BG.CU.99.Cu103	------------------A--------GA------G-----------G-------------------------G-A-G------C-C--------------------G---A----T---C----C--------T---------GT------T---CA-------A-AAA	449
21_A2D.KE.91.KNH1254	------------------A--------GCC------------------C-------A----------------G-C----G---C-------------------------------T---A----C--------T----------------------A-------A-A-T	179
22_01A1.CM.01.01CM_0001BBY	---------C-T------A--------GC------G--------------A---------------C------G-A-G------T----------------A---T----CA--------C--G-C--------T---------G-------A---CA-------A-A-A	179
23_BG.CU.03.CB118	------------------A--------GC------G-----------G-------------------------G-A-G------C-C--------------------G--CA----T---C----C--------T---------G-------T---CA-------A--A-	423
24_BG.CU.03.CB378	-----------T------A--------TC------G------------------------------C------G-A-G--G---C----------------------G---A----T---C-C--C--------T---------G-------T---CA-------T--AA	423
24_BG.CU.03.CB471	------------------A--------GC------G-----------G-------------------------G-A-G------C-T--------------------G---A----T---C-C--C--------T---------G---G---T---CA-------T--AA	423
25_cpx.CM.01.101BA	-----G-----T------A--------GC------G-------------------------------------G-C--------C----------------------G---A-G------C----C-----T--T---------C-------A---CAG------A--AA	179
27_cpx.CD.97.97CDKTB49	----A-------------C--------GC------G-------AT---------------------C------GGA--------C----------------------G---A----T---C----C--------T---------G-----------CT-G-T-----A--	185
28_BF.BR.99.BREPM12609	-----------------T-----G---A---G---------------------------------------T--------------------C------------------------------A-A--	343
29_BF.BR.02.BREPM119	................-----------A-----------------C-------------------------C-G----------G---A--------------G---------A---------A	154
31_BC.BR.02.110PA	---A-------A--C---A--------ACT------------AA-----A----------------C------TGA-G-----CC----C-----------------G--GA----T---C--G-C--------T-----G-------C--------A-------A--AA	429
32_06A1.EE.01.EE0369	-----------T------A--------GA------G-------------------------G---------------G-----C-----------------------G---A-------GC----C--------T---------G-G-----A---CAG--G---G---A	624
33_01B.MY.05.05MYKL007_1	-----------T------A--------GC-------------------C------------------------GGC-------------------------A-AT--G--CA--------C----C--------T---------G-------A---CA----C--A-A-A	359
34_01B.TH.99.OUR2478P	-----------T------A--------A--------------------C-------A----------GT---C-GA-G------C--------------------T-----A--------C----C--------T---------G-------A---CA-------A-A-A	179
35_AD.AF.05.05AF095	---A-------T--A---A--------GC--G-C--------C----------------------G---A--------C----C---A----T---------A-------A--CCA-------A---A	179
36_cpx.CM.00.00CMNYU830	-----------T------A--------GC------G-------------------------------------G-A--------T----------------A---T-----A-----------G-C--------T------T--G----G--A---CA-------A-A-A	179
37_cpx.CM.00.00CMNYU926	-----G----CT------A--------GC-------------------------------------C------G-A-G------T----------------------G---A--------C----C--------T---------G-------A--CCAG------A-A-A	179
38_BF1.UY.05.UY05_4752	---A------C-------A----G---GC-------------------C-------A----------------GG---------C--------------------------------------G-C--------A---------C--------------A-G---A-A-T	400
39_BF.BR.03.03BRRJ103	--------G------------------A-G---GG-----G---C--A-A--	448
40_BF.BR.04.04BRRJ115	---------------------------AAG-----G---------C----A-----A-G------------C-G------------------------------AT----------------------------T---------------------------------AA	514
42_BF.LU.03.luBF_05_03	--G------------------------A--A--------------G-----C--------------	503
43_02G.SA.03.J11223	-----------T------A--------GC------G-----------G--R----------------------G-A-G------Y----------------------G---A----T---A----C-----T--T------T--G-------A---CA-------A---A	508
U.CD.83.83CD003	------------------A--------GC----------------C--------------CGC----------G-C--------C----------------------G---AA-------A----C-----T--T-----G--GG----G------CA-------A-T--	472
U.CD.90.90CD121E12	------------------A--------GC------G------------------------CGC----------G-C--------C----------------------G---A--------C----C--------T-----G--GG-----------CAG------A-T-T	176
U.GR.99.GR303	-----G-----T------A--------GCC------------------------------------C----C-G-C-G------T----------------A-----G---A----T---C----C--------T-----------------A---CA-------A---A	263
U.NL.95.U_NL_95_H10986_D1	---------C-T--C---C--------AC------------C--------------------------------GC-------CC----------------------G---A----T--------C--------TC----G---G-----------CA-------A-AAT	871
CPZ.CD.90.ANT	--T--T--G--G--A-AGA-GC-----AC--------GT--CA--C-TC----C--T--C--------G--C-TGA--------C-G--T--------ATC------GC-G--T--T--GC-C-GCT-CTC---TC----------------T---GA-A-GGCTA-CCA	338
CPZ.CM.01.SIVcpzCAM13	-----G----CA-----GA----G-----C---------G-----C-T--A--C-------G-----------TGA-G------T----------------A-----G---A--------TG---C--C--G--AA-G----GTG-------A---CTTA--T-G--ACA	536
CPZ.CM.05.SIVcpzEK505	----------CA--A---A---------A-------------TAT--G--A--------------------C-TGA-G-----CT----------------------G---A-------GTG---C--A--T--CA-G--C---G-------T---GCC---T----CA-	513
CPZ.CM.05.SIVcpzLB7	---A-C----CA--C---A-G------GAT---------G-----C-------C-------GA----G-----TGA-G------T----------------------G--GA----T--TTG---C--C--G--TA-G------A-----------GAG---T-GA-TAC	520
CPZ.CM.05.SIVcpzMB66	-----T----CA------A-------CA-------G------TATC-G-------------------------TGA-G------T----------------------G---A-------TTG---C--CTC---CA-G------A------------AG--GC-TT--CA	523
CPZ.CM.05.SIVcpzMT145	---A------C-------AG-------GCT------------A--C----A--------------------C-TGA-G------T----------------A-----G---A-------TTG---C-----T--AA-G----ATG-------T---GAGA-GT--T--AA	543
CPZ.GA.88.GAB1	-----TC---CA------A----------C---------G-----C-T-----C-------GA----G-----TGA-G------T----------------------G---A--------TG-G-C--C--G--AA-G----GTAAG-----A----CTA--T-GT-ACA	1011
CPZ.TZ.00.TAN1	-----G--G--G--A-ATA-GC-G---AC-------TCC--AA--C-T-AAT-CA-A--C-G----------TT-A--------C----------CG-A--C--A---C-G--T-----GA-G-----C--T--CA-G--G-ACGT---------CT-GA----CA-CCT	537
CPZ.US.85.CPZUS	-----TC-G-CA------CG-------GCT--------G------C-T--A--T--------------G----TGA-G------T----T-----------------G---A----T--TTG---C--A--T--AA-G------G----C------CTC---T--T-AAA	1016
CPZ.CM.x.SIVcpzDP943	---A------CG------CG-------GCT------------A--C---------------------------TGA-G------T----C-----------------G---A-------TTG---C-----T--AA-G-----GG----G--A---GA---GC--T-AAA	542
CPZ.CM.x.SIVcpzMB897	----------CG--A---A-G------GAT------------TATC-CC------------------------TGA-G------T----------------------G---A-------TTG---C--C-----AA-G--C---A-------A---GA----T--A-TAA	519
N.CM.02.DJO0131	-----G----CA------A---------A-------GC----TAT--G--A--------------------C-G-A-G------T----------------------G---A-------TTG---C--A--T--CA-G--C---G-GAGC------GCC--GT----AAA	444
N.CM.04.04CM_1015_04	-----G----CA------A--------AA-------GC----TAT--G--A--------------------C---A-G------T----------------------G---A-------TTG---C--A--T--CA-G--C---G-GC-C-------AC--GT----AAA	446
N.CM.04.04CM_1131_03	-----G----CA------A--------AA-------GC----TAT--G--A----------------------G-A-G------T----------------------G---A-------TTG---C--A--T--CA-G--C---G-GA-G------GCC--GT----AAA	445
N.CM.95.YBF30	-----G----CA------A---------A-------TC----TAT--G--A--------------------C-G-A-G------T----------------------G---A-------TTG---C--A--T--CA-G--C---G-G--C------GCCA-GT----AAA	532
N.CM.97.YBF106	-----G----CA------A---------A-------GC----TAT--G-----------------------C-G---G------T----------------------G---A-------TTG---C--A--T--CA-G--C---G-GA-T------GCC--GT--A-AAA	533
O.BE.87.ANT70	--T--G--G-CA--AA-CA----G---GC-------C-----A------A------ATCT--A--G-------G-C-------CT----------------A-----G---A--------TG--------AG--AC-------TG----G--TAA-GAGA-GC-G--ACA	1025
O.CM.91.MVP5180	--T--G--G-CA---A-TA----G---GC-------CG----A-------------ATCT--A--GGC-----GGC--------T----------------------G---A---A----TG---------T--A--------TG-------TAC-GAG---C-G--ACA	1000
O.CM.96.96CMABB637	--T--G--G-CA---A-TA----G---GC-------C-----A------A------ATCT--A--G-------GGC--------C----------------------G--CA----T---TG----Y---AG--A--------TG-------TAA-GAGAC-C-GT-ACA	452
O.CM.98.98CMA104	--T-----G-CA---A-CA----G---GC-------C-----AA-----A------AT-T--A--G-------G-C--------T----------------------G---A-------TTG--------AG--AC-------TG-------TAA-GAGA--T-GT-ACA	457
O.CM.99.99CMU4122	-----G----AA---A-CA----G---GC-------C-----A------A------ATCT--A--G-------G-C--------T----------------------G---A--------TG--------AG--AC-------TG-C--G---AACGAGA--C-GT-ACA	456
O.FR.92.VAU	--T--G--G-CA---ACTA----G---GC-------C-----A------A------ATCT--A--G--------GC--------T--A---------------G---G--CA----T---TG--------AG--AC-------TG-------TAA-GA----C-GT-ACA	534
O.SN.99.SEMP1299	--T--G--G-CA---A-CA----G---GC-------C-----A------A------AT-T--A--G-----C-G-C--------T----------------------G--CA---C----TG--------AG--AC-------TG----G---AA-GAGA--C-GT-ACA	1027
O.US.99.99USTWLA	--T-----G-CA---AA-A--------AC--A----C-----A--------T----ATC---A--G-----C---C--------T---------------A----------A----T---TG--------AT-----------TG-------TAA-GAG---T--T--CA	448

```
B.FR.83.HXB2           ACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAAGCACAGCAAGCAG 1144
Gag                    __Q__L__Q__P__S__L__Q__T__G__S__E__E__L__R__S__L__Y__N__T__V__A__T__L__Y__C__V__H__Q__R__I__E__I__K__D__T__K__E__A__L__D__K__I__E__E__E__Q__N__K__S__K__K__K__A__Q__Q__A__
A1.GE.99.99GEMZ011     ----T----G---G-TT-CAG-------A-------T----A----------------A-----GT------------A--------------TG-------------------------Y--A-----R---ATAM-R--T--G-RC---C---RGAMM--A--G---.  348
A1.KE.00.KER2008       C--AA--------G-T------------A---------------------T----G----------------------A-----GC--------G-------------------------T--AG--------ATA--------G--C--AC----GA----A--G----  349
A1.KE.00.KNH1144       ----A--------G-T--CA--------A---------------------T-------A----------------C--A-------A------TG-G------------------C----T--A---------CT------T--G--C---C----GA----A-------  349
A1.KE.00.KSM4024       ----T--------G-T---A--------A------------A----C-G-T---------------------------A------C--------G--------------------C----T--A---------CT-...C-T--G------C----GA----A--G----  346
A1.KE.00.MSA4069       ----T----G-----T--CAC------GA----------G------------------------GT------------A------C-A-----TG--------------------C----T--A---------AT------T--G--CC--C-G--GA----A--G----  349
A1.KE.00.NKU3005       ----T----G---G-T--C---------A--------------------------------------------------A------C-------TG--------------------C----T--A---------ATAA----G-G---C---C----GA----A-------  349
A1.RU.00.RU00051       ----T--------G-T--CAG------G-Y-----------A------------C---R----------------C--A----T-C-------TG-------------------------T--A---------ATA-----T--G--C---C----GA-C--A--GA---  448
A1.RW.93.93RW_024      ----T--------G-T---A--------A--------A---A------------------------------C-C---A----G---------TG--------------------C----T--AG-------GATA-----T--G--C---C----GA-T--A--G----  349
A1.SE.95.SE8891        TA-AT--------G-T--CAG-------A-----------------C----------------------------C--A-------A-----CTG----------------------------A---------ATA-----T--G-AC---C----GA----A-------  349
A1.SE.95.UGSE8131      ---AT-----T--G-T-CCA-A------A--------------------------------------------------A--------------TG--------------------C----T--A--------GATA-----T-----C---C----GAA-G----G----  540
A1.TZ.01.A173          C---T--------G-TG-CA--------A--------A---A--------T---------------------------A------C-------TG-------T------------C----T--A---------ATA--G--T--G--C---C----GA----A--G----  349
A1.UA.01.01UADN139     ----T--------G-T--CA---------------------A--------------------T---------------A------C------------G--------------G------T--A---------ATA-----T--G--C---C----GA-------GT---  349
A1.UG.92.92UG037       ---AT-----T--G-T--CAGA------A-----------------------------------------------C--A------C--------G--------------------C----T--A--------GATA-----G-----C---C----GA-------G----  1163
A1.UG.99.99UGA07072    ---AT--------G-TA-CA--------A--------A---A--------T------------------------C--A--------------T--------------------------T--A--------GAT----G-T--G--C---C----GA----A--G----  349
A2.CD.97.97CDKTB48     G--AT--------G-T--CG--------A-----------------------------A-----GT------C-T---T-------A-------G--------------------C----T--A-----------A---------T-C---C-G--GA-------G----  491
A2.CY.94.94CY017_41    G---T--------G-T--C--A------A------------A--------------------T-GT------C-G---A---------G----TG--------------------C----T--A-----------A--------GCAG--A-C-C-GCATGCAGC---T.  512
A.SN.01.DDI579         ----T--------G-T--CA--------A-------G---------C----------------------------C---------C-------TG-------------------------T--A---------ATA-----T--G-AC---C----GA-------G----  351
A.SN.01.DDJ369         ----T-----T--G-T--CA--------A------------A--------T----------------------------------C--------G-G-----T-----------------T--A---------ATA-----G-GGCAC---C-----A-------G----  351
A.SN.96.DDJ360         T---T--------G-T--CA--------A-------G---------C---T------------------------C--A--------------TG------------------T------T--A---------ATA-----AGGG--C---C----GA-------G----  351
A.CD.02.02CD_KTB035    GA--T--------GTT--CAGC-----GA--------A-C-A-------------TT------------------C--A---------------G---G---------A------C----T--A-------C-CT------------C---C--G-A-----A--G----  1141
A.CD.97.97CD_KCC2      ----T----GT--G-T--C---------A----------C-A---------C--C--------------------C--A----G-A----------------------------C-----T--A-----A---CT------------CC--C----GA----A--G----  1146
A.CD.97.97CD_KTB13     ---AT----G---G-T---A-T------A--------A-C-A---------------------------------C--A-----G--------C-----G--------------------T-CA---------AT------------C---------CAGGCAGC---T.  359
B.AR.04.04AR151516     -------------G-----A-----------------T-----------------------------------------------A---G--TG-------------------------T----------------------GT-----------G-------------  351
B.AU.87.MBC925         -------------G----------------------CT--------------------------------------------------------G-------------------------G-G-G--------------------------G------------------  1146
B.BO.99.BOL0122        ------------G--------A------------------------------------------GT-----T--------------A-------G-----------------------------------------------------------------------------  349
B.BR.03.BREPM2012      -------------------A--------------------------------------------GT---------------------AG----TG-------------------------G-----------------------------------G-----------C-  539
B.CA.97.CANB3FULL      ----------------------------------------------------------A-----GT--------------------A-------G-------T-----------------T------------------------------------------------  424
B.CN.05.05CNHB_hp3     -------------------------------------T-G-A--------TC---G--------GT----------------------------G-------------------------G-----------------------------------G-------------  1140
B.CO.01.PCM001         ------------------------------------G----A--------T-------A-----GT------------A---------------G--------------------C----T-G-----------------------------------------------  349
B.GB.83.CAM1           ------------G------------------------------------------------------------------------------A------TG-------------------C-----G--A--------------------------------------------G-  1145
B.GB.86.GB8            --------------------------------------G-----------T---C-----------T-----------A--------------TG-------------------------T-------------------G-----------------------------  1143
B.GE.03.03GEMZ010      ----T---------------C--------------------A------------C---------GT------------A-------------G-G-----------C-------------A------C--------A--------------------G-------------  349
B.IT.05.SG1            -----------------------------------------A----------------------GT------------A------G-A--------------------------C-----G-----------------------------------G----------T--  943
B.JP.05.DR6538         -------------G---------------------------A--------T-------A-----GT------------A-------AA----C----C----------------C-----G-----------------------------------G-----A-------  1148
B.KR.05.05CSR3         ------------G----------------------------A-------C-----G--A-----------------------AG---------TG-------------------------G---G---C-----------G--------------C--------------  1169
B.NL.00.671_00T36      -------------G---------------------------A--------T-------------------------------GC----------G-------------------------A-A------------A------------------CGG------------C  670
B.RU.04.04RU128005     -------------AG---------------------R-----G---------------A-----GTT-----------A-------AA------G-------------A-----------G--A--------G--A--------------------G-------------  635
B.TH.00.00TH_C3198     ----------------------------------------------------TC--------------------------A--------------TG-------------------------G---G--------------------------C----G---------A---  349
B.UA.01.01UAKV167      ----T------------------------------------A----------------------GT------------A-------A---------------------A----------------------------------C--------G-------------  349
B.US.04.ES10_53        ---T-----T--G--------------------------A-----------------G-------------------------------------TG---G----------------C------G-G--------------------T----------G-------------  1144
B.US.99.PRB959_03      ---A--------G----------------------------A----------------A----------------------------A-----TG----------------------------------------------------G-------------  355
C.AR.01.ARG4006        -------------G-T---------------A-G-------T------------------T-----T-----------A---GC--AA---------CG---------A-----C--------A---A-----------G-A------C--C-----A-------G---.  348
C.BR.04.04BR013        -------------G-T--CA--------A--A-G-------T------------------T-----T-----------A-------A----------CG---T-----------C--------A---C-------A------------C--CT----AT------G---.  604
C.BW.00.00BW07621      -------------G-T------------A----G--------------------C-----------T-----------A---A--A-------TG--CG------------G--A------C-------A----A------C-----C--C-----A-------G---.  520
C.CN.98.YNRL9840       -------------G-T------------A----G-----------------TC--C-----------T-----------A---G--G-------AG--CG---------A-----C--------------------A----------T-C--C-----A-------G---.  354
C.ET.02.02ET_288       ---AT--------G-T---A-A------A----G-----------------T---C-----------T-----------A---G--G-------T---CG---------------C-------G------------A---------GTGC--C-----A-------G----  349
C.GE.03.03GEMZ033      -------------G-T------------A----G------------------C--C-----------T-----------A---A--G--------G--CG------------C--C-----T--------------A-----GG--T--C--CG----A-------G---.  348
C.IL.99.99ET7          ---------C---G-T------------A----G-------A------------C-----G-----T--A--C-----A---A--AC------TG-------T-----------C--------------------A------------C--C-----A-------G---.  348
C.IN.99.01IN565_10     ---------T---G-T------------A----G--------G--------C--C-----------T-----------A---GC-GAT------G--CG-------------T-C--------------------A---------T--C--C-----A----A--G---A  525
C.KE.00.KER2010        -------------G-T------------A----G------------------C--C-----------T-----------A---A-T-A------AG--CG---------------C--------------------A------------C--C-----A-------G-G-.  348
C.MM.99.mIDU101_3      -------------G-T------------A---------------------TC--C-----------T-----------A---TC-G-------AG--CG---------A-----C--------------------A----------T----C-----.-------G---.  521
C.MW.93.93MW_965       -------------G-T------------A----G---T-G---------C----C-----------------------A---GC--AC---------CG---------------C--------------------A------------C--C-----A----AA-----.  348
C.SN.90.90SE_364       -------------G-T------------A----G-----------G--------C-----------T-----------A------GAC---A--G--C----------------C--------------------A------------C--C-----A-------G---.  348
C.SO.89.89SM_145       -------------G-T------------A----G-----------------TC--C-----------T-----------A---G---A-------G--CG---------------C--------------------A------------C--C-----A-------G---.  348
C.TZ.02.CO178          ---------C---G-T------------A----G-----------------TC---G-----T--GTT-----------A------G-----------CG---T-----------C--------------------A---------T--C--C-----A-------G---.  348
C.UY.01.TRA3011        ---A---------G-T--C---------A--A-G-----C-T----C-------------T-----T---------------G---A-------G--CG---------------C--------A---A-------A----------AGC--C-GGC--A-..........  339
C.YE.02.02YE511        -------------G-T------------A---CG-------A------------C-----G-----T--T--C-----A---A--GAC----CAG-G-----------------C--------A-----------A------------C--C-----A-------G---.  348
C.ZA.04.04ZASK164B1    ------C--------T------------A----G-----------------TC--C---------GTT-----------A---A--G-----A--G--CT---------------CC-------------------A---------T--C--C-----A--A----G----  579
C.ZA.05.05ZAPSK240B1   -------------G-T------------A----G--------------------C-----------T-----------A---AG---------TG--CG------------------------A-----------A------------C--C-----A-------G....  574
C.ZM.02.02ZM108        ------C--------T------------A----G--------------------C---------GTT--T--------A---G-----------G--CG---------------C--------------------A-----G---T--C--C-----AT---AGC-CAGC  1143
D.CD.83.ELI            G------------G-TA-----------A---------------------------------------------A---A-G-A------TG-------------------------A-----G-----------------------------G-------------  690
D.CM.01.01CM_0009BBY   C---T-G--G-----T---A----------------G-------------T---------------A-----------A---A-C--A------G-------------A-----------T--------------A--G...----AAGCAC-GC-----GCAGCT....  342
D.KE.01.NKU3006        -------------G-TA----------------G-G-A---A------------------------------------A---A-G--A---C--G--GC----------------------------------A----C-------------------------A--A  349
D.KR.04.04KBH8         ---------------TA--A--------------A------A--------T---------------------------A---G----A---A-TG----G-----------------------A--------G--A--------------A------------------A  1103
D.TD.99.MN011          C--------G---G-TA--A-------------------------------T---C-----------A-----------A---G---A----C--G----G--------------------T--------------A-------------------------------T-C  362
D.TZ.01.A280           ....---------G-TA------------------------A--------T-------------------------------AG-A--------G-------------------------A--AT---------------C---------------G------------A  339
D.UG.99.99UGK09259     ----T----------TA----------------G---G---A--------T-------------GT----CT------A---GC---A---A--G--GC---------------C--------------------A---GC-----------G---G------------A  349
D.YE.01.01YE386        -A-----------A-TA----------------G-G-G---A------------------------------------A---G--G-----A-A--------------A---------------C----------A----C----------G------------------  349
D.YE.02.02YE516        C------------G-TT-GA-------T----------------------T---------------------------A---G---------G-G----G--------------------T--------------A----------T---------G----C--------  349
D.ZA.90.R1             ---------------TA--------------------------------------------------------------A---G----------TG-----------C----------------------------A-----------------G--G-------------  490
F1.AR.02.ARE933        ---------C------A-------------------G----A---G------------------GT---------C--A--------------AG----G--------------------G---C----------C-----T------C--C----GAG---------G-  453
F1.BE.93.VI850         ------------------------------------G----AG-------T-------------GT---------A--A---------GC--G-G---C-------------------------C----------A------------C--C----GA----------G-  499
F1.BR.01.01BR125       ---A-----------------------------G--G----A--------T-------------GT---------A--A---------G-----G----G--------------------G---C----------A--------------A-GTC-GCA-A--AC-CAGC  598
F1.ES.x.P1146          ------T--------A--------------------G-----------G---------A-----GT---------A--A---------G-----G----G-----------C--------G---C----------A---------------C----GA----------G-  442
F1.FI.93.FIN9363       ----A----G------A-------------------G-----------------C---A-----GT---------T--A-------A-------G----G--------------------T---C----------A------------C--C----GA----------G-  485
F2.CM.02.02CM_0016BBY  ---A------T-------------------------GA---A------------C---------GT---------A--A-------A----C-A----G---------------------T---C--C-------A--G------TA-C--C-----A----A-C-----  349
F2.CM.95.MP257         ---A-----------------------G-----G-----G-A--------T---C---A---T-GTT--------A--A-------A-------G---G---------------------T---C--C-------A--G------CA-C--C-----A----A------A  356
F2.CM.97.CM53657       ---A--------------------------------G----A--------TC--C---A---T-GT---------A--A------------A-A---GG---------------------T---C--C-------A--G---------C--C-----A----A-C-----  349
G.BE.96.DRCBL          ----T-G------G-T--C--A------A----G--GA---A----C---T---------------------------A-------A-------G---G---------A--G---C----GG-AG-G--AA-GATA-----G--G---C--C-----.............  1116
G.CM.01.01CM_4049HAN   ----A--------G-T--C---------A----G--GA-----------CT-------C-----GT------------A-------A-------G-------------A------C----GG-AG-G--AA-GATA-----G-G---CC--C-----AT------G----  349
G.CU.99.Cu74           C--AT-G------G-T--C--A----C-A----G--G----AG-----------------------------------A-------A-------G-------------A------C----GG-AG-G--AA-G-CA-----G------C--C-----A-G----------  757
G.ES.99.X138           ----T-G------G-T--C---------A----G-G--------------T----G----------A-----------A-------------G-G-------------A------C----GG-AG-G--AA-G-CA--G--G---------C-----CA------G----  594
G.GH.03.03GH175G       G---T-G------G-T--C---------A----G-G--------------TC--------------------C-----A--------A-----AG-------------A------C----G--AG-G--AA--CTA-----G-T----C--C-----CAG---GC----A  1187
G.KE.93.HH8793_12_1    C---T-G------G-TA-C--M------A----G--GA------------T---------------------------A----CC-A-------G----G--------A------C----GG-AG----AA-GATA-----G------C--C-----AT------G----  546
G.NG.01.01NGPL0669     C-----G------G-T--C---------A----G--GA---A----C--------G-T--------------------A---G-----------G---G---------A-----------TG-AG-G--AA--TTA-----G-C----C----G--A-----A-G-----  349
G.PT.x.PT2695          ----T--------G-T--C---------A----G-G--------------T---------------A-----------A---------------G-------------A------C----GG-AG-G--AA-GACA--G--G------C-----C--CAG---GC----A  1095
G.SE.93.SE6165         ----T-G--------T--C---------A----G--A----A-------------G----------------------A---------------G-------------A------C----AG-AG-G--AA-GATA-----G------C--G-----AT------G----  547
H.BE.93.VI991          ---------G---G-T---A--------A-------C---CA----------------A-----GT---------C--A-------A------TG-G----------------------GG------------ATA--G--T--G-AC---C---G-A-------G--CC  530
H.BE.93.VI997          ---A----GG-----TA--A--------A-----------CW--------T------C-----G-----------C--A-T------A------G------------------------GG------------ATA-----------GC--C-----A-----------A  471
H.CF.90.056            ----A----G---G-TA--A--------A------------A--------T----CT-------GT---------C--A----G-AA------TG-G--------------G--------T------------ATA------------C--C-----A------------  491
J.CD.97.J_97DC_KTB147  ----A-C------GAT---T-A------G----G--GA-A-A--------T---C---A-------------------A---T-G--------C----G------------G------------------------CTTT----------C----G-G----TA--A---A--  342
J.SE.93.SE7887         ------C------G-T--C---------A--C-----A---A------G--------------------------C--A-----G--------A-------------T------------G------------ATT--------G-AC--AC-GC-G-----A-------  467
K.CD.97.EQTB11C        ---A-----------------A-----C--G-----G---------C---T------------------------G-----------T---C--G--GG---------------C--------AC----------A-------G---C-C-C----GA-----.......  348
K.CM.96.MP535          ---AA----G------A----A--------------GA---A----C-----------A-----GT---------T--A-------A-------G-------------------C--------AC----------A------------C--CG---GA----A.......  348
```

```
B.FR.83.HXB2                  ACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAAGCACAGCAAGCAG 1144
Gag                           __Q__L__Q__P__S__L__Q__T__G__S__E__E__L__R__S__L__Y__N__T__V__A__T__L__Y__C__V__H__Q__R__I__E__I__K__D__T__K__E__A__L__D__K__I__E__E__E__Q__N__K__S__K__K__K__A__Q__Q__A__
01_AE.CF.90.90CF11697         ----T-----T--AAT--CA-----------------A--A--------T------CA------------GG-----A---------------TG---------------------------T--A----------TA-----T--G-AC---C----GA-------G---- 1099
01_AE.CN.05.FJ051             ----T----GT--A-T--CA--------------------A--------T-------A----GT-----GG--C--A--C-----------T----------------------------T--A----------CA----GG--G--CC--C----GA-------G---- 1160
01_AE.CN.06.FJ054             ---------GT--A-T--CA--------------------A------------------T------T-GG--C-----C-----------TG-------------------A-----T--A--G-------TA-----G--G--C---C----GA-------G---- 1166
01_AE.HK.x.HK001              ----T----GT--A-T--CA--------------------A--------T----------------------GG--C--A--C------------G----------------------------T--AT--------ATA-----T--G-ACC-------GA-------G----  521
01_AE.TH.01.01TH_R2184        ----T----GT--A-T--CA--------------------AG-------T---------------A---GG--C--A--C-T----------G----------------------------T--A----------TA-----G--G--AC--C----GA-------GA---  349
01_AE.TH.02.OUR769I           ----T----G---AGT--CA-----------------------------T---------G-T-------GG--C--A--C----A-------G----------------------------T--A----------TA-----T--G---C--C----GAT------G----  349
01_AE.TH.90.CM240             ----T----GT--A-T--CA---------T----------A--------T------------------------GG--C--A--C------------G----------------------------T--A----------TA-----T--G--CC--CG---GA-------G----  709
01_AE.US.00.00US_MSC1164      ----T----G---A-T--CA--------------------A--------T------C------------GG--C-----C------------G----------------------------T--A--G-------TA--------G--CC--C----GA-------G----  349
02_AG.CM.02.02CM_4082STN      C---T----G---A-T--CA--G-------------GG---A---------------T---------T-GG----A--C-----------A-------------A-----C-----T--A--G------ATA-----T--G--C---C----GA-------G-A--  349
02_AG.EC.x.ECU41              G---T-----T--CAT--CA---T----------------A--------T-------A---------T-GG--C-----------------A-----------------G--C-----T--AT-G-------TA-----T--G-------C--C-GAA------G----  255
02_AG.GH.03.GHNJ196           ----T-----T--A-T--CAG-----------G-------A---C------------A---------T-GG--C--A---------------C---G------------------C-----T--A----------CT-----T--G--C--AC---GA-------G---- 1167
02_AG.NG.01.PL0710            ----T-----T--A-T---AG-------------------A--------T-------A---------T-GG----A-------A----TG-------------A--G--C-----T--A--------GATA-----T--G-AC----C---GCAGGCAGC---T.  348
02_AG.NG.x.IBNG               ----T-----T--G-T--CAG-------------------A--------T-------A---------T-GG--C-----------------C---------------------C-----T--A--G-----ATA-----T--G--C---C----G-T------GA---  684
02_AG.SE.94.SE7812            ---AT-----T--A-T--CAG-------------------A--------T------CA---------T-GG--C--A---------------C--------------A----C-----T--AG-------G-TA-----G--G--C---C----GA-------G----  531
02_AG.SN.98.MP1211            ----T-G---T--A-T--CA--------------------A--------T---------------T--------A----G--A----AG------------------C-----T--AT--------ATA---G-G--G--C--AC----GA-------G----  356
03_AB.RU.97.KAL153_2          ----T--------A-T--CA--------------------A-------------------------------A-----C-----------------------------T--A---------ATA-----T--G--C---C----GA-C--A--G----  376
04_cpx.CY.94.CY032            ----T-----T--A-T--CA-A-----------------------------TA--A--------GG--C--A--------A----TG--C---------------------T--A---------ATA----GT--G--C---C----GA-------G----  510
05_DF.BE.x.VI1310             C---T--------G-TA----------A------------A---------C-------------------------C---G-----------G-G-------------------------A---------------------------A-G----AAGGCA--GCA--  520
06_cpx.AU.96.BFP90            G---T-----T--A-T--CA--------------------A--------T-------------------------A-----------A-AG---C--------------CA----T--A---------ATA-----G--G--C---C----G-----T--G---- 1172
08_BC.CN.97.97CNGX_6F         -------------G-T-----------A---G-------------TC--C----------T----------A---G--GA-----AG--CG--------A-----C-------------------A-----T---T-C--C-----A-G-----G....  347
09_cpx.GH.96.96GH2911         ----T----G-----T---AG-------------------A----------------------------------A-------A-----G--G-----------G-----T--A---------ATA---GGT--G--C---C----GA-------G---A  355
10_CD.TZ.96.96TZ_BF061        -------------G-TA----------------G---A--A-----------------------------------A---G-------A--G-GC-----------------------------------A---C---------------G-----------A  523
11_cpx.GR.x.GR17              ---AT--------G-T--CGG-------A---------------------------------G-------------A----TC--------G---G----------G--------T--AG---------ATA-----T--G--C------C-GCAGGCAGC---T.  460
12_BF.AR.99.ARMA159           -------------------------------------G------------T--------------GT------------A-------A------G---G-------------------G---C----------A-----------C--C-----A----------G- 1164
13_cpx.CM.96.1849             ----T-G--------T--C--------------G--G----A--------------------------------A--C-----------G------------A-----C----GG-AG-G--AA-G-TA-----T-GC--CC--C---G-A-------G----  522
14_BG.DE.01.9196_01           ----T-G--------T--CA--------A----G--GA-----------T-----------------A-----------A-------A------G------------A-----C----GG-AG-G--AA-G-CA--G--G-----KC--G-----CAG---GC----A  670
15_01B.TH.99.99TH_MU2079      ----T----GT--A-T--CA--------------------A--------------------------GG--C--A--AA-C--------G----------------------------T--------------TA-----T--G-ACC--C----GA-------G----  552
16_A2D.KR.97.97KR004          -------G-------TA--A-----------------A--AG----G-T----G-----T---------------A--------A-----G--G-----------------C----T--AT-----------A---------CAC---C-G--GA------CT---.  510
17_BF.PE.02.PE02_PCR0155      -------------------------------------G------------------------GT-------------A-------A-G----G----G-------------------G---C----------A-----S---T--C--C----G-----------G-  349
18_cpx.CU.99.CU76             ----T-----T----TA-CA-A---------------A--AG--------------------T----------A----GG------TG------------------C-----T--A---------ATA-----T--G--C---C----GA-------G----  461
19_cpx.CU.99.CU7              ----T-----T--A-T--CA---------------------------------T--------------------GG--C--A----GCA------G----------------------------T--AT---------TA--G--T--G--C---C----GA-------G-A-A  370
20_BG.CU.99.Cu103             C---T--------G-T--CA-A------A----G--G----AG------T------------------------A--------------G-G-----------A-----C----TG-AG-G--AA-GTTA-----G-------C--C-----ATG--A--G----  619
21_A2D.KE.91.KNH1254          ---------------TA-----------------------A--------T------------------------A---GR---------TG---G------------------------A-----------------A----------------C---AAGGT---GCA--  349
22_01A1.CM.01.01CM_0001BBY    ----T--------AGT--CA-----------------A--A---------------A------------------A-----C-------TG----------------------------T--A---------GATA-----T--G--C---C----GA-------G-G--  349
23_BG.CU.03.CB118             C---T-G------G-T--CA-A------A----G---A--AG-------T----G--------------------A---------------G------------A-----C----TG-AG-G--AA-GTTA-----G-------C--C-----ATG--A--G----  593
24_BG.CU.03.CB378             C--A--G------G-T--CA-A------A----G--GA--AG-------T-----------------------------A------------G-G-----------T-----C----GG-AG-G--AA-GTTA-----G-------C--C-----ATG--A--G----  593
24_BG.CU.03.CB471             C--A--G------G-T--C--A------A----G--G----AG-------T------------------------A-------AA-----G------------A-----C----GG-AG-G--AA-GTTA-----G-------C--C-----ATG-----G----  593
25_cpx.CM.01.101BA            C---T-----T--G-T--CA----------C---------A-----------G--A----GT-----------A---T--A----T---------------------------T--A---------GATA-----T--G--C---C----GA-------G-A--  349
27_cpx.CD.97.97CDKTB49        -------------G-T---A-------------------CCA------------CT-----------------C--A---G--A-----G------------A-----C----GG-AG-G--AA-GATA----GG--C------C-----A-------G-A--  355
28_BF.BR.99.BREPM12609        ---------G--------A----------------------------------C-----------GT--------------C---A-----------------------------------------------G--------------------G-----------  513
29_BF.BR.02.BREPM119          ---------------T------------------------A--------T-----------GTT----------------A-----TG----------------C-----G-----------------------------A-GT---CAG--AGC-----  324
31_BC.BR.02.110PA             T---A--------G-T-----------GA--A-G-------T---C-------------------T----------A---G---AC-----G--CG-A-----------C-------A---A-------A---------T-C--C-----A-------G---.  598
32_06A1.EE.01.EE0369          ----G--------G-T--CAG------G-------------AG------T-----------------------------A------------A-AG--GC-------------C----T--A---------ATA-----T--G--C---C----G--------G----  794
33_01B.MY.05.05MYKL007_1      ----T--------A-T--CA--------------------A--------T-------A-------------GG--C--A--C----------------------------A----T--A------------TA----CT--G---C--C----GAA------G----  529
34_01B.TH.99.OUR2478P         ----T----GT--A-T--CA------------G-------A--------T-------A-------------GG----A--C---G------TG---------------------------T--AT------------TA-----T--G---C--C----GA-------G----  349
35_AD.AF.05.05AF095           ----T--------G-T--CA--------A--------------------T--------------------------C--A-------A----CA--------------------C----T--A---------ATA-----T-----C---C----------A--G----  349
36_cpx.CM.00.00CMNYU830       -------------A-T--CA--------------------A---------------------------------A-----C------TG------------------------------T--A------A-GTTA----GT--C--C---C----GA-------G----  349
37_cpx.CM.00.00CMNYU926       ----T-----T--A-T--CA--------------------A--------T--------------------GG----A--C---C------TG---G-------------------------T--AT---------TA---...--T-AGC-A--G-C-CAG---GC-----  346
38_BF1.UY.05.UY05_4752        -------------------------------------G----A---------------A--T-GT------C-T---A----G--A-G-----G----G-------------C----T---T-----------A-----T------C--C----GA----------G-  570
39_BF.BR.03.03BRRJ103         ---A-----------------------------------T-A--------------------GT------------------A------TG----------------------------T--------------------------------------G-----------  618
40_BF.BR.04.04BRRJ115         ----------------------------------------A------G-T----------------------------------------C------------------------------------------------------------------TG--G-----G----  684
42_BF.LU.03.luBF_05_03        -------------G--------------------------A-----------------------------------A-------A-----TG---G----------------------------------------------------------------R-------------  673
43_02G.SA.03.J11223           ----Y-M--RT--A-T---AG-------------------A--------T-------A---M-----T-GG--Y--A--------A-----C--------------------C----T--AG--------AT------T--G--C--AC----GA-------G----  678
U.CD.83.83CD003               -------------G-T-----------------G-------A----A--------CTT-----------------C--A--------------G-----------A-------------T--------------TA---------G-C--AC-------T--A--G----  642
U.CD.90.90CD121E12            -------------G-TT----------------G-------A----C------------------------------C--A-----------------------------A-------------A-------------CA---------G-----C----------A--G----  346
U.GR.99.GR303                 ----T--------G-T--CA--------------------A-----------------------------------C--A--------------TG-------T----------------T--A---------CTA-----A-G-------C----GA-----------  433
U.NL.95.U_NL_95_H10986_D1     G------------G-T-----------A---------------------------------T------------------A---TGT-A----------------------------C-----------------GATT-----------AC---------GA--G-AGCT.... 1037
CPZ.CD.90.ANT                 T--AT-GAGC--T---A-AG-A-T-A---CCCT---A-A-T---T--G-T---C--CA-TTGTGTT--G-GG--C--A---A--G--GA-A------------AG-AC----CG-TA-A-CAG-GA-AATGA-AGT--TGC-G-CAC-AGC-G--A--GGAAGTAGCC  508
CPZ.CM.01.SIVcpzCAM13         G---T--G-G---G-T--GA-A--C--C--G----GG--GC----T----T---C--TC-----GTGT-A-GG-----T--CAGT-AC--C---G--G------AC--A----A-----AC--C-TA--A--C-C--TGGAG--CAAGTA-G----AATG--A-CA---  706
CPZ.CM.05.SIVcpzEK505         ---AT--G-----G-T--CA-A-----G-------G---GC-C--TC---T---C--CC-G---GTT--T--C-----T---A-T-AC---A--G--C-GA-T--AC-------------A--AT--AGA--GA-AATG--AGC-GAAC-A---G--C-TG-A-C--A-C  683
CPZ.CM.05.SIVcpzLB7           ----T--G-----G-T---A-------G------A-C----A---T----T---C--CA---TGGT-----GG--C---------G--A--C-C--------------------G--------T-CAC--ACA-C--TTAG-C-G---CAGC-A--GG--AAGG-AGC---T-  690
CPZ.CM.05.SIVcpzMB66          G------G-----G-T--CA-A-----T-------G----------T--G-T---C--CA---T-GT---T-GG--C---------G---CCCAG-------T--A--A----------TC--C--C-A----C---GC-G---G-A--AC-GG---TGGCTGC---TA  693
CPZ.CM.05.SIVcpzMT145         -------G-----G-T---A-------G-------G---GC-C--T----T---C--TT-G---GT---T-GG----A---A-G--A----AG-GG------AC-AC-G-----C-G--AC-GA-A-CT-CCGTG---CC-GAAC--GGTG--A---CATC-....  709
CPZ.GA.88.GAB1                ---AT--G-G---G-T--CA-A-----C-------G---GC-G--C--G-T---C--TC-G---GTA--G-GG--CA-A---AGTGAC--CACTG--G-------AC--A-----C----AC--C--A--CGGC-T--TGGAG--CAAC---GC---A-TG-AAGTAACT 1181
CPZ.TZ.00.TAN1                C-----G--G--T--GG-AG-C-TT--T--TCC----A-C-TT--T--G-T------CA-CTGTGTA-----C--C--A--CGC-G-AGA-AGAG-CC----T--GG-A-----AG-CA-A-TTG-GA-AATGA-A-T--CTGT-CAG--A--T--CT-CACAGCGA--T  707
CPZ.US.85.CPZUS               ----T--G-----G-T--AA-A--------G----GC--GC----GC-C-T---C--CT----GGT---T-GG-----C--CAGT---G-GACAG-GG-G--T--GC--C-G--C----TG--AC-GA-A--G-TAGTGC-GTTGCAG--AC--C-G-A-G-AA---A-C 1186
CPZ.CM.x.SIVcpzDP943          ----T--G-----G-T--CA-A-------------GC--GC----GC-T-T---C--TC----GGTA--T-GG-----T---A----A--TCCAG-GG------AC--C-G-----ACT--AC-GA-A---ACAGCGGC-CGRGAAC--G--GTG--C--A-C-....  708
CPZ.CM.x.SIVcpzMB897          C---T-G--G---G-T--CACA-------------G---A-----T----T---C--CC-----GT-----GG--C--A--------A--TAGAG----------A--A--G-------------AC----A----T---GC-A-----C--AGTG----T-A-A--G-A-.  688
N.CM.02.DJO0131               C--AT--G-----G-T--CGCA-----G-------G---GC-C--C--------CG-CC-----GTT--T--------C---AGT-AC---C--G--C-CA-T--AC-A--G--G--G---------A-A--GATAA--G-G---AGG-ACGGC-CAAGTCCG-GC--A  614
N.CM.04.04CM_1015_04          C--AT--G-----G-T--CA-T-----G-------G---GC-C--C--------Y--CA-----GTT--T--------C---AGT-CA--CC-G--CTCA-T--AC-A--G-----G----------A-A--GA-AA-GG-AC-GCAC---TCTG-G---A-AA-TC---  616
N.CM.04.04CM_1131_03          C---T--G-----G-T--CA-T-----G-------G---GC-C--C--------C--CT-----GTT--T--------C---AGT-CA--CC-G--CTTA-T--AC-A--G-----G----A-----A-A--GA-AA-GG-AC-GCAC---CCTG-G---A-AA-TC---  615
N.CM.95.YBF30                 T--AT--G-----G-T--CA-------G-----------GC-C--T--------CG-TC-----GTT--T--------C---AGT------C-----C-CA---AC----------G--------------A-A--GA-A--GG-AC-GCAC---CCCG-GC--A-AA-CC---  702
N.CM.97.YBF106                C--AT--G-----G-T--CA-------G-------G---GC-C--T--G-W---C--CC-G---GTT--T--------C---AGT-AC---CC-G--C-CA-T--AC-A--G-------------------A-A--GA-A--GG-AC-GCAC---TCCG-GC--A--A--C---  703
O.BE.87.ANT70                 G---T--G-G---G-T--CA-------G-----CAGC--GCAG---C-C-GG--CG--A---T-GTG----GG-----T--CA-C--ATATA-A--TGG---T--GC--C-G--AA--C-A---T--A------TAATGGGG-GC--G---TCTGCG-ACGCCGCTAAG. 1194
O.CM.91.MVP5180               G---T--G-G---G-T--CA-------G-----G--C--G-A---TC-C-GG--CG--A-----GTA----GG--C--T--CA-C--AT-T--C--CCG---T--AC--C-G--AA--C-A---T--A------TAATGGCA-GC--G---TCTGC--AGGCCGCTAAG. 1169
O.CM.96.96CMABB637            ----T--G-----G-T--CA-------G-----M-RT--G-AG---C-C-GG--C---A----GGTA----GG--C--T---A-C--AT-TA-AG-TG----T--GC--C-G--AA--C-G--AT--A-A----TAATGGGA--C--G---TCTGC--ATACCGCTAA-.  621
O.CM.98.98CMA104              ----T--G-G---G-T--CA-------G------AGC--GCAG---C-C-GG--C---A-----GTG----GG-----T--CA-C--AT-TA-AG-TG-G--T--GC--C-G--AA--C-A---C--AGA----AA-GGGGGGC--G---TCTGCG-GTGCCGCTAAG.  626
O.CM.99.99CMU4122             G---T--G-G---G-T--CA-A-----G-----CAGC--GCAG---C-C-GG--C---A---T-GTG----GG-----T--YA-C--ATATAGA--TG----T--AC--C-G--AA--A-A--AT-GAG-----TAATGGGR-GC-RG--ATCTGC-AGTG-TGCTAAG.  625
O.FR.92.VAU                   ----T--G-G---G-T--CA-------G------AGC--GTAG---C-C-GG--CG-TA-----GTG----GG-----T---A-C--ATATA-AG-TG----T--AC--C-G--AG--AGA--AT--A----G--AATG--G-----A---CCTGTCAGTAGTACTAAG.  703
O.SN.99.SEMP1299              G---T--G-G---G-T--CA-------G------AGC--GCAG---C-C-GG--C--TA-----GTG----GG-----T--CA-C--A--TA-AG-TG----T--GC--C-G--AA--C-G--AC--A------TAATGGGG-GC--G---TCTGC--GTACCGCTAAG. 1196
O.US.99.99USTWLA              G---T--G-----G-T--CA-------G-----CAGC--GCAG---C-C--G--CG--A---T-GTA--TAGG-----T--CA----ATATAGAG-T--------AC--C-G--AA--C-A--AT--AR-----TAATRRGG-GC-RG---TCTGCG-ACA-TA-R....  614
```


HIV-1/SIVcpz
Complete Genomes

Gag p17 Matrix end Gag p24 Capsid start

```
B.FR.83.HXB2               CAGCT..............................GACACAGGACACAGCAATCAG.......................................GTCAGCCAAAATTACCCTATAGTGCAGAACATCCAGGGGCAAATGGTACATCAGGCCATATCACCTAGAACTTTA 1245
Gag                        A__A__.__.__.__.__.__.__.__.__.__.__D__T__G__H__S__N__Q__.__.__.__.__.__.__.__.__.__.__.__.__.__V__S__Q__N__Y__P__I__V__Q__N__I__Q__G__Q__M__V__H__Q__A__I__S__P__R__T__L_
A1.GE.99.99GEMZ011         ...................................ACA--T-RCACA..................................................-G---T-----------C--------A--TGCA--A---------A----C---T----G--------G-----G 435
A1.KE.00.KER2008           -----..............................--T------A---A--GCA--.......................................--------T-----------------A--TGCA--A---------A--------T--T-------A--G-----G 450
A1.KE.00.KNH1144           -----..............................--T------AGTG---GCA--GTCAGC.................................--------------------------A--TGCA--A---------A----C---A-TT-G--------G--C--G 456
A1.KE.00.KSM4024           -----..............................--T------A------GCA--.......................................--------------------------A--TGCA--A---------A----C---T--T-G--------G-----G 447
A1.KE.00.MSA4069           -----..............................--T------A------GCACA.......................................--T--T--------------------A--TGCA--A---------A--------C--T-G--------------- 450
A1.KE.00.NKU3005           -----..............................---------AG-----GC---.......................................-----T--------------------A--TGCA--A---------A----C---AA-T-G--------G-----G 450
A1.RU.00.RU00051           --A--..............................-G-G-----A-------CA--.......................................-----T--------------------A--TGCA--A------------Y-----T--C-R--------G-Y---G 549
A1.RW.93.93RW_024          -----..............................---------A------GCAG-AGC....................................--------------------------A--TGCA--A--------------C---C--T----------G-----G 453
A1.SE.95.SE8891            -----..............................--------GA------GCA-T.......................................-------G------------------A--TGCG--A--------------C--AT--T----------G-----G 450
A1.SE.95.UGSE8131          -----..............................---------A------GCA--.......................................---------------------A----A--TGCA--A--A-----------C--A-T---G--------G-----G 641
A1.TZ.01.A173              -----..............................---------G------GCA--.......................................---------------------A-A--A--TGCA--A---------A----C---AA-T-G--------G------ 450
A1.UA.01.01UADN139         --A--..............................-G-------AG-----GCA--.......................................-----T--------------------A--TGCA--A---------AC---C---T----G--------G-----G 450
A1.UG.92.92UG037           -----..............................---------AGT----GCA--.......................................--------------------------A--TGCA--A---------A-C--C---T--T-G--------G-----G 1264
A1.UG.99.99UGA07072        ---T-..............................--T------AG-----GCA--.......................................--------------------------A--TGCA--A---------A----C---C--G-------A--------G 450
A2.CD.97.97CDKTB48         -----..............................---------AG-----GCAGTCAAAATTACAGAGGTAGC.....................AG---T--------------------A--TGCA--A--------------C------G-G--------G-----G 610
A2.CY.94.94CY017_41        ...................................--------GA------GC..........................................-...-T--------T--C--------A--TGCA--A--------------C-----T-----------G--G--G 602
A.SN.01.DDI579             --A--..............................---------A------GA---.......................................-----T--------------------A--TGCA--A------------------T--T----------G-----G 452
A.SN.01.DDJ369             ----AGCT...........................---------AG-----GA---.......................................--------------------------A--TTCA--A--------------C---T--T----------G-----G 455
A.SN.96.DDJ360             -----..............................---------A------GA---.......................................-----T--------------------A--TGCA--A------------------T--T----------G-----G 452
A.CD.02.02CD_KTB035        ----CGCT...........................-CAGG--ACAG------A..........................................-----T-----------C--------A--TGCA--A-----G------------T----------C--G-----G 1242
A.CD.97.97CD_KCC2         -----AGT...........................-CAGG--ACAG-----GA..........................................-----T--------------------A--TGCA--A--------------C------T----------G-----G 1247
A.CD.97.97CD_KTB13         ...................................A-TG-----G---A--GCA--.......................................-----T-----------A--------A--TGCA--A--------------C------T----------G-----G 455
B.AR.04.04AR151516         --A--..............................---------A-T-A--GC---.......................................-----T------------------------T-A--A---------------------T----------------- 452
B.AU.87.MBC925             -----..............................---------A------GCA-ACAG....................................--------------------------------G------------------------------------------ 1250
B.BO.99.BOL0122            -G---..............................---------A--------GGAAACAGCAATCAG...........................--------C----T----------------C-T--------G------------T--C----------------- 462
B.BR.03.BREPM2012          ----ACAGCAAGCGGCAGCT...............A--------A---A--GC---.......................................--T--T-----C------------------C-T--------------------AC-------------------- 655
B.CA.97.CANB3FULL          -C---..............................-G-----A-A------GC---.......................................------------------------------C-A------------------------C----------------- 525
B.CN.05.05CNHB_hp3         -----GCT...........................---------A---A--GC---.......................................--------C-----------------A---C-T---------------------C-------------------- 1244
B.CO.01.PCM001             -----..............................---------A------GC---.......................................--------G---------------------C--------------------------T----------------- 450
B.GB.83.CAM1               -----..............................-G-------A------GC---.......................................------------------------------C-A------------------------------------------ 1246
B.GB.86.GB8                -----..............................-----AG---A-------CAGCAACCAG................................------------------------------C-------------------------------------------- 1250
B.GE.03.03GEMZ010          -----..............................-------A-A---AA--C---.......................................-----------C------------------C-T-----------------C--A-----------A--------- 450
B.IT.05.SG1                -----GCT...........................---------A---A--GC---.......................................-G------------------------------T------------------------C---------------N- 1047
B.JP.05.DR6538             -----..............................---------A------GC---.......................................------------------------------C-A--------------------A--------------------- 1249
B.KR.05.05CSR3             -----..............................-------C-AG--A--GC---.......................................--------------------------------------------------------------------------- 1270
B.NL.00.671_00T36          AGCAAGCAGAAGCT.....................---G-----A-A-A---C-C-.......................................--------G--------------------TC----A--------------------------------------- 780
B.RU.04.04RU128005         -G---..............................AG-------A------GC---.......................................------------------------------C----A-----G--R--------------------------C--- 736
B.TH.00.00TH_C3198         ----AGCAGCT........................-G-------A------GC---.......................................--------------------------A---------------------------C-------------------- 456
B.UA.01.01UAKV167          -----..............................---G---C-A---AA-GC--A.......................................-----------------------------TC----------------------A---C-G--------------- 450
B.US.04.ES10_53            -----AACACAGAAACT..................A--------A------GC---.......................................--T-----------------------------T-G------------------------------------------ 1257
B.US.99.PRB959_03          -----..............................-C-------A---A--GC---AGCAGCCAG..............................------------------------------C-A--A------------------C-AC-------C--------- 465
C.AR.01.ARG4006            ...................................--AG-G-CTG---AAGGAA--.......................................-----T--------T--------------TC----A--------------C---C--C-G--------------G 444
C.BR.04.04BR013            ...................................--AG-G-CTG---AAG-AA--.......................................A----T--------T--------------TCA---A------TG----------C----C---G-C--------G 700
C.BW.00.00BW07621          ...................................--A--G-CTG---AA--G...........................................-----T--------T--------A-----T-----A--------------C---T--T----------------G 613
C.CN.98.YNRL9840           ...................................A-AGAG-ATG--G-G--G...........................................-----T--------T--------------TC----A--------------C-----------------------G 447
C.ET.02.02ET_288           G-.................................--AG-T-ACA-AG-A--G...........................................-----T--------T--------------TC-------------------C---T--C--------------C-G 444
C.GE.03.03GEMZ033          ...................................--AG-G-CTG---AG--G...........................................-----T--------T--------------TC----A--------------C-----------------------G 441
C.IL.99.99ET7              ...................................ACTG-GACGG--GAA-CA...........................................-----T--------T--------A------C----A------------T-C--AAATT----------------G 441
C.IN.99.01IN565_10         A--GGGCT...........................---GGGAA-GT---TC-AA-AGGGGCTGACGGGAAG........................-----T-T------T--------A-----TC-A--A--------------C-----T-----T-----------G 644
C.KE.00.KER2010            ...................................A-AG-G-CTG-TG-A--G...........................................-----T--------T--------A-----TC----A--------------C------C----------------G 441
C.MM.99.mIDU101_3          ...................................A-AGAG-CTG--G-G--G...........................................-----T--------T--------------TC----A--------------C--------------C--------G 614
C.MW.93.93MW_965           ...................................--AG-G-CTG-T-AAGGAA--.......................................-----T--G-----T--A-----------TCA------------------C--------------G--------- 444
C.SN.90.90SE_364           ...................................--AG-G-CTGG-G-A--G...........................................-----T--------T--------A-----TC----A------------------C-TC----------------G 441
C.SO.89.89SM_145           ...................................A-AG-G-CTG--G-A--G...........................................-----T--------T--------------TC-A--A--------------C-----------------------G 441
C.TZ.02.CO178              ...................................A-AG---CTG--G-G--A...........................................-----T--------T--------------TC----A--------------C---C-------------------G 441
C.UY.01.TRA3011            ...................................-CGG-T-ACA-AG-A--G...........................................-----T--------T--------------TG----A--------------C---C--T-----G-C--------G 432
C.YE.02.02YE511            ...................................ACAG-G-AGG---AA--A...........................................--T--T--------T--------------T-----A--------------C--AT-TC----------------G 441
C.ZA.04.04ZASK164B1        A-.................................A-AG-G-CTG--G-G--G...........................................-----T--------T--------------TC-A--A-----------------A--------------------G 674
C.ZA.05.05ZAPSK240B1       ...................................-CA-AG-CTG--G-A--G...........................................-----T--------T--A-----------TC----A--------------C---C----------G--------- 667
C.ZM.02.02ZM108            AG--A..............................--AG-G-CTG---AA--G...........................................-----T--------T--------------TC----A--------------------------------------- 1241
D.CD.83.ELI                -----..............................---------A---A--GC---.......................................--------------T-------------------C-A-----------------------------------------G 791
D.CM.01.01CM_0009BBY       ...................................A---A----A---A--GC---.......................................------------------------------T-A-----A--------------A-----------------C--G 438
D.KE.01.NKU3006            -----..............................------AA-AG-----GC---.......................................-----------C--T-----------A---C-A-----------------C------T-------------C--G 450
D.KR.04.04KBH8             -----..............................------AA-A---A--GC---.......................................--------------T-----------A---C-A-----------------C---C--C--------------C-G 1204
D.TD.99.MN011              T----..............................---------A-------C---.......................................--------C---------------------T-A-----------------C--A---C-------------C--G 463
D.TZ.01.A280               -----..............................---------AG-----GC---.......................................--------------T-----------A---C-A-----------------C--------------------C--G 440
D.UG.99.99UGK09259         -----..............................---------A------GC---.......................................------A------------------GA-----A-----------------C---C-------------------G 450
D.YE.01.01YE386            -----..............................------A--A-----CCG..........................................--------------T-----------A---C-A-----------------C--------------------CC-G 447
D.YE.02.02YE516            -----..............................---------A------GC---.......................................------------------------------T-A-----------------C--A-----------------C--G 450
D.ZA.90.R1                 -G---..............................---------A------GC---.......................................--------------T-------------------C-A---------------------C-------------------G 591
F1.AR.02.ARE933            -----..............................--------G............................................-----T------------G----A-----TCCT-----A---------T-C---T--T----------------- 542
F1.BE.93.VI850             -----..............................----A---G............................................-----T-----------------A-----TC-T-----A-----------C---T-TC----------------- 588
F1.BR.01.01BR125           A---GGCAGCT........................--A-A---G............................................-----T-----------------A-----TC-T-----A-----------C---T-TT----------------- 693
F1.ES.x.P1146              -----..............................----A---.............................................-----T-----------------A--A--TC-T-----A-----------------TC----------------- 531
F1.FI.93.FIN9363           -----GCAGCT........................----A---G............................................-----T-----------------A-----TC-T-----A-----------------T------------------ 580
F2.CM.02.02CM_0016BBY      -G---..............................--T-A---G............................................-----T-----------------A-----TC-T-----------------------T---------------C-- 438
F2.CM.95.MP257             -G---GACAAAGGGGTC..................AGT-A---G............................................-----T---------------C-A--A--TC-T-----------------C---AGTC----------------- 457
F2.CM.97.CM53657           -G---..............................----A---G............................................-----T-----------------A-----TC-T-----A-----------C---T-TC----------------- 438
G.BE.96.DRCBL              ......................................GA-AACAG-----GC--A.......................................-----T-----------------------TGCA--A--------------C-----------------------G 1209
G.CM.01.01CM_4049HAN       --AGA..............................--TGA----G-------C--A.......................................AG-----------------------------TGCA--A---------A----C-----A---A----C--------- 450
G.CU.99.Cu74               --AGG..............................--TGA----A---A--GC--A.......................................-----T--------T--------------TGCA--A------CCA-----C------------------------ 858
G.ES.99.X138               --ATG..............................---GA----A---A--GC--A.......................................-C------------T--------------TGCA--A-----------------------------C--------- 695
G.GH.03.03GH175G           AG.................................-CTGA----A---A---C--A.......................................--------------T--------------TGCA--A--A-----------C--AC----------------C--- 1285
G.KE.93.HH8793_12_1        --AGG..............................--TGA----A------GC--A.......................................--------------T---------------GCA--A--A--G--------C------------------------ 647
G.NG.01.01NGPL0669         --A-A..............................AGT------A-T-A---C--A.......................................--T--T--------T--------------TGCA--A--------------C------------------------ 450
G.PT.x.PT2695              TG.................................---GA----A------GC--A.......................................--------------T--------------TGCA--A-----------------------------C--------- 1193
G.SE.93.SE6165             --ATG..............................--T-A----A-------C--A.......................................--------------T--------------TGCA--A--------------C---------A-------------G 648
H.BE.93.VI991              ----AGCAGCT........................--T-A--A-A-GGA--GCA--.......................................A----T--------T--------A-----TGC-------------------C-----A-----------G--C--- 637
H.BE.93.VI997              -----..............................A-T-AG-A-AGAGA---CA--.......................................-----T--------T--------A-----TGCT--A-----G--------C---C---------V---G--C--- 572
H.CF.90.056                -----..............................--T-AG-A-A-AGA---CA--.......................................-----T--------T--------A-----TGCT--A-----G--------C-----------------G--C--- 592
J.CD.97.J_97DC_KTB147      -T.................................-T--A-AA-T---A--G----.......................................--G--T--C-----T------T----A--T-GT-A----G--C-A-----C------C-------------A--- 440
J.SE.93.SE7887             A-A--..............................----A-AA-G---A--G----.......................................-----T--------T--------------TC-G--A------CC------C------C----------------- 568
K.CD.97.EQTB11C            ...................................C-AGG-AA-GCTGA---AGG-.......................................--T--T-----------------A-----TC-T-----------------C------C----------------- 444
K.CM.96.MP535              ...................................C-AGA--C-GCTGA---AGG-.......................................-----T-----------------A-----TC-G-----------------C------C----------------- 444
```

```
                                                                                             Gag p17 Matrix end_Gag p24 Capsid start
B.FR.83.HXB2                CAGCT..............................GACACAGGACACAGCAATCAG....................................GTCAGCCAAAATTACCCTATAGTGCAGAACATCCAGGGGCAAATGGTACATCAGGCCATATCACCTAGAACTTTA 1245
Gag                         A__A__.__.__.__.__.__.__.__.__.__.__D__T__G__H__S__N__Q__.__.__.__.__.__.__.__.__.__.__.__.__.__V__S__Q__N__Y__P__I__V__Q__N__I__Q__G__Q__M__V__H__Q__A__I__S__P__R__T__L_
01_AE.CF.90.90CF11697       -----...............................-G-------AG--A--GCA-A....................................--------------------------A--TGCA--A------------------C--T----------------G 1200
01_AE.CN.05.FJ051           -----...............................-G-------AG-----GCA-A....................................--------------------------A--TTCA--A------------------C-TC-T--------------G 1261
01_AE.CN.06.FJ054           -----...............................---------AG-----GCA-A....................................--------------------C-----A--TGCA---------------------T-TT-------C--------- 1267
01_AE.HK.x.HK001            -----...............................-G-------AG--A--GCA-A....................................--------------------------A--TGCA--A--A------A--------T-TT----------------G 622
01_AE.TH.01.01TH_R2184      -----...............................---------AG-----GCA-A....................................--------------------------A--TGCA--A------------------C-T-----------------G 450
01_AE.TH.02.OUR769I         -----...............................-G-------AG-----GCACA....................................--------------------------A--TGCA--A------------------C-TG----------------G 450
01_AE.TH.90.CM240           -----...............................-G-------AG-----GCA-A....................................--------------------------A--TGCA--A----------C-------C-TT----------------G 810
01_AE.US.00.00US_MSC1164    -----...............................-G-------AG----GGCA-A....................................--------------------------A--TGCA--A--------------C---C-TG----------------G 450
02_AG.CM.02.02CM_4082STN    -----...............................-C-------AG-.............................................AG------------------------A--TGCA--A------------------AG---G--------------G 441
02_AG.EC.x.ECU41            -----...............................-C-------AG-.............................................AG---------C--------------A--TGCA--A------------------T-------------G-----G 347
02_AG.GH.03.GHNJ196         -----...............................-C-------AG------C--AGTAAGAGTC...........................AG-----------T------------A--TGCA--A---------A--------T----G--------G------ 1277
02_AG.NG.01.PL0710          ....................................-C-------GG-.............................................AG---------------------------TGCA---------------------T----G--------G------ 435
02_AG.NG.x.IBNG             -----...............................-C-------AG-.............................................AG------------------------A--TGC-A-A---------AC-------T----G--------------- 776
02_AG.SE.94.SE7812          -----...............................-C-G-----AG-.............................................AGT-----------------------A--TGCA--A---------AC-------T----C--------G------ 623
02_AG.SN.98.MP1211          -----...............................-C-------AG-.............................................AG-----------T------------A--TGCA--A-----G------------T-------------G-----G 448
03_AB.RU.97.KAL153_2        --A--...............................-G-------AG-----GCA--....................................-----T-----------C--------A--TGCA--A---------AC---C---T----G--------G-----G 477
04_cpx.CY.94.CY032          -----...............................-C-G----GGT-----GCA-T....................................--------------------------A--TGCA--A------------------AG---T--------------G 611
05_DF.BE.x.VI1310           ---AAGCT............................-G-G----GA------GC---....................................-C------------T-----------------A-----------------C------C----------------G 624
06_cpx.AU.96.BFP90          -T..................................-C-------A------G-A-T....................................C----T-----------C--------A--TGCA--A-----------------------G-----AC-------- 1270
08_BC.CN.97.97CNGX_6F       ....................................-CA-A-AA-GCTGAAG-GA--....................................-----T--------T--------------TCC---A------------------C--C----------------- 443
09_cpx.GH.96.96GH2911       -----...............................A--------A---A--GCA--....................................-----------------A-----------TGCA--A-----G-----------------C--------------R 456
10_CD.TZ.96.96TZ_BF061      -----...............................---------A------GC---....................................--------------T-----------A--TT-A-----------------C---C--T--A-------------G 624
11_cpx.GR.x.GR17            ....................................---T-----A-------CA--....................................-----T-----------------------TGCA--A--------------C--------G-----------C--G 556
12_BF.AR.99.ARMA159         -----...............................----A---G................................................-----T--G--------------A-----TC-T-----A-----------C---C-TT----------------- 1253
13_cpx.CM.96.1849           --AAG...............................A-TGA----A-------C--A....................................-----T--------T--------------TGCA--A-----------G--C---T-------------------G 623
14_BG.DE.01.9196_01         TG..................................---GA----A-T-A--GC--A....................................-C------------T--------------TGCA--A---------------------T-------C--------- 768
15_01B.TH.99.99TH_MU2079    -G---...............................-G-------AG-----GCACA....................................--------T-----------------A--TGCA--A------------------C-TG----------------G 653
16_A2D.KR.97.97KR004        ....................................-CAG-T-ACACAG-A-GCAGC....................................-G---T-----------C--------A--TGCA--A---------AC-T-----AA-T----------G-----G 606
17_BF.PE.02.PE02_PCR0155    -----...............................A---A---G................................................--T--T-----------------A-----TC-T-----A--------G--C---T-TT----------------- 438
18_cpx.CU.99.CU76          -----...............................-C-G-----AG-----GC---....................................-----------------C--------A--TGCA--A------------------T-------------G-----G 562
19_cpx.CU.99.CU7           -----...............................------A-GA------GC--C....................................--------------T-----------C--GC-G------------------------C-G--------------G 471
20_BG.CU.99.Cu103          --A-G...............................-GTGA----A------GC--A....................................--------------T--------------TGCA--A------CCAA----C-----------------G------ 720
21_A2D.KE.91.KNH1254       --A-AGCT............................------AAGA------GC---....................................-----------------C--------A--TGCA--A--A---TG------C--A--------------G--C--G 453
22_01A1.CM.01.01CM_0001BBY -----...............................-G-------AG-----GCA-C....................................--------------------------A--TGCA--A-----G---A----------TT----------------G 450
23_BG.CU.03.CB118           --A-G...............................-GTGA----A-----GGC--A....................................--------------T--------------TGCA--A------CCA-----C-----------------G------ 694
24_BG.CU.03.CB378           --A-A...............................-GTGA----A------GC--A....................................-----------------------------TGCA--A------CCAA----C--------------A--G------ 694
24_BG.CU.03.CB471           --A-G...............................-GTGA----A------GC--A....................................-----------------------------TGCA--A------CCAA-G--C------T----------G------ 694
25_cpx.CM.01.101BA          -----...............................---------AGT----GCAGCCAG.................................-----------------C-----------TGCA--A------------------T-TC----C------------ 453
27_cpx.CD.97.97CDKTB49      --AAG...............................----AG---AT--A--GC--A....................................--T-----------T-----------A--TGCA--A--------------C--A------A----------C--G 456
28_BF.BR.99.BREPM12609      -----...............................---G-----A------GC--A....................................------------------------------C--------------------------C-G--------------- 614
29_BF.BR.02.BREPM119        -T..................................-G-------A-T----GC---....................................--------G--------------------------------------G-----AC-----G-------------G 422
31_BC.BR.02.110PA           ....................................--AG---CTG---AAGGAA--....................................-----T--------T--------A-----TC----A--C-----------C-----------------------G 694
32_06A1.EE.01.EE0369        --A--...............................-C-------A--.............................................AG---T-----------C--------A--TGCA--A---------AC---C---T-------------G-----G 886
33_01B.MY.05.05MYKL007_1    -----...............................-G-------AG--A--GCA-A....................................--------------------------A--TGCA--A------------------T-TT----------------G 630
34_01B.TH.99.OUR2478P       -----...............................-G-------AG-G--RGCA-A....................................--------------------------A--TGCA--A--A---------------C-TG----------------G 450
35_AD.AF.05.05AF095         -----...............................--T------AG------G.......................................-----------------------A--A--TGCA--A--------------C---T--T-G--------G--C--G 447
36_cpx.CM.00.00CMNYU830     -----...............................---------AG-----GCACA....................................-------G---------C--------A--TGCA--A--------------------TT----------------G 450
37_cpx.CM.00.00CMNYU926     -T..................................-G-------AG----TGCA--....................................A-------------------------A--TGCA--A---------A--------T--T-G-----C--G-----G 444
38_BF1.UY.05.UY05_4752      -----...............................-GACAG--G................................................-----T-----------C-----A-----TC-T-----A-----------C---T-T------------------ 659
39_BF.BR.03.03BRRJ103       -----...............................---------A---A--G----....................................-------------T----------------C-T------------------------C----------------- 719
40_BF.BR.04.04BRRJ115       -C--AGCAGCT.........................---------A------GCA--....................................-----------------------------TC-G------------------------C----------------- 791
42_BF.LU.03.luBF_05_03      -----...............................---------A-------C---....................................-----------------------A------C-------------------------------------------- 774
43_02G.SA.03.J11223         -----...............................-C-------AG-.............................................AG------------T--------------TGCA--A------------------T--T----------G------ 770
U.CD.83.83CD003             -----...............................----A----AG--A--G----....................................--------------T--------A-----T--G--A--------------------------------------- 743
U.CD.90.90CD121E12          -----...............................----A----G---A--GCA--....................................-----T--------T--------A-----T--G--A--------------C--AC-------------------- 447
U.GR.99.GR303               -----...............................---------GG--C--GCA--....................................--------------------------A--TGCA--A---------A--------C-T--G--------G-----G 534
U.NL.95.U_NL_95_H10986_D1   ....................................---------A-A-C--G----....................................-----------------C-----A-G---TC-T-----------------C------T-------------C--- 1133
CPZ.CD.90.ANT               A-A-CGCAAGCAGAGGCATGCTTCTGCGGCTGCTCCTGTT-AACA-ACAGTGGTGTCAGCGACATCTTAGTGGC..................-AAG-GAG---------C--CA-AGT-G-TGCAGGA--AAT-GCAAGG------C-AC-GA----A-----C--- 657
CPZ.CM.01.SIVcpzCAM13       A-AACAATGAAAGCAGACAGGGC.............-CT-GT-CTTC-TCTGGC.......................................ACT--TGGG--------------A-----TGCT--A-----G---------ACTC----C--C--------C--- 822
CPZ.CM.05.SIVcpzEK505       A---AGCA............................-GGG---CGGCAGCACC-G-AAGCAGT..............................A----TAGG--C--T---C-G--A--A--TGCT--A------------------C-GC-G-----C-----C--- 793
CPZ.CM.05.SIVcpzLB7         ----CGCGGCAAGGGAC...................A-TGACA-T................................................-GT---AG---C------------ACA--TGCA--A-----GCC------C------C-C--------------- 791
CPZ.CM.05.SIVcpzMB66        -C--C...............................---GGTACTAG--CT..........................................-----TAGG--C-TT--------AGCA--TGCT--------G-----------AC----C-G------------- 788
CPZ.CM.05.SIVcpzMT145       .............................................................................................-CT---GCG--C-T----G----A-GA---GCT-----A--GT-ACG----A-TC-A--------C-----CC-- 784
CPZ.GA.88.GAB1              ---GAAGCCGTGAAGGGGGAGCC.............AGTCA---CGCT--TGCCTCTGCTGGC..............................A-T--TGG---------CC-C--A--A--TGCT--A-----G-----G-----------C-----C-----C--G 1306
CPZ.TZ.00.TAN1              -TAG-GGACAAAGACAGAATGCAGGTGAAAAAGAG--A----TG-CACCT-G-GGCAATACAGGAAACACAGGGAGAGCAACAGAGACACCTAGT-GG--TAG-CTA-----AG-GA-AACTG-TGCA-----AGTTGCAAGG------C-T--T------------C-- 877
CPZ.US.85.CPZUS             A-CAGCAACAGGAGGCA...................AGTGGCA-TA---TT..........................................-GT---AGC--------AG--A-A------GCT-----A--G--------------A--G--T--C--------- 1293
CPZ.CM.x.SIVcpzDP943        ....................................C-GC-GCAG--AGA--GCA-T....................................-----TAG---C-T--AG-----A-----TGCA--A------T--A-G------C----G--------------- 804
CPZ.CM.x.SIVcpzMB897        ....................................-C-G-G-CGGCAGA-CCCAGCGGCGGCGAG...........................---AG---C-T---G-----AACA--TGCT-----A--------C------T-GC-C-----C--------- 790
N.CM.02.DJO0131             A-AACCCAGAAGCA......................-GGG---CGGCAGCAGC-G-TAACAAT..............................A----TAGG-----T---C----C-----TGCT--A--A------A--------C-GC-GA----C--------- 730
N.CM.04.04CM_1015_04        A---A...............................-GGG---CGGCAGCAGC-G-TAACAAT..............................A----TAGG--C--T---C----C-----TGCT--A--A--------------TC-GC-GA----C--------- 723
N.CM.04.04CM_1131_03        A---A...............................-GGG---CGGCAGCAGC-G-TAACAAC..............................A----TAGG-----T---C----T-----TGCT--A--A------A--------C-GC-GA----C--------- 722
N.CM.95.YBF30               A---A...............................-GGG---CGGCAGCA-C-G-TAGCAAT..............................A----TAGG-----T---C----C----CTGCT--A--A---------------C-GC-GA----C-----C--- 809
N.CM.97.YBF106              A---A...............................-GG----CGGCGGCAGC-G-CAGCAGT..............................A----TAGG-----T---C----C-----TGCT--A--A---------------C-GC-GA-------------- 810
O.BE.87.ANT70               ....................................--AGACAC-AG-GCA-GG---....................................-CGG-T-----------------ATCA--TGCG-----A--------------------C--C--C--G------ 1290
O.CM.91.MVP5180             ....................................--AGA-AC-AG-CCT-GG---....................................ACA--T-----------------AACA--TGCA-----A--------------A-----C--C--C--G------ 1265
O.CM.96.96CMABB637          .......................................GA-AACA--TCA--G---....................................ACA--T-----------AG----AGCA-----G--A------------------C--C-C--C--C--G------ 714
O.CM.98.98CMA104            ....................................--AGACAC-AG-GCA-GG---....................................ACGG-T------------G----A--A--TGCG-----A------------------C-C--C--C--G------ 722
O.CM.99.99CMU4122           ....................................--AGACAC-AGYGCA-GG---....................................ACGG-T---------------A-A-CA--TGCA--A--A--------G------C----C--C--C--G-----G 721
O.FR.92.VAU                 ....................................--A-ACAC-AG-TCA--G---....................................ACAG-T------------G-----GCA--TGCA-----A--------------A---C-C--C--C------C-- 799
O.SN.99.SEMP1299            ....................................--AGACAC-AG-GCA-GG---....................................ACGG-T-----C-----------AACA--TGCA-----A---------------T--C-C--C--C--G------ 1292
O.US.99.99USTWLA            ....................................--AGACAC-AR--CA--G---....................................ACA--T------------------ACA--TGCG---R-A---GCT---------C--C-C--C--C--------- 710
```

```
B.FR.83.HXB2     AATGCATGGGTAAAAGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTGATACCCATGTTTTCAGCATTATCAGAAGGAGCCACCCCACAAGATTTAAACACCATGCTAAACACAGTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGC 1415
Gag              N  A  W  V  K  V  V  E  E  K  A  F  S  P  E  V  I  P  M  F  S  A  L  S  E  G  A  T  P  Q  D  L  N  T  M  L  N  T  V  G  G  H  Q  A  A  M  Q  M  L  K  E  T  I  N  E  E  A
A1.GE.99.99GEMZ011   ----------G-----A----------A----------A----------G----------TG----G----T----------C-G----T----------G----T----------  605
A1.KE.00.KER2008     ----------G---A------A-G--G---T----------A----------C----------T-TG----G----T----------C-G----T----------A----------  620
A1.KE.00.KNH1144     ----------A------A----------A----------T----------T-TG----G----T----------C-G----------G--T----------  626
A1.KE.00.KSM4024     ----------A------A----------A----------T----------T-TG----------T----------C-G----T----------T----------G--  617
A1.KE.00.MSA4069     ----------A------A---T----------A----------C----------T-TG---T-G----T----------C-G----T----------T----------  620
A1.KE.00.NKU3005     ----------G-----A------A-G----------A----------C----------C----------C-------T-TG----G----T----------C-G----T---G----------T----------  620
A1.RU.00.RU00051     ----------R-----A------A----------A----------CA----------Y-TG---Y-G----T----------C-G----T----------G----T----------  719
A1.RW.93.93RW_024    ----------C-----A------A----------A----------C-----C----------T-TG----G--TGT----------C-G----T----------T----------  623
A1.SE.95.SE889I      ----------G-----A------A---G----------A----------C----------GGGT--C----T-TG----G----T----------C-G----T----------T----------  620
A1.SE.95.UGSE8131    ----------G-----A------A-G----------C---A--T-------C----------T-------T-TG----------T----------C-G----T---G----------T----------  811
A1.TZ.01.A173        ----------A------A----------A----------C----------TGTG----G----T----------C----------A-----  620
A1.UA.01.01UADN139   ----------G-----A------A----------A----------G----------TG----G----T----------C-G----T----------G----T----------  620
A1.UG.92.92UG037     ----------G-----A------A----------A----------C----------G--TG----G----T----------C-G----T----------T----------1434
A1.UG.99.99UGA07072  ----------G----------A-G--G----------T-TG----G----T----------G----T----------T----------  620
A2.CD.97.97CDKTB48   ----------C-------T----------T----------A----------A----------C----T--T----------T----------T----------  780
A2.CY.94.94CY017_41  -----C-----C----------A----------A-----T-----A----------C----T--T----------T----------T----------  772
A.SN.01.DDI579       ----------G-----A------A----------A----------G---A----------T-TG--T-G--T-T----------T----------T----------A----  622
A.SN.01.DDJ369       ----------G-----A------C---G--T----------A----------G----------C----------T-TG----G--T-T----------C----------T----------T----------  625
A.SN.96.DDJ360       ----------G-----A------A----------A----------G----------T-TG--T-G--T-T----------C----------T----------T----------A----  622
A.CD.02.02CD_KTB035  -----C----G----------A-------T----------A----------G----------TG----------T--T----G----------T----------T----------1412
A.CD.97.97CD_KCC2    ----------G--G-------G--A----------A----------TG----------T-TG-----G----------T----------G--T----------1417
A.CD.97.97CD_KTB13   ----------G---C----------A----------C----------TG----------T-------G----------G----T----------T----------  625
B.AR.04.04AR151516   ----------A----------A----------  622
B.AU.87.MBC925       ----------A----------1420
B.BO.99.BOL0122      ----------A----------A----------  632
B.BR.03.BREPM2012    ----------C-------T----------  825
B.CA.97.CANB3FULL    ----------A-------T----------A----------  695
B.CN.05.05CNHB_hp3   ----------G----------T-AT-------A----------A----------G----------G----------G----------A----------1414
B.CO.01.PCM001       ----------A----------A----------G----------  620
B.GB.83.CAM1         ----------A----------1416
B.GB.86.GB8          ----------A----------G----------1420
B.GE.03.03GEMZ010    ----------G----------A-------T----------A----------A----------T----------A----------C----------  620
B.IT.05.SG1          ----------G----------N----------1217
B.JP.05.DR6538       ----------A----------T----------1419
B.KR.05.05CSR3       ----------G----A----C----------A----------A----------A----------1440
B.NL.00.671_00T36    ----------C----------A----------  950
B.RU.04.04RU128005   -------R----------A------A----C----------A----------C----------A--A----------  906
B.TH.00.00TH_C3198   ----------T----------A----------A----------A----------  626
B.UA.01.01UAKV167    ----------A----------G----------T----------  620
B.US.04.ES10_53      ----------A----------A----------GCT---C----------T----------T----A-----1427
B.US.99.PRB959_03    ----------A----------G----------  635
C.AR.01.ARG4006      ----------G---A----G--A----------G--C----------A----------T---T----------T----------G--  614
C.BR.04.04BR013      ----------G----------G--A----------A----------T---T----A----------A----------T----------G--  870
C.BW.00.00BW07621    ----------A----G----------G--A--T----------A----------C----------T----------A----------G--  783
C.CN.98.YNRL9840     ----------A----G----------T-------G--A----------A----------T---T----------T----------A--G--  617
C.ET.02.02ET_288     ----------A------A----------GA-A----------A------G----------C--T----------T----------A--G--  614
C.GE.03.03GEMZ033    ----------A------A---G--T-A------G--A----------A----------G----------C----------T---T----------T----------G--  611
C.IL.99.99ET7        ----------A----G----------G--A----------A----------C----------T---T----------C------T----------G--  611
C.IN.99.01IN565_10   ----------G-----A----G----------T-------G--A----------A----------T---T----AA----------T----------A--G--  814
C.KE.00.KER2010      ----------A----G--A-------T-------G--A----------A----------T---T----A----------G----------T----------G--  611
C.MM.99.mIDU101_3    --C----------A----G----------T-------G--A----------A----------T----------T---T----A----------T----------A--G--  784
C.MW.93.93MW_965     ----------A----G----------C--G----------A----------T---T----------T----------G--  614
C.SN.90.90SE_364     ----------A----G------G----A-----G--A----------A----------C----------T---T----------T----------G--  611
C.SO.89.89SM_145     ----------A----------C--T-------GA-A----------A----------T---T----------T----------G--  611
C.TZ.02.CO178        ----------G--A----G--T-AT-----G--A----------A-----C----------G----------T-G-T----------T----------G--  611
C.UY.01.TRA3011      -----G-------G-------G----------T-------G--A----------T---T----------C----------T----------G--  602
C.YE.02.02YE511      ----------A----G----------A----------A----------C----------T---T----------C------T----------G--  611
C.ZA.04.04ZASK164B1  ----------A----G--A---GC--T-A--T--G--A----------A----------T---T----A----------G--T----------G--  844
C.ZA.05.05ZAPSK240B1 ----------A----------T--T------A-A----------A----------T----------T---T--G----------T----------G--  837
C.ZM.02.02ZM108      ----------G-----A--C-----T--------A----------CGTG---C----T-----T---T----------C----------T----------G--1411
D.CD.83.ELI          --C----------A----------A----------A----------C----------A-----  961
D.CM.01.01CM_0009BBY --C----------A----G--A---------A--------A----------G--G----------A----------  608
D.KE.01.NKU3006      --C----------A----G----------A----------T----------  620
D.KR.04.04KBH8       --C----------AT-----G--A----------A----------TT----------G----------G----------1374
D.TD.99.MN011        --C----------A----G--A-----C---A--------A----------C----------T----------A----------  633
D.TZ.01.A280         --C----------A----------A----------T----------  610
D.UG.99.99UGK09259   --C----------A----G----------AT-------A----------C----------T----------  620
D.YE.01.01YE386      --C----------A----G----------A----------G----------  617
D.YE.02.02YE516      --C----------A----G--A-------T-A--------A----------T----------C-----T----------A----------  620
D.ZA.90.R1           ----------G---A------A-----C----------A----------G---C----------  761
F1.AR.02.ARE933      ----------G--G--G----------T----------A----------G----T----------T---T----------C----------  712
F1.BE.93.VI850       ----------G--GA----------T----------A----------G----T--CAC----------T----------C----------  758
F1.BR.01.01BR125     ----------GA----------T----------A----------G----T----------T---T----------C----------  863
F1.ES.x.P1146        ----------G--G----------T----------A----------G----------T----------C-----T--------G--  701
F1.FI.93.FIN9363     ----------G--GA----------T----------T----------G----T----------T----------C-----T----------  750
F2.CM.02.02CM_0016BBY ----------A----------C----------G----------A----------T----------  608
F2.CM.95.MP257       ----------A----------A----------G----------T----------  627
F2.CM.97.CM53657     ----------A----------A----------T----------  608
G.BE.96.DRCBL        ----------A-----C-----T-------A----------A----------T----------T---------G----------T----------G----T-------T-----1379
G.CM.01.01CM_4049HAN ----------A-----C-----T-------C----------G----------CTCC-------T----------G----------T---------C-T--G--T--T----------  620
G.CU.99.Cu74         ----------C----------A-----C--T--T-------A----------G----------T-----T----------G----------C----G--T--T----------1028
G.ES.99.X138         ----------R-----C-----T-------A-----T----------G----------T----------G----------T---------C----G--T--T----------  865
G.GH.03.03GH175G     ----------A-----C-----T-------A----------G----------T----------G--T--T----------1455
G.KE.93.HH8793_12_1  ----------A-----C-----T-------A----------T----------G----------T---------C----------T----------  817
G.NG.01.01NGPL0669   ----------A------A-----C--T--T-------A----------CA----------T----------T------G----------T---------C----G--T--T--T-----T-----  620
G.PT.x.PT2695        ----------A-----C-----T-------A----------G----------T----------A----G----------T---------C----G--T--T--T----------1363
G.SE.93.SE6165       ----------A-----C-----T-------A----------TCT---T----------G----------T---------C----G--T--T----------  818
H.BE.93.VI991        ----------T----------A----------C----TG----------T----------T--A----------  807
H.BE.93.VI997        ----------G----------T----------A----------C----TG-T-------T----------T--A----------  742
H.CF.90.056          ----------A----------T----------A----------C----TG-T-------T----------G----------T--A----------  762
J.CD.97.J_97DC_KTB147 ----------G--GA----------A-G-----A----------T----------T----------G----------G----------T----------  610
J.SE.93.SE7887       ----------GA------A--A----------G----------T----------A-A-------C-------T----------T--T----------  738
K.CD.97.EQTB11C      ----------T------A----------A----------T----------T--T----------  614
K.CM.96.MP535        ----------G--G---A----G----------A----------A----------T---------C----------T--T-------T-----  614
```


HIV-1/SIVcpz
Complete Genomes

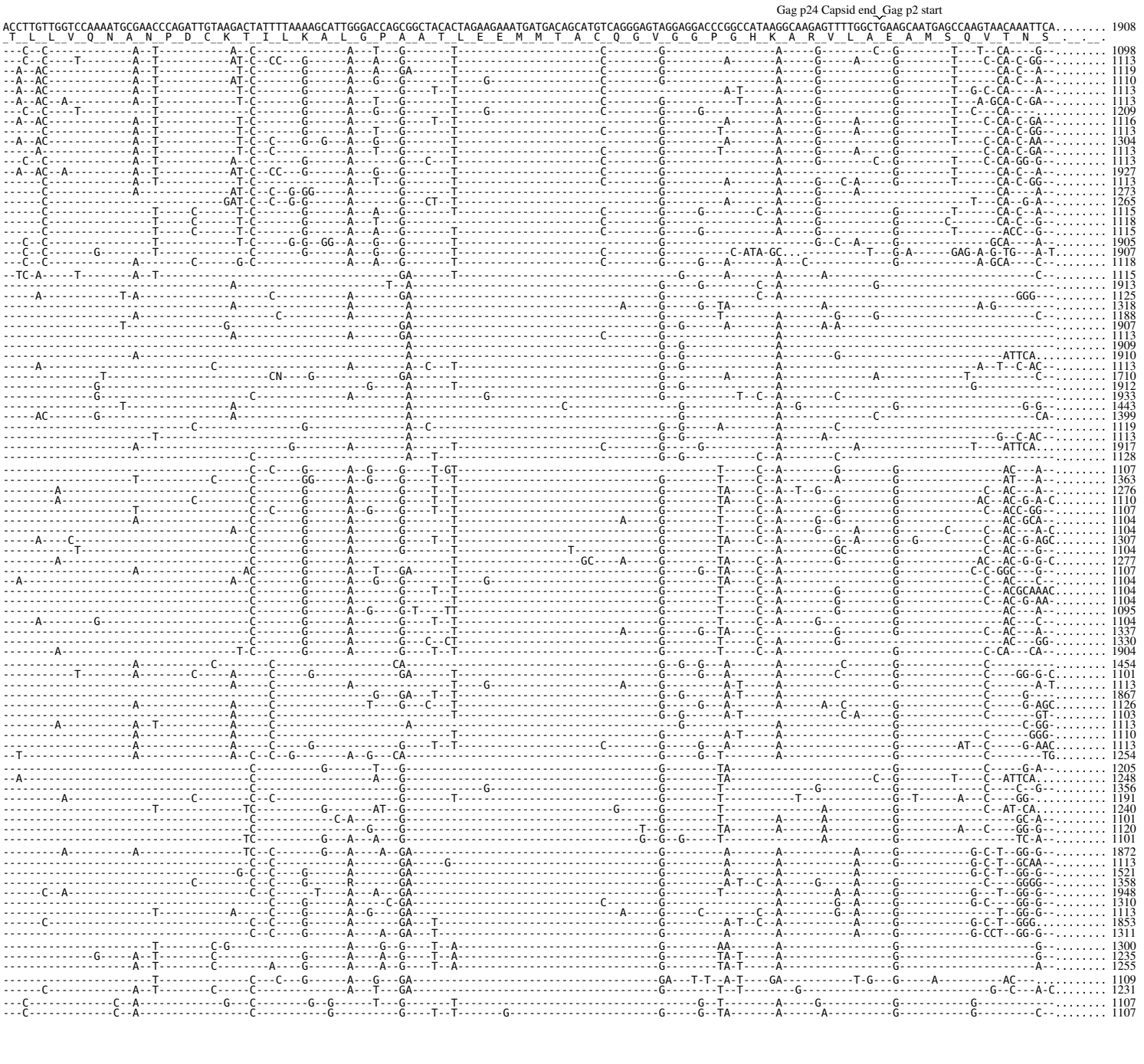
```
B.FR.83.HXB2                  AATGCATGGGTAAAAGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTGATACCCATGTTTTCAGCATTATCAGAAGGAGCCACCCCACAAGATTTAAACACCATGCTAAACACAGTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGC 1415
Gag                           _N__A__W__V__K__V__V__E__E__K__A__F__S__P__E__V__I__P__M__F__S__A__L__S__E__G__A__T__P__Q__D__L__N__T__M__L__N__T__V__G__G__H__Q__A__A__M__Q__M__L__K__E__T__I__N__E__E__A
01_AE.CF.90.90CF11697         --------------------------A----G---T-A---------A-----------C--------------------------------------T-TG----------T---------------G-----A-----------------T----------------- 1370
01_AE.CN.05.FJ051             --------------------------A----G---T-A---------A-----------C--------------G-----------------------T-TG--------T-T------------C--G-----A-----------------A-----------A----- 1431
01_AE.CN.06.FJ054             --------------------------A----G---T-A---------A-----------CA-------------G-----------------------T-TG----------T------A-----C--G-----A-----------------A----------------- 1437
01_AE.HK.x.HK001              -----------G--------------A----GC--T-A---------A-----------C--------------G------G-T--------------T-TG--------T-T------------C--G-----A-----------------A----------------- 792
01_AE.TH.01.01TH_R2184        --------------------------A----G---T-A---------A-----------C--------------G-----------------------T-TG--------T-T------------C--G-----G-----------------A----------------- 620
01_AE.TH.02.OUR769I           --------------------------A----G---T-AT--------A-----------C------C-------G-----------------------T-TG--------T-T---------G--C--------A-----------G-----A----------------- 620
01_AE.TH.90.CM240             -----------G--------------A----G---T-A---------A-----------C--------------G-----------------------T-TG--------T-T------------C--G-----A-----------------A---------------C- 980
01_AE.US.00.00US_MSC1164      --------------------------A----G---T-AT--------A-----------C--------------G-----------------------T-TG--------T-T------------C--G-----A-----------------A----------------- 620
02_AG.CM.02.02CM_4082STN      -----------G------A-------A--------T-----------A------------A-------------G-----------------------T-TG----------T------------C--G-----A-----G-----------T----------------- 611
02_AG.EC.x.ECU41              -----------G--G---A-------A--A-----------------A-----------------C--------G-----T-----G-----------T-TG---T------T------------C--G-----A-----G-----------T--T--------A----- 517
02_AG.GH.03.GHNJ196           -----------G--G---A-------A----G---------------A---------------------------------G----------------T-TG----------T---------------G-----A-----G-----------T----------------- 1447
02_AG.NG.01.PL0710            --C-----------G---A-------A--------------------A------------A-------------------T-----G-------------TG----------T------------C--G-----A-----G-----------T----------------- 605
02_AG.NG.x.IBNG               -----------G------A-------A----G---------------A--------------------------G-----------G-----------T-TG----------T------------C--G-----A-----G-----------T----------------- 946
02_AG.SE.94.SE7812            --------------G-----------A--------------------A--------------------------G-----------------------T-TG----------T------------C--G-----A-----G--------G--T----------------- 793
02_AG.SN.98.MP1211            -----------G--G---A-------A-----C--------------A--------------------------G-----------------------T-TG----------T------------C--G-----T-----G-----------T----------------- 618
03_AB.RU.97.KAL153_2          -----------G------A-------A--------------------A--------------------G-----G-------------------------TG-----G----T------------C--G-----T-----------G-----T----------------- 647
04_cpx.CY.94.CY032            ------------------A-------A--------------------A-----------C--------------G--------------------G----TG--------T-T------------C--G-----A-----------------T----------------- 781
05_DF.BE.x.VI1310             --C---------------A-------A--------------------A-----------------------G--G--------------------------------------------------C-----------------------------T-------------- 794
06_cpx.AU.96.BFP90            -----------G------A-------T--------------------A-----------CA---------------------------------------TG-----G----T------------C--G-----T-----------------T----------------- 1440
08_BC.CN.97.97CNGX_6F         --------------------G--------------T-----------A-----------CA--------------------------------------------T----T-----A-----------------------------------T-----------A--G-- 613
09_cpx.GH.96.96GH2911         ------------------A-------A--------------------A-----------CA-------------G------------------C----T-TG----------T---T--A--------G-----T---------C-------T----------------- 626
10_CD.TZ.96.96TZ_BF061        -----G-----------------G-----------T-----------A-----------------------------------------------G-------------------------------------------------------------------------- 794
11_cpx.GR.x.GR17              -----------G------A-------A-----------T--------A--------------------------------------------------T-TG-----G----T------------C--G-----T-----------------T----------------- 726
12_BF.AR.99.ARMA159           -----C--------G--G-----------------T-----------A-----------------------------G-----T---------------------T----T-----A--A--------------T-----------------C----------------- 1423
13_cpx.CM.96.1849             --------------------------A-----C-----T--------------------------------------G--T-----------------------------------------G-----------T---------C----G--T--T-------------- 793
14_BG.DE.01.9196_01           --------------------------A-----C-----T--------A--------------------------------------G-----------T-----------------A-----G-----------T---------C----G--T--T--------A----- 938
15_01B.TH.99.99TH_MU2079      -----------G--------------A--A-G---T-A---------A-----------C--------------G-----------------------T-TG--------T-T------------C--G-----A-----------------A----------------- 823
16_A2D.KR.97.97KR004          -----------C--------------A--------T-----------A--------------------------------------------C-----T-----------------------------G--T--T-----------------T----------------- 776
17_BF.PE.02.PE02_PCR0155      --------------G--GA------------G---T-A---------A------------A-------G--------G-----T---------------------T----T-----------------------------------------T----------------- 608
18_cpx.CU.99.CU76             ----------------C-A-------A--------T-----------A-----------C--------------G-----T-----------------T-TG-----G----T---------------G-----T-----------------T----------------- 732
19_cpx.CU.99.CU7              --C--G------------A------------------A---------A-----------------G--G------------------------------------T----T----------------------------------------------------------- 641
20_BG.CU.99.Cu103             --------------G-----G-----A-----C-----T--------A--------------------------------------------------T--------------C--------G-----------T---------C----G--T--T-------------- 890
21_A2D.KE.91.KNH1254          -----------C--------------A----------A---------A--------------------------------------------------------------------------T---------------------C------------------------- 623
22_01A1.CM.01.01CM_0001BBY    -----------G-------------------GA--T-A---------A-----------C--------------G---A--------AC---------T-TG----------T-A-A--------C--G-----G-----------------C----------------- 620
23_BG.CU.03.CB118             --------------------------A-----C-----T--------A--------------------------------------------------T-----------T-----------------------T---------C----G--T--T-------------- 864
24_BG.CU.03.CB378             --------------------------A-----C-----T--------A--------------------------------------------------T--------------C--------------------A---------C----G--T--T-------------- 864
24_BG.CU.03.CB471             --------------------------A-----C--T--T--------A--------------------------------------------------T--------------C--------------------A---------C----G--T--T-------------- 864
25_cpx.CM.01.101BA            --------------------------A-----------T--------A--------------------------------------------------T-----------------------G-----------T---------C----G--T--T-------------- 623
27_cpx.CD.97.97CDKTB49        --------------------------A-----C-----T--------A------------------C-------G-----------------------------------------------G-----------T---------C----G--T--T--T----------- 626
28_BF.BR.99.BREPM12609        --------------G--GA----------------T--------G--A------------------C-------------------------------------------------------------------------------------C----------------- 784
29_BF.BR.02.BREPM119          -----------------------G-----------T--T--------A------------------C-------------------------------T------T----TT----------------------------------------T----------------- 592
31_BC.BR.02.110PA             -----G--------G---A----G--------------------G--A------------A-------------G------------------------------T----T--------------C--------------------------T--------------G-- 864
32_06A1.EE.01.EE0369          -----------G------A-------A--------------------A-----------------G--G-----G-------------------------TG-----G----T------------C--G-----T-----------G-----T----------------- 1056
33_01B.MY.05.05MYKL007_1      -----------G--------------A----G---T-A------G--A--------------------G-----G-----------------------T-TG---T----T-T------------C--G-----A-----------------A----------------- 800
34_01B.TH.99.OUR2478P         --C--------G--------------A----G---T-A---------A-----------C--------------------------------------T-TG--------T-T---------------------------------------A----------------- 620
35_AD.AF.05.05AF095           -----------G------A-------A--------------------A-----T-----C--------------------------------------T-TG-----G----T------------C--G-----T-----------------T--T-------------- 617
36_cpx.CM.00.00CMNYU830       -----------G--------------A----G---T-A---------A-----------C--------------G-----T-----------------T-TG----------T------------C--G-----A-----------------T----------------- 620
37_cpx.CM.00.00CMNYU926       ------------------A-------A--------------------A-----------CA-------G-----T--------------C--------TG-G--------TGT------A--------G-----T--------------G--T----------------- 614
38_BF1.UY.05.UY05_4752        --------------G--GA----------------T-----------A--------------------------G------------------------------T----T-----------------------------G-----------T-----------A----- 829
39_BF.BR.03.03BRRJ103         ------------------A----------------------------A--------------------------------------------------T-----------------------------------T-----------------C----------------- 889
40_BF.BR.04.04BRRJ115         ------------------A----------------------G-----A--------------------------------T------------------------T----T-----A-----------------T-----------------C----------------- 961
42_BF.LU.03.luBF_05_03        --------------G--G-----------------T-----------A---------------------G----------------------------------------Y--------------------------------------G--C----------------- 944
43_02G.SA.03.J11223           ------------------A----G--A-----C-----T--------A------------A-------------G-----------------Y-----T-----------------------G-----------T---------C----G--T--T--------A----- 940
U.CD.83.83CD003               ------------------A-------A-----C--------------A---------------------------------------------------------T----------------------------------------------T--T--A----------- 913
U.CD.90.90CD121E12            ------------------A-------A--------------------A---------------------------------------------------------T----------------------------------------------T--T--A----------- 617
U.GR.99.GR303                 -----------G------A-------A--------T-----------A-----------CA-------------D-----------------------T-TG-----G--T-T-A-T--------C--G-----T-----------------T-----------A----- 704
U.NL.95.U_NL_95_H10986_D1     -----G-----G--------------A-----------T--------A-----------------------------------T-----G--C-----------------------------------------------------------T--T-------------- 1303
CPZ.CD.90.ANT                 -----C--------GTGT-----------AAA-----AT--------C--C--T--------T--T-----------G--A--T--T--T-----------------T--TG----T----AC-------G---------GG--C-------AGTA-------------- 827
CPZ.CM.01.SIVcpzCAM13         --------------------G--------------T-AT--T--G--A-----A-----------TC-G-----------TCTA--TTTG---G-T-----A-----G--TG-T--A-----------G-G----------G-TC---G---GT----------A----- 992
CPZ.CM.05.SIVcpzEK505         -----T-----------G--------A-----C--T-AT--------A-----A------ATG--C--G--------G--A--G--C-----C------------T----T-----A-----------G-----A-----G---C-G--G--AGTT-------------- 963
CPZ.CM.05.SIVcpzLB7           -----C-----C--G-----------A-----A--T-----------A-----------CG----C---G----------------T-----------T-----------T-----A--------C--G-----------G--C--------GT----T----------- 961
CPZ.CM.05.SIVcpzMB66          --------------------------A-----A--T-----------A--T--------CATG--TC-G-----------------T------C-------------------------------C--G-----A---------C-------AGTA-------------- 958
CPZ.CM.05.SIVcpzMT145         -----C----------C---------A--A-------AT-----G--C-----A-----CATG--------------G--A--G--T-----CC-C--T--T--------T---A-A--A--------G--------------TC-------AGTA--------T--G-- 954
CPZ.GA.88.GAB1                -----------------G-----G--A-----A--T-----T--G--A-----A-----------TC-------G------TTA--T--G---G-T---------T----TG----A-------------G----------G--C----G--AGT---T----------- 1476
CPZ.TZ.00.TAN1                -----C-------GG---A-------A--A-GG--T-AT--------A-----A-----C--------G--T--G-----A-----TT-T---C----T-GT-----C--TG-T--T----A------------A-----------G--G--AGT--------------- 1047
CPZ.US.85.CPZUS               -----C----------CG--------A-----C--T-A---G-----C--C--G-----CATG--C-----------G--G--T--C------G-C-----T-----C--TG-TA-A-----------G-G---A-----G-T------G--AG---------------- 1463
CPZ.CM.x.SIVcpzDP943          -----T-----T--G-----G-----A---AA---T-AT--------------A------ATG--CC-G-----G-----G--T--C--G---G-C--T--T-----T--TG-TA-A--A--G--C--G-G---T-----GG-------G--AGT--------------- 974
CPZ.CM.x.SIVcpzMB897          --------------------------A-----A--T-----------C------------ATG--TC-------------------T--G--------------------T---A-A-----------G-----A--------C--------AGTG-------------- 960
N.CM.02.DJO0131               -----T-----------GA----G--------C--T-AT--------A-----A------ATG--C--G--------G--A--A--CGC----C----T--T---T----T-----A--A--------G-----T-----G---C----G--AGT--------------- 900
N.CM.04.04CM_1015_04          -----T------------A----G--------C--T-AT--------A-----A------ATG--C--G--------G--A--A--CTC----C----TT-T---T----T-----A--A--------G-----T---------C----G--AGT--------------- 893
N.CM.04.04CM_1131_03          -----T-----------GA----G--------C--T-AT--------A-----A------ATG--C--G--------G--A--A--CTC----C----TT-T---T----T-----A--A--------G-----T---------C----G---GT--------------- 892
N.CM.95.YBF30                 -----T-----G-----GA----G--------C--T--T--------A-----A------ATG--C--G--------G--A--G--CTC----C----T--T---T----T-----A-----------G-----A-----G---C-G--G--AGT--------------- 979
N.CM.97.YBF106                -----T------------A----G--A-----C--T-AT------A-A-----A------ATG--C--G--------G--A--G--CTC----C----T-GT---T----T-----A--A--------G-----T-----G---C----G--AGT--------------- 980
O.BE.87.ANT70                 --------------GC----------A--A--C--T-A---T---A-C--T--T-----CATG-----G-----G-----A-TTT-CT-T---A-T--T--T--------TG-CA-A--A--------G-----TT-A---G--C----G--AGTA-------------- 1460
O.CM.91.MVP5180               --------------GC----------------C--T-A---T---A-T--T--T------ATG--------------G--TGT---CT-T---A-C--T--------G--TG-CA-A--------C----GG--TT-A---G----G--G--AGTA-------------- 1435
O.CM.96.96CMABB637            --------------GC----------A-----C--T-A---T---A-T--T--T------ATG-----G-----------T-TT--CT-T--CA-C--T--T--------TG-CA-A--A-----C--G-GG--TT-A---G----G--G--AGT--------------- 884
O.CM.98.98CMA104              --------------GC----------A-----C--T-A---T---A-T--T--T-----CATG-----------R-----T---T-CT-T---A-Y--T--T---T----TG-CA-A--A----------GG---T-A---G-AC----G--AGTA--------A----- 892
O.CM.99.99CMU4122             --------------GC----------A--A--C--T-A---T---A-C--T--T-----CATG-----G-----G-----T-TT--CT-T---A-T--T------T----TG-CA-A--A--C-------GG---T-A---G--C-------AGTA-------------- 891
O.FR.92.VAU                   --------------GC--------A-A--A--C--T-A---T--GA-C--T--T------ATG-----G-----G-----T--T--CT-T---G-C--T--T--------TG-CA-A--------C----GG--TT-A---G-A--GG-G--AGTA--------T--G-- 969
O.SN.99.SEMP1299              --------------GC----------A-----C--T-A---T---A-C--T--T-----CATG-----G-----G-------TT--CT-T---ACT--T-----------TG-CA-A--A----------GG--TT-A---G--C-------AGTA-------------- 1462
O.US.99.99USTWLA              --------------GC----------A-----C--T-A---T---A-T--C--T------ATG--C--------------T-TTT-CT-T---A-C--TGTT--------TG-CA-A--A-----C----GG--TT-A---G-A--------AGTA-------------- 880
```

```
B.FR.83.HXB2              TGCAGAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAACTACTAGTACCCTTCAGGAACAAATAGGATGGATGACA.........AATAATCCACCTATCCCAGTAGGAGAAATTTATA 1576
Gag                       __A__E__W__D__R__V__H__P__V__H__A__G__P__I__A__P__G__Q__M__R__E__P__R__G__S__D__I__A__G__T__T__S__T__L__Q__E__Q__I__G__W__M__T__.__.__.__N__N__P__P__I__P__V__G__E__I__Y__
A1.GE.99.99GEMZ011        ---T--G--------GT-A-------CA--G---------T--C-------G-----------------------------------------------------A---------------------.........-GC--C--------------G-----C--C----  766
A1.KE.00.KER2008          ------------C---T-A-------CA--G------------C-------------A-----C--------------------------------C----C---A---------AC----------.........--C--C--------------G-----C--C----  781
A1.KE.00.KNH1144          ------------C--GACA--------A---------------C----------------------------------------------------C----C---A---------------------.........-GC--C--------------G-----C--C----  787
A1.KE.00.KSM4024          ------------C--G--A--------A--------------CC-------------------G------------------------------------A----A---------------------.........-GC--C------G-------G-----C--C----  778
A1.KE.00.MSA4069          ------------C---T-A--C----CA--G------------C-------------A--G--------------C--------------------C----C---A---------AT----------.........GCA-----------------G-----C--C----  781
A1.KE.00.NKU3005          ------------C--GC-A------A-A-----------C--------------------------------------T----------------------C---A----------C----------.........GGC--C--------------G-----C--C----  781
A1.RU.00.RU00051          ---T-----------GT-A-----C-CA--G---------T--C-------G---C----------------------T-----------A--------------A---------------------.........-GC-----------------G-----C-------  880
A1.RW.93.93RW_024         ------------C--GT-A--------A---------------C-------T-------------------------------------------------C---A---------------------.........GGC-----------------G-----C--C----  784
A1.SE.95.SE8891           ------------C--GT-A--------A---------------C----------------------C----------------------------------C---A---------------------.........-GC-----------------G-----C--C----  781
A1.SE.95.UGSE8131         ------------C--GACA------A-A------------G-------------------------------------T--------------------------A----------C----------.........TCC--C------G-------G-----C--C----  972
A1.TZ.01.A173             ------------C--GT-A--------A--G------------C-------------A-----------A--------------------------C----C---A------T--CA----------.........GG---C--------------G---A-C-------  781
A1.UA.01.01UADN139        ---T-----------GT-A-------CA--G---------T--C-------G-----------------------------------------------------A---------------------.........-GC--C--------------G-----C--C----  781
A1.UG.92.92UG037          ------------C--GC-A--------A------------G-------------------------------------T----------------------C---A----------C----------.........GGC--C--------------G-----C--C---- 1595
A1.UG.99.99UGA07072       ------------C--G--A--------A---------------C------------------------------------------------------ATA----A-----------G---------.........-GC--C--------------G-----C--C----  781
A2.CD.97.97CDKTB48        ------------C--GT-A--------A--G------------C----------------------------------------G---G---------A---------------------------C.........-GC--C--------T-----G--------C----  941
A2.CY.94.94CY017_41       ------------C--G--A--------A---------------C-------------------------------------------------------------------------T--------C.........-GCG-------C--------G-------------  933
A.SN.01.DDI579            ------------C--GA----------A---------------C-------------------------------------------------G-----------A---------------------.........-GC---------G-------G--------C----  783
A.SN.01.DDJ369            ------------C-----A--------A---------------C------------------------------------------------------A------A---------------------.........-G----------G----G--G-----C--G----  786
A.SN.96.DDJ360            ------------C--GA-A--------A---------------C-------------------------------------------------G-----------A---------------------.........-GC---------G----G--G--------C---- 783
A.CD.02.02CD_KTB035       ------------C--GT-A--------T------------C--C------------C---------------------T------------------------------------------------.........-GC-----------------G--------C---- 1573
A.CD.97.97CD_KCC2         ------------C--GA----------A--------------AC----------------G-----------------T------------------------------------------------.........-GC-----------------G------------- 1578
A.CD.97.97CD_KTB13        ------------C--GT-A--------A--------------AC----------A-----G--------A--------T-----------------------G------------------------.........-GC-----------------G--------C----  786
B.AR.04.04AR151516        ----------------T-A------A-A------------G----G-----------------------------------------------------------------------------------.........--------------------------G--C----  783
B.AU.87.MBC925            -------------------------------------------------------------------------------------------------------------------------------.........-----------------------------C---- 1581
B.BO.99.BOL0122           ---C------------T----C----CA--G--------------------------A--------G-----------------------------C------------------AC----------.........-----------------------------C----  793
B.BR.03.BREPM2012         ----------------T--------------------------------------------------------------------------------------------------------------.........----------------------------------  986
B.CA.97.CANB3FULL         ----------------T-----------------------G--------------------------------------------------------------------------------------.........-----------------------------C----  856
B.CN.05.05CNHB_hp3        ------------------A--C--------G--------AG--------------C-------T-----------------------------------T--------G------------------.........GC----------G----------------C---- 1575
B.CO.01.PCM001            ----------------T--------------------------------------------------------------------------------------------------------------.........C-------------------------G--C----  781
B.GB.83.CAM1              ----------------T-------------------------------------------------------------------------------------------G------------------.........-----------------------------C---- 1577
B.GB.86.GB8               ----------------------------------------G--------------------------------------------------------------------------------------.........--------------------------G--C---- 1581
B.GE.03.03GEMZ010         ------------------A-----------------------A--T-----------A------------------------------------------------------G--AC----------.........-G------------------------G--C----  781
B.IT.05.SG1               ----------------T----------------------CG---------------------------------------------------------------------------C----------.........-----------C-----------------C---- 1378
B.JP.05.DR6538            ----------------T--------------------------------------------------------------------------------------------------------------.........--C--------------------------C---- 1580
B.KR.05.05CSR3            ----------------T-------------------------------------------------------------T------------------------------------------------.........---------------------------------- 1601
B.NL.00.671_00T36         ----------------T----------------C---------------------------------------------------------------------------------------------.........-----------------------------A---- 1111
B.RU.04.04RU128005        ------------C---T----------------------------------A-----T-----G-----R----------------------------------------------C----------.........-----------------------------C---- 1067
B.TH.00.00TH_C3198        G---------------T----------------------A--------G--A-----------------------------------------------T--------G------------------.........-G-------G-------------------C----  787
B.UA.01.01UAKV167         ----------------T---------------------------------------------------------------------------------------------------C----------.........-----------------------------C----  781
B.US.04.ES10_53           ----------------T---------CA--G---------G-G--------------A--G------------------------------------------------------AC----------.........-----C--------------------C--C---- 1588
B.US.99.PRB959_03         ------G---------T---------C---G--------------------------A---------------------------------------------------------AC----------.........-----------------------------C----  796
C.AR.01.ARG4006           ------G-----C---T-A--------R------------G-C-----------A--------------------------------------C----A-----------------C----------.........-G---C--------T-----------C--C----  775
C.BR.04.04BR013           ---------------GT-A-----------------------C--------T--A-------------------------------------------------------------C----------.........GG---C-A------T--------------C---- 1031
C.BW.00.00BW07621         ---------------GT-A--------A--------------------------A-------------------------------------------------------------C----------.........-G---C------G----------------C----  944
C.CN.98.YNRL9840          ----------------T-A--------A--------------------------A--------------A----------------------------G-----------------C---------G.........GG---C---T--G-T-----G-----C--C----  778
C.ET.02.02ET_288          ---------------GA-A------------------------CA---------A-------------------------------------------------------------C----------.........-G------------T-----G-----C--C----  775
C.GE.03.03GEMZ033         ---------------GC-A---------------------G-------------A-------------------------------------------------------------C----------.........-G---C--------T--------G--C--C----  772
C.IL.99.99ET7             ---------------GC-A---------------------G-A--G--C-----A----------------------------------------------C----------C---C----------.........-G---C------G-T-----G-----C-------  772
C.IN.99.01IN565_10        ----------------T-A--C-------C------------------------A--------------------------------------------------------GG---C----------.........GG---C------G-T--------------C----  975
C.KE.00.KER2010           ----------------T-A--------------G-------A-C----------A-------------------------------------------------------------C----------.........-G------------T--------------C----  772
C.MM.99.mIDU101_3         ----------------T-A--------A------------G-------------A-------------------------------------------G-----------------C---------G.........GG---C------G-T-----G-----C--C----  945
C.MW.93.93MW_965          ---------------GT-A--------A--------------------------A--------------------------------------------------------G----C----------.........-----C--------A-----------C--C----  775
C.SN.90.90SE_364          ---------------GT-A---------------------G-------------A-----------C-------------------------------A-------------G---C----------.........GG---C--------T-----G-----C--C----  772
C.SO.89.89SM_145          ----------------T-A--------A--G--G------G-------------A-----------------------T------------------------------------------------.........-G---C--------T-----G-----C--C----  772
C.TZ.02.CO178             ----------------T-A--------A--------------------------A-----------------------------------------------C--------G----C----------.........-G---C--------T-----G-----C--C----  772
C.UY.01.TRA3011           ------------C---T-A-----------------------A--------A--A--------------------------------------C--------------------G-C----------.........-----C------G----------------C----  763
C.YE.02.02YE511           ----------------T-A--------A--C---------G-------T-----A--------G----------------------------------G-----------------C----------.........GG---C-----CG-T-----G-----C-------  772
C.ZA.04.04ZASK164B1       ------------------A-----------------------------------A------------------------------------------------------------------------.........--C--C--G-----------G-----C--C---- 1005
C.ZA.05.05ZAPSK240B1      ---------------GT-A--------------------------------T--A-------------------------------------------------------------C----------.........-----C------G-T-----G-----C--C----  998
C.ZM.02.02ZM108           ---------------GT-A------CCA--G---------G-------------A--A-----------------C---------------------------------------AC----------.........-G---C--------T-----------C--C---- 1572
D.CD.83.ELI               ---------------GT-A-----------------------------------------------------------T-------------------------------------C----------.........-G---C-----------------------C---- 1122
D.CM.01.01CM_0009BBY      ------------C--GC-A--------------------CG-------------------------------------T--T--------------------------------G-C------A---.........GGC---------G----G-----G-----C----  769
D.KE.01.NKU3006           ---------------GC-A------A-A-------------A------------A-----------------------T-------------------------------------C----------.........-GC-------------------------------  781
D.KR.04.04KBH8            -T-------------GC----------------------C-----------T--AC----------------------T-----------------------------C------------------.........-GC--C-----------------------C---- 1535
D.TD.99.MN011             ------------C--GC-A---------------------G----------------------G--------G-----T--T--------C------------------------------------.........-GC-----------T--------------C----  794
D.TZ.01.A280              ---------------GC-A--------------------------------A--------------------------T-------------------------------------C----------.........------------G----------------A----  771
D.UG.99.99UGK09259        ---------------GC-A-----------------------------------A-----------------------T--------------------------------------AC---------.........GGC--------------------------C----  781
D.YE.01.01YE386           ---------------GC-A--------------------CG-------------A--------G--------------T-------------------------------------C----------.........--C------G-------------------C----  778
D.YE.02.02YE516           --G---------C--GC-A-------CA--A--------------------------A--------------------T--T-T------C------------G-----------A-----------.........-GC--------------------------C----  781
D.ZA.90.R1                ---------------GC-A-----------------------------------------------------------T-------------------------------------C----------.........--C--C--------------------T--C----  922
F1.AR.02.ARE933           ------------C---T-A-------C---------------CC----C-----------G-----T-----------T-----T--------------T-C-------------CA----------.........-GC--C------G-------G--------C----  873
F1.BE.93.VI850            ------------C---T-A---------------------GCCC----------------G-----T-----G-----T-----T------------------------------CA---------G.........GGC--C------G-------G-----C--C----  919
F1.BR.01.01BR125          ------------C---T-A-----------------A-----CC-------T--------G--G--T-----------T-----T-------------AT-----A---------CA----------.........-GC--C------G-------G--G-----C---- 1024
F1.ES.x.P1146             ------------C---C-A-----------------------CC----------------G-----T-----------------T------------------------------CA----------.........-GC--C------G-------G--------C----  862
F1.FI.93.FIN9363          ------------C---T-A-----------------A------C----------------G-----T--------C--T-----T------------------------------CA----------.........-G---C------G-------G-----C--C----  911
F2.CM.02.02CM_0016BBY     ------------C--GT-A-----------G-----A-----CC-------T-----A--------T-----------T-------------------------------------C----------.........-GC--C------G---------------------  769
F2.CM.95.MP257            ------------C--GT-A-----------------A-----CC-------T--------------------------T-----------C-------------------------C----------.........-GC--C------G----------------C----  788
F2.CM.97.CM53657          ------------C---T-A-----------G-----A--C--CC-------T-----A--------T-----------T-----------------C-A---A-------------C----------.........-GC--C------G----------------C----  769
G.BE.96.DRCBL             ------------C--GC-A------CA---G--------------------------A--G--C----C---------T---------G-------------G--------G---A----------C.........-GC--C------G-------G-------------  1540
G.CM.01.01CM_4049HAN      ------------C--GT-A------CC---G------------C-------------A--------------------T-----------------------G------------A----------C.........-GC--C--------------G-----T--C----  781
G.CU.99.Cu74              ------G-----C--GC-A--C---CC---G-----------CC-------------A--------------------T-----------------------A-----------------------T.........--C-----------------G-------------  1189
G.ES.99.X138              ------G-----C--GA--------CAR--K-----------CC-------------A-----------A--------T-----------------------G------------A----------C.........-GC-----------------G-------------  1026
G.GH.03.03GH175G          ------G-----C--GT-A------CA---G--------------------------A-----T----CA--------T--------------------A--G-----G------A----------C.........-G---C--------------G--------C---- 1616
G.KE.93.HH8793_12_1       ------G-----C---A-A------CCA--G------------C----------A--A--------------------T-----------C-------A---G------------A----------C.........-GC--C--------------G-------------  978
G.NG.01.01NGPL0669        ------G--------GC-A------CAA--G------------C-------------A--------C-----------T-----------------------G------------AC---------C.........--C-----------------G-------------  781
G.PT.x.PT2695             ------G-----C--GA-A------CAA--G-----------CC-------------A--------------------T-----------------------G------------A----------C.........-GC-----------------G------------- 1524
G.SE.93.SE6165            ------G-----C--GA--------CAA--G---------T--C-------------A--------------------T------------------T----G------------ACT--------C.........GGC--C--------------G-------------  979
H.BE.93.VI991             ---------------GC-A--------A---------------C----------------------------------T---------------------------------G---C----------.........GGC-----C--A--T-----G-----C--C----  968
H.BE.93.VI997             ---------------GC-A------------------------C-------------------------------C--T-----T-------------------------------C----------.........GGC------AG---------G-----C--C----  903
H.CF.90.056               ------------C--G--A------------------------C----------A--------------------C--T-----------------------G-------------C----------.........GGC------G----------G-----C--C----  923
J.CD.97.J_97DC_KTB147     ------------C--GC-C--------A--------A---G----------T--------------G-----------T-------------------A-----------------C------A---GGCAACGGTGGC--C-A---------------------C----  780
J.SE.93.SE7887            ------------C--G--A--------A---------------------------G----------------------T-----------------------C------------------------.........GGC-----------------------G-------  899
K.CD.97.EQTB11C           ------------C--GA----C--------A-----------CC----------A--A--------T-----G-----T-----------------C--T---------------AC----------.........-GC--C--------------G--------C----  775
K.CM.96.MP535             ------------C--GT-A--C--------------------CC----------A-----------G-----G--------------------C--C-------------------C----------.........-GC--C------G-------G--G-----C----  775
```

```
B.FR.83.HXB2                 TGCAGAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAACTACTAGTACCCTTCAGGAACAAATAGGATGGATGACA.........AATAATCCACCTATCCCAGTAGGAGAAATTTATA 1576
Gag                          __A__E__W__D__R__V__H__P__V__H__A__G__P__I__A__P__G__Q__M__R__E__P__R__G__S__D__I__A__G__T__T__S__T__L__Q__E__Q__I__G__W__M__T__.__.__.__N__N__P__P__I__P__V__G__E__I__Y__
01_AE.CF.90.90CF11697        ---------------GC-A--C-----A-------------C----------------------------------------S------------C-A-----A-------------------------.........-GC--C-------------G-----C--C---- 1531
01_AE.CN.05.FJ051            -----------------A--C-----A-------------C-------A--------G---------------------------------------------A-------------------------.........--C---------------G-----C--C---- 1592
01_AE.CN.06.FJ054            ---------------G--A--C---A-A--G---------C------------A--G---------------------------------------------A----------C---G-----.........--C--------G---T-------T--C---- 1598
01_AE.HK.x.HK001             ---------------G--A--C-----A-------------C----------A----G---------------------------C----------------A-------------------------.........-GC------Y--------G-----C--C----  953
01_AE.TH.01.01TH_R2184       ---------------G--A--C-----A-------------C---------------G---------------------------------------------A-------------------------.........--C---------------G-----T--C----  781
01_AE.TH.02.OUR769I          ---------------G--A--C-----A--G---------C------------T--G--G---------------------------------A-----A--------A-----------.........--C---------------G--G--C--C----  781
01_AE.TH.90.CM240            ---------------G--A--C-----A-------------C---------------G---------------------------------------------A-------------------------.........--C-------C--------G-----C--C---- 1141
01_AE.US.00.00US_MSC1164     ---------------G--A--C-----A--G---------C-----------C-A--G---------------------------------A-----A-------------------------.........--C---------------G-----C--C----  781
02_AG.CM.02.02CM_4082STN     ------------C--G--A--------A------------CC----------------------------------------------------------A-------------------------.........-GC---------G-------G--------C----  772
02_AG.EC.x.ECU41             ------------C--G-CA--------A--G---------C---------------G---------------------------------------------A-------------------------.........-GC-----------------G--------C----  678
02_AG.GH.03.GHNJ196          ------------C-----A--------A-------------C---------------G---------------------------------------------A----------G-------------.........-GC-----------------G------------- 1608
02_AG.NG.01.PL0710           ------------C--G--A--------A-------A--GC-C---------------G----T-----------------------------------------A-------------------------.........-GC---------G-------G--------C----  766
02_AG.NG.x.IBNG              ---------------G--A--------A-------------C---------------G---------------------------------------------A-------------------------.........-GC-----------------G--------C---- 1107
02_AG.SE.94.SE7812           C-----------C--G--A--------A-------------C---------------G---------------------------------------------A--G----------------------.........-GC-----------------G--------C----  954
02_AG.SN.98.MP1211           ------C-----C--GACA--------A-------------C---------------G---------------------------------------------A-------------------------.........-GC--------------------C--C----  779
03_AB.RU.97.KAL153_2         ---T-----------GT-A-------CA--G---------T--C-------G---------------------------------G--G-----------A-------------------------.........-GC--C-------------G-----C--C----  808
04_cpx.CY.94.CY032           ------C-----C--GACA--------A-------------C-------------------------T-----------------------------------A-------------------------.........-GC--C------G-------G--------C----  942
05_DF.BE.x.VI1310            ---------------GT-A----------G---------G-G----------A-------T-----------T-------------C--C--------------------C-------------.........--C--C--------T-------------C----  955
06_cpx.AU.96.BFP90           ------------C--G--A--------A-------A-----C----------A-----------------T--------------------------------------------------------.........-GC--C-------------G--------C---- 1601
08_BC.CN.97.97CNGX_6F        ----------------T-------------------AG-G--------------------------------------------------T------G-----------------------------.........-------------------------------C----  774
09_cpx.GH.96.96GH2911        ---------------G--A--------A----------A-C---------A---------------------------------------------G---A--------C-----------.........GG---C-----C--------G--------C----  787
10_CD.TZ.96.96TZ_BF061       ---------------GC-A--------A--G---------G------------A--A---------------T----------------------------------------A-----------.........-GC----------------------C----  955
11_cpx.GR.x.GR17             ------------C--G--A--------A------------GCCC----------------G-------------------------A---G---A-----A--G---------------------.........GGC--C-----------------------C----  887
12_BF.AR.99.ARMA159          ------C-----C---T-A----------------A-----CC----------------G-----T-----------T-----T---------------------------CA-------------.........-GC--C------G-------G--------C---- 1584
13_cpx.CM.96.1849            ---------------GA--------CAA-GA---------T--C---------A--A-------------------T-----G-----C-----------G-----------AC---------C.........-----C-------------G-------------  954
14_BG.DE.01.9196_01          ------G-----C--GA--------CAA--G---------CC-------------A-------------------T--------T---------A---G-----------A----------T.........-GC----------------G------------- 1099
15_01B.TH.99.99TH_MU2079     ----A-------------A--C-----A-------------C---------------G---------------------------------------------A-----G-----------------.........-GC------G---------G-----C--C----  984
16_A2D.KR.97.97KR004         G-----------C--GC----------A-------------C---------------G---------------T---------------------------------------------------C.........-GC--C------G-------G-----G--C----  937
17_BF.PE.02.PE02_PCR0155     ------------C---T-A----------A-----A-----CC-----------------------T-----------T-----G-----------A---------------CA----------T.........-GC--C------G-------G--------C----  769
18_cpx.CU.99.CU76            ------------C---A-A--------A--G--------------------A-----T-------------T--------------------------------A--------A-G---------.........--C--C-------------G-----T--C----  893
19_cpx.CU.99.CU7             ---------------GT-A------------------------------------------------------T-------------------------------------------CC----------.........TCC------------------G--A----  802
20_BG.CU.99.Cu103            ------G-----C--GACA--C---CA---G---------CC-------T----A----G----------C--T--------G-------------A-----------AC---------C.........--C---------------G------------- 1051
21_A2D.KE.91.KNH1254         ---------------GT-A-----A-A------------------------A-----------------------------------------------------------C-------------.........GGC--C-------------G-----C----  784
22_01A1.CM.01.01CM_0001BBY---------------GA---C-----A--G--------TA-C--------------------G---------------------------------------A---------C-------------.........-GC--C--T---G-------G--------C----  781
23_BG.CU.03.CB118            ------G-----C--GC-A--C---CA---A----------C-------T--A--A----G----------C--T----------C---------A-------G--A----------C.........--C---------------G------------- 1025
24_BG.CU.03.CB378            ------G-----C--G--A--C---CC---G---------CC-------T--A--A----G----------C--T---------------------A-------G--A----------C.........-GC---------------G------------- 1025
24_BG.CU.03.CB471            ------G-----C--GT-A--C---CC---G---------CC-------T--A--A----G--------------T--------------------A-T-A-------G--A----------C.........-GC---------------G------------- 1025
25_cpx.CM.01.101BA           ------G-----C--GA-A--------A--G----------C------------A--------------------T-------------------A---G--A--------A----------T.........-GC--C-------------G-----------C-  784
27_cpx.CD.97.97CDKTB49       ------G--------GACA------CCA--A---------G--C-------G-----A-----C--------------T------------C-------G-----------C----------T.........------------------G-------------  787
28_BF.BR.99.BREPM12609       ------------C---T-A--------A-------A-----CC-------T-------G-----T-----------T-----T-----C-------A------------------C-------------.........-GC--C------G-------G--------C----  945
29_BF.BR.02.BREPM119         ------------C-----A---------------A------C---------------G-----T-----------T-----T-----------------------------CA-------------.........-GC--C------G-------G--------C----  753
31_BC.BR.02.110PA            -----------------A----------G-------------C----------A---------------------------------------C--------------------AC-------------.........GG---C------G---------G--G--C---- 1025
32_06A1.EE.01.EE0369         ---T-----------GT-AT-----ACA--G---------T--CA------G--------------------------T-------------------------------A-------------------------.........-GC--C-------------G-----C--C---- 1217
33_01B.MY.05.05MYKL007_1     ------T-------AGACA--C-----A-------------C---------------G---------------------------------A-----A-------------------------.........--C---------------G-----C--C----  961
34_01B.TH.99.OUR2478P        ------C--------G--A--C-----A--G---------C----------A-----G---------------------------------------------A--G-----------------------.........-GC---------------G-----C--C----  781
35_AD.AF.05.05AF095          ------------C--GT------G--A---------------------------------------------T----------C-----C---C--A----G---C---------------.........GG---C-------T-----G-----C--C--C-  778
36_cpx.CM.00.00CMNYU830      ---------------G--A--C-----A-----------CA-C-------------A-------------------------A-----------------A--------C-------------......AAC--CCCA--TAT-TCT--------------C----  784
37_cpx.CM.00.00CMNYU926      ------C-----C--G--A--------A-----------CA-C----------T---------------------------------------------------A--G--G--C-------------.........-GC--C-------------G-----C--C----  775
38_BF1.UY.05.UY05_4752       ------------C---A-A----------G---------CC-------------A--G-----T-----------T-----------------A----------------CA-------------.........-GC--C------G-------G--------C----  990
39_BF.BR.03.03BRRJ103        ----------------T---------A-----------------------------------------------------------------------------------------------------.........-------------------------------C---- 1050
40_BF.BR.04.04BRRJ115        ------------C---C-----------------------CC-------------A--G-----T-----------T--------T--------------------G---CA-----A---.........GGC--C------G-------G--------C---- 1122
42_BF.LU.03.luBF_05_03       ------------C---T-A---------------------CC-------G--------G-----T-----------T-----T-----------A----------G-----CA-------------.........-G---C------G-------G-----G--C---- 1105
43_02G.SA.03.J11223          ---------------GC-A------CCA--G----------C-------T-----A--------------------T---------------------G----------WAC---------C.........-GC--C-------------G------C------ 1101
U.CD.83.83CD003              ------T-----C--GT-A---------------------CC---------------G----------------T----------C---------------------------------------.........-GC--------C-------G-----C--C---- 1074
U.CD.90.90CD121E12           ------T-----C--GT-A---------------------CC---------------------------------T----------C---------------------------------------.........-GC---------------G-----C--C----  778
U.GR.99.GR303                ---------------GT-A------ACA--A---------T--C----------------------------------------------------------A-------------------------.........-GC--C-------------G-----......  859
U.NL.95.U_NL_95_H10986_D1    ------------C--G--A--C-----A------------CC---------------------T----G----T----------------C-----------------C---------------.........-GC--C------G-------G--------C---- 1464
CPZ.CD.90.ANT                A--T--G--------GT-A--C--CACT--------A--AG-ACAGG---A--AT-A--G-------CA--------T-------G--A--A--C--AG-G-----G--G--GCA-------T--ACACCTCAAC-G---GG-GGAG----------G--C--C----  997
CPZ.CM.01.SIVcpzCAM13        -------------C-GA-------T--T-----C--C--C--A--T----A---C-A---------C----T-----T-------G-----C--C--AT-G-----G-----T------GCA--.........GCA--C-----C-------------C--C---- 1153
CPZ.CM.05.SIVcpzEK505        A-----G-----C--G-GA--------A---ATG-----A--AC-------A--AG-A--G--G-----A-------T---------A---------GGC--G--G--C---------T.........GCC-----T---G-T-----------T------- 1124
CPZ.CM.05.SIVcpzLB7          ACA------------GC-------TA-C---ATG-----A--A-----------A-------------------------------C--------TG-A--A-----------------------T.........-GC--C--T-----A-----------------C- 1122
CPZ.CM.05.SIVcpzMB66         ------G-----C--GC-C--------T---------A---G--C------A--A---------------------------------C--C--C---A-C--A--G--GG---G---------T.........-GC--C--T-----T----------T--C---- 1119
CPZ.CM.05.SIVcpzMT145        ---T--G-----C--GACA--C--TACC-----------------A--T----A--A----G-------A--C---------------C--------CT---AG---G--G---G-----------.........GCA--C--T-----T----------T------- 1115
CPZ.GA.88.GAB1               C--C--G------C-CT-A-----CACT-----T--T--G--A--C----A---C-A--G--G--------T-----T-------G--C--C--C--A--G-----------T--G----CA--.........GCA-----T--C-----------G--TG------ 1637
CPZ.TZ.00.TAN1               A-----G-----C----CA-----C-CT--------A--CCAGCA-G---GAT-C-A-----G--C-CA--GGCA--T--T-----G--C--------G--A--A------G--CTG---------ACCCCACAGGCAC-AGG-GGAG-G------------C--C---- 1217
CPZ.US.85.CPZUS              A-----T--A-----ACA-----TC----------C-----A-------G----------------C-------------T--------C--C--C---T-G--A-----G-G--C-------------.........GCA--C-----C--T-----------C--A---- 1624
CPZ.CM.x.SIVcpzDP943         ACTG--------C--GACC-----T-----------T--A--A-------A--AC------GG--C-A--------------T-----C-----C--T--AGCA-----G--T-CT-------------.........GCA--------CG-T-------------C---- 1135
CPZ.CM.x.SIVcpzMB897         ------G-----C--GACA--C--TA-A-----------A---C-T-------A-----G---------A-------------------G--A-----C--A-----------G-G-----------C.........-G---C--T----C---G---------C-----C- 1121
N.CM.02.DJO0131              A-----------C--GAC-------R-C-C--YG--A--AT-AC-C-----A--AC----------T--A-------T-------G--A--------T--GMT------GG-GACC--------T.........TC------T---G-T---------------C---- 1061
N.CM.04.04CM_1015_04         A-----------C--GAC----------C-C--TG--A--AC-AC-C-----G--AC----------T--A--R-----T-------G--A--------T--G-CA----G---AC---------T.........TC------T-----T-------------T------- 1054
N.CM.04.04CM_1131_03         A-----------C--GAC----------C-C--TG--A--AC-AC-C-----G--AC------G--T--A--C-----T-------G--A--------T--G-CA----G---AC---------T.........TC---C--T-----T-------------T------- 1053
N.CM.95.YBF30                A-----C--------GACA---------C-C--TG--A--AC-AC-C-----G--AC------C--T--A---------T-----------A-----C-----GGCA-----GG-G-CT--------T.........GC------T---G-T-------------T------- 1140
N.CM.97.YBF106               G--G--G-----C--GAC---------CC-C--TG--A--AC-AC-C-----G--A--------C--T--A---------T-------G--A--------T--GGCA-----GG-G-C---------T.........TC------T-----T-------------T------- 1141
O.BE.87.ANT70                A-T---G---------ACT--C---CCA-CG-T------GT-GC--------G-----A--G-------CA----------T--T--G--A-----C----AG--A--G-----TCAC----CT--C......AGGCCC--C-A-------------------C--C---- 1624
O.CM.91.MVP5180              A---------------ACT------CCAGCAATG-----GT-AC--------G-----A--G-------CA----------T--T-----A-----C--A-AG--A--G-----TAT-----CT--T......AGAGGGGC-AACT----------------C--C---- 1599
O.CM.96.96CMABB637           A-----T---------ACC--C---CCA-CAAT------GT-AC--------G--A--A--G-------CA----------T--T-----A-----C-A--AG--A--G---G-CCAC-----T--C......AGGGC---CG-C------------T-----T--C---- 1048
O.CM.98.98CMA104             A---------------ACT--C---A-AGCG-T------AT-GC----G--A-----A--G-------CA----------T--T--G--A-----C-----G--A-----GG-GCAC----CT--C......AGAGCC--CAAC-----------------C--C---- 1056
O.CM.99.99CMU4122            AT--------------ACT--C---CCAGCGATG-----AT-AC--------G---------G-------CA----------T--T--G--A-----C----AG--A--G---G-TCAC----CT--C......AGGGCC--CAAC-----------------Y--C--C- 1055
O.FR.92.VAU                  A-----C--------A-ACT------CCAGTA-T---A--AT-AC--------G-----A--G-------CA----------T--T-----A--A--C-----G--A--G-----TCAC----CT--C......AGGGC---C-AG-----------T--G--T--C---- 1133
O.SN.99.SEMP1299             A---------------ACT--C---CCAGCG--------AT-GC-TGT---G-----A--G-------CA----------T--T-----A-----C----AG--A--G---G-TCAC-----T--C......AGGCCC--C-A-------------------C--C---- 1626
O.US.99.99USTWLA             A-----T---------ACT--C---CCA-CGAT---A--GT-AC--------G-----A--G-------CA----------T--T-----A--C--C--A-AG--A--G-----TCAC----CC-GT......AGGCCAGG-AAC-----------TR----C--C---- 1044
```

```
B.FR.83.HXB2               AAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGAATGTATAGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAACTCTAAGAGCCGAGCAAGCTTCACAGGAGGTAAAAAATTGGATGACAGAA 1746
Gag                        K__R__W__I__I__L__G__L__N__K__I__V__R__M__Y__S__P__T__S__I__L__D__I__R__Q__G__P__K__E__P__F__R__D__Y__V__D__R__F__Y__K__T__L__R__A__E__Q__A__S__Q__E__V__K__N__W__M__T__E_
A1.GE.99.99GEMZ011         ----------------A---------------------------------T-T------T-A--------------G-----A--------C--------------TA------T---------T-----T---------A-------T-----G--C-----------C  936
A1.KE.00.KER2008           ----------------A---------------------------------GTT-----------T----A------G-----A--------C--------------TA------T----GT---C--------A------A-------A------GG------------G  951
A1.KE.00.KNH1144           --------------T-----------------------------------GTT------T----T----A------G-----A--------C-----T--------TA------T---------C-----T---------A--------------GG-------------  957
A1.KE.00.KSM4024           --------------------------------------------------GTT------T----T----A------G-----A--------C-----T--------TA-A----T---------C-----T---------A--------------GG-------------  948
A1.KE.00.MSA4069           --------------------------------------------------GTT-----CT----T----A------G-----A--------C-----T--------TA------T---GGT---C-----T-----G---A--------------GG------------C  951
A1.KE.00.NKU3005           -------------------------------------------C------GTT------T----T----A------G-----A--------C-----T--------TA------T----G----C-----T-----G---A----------G---GG-------------  951
A1.RU.00.RU00051           ----------------A--------------------------C--T---GT---------A--------------------A--------C-----T--------TA------T---------C-----T---------A-------------G--C------------ 1050
A1.RW.93.93RW_024          --------------------------------------------------GTT------T----T----A------G-----A--------C--G--T--------TA-A--T-T------A--C-----T---------A--------------GG-------------  954
A1.SE.95.SE8891            ----------------A---------------------------------GTT--T---T---------A------G-----A--------C--------------TA------T---------C--------A------A--------------GG-------------  951
A1.SE.95.UGSE8131          ----------T-----A---------------------------------GTT------T----T----A------------A--------C--------------TA------T----G----C-----T---------A-------A------GG------------T 1142
A1.TZ.01.A173              ----G---------------------------------------------GTT--T---T----T----A------G-----A--------C--------------TA------T---GG----C-----T---------A-------T------GG-------------  951
A1.UA.01.01UADN139         ----------------A-----------G---------------------GTT------T-A--------------G-----A--------C--------------TA------T---------T-----T---------A-------T-----G--C------------  951
A1.UG.92.92UG037           --------------------------------------------------GTT------T-A--T----A------G-----A--------C--------------TA----T-T---------C-----T---------A--------------GG------------- 1765
A1.UG.99.99UGA07072        -------------------------------------------C------GTT-----CT----TG---A------G-----A--------------A---------A------T-----T---C-----T--A------A--------------GG------------C  951
A2.CD.97.97CDKTB48         --------------------------------------------------GT-------T---------A------G-----A-----------------------TA------T---------C-----T---------A----------------------------C 1111
A2.CY.94.94CY017_41        --------------------------------------------------GT-------T----------------G-----A--------------T-----G--TA------T---------------T--------CA----------------C--------G--C 1103
A.SN.01.DDI579             G------------------G--------G---------------------GTT------T----------------G-----A--------C--------------TA-A----T----TG---C--G--T---------A-------T-----G--C------------  953
A.SN.01.DDJ369             G---------------A-----------G---------------------GTT------T-------C--------G-----A-----------------------TA------T----TG---C-----T-----G---A-------A--------C------------  956
A.SN.96.DDJ360             G---------------------------G---------------------GTT------T----------------G-----A--------C--------------TA------T----TG---C--G--T---------A-------A--------C------------  953
A.CD.02.02CD_KTB035        --------------------------------------------------GT-----------------A------G-----A--------C-----T--------TA-A----T---------C-----T---------A-C--------------------------C 1743
A.CD.97.97CD_KCC2          ----------------------------G---------------------GT-------T---------A------G-----A-----T--------T--------TA-A----T---------C-----T--A------A-------T--------C-----------C 1748
A.CD.97.97CD_KTB13         --------------------------------------------------GT------------------------G-----A--------------T--------TA-A----T----GT---C-----T-----------------T--------------------C  956
B.AR.04.04AR151516         --------------------Y-------------------------T---GT------------------------------------------------------------------------------------------------A---------------------  953
B.AU.87.MBC925             -------------------------------------G--------T-----------------------------------------------------------------------------------------------------T--------------------- 1751
B.BO.99.BOL0122            --------------------C----------------G------------GT-------T-------------G-----------G--T--------------------A------------T----G--T--A------------------------------------  963
B.BR.03.BREPM2012          -------------T--------------------------------T-----------C-----------------G--------------------T--------------------------------------G--C--------T--------------------- 1156
B.CA.97.CANB3FULL          --------------A-------------------------------------------------T--------------------------------T------------------------------------------------------------------------ 1026
B.CN.05.05CNHB_hp3         --G--------------------------------------------------G-T---------------------------------------------------A----------G--------------A-----------A--T--------------------- 1745
B.CO.01.PCM001             --------------T-------------------G--------------C---------T-------------------------------C--------------------------------------A---------------------------------------  951
B.GB.83.CAM1               -------------T--------------------------------------------------------G------------------------------------A----------------------T--A------------------------------------ 1747
B.GB.86.GB8                --------------------------------------------------GT---------------------------------------------T--------------------------------------------------T--------------------- 1751
B.GE.03.03GEMZ010          --------------T-----C-------------------------------T-----------------------------------T---------------------------------T-------------------------T---------------------  951
B.IT.05.SG1                ----------------A--C-----------------G-----------NT-T----------------A----------------T-N---------------------------------A----------------------------------------------- 1548
B.JP.05.DR6538             -------------------------------------------------------T--------------------------A--------------T--------T--------------------------------------------------------------- 1750
B.KR.05.05CSR3             --------------A-------G--C-----------G--------------------------------------------A--------------T---------A-A--------G--------------------C--C--------------------------- 1771
B.NL.00.671_00T36          -G-------------------------------------------------------------------A------------------------------------T---------------T-------T--------C--------A--------------------- 1281
B.RU.04.04RU128005         -------------T--RR-------------------G------------GT-------------------------------------------R--------------------------------------------A-------A--------------------- 1237
B.TH.00.00TH_C3198         ----------------A--------------------------------------------------------------------------------T-----------------------------------------------A--T---------------------  957
B.UA.01.01UAKV167          --------------------------------------------------GTT--------A--------------------A---------------------------------------T----G--T--------------A--T---------------------  951
B.US.04.ES10_53            -G-----------------G-----------------G-----------C--------------------------------------------G-----------------------------------------------------A--------------------- 1758
B.US.99.PRB959_03          ----------------------------------------------------------------------------------A--------------T--------------------------------T---------------------------------------  966
C.AR.01.ARG4006            -------------T-----G------------------------------GT-------T---------A------G-----A--------------T--------------T-T-------T-------A--A------A-C--A-----------------------G  945
C.BR.04.04BR013            -------------T--A--G------------------------------GT-------T---------A------G-------------------------------------T-------T-------A--A------A-C--A--T-----G--------------T 1201
C.BW.00.00BW07621          -------------T-----G------------------------------GT-------T---------A------G-----------------------------T-------TC-----GT-------T--A------A----A--------G--------------C 1114
C.CN.98.YNRL9840           -------------T-----G--------------------------T---GTT------T----------------G-------------------------------------T-------T-------T--A------A----A--T--------------------C  948
C.ET.02.02ET_288           -------------T-----G------------------------------GT-------T----------------G-----A--G--T-------------------------T------CT-------T--A------A----A--T---------------------  945
C.GE.03.03GEMZ033          -------------TA----G------------------------------GT-------T---------A------G-------------------------------------T---------------T--A------A----A--T--------------------T  942
C.IL.99.99ET7              -------------T------------------------------------GT-------T---------A--------------------------------------------T-------T-------T--A------A----A--T--------------------C  942
C.IN.99.01IN565_10         -------------T-----G------------------------------GT-------T---------A------G-----A-----------------------T-------T--G----T-------T--A------A----A-----------------------C 1145
C.KE.00.KER2010            -------------T-----G------------------------------GTT-----------------------G-------------------------------------T-------T-------T--A------A----A--A--------------------C  942
C.MM.99.mIDU101_3          -------------T-----G------------------------------GTT------T---------A------G-------------------------------------T-------T-------T--A------A----A--T--------------------- 1115
C.MW.93.93MW_965           -------------T-----G------------------------------GT-------T---------A------G-----A--------------------------A----T-------T-------T--A---T--A----A-----------------------C  945
C.SN.90.90SE_364           -------------T-----GC-----------------------------GT-------T---------A---------G----------------------------------T-------T-------T--A------A----A--T--G-----------------C  942
C.SO.89.89SM_145           -------------T-----G------------------------------GT-------T---------A------G-------------------------------------T-------T-------T--A------A----A--T--------------------C  942
C.TZ.02.CO178              -------------T-----G------------------------------GT-------T---------A------G-----------------------------T-------T-------T-------T--A------A----A--T--------------------C  942
C.UY.01.TRA3011            -------------T-----G-----------------G------------GT-------T----------------G-------------------------------------T-------T-------A--A------A-C--A-----------------------T  933
C.YE.02.02YE511            -------------TA----------------------G------------GT-------T---------A---------------------------T--------T-------T-------T-------T--A------A----A--T--------------------C  942
C.ZA.04.04ZASK164B1        -------------T-----G--------G---------------------GTT------T---------A------G-----------------------------T-------T-------T-------T--A------A----A--T-----G--------------C 1175
C.ZA.05.05ZAPSK240B1       -------------T--------------------G---------------GT-------T----------------G-----------------------------T-------T-------T-------T--A------G----A--C--------------------C 1168
C.ZM.02.02ZM108            -------------T-----G------------------------------GT-------T---------A------G--------------------T----------------T------CT-------T--A---T--A----A-----------------------C 1742
D.CD.83.ELI                -------------TG-----------------------------------GT-------T-------------G--------------T-----------------------------------------------------------T--------------------- 1292
D.CM.01.01CM_0009BBY       -G------------T-----C-----------------------------GT---T---T----T-----------------A--G-----------------------------------------------A-----------A--A---------------------  939
D.KE.01.NKU3006            ----------------A---------------------------------GT-------T---------A---------------------------------G--T-----------------------------------------T-----------------C---  951
D.KR.04.04KBH8             ----------------------------------------------T---GTTG-----T---------A------------A-----------------------T-----------------------------------------------------------C--- 1705
D.TD.99.MN011              --------------T-----------------------------------GT-------T----T-----------------A--------------T-----------------------------------------------A--T---------------------  964
D.TZ.01.A280               --------------------------------------------------GTT------T-------------------------G--------G-----------T--------------------------------------A--T-----------------T---  941
D.UG.99.99UGK09259         ----------------A---------------------------------GT-------T----------------------A-----------------------T--------------------------------C--------T-----------------T---  951
D.YE.01.01YE386            ----------------A---------------------------------GT-------T----------------G-----A--------------T--------T-----------------------T-----------------T---------------------  948
D.YE.02.02YE516            --------------T-----------------------------------GT-------T----T-----------------A--------------------------------------C-----------A-----------A--A---------------------  951
D.ZA.90.R1                 --------------------------------------------------GT------AT-------------G--------------T--------T--------------------------------A-----------------T--------C------------ 1092
F1.AR.02.ARE933            ----------C-----A---------------------------------GT-------T---------AG-----G-----A-----T------------------A------T----G-C--------T--A--G---A-------T--------------------C 1043
F1.BE.93.VI850             ----------C-----A-----------------------------T---GT-------T---------A------G-----A------------------------A-A----T----GTC--------T-----------------C-----GGG------------C 1089
F1.BR.01.01BR125           ----------C-----A--C------------------------------GT------CT-A-----------G--G-----A------------------------A----T-T------C--------T--A------A-------A-----GGG------------C 1194
F1.ES.x.P1146              ----------C-----A--G--------------G---------------GT-------T----------------G-----------------------------TA------T---------------T--A------A-------A-----GGG------------C 1032
F1.FI.93.FIN9363           ----------C-----A--------------------G------------GT-------T----------------G-----A------------------------A------T----G----------T---------A-------A-----GGG------------C 1081
F2.CM.02.02CM_0016BBY      ----------------A---------------------------------GT-------T----------------G-----A------------------------A------T---------------T---------A-------A------GGC-----------C  939
F2.CM.95.MP257             ----------------A---------------------------------GT-------T---------A------G-----A------------------------A------T---------------T--------CA-G-----A------GGC------------  958
F2.CM.97.CM53657           ----------------A---------------------------------GT-------T---------A------G-----A------------------------A------T---------------T-----------------A------GGC-----------C  939
G.BE.96.DRCBL              -------------------G-----------------------C------GT-------T-------------G--G-----A--------------T-----G--TA-A----T------C--G-----T---------A-------A-------GC-----------C 1710
G.CM.01.01CM_4049HAN       -------------------G------------------------------GT-------T---------A------G-----A--------------T---------A------T-------T-G-----T--A------A----A--A------GGC-----------T  951
G.CU.99.Cu74               --------------------------------------------------GT-------T----T----A------G-----A-----T--------T---------A------TC------T-C-----T---------A-------A------GGC-----------C 1359
G.ES.99.X138               --------------------------------------------------GT-------T----T-----------G-----------T--------------T--TM------T-------T-------T--A------A-------W------GGC------------ 1196
G.GH.03.03GH175G           -------------------G-----------------------C------GT-------T----------------G-----A--------------T---------A-A----T----G----------T--A------A-------A------C-------------- 1786
G.KE.93.HH8793_12_1        --------------------------------------------------GT-------T----------------G-----A--------------T---------A------T-------T-------T---------A-------A------GGC-----------C 1148
G.NG.01.01NGPL0669         -------------------GC-G-----------------------T---GT-------T---------AG-----------A--------C-----T--------TA----T-T-------T-G-----T---------A-------T--------C-----------C  951
G.PT.x.PT2695              --------------------------------------------------GT-------T----T-----------G-----A-----T--------T--------TA------T-------T-------T--A------A-------A------GGC-----------C 1694
G.SE.93.SE6165             -------------------G------------------------------GT-------T-A--------------G-----A-----------------------TA------T----TG-T-G-----T-----------------T------GGC-----------C 1149
H.BE.93.VI991              -G------------------------------------------------GT-------T---------A------------A--------C---------------A------T--G-GT-T-------T---------A-------T--------C-----------C 1138
H.BE.93.VI997              ----------------------------G-----------------T---GTT--T-------------A------G-----A------------------------A------T-------T-------T--------CA----------G--G--------------C 1073
H.CF.90.056                ----------------------------G-----------------T---GT-----------------A------G-----A------------------------A----T-T-------T-------T--------CA-------T--G--G--------------- 1093
J.CD.97.J_97DC_KTB147      -------------TT-A-----------------G---------------GT-------T----------G-----------A-----T--------------G--TA------T-------------------------A----------------------------T  950
J.SE.93.SE7887             -------------T------C-----------------------------GT---T---T----T-----------------A-----T------------------A------T----G----------T---------A-------T--------------------T 1069
K.CD.97.EQTB11C            -------------------G--------------G---------------GT-------T----------------G-----A-----------------------TA------T--G-GT---------T--A-----CA-------A---------------------  945
K.CM.96.MP535              -------------------T-----C------------------------GT-------T----------------G-----A-----------------------TA------T------C--------T--A-----CA-------A-----G--------------C  945
```

```
B.FR.83.HXB2                 AAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGAATGTATAGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAACTCTAAGAGCCGAGCAAGCTTCACAGGAGGTAAAAAATTGGATGACAGAA 1746
Gag                          K__R__W__I__I__L__G__L__N__K__I__V__R__M__Y__S__P__T__S__I__L__D__I__R__Q__G__P__K__E__P__F__R__D__Y__V__D__R__F__Y__K__T__L__R__A__E__Q__A__S__Q__E__V__K__N__W__M__T__E_
01_AE.CF.90.90CF11697        ----G-----------A---------------------------------GTT------T----------------G-----A--------C-----------G--TA------T---------C-----G--A------A----------------------------- 1701
01_AE.CN.05.FJ051            --------------------------------------------------GTT------T-A--------------G-----A--------C--------------TA-A--------------C-----A--A------A----A-----------C------------ 1762
01_AE.CN.06.FJ054            ----G--------------G------------------------------GTTG----------------------G-----A--------C--------------TA----------------C-----A--A--------------A--------------------- 1768
01_AE.HK.x.HK001             ----G--------T------------------------------------GTT------T----------------G-----A--------C--------------TA----------------C-----G--A------A-------A--------------------- 1123
01_AE.TH.01.01TH_R2184       ----G---------------------------------------------GTT------T-------C--------G-----A--------C--------------TA----------------C-----A--A------A-----------------------------  951
01_AE.TH.02.OUR769I          ----G---------------------------------------------GTGG-----T----------------G--------------C-----T--------TA----------------C-----A--A-----------------------C------------  951
01_AE.TH.90.CM240            ----G---------------------------------------------GTT------T----------------G-----A--------C--------------TA----------------C-----G--A------A----------------C------------ 1311
01_AE.US.00.00US_MSC1164     ----G-----------A--G------------------------------GTT------T----------------G-----A--------C--------------TA----------------C-----G--A------A-----------------------------  951
02_AG.CM.02.02CM_4082STN     -----------G-A--------------------G---------------GT-------T-A--------------G-----A--------------T--------TA------T-------T-G-----T--A-----CA-G-----T--------C------------  942
02_AG.EC.x.ECU41             -----------G-----------------------------------------------T----------------------A--------------T--------TA------T-------T-------T--A------A----------------C------------  848
02_AG.GH.03.GHNJ196          -----------G-T------------------------------------GT-------T----------------G-----A--------------T---------A------T-------T-------T--A------A-------T-----G--C------------ 1778
02_AG.NG.01.PL0710           -----------G--------------------------------------GT------CT----------------G-----A-----------G--T--------TA------T-------T-G-----T--A-----CA-G------------GGC------------  936
02_AG.NG.x.IBNG              -----------G-------------------------G------------GT-------T----------------G-----A--------------T--------TA------T-------T-G-----T--A------A-G--------------C------------ 1277
02_AG.SE.94.SE7812           -----------G--------------------------------------GT-------T----------------G-----A--------------T--------TA------T-------T-------T--A------A-------T--------C------------ 1124
02_AG.SN.98.MP1211           --------------------C----------------G------------GT---T---T----------------G-----A--------------T--------TA------T-------T-G-----T--A------A-G--------------C------------  949
03_AB.RU.97.KAL153_2         ----------------A-----------G---------------------GTT------T-A--------------G-----A--------C--------------TA------T---------T-----T---------A-------T-----G--C------------  978
04_cpx.CY.94.CY032           --------------T----G-------------------C---------C-TT-----CT----------------------A--------C-----T--------TA------T----TG---C-----A--A------A-C--------G------------------ 1112
05_DF.BE.x.VI1310            -------------T------------------------------------GT-------T----T--------G--------------T---------------------------------------------------A-------A--------C------------ 1125
06_cpx.AU.96.BFP90           ----------------A---------------------------------GTTG-----T---------A------------A--------C--G--T--------TA------T---------------T---------A-------A--------------------C 1771
08_BC.CN.97.97CNGX_6F        ---------------------------------------------------------------------A------G-------------------------------------T-------T-------T--A------A----A--T--------------------C  944
09_cpx.GH.96.96GH2911        -----------G--------------------------------------GT---T---T----------------G-----A--------C--G-----------TA-A----T---------C-----T---------A-------T-----------A---------  957
10_CD.TZ.96.96TZ_BF061       -------------A--A--G------------------------------GT-------T-A--------G--------------------------T--------T-----------------------------------G-----T-----------------C--- 1125
11_cpx.GR.x.GR17             G------------T------------------------------------GTT------T----------------G-----A--------C-----T---------A-A--T-TC--------T-----T--A--------------------G--C------------ 1057
12_BF.AR.99.ARMA159          ----------C-----A------------------------------------------T----------------G-----A------------------------A------T------C--------T---------A-------A-----G--------------C 1754
13_cpx.CM.96.1849            -------------------G--G---------------------------GT------------------------G-----A--------------T---------A------T-------T-G-----T---------A-------T--------C-----------C 1124
14_BG.DE.01.9196_01          --------------------------------------------------GT-------T----T----A---G--G-----A-----G--------T--------TA------T-------T-------T--A------A-------A-------GC-----------C 1269
15_01B.TH.99.99TH_MU2079     ----G---------------------------------------------GTT------T----------------G-----A--------C--------------TA----------------C-----A--A------A----------------------------- 1154
16_A2D.KR.97.97KR004         --------------T-----------------------------------GT------------------------G-----A-----------------------TA------T---------C-----T---------A-------A--------------------- 1107
17_BF.PE.02.PE02_PCR0155     ----------C-----A---------------------------------GT-------T---------A------G-----A------------------------A------T------C--------T---------A-------A-----GGG------------G  939
18_cpx.CU.99.CU76            -G-----------A------------------------------------GTT------T----------------G-----A-----T--------------G--TA------T---------C-----T-----G---A----------------------------- 1063
19_cpx.CU.99.CU7             --------------------------------------------------GT-------T----------------------------------------------T-----------------------------------------T-----------------T---  972
20_BG.CU.99.Cu103            --------------------------------------------------GTA------T----T----A------G-----A--------------T--------TA------TC------T--C----T---------A--------------GGC-----------C 1221
21_A2D.KE.91.KNH1254         -------------T------------------------------------GT---------A-----------G--------------T-----------------------------------------------------------A--------------------C  954
22_01A1.CM.01.01CM_0001BBY   ----G---------------------------------------------GTT------T----------------G-----A--------C-----------G---A----T-T---------C-----G--A--G---A-------T--G------------------  951
23_BG.CU.03.CB118            --------------------------------------------------GTA------T----T----A------G-----A--G--T--------T--------TA-A----TC------T--C----T---------A-------A------GGC-----------C 1195
24_BG.CU.03.CB378            --------------------------------------------------GTA------T----T----A------G-----A--------------T--------TA------T-------T-------T--A------A-------A------GGC-----------C 1195
24_BG.CU.03.CB471            --------------T-----------------------------------G-A------T---------A------G-----A-----------G--T--------TA------TC------T-------T--A------A-------A------GGC-----------C 1195
25_cpx.CM.01.101BA           --------------------------------------------------GT-------T----------------G-----------T-----------------TA------T-------T-------T---------A-------A------GGC------------  954
27_cpx.CD.97.97CDKTB49       --------------------------------------------------GT---T---T----------------------A--------------T--------TA------T-------T-G-----A--------------A--A--------------------C  957
28_BF.BR.99.BREPM12609       ----------C-----A--------------------G------------GT-------T----------------G-----A-----------------------TA------T------C--------T------TG-A-------A------GG------------C 1115
29_BF.BR.02.BREPM119         ----------C-----A---------------------------------GT-------T----------------G-----A--------------T-----G---A------T------CT-------T---------A-------A-----G--------------C  923
31_BC.BR.02.110PA            -------------T-----G------------------------------GT-------T---------A------G------------------------------A------T--G---AT-------A--A------A-C--A--T--G--G--------------T 1195
32_06A1.EE.01.EE0369         ----------------A-----------G--------------------GGTT------T-A--------------G-----A--------C--------------TA------T---------T-----T---------A----C--T-----G--C------------ 1387
33_01B.MY.05.05MYKL007_1     ----G---------------------------------------------GT-------T----------------G-----A-----T--C--------------TA-------C--------C-----G--A-----AA-------------------------G--- 1131
34_01B.TH.99.OUR2478P        ----G---------------------------------------------GTT------T----------------G-----A--------C--------------TA----------------C-----G--A------A-------A--------C------------  951
35_AD.AF.05.05AF095          -------------------------------------G-----------CGTTG-----T----T----A------G-----A-----T--C--------------T-------T----GT---C-----T---------A-------------GGG-------------  948
36_cpx.CM.00.00CMNYU830      ----G---------------------------------------------GTT------T----------------G-----A--------C--------------TA------T---------T-----G--A------A-------T---------------------  954
37_cpx.CM.00.00CMNYU926      ----------------A---------------------------------GT-------T----------------G-----A--------C--------------TA------T---GGT---C-----T-------T-A----------------C------------  945
38_BF1.UY.05.UY05_4752       ----------C-----A---------------------------------GT-------T----------------G-----A------------------------A------T----G-C--------T--------CA-------A-----G--------------C 1160
39_BF.BR.03.03BRRJ103        ----------------A-----------------------------T---GT------------------------G--------------C--------------------------------------------------------T--------------------- 1220
40_BF.BR.04.04BRRJ115        ----------C-----A---------------------------------GT-------T-----------------------------------------------A------T------C--------T--A--------------A--------------------C 1292
42_BF.LU.03.luBF_05_03       ----------C-----A---------------------------------GT------CT----------------G-----R------------------------A------T---G--CT-------T---------A-------A-----GGG------------C 1275
43_02G.SA.03.J11223          -------------T-----G-----------------------------AGT-------T---------A------------A--------C-----T--------TA------T-------T-G-----T--------CA-------T--------------------C 1271
U.CD.83.83CD003              -G-----------------------------T--------------T---GT-------T---------A------T-----A--------------T-----G---A-A----T----GT-T-------T------T----------A--------------------C 1244
U.CD.90.90CD121E12           -G-----------------------------T--------------T---GT-------T----------------------A-----T--------T-----G--TA-A----T------CT-------T------T----------A--------------------C  948
U.GR.99.GR303                -------------T------------------------------------GT-------T----------------G-----A--------C--G-----------TA------T---G-----C----GHA--------A----------------------------- 1029
U.NL.95.U_NL_95_H10986_D1    ----------------A--T------------------------------GT-------T---------A------G-----A-----------------------TA------T----TA---------T--A-----------A--T--------------------T 1634
CPZ.CD.90.ANT                -G--------C--A--------------GG-G--C--G-N------T--AGT---------A--G----A------------A--G-----C-----T-----G--TA-A-----------AA-T-----A--A--G---------CCT--G---GCC------------ 1167
CPZ.CM.01.SIVcpzCAM13        G-------G----T--A--GC----C-G-G------A-------T-T--AGTA-----C--T--T----AG-----------A--G-------A---T---------A----T--------CA-G-----T-----------T-----A--G----C------------C 1323
CPZ.CM.05.SIVcpzEK505        G----------G--T-------------G--T--------------T--G-----------A-----C-A---G--------A--GA-T--------T--C-----TA----T--C--------------A-----G--CA-T-----A-----G--------------G 1294
CPZ.CM.05.SIVcpzLB7          -------------TA----GC-G-----GG----C--------C------GTT-----C----------AG-----G-----A-----A--C-----------G---A----T--C-----AT-------A--A-----CA-------A--------------------C 1292
CPZ.CM.05.SIVcpzMB66         -------------TT----GC-------G-------A-------T-----GT------A--A-------A------G-----A--GT-A--C-----T-----G---A-A--T---------T----G--A--A-----A-----A--A--------------------C 1289
CPZ.CM.05.SIVcpzMT145        G----------G-------G--------------T-A------C--T---GTT------T-------T-A------------A--G--T--C--G------------A-------C--G--CT----G--T--A-----A--------T--T-----------------C 1285
CPZ.GA.88.GAB1               G-------G----TT-A--G-----C---G-G--T---------T----AGTA--T--C--T-----C--------G--------G-----------T-----G---A-A-----C------T----G--T-----------------A--G--------------T--C 1807
CPZ.TZ.00.TAN1               ----G--------TT-A------------T----C--------C------GTT------T---------A---G--------A-----A--C-----T---------A-A-----C-----AA-C-----A--A-----A--T--ACCA-------C------------- 1387
CPZ.US.85.CPZUS              G-------G-GG----A--GC-T-----G-------A------CT-T--CGTTG-T--CT-A--T--C-AG--G-----T--A--G---------------------A-------C-----AT-------A--A--------T-CA--A-------CC-----------G 1794
CPZ.CM.x.SIVcpzDP943         G-------G-GG-T--A--GC-T--C---G-G-----------------AGT---T--C--A-----T-AG--G-----T-----G--T--C--G-----------TA-A-----C-----C--G--G--A--A--G--AA---------------CC--------T--T 1305
CPZ.CM.x.SIVcpzMB897         -G---------G-TT-A--GC-T---------------------------GT---T-----A--T----AG-----------A-----A--------T-----G---A-A--T--C--------------T--A-----AA-C--A--T--------------------G 1291
N.CM.02.DJO0131              ---A-------G-T-----G-------G---T--G-----------T---GT---------A--G--C-A------------A--------------T---------A-------C------T-------A-----G--A--------T-----G--------------- 1231
N.CM.04.04CM_1015_04         G----------G-TT----G-----C-G---T--G-----------T---GT-------T-A--G--C-AG-----------A--------------T---------A-------C------T-------A-----G--AA-------T-----G--------------- 1224
N.CM.04.04CM_1131_03         G----------G-TT----G-----C-G---T--G-----------T---GT---------A--G--CGAG-----------A--------------T---------A-------C---GT-T-------A-----G--AA-------T-----G--------------- 1223
N.CM.95.YBF30                G----------G-------G-----C-G---T--G-----------T---GT---------A--G--C-A------------A--------C---------------A-------C--------------A-----G--AA-------A-----G--------------- 1310
N.CM.97.YBF106               G----------G-T-----G-----C-G---T--G-----------T---GT---------A--G--C-AG-----------A--G-----------T---------A-------C------T-------A-----G--AA-------T-----G--------------- 1311
O.BE.87.ANT70                G--A-------G-GT-A---C----C-----G----A------C-----AGTG-----CT-A--T--T-AG--G--------------A-----------------T--------C-----AT-------T--------A--T--A--A--------------------- 1794
O.CM.91.MVP5180              G--A-------G-G--A---C----C-----G----A------C--T--AGTG-----CT-A--T--T--G--G--------A-----A--C-----T--------T-----T--C-----AT-------T--------A--T--A--A-----G--------------- 1769
O.CM.96.96CMABB637           G--A-------G-G--A---C----------G----A------C--T--AGTG-----CT-A--T--T-A---G--------A-----A--C-----T--------TA-------C-----AT-------T--A-----AA-C-----A--R------------------ 1218
O.CM.98.98CMA104             G--A-------G-GT-A--GC----C-----G----A------C--T---GTG-----CT-A--T--T-A---G--------A-----G-----------------T--------C-----AT-------A--A-----A--T--A--A--------------------- 1226
O.CM.99.99CMU4122            G--A-------G-GT-A-----R--C-----G----A------C--T--AGTG-----CT-A-----T-----G--------A-----A-----------------TA-------C-----AT-------T--------A--T--A--A--------------------- 1225
O.FR.92.VAU                  G--AG------G-AA-A---C----C-----G----A------C--T--AGTG-----CT-A-----T-AG---A-----G-A-----A--------T-----G--T-A------C-----AT-------T--------AA-T--A--A--------------------- 1303
O.SN.99.SEMP1299             G--A-------G-GT-A--------C---G-G----A------C-----AGTG-----CT-A--T--T-----G--------A-----A-----------------T--------C-----AT-------T--A--G--AA-T--A--A--------------------- 1796
O.US.99.99USTWLA             G--A-------G-A-----TC----C-----G----A------C--T--AGTG-----CT-A--T--T--G--R--------A--G--A--------T--------TA-------C-----AT-------T--------A--T-----A--------------------- 1214
```



HIV-1/SIVcpz Complete Genomes

```
                                                                                                                                                        Gag p24 Capsid end Gag p2 start
B.FR.83.HXB2                    ACCTTGTTGGTCCAAAATGCGAACCCAGATTGTAAGACTATTTTAAAAGCATTGGGACCAGCGGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGGACCCGGCCATAAGGCAAGAGTTTTGGCTGAAGCAATGAGCCAAGTAACAAATTCA........ 1908
Gag                             _T__L__L__V__Q__N__A__N__P__D__C__K__T__I__L__K__A__L__G__P__A__A__T__L__E__E__M__M__T__A__C__Q__G__V__G__G__P__G__H__K__A__R__V__L__A__E__A__M__S__Q__V__T__N__S__.__.__.
01_AE.CF.90.90CF11697           -----AC-A--------------T-----C------T-C--C-----------A---A---G-------T----------------------C--------G--------T--------A-----------------G--------T------CA-C--G--........ 1863
01_AE.CN.05.FJ051               ------C-A--T-----------T------------T-C-C------------A---AT--GA------T----------------------C--------G--------TA-------A--G--G---C-------G-------------CGCA-C--A--........ 1924
01_AE.CN.06.FJ054               --------A--------------------C------T-C--------------A---A-G-GA---T--T----------------------C--------G--------TA-------A-----G-----------G---------AGC---CA-C--A--........ 1930
01_AE.HK.x.HK001                ------C-A--------------T-----C------T-C--------------A---G---GA-----TT----------------------C--------G--------TA-------A-----G-----------G-------A-----C-CA-C-GA--........ 1285
01_AE.TH.01.01TH_R2184          ------C-A--------------T-----C------T-C--------------A---G----A------T----------------------C--------G--------TA-------A-----G-----------G-------------CTCA-C--G--........ 1113
01_AE.TH.02.OUR769I             ------C-A--------------T-----C------T-C---C----------A---A---GA------T----------------------C--------G--------TA-------A-----G-----------G-----------C-C-CA-C-GG--........ 1113
01_AE.TH.90.CM240               ------C-A--------C-----T-----C------T-C--------------A---A---GA------T----------------------C--------G--------TA-------A-----G-----------------------C-C-CA-C--G--........ 1473
01_AE.US.00.00US_MSC1164        ------C-A--------------T-----C------T-C--------------A---A---GA---T--T-------------------G--C--------G--------TA-------A--T--G-----------G-----------C---CA-C--G--........ 1113
02_AG.CM.02.02CM_4082STN        ------C----------------T-----C--------C-------G--G---A-------G-------T-------------------------A-----G--------T--------A-----G-----------G--------T----C-CA-C-GG-C........ 1104
02_AG.EC.x.ECU41                ------C----------------T-----C--------C-------G------A-------G-------T-------------------------------G--------T--------A--T--G-----------G--G-----T----C-CA-C-GG-C........ 1010
02_AG.GH.03.GHNJ196             ------C-------------T--T-----C--------C-------G------A-------G----T--T-------------------------------G--------TA-------------G--------------------T----C-CA-C-G--C........ 1940
02_AG.NG.01.PL0710              ------C----------------T-----C------T-C-------G------A-------G-------T-------------------------------G-----------------A-----G-----------G--------T------CA-C-GA-C........ 1098
02_AG.NG.x.IBNG                 ------C----------------T-----C------T-C-------G------A---A---G-------T-------------------------------G--------T--------A-----G-----------G--------T------CA-C-GG-C........ 1439
02_AG.SE.94.SE7812              ------C----------------T-----C------T-C-------G------A-------G-------T-------------------------------G--------TA-------A-----G-----------G--------T------CA-C-GC-C........ 1286
02_AG.SN.98.MP1211              ------C-----------T----T-----C----G---C---C----------A-------G-------T--------------T-------------------------T--------A-----G-----------G--------T----C-CA-C-G--C........ 1111
03_AB.RU.97.KAL153_2            ---C--C-------------A--T-----------A--C-------G------A---T---G-------T------------------T---C--------G-----------------A-----G--------C--G--------T------CA----G--........ 1140
04_cpx.CY.94.CY032              ---C--C----------------T-----C------T-C--C-----------A---A---G-------T-------------------------------G---------A-------A-----------A-----G-----------G-C-T-----G--GCAGCA.. 1280
05_DF.BE.x.VI1310               --------------------A--------------A-----C-----------A----A-CA---------------------------------------G--G--G---A-------A-----------------G-------------C----------GCTGCT.. 1293
06_cpx.AU.96.BFP90              --------------------A-----------------C-------G------A-------GA--------------------------------------G--------T--------------------A-----G-----------G-C-T--GT-GG-........ 1933
08_BC.CN.97.97CNGX_6F           --------A-----------------------------C-------G------A---T---G----T--T----------------------------------------TA----C--A--------G--------G------------AC--AC---A-C........ 1106
09_cpx.GH.96.96GH2911           ------C----------------T-----C--------C--C----G------A-------G-------T-------G-----------------------G--------T--------A--------------Y-----------T------CA----A--........ 1119
10_CD.TZ.96.96TZ_BF061          -----------------------------------A-----C---------------------------T-------------------------------G--G------A-T-----A-----------A-----G-----------G-C----GG-GG-........ 1287
11_cpx.GR.x.GR17                ------C----------------T-----C------T-C-------G------A---GT--GA-----------------------------C------------------A-------A-----G-----------G-------------C-CA-C--A--........ 1219
12_BF.AR.99.ARMA159             -----------------------------------A--C------------------T---G---------------------------------------G--------T----------------------------------------C-----G-A--........ 1916
13_cpx.CM.96.1849               -----------------------T--------------C--C----G------A-------G---------------------------------------G-----------------A-----------A-----------------G-CCT---G-G--GCAGCA.. 1292
14_BG.DE.01.9196_01             --------------------------------------C--C----G------A-------GA--------------------------------------G---------A-T--C--R-----------A-----G-----------G-C-T--GGGG--........ 1431
15_01B.TH.99.99TH_MU2079        --A---C-A-C------------T-----C------T-C--------------A---T---GA---T--T-------------------------------G--------TA-------A-----------------G-----------T---CA-C-GA--........ 1316
16_A2D.KR.97.97KR004            ------C----------------------------AT-C--C--G-G---C--A-------G----T--T-------------------------------G---------A-------A-----G---------------------------CA----A--........ 1269
17_BF.PE.02.PE02_PCR0155        -----------------------------------A--C--C----G----------T---G---------------------------------------G--------T--------------------------G------------------GG-A--........ 1101
18_cpx.CU.99.CU76               -----AC----------------T-------------AC-------G------A-------G---------------G-----------------------G---------A-------A--T--G--------------------T--T---CA----A-C........ 1225
19_cpx.CU.99.CU7                --------------------A--------------A-----CC------------------G-------T-------------------------------G--------TA-T-----A-----------------G-------------C----------........ 1134
20_BG.CU.99.Cu103               -----------------------------C--------C--C----G------A-------GA-----------------------------C--------G-----------------A-----------A-----------------G---T--GG-G--........ 1383
21_A2D.KE.91.KNH1254            ------C----------------------------AT-C--C--G--------A-------G-------T-------------------------------G---------A-------A---------C-A-----G-------------C---G-GC---........ 1116
22_01A1.CM.01.01CM_0001BBY      ------C-A--------------------C-----AT-C-------G------A--GA---G-------T-G--------------------C--------G--------G--------A-----G-----A-----G--------T------CA-C--A--........ 1113
23_BG.CU.03.CB118               --------------------------------------C--C----G------A-------GA----GT-----------------------------------------T--------A-----------A-----------------GA--C--GG-G--........ 1357
24_BG.CU.03.CB378               ---------A-------------------C--------C--C----G------A-------GA--------------------------------------G--------T--------A-----------A-----------------G---T--GG-A--........ 1357
24_BG.CU.03.CB471               -----------------------------C--------C--C----G------A-------GA--------------------------------------G--------T--------A-----------A-----------------G---T--GG-A--........ 1357
25_cpx.CM.01.101BA              --T-----------------A--T--------------C--C----G------A-------GA---T-------------------C--------------G-----------------A-----------A-----G-----------G---T--GG-A--........ 1116
27_cpx.CD.97.97CDKTB49          --------------------------------------C--C-----------A---T---GA-----------G--------------------------------G---A-------A-----------------------------G-CCT--GG-G--GCA..... 1122
28_BF.BR.99.BREPM12609          --------------------------------------C----------------------G---------------------------------------G--------TA-------------------------G-------------C-------A--........ 1277
29_BF.BR.02.BREPM119            --------------------------------------C--C-------------------G------------------------------------------------TA-------------G-----------G-----------C-C-------A--........ 1085
31_BC.BR.02.110PA               --------A-----------------------------C-------G------A-G-----G----T--T-------------------------------G--------T-----C--A--------G--------G----------------AC---A--........ 1357
32_06A1.EE.01.EE0369            ---C--C-------------A--T-----------A--C-------G------A-------G-------T----------------------C--------G-----------------A-----G--------C--G--------T----C-CA----G--........ 1549
33_01B.MY.05.05MYKL007_1        -----AC-A-----------A--T-----C------T-C--------------A---A---GA------T-G--------------------C--------G--------TA-------A-----G-----------G-----------C-C-CA-C-CA--........ 1293
34_01B.TH.99.OUR2478P           ------C-A--------------T-----C------T-C--------------A---A---GA------T----------------------C-----------------TA-------A-----G-----------G------------AC-CA-C-GG--........ 1113
35_AD.AF.05.05AF095             --A--A-----T--------A--T------------T-C--CC-G-G------A---A---G-------T----------------------C--------G---------A-------A-----------A-----G--------T----C-CA-C--........... 1107
36_cpx.CM.00.00CMNYU830         ------C-A--------------------C------T-C--------------A---A---G-------T----------------------C--------G--------T--------------G-----A-----G---------------CA-C--........... 1113
37_cpx.CM.00.00CMNYU926         ---C--C-A--------------T-----C------T-C--------------A---A---G-------T----------------------C-----------------T--------A-----G-----A-----G-------CT----C-CA-C-G........... 1104
38_BF1.UY.05.UY05_4752          --------------------A--------------A--C---C------G-------AT--G---C-----------------------------------G--------T--------------------------G-----------G---------C--........ 1322
39_BF.BR.03.03BRRJ103           ------C-A--T-----------T--------------C--------------A--------A--------------------------C-----------G--G------A-------A-------------------------------C-------CA-........ 1382
40_BF.BR.04.04BRRJ115           --------------------------------------C----------------------G---------------------------------------G--------T--------------G-----------G--------T---A--------A--GCA..... 1457
42_BF.LU.03.luBF_05_03          --------------------------------------C----------------------G---------------------------------------G-----G-T---------------------------G-----------G-C----GG-........... 1434
43_02G.SA.03.J11223             -------------------------------------AC--CC---G------A---T---G---------------------------------A-----G---------A----C--------G-----A-----G-----------G-C-T--GG-G--........ 1433
U.CD.83.83CD003                 ---------A-T--------A-----------------C--------------A---A---G------------G--------------------------G---------A-------A--------G--------G--------T-GAC--AT............... 1400
U.CD.90.90CD121E12              --------AACT--------A-----------------C--------------A---A---G-------T----G--------------------------G--------T--------A-----------------G-----------GA-G------........... 1107
U.GR.99.GR303                   ------C-----A-G--------T-----C--------C--------------A---T---GA------T----------------------C--------G--------T--------A-----G--------------------T--G---CA--G-........... 1188
U.NL.95.U_NL_95_H10986_D1       -----------------C--A-----------------C-----------G-----------A---T----------------------------------G--------T-------------GA-----------G-----------G-C------G........... 1793
CPZ.CD.90.ANT                   --------AA----------C--T--------C--ACAC--CC-G--G--T------A---GA--CT-CT-----------T-A-----T-----A-----------C--A-C---------------G-----A-----T---GCTTCT-CT-AT---G--CAG..... 1332
CPZ.CM.01.SIVcpzCAM13           ---C-T-----G--------T-----------C---CAA---C----G-CC-----T-G-G----CT-C--------------------C-----A--G-----------AC----------------C--A--------T----CAATG--CCA--G-CA-GCAAGAAC 1493
CPZ.CM.05.SIVcpzEK505           --AC-C--A--A--------A--T--T--------A-A--CC-G-G-------A-------GA--CT-C--G--G-----------------C--A--G--G--------ACT---------------AC-A--A-----C----CA--G-C-CA--CAG--........ 1456
CPZ.CM.05.SIVcpzLB7             ------C-T--A--G-----T-----------C--A-G---CC-G-G---------T----GA--CT-CT----G-----------------C--------C--G--T--AC-------A--C--G--C--A--A-----T-----T----C--ATGCAG--........ 1454
CPZ.CM.05.SIVcpzMB66            --T--A-----G--------T--T------------T-C--A---G-------A-G-----GA-----CT------GGGATGA-C----C-TC--------G---------A-------A--T-----CC----A-----------T----C-CA-C-CAGT........ 1451
CPZ.CM.05.SIVcpzMT145           ---C----A-A--G------T--------C-----A-AC--C----G------A-------GA-C--T---------------------T-----A--G----------TCA----------T-----GC----A-----C---------G---AC--AA--GAA..... 1450
CPZ.GA.88.GAB1                  --TC-C--A--G--G-----A---------------CAA--C-----------------T-G--C--C------G--------------T-----A-------------ATCT------------G---C-A--------C---...TC-A-GGTTC-GAATCAA..... 1969
CPZ.TZ.00.TAN1                  --T--AC---A---------A---------------CA---C--------C-----G-A--GA--A---T-----------C-C-----C-----A-----G-------TCT------------AGA--C--------------GC-TC--C----GCAGGGGGA..... 1552
CPZ.US.85.CPZUS                 ---C-T-----A-----C--T--T-----------A---------G----C--------T-GA--C--C------------G-------------A-----G--------CT-------------G--C--A--------C---T-----A-G-A---CC-C........ 1956
CPZ.CM.x.SIVcpzDP943            ---C-CY----A--------T--------C-----A-A--C--G-GG--C--A--------GA--C--CT-------G--------------C--A-----G--------A-C---------T--G--------------T------------CA---C--C........ 1467
CPZ.CM.x.SIVcpzMB897            ------C-A--G--------C---------------CAA-------G---TC---T---G--A--CT----------------------C--C-----------G--C--TCT------A--C-----A-----A-----------T--G-C--AC-CC........... 1450
N.CM.02.DJO0131                 --AC-C--A--A--------A--T-----------ACAG--AC----------A-G------A-----CT----------------G--C-----------G--G-----ACG---------------GC-A--A--G--T---TCA--G---CAG-CAG--........ 1393
N.CM.04.04CM_1015_04            --AC-C--A--A--------A--T-----------RCAG--CC----------A------T-A---GCT--------G--------G--C--C--A-----G--G-----A--G--------------GC-A--A--G--T---GCA--G-C-CAGGCAG--........ 1386
N.CM.04.04CM_1131_03            --AC-C--A--A--------A--T-----------ACAG--CC----------A-G-----GAT-CT-CT-------G--------G--C--C--------G--G-----A--G--------------GC-A--A-----T-C-GCA--G-C-CAGGCAG--........ 1385
N.CM.95.YBF30                   --AC-C--A--A--------A--------------ACAGC-CC----------A-G-----GA-----CT-------G--------G--C--C--------G--G-----ACA---------------GC-A--A--G--T---TCA--G--GCAGC-GC--........ 1472
N.CM.97.YBF106                  --AC-C--A--A--------A--------------ACAG--CC----------A-------GA-----CT-------G--------G--C--C-----------G-----ACA---------------GC-A--A--G--T---GCA--G-C-CA--CAG--........ 1473
O.BE.87.ANT70                   --TC--C-T--T--------C--T--------C--ACAG-----G---T----A-G-----GA-----CT-------G------GT---C--------------------AACT--------C-----GC-A--A---------GCTAC--CCCAGC-AGATTTGAAAGG 1964
O.CM.91.MVP5180                 ------C-T--T--G---T-A--------------ACAA---C-G--------A-------AA-----TT--------------GT---C-----A-----------G--AACT--C-------A-A-AC-A--A---------GCTTCT-CCCAGC-AGATTTAAAAGG 1939
O.CM.96.96CMABB637              ---C-AC-T--T--G-----C--T--------Y--ACAA---------T-C--A-G-----GA-----CT--------------GT---------------------G--AACT------------A-AC-A--A--R------GC-TC---C-ACC-AGATCTAAAAGG 1388
O.CM.98.98CMA104                --TC--C-T--T--------C--------------ACAG-----G------C-A-G-----GA-----CT--------------GT---C--------------------AACA--------C------G-A--A---------GC-TC--CCCAGC-AGATTTAAAAGG 1396
O.CM.99.99CMU4122               ------C-T--T--------C-----------C--ACAG-----G---T----A-G-----GA-----CT----G--G------GT---C--------------------AACT--------C--G--GC-A--A---------GCTAC--CCCAGC-AGATCTGAAAGG 1395
O.FR.92.VAU                     ---C--C-T--T--------C-----------C--ACAA-----G--------A-G-----GA----GC---------------GT---T-----------G--------AACT--------T-----CC-A--A---------GCTTC--CCCAGC-AGATCTAAAAGG 1473
O.SN.99.SEMP1299                ---C--C-T--T--------C-----------C--ACAG-----G---T----A-G-----GA-----CT---------------T---T--------------------AACT--------C--G--A--A--A---------...GCT-C-G-CC-AGATCTGAAGGG 1963
O.US.99.99USTWLA                --TC--C-T--T--G---T-C-----------C--ACAG-----G--------A-------GA-----CT-------G------GT---C-----R-----------G--AACT----------A-A-A-----A---------GC-TC--CCCAGC-AGATCTAAAAGG 1384
```

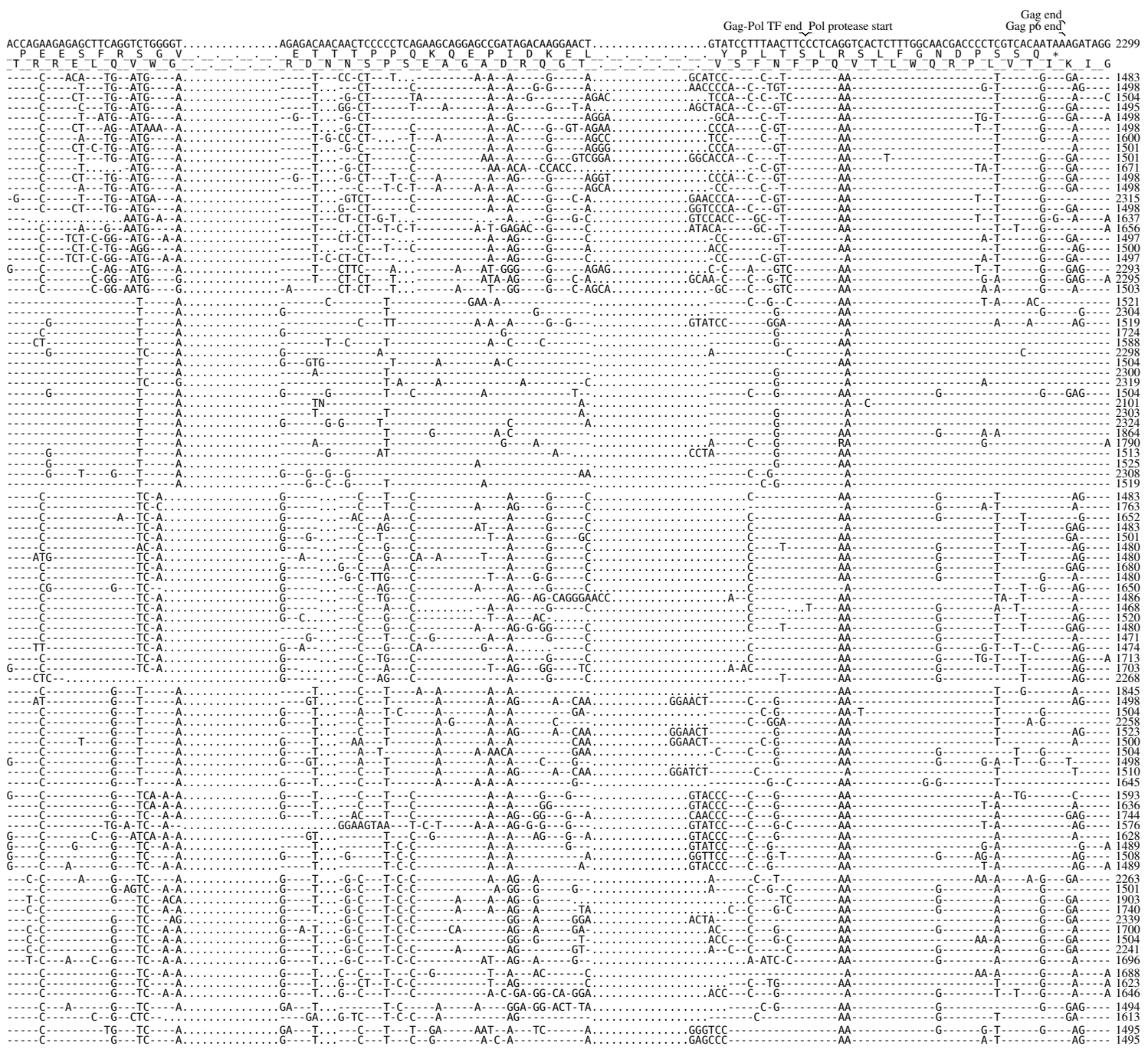
Gag-Pol fusion TF protein start
Gag p2 end Gag p7 nucleocapsid start

```
B.FR.83.HXB2              .......GCTACCATAATGATGCAGAGAGGCAATTTTAGG...AACCAAAGAAAGATTGTTAAGTGTTTCAATTGTGGCAAAGAAGGGCACACAGCCAGAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGATTGTA 2068
Gag                       __.__.__A__T__I__M__M__Q__R__G__N__F__R__.__N__Q__R__K__I__V__K__C__F__N__C__G__K__E__G__H__T__A__R__N__C__R__A__P__R__K__K__G__C__W__K__C__G__K__E__G__H__Q__M__K__D__C__
A1.GE.99.99GEMZ011        ..........-AT-----------A-A-A-T---------...GG--C-----GA...A-------------C--------------A---CT---------------------------------------------------------C--T-----------C--C- 1252
A1.KE.00.KER2008          ..........-ATG-------------G----------A-...GG---G-A--GA...A-------------C-----------G--A---CT---------------A--------------A-----------G-----G--A-----------------------C- 1267
A1.KE.00.KNH1144          ..........-A------------A--------C------...GG---G-A--GA...A-------------C--------------A---CT------------------------------------------------G-----T-----T-----------C--C- 1273
A1.KE.00.KSM4024          ..........-A----------------------------...GG---G-A--G-...A-------------C--------------A--TCT------------------------------------------------G--A--G-----T-----------C--C- 1264
A1.KE.00.MSA4069          ..........-AAG------------A-------------...GG---G----G-...A-------C-----C-----T--------A---CT------------------------------------------------G-----G-----------------C--C- 1267
A1.KE.00.NKU3005          ..........-A----------------------------...GG---G-A--G-...A-------------C--------------A---CT------------------------------------------------G-----G-----------------C--C- 1267
A1.RU.00.RU00051          .......A-A-AT-------------A-------------GGCCC--CG-A--GA...A-------Y-----C--------------A---CT---------Y-----------------------------------------------C--T-----------C--C- 1369
A1.RW.93.93RW_024         ..........-A----------------------------...GG---G-A----...A-------------C--------------A---CT---T--------------------------------------------G-----G-----------------C--CT 1270
A1.SE.95.SE8891           ..........-A-G------------------------A-...GG---G-A--GA...A-------------C--------------A---CT------------------------------A-----------------G-----------T-----------C--C- 1267
A1.SE.95.UGSE8131         ..........-A----------------------------...GG---G-A----...A-------------C--------------A---CT------------T--A--------------------------------G-----G-----------------C---- 1458
A1.TZ.01.A173             ..........-A------------------T---------...GG---G-A--G-...A----------T--C--------G-----A---CT------------------------------------------------G--A--G-----------------C--C- 1267
A1.UA.01.01UADN139        ..........-A--------------A-A-T--------A...GG--C--A--GA...A-------------C--------------A---CT---------------------------------------------------------C--T-----------C--C- 1267
A1.UG.92.92UG037          ..........-A--------------------------A-...GG---G-A--G-...A-------------C--------------A--TCT-----A------------T-----A-------------------------G------G-----------G--C--C- 2081
A1.UG.99.99UGA07072       ..........-A-------G----------------C---...GG---G-A--G-...A-------------C--------------A---CT---T--------------------------A-----C-----------G--A--------------------C--C- 1267
A2.CD.97.97CDKTB48        ..........-A---------A----------------A-...GGT----A--GA...A-------------C--------G-----A---CT---T--------------------------------C-----------G-----------T-----------C--C- 1427
A2.CY.94.94CY017_41       ....AATA-A-A----------------------------...GGT----A--GA...A-------------C--------G-----A---CT------------------------------------C-----------------------T--------------C- 1425
A.SN.01.DDI579            ..........-A-G--------------------------...GGT--G-A--GA...A-------------C--------------A---CT---T--------------------------------------------G-----G-----------------C--C- 1269
A.SN.01.DDJ369            ..........-A--------------A-----------A-...GGT--G-A--GA...A----------------------------A---CT------------------------------------------------G-----G-----------------C--C- 1272
A.SN.96.DDJ360            ..........-A----------------------------...GGT--G-A--GA...A-------------C--------------A---CT---A--------------------------------------------G-----G-----------------C--C- 1269
A.CD.02.02CD_KTB035       ..........-AT-------------A------------A...GG--CG----GA...A-------------C-----------G--A--TCT------------------------A--G--------------------G-----------------------C--C- 2059
A.CD.97.97CD_KCC2         .......A-C-TA--G---C--A-AG-------C-----A...GG---G-A--GA...A-------------C-----------G--A---CT------------------------A--G--------------------G-----------T-----------C--C- 2064
A.CD.97.97CD_KTB13        ..........-A--------------A------------A...GG--CG-----A...A-------------C--------------A--TCT---------------------------G--------------------G-----------T-----------C--C- 1272
B.AR.04.04AR151516        .......----A----------------------------...--T----------C---------------------------G-------T--------------A---------------------------------------------T---------------- 1275
B.AU.87.MBC925            .......---------------------------------...-------------C------A----------------------------T---T-A---C------------------------------------------------------------------- 2073
B.BO.99.BOL0122           .......--C---G-----------C------------A-...T-TAC--A-----C------A--------------------G-------T---------------A-----G--A-----------------------------------------------G---- 1285
B.BR.03.BREPM2012         .......--C-A----------------------------...CC---------------------C-----------------G-------T-----A----------------------------------------------------------------------- 1478
B.CA.97.CANB3FULL         .......AA-G-TG--------------------------...------------CC---------------------------G-------T---A-A----------------------------------------------------------------------- 1348
B.CN.05.05CNHB_hp3        .......----AT-----------------------C---...-----------------C-------------------------------T-----A------T--------C-------------------G---------------------------G--G---- 2067
B.CO.01.PCM001            .......A--G------------------------A--A-...-----------ACC---------------------------G-------T---T-A------------------A--G-----T------------------------------------------- 1273
B.GB.83.CAM1              .......---------------------------------...-------------C---------------C----------TC-------T---T-A------------------------------------------------------T---------------- 2069
B.GB.86.GB8               .......-------------------A-------------...-G-----------C-----------------------------------T------------------A--------G-----------------------------------G-----G--G---- 2070
B.GE.03.03GEMZ010         .......C--------------------------------...-------------C-----------------------------------T-----A--------------------G---A-----------------G--A-----------------------C- 1273
B.IT.05.SG1               .......A------------------------------N-...----------N--------------------------------------T------------------------------A---------------------------------------------- 1870
B.JP.05.DR6538            .......----G-------G-A-----------------A...---------------A---------------------------------T-T---A------------------------------------------------------------------C---- 2072
B.KR.05.05CSR3            .......--C----------------------------A-...GG---------A--G-----A--------C-------------------T-----A------------------A--------------------------------------------G------- 2093
B.NL.00.671_00T36         .......C--G-----------------------CA---A...----------G--C-----------------------------------T------------------------N----------------------N----N------------------------ 1603
B.RU.04.04RU128005        .......----R--------------------------A-...--Y--G-A--G--G-----------------------G----R------T------------------T-----------R-------R-------R-------------T---------------- 1559
B.TH.00.00TH_C3198        ....GCA------G--------------------------...--T----------C---C-------------------------------T-----A------------------A----G----------------------------------------------- 1282
B.UA.01.01UAKV167         .......---------------------------------...-----------A-C---------------C-------------------T-----AG-----------------------------------------T-G-------------------------- 1273
B.US.04.ES10_53           .......--AG---------------A-------------...-----------A-C---------------C------------------CT--------------------------------------------------G-------------------------- 2077
B.US.99.PRB959_03         .......AAC--------------------T---------...--T--------A-------------------------------------T------------T-----------------A-----------------G-----------------------A---- 1288
C.AR.01.ARG4006           ..........-A--------------------------A-...GG--CT-A--GA-C-A----A--C-----C-------GG-----A---TT------------------------------A-----------------------------------------C---- 1264
C.BR.04.04BR013           ..........-A--------------------------AA...GG--CT-A--GACC-A----A--C--T--C--------G---------TT-----A------------------------A-----------------------------------------C---- 1520
C.BW.00.00BW07621         ..........-A-------T------A-AA--------AA...GG--CT----GA--------A--------C--------G--G-----T-T------G-----------------------A-----------------------------------------C---- 1433
C.CN.98.YNRL9840          ...............................A-T-------AA...GG--CT-A--GA--------A--------C--------G----------T------------------------------A-----------------------------------------C---- 1264
C.ET.02.02ET_288          ..........-A--------------A-----------A-...GG--CT----GA-C-A----A--C-----C--------G---------CT------------------------------A-----------------------------------------C---- 1264
C.GE.03.03GEMZ033         ..........-A------------------------C-AA...-G--CT-A--GA--------A--------C--------G--------T-T------------------------------A--------------------------------G--------A---- 1261
C.IL.99.99ET7             ..........CA----------------AAT-------AA...GGT-CT-AG-GA--------A--------C--------G---------CT------------------------------A-----------------------------------------C---- 1261
C.IN.99.01IN565_10        .............---C-----------A---------AA...GG--CT-A--GA---A----A--------C--------G----------T---A-A---------------G--------A-----------------------------------------A---- 1461
C.KE.00.KER2010           ..........-A---------T------A---------AA...-G--CT-A--GA--------A--------C--C-----G----------T------------------------------------------------------G-----------------C---- 1261
C.MM.99.mIDU101_3         .............---C-----------A---------AA...GG--CT-A--GA-C------A-----T--C--------G----------T------------------------------A------------------C----------------------C---- 1431
C.MW.93.93MW_965          ..........-A----------------A---------AA...GG--CT----GA-CA-----A-----T--C--------G---------CT------------------------------A--------------------A--------------------C---- 1264
C.SN.90.90SE_364          ..........CA----C----------GAA--------AA...GG--CT----GA--------A--------C--------G----------T-----A---------------------G--A-----------------------------T-----------C---- 1261
C.SO.89.89SM_145          ....AATATA-A----C-------A-------------AA...GG--CT-A--GA--------A--------C-------------------T---T--------------T-----------------------------------------------------C---- 1267
C.TZ.02.CO178             ..........-A----------------A---------AA...GGA-CT-A--GA-C------A--C-----C--------G----------T---------------------------G--A-----------------------------------------C---- 1261
C.UY.01.TRA3011           ..........-A----------------A---------AA...GG--CT-A--GA-C-A----A--C-----C--------G---------CT------G--C--------------------A-----------------------G--G--------------C---- 1252
C.YE.02.02YE511           ..........-AT---------------AA--------AA...GG--CT-AG-GA--------A--------C--------G----------T-----A---C--------------------A-----------------------------------------C---- 1261
C.ZA.04.04ZASK164B1       ..........-A----------------A-------C-AA...GG--CT-----A---A----A--------C--------G----------T-----A------------------A--------T--------------------------------------C---- 1494
C.ZA.05.05ZAPSK240B1      ..........-AT-------------A-A---------AA...GG--CT-A--GA...A----A--------C--------G---------CT-------------------------------------------------C----------------------C---- 1484
C.ZM.02.02ZM108           ..........-A----------------AA--------AA...GG---G----GA--A-----A--------C--------G----------T------------------T-----------------------------------------------------C---- 2061
D.CD.83.ELI               ....GTTA----AGC-----------------------A-...GG--C------A---A---------------------------------T---A-A-----------------------------------G----------------------C-A--------C- 1617
D.CM.01.01CM_0009BBY      ....TCT---G-A-------------A-----------A-...GG--C---G-GA-C---------------C-------------------T---A-A------------T--------------T-----------------------------------------C- 1264
D.KE.01.NKU3006           ....GCTAA-G-T-------------------GG----A-...GG-AC---G--A--CA-------C-----C-------------------T---A-A------------T----------G----------------------------------------------- 1276
D.KR.04.04KBH8            ....AAT--AG---------------------------A-...GG--C------A--A-----------------------------A----T---A-A------------------A----G---------------------------------------------C- 2030
D.TD.99.MN011             ....GCT---G-A--------------G----------A-...----C--AG-GA--------A--------C-------------------T---A-A------------------------------------------------------------------G--C- 1289
D.TZ.01.A280              ....AAT---G-------------------T-------A-...GG--C---G--A--CA-------------C--------------A----T---A-A---------------A--A--------------------------------------------------C- 1266
D.UG.99.99UGK09259        ....AAT---GGTG------------------------A-...GGT-----G-GA--CA-------------C------------------TT---A----------A---T-------------------------------G------------------------C- 1276
D.YE.01.01YE386           .......AA-G-TG-T----------------------A-...GG--C---G--A-CCA-------------C-----------------------A-A------------T----------G---------------------------------------------C- 1270
D.YE.02.02YE516           ....AAT---G-A--------------G----------A-...----C---G--A--------A--------C-------------------T---A-A---C--------------A--------------------------------------------------C- 1276
D.ZA.90.R1                ....GCTA--G-AG------------------------A-...GG---------A---A-------------C-------------------T---A-A--------------------------------------------G---------------------C--C- 1417
F1.AR.02.ARE933           ..........--TG------------A-A---------A-...GG---G-A--G---A-----A-----T-------------T---A----T-----A---------------C--A-------------------------GA--------------------C--C- 1362
F1.BE.93.VI850            ..........G---------------A-A-T-------A-...GG--------GAG-------A-----T-----------------A----T------------------------A-------------------------GA--------------------C--C- 1405
F1.BR.01.01BR125          ..........--TG------------A-A-T--C----A-...GG--------GA-G-A----A-----T-----------------A---CT--------------A---------A-------------------------GA--------------------C--C- 1513
F1.ES.x.P1146             .......A-AG---------------A-A-T--C----A-...GG---G----GA--------A-----T-----------------A----T-----A------------------A-----------C-------------GA--------------------C--C- 1351
F1.FI.93.FIN9363          ..........----------------A-A-T---------...GG--------GA--------A-----T-----------------A----T---------------------C---------------------------C-A-----G--------------C--C- 1397
F2.CM.02.02CM_0016BBY     ..........T--G--C---------A-A----C----A-...GG-----A--GA-----C-----------C--------------A--T-T---T-A------------------A----G------------------------------------------C--C- 1258
F2.CM.95.MP257            ..........G---------------AGA----C----A-...GG--------GA--------------T--C--------------A--T-T---T--------------------A----G------C------------C----------------------C--C- 1277
F2.CM.97.CM53657          ..........T-----T---------A-A----C----A-...GG--------GA-A---------------C--------------A--T-T---T-A------------------A----G------------------------------------------C--C- 1258
G.BE.96.DRCBL             ....GCA--AG---------------A-A-------C-A-...GG--C-----GA-CAA----A--------C--------G-----A--TCT------------------------------------------------------G-----T-----------A--C- 2035
G.CM.01.01CM_4049HAN      .......A-AG---------------A-A---------A-...GG--CG----GA-C-A-------------C--------G-----A--TCT---------------------C--------------------------------G-----T--G--------C--C- 1273
G.CU.99.Cu74              ....GGA--AG-TG--C-------A---A---------A-...GG--C-----GA-A-A----------T--C--------------A---CT------------------------------------------------------G-----T-----------C--C- 1684
G.ES.99.X138              ..........--A--M----------A-A---------A-...GGT-C-----GA-CGA-------------C--------G-----A---CT---T-A---C--------------A-----------------------------G-----T-----------C--C- 1515
G.GH.03.03GH175G          .......A-AT---------------ACA-----------...GG--C--A--GA-A-A-------------C--------G-----A--TCT------------------------------------------------------G-----T--G--------A--C- 2108
G.KE.93.HH8793_12_1       ....GCA--AG---------------A-A----C----A-...GG--CG-A--GA-A-A-C--------------------G-----A---TT------------------------A-----------C-----------------G-----T-----------C--C- 1473
G.NG.01.01NGPL0669        ....GCA--AG---------------A-A---------A-...GG--CG----GA-A-A-C-----------C--------G-----A--TCT------------------------------A-------------------G---G-----T-----------C--C- 1276
G.PT.x.PT2695             .......--A--A-----------A-A-A----C----A-...GGT-C--A--GA--GA-------------C--------G-----A---CT---T--------------------A-----------------------------G-----------------C--C- 2013
G.SE.93.SE6165            ....GCA--AG-----------------A---------A-...GG--C-----GA-C-A-C--A-----------------G-----A--TCT------------------------------------------------------------T-----------C--C- 1474
H.BE.93.VI991             ....AGT--AG---------------A------C----A-...GG--C-----GA-C------A----C---C--------------A----T------------------------A------------------------C-------------G--------C--C- 1463
H.BE.93.VI997             ....AAT--AG---------------A-A----C----A-...GG--C------A--------A--------------------G--A----T--------------------------------------------------G---------T--G-----G--C--C- 1398
H.CF.90.056               ....AATA-AG---------------A------C----A-...GG---------AT-------A--C-----C-----------G--A----T--------------------------------------------------GA--------T--G--------C--C- 1418
J.CD.97.J_97DC_KTB147     .......A-A-A------A-----A-----T--C------...GG---G-AG-GA-G------A--------C-----T-----G--A----T---A-A------------------A--------------A--G-----------------------A---A-C--C- 1269
J.SE.93.SE7887            ..........-A------------A-----T--C------...G----T-A--GA---------------------------C----A----T---A-A---C--------------A-----------------------------------------------C--C- 1388
K.CD.97.EQTB11C           ..........G--G-------------------C----A-...GGT-------GA---A-------C-----C--------------A---CT------------T-----------A-----------------------G--A--------T--G--------C---T 1264
K.CM.96.MP535             ..........GTTG------------A------C----A-...GG---T-----A-----------C-----C-------------------T------------------------A-----------------------G-----------T--G--------C--C- 1264
```

```
                                    Gag-Pol fusion TF protein start
                               Gag p2 end Gag p7 nucleocapsid start
B.FR.83.HXB2                .......GCTACCATAATGATGCAGAGAGGCAATTTTAGG...AACCAAAGAAAGATTGTTAAGTGTTTCAATTGTGGCAAAGAAGGGCACACAGCCAGAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGATTGTA 2068
Gag                         __.__.__A__T__I__M__M__Q__R__G__N__F__R__.__N__Q__R__K__I__V__K__C__F__N__C__G__K__E__G__H__T__A__R__N__C__R__A__P__R__K__K__G__C__W__K__C__G__K__E__G__H__Q__M__K__D__C__
01_AE.CF.90.90CF11697       ..........-AT-----------A-------------A-...GG---G-A--GA...A-------C--T--C--------------A---CT----------------------A----G---T-------------GC----------T-----------C--C- 2017
01_AE.CN.05.FJ051           ..........-AT--------------------------A-...GG---G-A--GA...A-------C-----C------------------CT----------------------A--------T-----------C--G-----------T-----------C--C- 2078
01_AE.CN.06.FJ054           ..........-AT--------------------------A-...GG---G-A--GA...A-------C-----C--------------A---CT-----------------T-----A---G----T-------------G-----------T-----------C--C- 2084
01_AE.HK.x.HK001            ...........--TG--------------------YC-A-...GG---G-A--GA...A-------C-----C------GG-----A---CT----------------------A--------T-----------C--G-----------T-----------C--C- 1439
01_AE.TH.01.01TH_R2184      ..........-AT---T------------T-------A-...GG--GG-A--CA...A----A--C-----C-------G------A---CT---T------------------A----------------------------------T-----------C--C- 1267
01_AE.TH.02.OUR769I         ..........-AT--------------------------A-...GG---G-A--GA...A-------C-----C--------------A---CT----------------------A--------T--C-----------G-----G-----T-----------C---- 1267
01_AE.TH.90.CM240           ..........--T---------------------C-A-...GG---G-A--GA...A-------C-----C-----T-G------A---CT----------------------A---C----T-----------C--G-----------T-----------C--C- 1627
01_AE.US.00.00US_MSC1164    ..........-AT--------------------------A-...GG---G-A--GA...A-------C-----C-------G------A---CT-----------------C--A--------T-------------G-----------T-----------C--C- 1267
02_AG.CM.02.02CM_4082STN    ..........-A---------------------------...GG---G----CA...A-A-----------C--------------A---CT-------------A------C------G------C------------G-----------------------C---- 1258
02_AG.EC.x.ECU41            ..........-A---------------------------...GG--CG----CA...A-A--------------------------G--A---CT-----A-------A-------------G--------------G--A--------------------C---- 1164
02_AG.GH.03.GHNJ196         ..........-ATG-------------------------...GG---G----CA...A-A-----------C--------------A---CT-------------A-------------G---T-------G-----G-----------------------C--C- 2094
02_AG.NG.01.PL0710          ..........-A-G---------------T---------...-G---G----CA...A-A--------------------------A---CT-------------A------------------------------------------------------C--C- 1252
02_AG.NG.x.IBNG             ..........-A-G-------------------------...GG---G----CA...A-A-----------C--------------A---CT-------------A-------------G-----------------G--A--------------------C---- 1593
02_AG.SE.94.SE7812          ..........-AT--------------------------...GG---G---CCA...A-A-----------C--------------A---CT-------------A-------------------C-----------G-------------------------- 1440
02_AG.SN.98.MP1211          ..........-A---------------------------...GG---G----CA...A-A-----------C--------------A---CT-------------A---T----A---G---T-------------G-----------------------C---- 1265
03_AB.RU.97.KAL153_2        ..........-A-------------A-A-T---------...GG--C--A--GA...A-----------C--------------A---CT-----------------------------------------------------C--T-----------C--C- 1294
04_cpx.CY.94.CY032          .........G---------------A-A----A----A-...GG--------GA-C-A------------C--------G-----A--TCT---------------------------------C-----G---------G-----T-----------C--C- 1437
05_DF.BE.x.VI1310           ..........--AGC------------------------A-...GG--C------A---A------------C-----T-------------T-----A-----T-----T--G-A-----------------------G-GA--------------------C--C- 1450
06_cpx.AU.96.BFP90          .........G---------------A-A----------A-...GG--CG-A--GA-G-A-------C-----C--------G-----A--TCT--------------------------A----------------------G-----T----------------- 2090
08_BC.CN.97.97CNGX_6F       ..............---C-----------A----------AA...GG-TCT-A--GA--------A--------C--------G---------T-----A------------------------A-------------------------------------C---- 1260
09_cpx.GH.96.96GH2911       ..........-A----C-A--------------------A-...GGT--G-A--GA...A------------C--------------A--MYT------------------------------------------G-----------------------C--C- 1273
10_CD.TZ.96.96TZ_BF061      .......AA---T--------------------------A-...GG--C--AG--A-GCA------------C--------------------A-A-----------T------G--G-------------------G-------------C-----------CC 1447
11_cpx.GR.x.GR17            ..........-AT--------------------------A-...GG---G-A--GA...A------------C--------------A--TCT----------------------A-------------A--------G--A--------------------C--C- 1373
12_BF.AR.99.ARMA159         .........G-TG--C---------A-A-T-GC----AT...GG--------GA--------A-----------------AC---A----T-----A--------A---T-----A-------------------------A--------------------C--CC 2073
13_cpx.CM.96.1849           ........--AG--------------A-A---------A-...GG--C--A--GA---A------------C-------GG-----A--TCT--------------------------------------------------G-----T-----------C--C- 1452
14_BG.DE.01.9196_01         ..........--A------C----A-A-A----------A-...GGT-C-----GACAGA------------C--------G-----A---CT---T-------------------A---------------------------G-----T-----------C--C- 1588
15_01B.TH.99.99TH_MU2079    ..........-ATG-------------------------A-...GG---G-A--GA...A-------C-----C--------------A---CT-----G-----------------A--------T-------------GG----------T-----------C--C- 1470
16_A2D.KR.97.97KR004        .AATTCAAAC-TA--G----------A-------------...GGG----A---....A-------------C--------G-----A---CT-------------------------------C-----------G-----------T-----------C--C- 1431
17_BF.PE.02.PE02_PCR0155    ..........--TG-----------A-A-T--C----A-...GG--------GA--------A---------------CG------A----T-----A------------------A-------------------------GA--------------------C--C- 1258
18_cpx.CU.99.CU76           ..........-A-----------------A----C-A-...GG---G-A--GA...A--------------------------A----T------------------------------------------------G--A--------------------C--C- 1379
19_cpx.CU.99.CU7            ....CAT------------------A---T-------A-...GG--T------AG-----------------C-------------------T--A-----------A-------A-------------------------------G------------------C- 1297
20_BG.CU.99.Cu103           ....GGA--AT-----C-------A-A-A----------A-...GG--C--A--GA-A-A----------T--C--------------A---CT----------------------------------------------------G-----T--------G--C--C- 1546
21_A2D.KE.91.KNH1254        ....GCT---G-AG-------------------------A-...GG--C------A---A------------C------------------T--A-AG------------------------------------------G------G------------------ 1279
22_01A1.CM.01.01CM_0001BBY..........GGT--------------------------A-...GG---G-A--GA...A------------C--------------A---CT-A---------------------A-G------T--C-----------G-----------------------C--C- 1267
23_BG.CU.03.CB118           ....GGA--AT-------------A---A----C----A-...GG--C--A--GA-A-A---------T--C-----------T--A---CT---------------------------------A-------------------T-----T-----------C--C- 1520
24_BG.CU.03.CB378           ....GGA--AT-------------A-A-A------------...GG--C-----GA-A-A---------T--C--------------A---CT-----------------------------------------------------G-----T-----------C--C- 1520
24_BG.CU.03.CB471           ....GGA--AT-------------A-A-A----------A-...GG-AC--A--GA-A-A---------T--------------A---CT--------C-----------------A---------A----------G-------G--G-----T-----------C--C- 1520
25_cpx.CM.01.101BA          ....GCAA-AG----------------------------A-...GG--C--A--GA-A-A------------C--------G-----A--TCT----------------------A---------------------------G-----T-----------C--C- 1279
27_cpx.CD.97.97CDKTB49      ........--AG--------------A-A---------A-...GG--CG-----A-C-A------------C--------G-----A--TCT-------------------------G------------G-----------G-----T-----------C---- 1282
28_BF.BR.99.BREPM12609      ..........--TG----------------T--C----A-...GG--------A--------A----T----------------A---T---------------------A-------------------------GA--------------------------C- 1434
29_BF.BR.02.BREPM119        .........G-T---C-A-------A-A-T--C----A-...GG--------GAG-------A----T----------------A---T----A------------------A---------------------------GA--G-----------------C--C- 1242
31_BC.BR.02.110PA           ..........-A-----------------A---------AA...GG--CT-A--GACC-A----A--C-----C--------G---------T-----------------------A---------A-----------------A---------------------A---- 1514
32_06A1.EE.01.EE0369        ..........-A-------------A-A-T----------...GGA-C--A--GA...A------------C--------------A---CT----------------------------------------------C------C--T-----------C--C- 1703
33_01B.MY.05.05MYKL007_1    .........GGA---------------------------A-...GG---G-A--GA...A-------C-----C--------------A---CT----------------------A--------T-------------G-----------T----------------- 1447
34_01B.TH.99.OUR2478P       .......AA--TG--G-----------------------A-...GG---G-A--GA...A-------C-----C--------------A---TT----------------------A--------T-------------G-----------T-----------C--C- 1270
35_AD.AF.05.05AF095         .......A-A-A-------------A-------------A-...GG---T-A--GA...A------------C-----------C--A---CT----------------------A------------------------G-GA--------------------C--C- 1264
36_cpx.CM.00.00CMNYU830     ........--AGGT--------A------A----------...GG--------CA...A-A--A---------C--------G-----A---CT---------------------------------------------G-----------------------C--C- 1270
37_cpx.CM.00.00CMNYU926     .......A-A-G-----------A---A-----------A-...GGT-C--A--GA-A-A---A---------C--------G-----A--TCT----------------------A-------------------------------------------------C--C- 1264
38_BF1.UY.05.UY05_4752      ..........-TTG-----------A-AAT--C----A-...GG--------GA-G-A----A-------------------A--TCT----------------------A-------------------------GA--G--------G---------C--C- 1479
39_BF.BR.03.03BRRJ103       .......A---ATG--------------------------...-------------C---------------------------------GT-----------------------------------------------------------------------A--C- 1542
40_BF.BR.04.04BRRJ115       ....AATA-AG-TG-C---------A-A-T--C----A-...GG--------GA--------A----T----------------A---T----------------------A---------------G-------GA-------------------------- 1620
42_BF.LU.03.luBF_05_03      .......A-AG-TG--C----------A-T--C----A-...GG--C------A--------A----T--C--------------A---T----A------------------A-------------------------GA--G-----------------C--C- 1594
43_02G.SA.03.J11223         ....GCA--AG-------------A-------------A-...GG--C-----GA-A-A---------T--C--------G-----A--TCT-----------------------------------------------G-----T-----------C--C- 1596
U.CD.83.83CD003             .......A-AG--------------------C----A-...GG--C-----GA--------A--------C-------------------T-T-----A---------------------------------------G--------------------C--C- 1560
U.CD.90.90CD121E12          .......A-A--TG-----------------C----A-...GG---G-----A--------A--------C-------------------T-T-----A---------------------------------------G--------------------C--C- 1267
U.GR.99.GR303               .......C-A-A-------------A-------------...GG----A--GA...A----------T--C--------------A---CT---------------------------R-------------------G--------G--T-----------C--CC 1345
U.NL.95.U_NL_95_H10986_D1   .......A-AGTTG-------------------C----AA...GG-AC--A--GA---A-------------C--C-----G--G--A---T-------------------------------------------------------------------------C--C- 1953
CPZ.CD.90.ANT               ....GGAA-CG-AG-CT-TC-------------GGA-ATAGAGGAGG--A--GACC-C-C--A-----T--C--C--T-----G--C--T--T--A----------A------A--A-GG--A------C----G-------C-------------GC-T---A-C---C 1498
CPZ.CM.01.SIVcpzCAM13       A.........GA---TT-CT-T----A---GCCAGGGGCAACC...-C-----GA-AAA-A--A--C-AT--------A--G-----C--TCTG--------C--T-A---A--A----G---------C-----------GC----G-------------------C---T 1651
CPZ.CM.05.SIVcpzEK505       .......A---GTG-TT-T-----A-----A-------AA...GG-AT--------CCA-C--A--C--T--------T-----G--C--TCT---A------C--T-A---A------A----------C-----------GC-A---G-----T--------G--C--C- 1616
CPZ.CM.05.SIVcpzLB7         .......A-AGT---T--------A--G--A-------AA...GGA-C--A--GAG--A----A-------------------------T---A--------T-AA-----C----G---A---------------------GC-----------T-GG----C--------- 1614
CPZ.CM.05.SIVcpzMB66        ..................-ATGATGCA-AGA-GC-A----AA...GG--C--A--GA-----C--A--C--T-----------G-----T----T---T--G--------A------C--A-G---A-----------G-----GC-A-----G--T--G----G-A-----C- 1602
CPZ.CM.05.SIVcpzMT145       ..............--CT-T-----A-----GGGACAA-ATAGACCT-CT-AT--A-AGA----A--C--T--------A--------A---CT---A--G--C--------A--CC-A-G-----------C-----G-----G-----T--G--------------- 1607
CPZ.GA.88.GAB1              .GGGAGA--AGATG-TT-CT-C----AG--AC-AGG-GC-...GG--C--A--GA-AAA-A--A--C--T--------G---------T---CT---T---------T-AA--A--A--A-G---A----------CGG-----GC-A--G--------------------C--C- 2135
CPZ.TZ.00.TAN1              ..........GTA-AT---C-----G----A--AAGACCACCCTTAA---AGGGTCAGC-GC--------T--C-----G----T---C--T-----A--------T-----T--A--A---G--A--T--C----GG-----C-A--G-----T--------G--C--C- 1712
CPZ.US.85.CPZUS             .......T-A-G-G--T-TC-A----A---A--CGC-G--...--G-C-G-G-GA-AAA-C--A-----T----------------G--A--TCT--------------------T--A--A-G---A--G--C-----G-----GC------------T-----------C--C- 2116
CPZ.CM.x.SIVcpzDP943        .......A-A-GTG-GT-TT-------G------AA-G--...-GA-C--CT-G--AAA-C--A--C--T------------------------TTT---A-----C-----------A--A-GG-----G--C-----------C-A--G--G--T--G--A----------- 1627
CPZ.CM.x.SIVcpzMB897        ...........--TG-T--------A-----T---G---A-...-TT--G-A--GA-----G--A--C--T--C--------G-----T-----T---A-----C--T--A--------A-G---A---------------------GC-A-------------CA----A-C--CC 1607
N.CM.02.DJO0131             .......A---GTG-CT-TGCA--AG-GA-A--C----AA...GG-AT---G-GA-CCA---GA------------------G--C--TTTG--A-----C--T-A---------A-G-GG------------------GC-A--------T----------------- 1553
N.CM.04.04CM_1015_04        .......A---GTG-CT-TG-A--A-----A--C----AA...GG-AC---G-GA-CCA----A---------C------------G--C--TTTG--A--G--C--T-A---T-----A-G-RG------------G-----GC-A--------T----------------- 1546
N.CM.04.04CM_1131_03        .......A-C-GTG-CT-TG-A--A--G--A--C----AA...GG-AC---G-GA-CCA----A------------------------G--C--TTTG--A--G--C--T-A-----------G-GGA-------------------GC-----------T----------------- 1545
N.CM.95.YBF30               ....ACAA---GTG-CT-TGCA--A--G--A--C----AA...GG-AT---G--ACCCA----A------------------------G--C--TTTG--A-----C--T-A----------A-G-GGA----------G-----GC-A--------T----------------- 1635
N.CM.97.YBF106              .......A---GTG-CT-TG-A--A--G--A--C----A-...GG-AT-------A-CCA----A-----T-------------G--C--TTTG--A-----C--T-A----------A-G--G------------G-----GC-A--------T----------------- 1633
O.BE.87.ANT70               AGGATACA-AG-AG--T-C-----A-----GC-AAA-CCA...-TTAGG-A-GGA-C-A-A--A---------C-----A-----G--A--T-T---A--------TC-A--A-----------T--C------------C-------T----------------C- 2131
O.CM.91.MVP5180             AGGATACA-AG-AG--T-C-----A-----GC-GAA-CCA...--TAG--A-GG-CCCA-A--A--C-----------A-----G--A--T-T---A-A---C--TC-A--A-----A----G---T--C------------C-------T----------------C- 2106
O.CM.96.96CMABB637          AGGCTACA-AG-AG--T-T-----A-----AC-AAACCCA...-CTAGGCAG-G-CC-A-A--A--------------A-----G--A--TCT---A---------T--A--A-----A--G--A--T--C------------C-------T----------------C- 1555
O.CM.98.98CMA104            AGGATACA-AG-AG--T-C-----A-----GC-AAA-CCA...-TTAGG---GGACC---A-----------C-----G-----G--A--T-T---A---------TC-A--A-----A--G--A-----C-----C-----GC----T--T-----G-------------C- 1563
O.CM.99.99CMU4122           AGGATACA-AG-AG--T-T-----A-----GC-AAA-CCA...GGAAGG-A-GGACC-A-A--A--------------A-----G--A--TCT---A---------TC-A--A-------G--A--T--C------------C-------T----------------C- 1562
O.FR.92.VAU                 AGGATACA-AT-AG--T-C-----A-----GC-AAAACCA...-TTAG--A-GGACC-A-A--A--------------A-----G--A--TGT-A-A-A------T--A--A-----A--G--A--T--C------------C-------T--T-------------C- 1640
O.SN.99.SEMP1299            AGGATACA-AG-AG--T-T-----A-----GC-AAACCCA...-GTAGG-A-GG-CC-A-A--A--------------A-----G--A--TCTG--A---------TC-A--A-----A--G--A--T--C------------C-------T----------------C- 2130
O.US.99.99USTWLA            AGGATAYA-AG-AG--T-C-----A-----GC-AAACCTA...GGTAGG-A-GGACC---A--A--------------A-----G--A--TGT---A---------TC-A--A-----A--G--A--T--C------------C-------T-----------G-----C- 1551
```

```
                                                  Gag-Pol -1 ribosomal slip site
                                         Gag p7 nucleocapsid end Gag p1 start                              Gag p1 end Gag p6 start
B.FR.83.HXB2                  CT............GAGAGACAGGCTAA.TTTTTTAGGGAAGATCTGGCCTTCCTACAA............GGGAAGGCCAGGGAATTTTCTTCAGAGCAGACC.........AGAGCCAACAGC.............................................CCC 2162
Gag                           T__.__.__.__.__E__R__Q__A__N#_F__L__G__K__I__W__P__S__Y__K__.__.__.__.__G__R__P__G__N__F__L__Q__S__R__P__.__.__.__E__P__T__A__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__P
Pol                                                         _F__F__R__E__D__L__A__F__L__Q__.__.__.__.__G__K__A__R__E__F__S__S__E__Q__T__.__.__.__R__A__N__S__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__P_
A1.GE.99.99GEMZ011            --............--------------.-----------GA--T---------AG---............A--G--------A-------C------------.........-------T----.............................................--- 1346
A1.KE.00.KER2008              T-............--------------.------------A--T---------AG---............A--G----------------C------------.........G-----------.............................................--- 1361
A1.KE.00.KNH1144              --............--A-----------.------------A--T---------AG---............---G--------A-------C------------.........G--A--------.............................................--- 1367
A1.KE.00.KSM4024              --............--A-----------.------------A--T---------AG---............---G--------A-------C-----A-----T.........G--A--------.............................................T-- 1358
A1.KE.00.MSA4069              --............--A-----------.------------A--T---------AG---............---G--------A-------C--T--A---C-T.........G--A--------.............................................--- 1361
A1.KE.00.NKU3005              --............--------------.------------A--T---------AG---............---G--------A-------C------------.........G-----------.............................................--- 1361
A1.RU.00.RU00051              --............--------------.------------A--T---------AG---............R--G--------R-------C------------.........------------.............................................--- 1463
A1.RW.93.93RW_024             --............--A-----A-----.-----------GA--T---------AG---............---G--------A-------C------------.........------------.............................................--- 1364
A1.SE.95.SE8891               --............--A-----------.------------A--T---------A---G............---G--------A-------C------------.........G-----------.............................................--- 1361
A1.SE.95.UGSE8131             --............--A-----------.------------A--T---------C----............---G-----G--A-------C------------.........G------T----.............................................--- 1552
A1.TZ.01.A173                 --............--------------.------------A--T---------AG---............A--G--------A-------CC-----------.........G-----------.............................................--- 1361
A1.UA.01.01UADN139            --............--------------.-----------GA--T---------AG---............A--G--------A-------C------------.........-------T----.............................................--- 1361
A1.UG.92.92UG037              --............--------------.------------A------------AG---............A--G--------A-------C------------.........---A--------..........................................CCCA-- 2178
A1.UG.99.99UGA07072           --............--A-----------.------------A--T---------AG---............---G--------A-------C------------.........G-----------.............................................--- 1361
A2.CD.97.97CDKTB48            --............--------------.------------A--T---------A----............A--G----------------C---------GA-.........------------.............................................--- 1521
A2.CY.94.94CY017_41           --............--------------.------------A--T---------A----............A--G--------A-------C----------A-.........------------.............................................--- 1519
A.SN.01.DDI579                --............--------------.------------A------------A----............---G--------A-------C------------.........------------.............................................--- 1363
A.SN.01.DDJ369                --............--A-----------.------------A--T---------AG---............---G----------------C------------.........------------.............................................--- 1366
A.SN.96.DDJ360                --............--------------.------------A--T---------AG---............A--G-----------C----C------------.........------------.............................................--- 1363
A.CD.02.02CD_KTB035           -C.........AAT--------------.------------AG-T---------A---G............A--------------C----CC----AA----T.........G-----------.............................................--- 2156
A.CD.97.97CD_KCC2             --............--A-----------.------------A--T---------A----............A--G-----------C----CC----AAG---T.........------------.............................................--- 2158
A.CD.97.97CD_KTB13            --............--------------.------------AG-T---------A----............A--------------C----CC--A-A-----T.........------------.............................................--- 1366
B.AR.04.04AR151516            --............--A-----------.-----------GAC-----------C----............---G--------A-----C--------------.........------------CCCAGAGCC..............................AACAGC--- 1384
B.AU.87.MBC925                --............--------------.-------------------------C----............---------------------------------.........------------.............................................--- 2167
B.BO.99.BOL0122               --............--A-----------.------------A------------C----............A--G-----------------------------.........------------.............................................--- 1379
B.BR.03.BREPM2012             --.........TTT--------------.-------------------------C----............---G--------------C-------A---G--.........------C-G-C-AGAGCC.................................AACAGC--- 1587
B.CA.97.CANB3FULL             --............--------------.-----------GA------------C----............---------------------------------.........------------CCCACC....................................AGT--A 1451
B.CN.05.05CNHB_hp3            --............--------------.------------A------------C----............---G-----------------------------.........-------T----.............................................--- 2161
B.CO.01.PCM001                TC............--------------.------------A------------A----............---G----------------C-----A-----T.........------------.............................................--- 1367
B.GB.83.CAM1                  A-............--------------.------------A------------C----............---G-----------------------------.........------------.............................................--- 2163
B.GB.86.GB8                   --............--------------.-------------------------C----............---------------------C-----------.........--------T---CCCACCAGA...........................GCCGACAGC--- 2182
B.GE.03.03GEMZ010             --............--------------.------------A------------C----............---------------------------------.........------------.............................................--- 1367
B.IT.05.SG1                   --............--------------.--------------------------C---............------------N-----------------G--.........------------.............................................--- 1964
B.JP.05.DR6538                --............--------------.-------------------------C----............A-----------------C--------------.........------------.............................................--- 2166
B.KR.05.05CSR3                --............--A-----------.------------A------------C----............---G-----------C--C-------------T.........------------.............................................--- 2187
B.NL.00.671_00T36             --............N------------N.---------------T---------C----............---G-----------------------------.........------------CCCTTCTCAGAGCAG...............ACCAGAGCCAACAGC--- 1727
B.RU.04.04RU128005            --............A-------------.------RR-----------------C----............------------------C--------------.........------------.............................................--- 1653
B.TH.00.00TH_C3198            --............--------------.------------A---------G--C----............---G-----------------------------.........------------.............................................--- 1376
B.UA.01.01UAKV167             --............--------------.------------A------------C----............------------------C--------------.........--G---------CCCACGAACAGA........................GCCAACAGC--- 1388
B.US.04.ES10_53               --............--A-----------.------------A--T---------C----............---------G--------C--C----A----A-.........-------T----.............................................--- 2171
B.US.99.PRB959_03             --............--------------.------------A------------C----............------------------C--------------.........------------.............................................--- 1382
C.AR.01.ARG4006               --............--------------.------------A--T---------A----............------------A-----Y--C-----------.........------------.............................................--- 1358
C.BR.04.04BR013               --............--A-------T---.------------A------------C----.........GGG---G--------------C--A-----------.........------------CCCACCACTAGA........................GCCAACAGC--- 1638
C.BW.00.00BW07621             --............--A--G--------.------------A--T---------C----............---G--------------C--C-----------.........------------.............................................--- 1527
C.CN.98.YNRL9840              -G............--A--G--------.------------A--T---------C----............---G--------------C--C--------G--.........------------.............................................--- 1358
C.ET.02.02ET_288              --............--------------.------------AC-T---------C----............---G--------------C-------A------.........--C---------CCCACTAGA...........................GCCAACGGC--- 1376
C.GE.03.03GEMZ033             --............--------------.------------A--T---------C----............---G-----------C--C--C--------G--.........------G-----.............................................--- 1355
C.IL.99.99ET7                 --............--A--G--------.------------AT-T---------C----............---G--------------C--------------.........------------.............................................--- 1355
C.IN.99.01IN565_10            --............-----G--------.------------A--T---------C----............------------------C-------A------.........------------.............................................--- 1555
C.KE.00.KER2010               --............--A--G--------.------------A--T---------C----............---G--------------C-------A---G--.........------------.............................................--- 1355
C.MM.99.mIDU101_3             -A............--C--G--------.--------A---AC-T---------C----............---G--------------C--C-----------.........------------.............................................--- 1525
C.MW.93.93MW_965              --............--A--G--------.------------A--T--------...---............---G--------------C--------------.........------------.............................................--- 1355
C.SN.90.90SE_364              --............-----G--------.------------A--T------G--C----............---C--------------C--------------.........------------.............................................--- 1355
C.SO.89.89SM_145              --............-----G--------.------------A--T---------C----............---G--------------C-------A---G--.........------------CCCACCAGCAGAGAGCTTCAGGGTCAGGCCAGAGCCAACAGC--- 1403
C.TZ.02.CO178                 -A............-----G--------.------------A------------C----............---G--------------C-----------G--.........------------.............................................--- 1355
C.UY.01.TRA3011               --............-----G--------.---------------T---------C----............---G--------A-----C-CC----------T.........------------.............................................--- 1346
C.YE.02.02YE511               A-............-----G--A-----.------------A--T---------C----............------------------C--------------.........------------.............................................--- 1355
C.ZA.04.04ZASK164B1           --............--A-----------.------------A------------C----............---G-----------C--C--------------.........------TT----.............................................T-- 1588
C.ZA.05.05ZAPSK240B1          --............--------------.------------A--T---------C----............---G-----------C-----C----A---G--.........------------.............................................--- 1578
C.ZM.02.02ZM108               --.........AAT--------------.------------A--T---------C----............---G--------------C--C-----------.........------------.............................................--- 2158
D.CD.83.ELI                   --............--------------.-----------GA--T---------C----............---------G-----C-----C--A--------.........------------.............................................--- 1711
D.CM.01.01CM_0009BBY          --............--A-----------.------------A--T---------C----............---------G-----C-----------------.........------------.............................................--- 1358
D.KE.01.NKU3006               --............--A-----------.------------A--T---------C----............---------G-----C--C-------A------.........------------.............................................--- 1370
D.KR.04.04KBH8                --............--A-----------.-------------------------C----............T-----------A--C--C-C----------A-.........---------C--.............................................--- 2124
D.TD.99.MN011                 --............--A-----------.------------A--T---------AG---............---------G--A--------------------.........------------.............................................--- 1383
D.TZ.01.A280                  --............--A-----------.------------A--T---------C----............A--------------C--C--------------.........------------.............................................--- 1360
D.UG.99.99UGK09259            --............--A-----------.------------A--T---------C----............---G--------------C--------------.........------------.............................................--- 1370
D.YE.01.01YE386               --............--A-----------.------------A--T---------C----............---------------C--C--------------.........------------.............................................--- 1364
D.YE.02.02YE516               --............--------------.-----------GA------------AG---............---G-----G--A--C-----------------.........---------T--.............................................--- 1370
D.ZA.90.R1                    --............--A-----------.------------A--T---------C----............---------G-----------------------.........------------.............................................--- 1511
F1.AR.02.ARE933               --............--A-----------.------------A--T---------A----............---G-----C--A-----C--------------.........------------.............................................--- 1456
F1.BE.93.VI850                --............--A-----------.------------A--T---------A----............---G-----C--A-----C--------------.........------------.............................................--- 1499
F1.BR.01.01BR125              --............--A-----------.------------A--T---------C----............---G-----C--------C-------A---G--.........---A--------.............................................--- 1607
F1.ES.x.P1146                 --............--A-----------.------------A--T---------C-G--............---G-----G--A-----C-------A------.........------T-----.............................................TG- 1445
F1.FI.93.FIN9363              --............--A-----------.------------A--T---------A----............---G-----C--A-----C--------T-----.........------------.............................................--- 1491
F2.CM.02.02CM_0016BBY         --............--A-----------.------------A--T---------A----............A--G-----T--------CA-------------.........---------G--.............................................--- 1352
F2.CM.95.MP257                --............--------------.------------A--G---------A----............---G-----C--A-------------A------.........------------.............................................--- 1371
F2.CM.97.CM53657              --............--------------.------------A--T---------A----............---G-----T--A-------------A------.........------------.............................................--- 1352
G.BE.96.DRCBL                 -A............--A-----------.------------A--T---------A----............---G--------------C-------A---G--.........------------.............................................--- 2129
G.CM.01.01CM_4049HAN          -A............--------------.-----------GA------------A----............A--G-----------C-----C----A---C--.........------------.............................................--- 1367
G.CU.99.Cu74                  -A............--------------.------------AC-T---------A----............A--G-----G--------------A-A---G--.........-------T----.............................................--- 1778
G.ES.99.X138                  -A............--------------.------------A--T---------A----............---G-----------------C----A---G--.........G-----------.............................................--- 1609
G.GH.03.03GH175G              -A............--A--------C--.------------A------------A----............---G----------------------A---G--.........-------T----.............................................T-- 2202
G.KE.93.HH8793_12_1           -G............--A-----------.---.--------A--T---------A----............---G----------------------A---G--.........------------.............................................--- 1566
G.NG.01.01NGPL0669            -A............--------------.------------A--T---------A----............---G--------A-------------A---G-T.........------G-----.............................................--- 1370
G.PT.x.PT2695                 -A............--------------.------------A--T---------A----............---G-----------------C----A---G--.........C-----------.............................................--- 2107
G.SE.93.SE6165                -G............--------------.------------A--T---------A----............---G----------------------A---GA-.........---A--------.............................................--- 1568
H.BE.93.VI991                 -A............-GA-----------.------------A--T---------AG---............------------------C-CC----AG--G-T.........------------.............................................--- 1557
H.BE.93.VI997                 -A............--------------.------------A------------AG---............A--G--------A--------C--------G--.........------------.............................................--- 1492
H.CF.90.056                   -A............--------------.------------A--T---------AG---............A--G--------A--------C--------G--.........---A--------.............................................--- 1512
J.CD.97.J_97DC_KTB147         --............-----------C--.------------AG-T---------AG---............---G-----------C-----C--------G--.........---A--------.............................................--- 1363
J.SE.93.SE7887                --............--------------.---------------T---------AG---............A--G-----------C-----C-----------.........------------.............................................--- 1482
K.CD.97.EQTB11C               --............--------------.------------T---------CT-A----............A-AG--------A-------------A------.........------------.............................................--- 1358
K.CM.96.MP535                 --............--------------.------------A------------C----............---G-----------------------------.........------------.............................................--- 1358
```

```
                                                 Gag-Pol -1 ribosomal slip site
                                          Gag p7 nucleocapsid end Gag p1 start                               Gag p1 end Gag p6 start
B.FR.83.HXB2                 CT............GAGAGACAGGCTAA.TTTTTTAGGGAAGATCTGGCCTTCCTACAA............GGGAAGGCCAGGGAATTTTCTTCAGAGCAGACC.........AGAGCCAACAGC..........................................CCC 2162
Gag                          T__.__.__.__.__E__R__Q__A__N#_F__L__G__K__I__W__P__S__Y__K__.__.__.__.__G__R__P__G__N__F__L__Q__S__R__P__.__.__.__E__P__T__A__.__.__.__.__.__.__.__.__.__.__.__.__.__.__P
Pol                                                       _F__F__R__E__D__L__A__F__L__Q__.__.__.__.__G__K__A__R__E__F__S__S__E__Q__T__.__.__.__R__A__N__S__.__.__.__.__.__.__.__.__.__.__.__.__.__P_
01_AE.CF.90.90CF11697        --............--------------.------------A--T------CT-A----............---G----------------C-----------T.........------------..........................................--- 2111
01_AE.CN.05.FJ051            --............--------------.------------A--T---------A----............---------G--A-------C---------------------------------..........................................T-- 2172
01_AE.CN.06.FJ054            --.........AAT--------------.------------A--T---------A----............---------G----------C---A--------.........------------..........................................--- 2181
01_AE.HK.x.HK001             --............--------------.------------A--T------C--A----............---G----------------C---------------------------TT----..........................................--- 1533
01_AE.TH.01.01TH_R2184       --............--------------.-------------C-T---------A----............---------G--A-------C-----A------.........------------..........................................--- 1361
01_AE.TH.02.OUR769I          --............--------------.------------A------------A----............---------G----------C---------------------------A--------..........................................--- 1361
01_AE.TH.90.CM240            --............--------------.------------A--T---------A----............---------G----------C---------------------------------..........................................--- 1721
01_AE.US.00.00US_MSC1164     --......GAGAGAC--G----------.------------A--T---------C----............---------G--A-------C---------------------------------..........................................--- 1367
02_AG.CM.02.02CM_4082STN     --............--------------.---------A--A--T---------A----............---G--------A-------C---------------------------A--------..........................................--- 1352
02_AG.EC.x.ECU41             --............--------------.------------A--T---------AG---............---G--------A-------CC-----------.........G--A--------..........................................T-- 1258
02_AG.GH.03.GHNJ196          --............--------------.------------A--T---------A----............---G--------A-------C------------.........G--A---T----..........................................--- 2188
02_AG.NG.01.PL0710           --............--------------.------------A--T---------AG---............------------A-------C------------.........G-----------..........................................--- 1346
02_AG.NG.x.IBNG              --............--------------.------------A--T---------AG---............---G--------A-------C------------.........G--A--------..........................................T-- 1687
02_AG.SE.94.SE7812           --............--------------.------------A--T---------AG---............---G--------A-------C------------.........G--A--------..........................................--- 1534
02_AG.SN.98.MP1211           --............--A-----------.------------A--T---------AG---............---G--------A--C----C------------.........G--A--------..........................................--- 1359
03_AB.RU.97.KAL153_2         --............--------------.-----------GA--T---------AG---............A--G--------A-------C---------------------------T----..........................................--- 1388
04_cpx.CY.94.CY032           --............--------------.-----------GA--G---------AG---............A--G--------A-------------A---G-----------------------..........................................--- 1531
05_DF.BE.x.VI1310            --............-----G--------.------------AG-T---------C----............---G-----------C--C--C-----------.........G-----CT----..........................................--- 1544
06_cpx.AU.96.BFP90           -A............--------------.------------A------------A----............---G----------------------A---G-----------------------..........................................--- 2184
08_BC.CN.97.97CNGX_6F        --............--A--G-----G--.------------AT-T---------C----............---G--------------C--C--------------------------------..........................................--- 1354
09_cpx.GH.96.96GH2911        --............--------------.------------A--T---------A----............A--G-----------C----CY-----------.........G------T-G--..........................................--- 1367
10_CD.TZ.96.96TZ_BF061       -A.........GAAAGA-----------.------------A--T---------A----............A--------------C--C-----------------------------------..........................................T-- 1544
11_cpx.GR.x.GR17             -C............--------------.------------A--T---------AG---............---G--------------C--A--------G-----------------------..........................................--- 1467
12_BF.AR.99.ARMA159          --............--A-----------.------------A--T---------A----............---G-----C--A-----C-------A---G-----------------------..........................................--- 2167
13_cpx.CM.96.1849            TG............--A-----------.------------A------------A----............A--G-----------------------------.........--GA--------CCCACCAGAGAGCAG............ACCAGAGCCAACAGC--- 1576
14_BG.DE.01.9196_01          -A............--------------.------------A--T---------A----............---G-----------------C----A---G-----------------------..........................................--- 1682
15_01B.TH.99.99TH_MU2079     --............--------------.-----------GAC-T---------A-...............--------------------C---------------------------T----..........................................--- 1561
16_A2D.KR.97.97KR004         --............--------------.------------A--T---------C---G............C--G----------------C----------A-.........---A--------..........................................--- 1525
17_BF.PE.02.PE02_PCR0155     --............--A-----------.------------AG-T---------A---G............---G-----C--A-----C-------A---GA-.........------------..........................................--- 1352
18_cpx.CU.99.CU76            --............--A-----------.------------A--T---------A----............---G-----------------C----AA--G-----------------------..........................................--- 1473
19_cpx.CU.99.CU7             --............--A-----------.------------A--T---------C----............---G--------A--C--C-----------------------------------..........................................--- 1391
20_BG.CU.99.Cu103            -A............--------------.------------AC-T---------C----............---------------------C--------G-----------------------..........................................--- 1640
21_A2D.KE.91.KNH1254         -G............--------------.-------------G-T---------C----............---G-----G-----C--------------G-----------------------..........................................--- 1373
22_01A1.CM.01.01CM_0001BBY   --CTT...ACTCTT--------------.------------A--T---------AG---............---G-----------C----C----G-------.........C-----------..........................................--- 1370
23_BG.CU.03.CB118            -A.........GAA-G------------.------------AC-T---------C----............---------------C----C-----A---G-T.........------------..........................................--- 1617
24_BG.CU.03.CB378            -A............--------------.------------AC-T---------C----............---------G--A--C----C-----A---G-T.........------------..........................................--- 1614
24_BG.CU.03.CB471            -A............--------------.------------AT-T---------C----............---G----------------C-----A---G-T.........------G-----..........................................--- 1614
25_cpx.CM.01.101BA           -A............--------------.---------------T---------AG---............A--G--------A--C----------A---G--.........------G-----..........................................--- 1373
27_cpx.CD.97.97CDKTB49       -A............--------------.------------A--T---------A----............---G-----------------C--------G--.........-------T----..........................................--- 1376
28_BF.BR.99.BREPM12609       --............--A-----------.------------A--T---------C----............---G-----C--A-----C-------A---G-----------------------..........................................--- 1528
29_BF.BR.02.BREPM119         --............--A-----------.------------A--T---------C----............---G-----C--A-----C-------A---G-----------------------..........................................--- 1336
31_BC.BR.02.110PA            --............-----G-----C--.------------A--T---------C----............---G--------R-----C--C--------------------------------..........................................--- 1608
32_06A1.EE.01.EE0369         --............--------------.------------A--T---------AG---............---G-----------C----------A---G-----------------------..........................................--- 1797
33_01B.MY.05.05MYKL007_1     --............--------------.-------------C-----------C----............---------G-----C--C-----A---C----.........------------..........................................--- 1541
34_01B.TH.99.OUR2478P        --............----A----AT---.------------AT-T---------A----............---------G--------C--C-----------.........------------..........................................--- 1364
35_AD.AF.05.05AF095          --............--------------.------------A--T---------C----............---G--------A-----C-----------------------------------..........................................--- 1358
36_cpx.CM.00.00CMNYU830      --.........AAT--A-----------.------------A--T---------AG---............---G----------------C---------------------------A--------..........................................--- 1367
37_cpx.CM.00.00CMNYU926      --............--------------.-----------GA--T---------AG---............A--G--------A-------C------------.........G-----------..........................................--- 1358
38_BF1.UY.05.UY05_4752       -A............--A-----A-----.------------A--T---------A----............---G--------A-----C-------A---G-----------------------..........................................--- 1573
39_BF.BR.03.03BRRJ103        -A............--------------.------------A------------C----............---G----------------------C---------------------------CCCACC.................................AGC--- 1645
40_BF.BR.04.04BRRJ115        --............--------------.------------A------------C----............---G--------------C-------A---G-T.........------G-----..........................................--- 1714
42_BF.LU.03.luBF_05_03       --............--A-----------.------------A-----------------............---G-----C--A-----C--C----A---G--.........--------GGC-AGAGCC..............................AACAGC--- 1700
43_02G.SA.03.J11223          -A............--A-----------.------------A--T---------C----............---G-----------C--C-----------G-----------------------..........................................--- 1690
U.CD.83.83CD003              --............--------------.------------A--T---------A----............---G----------------------A---------------------------..........................................--- 1654
U.CD.90.90CD121E12           --............--------------.------------A------------A----............--------------------------A----A-.........------------..........................................T-- 1361
U.GR.99.GR303                --............--------------.-----------G---T---------AG---............--------------------C---------------------------G-----..........................................--- 1439
U.NL.95.U_NL_95_H10986_D1    --............--------------.-----------GAG-T---------A----............A--G--------------C-----------------------------------..........................................--- 2047
CPZ.CD.90.ANT                -AGCA...ACAAATACAG--A-A-TA--.------------ACCGACC--CA-G-GGTG.........GGG-T-C--A--------C---G-G----AGGAGGA...AGTAGTG--------.........................................AGCT-- 1610
CPZ.CM.01.SIVcpzCAM13        -A............-G------A-TA--.-------------GG----------CG--G............C--G--------C--C---G-A----A---GA-.........------------..........................................--- 1745
CPZ.CM.05.SIVcpzEK505        GA......AGTGGA--A--G-----G--.-------------G--------CT--C---............A--G--------A-----C-CC----CA-C-A-...AAGGAA------------..........................................--- 1722
CPZ.CM.05.SIVcpzLB7          AC............AT-G-GA---T---.------------AC------------C-C-............------------A--C----C------------.........------C-----..........................................--- 1708
CPZ.CM.05.SIVcpzMB66         -A.........AAT--A-----------.------------AC-----------AG---............------------------C-------AG-----.........---A--------..........................................T-- 1699
CPZ.CM.05.SIVcpzMT145        A-......CAAAATA-TG--AG-AT---.---------A--A-AT------C---GGGG............---G-----------C-A-G-G--AG----GAA.........G--A--------..........................................T-- 1707
CPZ.GA.88.GAB1               -A............-GA-------TG--.------------AGGT---------CG--G............C--G--------C--C---G-G----A----A-.........---A-----G--..........................................--- 2229
CPZ.TZ.00.TAN1               -CACCAGAAACAACAGC-CTGG--TA--.------------ACG-ACC--C-TG-GGGG............-T-C--A--------C---G-G----A--C......................................................................--- 1803
CPZ.US.85.CPZUS              -AGCA......GGAA-C-----------.-------------ACAT--T-CC-A-CGTGGAGCGGGGGATCAAAG--A--------C--C--GG-A-A---GAA.........G-----C-----..........................................T-- 2228
CPZ.CM.x.SIVcpzDP943         -AGCT......GGA--C-----------.------------AGGT---T--C-G-GG-G...CGGGGGATCAAAG--A--------C--C--G---G----GAA.........G--A--C-----..........................................T-- 1736
CPZ.CM.x.SIVcpzMB897         -C.........GATAT-CA--G---G--.------------A---------C--CG---............A--G--------------C--C----AG-----...AGAACC-CCAGA-C--AG.......................................TGC--- 1713
N.CM.02.DJO0131              AA......AATGAA-GA-----------.---------A---GG----T--C---T---............A--G--A-----C--C--C-CC----CA-C-A-AACAAGGAA------C-----..........................................--- 1662
N.CM.04.04CM_1015_04         AA......AATGAA-GA-----------.---------A---GG----T--C---T---............A--G--A--------C--C-CC----CAGC-A-...AAGGAA------C-----..........................................--- 1652
N.CM.04.04CM_1131_03         AA......AATGAA-GA-----------.-------------GG----T--C---T---............A--G--A--------C----CC----CA-C-A-...AAGGAA------C-----..........................................--- 1651
N.CM.95.YBF30                AA......AATGAA-GA-----------.--------------G----T--C---T---............A--G--A-----A--C--C-CC----CA-C-A-...AAGGAA------C-----..........................................--- 1741
N.CM.97.YBF106               AA......AATGAG-GA--M--------.---------A---GG----T--C---T---............A--G--A-----A--C--C-CC----CA-C-A-...AAGGAG------C-----..........................................--- 1739
O.BE.87.ANT70                GA.........AAT-GA-A------A--.---------C--ATA-------C-GGGGGG............CACG--------C----A-G-G-----ACC-G-.........-C-C---T----..........................................--- 2228
O.CM.91.MVP5180              AA.........AAT-GA--------A--.-------------TA-------C-GGGGGG............CACG--------C----A-G-G----AACA-GT.........GTCC---T----..........................................--- 2203
O.CM.96.96CMABB637           GA.........AAT-GA-A------A--.---------C-G-TA-------C-GGGAGG............CACG--------C--C-A-G-A----AGCA-GT.........GCCC---T----..........................................--- 1652
O.CM.98.98CMA104             GA.........AAT-GA-A------A--.---------C--ATA-------C-GGGGGG............CACG--------C------A-G----AACA-GT.........G-CC---T----..........................................--- 1660
O.CM.99.99CMU4122            GA.........AAT-GA-A------A--.------------ATA-------C-GGGGGG............CACG--------C----A-GCA-----ACA-GT.........GTCC---T----..........................................T-- 1659
O.FR.92.VAU                  AAAAT.........AG-------A-A--.--------AA--ATAT------C-GAGAGG............CACG--------C----A-G-G---GAACA-GT.........GTCC---T----..........................................--- 1737
O.SN.99.SEMP1299             AA.........AAT-GA--------A--.------------ATA-------C-GGGGGG............CACG--------C----A-GCG-----ACA-GT.........GTCC---T----..........................................--- 2227
O.US.99.99USTWLA             GG.........AAT-GA--------A--.-------------TA-------C-GGGGGG............C-AG--------C------A-G----AGCA-ATGCCCCCTCC--CC---T----..........................................--- 1657
```

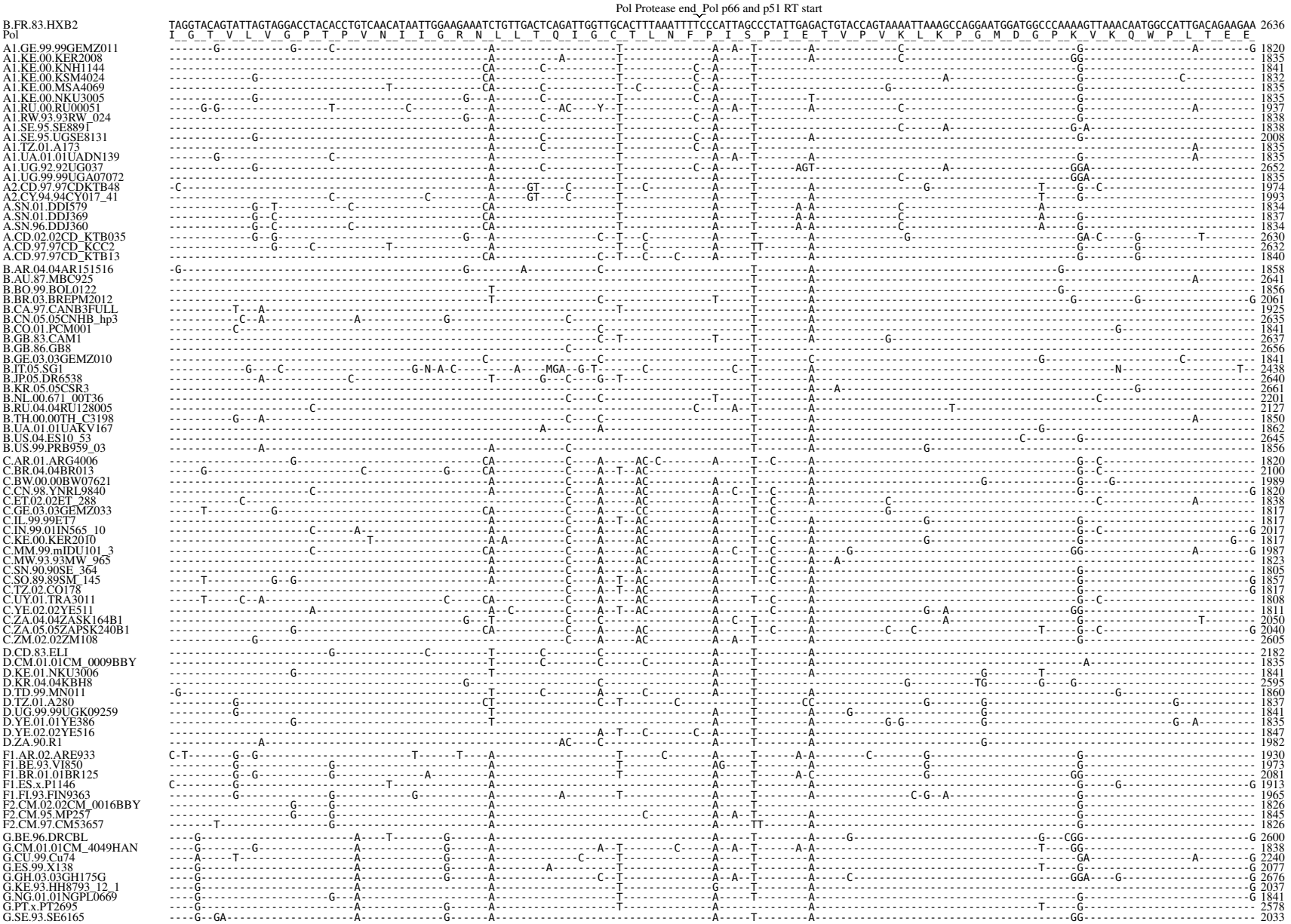


HIV-1/SIVcpz
Complete Genomes

```
                                                
```

```
B.FR.83.HXB2            GGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAA...ATGAGTTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGTGGACATAAAGCTA 2466
Pol                     __G__Q__L__K__E__A__L__L__D__T__G__A__D__D__T__V__L__E__E__.__M__S__L__P__G__R__W__K__P__K__M__I__G__G__I__G__G__F__I__K__V__R__Q__Y__D__Q__I__L__I__E__I__C__G__H__K__A__
A1.GE.99.99GEMZ011      A--A--G-----A--------------------------------------------C...--AGA-----------A------------------------------------------G---------------G-------T--------T------A-A--G---- 1650
A1.KE.00.KER2008        A--AG-GA-G-GA--------------------------------------------T...--A-A-----------A------------------------------------------G----A---A---------G--TCT--------T------A-A--G---- 1665
A1.KE.00.KNH1144        ---A--G----GA--------------------G-----------------------C...--AGA---------T-A------G-----------------------------C-----G----A---------G---G----T--------T------A-A--G---- 1671
A1.KE.00.KSM4024        ---A--G--------------------------------------------------C...--A-A-----------A------------------------------------C----------A------------------T--------T------A-A--G---- 1662
A1.KE.00.MSA4069        ------G-----A-----C--------C-----G-----------------------C...--AGA---A-------A---------------------------C--G-----------G-----------------------T--------T------A-A--G---- 1665
A1.KE.00.NKU3005        ---A--G--------------------------------------------------C...--A-A-----------A------------------------------------C-----G-----------------------T--------T------A-A--G---- 1665
A1.RU.00.RU00051        A--A--G-----A------T-------------------------------------C...--A-A-----------A------------------------------------------G-----------------------T--------T------A-A--G---- 1767
A1.RW.93.93RW_024       ---A--G-----A-----------------------------C-----T--------C...--A-A-----------A------------------------------G-----C-----------------------------T--------T------A-A--G---- 1668
A1.SE.95.SE8891         ---AATG-A---A--------------------------------------------T...--A-A-----------A-----------------------------------------------AG-----------------T--------T-----GA-A--G---- 1668
A1.SE.95.UGSE8131       ---A--G-----A------------------------------------C-------C...--A-A-----------A----------------------------------------------------------------T-A--------T------A-A--G---- 1838
A1.TZ.01.A173           A--A--G-----A-----C--------------------------------------C...--A-A-----------A---------------------------------G--------G----A-----------------CT--------T-----GA-A-GG---- 1665
A1.UA.01.01UADN139      A--A--G-----A---------C----------------------------------C...--AGA-----------A------------------------------------------G-----------------------T--------T------A-A--G---- 1665
A1.UG.92.92UG037        ---A--G-----A--------------------------------------------C...--A-A-----------A------------------------------------C-----G----A------------------T--------T------A-A--G---- 2482
A1.UG.99.99UGA07072     A-----GT----A------T----G--------------------------------T...--A-A---A-------A-----------G---------------------C--------G----A------------------T--------T------A-A--G---- 1665
A2.CD.97.97CDKTB48      A--A--G----GA------T----------------------------G--------C...--A-A-C-----T-G-A---------------------------------------------------A------------G-T--------T------A-A-GG---- 1804
A2.CY.94.94CY017_41     ---A--G-----A-----------------------------------G---------...--A-A-----------A---------------------------------------------------A------------GCT--------T------A-A-GG--C- 1823
A.SN.01.DDI579          A--A--G-----A-----------------------------------G--------C...--A-A-----------A------------------------------------------------------------------T--------T------A-A--G---- 1664
A.SN.01.DDJ369          A--A--G-----A--G--G--------------------------------------C...--A-A---------G-A---------------------------------------------------A--------------T--------T------A-A--G---- 1667
A.SN.96.DDJ360          A--A--G-----A--------------------------------------------C...--A-A-----------A------------------------------------------G-----------------------T--------T------A-A--G---- 1664
A.CD.02.02CD_KTB035     A--A--G-----A--G-----------------------------------------C...--AGA-----------A---------------------------------------------------A-----------T--T--------T------A-A--G---- 2460
A.CD.97.97CD_KCC2       A--A--G-----A-----------------------------------G--------C...--A-A-----------A-----------------------------------------------A------------------T--------T------A-A--G---- 2462
A.CD.97.97CD_KTB13      A--A--GT----A--------G--------------------------G--------C...--A-A-----------A------------------------------------------------------------------T--------T------A-A--G--C- 1670

B.AR.04.04AR151516      ---A---A--------------C----C--------------------M--------C...----A-----------A---------------------------------C------------------------------TCT--------T--C------------G 1688
B.AU.87.MBC925          ---A------------------------------------------------------...----A-------------------------------------------------------------------------------------------------------- 2471
B.BO.99.BOL0122         -------------------------------------------------C-------C...----A-----------A-----------------------------------------------AG----------------C------------C--G---------- 1686
B.BR.03.BREPM2012       ------------------A------------------------------G--------...----A-----------A-----------------------------------------------------------------CT--------T---------------- 1891
B.CA.97.CANB3FULL       ----------------------------------------------------------...----A----------------------------------------------------------------------------AG------------------------AG 1755
B.CN.05.05CNHB_hp3      -------A-------------G-----------------------------------C...----A-----------------------------------------------------------------------------C------------C-----C--G---G 2465
B.CO.01.PCM001          ---A--G-----------------------------------C---------------...--------------G-A-----G---------------------------G--------------------------T------G-------------G---------- 1671
B.GB.83.CAM1            ----------------------------------------------------------...----A-----------------------------------------------------------------------------C-------------------------- 2467
B.GB.86.GB8             ------G--------------------------------------------------C...----A-----------------------------------------------------------A---A---------------G----------C------------- 2486
B.GE.03.03GEMZ010       ----------------------------------------------------------...----A---A-------------------------------------------------------------------------C---------T--C-----C------- 1671
B.IT.05.SG1             ------------A-------------------------------------G-------...----A---------G-----------------------A---G---------------------------------------C----------------N--------- 2268
B.JP.05.DR6538          ----------------------------------------------------------...----A-----------A---------------------A-------------------------------------------C-------------------------- 2470
B.KR.05.05CSR3          ------G------------------------------A--------------------...--------A------------------------------------------------------------------------A-G------------------------- 2491
B.NL.00.671_00T36       ------G--------------------C------------------------------...---GA---------------------------------------------------------------------------------------TA-----------A--- 2031
B.RU.04.04RU128005      ----------------------------------------------------------...--------A-------A---------------------R-----------------T------------------------ACT-----C--A---------------- 1957
B.TH.00.00TH_C3198      -------T-------------------------------------------------C...----A---------------------------------A-------------------------------C------------------C------G 1680
B.UA.01.01UAKV167       ----------------------------------------------------------...----A-------------------------------------------------------------------------------G----------C-----C------- 1692
B.US.04.ES10_53         ----------------------------------------------------------...--A-A-----------------------G---------------------A-------------------------------C---------T---------------- 2475
B.US.99.PRB959_03       ----------------------------------------------------------...----A-----------------------------------------------------------------------------C--------------------G----- 1686

C.AR.01.ARG4006         ---A--G--------G-----C-----C------------------------------...----AA----------AT--------------------A--------------------------------------A---G-T--G-----T------A-A--G---- 1650
C.BR.04.04BR013         ---A--G----G---G-----CA----C------------------------------...--AGAA----------AG--------------A----TA--------G----------G------------------AG---CT--------T------A-A--G---- 1930
C.BW.00.00BW07621       ---T---A-------G-----T---------------------------C-------C...--A-A-----------AG--------------------A--------------------------------------A-----T-----G--T------A-A-GG---- 1819
C.CN.98.YNRL9840        ---C--GAC---A--G-----C-----C------------------------------...--A-A-----------A---------------------A-----------------------------A--------A----CT--------T------A-A--G---- 1650
C.ET.02.02ET_288        ---A--G--------------C-----C------------------------------...--A-A-----------A---------------------A-----------------T-----------A--------A----AT--------T------A-A--G---- 1668
C.GE.03.03GEMZ033       ---C--GA-------G-----C-----C------------------------------...--A-----------G-A------------------------------------------------------------A---G-T--------T------A-A--G---- 1647
C.IL.99.99ET7           ---C--GA---G---G-----C------------------------------------...--A-A-----------A---------------------A-----------------------------A--------A-----T--------T------A-A--G---- 1647
C.IN.99.01IN565_10      ---C--GA----A--G-----C-----C-----G------------------------...T-AGC-----------A---------------------A-----------G-----------------A--------A---ACT--------T------A-A--G---- 1847
C.KE.00.KER2010         ---C---A---G---G-----C-----C------------------------------...--A-A-----------A---------------------A--------------------------------------AG-CA-T--------T-----GA-A--GA--- 1647
C.MM.99.mIDU101_3       ---C--GA---TA--G-----C-----C------------------------------...--A-A-----------A---------------------------------------------------A--------A---TCT--G-----T-----GA-A--G---- 1817
C.MW.93.93MW_965        ---C--GA-------G-----C------------------------------------...--A-A-----------A---------------------A-----------------------------------G--AG---CT--------T------A-A--G---C 1653
C.SN.90.90SE_364        ---C--GA-------G-----T------------------------------------...--A-A-----------A---------------------A-----------A--------------------------A----CT--------T------A-A--G---- 1635
C.SO.89.89SM_145        ---C--GA----A--G-----C-----C--------------------G--------C...--A-A-----------A---------------------A--------------------------------------A----CT--------T------A-A--G---- 1687
C.TZ.02.CO178           ---C---G-------G-----C-----C------------------------------...--A-A-----------A---------------------A--------------------------------------A---TCT--------T------A-A--G---- 1647
C.UY.01.TRA3011         ---A--G--------G-----C-----C--------------C---------------...----AA--A-------AT--------------------A--------G--A--------------------------A---G-T--------T------A-A--G---- 1638
C.YE.02.02YE511         ---C--GA-------G-----C------------------------------------...--A-A-----------A---------------------A--------------------------------------A-----T--------T------A-A-GG---- 1641
C.ZA.04.04ZASK164B1     ---C--GA-------G-----C-----C------------------------------...--A--C----------A---------------------A--------------------------------------A---TCT--------T------A-A--G---- 1880
C.ZA.05.05ZAPSK240B1    ---T--GG----A--G-----C-----C--------------------G--------C...--A-A-----------A-----------G---------A-----------------------------A--------A-----T--------T------A-A--G---- 1870
C.ZM.02.02ZM108         ---C--GA-------------C-----------G------------------------...--A-A-----------A---------------------A-----------------------------A--------A----CT--------T------A-A--G---- 2435

D.CD.83.ELI             ---A--G---------------------------------------------------...----A-----------A------------------------------------------------------------A----C------------------G------- 2012
D.CM.01.01CM_0009BBY    ---A--GT----A---------C----------------------------------C...--A-A-----------A-------------------------------------------------------------G---CT-----------------G--G---- 1665
D.KE.01.NKU3006         ---A--G----G----------------------------------------------...--AGA-----------A------G-----------------------------------------------------A-----AG----C-------------G----- 1671
D.KR.04.04KBH8          ---A--GA-------------------------------------------------C...--AGAA----------A---------------------------------------T--------------------A------------------------------- 2425
D.TD.99.MN011           ---A--G-A---A--------------C-----------------------------C...--A-A-----------A---------------------------------------------------A-----G-----C-CTG------------------CG---G 1690
D.TZ.01.A280            ---A--G---------------------------------------------------...-----C----------A------------------------------------------------------------A-----------------------A------- 1667
D.UG.99.99UGK09259      --AA--G--------G------------------------------------------...--AGA-----------A---------------------------------C--------------------------A-----TG------------------G----- 1671
D.YE.01.01YE386         ---A--G---------------------------------------------------...----A-----------A------------------------------------------------------------A-----TG-------T--------------C- 1665
D.YE.02.02YE516         ---A--G-A---A--------------------------------------------C...--AGA-----------A---------------------------------------------------------A---G-C--TG-------------G-----G---- 1677
D.ZA.90.R1              ---A--G--------------C------------------------------------...----A-----------A------------------------------------------------------------A----C-T------------------------ 1812

F1.AR.02.ARE933         ---A--G-----------G--------------------------------------C...--A-A-----------A-----------------------------------------------A----------A-C--------------T--------C--G---- 1760
F1.BE.93.VI850          ---A--GA-------G-----------------------------------------C...--A-A-----------A-----------------------------------------------A----------A-C--------------T--------C--G---- 1803
F1.BR.01.01BR125        ------G--------------T-----------------------------------C...--A-A-----------A------G----------------------------------------A----------A-C--------------T--C-----C--G---- 1911
F1.ES.x.P1146           ---A--G----G---------------------------------------------C...--A-A-----------A------G----------------------------------------A------------A---AC------T--T--------C--G---- 1743
F1.FI.93.FIN9363        ---A--G--------------------------------------------------C...--A-A-----------A---------------------------------------T-------A------------C--------------T--------------C- 1795
F2.CM.02.02CM_0016BBY   -----------G---G-----------------G-----------------------T...--A-A----T------A-----------G------------------------------------------------AG---C---------T--------A--G---- 1656
F2.CM.95.MP257          A----------G---G-----------------G-----------------------T...--A-A-----------A---------------------------A--G-----------------------------AG--TC---------T--------A------- 1675
F2.CM.97.CM53657        -----------G---G--A--------------G-----------------------T...--A-A----T------A-----------GG--------------------------------------------C--A----C---------T--C-----A--G---- 1656

G.BE.96.DRCBL           ---A--G----TA-----C--------C------------------------------...--AGA---A-------A------------------------------------------------------------A-----T--------TA-----A-A-GG---- 2430
G.CM.01.01CM_4049HAN    ---A--G----TA-----C--G-----C--------------C--------------C...--A-A---A-------A------------------------------------------------------------A-----T--------TA-----A-A--G---- 1668
G.CU.99.Cu74            ---A--G----TA-----C--------C-----------------------------T...--A-A---A-------A------G-----------------------------------------------------A--------------T------A-A--G---- 2070
G.ES.99.X138            ---A----C--TA-----C--G-----C--------------------G--------T...--A-A---A-------A-------------T----------------T-----------------------------A-----TG-------T------A-A--G---- 1907
G.GH.03.03GH175G        ---A--G----TA-----C--------C--------------C---------------...--A-A---A-------A------------------------------------------------------------------T--------T------A-A--G---- 2506
G.KE.93.HH8793_12_1     ---A--G----TA-----C--------C-----G-----------------G-----C...--A-AG--A-------A------------------------------------------------------------A-----T--------TA-----A-A--G---- 1867
G.NG.01.01NGPL0669      ---A--G----TA-----C--------C-----G---------------------C--...--A-A---A-------A---------------------------------A-------------A------------A----CT--------TGAG--GA-A------- 1671
G.PT.x.PT2695           ---A--G----TA-----C--G-----C-----------------------------T...--A-A---A-------A---------------------------------------T--------------------A-----T--------T------A-A--G---- 2408
G.SE.93.SE6165          ---A--G----TA-----C--------C--------------C---------------...--A-A---A-----------------------------------------------------------A--------AG---CT--------TA-----A-A--G---- 1863

H.BE.93.VI991           ---A--GT-G-----G-----------C-----------------------------C...--A-A-----------A---------------------A-----------------------------A-----G--AG--GC---------T-T----A-A--G---- 1855
H.BE.93.VI997           ---A--GT---G----------------------------------------------...--A-A-----T------------------------------------------------------------------AG--GC----------------A-A--G---- 1790
H.CF.90.056             ---A--GT---G---------------------------------------------G...--A-A------G----A---------------------------------------------------------G--AG--GC----------------A-A--G---- 1813

J.CD.97.J_97DC_KTB147   ------GT--------------------------------------------------...--AGA-----------A-------------------------------------------------------A--G-T--------------TGAG---A-A--G---- 1661
J.SE.93.SE7887          ------G--G-G---------------------------------------------C...--AGAC------C---A-------------------------------------------------------A-CG--G---CG--------TGAG---A-A--G---- 1780

K.CD.97.EQTB11C         ---A--GT---GA---------------------------------------------...--A-A---A--G----A---------------------------------------------------A--------AG--TGT--G-----T-----G--A--G---- 1662
K.CM.96.MP535           ---A--GT---GA---------------------------------------------...--A-A-----------A---------------------------------------------------A--------AG----T--------T--------A--G---- 1662
```

```
B.FR.83.HXB2                GGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAA...ATGAGTTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGTGGACATAAAGCTA 2466
Pol                         __G__Q__L__K__E__A__L__L__D__T__G__A__D__D__T__V__L__E__E__.__M__S__L__P__G__R__W__K__P__K__M__I__G__G__I__G__G__F__I__K__V__R__Q__Y__D__Q__I__L__I__E__I__C__G__H__K__A__
01_AE.CF.90.90CF11697       A--A--G--G--A-------------------------------------------C...--A-A-----------------------------------------------------G--------A--------A-----T--------T------A-A-GG---- 2415
01_AE.CN.05.FJ051           A--A--G--G--A-------------------------------------------T...--A-A----------AG-------------------A--------G----------G--------A--------------T--------T------A-A--G---- 2476
01_AE.CN.06.FJ054           A--A--G--G-GA-------------------------------------------T...--A-A----------A------G--------------------------------------G--G--A--------G----T--------T------A-A--G---- 2491
01_AE.HK.x.HK001            A--A--G-----A-------------------------------------------T...--A-A----------A-----------------------------------------------------A--------------T--------T------A-A--G---- 1828
01_AE.TH.01.01TH_R2184      A--A--G--G--A-------------------------------------------T...--A-A----------A-----------------------------------------G--------A----A------CT--------T------A-A--G---- 1665
01_AE.TH.02.OUR769I         A--A--G-----A----------------------------------------G--T...--A-A----------A-----------------------------------------G-----G--A-----------------------T------A-A--G---- 1665
01_AE.TH.90.CM240           A--A--G--G--A-------------------------------------------T...--A-A----------A-----------------------------------------G----AG--A--------------T-----------------A-A--G---- 2040
01_AE.US.00.00US_MSC1164    A--A--G--G--A-------------------------------------------T...--A-A----------A-----------------------------------------------------A--------A-----T--------T------A-A--G---- 1671
02_AG.CM.02.02CM_4082STN    ---A--G----TA--G--C--------C--------------T--------------...--A-A--A-------A-------------------------------------------------------------------CT--------T------A-A--G--C- 1665
02_AG.EC.x.ECU41            ---A--G----TA-----C------------------------C--------------...--A----A-------A----------------------------------------G------------------------CT--------T------A-A--G--C- 1562
02_AG.GH.03.GHNJ196         -T-T--G----TA-----C--------C-----------------------------AAAG-AGAA--A-------A-------------------------------------------------------------------T--------T------A-A-GG--C- 2495
02_AG.NG.01.PL0710          ---A--G----TA-----C--------C----------------------------C...--A-A--A----G-A-----------G---------------------T-----------------A-----------GCT--------T------A-A--G--C- 1650
02_AG.NG.x.IBNG             ---A--G----TA-----C--------C----------------------------C...--A-A--A-------A--------------------------------------------------A--------A-----T--------T------A-A--G--C- 1991
02_AG.SE.94.SE7812          ---A--G----TA-----C--------C-----------------------------...--A-A--A-------A-------------------------------------------------------------------T--------T------A-A--G--C- 1838
02_AG.SN.98.MP1211          ---A--G----TAA-------------C---------A---------------A---...--A-A--A-------A-----------------------------------T-------------------------------T--------T------A-A--G--C- 1663
03_AB.RU.97.KAL153_2        A--A--G-----A-------------------------------------------C...--A-A----------A-----------------------------G---------------G---------------------T--------T------A-A--G---- 1692
04_cpx.CY.94.CY032          ---A--GA---G---G-----T-----------------------------------...--A-A----------A------G--------------------C-----------------------A-------------CT--------T------A-A--G--C- 1832
05_DF.BE.x.VI1310           ------G------------------------G------------------------C...--A-A----------A--------------------------------------------------------------------TG-----------------G------- 1848
06_cpx.AU.96.BFP90          ---A--G----TA-----C--------C----------------------------C...--A-A--A-------A------------------------------------------------A--------A-----T-----G--T------A-A--G---- 2494
08_BC.CN.97.97CNGX_6F       ---C--GA----A--G-----C-----C--C--------------------------...G-A-A----------A-------------------A-------------------------------A-----G--A---CT--------T------A-A--G---- 1646
09_cpx.GH.96.96GH2911       A--A---T-------------G-----C-----------------G-----------...--A-A----------A-----------------------------------------------------------A-----T-----G--T--------A--G--G- 1668
10_CD.TZ.96.96TZ_BF061      ---A--G-----------T--------------------------------------...----A----------A--------------G--A-----------------------------------------A-----T--------A------T--------- 1845
11_cpx.GR.x.GR17            ------G--G-----------------------------------------------...----------------------------------------------------------------A---A-----GG-T---AC------G--TGAA---A-A--G---- 1768
12_BF.AR.99.ARMA159         ---A--G-------------------------------------------------C...--A-A----------A-----------------------------------------------A----------A-CG--------------T--------C--G---- 2471
13_cpx.CM.96.1849           --AA--G----TA-----C-----------G-------------------A------...--A-A---------------------------------------------------------------------A-C--GAG---------TGAG---A-A--G---- 1871
14_BG.DE.01.9196_01         ---A--G----TA-----C--G-----C-----------------G--------T...--A-A--A-------A-------------------------T-----------------------------A-----T--------AG-----A-A--G---- 1983
15_01B.TH.99.99TH_MU2079    A--A--G--G--A--------------C--------------C-----------T...--A-A--C----G-A------C------------------------------------G-----G--A-------------CT--------T------A-A--G---- 1865
16_A2D.KR.97.97KR004        A--A--G----GA-----------------------------G--------T...--A-A-C-A-------A-----------------------------------------------G--A------------ACT--------T----A-A-A-G---C- 1829
17_BF.PE.02.PE02_PCR0155    ---A--G-------------------------------------------------C...--A-A--A-------A----G--------------------------------------------A----------A-C----C---------T--------C--G---G 1656
18_cpx.CU.99.CU76           ---A--G----TA-----C--------C----------------------------C...--A-A--A-------A-----------------------------------------------------A--------A-----T--------T------A-A-GG---- 1771
19_cpx.CU.99.CU7            ---A--G-------------------------------------------------C...--A-A--A-------A----G-----------------------------------------------------------A----C---------T--------G------- 1692
20_BG.CU.99.Cu103           ---A--G----TA-----C--------C----------------------------T...--A-A--A-------A--------------------------------------------------------------A-----T--------T------A-A--G---- 1932
21_A2D.KE.91.KNH1254        ---A--G--------------------------------------------------...----A----------A-------------------------G-------------------------------------A------------------------------- 1671
22_01A1.CM.01.01CM_0001BBY  A--A--G-----A---------C---------------------------------C...--AGA----------A-----------------------------------------G--------A--------A-----T--------T------A-A--G---- 1674
23_BG.CU.03.CB118           ---A--G----TA-----C--------C----------------------------T...--A-A--A-------A-----------------------------------------------------------A----CT--------T------A-A--G---- 1909
24_BG.CU.03.CB378           ---A--G----TA-----C--------C----------------------------G...--A-A--A-------A-----------------------------------------------------------A-----T--------T------A-A-GG---- 1891
24_BG.CU.03.CB471           ---A--G----TA-----C--------C----------------------------T...--A-A--A-------A-----------------------------------------------------------A-----T--------T------A-A--G---- 1906
25_cpx.CM.01.101BA          T------A---TA-----C--------C-----------C------C-------C...--A-A--A----G-A------G---------------------------------T-----------A--------------T--G-----TA-----A-A--G---- 1674
27_cpx.CD.97.97CDKTB49      ---A----------------------------------------------------C...--A-A----------A----------------------------------------------AG--------G--AG--G-TG---------------A-A--G--C- 1674
28_BF.BR.99.BREPM12609      ---A--G----G--------------------------------------------C...--A-A----------A-----------------------------------------------A----------G-C-----------C--T--------C--G---- 1832
29_BF.BR.02.BREPM119        ---A--G-------------------------------------------------C...--A-A----------A-----------------------------------------------A---------GA-C---------------T--------C--G---- 1640
31_BC.BR.02.110PA           ------G----GA--G-----C-----C-----------------------------...--AGAA---------AT-------------------A-------G-------------------------A---A---CT---------------A-A--G---- 1900
32_06A1.EE.01.EE0369        --AA--G----TA-----C-------------------------------------T...--A-A--A-------A-----------G-------------------------------------------------A---AT-----G--T------A-A--G---- 2101
33_01B.MY.05.05MYKL007_1    -------T------------------------------------G-----T...--A-A-----------------------------------------------------G-----G--A--------------T--------T------A-A-GG---- 1845
34_01B.TH.99.OUR2478P       A--A--G--G--A-------------------------------------------T...--A-A----------A-------------------------------------------------------------------CT-----------C-----C------- 1668
35_AD.AF.05.05AF095         ---A--G-A--G--------------------------------------------T...--A-A----------A-----------------------------C--T--G---A-------------A-----T--------T-----GA-A--G---- 1659
36_cpx.CM.00.00CMNYU830     ------GT---TA-----C--------C----------------------------C...--A-A--A-------A--------------------------------------------------A-------------CT--------T-----GA-A--G---- 1671
37_cpx.CM.00.00CMNYU926     ---A--G----TA-----CT-------C----------------------------C...--A-A----------A---------C---------------------A---------------A--------A----CT--------T------A-A--G---- 1656
38_BF1.UY.05.UY05_4752      ---A--G----T--------------------------------------------C...--A-A----------A-----------------------------------------------A----------A-C-----TG-------T------T-C--G---- 1877
39_BF.BR.03.03BRRJ103       ---------------------------------------------------------...----A----------A-------------A-----------------------G----AG--------AG-C----C---------T--------C-GG-T-- 1955
40_BF.BR.04.04BRRJ115       A--A--G----GA-----------------------A-------------------T...--A-A--A-------A------------------------------------------------G-A----------G-C----C---------T--------C--G---- 2018
42_BF.LU.03.luBF_05_03      --AA--GA--------------------------------------------------...----------------------------------------------T-------A----------A-C---AG------C--T--------C------- 2004
43_02G.SA.03.J11223         --AA--G----TA--------------C--------------------------C--...--AGA--A---------------------------------------A-------------A----------A---CT--------TGAA---A-A--G---- 1991
U.CD.83.83CD003             ---A--G-----A--------------------------------G--------...----C-----------A--------------------------------------------------------------A--------------T------A-G--G---- 1961
U.CD.90.90CD121E12          ---A--G-----A--------------------------------G--------...----C-----------A--------------------------------------------------------------A--------------T------A-A-GG---- 1665
U.GR.99.GR303               ---A--G----TA--------------C--------------------------G-...--A----A-------A--------------------A-----------------------------A-------------T--------T-----GA-A--G---- 1740
U.NL.95.U_NL_95_H10986_D1   -------G------------G----------G------------------------T...--A-A--------G-A----G----G-----------------------------------------------A-----G--AG--TGT--------T--------A--G---- 2351
CPZ.CD.90.ANT               ---A---AA-TGTC-G------------------T-----C-----G---G-G-...--TCA-----A-----AC-----G-------CA--T--T--------G--------TTCC---CA-------A-CA-AG---C----C----TG-A-ACAGG-C--TAC 1905
CPZ.CM.01.SIVcpzCAM13       A------T--TGT-----C--G-----C-----G--T-------------A--C---C...--TCAA--A-AG----TT----------------------------------------C-----T-----G-A---A------A-TG--TCA-----G--AGAG--GA-A----TGC 2046
CPZ.CM.05.SIVcpzEK505       A--AG--A----A--------TC-------------T-------------A----------...--ACAG---GA-----A----------------------------------------G------------A-----------A-TG-GA-A-----G--ACAA--GA-A-----AG 2038
CPZ.CM.05.SIVcpzLB7         ---CT-CT-G--A-G---A--G-----------G------------GC-------T...T-AGAA--AGGG-----G------------G-------A---------------------------------AG--A------A-T----CA---------AG-A---A-A--G--AC 2021
CPZ.CM.05.SIVcpzMB66        ---A--GT----A-----A--G-------------------C-----GC-----------...--AGAA---GG---T-------------------------------------------C--A------------A------A-TG-TACA-----C--A------A-A-G---AG 2012
CPZ.CM.05.SIVcpzMT145       A--A---T---TT-----CT-GC----C-----G--T-----C-------A-T-----C...T-A-A-------------A-----GT-------------------------------------T-------A---A-T---AAGAG--AA---------AGAA---A-A--GATA- 2008
CPZ.GA.88.GAB1              A---------TGT------T-GC-------------T-------------A---GAG-...--ACAA--A-A---CTT--------------------A------------------------C-A---A-T---A-TG---A-----G--AGAA--GAGA----TAG 2527
CPZ.TZ.00.TAN1              ---A---G-CTGTC-------T--------T-----------C-GT--T--TTGTA-C...--C-AA--AAAG---CAG-----C--------CC-----A-----A-----A--G-ACC---T--TG----CT--A-T--T-CAG--C-----TG-CAATA-AG---TC- 2110
CPZ.US.85.CPZUS             A-----GA--GTA--------T-----C--------T-------------C----CA-C...--ACAAA-TGA--G-C-----G------------G------------------A--------A---A--------CG-CAAT--------AGAG---AGA-----AC 2526
CPZ.CM.x.SIVcpzDP943        ------G--G-TA--------GC----C--------T-------------GA----TA-T...G-ACAA--AA-----------------------------------------------------A-----------A---A------A-C---GCA--------TGAA--CAGA-----A- 2052
CPZ.CM.x.SIVcpzMB897        A--AGTTA----A--G-TA--------C-----G----------------C----G---...--AGAA--AGAG----A-----GG--------------------------------A-------------A----------------C---------T------A-A-G---A- 2032
N.CM.02.DJO0131             AAAAG--G---GA--------TC-------------T-------------GA-------G...--ACAA--AGA-----A------G-------------A----------A-----------------A------A-TG--ACA-----C--ACAA---AGG-----AG 1972
N.CM.04.04CM_1015_04        AAATG--G---GA--------T--R-----------T-------------A---------...--ACAA--AGA---------------------------A----------A-----------------A------A-T---ACA-----C--ACAA---AGG-----AG 1971
N.CM.04.04CM_1131_03        AAATG--G---GA--------T-----C--------T-------------A---------...--ACAA--AGA---G---------------------------------A-----------------A------A-T---ACA-----C--ACAA---AGG-----AG 1970
N.CM.95.YBF30               -AAAG--G---GA--------T--------------T-------------A-------G...C-ACAA--AGAG----A---------------------A----------A----------G-----A------A-T---ACAG----C--ACAG---AGA-----AG 2060
N.CM.97.YBF106              AAA-G--A---GA--------T--------------T-------------A---------...--ACAA--AGA-----A---------------------A----------A-----------------A------A-T---ACA-----C--ACAA---AGA-----AG 2058
O.BE.87.ANT70               ---C--C---TGT----T-T-GC-G------------------C------C--A-CA-C...--ACAA---GA-----A-------------------------T--A-----------A--------A-G-A------A-TG-GACAG-------AGAA---AGGG-G-TAC 2520
O.CM.91.MVP5180             ---T--T---TGT--G---T--C-G--------G--------------------A-TA-C...--ACAA--AGA-----------C------------------T--A-----C-----A--------A-G-----A-CA-TG-GACAG-----G-ACAA---A-GG---TAC 2495
O.CM.96.96CMABB637          ---C--C---TGT--G-T-T--C-G--------G---------------C--A-CA-C...-CACAA---GA-----------C------------------T--A--R--G-----A--------A-G-A---G-A-TG-GACAG-------AGAR---A-A--G-TGC 1950
O.CM.98.98CMA104            ---C--T---TGT----TCT-GC-G--------G------------------A-T--C...--ACAA---GA-----A------------------------A-----A--G----G---A-GG----A-G-A-T-A--A-TG--ACAG-------AGAA---AGGG---TAC 1952
O.CM.99.99CMU4122           ---C--T---TGT----T-T-GC-G--C-----G---------------C--A-CA-T...--ACAA---GA-----A-----C------------------T--A-----------A--------A-G-A---T--R-TG-GAAAG----G--AGAG---AGAG---TAC 1951
O.FR.92.VAU                 ---C--G---TGT--G-T-T--C-G-----G--G---------------C--A-CA-C...--ACAA---GA-----A---GC-------------------T--A-----------A--G----A-G-A---T--A-TG-GGCTG-----T-AGAG---AGAG---TA- 2032
O.SN.99.SEMP1299            ---C--T---TGT----T-T--C-G--------G---------------C--ACCA-C...--ACAA---GA-----A-----C------------------T--A-----------A--G----A-G-A---A----AG-G-CAG-------AGAG---AGGG---TAC 2519
O.US.99.99USTWLA            ---C--T---TGT--G-T-T--C-G--------G------------------AGCA-C...--ACAA---GA-----A-----C------------------T--A-----C-----A--G----A-G-A---A-KG-TG-GACAG-------AGAA---AGG---ATGC 1949
```



HIV-1/SIVcpz
Complete Genomes

```
                                 Pol Protease end Pol p66 and p51 RT start
B.FR.83.HXB2                TAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCCATTAGCCCTATTGAGACTGTACCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAA 2636
Pol                         I__G__T__V__L__V__G__P__T__P__V__N__I__I__G__R__N__L__L__T__Q__I__G__C__T__L__N__F__P__I__S__P__I__E__T__V__P__V__K__L__K__P__G__M__D__G__P__K__V__K__Q__W__P__L__T__E__E_
01_AE.CF.90.90CF11697       --------------------------------------------C-----A-------------------T-----------C--A--A--T--------C-------------C---------------------A-----G----------------A---------- 2585
01_AE.CN.05.FJ051           ----------G---------------------------------C----CA-------------------T--------------A--C--T--------C-------------C------------------C--G-----G--------G------------------ 2646
01_AE.CN.06.FJ054           --------------------------------------------------A-------C--------C--T---C-------C--A-----T--------C-----------------------------------GT----G--------G------------------ 2661
01_AE.HK.x.HK001            -------------------------------------------GC-----A----------------C--T-----------C--A--C--T--------C-----------------------------------A-----G--------G------------------ 1998
01_AE.TH.01.01TH_R2184      ----------G---------------------------------C-----A-------------------T--------------A--------------C-------------C---------------------A----GG--------G------------------ 1835
01_AE.TH.02.OUR769I         ----------G------------------A--------------C-----A-------------------T-----------C--A-----T--------C-----------------------------------A----GG--------G------------------ 1835
01_AE.TH.90.CM240           --------------------------------------------C-----A-------------------T-----------C--A-----T--------C-------------C---------------------A-----G--------G------------------ 2210
01_AE.US.00.00US_MSC1164    ----------G------------------A--------------------A-------------------T-----------C--A-----T--------C-------------C-----GA--------------A-----G--------G------------------ 1841
02_AG.CM.02.02CM_4082STN    --------------------------------------------C-----A-------C-----------T--------------A-----T--------------G-----------G-----------------------G--------------------------- 1835
02_AG.EC.x.ECU41            --------------------------------------------C-----A-------------------T-----------G--A-----T--------A--C--G-----------------------------------G--C------------------------ 1732
02_AG.GH.03.GHNJ196         ----------------------------------------------G---A----------------C--T--------------A-----T--------A-----G-----------------------------T-----G--------------------------- 2665
02_AG.NG.01.PL0710          --------------------------------------------C-----A-------------C-----T--C-----------A-----T--------A-----G------------------------------------A-------------------------- 1820
02_AG.NG.x.IBNG             --------------------------------------------C-----A-------------------T--------------A-----T--------A-----G-----------------------------------G--------------------------- 2161
02_AG.SE.94.SE7812          --------------------------------------------C-G---A-------------------T--------------A-----T--------A--C-----------------------C--------------G--------------------------- 2008
02_AG.SN.98.MP1211          -G---T---------------------------------------A----A-T--------A--------T---C----------A-----T--------A-----G---T-------------------------------G--------------------------- 1833
03_AB.RU.97.KAL153_2        -------G-----------------C------------------------A-----------C-------T--------------A--A--T--------A-------------C---------------------------G-----------------A--------G 1862
04_cpx.CY.94.CY032          ----C-----G---------------------------------C----CA-----------C-------T--------------A-----T--------A-----------------------------------------G--------------------------- 2002
05_DF.BE.x.VI1310           --------------------------------------------------T-------C--------C-----------------A--------------A-----------------------------------------------------------A--------- 2018
06_cpx.AU.96.BFP90          -----------------------------A--------------C----CA-------C-----------T--------------A-----T--------A---------------C----------G--------------G--------------------------G 2664
08_BC.CN.97.97CNGX_6F       ----------------G-----C-----------------A---------A-------C---C----A-----AC----------A--C--T--C-----A-----------------------------------------G--------------------------- 1816
09_cpx.GH.96.96GH2911       --------------------------------------------C-G--CA----------A--------T--------------A-----T--------A------------------------------------------A-------------C------------ 1838
10_CD.TZ.96.96TZ_BF061      -------------------------------------------G------T----------------------------------A-----T--------A--G-----------------------------------------------------G-----------G 2015
11_cpx.GR.x.GR17            ----------G---A----------------------------------CA-------C--------C--T--------------A-----T---G----C--A--G--------------------------------------------------------------- 1938
12_BF.AR.99.ARMA159         ----------G---A-------------G---------------------------------C----------------------------T--------A--------------------------------------------------------------------- 2641
13_cpx.CM.96.1849           --------------A-----------------------------------A-------------------T--------------A-----T--------A--------------------------------------------------------------------- 2041
14_BG.DE.01.9196_01         ----G------------------------A-------------G------A-------------------T--------------A-----T--------A-----------------------------------T-----G--------------------------- 2153
15_01B.TH.99.99TH_MU2079    ----------G---------------------------------C-----A----------AC-------T-----------C--A-----T--------A-----------------------------------A-----G--------G------------------ 2035
16_A2D.KR.97.97KR004        --------------------------------------------------A-----GT---AC-G-----T--------------A--A--T--------A--------------------------------------------C------------------------ 1999
17_BF.PE.02.PE02_PCR0155    ----------------------W------------------------------------AT------------------------------T--------A-----------------------------------A--------------------------------- 1826
18_cpx.CU.99.CU76           --------------------------------------------------A-------------------T--------------A-----T--------A-----------------------------------------G--------------------------- 1941
19_cpx.CU.99.CU7            -------------------G------------------------------T----------------C-----------------A--------------A--------------------------G-----------------G------------------------ 1862
20_BG.CU.99.Cu103           ----G------------------------A-------------G------A-------------------T--------------A-----T-----C--T-----G--------------------------------------G------------------------ 2102
21_A2D.KE.91.KNH1254        ----------G-----------------------------C---------T-------C--------------------------A-----T--------A--------------------------------------------------------------------- 1841
22_01A1.CM.01.01CM_0001BBY  ----------G--------------------T------------C-----A----------------C--T--C--------C--A--A--T--------A------------GC---------------------A-----GA-------------------------G 1844
23_BG.CU.03.CB118           ----G------------------------AC------------G------A-------------------T---C----------A-----T-----C--A--------------------------------------------G------------------------ 2079
24_BG.CU.03.CB378           ----G-----------------C------A----------C--G------A-------------------T---C----------A-----T-----C--A-----G--------------------------------------G------------------------ 2061
24_BG.CU.03.CB471           ----G------------------------A-T--------C--G------A----------A--------T---C----------A-----T-----C--A-----G--------------------G-----------------G------------------------ 2076
25_cpx.CM.01.101BA          ----G------C-----------------A-------------G------A----A--------------T--G-----C-----A-----T--------A--------------------A--------------------G--------------------------- 1844
27_cpx.CD.97.97CDKTB49      ----------------------------------T-----------G---T-A--------A-----------------------A--C-----------A--------------------------G--------------G---------------C----------- 1844
28_BF.BR.99.BREPM12609      C---------G--------------------------------------CA-------------------T--------------------T--------A--------------------------------------------------------------------- 2002
29_BF.BR.02.BREPM119        -------------G-----------G------------------------A-------------------T--------------T------------------------------------------------------------------------------------ 1810
31_BC.BR.02.110PA           -----------C-------------C--------------------G--CA-----------C----A-----AC----------A-----T--C--C--A--G--------------------------------------G--C------------------------ 2070
32_06A1.EE.01.EE0369        --------------------------------------------C----CA-------C-----------T-----------G--A--A--T--------A----------------------------------------GG--------------------------G 2271
33_01B.MY.05.05MYKL007_1    --------------------------------------------C-----A-------------------T--------------------T--------A-----------------------------------T--------------G------------------ 2015
34_01B.TH.99.OUR2478P       --------A----------------------------------------------A------C-------T--------------------T--------A--------------------------------------------------------G------------ 1838
35_AD.AF.05.05AF095         -------------G--------------------------------G---A-------C---C-------T-----------C--A-----T------A-T-------------GG--------------------------G--------------------------- 1829
36_cpx.CM.00.00CMNYU830     ----------G------------------A--------G----GC-----A-------------------T--------------A-----T--------A------------GC---------------------A-----G--------------------------- 1841
37_cpx.CM.00.00CMNYU926     -------------G--------C---------------------C-----A-------------------T--------------A--C--T------A-C------------TC------------G--------------G------------------T-------G 1826
38_BF1.UY.05.UY05_4752      C-------------------------------------------------------------C-------------------C--------T--------A--------------------------------------------------------------------- 2047
39_BF.BR.03.03BRRJ103       ----------G--GA----------------A--T---G--------G-C--CA----------G--C-----------------------T--------A-----------G-----------------------------G--------------------------- 2125
40_BF.BR.04.04BRRJ115       ----------------------------------T--------G---G----T-----C---C----C------C----------------T--------A----------------------------------------------G---------------------- 2188
42_BF.LU.03.luBF_05_03      ----------------------------------T---------------------------C----C--------------C--------T--------A-----------------------------------------G--------------------------- 2174
43_02G.SA.03.J11223         ----G------C-----------------A-------------C------A-------------------T--------------A-----T--------------------------------------------------G----------------------C---G 2161
U.CD.83.83CD003             --------------------------------------------------A----------A-----------------------A-----T--------A-----------------------------------------G--G------------------------ 2131
U.CD.90.90CD121E12          -----------C-------G------------------------------A----A-----------------------------A--C--T--------A-----------------------------------------G--A--------------A--------- 1835
U.GR.99.GR303               --------------------T-----------------------C-----G-------C-----------T-----------C--A-----T--------A------------GC------------G--------T----GA-----------------R-Y------R 1910
U.NL.95.U_NL_95_H10986_D1   ----------------------------------------------G--CA-------------------T--------------A-----T--------A-----------------------------------------------G-----------A--------- 2521
CPZ.CD.90.ANT               ---CA------C-GT-G-----A-AC--A--A-----T--A--T------G-T--ATG--TTT-A--A--T--------------A-----TAAAG----A--A--G--C--T--------AGA------------A----G---A-----------GC-CT--A----G 2075
CPZ.CM.01.SIVcpzCAM13       A---G-----G--G--G-----A-----------T--------G------A-T--A--A-----A--G-----AC-GGTA--C--T--C--TT-A-----A-----C-----C-----G--A-----------C--G----G-A-A--G--------T--AT---C---- 2216
CPZ.CM.05.SIVcpzEK505       -T--C-----C--------G--A--------T--T--T------------T-TC----A--A--------T---C----------A--A--T--------A--CA-------G-----G--------T-----------G-G---A-----G-----T------------ 2208
CPZ.CM.05.SIVcpzLB7         ----------T-----------C-----A--A--T--C--A-----G--CA-T--A-----A-----A--T---C----------T--------------A------------C----------------------G--------A-----------C--A-----G--- 2191
CPZ.CM.05.SIVcpzMB66        ----------T-----------C--------A--T--C--A-----G---A-T--------A-----A--------------C-----C--T--------A-------------GT--------------------G----GG--C-----G--------A--------- 2182
CPZ.CM.05.SIVcpzMT145       CCA----T--------------A--C--------T-----A---------G-CC-A--AA-AT-A--G--T---C-------C--A--A--T--A--A--C--A---AA---GTTTC----------G-----C--G--C--G--A-----G-----TC--T--A----- 2178
CPZ.GA.88.GAB1              ----G-----------------A--------A--T---------------A-T-----A--AT-G-----T---C--GTG--C--A-----TT-A-----A-----C-----C-----G--A--------------T-----G--A--G--------C--AT---C---- 2697
CPZ.TZ.00.TAN1              G--CC--T--CC----G---GAA--C--CA-T--T-----A--T------A-T--A-AG--AT-A--A--T--C-----------T--------A--A---GTA---AA----C-------AGA------------G--------A--G--G-----CC-CT-CA-G--G 2280
CPZ.US.85.CPZUS             AG---T----T-----G-----A-----A--A--T--T-----C--G---A-T--A--A--A-----A--T--------------T--C---T-----A-A-------------G------A-----------------T-G---A--G----------------C---- 2696
CPZ.CM.x.SIVcpzDP943        C---C-----G--G--G-----A-----A--G--T--------C------A-A-----A--------C--T-----------C--T--A--T--C----GA-----G--------------------------------T-G---A-----------C--A--------G 2222
CPZ.CM.x.SIVcpzMB897        -------------------G--C--------G--T--T--A--G------A-C--A-----A-----G--T---C-----------------A----A--A--------------------------G-----C--T----G---------G-----CC-A--------- 2202
N.CM.02.DJO0131             -T--------G-----------A--------T--T--T--A--------CT-T--A--A--------C--T--------------A--A--T--------A--------------------A-----G-------------G---A-----------T-------C---- 2142
N.CM.04.04CM_1015_04        -T--------G--G--------A--------T--T--T--A--------CT-T--A--A--------C--T--CC----------A--A--T------A-A--------------------A-----------------------A-----------T-------C---- 2141
N.CM.04.04CM_1131_03        -T--------G-----------A--------T--T--T--A--------CT-T--A--A--------C--T---C----------AT-A--T------A-A--------------------A-----------------G-----A-----------T-------C---- 2140
N.CM.95.YBF30               -T--------------------A--------T--T--T--A-----------T--A--C--------C--T--------------A--A--T--------A--------------------A--------------------G--A-----------T------AC---- 2230
N.CM.97.YBF106              -T--------G-----------A--------T--T--T--A--------CT-T--A--A--------C--T--------------A--C--T--------A--------------------A-----G-------------GG--A-----------T-------C---- 2228
O.BE.87.ANT70               AG--A--------G--G--------T-----T--T--T------------A-A-----AGGAT-A-----T--AC----C--C--T--A-----C--A-CCC-A--G---------C----A--------------A--------A-----------CC-AT-TA----- 2690
O.CM.91.MVP5180             AG--A--------G--G--------T-----T--T--TC----G-----CA-A-----AGGAT-A--A--T--AC-------C--T--A--T--C--A-CCC-A--G--------GC----A--------------A--------A-----------CC-AT-TAG---G 2665
O.CM.96.96CMABB637          A---G--------G--G--------T-----T--T--T--------G---A-A-----AGGAC-C--------AC-------C--T--A-----C--A-CCC-A------------C----A--------------A------T-A-----G-----CC-AT-TAG---- 2120
O.CM.98.98CMA104            AG--A--------G-----------T-----T-----T-----------CA-A-----AGGAT-A--------AC-------C--T--A---T-C--A-CCC-A----------G-C----A--------------A--------A-----------CC-AT-TA----- 2122
O.CM.99.99CMU4122           AG--A--------G--G--------T-----T--T--T------------A-A-----AGGAT-A--C--T--AC-------C--T--A-----C--A-CCC-A--G--------RC-G--A--------------A------A-A-----------CC-AT-TA----- 2121
O.FR.92.VAU                 GG--G--------G--G--------T-----T-----TC-A---------A-A-----AGGAT-A-----T--AC-------C--T--A-----C--AACCC-A------------C----A--------------A--------A-----------CC-AT-CAG---G 2202
O.SN.99.SEMP1299            -G--A--------G--G--------T-----T--T--T-----------CA-A-----AGGAT-G--------AC-------C--T--A--T--C--A-CCC-A------------C----A--------------A------A-A-----------CC-AT-TA----- 2689
O.US.99.99USTWLA            AGA-G-----------G--------T-----T--T--TC-----------A-A-----AGGAT-A--G--T--AC-G--Y--C--T--A-----C--AACCC-A-----------GC----A--------------A--------A-----------CC-AT-TAG---- 2119
```

HIV-1/SIVcpz Complete Genomes

```
B.FR.83.HXB2                AAAATAAAAGCATTAGTAGAAATTTGTACA......GAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGAAAATTAGTAGATTTCAGAGAACTTAATAAGAGAAC 2800
Pol                         _K__I__K__A__L__V__E__I__C__T__.__.__E__M__E__K__E__G__K__I__S__K__I__G__P__E__N__P__Y__N__T__P__V__F__A__I__K__K__K__D__S__T__K__W__R__K__L__V__D__F__R__E__L__N__K__R__T
A1.GE.99.99GEMZ011          ---------------AC---C-------AT......--T--------------A--------------------------------------------------T-----------G-----C-----G-----G-----------------G-----C-----A----- 1984
A1.KE.00.KER2008            --------------GAC----------CA-......-----------------A------------------------------------------A-------A--------------T--C-----------G--------------------G--C-----A----- 1999
A1.KE.00.KNH1144            ---------------AC---------C---......-----------------A---------C-------A------------------------A-------A--------------T--C-----------G--------------------G--C-----A----- 2005
A1.KE.00.KSM4024            ---------------AC-------------......--A--------------A------------------------------------------A-------A--------------T--C-----------G--------------------G--------A----- 1996
A1.KE.00.MSA4069            ---------------AC-----------TG......--A--------------A------------------------------------------A-------A-----------G--T--C-----------G-----G-----C--------G--C-----A----- 1999
A1.KE.00.NKU3005            ---------------AC------A---G--......--A--------------A------------------------------------------A-------T-----------------C--------------G--------------------C-----A----- 1999
A1.RU.00.RU00051            -------GG------AC---C-------AG......-----------------A--------------------------C-----------------------T-----------G-----C-----------G-----------------G-----C-----A----- 2101
A1.RW.93.93RW_024           ---------------AC-------------......--A--------------A------------------------------------------A-------T-----------G-----C--------------------------------G--C-----A----- 2002
A1.SE.95.SE8891             ---------------AC-----------G-......-----------------A-----C--------------------------------------------T--------------T--C-----G--------------------T-----G--C-----A----- 2002
A1.SE.95.UGSE8131           ---------------AC-------------......--A--------------A------------------------------------------A-------T-----------GA-----GA--G---------------------------G--C-----A----- 2172
A1.TZ.01.A173               -------G-------AC-----------TG......-----------------A------------------------------------------A-------T-----------------C--------------------------------G--C-----A----- 1999
A1.UA.01.01UADN139          ---------------AC---C-------AG......-----------------A------------G------------------------------------TT-----------G-----CT----G-----G-----------------G--G--C-----A----- 1999
A1.UG.92.92UG037            ---------------AC----------G--......--T-------GA-----A------------------------------------------A-------T-----------G-----C--------------------------T-----G--C-----A----- 2816
A1.UG.99.99UGA07072         ---------------AC-A---------A-......-----------------A--------------------------------------------------T--------------T--C--------------------------T-----G--C-----A----- 1999
A2.CD.97.97CDKTB48          ---------------AC-----------A-......--A--------------------------------------------------C--------G-----T-----------------CGA---------------------------------C----------- 2138
A2.CY.94.94CY017_41         ---------------AC------C----A-......--A--------------A-----------------------------------C--------G-----T-----------------C--------------------------T--------C----------- 2157
A.SN.01.DDI579              --------------GAC--C--------AG......---------G-------A-----------------C--------------------------------T-----------------------------G-----G--------------G--C-----A----- 1998
A.SN.01.DDJ369              --------------GAC--C--------A-......---------G-------A--------------------------------------------------T-----------G--------C--------G--------------------G--C-----A----- 2001
A.SN.96.DDJ360              --------------GAC--C--------A-......---------G-------A--------------------------------------------------T-----------------------------G-----G--------------G--C-----A----- 1998
A.CD.02.02CD_KTB035         ---------------AC-----------G-......--------------------------------------------------T--C------A-------T-----------------C-----------G-----------------------------A----- 2794
A.CD.97.97CD_KCC2           ---------------AC-----------AT......-----------------A-----C--------------------------T---------A-------T----G------------C-----------G--------------T--------C-----A----- 2796
A.CD.97.97CD_KTB13          ---------------AC-------------......-----------------A----------G------------G--------T--C------A-------T-----------------C-----------G-----------------------C-----A----- 2004
B.AR.04.04AR151516          ------------------------------......--A-----------G--A-----C--------------------------------------------T-----------------------G---------C----------T--------C--------G-- 2022
B.AU.87.MBC925              ----------------------------T-GTTATG--A--------------A--------------------------------------------------T---------------G------------------------------------------------- 2811
B.BO.99.BOL0122             --------------------------C---......--A-----------------------------------------------------------------T-----------------------------------------------------------A----- 2020
B.BR.03.BREPM2012           ------------------------------......--A--------------A-----------------------------------------------------------------------------------------------T-------------------- 2225
B.CA.97.CANB3FULL           ------------------------------......--AC----------T--A--------------------------------------------------------------G---G-------------------------------G----------------- 2089
B.CN.05.05CNHB_hp3          ---------------A-G------------......--A--------------A-----------------------------------------------------------------------------------------------T--G-----------A----- 2799
B.CO.01.PCM001              ------------------------------......--A--------------A-------------------------------------------------------------------------------------------------------------------- 2005
B.GB.83.CAM1                ------------------------------......--A--------------A-------------------------------------------------------------------------------------------------------------------- 2801
B.GB.86.GB8                 --------------------G---------......--A--------------A-------------------------------------------------------------------------------------------------------------------- 2820
B.GE.03.03GEMZ010           ------------------------------......--A--------------A----------------------------------------------------------------------------------------------------G----------------- 2005
B.IT.05.SG1                 ----------------N-------------......--A-------------------N--------N------------N---N--------------N------N---------------------G---------N----------------------N-------- 2602
B.JP.05.DR6538              --------------G---------------......--A----------------------------------------------------------------------------------------------------------------------------------- 2804
B.KR.05.05CSR3              ------------------------------......--A----------------------------------------------------------------------------------------------------------------------------------- 2825
B.NL.00.671_00T36           ------------------------------......--AC-------------A------------------------------------------A-------------------------C-----------------------------------------A----- 2365
B.RU.04.04RU128005          --------------------------C---......--A--------------A--------------------------C--------C--------------------------------------------------------C-----------------A----- 2291
B.TH.00.00TH_C3198          ------------------------------......--A--------------A-----------------------------------------------------------------------------------------------T--------------A----- 2014
B.UA.01.01UAKV167           ------------------------------......--A--------------A----------G------------C-----------------------------------------------C----------------------------------T--------- 2026
B.US.04.ES10_53             ------------------------------......--A-------G---------------------------------------------------------------------G-----------------G----------------------------------- 2809
B.US.99.PRB959_03           ------------------------------......--A-----------------------------------------------------------------------------G----------------------------------------------------- 2020
C.AR.01.ARG4006             ---------------AC----------GA-......--A--------------A------A----------------------T--T--C--------------T-----A--G--G-----------G-----------------------G-----C-----A----- 1984
C.BR.04.04BR013             ---------------AC----------GAT......--A-----G--------A--------------------------------T--C------A-------------A--G--GA------AC-GG-----------------------G-----C-----A----- 2264
C.BW.00.00BW07621           ---------------AC--C-------GAT......-----------------A------A-------------------------T--C-----G--------------A--G--G--------A--G--------------------T--G-----C-----A----- 2153
C.CN.98.YNRL9840            ---------------AC--C-------GAT......--A-----G--------A------A-------------------------T--C------A-------------A--G--G-----------G-----------------------G-----C-----A----- 1984
C.ET.02.02ET_288            ---------------AC----------GAG......--A-----G--------A--------------------------C-----T--C--------------------A--G--G-----------G-----------------------G-----C-----A----- 2002
C.GE.03.03GEMZ033           ---------------AC--C-------GA-......--A-----G--------A------A-------------------------T--C--------------T-----A--G--G-----------G-----G-----------------G-----C-----A----- 1981
C.IL.99.99ET7               ---------------AC--C-------GA-......--A-----G--------A------A----------------------------C--------------------A--G--G--T-----C--G--------------------T--G-----C-----A----- 1981
C.IN.99.01IN565_10          ---------------AC--C---C---GAT......--A-----G--------A------A----------------------T--T---------A-------T-----A--GG-------------G-----------------------G-----CG----A----- 2181
C.KE.00.KER2010             ---------------AC----------GA-......--A-----G--------A------GA------------------------T--C--------------------A--G--------------G-----------------------G-----C-----A----- 1981
C.MM.99.mIDU101_3           ---------------AC--C-------GAG......--A-----G--------A------A-------------------------T--C------A-------------A--G--G-----------G-----------------------G-----C--C--A----- 2151
C.MW.93.93MW_965            ---------------AC--C-------GA-......--A--------------A------A----G--------------------T--C--------------------A--G--G-----------G-----------------------G-----C-----A--G-- 1987
C.SN.90.90SE_364            ---------------AC--C-------GA-......--A-----G--------A------A----------A--------------T--C--------------------A--G--G-----------G-----------------------G-----C-----A----- 1969
C.SO.89.89SM_145            ---------------AC--C-------GA-......--A-----G--------A------A-------------------------T--C--------------------A--G--G-----------G-----------------------G-----C-----A----- 2021
C.TZ.02.CO178               -----------T---A-G-C-------GA-......--A--------------A------A----------------------------C--------------T-----A--G--G-----------G--------------------T--G----------------- 1981
C.UY.01.TRA3011             ---------------AC--C-------GAC......--A-----G--------A------A-------------------------T--C--------------T-----A--G--G-----------G-----------------------G-----------A----- 1972
C.YE.02.02YE511             ---------------AC----------GA-......--C-----G--------A--G---A----------------G--------T--C--------------------A--G--------------G-----------------------G-----C-----A----- 1975
C.ZA.04.04ZASK164B1         ---------------AC----------CA-......--C--------------A------A-C-----------------------T-----------------T-----A--G--G-----------G-----------------------G-----------A----- 2214
C.ZA.05.05ZAPSK240B1        ---------------AC--C-------GAT......--A-------G------A------A-------------------------T--C------A-------------A--G--G-----------------------------------G-----C-----A----- 2204
C.ZM.02.02ZM108             ---------------ACT-C-------GAG......--A-----G--------A------A----------------------G--T--C--------------------A--G-----------A--G-----------------------G-----C-----A----- 2769
D.CD.83.ELI                 ---------------AC-------------......--T--------------A----------G-------------------------------A----------------------------C--G----------------------------------------- 2346
D.CM.01.01CM_0009BBY        ---------------AC-------------......--A--------------A----------G---------------------------------------T-----------------------G--------G-------------------------------- 1999
D.KE.01.NKU3006             ------------C--AC-------------......--A--------A------------------------------------------------A-------------------------------G----------------------------------------- 2005
D.KR.04.04KBH8              ---------------AC-------------......-----------------A----------G---------------------------A-----------------------G-----------G-----------------------------------A----- 2759
D.TD.99.MN011               ---------------AC-------------......--A--------------A----------G-------------------------------A-------T-----------------------G--------------------T-------------------- 2024
D.TZ.01.A280                ------------C--AC-------------......--A--------------A----------G-------------------------------A-------T-----------------------G----------------------------------------- 2001
D.UG.99.99UGK09259          ---------------AC-------------......--A--------------A------A---G------T------------------------A-------------------------------G----------------------------------------- 2005
D.YE.01.01YE386             ------------C--AC----------G--......--C--------------A----------G---------------------T---------A-------------------------------G----------------------------------------- 1999
D.YE.02.02YE516             ---------------A----------C---......--A--------------A----------G---------------------T-----C---A-------T-----------------------G--------------------T-------------------- 2011
D.ZA.90.R1                  -----------G---AC-------------......--A--------------A------------------------------------------A------------------------------------------------------------------------- 2146
F1.AR.02.ARE933             ---C----G------TC------A----TT......--A--------A-----A---------------------------------N----------------T-----------------------------G-----------C-----------------A----- 2094
F1.BE.93.VI850              --G------------AC------A---CT-......--A--------A-----A----------------------------------------------------------------------G---------G----------------A------------A----- 2137
F1.BR.01.01BR125            ---------------AC------A------......--A--------------A----------G---------------------------------------------------------------------------------------------------A----- 2245
F1.ES.x.P1146               ---------------AC------A---G--......--T--------------A-----------------------------------------------------------------------------------------------T-------------------- 2077
F1.FI.93.FIN9363            ---------------AC---G--A------......--T--------A-----A----------G---------------------------------------------------------------------G--G--------------------C-----A----- 2129
F2.CM.02.02CM_0016BBY       ------------C--AC------C------......-----------A-----A--------------------A-----------------------------------------G---------------------C----------------G--------A----- 1990
F2.CM.95.MP257              ------------C--AC------C------......--------G--A-----A--------------------A-----------------------------------A--G--G--------------------------------------G--------A----- 2009
F2.CM.97.CM53657            ------------C--AC------C------......-----------A-----A--------------------A-----------------------------------------G-----------------------------------------------A----- 1990
G.BE.96.DRCBL               ---------------AC-----------AT......-----------------A-----------G--------------C-----T--C------A------------------------------G------------G--------------G--C-----A--G-- 2764
G.CM.01.01CM_4049HAN        -----------G---AC-----------AT......--C--------------A-----------------------------------C------A-------------------------------------------------------------C-----A----- 2002
G.CU.99.Cu74                ---------------AC---C------G--......--T--------------A--G--------------------------------C------A-------------A--G-----------------------------------------G--C-----A----- 2404
G.ES.99.X138                -----------T---AC-------------......--A-------G------A-----------------A-----------------C------A-------------------------------------------G--------------G--C-----A----- 2241
G.GH.03.03GH175G            ---------------AC-----------A-......--C--------------A--------------------------C--------C--------------------------------------------------------------------C-----A----- 2840
G.KE.93.HH8793_12_1         -----------C---AC-------------......--A-------G------A------------------------------------------A------------G------G--------------------------------T-----G--C-----A----- 2201
G.NG.01.01NGPL0669          ------------------------------......--A-------------------------G------------------------C------A-------------------G-----------------G-----G--------------G--C-----A----- 2005
G.PT.x.PT2695               ------------------------------......--A-------G------A--G---------------------------------------A----------------------------------------------------------G--C-----A----- 2742
G.SE.93.SE6165              ---------------AC-----------A-......--A------G-------A-----------------------------------C------A-------------------G-----------------------G--------------G--C-----A----- 2197
H.BE.93.VI991               ---------------AC----------TT-......--A--------------T--------------A-----C--G-----------C------A-------------A--G--GA---------G---------------G--------------------A----- 2189
H.BE.93.VI997               ---------------AC-----------TG......--A--------------A--------------A--------G-----------C------A-------------A--G--G--------------------------G--------------C-----A----- 2124
H.CF.90.056                 ---------------ACG------------......-----------A-----A-----C----G---A--------G----------GC------A-------------A--G--G--T-----------------------G--------------C-----A----- 2147
J.CD.97.J_97DC_KTB147       ------G--------AC-C--------GA-......--A--------------A--------------------------------T--C--------G---------------------------------------.----------T--G--G--C-----A----- 1994
J.SE.93.SE7887              ---------------AC-C--------G--......--A------G-------A----------G-G-------------------T--C-----------------------------------------------------------T--------C-----A----- 2114
K.CD.97.EQTB11C             --------------------G---------......--A--------A-----A-----------------------------------C--C-----G----------------------------------T---------------T--------C-----A----- 1996
K.CM.96.MP535               ---------------AC-------------......--A--------A-----A-----------------------G--------T--C--------G--------------------------------------------------T--------C-----A----- 1996
```

```
B.FR.83.HXB2                AAAATAAAAGCATTAGTAGAAATTTGTACA......GAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGAAAATTAGTAGATTTCAGAGAACTTAATAAGAGAAC 2800
Pol                         _K__I__K__A__L__V__E__I__C__T__.__.__E__M__E__K__E__G__K__I__S__K__I__G__P__E__N__P__Y__N__T__P__V__F__A__I__K__K__K__D__S__T__K__W__R__K__L__V__D__F__R__E__L__N__K__R__T
01_AE.CF.90.90CF11697       ------T--------AC-----------A-......---------G-------A-----A-----------------------G--------------------T--T---G----------C-----------G--------------------G--C-----A----- 2749
01_AE.CN.05.FJ051           ----------------C-----------A-......--A------G-------A----------------A---------------T-----------------T-----------G-----C--C--------G-----------------------C-----A----- 2810
01_AE.CN.06.FJ054           -G-------------AC-----------AG......---------G-------A-----C--------C-----------------T-----------------T-----------G--T--C--C--------G--------------------G--C-----A----- 2825
01_AE.HK.x.HK001            ---------------AC-----------A-......---------G-------A--G--C--------------------------------------------T-----------G-----C--C--------G--------------------G--C-----A----- 2162
01_AE.TH.01.01TH_R2184      ---------------AC-----------A-......-----------------A-----C--------------------------------------------T-----------G-----C--C--------G--------------------G--C-----A----- 1999
01_AE.TH.02.OUR769I         --------------GAC-----------A-......--A------G-------A-----C--------------------------------------------T-----------G--T--C--C--------G--------------------G--C-----A----- 1999
01_AE.TH.90.CM240           ---------------AC-----------A-......---------G-------A-----C--------------------------------------------T-----------G-----C--C--------G--------------------G--C-----A----- 2374
01_AE.US.00.00US_MSC1164    ---------------AC-----------A-......---------G-------A-----C--------------------------------------------T-----------G-----C--C-----------------------------G--C-----A----- 2005
02_AG.CM.02.02CM_4082STN    ---------------AA---T------T-C......-----------------A--------------------------------T---------A-------T--------------T--------------------------------------C----------- 1999
02_AG.EC.x.ECU41            ---------------AC-------------......--T--------------A--------------C-----------------T--------------------------------T--C-----------------------------G-----C----------- 1896
02_AG.GH.03.GHNJ196         ---------------AC---C---------......--A--------------AC----------------C-----------C-----------------------------------T-----------------------------------G--------A---C- 2829
02_AG.NG.01.PL0710          ---------------AC-----------TG......-----------------A--G--------------A-----G-----------------------------------------T--------------------------------------C----------- 1984
02_AG.NG.x.IBNG             ---------------AC---C---------......-----------------A-----------------A--------------------------------------------G--T--------------------------------------C----------- 2325
02_AG.SE.94.SE7812          ---------------AC---C---------......--A--------------A----------G-------------------------------A----------------G---A-T-------G------------------------------C----------- 2172
02_AG.SN.98.MP1211          ---T----GCAT---AC---C------G--......-----------------A--------------------------------------------------T-------G------T-----A--------------------------------C----------- 1997
03_AB.RU.97.KAL153_2        ---------------AC---C-------AG......-----------------A-------------------------------------------------------------------------------------------G------------------------ 2026
04_cpx.CY.94.CY032          -----------C---AC---G--A------......--C--------------C--G---------------------------------------A-------T-----------------C-----------------------------------C-----A----- 2166
05_DF.BE.x.VI1310           ---------------AC-------------......--T--------------A-----C----G---------------C---------------A-------------------G----------GG----------------------------------------- 2182
06_cpx.AU.96.BFP90          --G------------AC-----------A-......--A--------A-----A--------------------------------------------------T--------------T--------------------------------------------A--G-- 2828
08_BC.CN.97.97CNGX_6F       ---------------AC--C-------GAT......--A-----G--------A------A----------------C--------T--C------A------------GA--G--G-------G---G-----------------------G-----C-----A----- 1980
09_cpx.GH.96.96GH2911       ---------------AC----------G--......-----------------A--------------------------------T-----------------T-----------G-----C-----G-----------------------------------A----- 2002
10_CD.TZ.96.96TZ_BF061      ------------C--AC-----------A-......--A--------------A--------G-GG--------------------T---------A-------------------------------G----------------------------------------- 2179
11_cpx.GR.x.GR17            ---------------AC------C------......--A--------------A--------------------------------T--C--------G-----------------------C--------------------T-----T--G-----C-----A----- 2102
12_BF.AR.99.ARMA159         ---------------AC------A------......--A--------A-----A--------------------------------------------------------------------------------G-----------------------------A----- 2805
13_cpx.CM.96.1849           ---------------AC------C----A-......--A--------------A-----------------A--------------T--C--------G-----T---------------------------------C----T-----T--G-----C-----A----- 2205
14_BG.DE.01.9196_01         ---------------AC-------------......--T-------G------A-----------------A-----------------C------A-------------------------------------------G--------------G--C-----A----- 2317
15_01B.TH.99.99TH_MU2079    ---------------AC-----------A-......---------G-------A-----C------------------------------------A-------T-----------G-----C--C--------G--------------------G--C-----A--G-- 2199
16_A2D.KR.97.97KR004        ---------------AC-----------A-......--A-----------G--A--G--------------------------------C--------G-----T-----------------C---------------C-------------------C----------- 2163
17_BF.PE.02.PE02_PCR0155    ---------------AC------A------......--A--------A-----A--------------------------------------------------------------------------------G-----------------------G-----A----- 1990
18_cpx.CU.99.CU76           --G------------AC------A------......-----------------A-----------G-----------------------C--C---A-------------------------------------------------------------------A----- 2105
19_cpx.CU.99.CU7            ---------------AC-------------......--T--------------A--------C--------------G--------T---------A-------------------------------G-----------G----------------------------- 2026
20_BG.CU.99.Cu103           ------------------------------......--A--------------A-----------------------------------C-----------------------G-----------------------------------------G--C-----A----- 2266
21_A2D.KE.91.KNH1254        ---------------AC-------------......--------------G--A----------G-------------------------------A-------T-----------G--------C--G-----------------------------------A----- 2005
22_01A1.CM.01.01CM_0001BBY  -----------T---AC-----------A-......-----------------A-----------G-----------G------------------A-------T-----------------C-----------------------------------C-----A----- 2008
23_BG.CU.03.CB118           ------------------------------......--AT-------------A-----------------------------------C--------------------------GA----------G-----------------------G----------------- 2243
24_BG.CU.03.CB378           ------------------------------......--A--------------A-----------------------------------C--------------------------------------G-----------------------G----------------- 2225
24_BG.CU.03.CB471           ------------------------------......--A--------------A-----------------------------------C--------------------------------------G-----------------------G----------------- 2240
25_cpx.CM.01.101BA          ---------------AC------C----AT......--T--------------A----------G---A--A------------------------A-------T-----------G-----------------------------------------A-----A----- 2008
27_cpx.CD.97.97CDKTB49      --------G------AC-------------......--A--------------A----------G---------------------T--C--------G-----------------------------------G--------T-----T--G-----C-----A----- 2008
28_BF.BR.99.BREPM12609      -----------------------C------......--A-----------------------------------------------T--------------------------------------------------------------------G-------------- 2166
29_BF.BR.02.BREPM119        ------------------------------......--A--------------A-------------------------------------------------------------------------------------------------------------------- 1974
31_BC.BR.02.110PA           ---------------AC--C-------GAT......--A-----G--------A-----GA-------------------------T--C--------------------A--G--G-----------G-----------------------G-----C-----A----- 2234
32_06A1.EE.01.EE0369        -------G-------AC-------------......--A--------A-----A------------------------------------------A-------T-----------G--T--------------------------------------------A--G-- 2435
33_01B.MY.05.05MYKL007_1    ---------------A--------------......--A-----------------------------C-----------------------------------------------G---G----------------G-----------T--------------A----- 2179
34_01B.TH.99.OUR2478P       ----------------C----------G--......--A--------------A------------------------------------------A-------T-----------G-----C--A--------G--------------------G--C-----A----- 2002
35_AD.AF.05.05AF095         ---------------AC-------------......--A--------------A-----------------------------C------------A-------T-----------------C--------------------G-----T--------C-----A----- 1993
36_cpx.CM.00.00CMNYU830     ------------C--AC----------G--......-----------A-----A--------------------------------------C----------------G---------T--------------------------------------C-----A----- 2005
37_cpx.CM.00.00CMNYU926     ---------------AC---------C-AT......--A--------------A----------G-G-------------------T--C------A----------------------------------------------------------G--C-----A----- 1990
38_BF1.UY.05.UY05_4752      ---------------AC------A------......--A--------------A-----------------------------------------------------------------T--C----------------------------------------------- 2211
39_BF.BR.03.03BRRJ103       ---------------------------G--......TTTT-----G-------A----------------------------------------------------------------G-G-----------------G---C--------------------------- 2289
40_BF.BR.04.04BRRJ115       ---------------AC------A------......--AT----G--A-----A--------------------C-----------------------G--------------G---A---------G----------A--AC---------------------A----- 2352
42_BF.LU.03.luBF_05_03      ---------------AA------A------......--A-----G--A-----A--------------------------------------------G-----T-----------------------------G-----------------------------R----- 2338
43_02G.SA.03.J11223         ---------------AC-----------TT......-----------------A-----------------------------------C------A----------------------------------------G--G-----------R--G--C-----A----- 2325
U.CD.83.83CD003             ---------------AC-----------CT......--A--------------------------------------------------C--------------T-----A--G--G-----------G-----------------------------C-----A--G-- 2295
U.CD.90.90CD121E12          --G------------AC-------------......--A--------A-----A-----------------------------------C--------------T-----A--G--G-----------------------------------------C-----A----- 1999
U.GR.99.GR303               ---------------AC-------------......-----------------A--------------------------------------------------T-----------G-----C----G------G--------------------G--C-----A----- 2074
U.NL.95.U_NL_95_H10986_D1   -------G-------AC---C---------......--A--------------A-----------------------------------C--------G-----------------------------------------------C--T--------------A----- 2685
CPZ.CD.90.ANT               ------G-G--CC--AA----------GAT......A--T-A---GCA---AAT--G-----T-G---A-----A--T--C-----------A---A-T-----A-----A--G-------C-T-A-----------GC----T------------T-A-----A----- 2239
CPZ.CM.01.SIVcpzCAM13       -----T-----C--GAC------A---CAG......--A-----------------G--A--T-----A--A--------C--------------TA-T-----A--T--A--G-----T---T-A------------------------------T-A-----A----- 2380
CPZ.CM.05.SIVcpzEK505       --------G--C---AC-------------......--A--------A-----A--------C-G---A-----------------T---------A-T-----T-----A--G-----T--C--------------------------------GT-A--C--A----- 2372
CPZ.CM.05.SIVcpzLB7         -----T-----T---AC------C------......--A------C-------A----------G---A-----------C-----T--------T--G-----T-----A--G-----T--C--A-----------------------T------T-A-----A----- 2355
CPZ.CM.05.SIVcpzMB66        -------G-------AC-------------......--A--------------A----------G-G-A-----------------------C---A-T-----T-----A--G-----T--C--------------------------------GT-A-----A----- 2346
CPZ.CM.05.SIVcpzMT145       ------G-G------ACT-----A----A-......--A--------------A--------------A-----A--------T--------A--TA-T-----A--T--------G-----C--C------------C-T--T-----------------------G-- 2342
CPZ.GA.88.GAB1              -----T-----C---AC----------CA-......--A--------------------A--G-----A--A-----------------------TA-T-----A--C--A--G--------------------------------C-----------A----------- 2861
CPZ.TZ.00.TAN1              -----TG-G------AC------A----AG......ACAT-------------A--------TGC-G----A--A-----C-----T--C--A--A--T--------T--------G--T-CCT----------------------------------G-----A----- 2444
CPZ.US.85.CPZUS             -----T--G------AC----------CCA......--A--------A-----A--G---A---G-G-A-----A--------T--------A--A--T-----T--C--A--G-----T--C-----------------------C--T------T-A-----A--G-- 2860
CPZ.CM.x.SIVcpzDP943        -----C-----T---AC----------CAG......--A--------A-----A--G-----------A-----A--------T--T-----A--A--T-----------A--G-----TG--GAA--------------G--------T------T-A----GA----- 2386
CPZ.CM.x.SIVcpzMB897        --------------GAC-------------......-----------------A----------G---A-----A-----------T-----A---A-------T--T--------G--T--C--A--------------------C---------T-A-----A----- 2366
N.CM.02.DJO0131             ------G-G------AG-C--------G--......--A------C-------A--------T-G---A--------------------C------A-C-----T-----A--G-----T--C-----------------------------G---T-A-----A--G-- 2306
N.CM.04.04CM_1015_04        ------G-G------AG-------------......--A--------------A--------T-G---A-----------------T--C------A-T-----T----GA--G-----T--C--C------------------------------T-A-----A--G-- 2305
N.CM.04.04CM_1131_03        ------G-G--G---AG-------------......--A--------------A--------T-G---A-----------------T--C------A-T-----T-----A--G-----T--C---------------------------------T-A-----A--G-- 2304
N.CM.95.YBF30               ------G-G------AG-------------......--A--------------A--------T-G---A--------G--------T--C------A-T-----T-----A--G-----T--C-----------------------------G---T-A-----A--G-- 2394
N.CM.97.YBF106              ------G-G------AG-------------......--A--------A-----A--------T-G---A-----------------T--C------A-T-----T-----A--G-----T--C-----------------------------G---T-A-----A--G-- 2392
O.BE.87.ANT70               ------G----C--GAC--C---A---CAG......--A------C-A-----A----------G---A--A-----------T--T-----A--TA-C-----T-----A--G-----TG-------------------G--------T--G---T-A----------- 2854
O.CM.91.MVP5180             ------G-----C--ACT-C---A---CA-......--A------C-------A-----C----G---A--A-----------T--T-----A--TA-T-----T-----A--G-----T--C-----G-----------G-----C---------T-A-----A----- 2829
O.CM.96.96CMABB637          ------G----C---AC--C---A---CAG......--A------C-A-----A----------G---A--A-----------T--T--Y--A--TA-C-----T-----A--G-----T--------G-----------G-----C-----G---T-A-----A----- 2284
O.CM.98.98CMA104            ------G----C--GAC--C---A---CAG......---------C-A-----A----------G-G-A--A-----------T--T-----A--TA-T-----T-----A--G-----T--------G--------------------T------T-A--C----A--- 2286
O.CM.99.99CMU4122           ------G----C--GAC--C---A---CAG......--A------C-A-----A----------G-G-A--A-----G-----T--T-----A--TA-C-----T-----A--G-----T----G---G--------GC-G-----C--T--G---T-A--C-------- 2285
O.FR.92.VAU                 ------G----CC-GAC------A----A-......--AT-----C----G--A----------G---A--A-----------T--T-----A--TA-C-----T-----A--G-----TG-A-----G-----------G-----C---------T-A--C--A----- 2366
O.SN.99.SEMP1299            ------G----C--GAC--C---A---CAG......--A------C-A-----A----------G---A--A-----------T--T-----A--TA-C-----T-----A--G-----T--C-----G--------GC-G-----C--T--G---T-A--C-------- 2853
O.US.99.99USTWLA            ------G----C---ACT-C---A---CAG......--A--R--GC-A-----A----------G---A------R-------T--T-----A--TA-C-----T-----A--G--R--T-----C--G-----R--RY-G-----C--------GT-A--C--R----- 2283
```

```
B.FR.83.HXB2          TCAAGACTTCTGGGAAGTTCAATTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGTATAAACAATGAGACACCAG 2970
Pol                   __Q__D__F__W__E__V__Q__L__G__I__P__H__P__A__G__L__K__K__K__K__S__V__T__V__L__D__V__G__D__A__Y__F__S__V__P__L__D__E__D__F__R__K__Y__T__A__F__T__I__P__S__I__N__N__E__T__P__
A1.GE.99.99GEMZ011    ---G-----T-------------------------C-----A--G--T-------------G-----------------A--------G-------------------T--------AG------A----------------C---T-----C----C------------------ 2154
A1.KE.00.KER2008      A--------T-----------------G-----G--C--A--G--AC------G-----------------------A--------G--C--------C---------C-----A--A----T--A-----------------------------T----T-----A------- 2169
A1.KE.00.KNH1144      A--------T-----------------G-----G-----A--G--CC------------------------G--A--------G--C--------C-----C--T---C----------T--A--------------C-------------C------------------ 2175
A1.KE.00.KSM4024      A--------T-----------------------G-----A--G--CC----------------------------A--C-----G------------------C--T---------A----T--A--A-----------C--------A----C---------------G- 2166
A1.KE.00.MSA4069      A--------T-----------------------G-----A--G--CC-----------G----------------A--------G--------------------T-----A--AGT--T--A--A-----------C-------------C------------------ 2169
A1.KE.00.NKU3005      ---------T--------G--------------G-----A--G--C-----------------------------A--------G--C--------C-----C--------------AG---T--A-----------A-----C-------------C-----------G----- 2169
A1.RU.00.RU00051      ---G-----T-----------------------------A--G--T-----------------------------A-----A--G--------------------T---------A----T--A--------------C--T--------------------------A------- 2271
A1.RW.93.93RW_024     A--------T--------------G--G-----G--C--A--G--CC----------------------------A-----A--G--C-----------------T--------------T--A--A-------------------------C------------------ 2172
A1.SE.95.SE8891       -------------..............------------T-----T-G---------------------------A-----A--G--C-----------------T--G---A--A----T--A--------------G--C------------------------------ 2159
A1.SE.95.UGSE8131     ---------T-----------------G-----------A--G--C-----------------------------A-----A--G--C---------------------------AG---T-----------------C-------------C---T-----A------- 2342
A1.TZ.01.A173         ---------T-----------G-----------G-----A--G--CC-------------------------T-A--------G--C-----------------TC--------AG---T--------------G--C-------------C------------------ 2169
A1.UA.01.01UADN139    ---G-----T-------------------------C-----A-----T-----------------------------A--------G--------------------T---------A-------A--------------C--T-----A-----------C---------- 2169
A1.UG.92.92UG037      ---------T-----------------------G-----A--G--C-----------------------------A--------G--C---------------------------AG---T--A--------------C-------------C------------------ 2986
A1.UG.99.99UGA07072   ---G-----T---------------A-------------------T-----------------------------A--------G--------------------T---------A-T--T--A--------------C------------------------------- 2169
A2.CD.97.97CDKTB48    ---------------------G--------------C--A--G--A-------------------------T-A--------G--C-----------C------C--C----------------A-----------C------------------------------- 2308
A2.CY.94.94CY017_41   ---------------------G-----------------A-----A---------------G--------------T--------G--C-----------C---------C-------------A--A-----------C-------------CC---------------- 2327
A.SN.01.DDI579        A--G-----------------------------G-----A--G--T--G-------------------------------------G--C-----------------T---------AG---T--A--------------C--------C----C------------------ 2168
A.SN.01.DDJ369        ---G-----------------------G-----G-----A--G--T--G---C---------------------------------G--C-----------------T----------G---T--A-----------G--C-------------C------------------ 2171
A.SN.96.DDJ360        A--G-----------------------G-----G-----A--G--T--G--------------------------A--------G--C-----------------T------AT-A----T--A--------------C-------------C------------------ 2168
A.CD.02.02CD_KTB035   ------------------C-----G--------------A-----T-----------------------------A--------G--C-----------------------------A-T--T--A-------------------T-----------C------------------ 2964
A.CD.97.97CD_KCC2     ------T--T--------------G--------------A----------G------------------------A--------G--C------------------------C-A----T------------------C--T-----------C------------------ 2966
A.CD.97.97CD_KTB13    ------T-----------------------------C--A-----------------------------------A--------G--------------------------------GT--T--A--A--------------T--------------------G------- 2174
B.AR.04.04AR151516    ------------------C---------------------------------------------------------------------------------------T---T------------------C------------------------C------------------ 2192
B.AU.87.MBC925        ---------------------------------------------------------------------------------------------------------------T--------------------------------------------C------------------ 2981
B.BO.99.BOL0122       ---------------------------------------A----------G-----------------------------------------------------------------------------------------------------------C------------------ 2190
B.BR.03.BREPM2012     ------T--------------G------------------------------------------------------------------------------------------------------------------------------------C------------------ 2395
B.CA.97.CANB3FULL     ----------------------------------------------------------------------------------------------------A-------------------A------------------------------------------------------- 2259
B.CN.05.05CNHB_hp3    ---------------------------------------G--------G--G-----------T--------C-----------------------C--G--C--T--------G-------------------------------------GG----------------- 2969
B.CO.01.PCM001        ---------T--------------------------------------------------------------------------G-----T--------------------------T--------A----------------------C--------------------- 2175
B.GB.83.CAM1          ---------------------------------------T------------------------------------------------------------------------A-------------------------------------------C------------------ 2971
B.GB.86.GB8           ------------------------------------C---T-G--A-----------------------------------------------C------------------A----A-----A----------------------------------------------- 2990
B.GE.03.03GEMZ010     ---------------------------------------T------------------------------------------------------C----------------CA----A--------------------C-------------C------------------ 2175
B.IT.05.SG1           -------------------------------------------------------------------------------------------N-N----------------------T--------------------------------------N------------------ 2772
B.JP.05.DR6538        ------------------------------------C--T-----------------------------------A------------------C--------A---------A----------A--------------C-------------G------------------ 2974
B.KR.05.05CSR3        ---------------------------------------T------------------------------------------------------------------------A--------------------------------------------------------- 2995
B.NL.00.671_00T36     ------------------------------------C--T--G------------------------------------C--------------------------------A-------------------------------------------C------------------ 2535
B.RU.04.04RU128005    ----------------------------------------T----R-------------------------------------------------------------A-----A--------------------------------------------------------- 2461
B.TH.00.00TH_C3198    ----------------------------------------------C-----------------C--------C--------------------C--C--------T-----A-------------------------------------------G------C---------- 2184
B.UA.01.01UAKV167     C--------------------------------------T------------------------------------------------------C----------------CA----------------C-----------------C---------------------- 2196
B.US.04.ES10_53       ------------T---------------------------------------------------------------------A----------------------------C---------------------------------------------------------- 2979
B.US.99.PRB959_03     ------------------------------------------G--------------------------------------------------------------------CC----T--------------------------------------------------------- 2190
C.AR.01.ARG4006       ---G-----T--------C--------G--------C--A--------------------------G-----------C-----G-----------------A--T--------------T-----A-----------C-----------------------A------- 2154
C.BR.04.04BR013       ------T--T-----------------G--------C--A---------------C----------G-----------------G-----------------A--T--G------AG---T-----A-----------C--------C----A----------------- 2434
C.BW.00.00BW07621     ---------T--------------------------C--A--------------------------G-----------------G--------------------TC--------A-T--------A-----------C-----------------------A------- 2323
C.CN.98.YNRL9840      ------T--T--------------------------C--A--------------------------G-----------------G--------------------T---T------G----------A-----------C-----------------------A------- 2154
C.ET.02.02ET_288      ---------T-----------------G--------C--A--------------------------G------T-A-----A--G-----------C--------T---------AG----------A--C--------C-------------C---------A------- 2172
C.GE.03.03GEMZ033     ---------T--------------------------C--A--------------------------G-----------------G--------C--------C--T---------------T-----A-----------C-----------------T-----A------- 2151
C.IL.99.99ET7         G--------T--------------------T-----C--A--------------------------G-----------------G------------------------------AGT---------A-----------------------------------A------- 2151
C.IN.99.01IN565_10    ------T--T--------C-----------------C--A--------G-----------------G-----------------G--------------------T---T----------------A-----------C-----------------------A------- 2351
C.KE.00.KER2010       ---------T-----G--------------------C--A--------G-----------------G-----------------G--------------------T-----A--A----------A-----------C-------------G---------A------- 2151
C.MM.99.mIDU101_3     ------T--T----------T---------------C--A--------------------------G-----G-----------G--------------------T---T----------------------------C-----------C-----------A------- 2321
C.MW.93.93MW_965      ---------T--------------------------C--A--------------------------G-----G-----------G--------------------T---------A----------A-----------C------------------------------ 2157
C.SN.90.90SE_364      ------T--T--------------------------C--A--------------------------G-----------------G--------------------T----------G---------A-----------C-----------------------A------- 2139
C.SO.89.89SM_145      C--------T-----G--------------------C--A--------------------------G-----------------G-----------C--------T----------G---------A-----------C-------------C---------A------- 2191
C.TZ.02.CO178         C--------T--------------------------C--A--------------------------G-----------------G-----------------C-------------G---------------------C-------------C------------------ 2151
C.UY.01.TRA3011       ------T--T-----------------G--------C--A--------------------------------------------G-----------------A--T---------------T-----A-----------C-------------C---------A------- 2142
C.YE.02.02YE511       ---------T--------------------------C--A--------------------------G-----------------G--------------------A------A----A---------A-----------C-----------------------A------- 2145
C.ZA.04.04ZASK164B1   ---------T-----G-----------------G--C--A---------------------------------T----------G-----G--------------T---------AG----------A--C--------C-----------------------A------- 2384
C.ZA.05.05ZAPSK240B1  ---------T-----G--------------------C--A--------------------------G-----------------G-----------------G--T----------G---------------------C--------C---G-----------A------- 2374
C.ZM.02.02ZM108       ---------T--------------------------C--A------------C-------------G---------A-------G--------C-----------T--------G-----------A-----------C-----------------------A------- 2939
D.CD.83.ELI           ------T--------------------------G-----T------C-G---------------------------------------------------------------------T--T-----A----C--C---------T------------------------- 2516
D.CM.01.01CM_0009BBY  ---------------------------------G------------C-G--------------------------------A-----------C------------------A----T--T--------C--------C-------------C------------------ 2169
D.KE.01.NKU3006       ----------------------C----------------T------C-------------------------------------------------------------------------T--A--------------C-------------C---T--------------- 2175
D.KR.04.04KBH8        ---------------------G-----------G--C--T------C-------------------------------------------------C-----------------------T--A---------C-----C--------------C----------------- 2929
D.TD.99.MN011         ---------------------------------G-----T------C-G-----------------G--------A---A-A--------------------------------C---T--T--A--A--C--------C--T---------------------------- 2194
D.TZ.01.A280          ------T---------------C----------G-----T------C----------------------------T-------------------C------------------A----T--T--A--------------C-----------C-C----------------- 2171
D.UG.99.99UGK09259    ----------------A-----C----------------T------C----------------------------A--------------------------C------C-------T--T-----A-----------C-------------C---------------G- 2175
D.YE.01.01YE386       ----------------------C----------------T------C----------------------------A--------------------------C------C-------A--T--A--A-------C-----C-------------C------------------ 2169
D.YE.02.02YE516       ---------------------------------------T------C-G-----------------G-----------------------C--------------T---------AGT--T--A-----C--------C-------------C------------------ 2181
D.ZA.90.R1            ---G--T--T--------A--G-----------G--C--T------C-G---------------------------------------------------------------TG---------T--------------------------------------------G---- 2316
F1.AR.02.ARE933       ------T--T-----G-----------------------T-----A--------------G-----------------------G------------------------------G--T-----------C--------C------------G-C---------------- 2264
F1.BE.93.VI850        ------T--------G-----------------------T--------------------G-----------------------G---------------------------A-G--T----A------C--------C------------G-C-------------T- 2307
F1.BR.01.01BR125      ------T--T-----G-----------------G-----T------C-------------G-----------------------G------------------------------G--T-----------C--------C------------G-C---------------- 2415
F1.ES.x.P1146         ---G--T--T-----------------------------T--------G-----------G-----------GT----------A---------------------------A-G--G-----A----T----------C-----------C-C---------------- 2247
F1.FI.93.FIN9363      ------T--T-----G-----------------------T--------------------G-----------------------G---------------------------A----T-----------C--------C------------G-C---------------- 2299
F2.CM.02.02CM_0016BBY ------T--T-----G-----G--------------C--T-----T---------------------------T----------G---------------------------A-G--A-----------C--------C--------------T---------------- 2160
F2.CM.95.MP257        ------T--T-----G--------------------C--T-------------------G------------------------G---------------------------A----G-----------C-----G--C--------------C---------------- 2179
F2.CM.97.CM53657      ------T--T-----G--------------------C--T-----------------------------------A--------G---------------------------A-G--G-----------C--------C--------------C----------G---- 2160
G.BE.96.DRCBL         ---------------G--C-C-------------T------GG---------C------G---------------A--------G--------------------T---------A----T--A--A-----------C--T----------C---T------------- 2934
G.CM.01.01CM_4049HAN  ------T--------G--C---------------T-------G-----G----------G------------------C-----A--------C--------------------G-----T--A--A-----------C--T--------------T------------- 2172
G.CU.99.Cu74          ---------T-----G--C---------------T--C----G--------------------------------A--------G---------------------------A-------T--A--------------C--T-----------C-C---T------------- 2574
G.ES.99.X138          ---------------G--C---------------T-----T-G--------------------------G--G--A--------G--------C--------G------------AG---T--A--------------C--T-----------C---T------------- 2411
G.GH.03.03GH175G      ------------------C---------------T-------G-----G----------G---------------A--------A--------C---------------------AG---T--A--------------C--T-----------------T------------- 3010
G.KE.93.HH8793_12_1   ------------------C---------------T-----T-G--------------------------------A--------G-----C-------------------------AGT--T--A-----------A-----C--T-----------C---T------------- 2371
G.NG.01.01NGPL0669    ------T--------G--C--------T------T-----T-G--A-----------G-----------------A--------A---------------------------AG---T--A--A-----------C--T-----C---------T------------- 2175
G.PT.x.PT2695         ---------------G--C---------------T-----T-------------------------------G--A-----A--G--------C--------G------------AG---T--A--------------C--T-----------C---T------------- 2912
G.SE.93.SE6165        ---------------G--C---------------T-----T------------------------------------A--------G------------------------------------T--A--------------C--T--------------T------------- 2367
H.BE.93.VI991         ------------------A--G-----------------A-------------------G-----T------------------G-G---------------C--T---C----------------------------C-------------C------------------ 2359
H.BE.93.VI997         ---------------------G-----------------A-------------------------T------------------G-----------------C--T-----A--------------------------------------------------------- 2294
H.CF.90.056           ---------------------G--------------C--A-------------------------T------------------G-----------------C--T-----A----A-----A--------------C------------------------------ 2317
J.CD.97.J_97DC_KTB147 ---------------------------------------A-------------------G------------------------A-----------------C--T---T-------------------C--------C--T---------------------------- 2164
J.SE.93.SE7887        ---------------------G-----------------A--------------------------C-----------------G-----------------C--T---T-------T--------A-----------C--T---------------------------- 2284
K.CD.97.EQTB11C       --C------------------------G-----------A-----------------------------------------A--G-----------------C--T-----A----T---------------------C--T-----------T---------------- 2166
K.CM.96.MP535         ---------------------------------------A-----A--------------------------------------G-----------------C--T-----A--------------------------C--T-----------T--------A------- 2166
```

```
B.FR.83.HXB2                TCAAGACTTCTGGGAAGTTCAATTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGTATAAACAATGAGACACCAG 2970
Pol                         __Q__D__F__W__E__V__Q__L__G__I__P__H__P__A__G__L__K__K__K__K__S__V__T__V__L__D__V__G__D__A__Y__F__S__V__P__L__D__E__D__F__R__K__Y__T__A__F__T__I__P__S__I__N__N__E__T__P__
01_AE.CF.90.90CF11697       ---G-----T------------------G----G-----A-----T-----------------------------A--------A--------------------T---T------G---T--A--------------C-----------------------A------- 2919
01_AE.CN.05.FJ051           ---G-----T-----------------------G-----A-----T-----------------------------A-----A--A--------------------T---------AG---T--A--------------C------------------------------- 2980
01_AE.CN.06.FJ054           ---G-----T-----------------------G-----A-----T--G--------------------------A--------A--C--------------C--T---------AG---T--A--------------C--A---------C------------------ 2995
01_AE.HK.x.HK001            C--G-----T------A-------------------C--A-----T-----------------------------A--------A--------------------A------A--Y-T--T--A--------A-----C------------------------------- 2332
01_AE.TH.01.01TH_R2184      ---G-----T-----------------C-----G-----A-----T-----------------------------A---A-A--A------------------------------AG---T--A--------------C-------------C----------------- 2169
01_AE.TH.02.OUR769I         ---G-----T-----------------------------A--------------------------G--G-----A--------A--------------------T---------AG---T--A--------------C------------------------------- 2169
01_AE.TH.90.CM240           ---G-----T-----------------------G-----A-----T-----------------------------A--------A--------------------T---------AG---T--A--------------C------------------------------- 2544
01_AE.US.00.00US_MSC1164    ---G--T--T-----------------------G-----A-----T-----------------------------A--------A--------------------T---------A----T--A--------------C------------------------------- 2175
02_AG.CM.02.02CM_4082STN    ---------------G--C--------------------T--G--AC----------------------------A-----A--G--------------G-----T-----CA-------T--A-----C--------C--T---------G----T------------- 2169
02_AG.EC.x.ECU41            ---------T-----G--C--------------T--------G--A-----------------------------A--------G---------------------------A----------A--------------C--T---------G----T------------- 2066
02_AG.GH.03.GHNJ196         ---------------G--C--------------T--C--A-----A-----------------------------A--------G------------------------C----------T--A--------------C--T---------G----T------------- 2999
02_AG.NG.01.PL0710          ---------------G--C-----------C--T-----T--G--A---------------------------T-A--------G---------------------------A-------T--A--------------C--T---------G----T-----A------- 2154
02_AG.NG.x.IBNG             ------T--------G--C--G-----------T--------G--A--------------------------G--A--------G---------------------------A-------T--A-----------------T---------G----T------------- 2495
02_AG.SE.94.SE7812          ---------------GA-C--------------T--------G--A-----------------------------A-----A--G---------------------------A-------T--A--------------C--T---------------------------- 2342
02_AG.SN.98.MP1211          ---------------G--C--------------T--------G--------------------------------A--------G-----------------G--A--------------T--A--------------C--T---------G----T------------- 2167
03_AB.RU.97.KAL153_2        ------------------------------------C--T-----------------------T------------------------------------------------C----------A----------------------------C----------------- 2196
04_cpx.CY.94.CY032          ---G-----------------G-----------G--C--A-----------G---------------------T----------G---------------------------CC---G-----------C--------C-------------CC---------------- 2336
05_DF.BE.x.VI1310           ------T--T--------------------------C--G--------------------G---A-------------------G---------------------------A-G--A-----------C--------C--------------C---------------- 2352
06_cpx.AU.96.BFP90          ---------T-----------------------------T--T-----G--G---------------------T----------G---------------A-----C--------A-A--T--A--------------C--T----------C---T-----A------- 2998
08_BC.CN.97.97CNGX_6F       ------T--T--------------------------C--A--------------------------------C-----------------------C--------T------A--------------------------------------G------------------ 2150
09_cpx.GH.96.96GH2911       ---G--T-----------------G--------------A--G--TC--------------------------T----------G-----------------C--T------A-------------------------C--T---------G-T--T-----------G- 2172
10_CD.TZ.96.96TZ_BF061      ---------T------------C----------------T------C--------------------------------------------------------------T----------------A-----------C------------------------------- 2349
11_cpx.GR.x.GR17            ----------------A-A--------------------A-----A--------------------------------------A-----------------C-TTC--------AG------A--------------C--T---------C------------------ 2272
12_BF.AR.99.ARMA159         ------T--T-----G--G--------------G-----T-----------------------------------A--------G--------C--------C-----------G-CT-----------C--------C---------T--G-C---------------- 2975
13_cpx.CM.96.1849           ---------------------------------------A-----T--------------------------------------G-----------------C--TC-----A-------------------------C--T---------G------------------ 2375
14_BG.DE.01.9196_01         ---------------G--C--------------C------T-G-----------------------------G--A--------G--------C--------G---------A--AG---T--A--------------C--T----------C---T------------- 2487
15_01B.TH.99.99TH_MU2079    ---G--T--T--------G--------G-----G-----A-----T-----------------------------A-T------A--------------------T---------AGT--T--A--------------C------------------------------- 2369
16_A2D.KR.97.97KR004        ---------------------G-----------------A-----A-------T---------------------A------A-G------------------------C-------------A--A--C----------------------C----------------- 2333
17_BF.PE.02.PE02_PCR0155    ---------T-----G--------------------------T---------------------------------------------------------------------A----A---------------------------------------------------- 2160
18_cpx.CU.99.CU76           ------T--T-----------G-----G-----------T--T------------------------------T----------G---------------------------CC---G-----------C--------C--------------C---------------- 2275
19_cpx.CU.99.CU7            C--------------------------------------T------C----G---------------------T-------A---------------------------------AG---T--A--------------C------------------------------- 2196
20_BG.CU.99.Cu103           ---------------G--C--------------C--------G--------------------------------A--------G-----------C--------------------T--T--A--A-----A-----C--T--------------T------------T 2436
21_A2D.KE.91.KNH1254        ------T-----------C--------------------T-----AC--------------------------------------------------------------AG---------T-----A--C----------------------C----------------- 2175
22_01A1.CM.01.01CM_0001BBY  ---G-----T-----------------------G-----A--G--A--------------------------G-----------G--------C-----------T---T------G---T--A--------------C-------------C---------A------- 2178
23_BG.CU.03.CB118           ---------------G--C--------------T--C-----G--T-----------------------------A--------G--------------------------------A--T--A--A-----A--------T--------------T------------- 2413
24_BG.CU.03.CB378           ---------------------------------------------------------------------------A--------G---------------------------A----A--T--A--------A-----C------------G----T------------- 2395
24_BG.CU.03.CB471           ---------------------------------------T-----------------------------------A--------G--------------------------------T--T--A--A-----A-----C--T--------------T------------- 2410
25_cpx.CM.01.101BA          ------T--------G--C--------------T--------G----------------G---------------A--------G-----G---------------------A----G--T--A--A-----------C--T----------C---T------G------ 2178
27_cpx.CD.97.97CDKTB49      ---------------------------------------A-----------------------------------A--------G-----------------C--TC--------A----------------------C--T---------G------------------ 2178
28_BF.BR.99.BREPM12609      ------------------------------------------------------------G---------------------------------------------------A----A---------------------------------------------------- 2336
29_BF.BR.02.BREPM119        ------------------C--------------------T-----A---C-------G-----------------------------------------------------CA----T---------------------------------------------------- 2144
31_BC.BR.02.110PA           ---------T-----------G-----G--------C--A-----------G--------------G-----------------G-----------------A--T------A--AG---T-----A-----------C-----------------T-----A------- 2404
32_06A1.EE.01.EE0369        ---------T-----------------------------G--T-----G--G---------------------T----------A-----------------C--T-----CA-------T--A-G------------C--T--------------T------------- 2605
33_01B.MY.05.05MYKL007_1    ---------------------------------------T-----------------G--------------G--A--------A-----------------C--T--------------T--A--------------C------------------------------- 2349
34_01B.TH.99.OUR2478P       ---G-----T-----------------------G-----A-----T----------G---G-----------C-----C-----------------C--------T------A--------------------------------------------------------- 2172
35_AD.AF.05.05AF095         ------T--T--------C--G-----------G-----A--G--T--------------------------------------G-----------------------------------T--A--------------C-------------C----------------- 2163
36_cpx.CM.00.00CMNYU830     ---------------G--C--------------T--------G--A-------------G---------------A--------G---------------------------A-------T--A--------C-----C--T--------------T------------- 2175
37_cpx.CM.00.00CMNYU926     ---------------G--A--------C-----C-----T--G--A-------------G---------------A--------G--C--------C--------T---------AG---T--A--------A-----C-------------C----------------- 2160
38_BF1.UY.05.UY05_4752      C--------------------------------------T--------------------G-----------------------------------------A-----------G--T-----------C--------C------------G------------------ 2381
39_BF.BR.03.03BRRJ103       ----------------------------------------------A----------C--G------------T----------A---------------A----A--------G--G-----------C--------C--T---------G-C---------------- 2459
40_BF.BR.04.04BRRJ115       A----------------------------------------G------------C-----G-----------------------G---------------------------A----A-----------C--------C------------C-C---------------- 2522
42_BF.LU.03.luBF_05_03      ---------T------------------------------------C-----------------------------------------------------------------A-------------------------------------C-C----------------- 2508
43_02G.SA.03.J11223         ---------T--------C--------C-----T--------G--------R--------------------------------A--------------------A---------AG---T--A--------A-----C--T----------CG--T------------- 2495
U.CD.83.83CD003             ------T-----------------------------C--A--------------------------------------------A-----------------C--G---------A-T--------A--------C--C-----------------T------------- 2465
U.CD.90.90CD121E12          ------T--------------G--------------C--A--------------------------------------------A-----------------C--G---------A-T--------A--------C--C-----------------T------------- 2169
U.GR.99.GR303               ---G-----------------------------G-----A-----T-----------------------------A--------G--C--------------C--T------A----T--T--A--A--C--------C--T--------C-C---------A------- 2244
U.NL.95.U_NL_95_H10986_D1   ------T--T--------G-----G--------------A--------------------------------G----G------G--------------------T------CC--GG--------A--C--------C--T-----------C---------------- 2855
CPZ.CD.90.ANT               A-----T--T-----GA-A-----------T--T-----A--C--A-----GC-A--G--------G------T----------A-----C-----C--CA-A---------C-G-----T--A--A-----A--T--C--A--T--A--CG-G-----------C---- 2409
CPZ.CM.01.SIVcpzCAM13       C-----T-----------A-----G--C-----C--C--A------------G----G-----------------A--------------------C--TTG---TC-----A-G-----------------A-----C--A--T------G----T-----A--T---- 2550
CPZ.CM.05.SIVcpzEK505       ------T--T--------A-----------C--------A------C----G--------------------G--A-----A--A-----T--C------TG-------------A-T--T--A--------A--T---------------G----T-----A-----T- 2542
CPZ.CM.05.SIVcpzLB7         ---------T-----------------------T--C--T--G--A-----G-------GG-----C------T-------A--A-----C--C------TGC---C--------A----T--A--A-----------C-----------------------A------- 2525
CPZ.CM.05.SIVcpzMB66        ---------T--------C--GC-------C--------T------------C-------------------TT-A-----A--G--C--T--C------TG----C-G------A----T--------C-----T-----------C---G------------------ 2516
CPZ.CM.05.SIVcpzMT145       ------A--T--------A---C-G--C--------C--A--T--------G--------------------CT----------A-----C---------TG---TC-T--CCC------T--A--A-T---------C-----T-----C-----T-----A--T---- 2512
CPZ.GA.88.GAB1              A--------T-----G--G--------C-----T--C--A--------------------------G------T-A-----A--A-----C--C--C--TTG----C-G---A----T--------A--------------A--T--C--------T--------C---- 3031
CPZ.TZ.00.TAN1              ------T--T-----GT-A--GC----------C-----G----------G---A-G---TATG--G--------------A--G-----C--C-----CA-----C-G--CC----------A--------A--T-----------C--C--C--T---A-C------- 2614
CPZ.US.85.CPZUS             A--------------------G-----C--------C--A-----A-----G-------G------------C--A-----A--G-----C--------TTG---AC-G---A-G--A--T--A--A--C--A-----------C------G-------------C---- 3030
CPZ.CM.x.SIVcpzDP943        A--G--------------C--------C--------C--G-----A--------------G--------------A--------A--C--T--------CTGC--AC--------AGT--T--A--A--C--A------------------G------------------ 2556
CPZ.CM.x.SIVcpzMB897        C-----T--T--------C--GC-------------C--T--G--AC-G-----------------------G-----C--A--G-----T---------TGC---C----------G--T--A--A--C-----C--C------------G----T------------- 2536
N.CM.02.DJO0131             C-----T--T--------A--GC-------T--------A-----A---C-G--------------------T--A-----A--A-----T---------TG---------------T--T--A--------A--C--C-------------C---T-----------T- 2476
N.CM.04.04CM_1015_04        C--------T--------G--GC-------T--------A-----A-----GC-------G-----------C--A-----A--A-----T---------TG---------------T--T--A--A-----A--T---------------G----T-----------T- 2475
N.CM.04.04CM_1131_03        C--------T--------G--GC-G-----C--------A-----A-----GC-------G-----------CT-------A--A-----T---------TG---------------T--T--A--------A-TT---------------G----------------T- 2474
N.CM.95.YBF30               C-----T--T--------G--GC-------T--------A-----A-----GC-------------G-----TT-------A--A-----T---------TG------G--CA----T--T--A--------A--T--------------------------------T- 2564
N.CM.97.YBF106              C-----A--T--------G---C-------T--------A-----A-----GC-------------------TM-A-----A--A-----T---------TG----------A----T--T--A--------A--T--------------------T-----------T- 2562
O.BE.87.ANT70               A-----G--------G--A--GC----T--C--------G-GG--T--G--GC-A--GC----T--T-----CT-A-----A--A-----T-----C---TG---------CCC---T--T--A--A--------T--C--T--T------G-G--T--------C---- 3024
O.CM.91.MVP5180             A-----T--------G--G--------T--T--------A-GG--T-----GC-A-GGC----T--T-----CT-A-----A--A-----T-----C---TGC--T------CC------T--A--A--C-----C--C--T--T------G-G-----------C---- 2999
O.CM.96.96CMABB637          A--------------G--A--------C--T--------G-GR--C-----GC-A-GGC----T--T-----TT-A-----A--A-----T-----C---TG----C-------G--T-----A--A--------T--C--T--T------G-G-----------C---- 2454
O.CM.98.98CMA104            A-----T--------G--A--G-----T--C--------G-GG--T-----G--A--GC----T--T-----CT-A--------A-----T-----C---TG---TC----CCC---T-----A--A--------T--C--T--T------G-G-----------C---- 2456
O.CM.99.99CMU4122           A-----T--------G--A--------T--C-----C--G-GG--C-----GC-A--GC----T--T-----C--A-----A--A-----T-----C---TGC--------CCC---T-----A--A--------T--C--T--T------G-------------T---- 2455
O.FR.92.VAU                 A-----T-----------G--G-----T--C--------A-GG--T-----GC-A---C----T--T-----CT-A-----AAAA-----------C---TGC---C----CCC---G--T--A--A--------T--C--T--T------G-G--------A--C---- 2536
O.SN.99.SEMP1299            A-----T--------G--A--G-----T--C--------G-GG--T-----GC-A--GC----T--T-----C--A-----A--A--C--T-----C---TGC--------CCC---T-----A--A--------C--C--T--T------G-G-----------C---- 3023
O.US.99.99USTWLA            A--------T-R---R--R--GY--R-T--C--------A-RGR-T-----GC-A-GGC----T--T-----CT-A-----A--A-----T-----C---TGC--T------CC---T--T--A--A--C-----T-----T--Y------G-------------C---- 2453
```

```
B.FR.83.HXB2                GGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAG 3140
Pol                         G__I__R__Y__Q__Y__N__V__L__P__Q__G__W__K__G__S__P__A__I__F__Q__S__S__M__T__K__I__L__E__P__F__R__K__Q__N__P__D__I__V__I__Y__Q__Y__M__D__D__L__Y__V__G__S__D__L__E__I__G__Q_
A1.GE.99.99GEMZ011          ----C-----------------A--------------------------------------GT-------------------------C------TC-A----------A--------------------------C-----------C-----T--------------A 2324
A1.KE.00.KER2008            -A-----G-----A-----------------A-----------------------------GT-C--T--------------------C------TC-A----------A---A-------------T--------CC-------------A--T--------------- 2339
A1.KE.00.KNH1144            -A-AA--G-----------------------------------------------------G-----T--------------------C------TC-A----------A---A----------------------C-----------------T--------------- 2345
A1.KE.00.KSM4024            -A--C--G-----------------C-----------------------G-----------G--------------------------C------GC-A--------------A-------C-----------------A--------------T--------------- 2336
A1.KE.00.MSA4069            -A---------TGC-----------------------------------G-----------G-----------------------A---------TC-A----C-----A---A----------G-----------C--------------------------------- 2339
A1.KE.00.NKU3005            -A-CC--G-----------------------------------------------------G------------------C-------C------TT-A----------A--GA----------------------------------------T-----CT-------- 2339
A1.RU.00.RU00051            --G-C--------------------G-----A------------------T----------GT------------G------------C------TC-A----------A--------------------------C-----------C-----T--------------- 2441
A1.RW.93.93RW_024           -A--C--G-----------------------------------------G-----------G--------------------------C------TC-AG-------------A----------------------C--A--------------T--------------- 2342
A1.SE.95.SE8891             -A--C--------------------------------------------G-----------G----------T---------------C------GT------------A---A----------------------C-----------------TC-------------- 2329
A1.SE.95.UGSE8131           -A--C--G-----------------------------------------G-----------G---------------------------------TC-A-G--------A---A-------C--------------C-----------------T----------AA--- 2512
A1.TZ.01.A173               -AG-C--G--------T--------------------------------------------G--------------------------C------TC-A----------A---A----------------------C-----------------T--------------- 2339
A1.UA.01.01UADN139          -A--C-----------------A---------------------------T----------G--------------------------C------TT-A----------A--------------------------C-----------C-----T--------------A 2339
A1.UG.92.92UG037            -A--C--G-----------------------------------------G-----------GGCC-----------------------C------TC-A-------------------------------------C-----------------T--------------- 3156
A1.UG.99.99UGA07072         -A--C--G-----------------------------------------G-----------G--------------------------C------TC-A----------A---A----------------------C-----------------T----------A---- 2339
A2.CD.97.97CDKTB48          -A-----G--------------A--------------------------------------G------------------C----T--C-----GGC-AG---------G--G--C-----C--------------C-----------------T----------AT--- 2478
A2.CY.94.94CY017_41         -AG----G--------------A--------------------------------------G--------------G-----------C------TC-A-G---A----AT--A-C-----C--------------C-----------------T---------A-C--- 2497
A.SN.01.DDI579              -AG-C--G-----------------A-----------------------------------GTA------------------------C------TC------------A---A----------------------C---------------A-T--------------- 2338
A.SN.01.DDJ369              -AG-C-----------------A--A-----------------------------------GTA------------------------C------TC------------A---A----------------------C--------------G--T--------------- 2341
A.SN.96.DDJ360              -AG-C--G-----------------A-----------------------------------GTA------------------------C------TT------------A---A----------------------C-----------------T--------------- 2338
A.CD.02.02CD_KTB035         -T-----G-----------------------A--------------------T--------GT-------------------------C------TC------------T-----------C--------------C-----------------T--------------- 3134
A.CD.97.97CD_KCC2           -T-----G-----------------------A-----------C--------T--------G----C---------------------C------TC-A----C-----G--------------------------C-----------------T--------------- 3136
A.CD.97.97CD_KTB13          -T-----G-----------------------A--------------------C--------G--------------------------C------GC-A-------------G--------C--------------C-----------------T-----------A--A 2344
B.AR.04.04AR151516          ----------------T--------------------------------------------G-----------------------------------------C--------------T--------------------------------A-----------------A 2362
B.AU.87.MBC925              --------------------------------------------------------------T----------------------------------------------A------------------------------------------------------------ 3151
B.BO.99.BOL0122             ----C--------------------G------------------------------------T----------------A-----------------G------------------------------------------------------------------------ 2360
B.BR.03.BREPM2012           -A-----------------------------------------------------------------T-----------T-----T--------------G--------------------------------------------------------------------A 2565
B.CA.97.CANB3FULL           --------------------------------------------------T-------------------------------------------------------------------------G-----------C--------------------------------- 2429
B.CN.05.05CNHB_hp3          ----C------------G--------------------------------------------T----------------------------------------------------------------------------------------------------------- 3139
B.CO.01.PCM001              --G-----------------------------------------------------------T----------------------------------------------------------------------------------------------------------- 2345
B.GB.83.CAM1                -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3141
B.GB.86.GB8                 ----------------------------------------------------------------------------------------C--------------------A------------------------------------------------------------ 3160
B.GE.03.03GEMZ010           -A--------------------A---------------------------------------------------------------------C----G-----------------------C---------------------------------C-------------- 2345
B.IT.05.SG1                 ---------------------------------------------------------G-----------------G---------------------------------------------------------------------------------------------- 2942
B.JP.05.DR6538              --------------------------------------------------------------T----------------------------------------------A-----------------T-----------A------------------------------ 3144
B.KR.05.05CSR3              -------------------------------------------------------------------------------T---------------------------A-------C--------------------------------------------G--------- 3165
B.NL.00.671_00T36           ----------------T-----------------------------------------------------------------------C--------------------------------------------------------------------------G------ 2705
B.RU.04.04RU128005          ----------C-----------A---------------------------T----------G--------------------------------------------------------------G-----------------------G--------------------A 2631
B.TH.00.00TH_C3198          ----------------------A------------------------------------------------------------------------------------------A-------------------------------------------------------- 2354
B.UA.01.01UAKV167           -A-------------------------------------------------------------------------G---------------------------------------------C------------------------------------------------ 2366
B.US.04.ES10_53             ----------------------A---------------------------------------------------------------------------------------C------------------------------------------------------AA--- 3149
B.US.99.PRB959_03           -A---------------------------------------------------------------------------------------------G-------------A---------------------------C-------------------------------T 2360
C.AR.01.ARG4006             -------------A--T---------------------------------T-G--------G-----Y-----------------A--C-----GGC------------A-----------------T-----------------------------------------A 2324
C.BR.04.04BR013             -A-----------A--T-----------------------------------G--------------------------A--------C------GC-A----------A-----------------T--------C--------------A-----------------A 2604
C.BW.00.00BW07621           -------G-----A--T-----T--------------------G-----------------G-------------G------------C-----G-C------------A--------------G--T--------C-----------------T--------------A 2493
C.CN.98.YNRL9840            -------G-----A--T--------------------------------------------G-------------G------------C-----GGC------------A-----C-----------T--------C-----------------------G--------A 2324
C.ET.02.02ET_288            -------------A--T-----------------------------------------T--G--------------------------C------GC-A----C-----A-----C--------------------C-----------------------G--------A 2342
C.GE.03.03GEMZ033           -------------A--T--------C-----------------------------------G--------------------------C-----G-C-AG---C--------G-----A--------T--------C--------------------------------A 2321
C.IL.99.99ET7               -------------A--T--------------A-----------------------------G-------------G------------C------GC------C-----A-----------------T--------C--A--------------T--------------A 2321
C.IN.99.01IN565_10          -------------A--T--------------------------------------------GGC------------------------C------GC------------A-----C-----------T--------C--------------------------------A 2521
C.KE.00.KER2010             -------------A--T-----------------T--------------------------G------------------C-------C------GC-A-G--------A-----C-----------T--------C--------------------------------A 2321
C.MM.99.mIDU101_3           -------G-----A--T--------------------------------------------G-------------G------------C-----GGC------------A-----C-----------T--------C-----------------T--------------A 2491
C.MW.93.93MW_965            -------------A--T-----------C--------------------------------G-------------G------------C-----GGC------------------------------T--------C--------------G-----------------A 2327
C.SN.90.90SE_364            -------------A--T---------------------------------T-------T--G------------------C-------C-----GGC------C-----A--------T--------T--------C--A--------------T---------AT---A 2309
C.SO.89.89SM_145            -------------A--T--------------------------------------------GT------------G------------C------GC-A----------A-----C-----------T--------C--------------------------------A 2361
C.TZ.02.CO178               -------------A--T--------------------------------------------G--------------------------C-----GGC------------A-----C-----------T--------C--------------------------------A 2321
C.UY.01.TRA3011             -------------A--T-----------------------------------G--------G-----T--------------------C-----GGC------------AG----------------T--------C-----------------T--------------A 2312
C.YE.02.02YE511             -------------A--T--------------------------------G-----------G--------------------------C------GC------C-----A-----------------T--------C--A--------------T---------AA---A 2315
C.ZA.04.04ZASK164B1         -------------A-----------------------------------------------G--------------------------C-----GGC------C-----------C--------------------------------C-----T-------------TA 2554
C.ZA.05.05ZAPSK240B1        -------------A--T--------------A------------------------------T----------------------T--C------GC-A----------A-----C-----C-----T--------C----------------------------A---A 2544
C.ZM.02.02ZM108             ----------------T--C--T--------A--G--------------------------GT-------------------------C-----G-C-A----C-----A-----------C--------------C-----------------T--G-----------A 3109
D.CD.83.ELI                 -------------------------------------------------G--------------------------------------C--------------------A--G--------------------------------------------------------- 2686
D.CM.01.01CM_0009BBY        ----------------------------------------G--C------------------T----T-----------------------------------------AC--A--------TC------------------------------T-----G--------- 2339
D.KE.01.NKU3006             -A-----------------------------A-----------------G-----------------------------------A-----------------------AG-G--------------------------------------------------------- 2345
D.KR.04.04KBH8              ----------------------------------------------------------------------------------------C--------------------A--------A------------------C-------------------------------- 3099
D.TD.99.MN011               ---C---------------------------A--G-----G-----G---------------T----T-----------------T--C--C-----------------AC----------------------------------------------------------- 2364
D.TZ.01.A280                ----------------------------------------------------------T--------------------------A--C--------------------A--------------------------------------------------G-----A--- 2341
D.UG.99.99UGK09259          -A-----G--------T--------------A-----------------------------------------------------A-----------------------AG-G--------------------------------------------------------- 2345
D.YE.01.01YE386             -A-C---------------------C-----------------------------------------------------------A-----------------------AG-G--C--T-----------------C--------------------------------- 2339
D.YE.02.02YE516             -------------------------------A--G-----G---------------------T-C--T--------------------C--------------------AC-------------G-----------------------------T--------------- 2351
D.ZA.90.R1                  -------------------------------------------------G------------T-------------------------C--------------------T-----------------------------------------------------------A 2486
F1.AR.02.ARE933             -A--------C--------------------A--G-----------------------T---T-------------------------C-------C-A-------------G--------C--------------------------G--------------------- 2434
F1.BE.93.VI850              -A-----G--C--------------------A------------------------------T-------------------------C-------TGA----C-----------C-----C--------------------------G--------------------- 2477
F1.BR.01.01BR125            -A-----G--C--------------------A-----------G------------------T----------------------T--C------GC-A---------G---------------------------C-----------G--------------------- 2585
F1.ES.x.P1146               -A--C-----C--------C--A--------A--------G---------------------T----------------------T--C------GC-A----------A-----------C---------------------------------------C-G------ 2417
F1.FI.93.FIN9363            -A-----G--C--------------------A------------------------------T-------------------------C-------C-AG---------------------C--------------------------G--------------------- 2469
F2.CM.02.02CM_0016BBY       -A-----------------------------------------G--------------------------------------------C------GC-AG---------G------G----C--------------------------G--------------------- 2330
F2.CM.95.MP257              -A-----------------------------------------G----------------------------T---------------C---------G----------A-----------C-----T---------C----------G--------------------- 2349
F2.CM.97.CM53657            -A-----------------------------------------G--------------------------------------------C------GT-A----------A--------T--C--------------------------G-----------------A--- 2330
G.BE.96.DRCBL               -A---------------.......A---------------------------------T--G--------------------------C-------C------------A-----G-----C-----------------A-----------A------------------ 3097
G.CM.01.01CM_4049HAN        -A--------------------------------------------------------T--G--------------------------C------GC-A----------A-----G-----C-----------------A------------------------------ 2342
G.CU.99.Cu74                ----------------T-----------------------------------------T--G--C--T--------------------C--------------------A----AA-----C-----T-----------A--------------------------A--- 2744
G.ES.99.X138                ----------------------------------------------------------T-----------------------------C-------T-A----------A-----G-----C-----T-----------A------------------------------ 2581
G.GH.03.03GH175G            -------------------------C--------------------------------T--G----------T------------A--C-------C-A----------------G-----C-----------------A--------G-----------------A--- 3180
G.KE.93.HH8793_12_1         ----------C---------------------------------------C-------T--G--------------------------C-------T-A----------A--G--G-----C-----T-----------A------------------------------ 2541
G.NG.01.01NGPL0669          -------------------------C-----------------------------------G--------------------------C-------C------------A--G--G-----------------------A------------------------------ 2345
G.PT.x.PT2695               ----------------------------------------------------------T-----------------------------C-----G-T-A----------A-----G-----C-----T-----------A------------------------------ 3082
G.SE.93.SE6165              --G-------------------------------------------------------T--G-------------G------------C------GC-A-T--------A--G--G-----C-----------------A------------------------------ 2537
H.BE.93.VI991               ----------------T-----------------------------C--------------G--------------G-----------C-----------------T--AG-GA----T--C--------------------------------------------A--A 2529
H.BE.93.VI997               ----------------T-----------------------------------------T--G--------------------------C-----------------T--A---A----T--C-----T-----------------------------------------A 2464
H.CF.90.056                 ----------------T--------------------------------------------G----------------------C---C------G----------T--A--G-----T--C-----------------------------------------------A 2487
J.CD.97.J_97DC_KTB147       ----C--------A-----C-----A-----A-----------------------------GT----------------------A---------GC-A----C-----AG----------C--G-----------C--------G--------T--G-------AA--- 2334
J.SE.93.SE7887              ----C--------------C-----A--------------------C-----------T--GT--------------------A-A---------G--AG---C-----A-----------C--G-----------C--------G-----------G-------AA--- 2454
K.CD.97.EQTB11C             -A--------------------A--A--------------------------------T---T-------------------------C-------GGA----------T--G---T-A--C--------------------C-----G--------------------A 2336
K.CM.96.MP535               --G----------------------A-----A------------------------------CA------------------------C-------T-A----------A--G-----A--C-----------------A--------G--------------------A 2336
```

```
B.FR.83.HXB2                 GGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAG 3140
Pol                          G__I__R__Y__Q__Y__N__V__L__P__Q__G__W__K__G__S__P__A__I__F__Q__S__S__M__T__K__I__L__E__P__F__R__K__Q__N__P__D__I__V__I__Y__Q__Y__M__D__D__L__Y__V__G__S__D__L__E__I__G__Q_
01_AE.CF.90.90CF11697        -A--C--------------------G-----------------------------------G--------------------------A-A----GC-A----------A--------------------------C-----------------T--------------- 3089
01_AE.CN.05.FJ051            -AG-C--------------------G-----------------------------------G--------------------------C-------T-A----------A-------------G------------C-----------------T--------------- 3150
01_AE.CN.06.FJ054            -A--C--------------------G-----------------------------------G-----T-------G------------C----A--T-A-------G--A--------------G-----------C-----------------T----------A---- 3165
01_AE.HK.x.HK001             -A--C--------------------G-----------------------G-----------G-----------------------A---------GC-A-C--------A---A----------------------C-----------------------------A--- 2502
01_AE.TH.01.01TH_R2184       -A--C--------------------G-----------------------G-----------G-------------G--------------------T-A----------A--G-----------------------C----------------------------A---- 2339
01_AE.TH.02.OUR769I          -A--C--------------------G-----------------------GT----------G-------------G------------C-------T-A----------A--G-----T-----------------C-----------------T--------------- 2339
01_AE.TH.90.CM240            -A--C--------------------G-----------------------G-----------G--------------------------C-------T-A----------A--G----------------A------C-----------------T--------------- 2714
01_AE.US.00.00US_MSC1164     -A--C--------------------G-----------------------------------G--------------------------C-------T-A----------A--------------------------C-----------------T--------------- 2345

02_AG.CM.02.02CM_4082STN     ----------------------------------------------C-----------T--GGCA--T--------------------C-A-----C-A-C--------G--G--G-----C-----T-----------A-----------A--------G--------- 2339
02_AG.EC.x.ECU41             -A-----------A--------------------------------------------T--GGCA-----------------------C------GC-A----------G-----G-----C-----T-----------A---A----------------G---A----- 2236
02_AG.GH.03.GHNJ196          -A--------------------------------------------------------T--GGCA-----------------------C-------C-A-C--------A--G--G-----C-----------------A--------------------G--------- 3169
02_AG.NG.01.PL0710           ----------------------------------------------------------T--GGCA--T--------------------C-------C-AGT--C-----G-----G-----------T-----------A--------------------G-----A--- 2324
02_AG.NG.x.IBNG              ----------C--------------------------------G-----G--------T--GGCA-----------------------C-------C-A----------G-----G-----C-----------------A--------------T--------------- 2665
02_AG.SE.94.SE7812           ----------------------------------------------------------T--GGCG-----------------------C-------C-A-------G--G-----G-----C-----T-----------A-----------A--------G-----A--- 2512
02_AG.SN.98.MP1211           ----------------------------------------------------------T--GGCA-----------------------C-------C-A----------AT-G--A-----C-----T-----------A--------------------G--------- 2337

03_AB.RU.97.KAL153_2         -------------------------------------------------------T-----------------------------------------------------G--------------------------------------------------G--------- 2366

04_cpx.CY.94.CY032           -A--------------------------------C---------------------------T-------------------------C------TTCA----C-----A-----C--A--C-----T--------------------G--------------------A 2506

05_DF.BE.x.VI1310            -A--C-----C--------------------A------------------------------T-------------------------C--------------------AG----------C--------------------------------------------A--A 2522

06_cpx.AU.96.BFP90           ----------------------A-----------------------------------T--G----------T---------------C-------T-A----C-----A-----G-----C-----------------A--------G--------------------A 3168

08_BC.CN.97.97CNGX_6F        --------------------------------------------------------------T-------------------------------------------------------------------------C------------------------------A-- 2320
09_cpx.GH.96.96GH2911        ----C--------------------A-----------------------------------G--------------------------C-------T-A----------A--------A--C--------------------------G--------------------A 2342

10_CD.TZ.96.96TZ_BF061       -A-----------------------------A-----------------G-----------------------------------------C---------------C-A--G--------------T------------------------------C----------C 2519

11_cpx.GR.x.GR17             -------------------C-----A-----------------------G--T--------G-----------------------A--A-------C------C-----AG-------------------------C--------G--G-----------------A--- 2442

12_BF.AR.99.ARMA159          -A-C---G--C-----------------------G-----------------------------C-----------G---C-------------------------G-----------------------------------------------T--------------- 3145

13_cpx.CM.96.1849            -------------A-----C-----G-----A--C--------------------------GT----------G-----------A--A-------C------C-----A-----G--T-----------------CC-------G------------------------ 2545

14_BG.DE.01.9196_01          ----------------------------------------------------------T-----------------------------C--Y----T-A----------A-----G-----C-----T-----------A-----G--------T--------------- 2657

15_01B.TH.99.99TH_MU2079     -A--C--------------------G-----A-----------------G-----------GT-------------------------C---------A-G--------A--G-----------------------C-----------------T--------------- 2539

16_A2D.KR.97.97KR004         -AG----G--------------------------------------------------------------------------------C--------------------A--------T----G------------------------------T--------------- 2503
17_BF.PE.02.PE02_PCR0155     ----------------------------------------------------------T--G--CC----------------R--A--------------------G--------------------------------------------------------------- 2330

18_cpx.CU.99.CU76            -A--C-----------------------------------------------------------------------------------C---------G----------A--------A--C-----------------A--------G--------------------A 2445

19_cpx.CU.99.CU7             -------------------------------------------------G--------T---------------------C----T--C--------------------A------------------------------------------------------------ 2366

20_BG.CU.99.Cu103            -------------------------------------------G----------------------------------------------------------------GA--------T-T------------------A--------------T--------------- 2606

21_A2D.KE.91.KNH1254         -------------------------------------------------G--------------------------------------C--------------------A--------------G--------------------------------------------- 2345
22_01A1.CM.01.01CM_0001BBY   -A--C--G-----------------G-----------------------G-----------G-------------G---------T--C------GC-A----------A--------------------------C-----------------T----------A---- 2348

23_BG.CU.03.CB118            -------------------------------------------G------------------T-----------------C----------------------------AT-------T--------------------A-----------------------T------ 2583

24_BG.CU.03.CB378            ----------------T---------------------------------------------T----------------------------C-----------------A--------T--------------------A------------------------------ 2565
24_BG.CU.03.CB471            ----------------T---------------------------------------------T----------------------------C-----------------A--------T--------------------A--------------T------C---A---- 2580

25_cpx.CM.01.101BA           ----------------------A-----------------------------------T--------------------T------A-C------GC------------A--------------------------C--------------------------------- 2348

27_cpx.CD.97.97CDKTB49       -A-----------------C-----A--------------------C--G-----------GT----------------------A--A----T--C------C-----A--------------------------C---C----G------------C----------A 2348

28_BF.BR.99.BREPM12609       ----------------------------------------------------------------C----------G---------------------------------------------------------------------------------------------- 2506

29_BF.BR.02.BREPM119         ---C------------------------G--A--G--------------G--------------------------------------------------------------------------G--TG------------------C------T--------------A 2314

31_BC.BR.02.110PA            ----------------------------------------------------------T----------------G---------------C-----------------AG----G-----------------------------------------------------A 2574
32_06A1.EE.01.EE0369         ----------------------A-----------------------------------T--G-----------------T--------C--------------------A--G--G-----------------------A--------------------G-----A--A 2775

33_01B.MY.05.05MYKL007_1     -A--C--------------------G-----------------------G--------------------------------------C-------T-A----------A--G--C--T-----------------C-----------------T--------------- 2519

34_01B.TH.99.OUR2478P        -A--C--------------------------------------------------------GT-------------------------C-----------------------G--------------------------A------------------------------ 2342

35_AD.AF.05.05AF095          -A-----------------------------A-----------------G--------------------------G--------A-----------------------A--------------G--T--------C-----------------T---------A-T--- 2333

36_cpx.CM.00.00CMNYU830      ----------------------------------------------------------T--GGC------------------------C-------C-A----------G-----G--T--C-----T-----------A---------------C----G--------- 2345

37_cpx.CM.00.00CMNYU926      -------G-----------------------------------G-----G-----------GT-------------------------C-----GGC------------A-----------C-----------------A--------------------------A--A 2330

38_BF1.UY.05.UY05_4752       ----------------------------------------------------------------C--T-----------------------------------------A--G--------------------------------------------------------- 2551
39_BF.BR.03.03BRRJ103        -A-----G--C--------------------A------------------------------T-----------------------------------G----------A-----------------TG----------------------A---C----------A--- 2629

40_BF.BR.04.04BRRJ115        -A--------------------A--------A------------------------------T----------------------T--C------GC-A-----------TCA--------C------G------------------C---------------------- 2692

42_BF.LU.03.luBF_05_03       ----------------------A-----------------------------------Y-----C--T-----------------R-----Y-----------------A---------------------------------R-------------------------- 2678

43_02G.SA.03.J11223          -A---------TGT-----------G--------------------------------T--G-----------------------A--C-------C-A----------A-----G-----C--G---G----------A---------------------C----A--- 2665

U.CD.83.83CD003              -------------------------------------------------------------G--------------------------C---------G----------G-----------C--------------------------C--------------------- 2635
U.CD.90.90CD121E12           --------------------------------------------------T----------G--------------------------C-----G-C-A----------A----C------C--G-----------------------C--------------------- 2339
U.GR.99.GR303                -------------------------------------------------------------G-----T-----------------A--C-------T-A----------A---A-C--A--C--------------------------G-----T--------------A 2414
U.NL.95.U_NL_95_H10986_D1    -A-----------A--------------------------------------------T---T----------------------A--C-A----GC-A----------AG-GA----------------------C--A-----G-----A-------------AA--- 3025

CPZ.CD.90.ANT                -A--A------TGT--T-----T--A-----A--C-----------------T--T--T---GCA--------C--G-------CA--A-----GG-TA-GT------CAG---AA--T------------------C-C--------------TA-G-----TACTGCA 2579
CPZ.CM.01.SIVcpzCAM13        ----C-----------T-----TT-A-----A-----------G--T-----T--C--T--------T-----------TC-G--A--C--C---GC-A----C--T--T--CAC-G----C--------------C--A--------G-----TC-T--C--TACC--A 2720
CPZ.CM.05.SIVcpzEK505        -T--------C--------------G-----A--C--------------------C--T--G--C-CT-----------A-----A--A-----G--GA-C--------AC----C--A-----G------------C-C--------C-----TC--------ACA--A 2712
CPZ.CM.05.SIVcpzLB7          -A--C--------A--------T--A--G-----------------------T-----T-----------------G--T--------A--C----G---G--C-----T-----A--T-----G--------------A--------G--A--T--GA------A---A 2695
CPZ.CM.05.SIVcpzMB66         -A--------C-----T--------A--G--A--G--------G--------------T--------------------T-----------C--------G--------A---A-------------T-----------A--------C---------A------A--TA 2686
CPZ.CM.05.SIVcpzMT145        -A-----------A--T-----A-----T--A--------G-----T--TT-C--C--T--G--------------G--------A--C-------C-A-GC-------A---A----T--C--G-----------C--------G--C--A--T---CCT----A-G-A 2682
CPZ.GA.88.GAB1               -AG----------A--T-----TT-G-----A-----------G--T---T-T--C--------C--T-----------TC----A--C--C---G--A-G-----T-----TAC---T--C--G-----------CC-A-----G--G-----TC-T-----T-AT--A 3201
CPZ.TZ.00.TAN1               ---AA----T------T--C---T-A--T--A--T-----G-----T--------T--T--G--C--T------------C----T-----C------G--C-C-----TG-GAC---T--C-----T---------C-T--CA----T--A--TC-TATGA---A-G-A 2784
CPZ.US.85.CPZUS              -A--------C-----T------T-A-----A-----------G-----------T-----G--C--T-----------TC----T---------------C-------TG-TA-A-----------T---------C-C--------G--A--TC--A-CT---AAA-- 3200
CPZ.CM.x.SIVcpzDP943         -A-----------A--T--------A-----A-----------G-----------T--T-----C--------------------A--A------C----G--------GT-GACA--T--C--G------------C-T--------C-----YC----------AG-A 2726
CPZ.CM.x.SIVcpzMB897         -A--A--G--C-----------A--C--G-----------------------T--T--T-----C--T-----T--G--TC-G--A----A---------G--------A-----C-----C--G------------C-A--------------T---C-T----A---A 2706

N.CM.02.DJO0131              -T--------C-----T------T-G--------------G--G-----------T--T------TCA-----------T-----A--A--------GA--C-------G---A-C--T-----G-----------CC-C-----G----------C----C---CA--A 2646
N.CM.04.04CM_1015_04         -T--------C-----T------T-A--------------G--------------T--T------TCA--------C--TY----A--A--------GA-GC-------A---A-C--T-----G-----------CC-C-----G--------------CC---CA--A 2645
N.CM.04.04CM_1131_03         -T--------C-----T------T-A------------G-G--------------T--T------TCA--------C--TC----A--A--------GA-GC-------G---A-C--T-----G-----------CC-C-----G-------------------CAA-A 2644
N.CM.95.YBF30                -T--------C-----T--------G-----A--C--------G-----------T--T--G----CA-----------TC----A--A--C---G-GA--C-------G---A-C--T--C--G-----------CC-C-----G---------------C---CA--A 2734
N.CM.97.YBF106               -T--------C-----T------T-A--------------G--------------T--T------TCA-----------TC----A--A-----G--GA--C-------G---A----T-----G-----------CC-C-----G-------------------CA--A 2732

O.BE.87.ANT70                -A--A-----C-----------C--C--G--A-----------T--------T------------TCA-----------TC----T--A--C--G-G-G-C--C-----AT---AA--T-G---G-----------CC-A-----------A--T---CCCC-GACAG-A 3194
O.CM.91.MVP5180              -AG-A-----C-----------C--C--G--A--G--------T--------C-----T--G---TCA--------G--TC----T--A---------AGC--C-----AG---AA--T-----G-----A-----C--A-----------A--T---CC-T-G-CAG-A 3169
O.CM.96.96CMABB637           -AG-G-----C-----------C--C-----A-----------G--------T------------TCA-----------TC----A--A--C--G---A-C--------A----AA--------G------------C-A-----------A--T---CCCT-G-TAG-A 2624
O.CM.98.98CMA104             -AG-A-----C-----------C--C-----A-----------G--------T-----------CTCA--------G--TC----C--A-----G--GG-C--C-----AT---AA--A-G---G-----------CC-A-----------A--T---CCCT-GACAG-A 2626
O.CM.99.99CMU4122            -C--A-----------------C--C--G--A-----------G--------T--------R---TCA---GT------TC-T--T--A--C--G---G-C--C-----AT---AA--T-G---G-----------CC-A-----------A--T---CCCC-GACAG-A 2625
O.FR.92.VAU                  --G-G-----C-----------T--C-----A-----------T--T-----T------------TCA-----------TC----T--A-----G---A-C--C-----AT---AA--T-----------------C--A-----------A--T---CC-C-G-CAG-A 2706
O.SN.99.SEMP1299             -A--A-----C-----------C--C--G--A-----------G--G-----T------------TCA-----------T-----T--A-----G---G-C--C-----AT---AA--T-G---------------C--A-----------A--T---CCCC-GACAG-A 3193
O.US.99.99USTWLA             -AG-A-----C-----T-----C--C--G--A--G-R------T--------T-----T------TCA-----------TC----T--A-----G---A-C--C-----A----AA--T-----R--------------A-----------A--T---CC-T-R--AG-A 2623
```

```
B.FR.83.HXB2            CATAGAACAAAAATAGAGGAGCTGAGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAAGACAGCTGGACTGT 3310
Pol                     _H__R__T__K__I__E__E__L__R__Q__H__L__L__R__W__G__L__T__T__P__D__K__K__H__Q__K__E__P__P__F__L__W__M__G__Y__E__L__H__P__D__K__W__T__V__Q__P__I__V__L__P__E__K__D__S__W__T__V
A1.GE.99.99GEMZ011      -----G---------------T-A---GCT-----AA----C-----GT----T--T-----------G--------------------T-----------A-----------------C-----------C---------A------T---T----------------- 2494
A1.KE.00.KER2008        ------G----------A--AT-A---GCT-----AC----C------T----T--------------G--------------------------------A-----G-----------C--G--------C----------A---------------A----------- 2509
A1.KE.00.KNH1144        ------G----------A--AT-A---GCT-----A-----C------T----T-----------G--G--------------------------------A-----G-----------C--G--------C----------A-------A-C-----A----------- 2515
A1.KE.00.KSM4024        ------G--------------T-----GCT-----A----AC------T----T--------------G--------G-----------------------A-----G-----------C-GG--------C------------------C------------------- 2506
A1.KE.00.MSA4069        ------G----GG----A---T-A---GCT---T-A-----C------T-A--T--------------G--------------------------------A-----G-----------C--G--------C--------------------------A----------- 2509
A1.KE.00.NKU3005        ------G-----G----A---T-A---A-T-----------C-----GT----T--------------G--------------------T-----------A-----------------C--G--------C--A-------A-T----------G--A----------- 2509
A1.RU.00.RU00051        ------------G--------T-A--GGCT-----A-----C------T--G-T--R-----T-R---G--------------------T-----------A-----------------C-----------C---------A----------T----------------- 2611
A1.RW.93.93RW_024       ------G----------A---T-A---GCT-----A-----C------T----T--------------G--------------------------------A-----G-----------C--G--------C--------------------G--G--A----------- 2512
A1.SE.95.SE8891         ------G-----G--------T-----GCT-----A-----T-----GT----T--------------G--------------------T-----------A-----------------C---------------------CA-------A-C-----A-A--------- 2499
A1.SE.95.UGSE8131       -----------------A---T----GGCT-----A-----C-----GT--TTT---------C----G--------------------T-----------A-----------------C--G--------C--A------AA---------------A----------- 2682
A1.TZ.01.A173           -----------------A---T-----A-T-----A----AA------T----T--C-----------G--------------A-----T-----------A-----------------C--G--------C-----------A--------C-----A----------- 2509
A1.UA.01.01UADN139      ---------------------T-A---GCT-----A-----C------T----T--------------G--------------------T-----------A-----------------C-----------C---------A----------T----------------- 2509
A1.UG.92.92UG037        -----------------A--AT-A---G-------A----AA------T----T-----------G--G--------------------T-----------A--------------------G--------C--A--G---A----------------A----------- 3326
A1.UG.99.99UGA07072     ------------G-G------T-----GCT-----A-----T------T----T--------------G--------------A-----T-----------A-----------------C---------------------CA-------A-C-----A----------- 2509
A2.CD.97.97CDKTB48      ------G--------------T-A--GGCT---T-A--A--A------T----T--------------G--------------------T--G--------A-----G--T--------C--G--------C---------AA--------------------------- 2648
A2.CY.94.94CY017_41     ------GT------------AT-A--GGCT--CT-A----AA------T--TAT-----------------------------------T-----------A-----G--T--------C-----------C---------AA--------------------------- 2667
A.SN.01.DDI579          --------------------A--A---GCT-----A----AA------T----T--------------G--------G-----------T-----------A-----------------C-----------C---------CA-T----------G-------------- 2508
A.SN.01.DDJ369          ---------------------------GCT-----A----AA-----------T--------------G-----------------------A--------A-----------------C-----------C---------CA-T-------C----------------- 2511
A.SN.96.DDJ360          -----------------------A---GCT-----A----AA------T----T--------------G--------G-----------T-----------A-----------------C-----------C---------CA-T------------------------- 2508
A.CD.02.02CD_KTB035     ---------------------------A-T-----A--A-AA-----GT----A-----------G--------A--G-----------T-----------A-----------------C-----------C---------CA---------C-----T--T-------- 3304
A.CD.97.97CD_KCC2       ------G--------------T-----GCT-----A--A-AA------T----A--------T--G--------------G--------T-----------G-----------C-----C-----------C---------CA------------G-------------- 3306
A.CD.97.97CD_KTB13      ------G----------A---T-A---G-T---T-A--A-AA------T----A-----T--T--G--G--------------C-----T--G--------G-----G-----------C-----------C-----C--CCAA-----------G-------------- 2514
B.AR.04.04AR151516      ------------G-------A------A-G--------A-A-------T-------C--------------------G-------------------------------------------------------------------------------------------- 2532
B.AU.87.MBC925          --------------------A---------------------------T----------------------------------------------------G-------------------------------------------------------------------- 3321
B.BO.99.BOL0122         --------------------A--A---------T------A------GTA------------T--------------------------------------------------------------------------C------------C-T----------------- 2530
B.BR.03.BREPM2012       --------------------A--T----------------A------GT-C--------------G-----------------------T-----------------------------------------------G---A---------------------------- 2735
B.CA.97.CANB3FULL       --------------------A----------------G--A-------T--TA------------------------------------------------G-----------------------------------------------------G-------------- 2599
B.CN.05.05CNHB_hp3      ------G-----G-------A------G-----T--------------T-------------------------------------------------------------------------------G--------------------------G-------------- 3309
B.CO.01.PCM001          ------------G----A--A------G------------A-------T--G-------------------C-----------------T-----------------------------C-----------------C---A-A-------------------------- 2515
B.GB.83.CAM1            --------------------A-----------------------------------------------------------------------------------------------------------------A------A-------------G-------------- 3311
B.GB.86.GB8             --------------------A--A--------------A--------GT---G------------------------------------T-------------------------------------------------------------------------------- 3330
B.GE.03.03GEMZ010       --------------------A------A----------A--T------T-----------------------------------------------------------------------------------------G------------------------------- 2515
B.IT.05.SG1             -----T-T------------A-----G-----N---------------T-------------------M-----------M------------------------------------------------------------AA-T----------N-------------- 3112
B.JP.05.DR6538          --------------------A------------T-----G--------T----------------------C--------------------------------------------------------------------------A--------G-------------- 3314
B.KR.05.05CSR3          ------G----------A--CT------------------A------GT----------------------------------------------------------------------------------G-----------------------G-----------A-- 3335
B.NL.00.671_00T36       ----------G------A--A--A---------T----A-A-------T--GA------------------------------C---------------------------------------------------------------------------CN----N---- 2875
B.RU.04.04RU128005      --------------------A--A--------C---------------T----------------------------------------T-----------------------------Y--R--------M---------AAA--------C----------------- 2801
B.TH.00.00TH_C3198      --------------------A-------------------A-------T--------------------------------------------------------------------------------------------A-A-------T------------------ 2524
B.UA.01.01UAKV167       --------------------A---------------C-----------T-------------------------------------G------------------------------------------------------AA--------------------------- 2536
B.US.04.ES10_53         -----------------AA-A---------------------------T-------------------------------------------------------------------------------------A-------A------------------------A-- 3319
B.US.99.PRB959_03       --------------------A------G-----T------A------GT-------------------G--------------C-----------------------------------C----------------------A---------------A----------- 2530
C.AR.01.ARG4006         ------G-----------A--T-A---G-------A----A------------------------G-----------------C-----T-----------G-----------------C---------------------CA------------G--T----------- 2494
C.BR.04.04BR013         ------G----------A---T-A---G-------A----A-------T-C-T----------C-G-----------------C-----T--G--------G-----G-----------C---------------------CA------------G--T----------- 2774
C.BW.00.00BW07621       ------GT-------------T-A---AG---C--A----AA------T----------------G-----------------C-----T-----------G--------------G--C---------------------CA---------G-----AGA--------- 2663
C.CN.98.YNRL9840        ------G--------------T-A---G----------A---------T----------------G-----------------C-----T-----------G-----------------C---------------------CA------------G--T----------- 2494
C.ET.02.02ET_288        -----------------A---T-A---G-------C----A-------T-C--------------G-----------------C-----T-----------G--------T--------C--G-----------A------CA------------G--A--T-------- 2512
C.GE.03.03GEMZ033       ------G----------------A---G-----T-A--A-AA------T----------------G--G--------------C-----T-----------G-----------------C---------------------CA----------G-G--G----------- 2491
C.IL.99.99ET7           -----GG-------------AT-A---GG---C--A-----A------T----T-----------G-----------------C-----T--G--------G-----------------C-----------G---------CA---A--------G--T----------- 2491
C.IN.99.01IN565_10      ------G-------------AT-A---GC---------A-A-------T----------------G-----------------C-----T-----------G-----------------C---------------------CA------------G--T----------- 2691
C.KE.00.KER2010         ------G--------------T-A---G-------A--A---------T----------------G--G--------------C-----T-----------G-----------------C---------------------CA---------C--G--T----------- 2491
C.MM.99.mIDU101_3       -----------------------A---G----------A-A-------T----------------G--------------------G--T-----------G-----------------C---------------------CA------------G--T----------- 2661
C.MW.93.93MW_965        ------G--------------T-A---GG----T-A-----A-----------------------G-----------------C-----T-----------G-----------------C---------------------CA------------G--T----------- 2497
C.SN.90.90SE_364        ------G-----G-----A--T-A---G----CT-A----AA------T----------------G-----------------------T-----------G-----------------C---------------------CA---------T--G--G----------- 2479
C.SO.89.89SM_145        ------G--------------T-A---G-------A----A-------T----------------G-----------------C-----T-----------G-----------------C---------------------CA------------G--T----------- 2531
C.TZ.02.CO178           ------G--------------------G-------A--A---------T----------------G--G--------------C-----T-----------G-----------------C-----------------C---CA------------G--T----------- 2491
C.UY.01.TRA3011         ------G-----G--------T-A---G------------A---------C--------------G-----------------C-----T-----------G-----------------C---------------------CA------------G--T----------- 2482
C.YE.02.02YE511         ------G-----G----A---T-A---G----C--A----AA------T----------------G-----------------C-----T-----------G-----------------C---------------------CA---A-----C--G--T----------- 2485
C.ZA.04.04ZASK164B1     ------GA----------C--T-A---G-C-----A--A-A-------T------------G---G-----------------C-----T-----------G---------------------------------------CA-------AC---G--T----------- 2724
C.ZA.05.05ZAPSK240B1    ------G--------------T-A---AC------AC-A-A-------T----------------G-----------------C-----T-----------G-----------------C---------------------CAA-----------G--T----------- 2714
C.ZM.02.02ZM108         ------G--------------T-----G----C--C----AA------T----------------G-----------------C-----T-----------G-----------------C--------------A--------A-----------G--G----------- 3279
D.CD.83.ELI             -----G------------A-AT-A---G-------A------------T-------G-----T--------------------C-----T---------------------------------------------T-----AAA-----------G--G----------- 2856
D.CM.01.01CM_0009BBY    ------T----------A--AT-A---G-----T-A----A---------------------T--G--------------C--------T---------------------------------------------------AAA--------------A----------- 2509
D.KE.01.NKU3006         ------G-------------AT-A---GG---C--A----A-------T-C-----------------G--------------------T---------------------------------------------------AAA--------------A----------- 2515
D.KR.04.04KBH8          ------G-----------A-AT-A---GG-----------A-------T-G--------------------------G--G--------T---------------------------------------------------ACA--A-----------A----------- 3269
D.TD.99.MN011           ------G-C---------A-AT-A---G-----T-A----A------------------------------------------------T---------------------------------------------------ACA--A-----------A----------- 2534
D.TZ.01.A280            --------------------AT-A---GG-C-CT-A----A-------T----------------------------------------T--------------------------------G--------G---------AAA-----------G--A----------- 2511
D.UG.99.99UGK09259      ----------C---------AT-A---GA---C--C--A-A-------T-A-----------------G--------------------T---------------------------------------------------AAA--------------A----------- 2515
D.YE.01.01YE386         -------T------------AT-A---GG---C--A----A-------T-------------------G--------------G----CT---------------------------------------------------AAA--------------A----------- 2509
D.YE.02.02YE516         -----------------AA-AT-A---G-------A-----------------------------------------------------T---------------------------------------------------ACA-----GC-------A-A--------- 2521
D.ZA.90.R1              ------G-------------AT-A---G-------A------------T----T--------T--------------------------T---------------------------------------------------AAA--------------A----------- 2656
F1.AR.02.ARE933         ---------------------T-A---A-------AC---CA------T----T--------------------A--------C-----------------G-----------------------------G--A------CAAT-------G--G-------------- 2604
F1.BE.93.VI850          ---------------------T-A---G-------AC----A------T----T-----------------------------------------------GC----------------------------G---------CAAT-----A-C--G-------------- 2647
F1.BR.01.01BR125        --C---G-------C------T-A---G-------AC----A------T----T--------------------A-G------C-----------------G-----------------------------G---------CAAT-----A-C--G-------------- 2755
F1.ES.x.P1146           ------G-------------AT-A---GG------AC---AA-----GT----T-----------------------------------------------G--------T-----A--------------G--------CCAA------A-C--G-------------- 2587
F1.FI.93.FIN9363        ---------------------T-A---G-------AC---AA------T----T--------------------------------------C--------G-----------------------------G---------CAAT-------C--G-------------- 2639
F2.CM.02.02CM_0016BBY   -----GG--------------T-A---G-------A-----A------T----T--------T--------------G-----C-----T-----------G-----------------C---------------G-----CAAT-----A-C--GAG------------ 2500
F2.CM.95.MP257          -----GG--------------T-A---GG------A-----A------T----T--------T--------------------C-----T-----------G---------------------------------G-----CAAT-------C--GAG------------ 2519
F2.CM.97.CM53657        -----G-----------A---T-A---G-------A-----A------T----T--------T--------------G-----C-----T-----------G-----------------C---------------G-----CAAT-------C--GAG------------ 2500
G.BE.96.DRCBL           ------G----------A---T-A---G-------AC----A------T-------------T--G-----------------------T-----------A-----G-----------C--------------A------CA-------A-C--G--A-A--------- 3267
G.CM.01.01CM_4049HAN    ------G--------------T-A---A-------AC--GAA------T-------------T--G-----------------------------------A-----G-----------C--------G-----A------CA---------C--G--A----------- 2512
G.CU.99.Cu74            ------G-------------AT-A-----------------C------T-------------T-----------------G-----------C--------G-----G-----C-----C--------------A------CA---------T-----A----------- 2914
G.ES.99.X138            ------G--------------T-A---A----------A--C-----GT-------------T--------------------------------------A-----G-----------C--------------A------CA---------C-----A----------- 2751
G.GH.03.03GH175G        ------G-----G--------T-A---G-------AC---AA------T-------------T-----------------------------C--------A-----------------C--------G------------CA---------T--G--A----------- 3350
G.KE.93.HH8793_12_1     ------G--------A----AT-A---G-------A-----A------T----------------------------------------------------A-----G--T--------C--------G------------CA------------G--A----------- 2711
G.NG.01.01NGPL0669      ------G----------A---T-----A-T-----AC----A------T-------------T--------------------------------------A-----G--------------------G-----A------CA---------T--G--A-AT-------- 2515
G.PT.x.PT2695           --C---G--------------T-A---A----------A-AC-----GT-------------T-----------A--------------------------A-----G--------C--C--G-----------A------CA---------C--G--A----------- 3252
G.SE.93.SE6165          ------G--------------T-A---G-------A----AA-----GT-------------T--------------------------------------A-----G-----------C--------G-----A------CA---------C--G--A----------- 2707
H.BE.93.VI991           ------GA---------A---T-A---GCT---T--------------T-C------------C-------------------C-----T-----------A-----------------C------------------G--AAAT------------------------- 2699
H.BE.93.VI997           ------G-G------------T-A---GCT---T--------------T-------------------G--------------C-----T-----------A-----------------C-----------G------G--AAA-------------------------- 2634
H.CF.90.056             ------G--------------T-A---GCT---T------AA------T----------------------------------C-----T-----------A-----------------C---------------A--G--AAA-------------------------- 2657
J.CD.97.J_97DC_KTB147   -------AG------------------G-G-----A------------T----------------------------------C-----T-----------A-----G-----------C--------G-----A---G-CCA---A-----------A----------- 2504
J.SE.93.SE7887          -------G-------A----------GG-------A----A-------T-------------T--------------------------T-----------A-----G--T--------C-----------C--A------CAA--------------AGA--------- 2624
K.CD.97.EQTB11C         ------G-------------A--A---G-------A-----A------T-------------------G--------------C-----T-----------G-----------------C-----------G---------CAA--------C--G--T----------- 2506
K.CM.96.MP535           -C------------------A--A---G-------A----AA------T-------------------G--------------C-----T-----------A-----------------C-----------G---------CAA--------C--G--T----------- 2506
```

```
B.FR.83.HXB2                    CATAGAACAAAAATAGAGGAGCTGAGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAAGACAGCTGGACTGT 3310
Pol                             _H__R__T__K__I__E__E__L__R__Q__H__L__L__R__W__G__L__T__T__P__D__K__K__H__Q__K__E__P__P__F__L__W__M__G__Y__E__L__H__P__D__K__W__T__V__Q__P__I__V__L__P__E__K__D__S__W__T__V
01_AE.CF.90.90CF11697           -----------------A-----A---G-T-----A----AA------T----T--------------G--------------------------------G-----------------C-G---------C----------AA-------------------------- 3259
01_AE.CN.05.FJ051               --C----T-------------------GCT--C--A-----C------T----T-----------------------G-----------------------A-----------------C-----------C--A-------AA-----------------T-------- 3320
01_AE.CN.06.FJ054               --C---G--------------------GCT-----A--C--C------T----T--------------G-----A--G--------------C--------G-----------------C-----------C----------AA--------T----------------- 3335
01_AE.HK.x.HK001                -------T-------------------GCT---T-A--A--C------T----T--------------G--------G-----------------------G-----------------C-G---------C----------AA-------------------------- 2672
01_AE.TH.01.01TH_R2184          --C------------------------GCT-----A-----C-----------T--------------G--------G-----------------------A-----------------C-G---------C----------AA-------------------------- 2509
01_AE.TH.02.OUR769I             --C--------------A-----A---GCT-----A-----C-----GT----T--------------G--------G-----------T-----------A-----------------C-G---------C--A-------AA--------------T--T-------- 2509
01_AE.TH.90.CM240               --C--------------------A---GCT-----A-----C------T----T--------------G--------G-----------------------A-----------------C-G---------C----------AA-------------------------- 2884
01_AE.US.00.00US_MSC1164        --C--------------------A---GCT-----A-----C------T----T--------------G--------G-----------------------A-----------------C-----------C------G---AA-------------------------- 2515
02_AG.CM.02.02CM_4082STN        ------G--------------T-----G-------AC-----------T----------------------------------------T-----------A-----------------C-----------C----A----CA------------------------C-- 2509
02_AG.EC.x.ECU41                ------G-----G--------T-----G-------AC---AA------T----------------------------------------T-----------A-----G-----------C-----------C---------CA--------------------------- 2406
02_AG.GH.03.GHNJ196             ------G--------------T-----G-------AC---AA------T----------------------------G-----------T-----------A-----------------C-----------C--A-A-----AA-------------------------- 3339
02_AG.NG.01.PL0710              ------G--------------T-----GG------A--A-AA------T----------------------------------------T-----------A-----------------C-----------C---------AA---------G-----AGA--------- 2494
02_AG.NG.x.IBNG                 ------G--------------T-----GG------AC---AA------T-------------------G--------------------T-----------------------------------------C------G---AA--A----------------------- 2835
02_AG.SE.94.SE7812              ------G--------------T-----G-------AC----A------T-------------TG----G--------------------T-----------A-----------------C-----------C-------------------------------------- 2682
02_AG.SN.98.MP1211              ------G--------------T-----G-------AC---AA------T----------------------------G-----------T-----------A-----------------C-----------C---------CA--------------------------- 2507
03_AB.RU.97.KAL153_2            ---------G----------A------G--------C-----------T-------------------------------------------------------------------------------T---------------T------------------------- 2536
04_cpx.CY.94.CY032              ------G----------A-----A---G-G-----A-----A------T-C--------------------------------C-----T-----------G-----------------C-----------G---------CAA-C-G-------G--T----------- 2676
05_DF.BE.x.VI1310               -----GG--------A-----T-A---G-------AC--GCA-----GT----T-----------------------------------------------G-----------------------------G---------CAAT-------C--G-------------- 2692
06_cpx.AU.96.BFP90              ------G--------------T-A---G-G-----A----AA--------------------T--------------------C-----T-----------G-----------------C-----------G---------CAA--C-----C--G--G----------- 3338
08_BC.CN.97.97CNGX_6F           --------------------A------G----------A-A-------T----------------G-----------------------T-----------G-----------------C---------------------CA------------G--T----------- 2490
09_cpx.GH.96.96GH2911           ------G--------A----A--A---G-------A----AA------T----T--------------G--------------C-----T-----------G-----------------C-----------G---------CAA--------C--G--A----------- 2512
10_CD.TZ.96.96TZ_BF061          ----C--T------------A--A--GGG---C--A--A-A-------T-----G-------------G-----------G--------T--G--------G---------------------------T-----------CA-------A----GA-T--T-------- 2689
11_cpx.GR.x.GR17                ------GA---GG-------AT----GA-G-----A--A-AA------T-C-----------T-----------------------------C--------A-----G-----------C--------G-----A------CA---------C-----G----------- 2612
12_BF.AR.99.ARMA159             --------G-----------A--A--------------A---------T-------------------------------------------------------------------------------G----------------------------------------- 3315
13_cpx.CM.96.1849               ------GA---GG-------AT-A---AG-------------------T-------------T--G-----------------------------------A-----G-----------C--------------AG-----CA-T-------C--G-------------- 2715
14_BG.DE.01.9196_01             ------G-------------RT-A---A----------A-AC------T-------------T--------------------------T-----------A--Y--G--T--------C--------------A------CA----H----C-----A----------- 2827
15_01B.TH.99.99TH_MU2079        --C-----------------A--A---GCT-----A-----C-----GT----T--G--------------------G-----------------------A-----------------C-G---------C----------AA-------------------------- 2709
16_A2D.KR.97.97KR004            --------------------AT-A---A-T-----AC---AA------T--TA---------T--------------------------T----------AA-----------------C-----------C---------AA------------G-------------- 2673
17_BF.PE.02.PE02_PCR0155        --------------------A-----------------A----------------------------------------------------------------------------------------------------------------------------------- 2500
18_cpx.CU.99.CU76               --------------------A--A-----------A----AA------T----------------------------------------T-----------A-----------------C-----------C-----G---CA--------------------------- 2615
19_cpx.CU.99.CU7                ---------------------T-A--G-------------A-------T-------------------------------------G--------------G--------------------G------------------TCA--------------A----------- 2536
20_BG.CU.99.Cu103               --------------------A----------------G--A-------T--TA---------T-----------------G--C-----------------A-----G-----C-----C--------------A------CA---------C----------------- 2776
21_A2D.KE.91.KNH1254            ------G----------A--AT-A---G-------A-----A------T-------------T-----------A--------A------------------------------------------------------G--AAA------A----G--G----------- 2515
22_01A1.CM.01.01CM_0001BBY      -------T---------C---T-----GCT-----A-----C-----GT----T--------------G-----A--------------T--C--------A--C--------------C-----------C--A------CA-------A-C----------------- 2518
23_BG.CU.03.CB118               ------G-------------A-----------------A-A-------T----------------------------------------T--------------------A---------------------------------------------------------- 2753
24_BG.CU.03.CB378               ------G-------------A-------------------A-------T------------------------------------------AG-----------------------------------------------------------G-----A----------- 2735
24_BG.CU.03.CB471               ------G-------------A-------------------A-------T--------------------------------------------------------------------------------------------A----------------A----------- 2750
25_cpx.CM.01.101BA              ---------------------T-----G-------AC---AA------T-------------------G--------G-----------T-----------A-----G-----------C---------------------AAT--------C----------------- 2518
27_cpx.CD.97.97CDKTB49          ------G--------------T-A--GACT---T----A-A-------T-C-----------------C--------------C-----T-----------A-----G-----------C--------G-----A------AA------------G--A----------- 2518
28_BF.BR.99.BREPM12609          ------------G-------A-----------------AGA------GT-----------------------------------------------------------------------------------------G----------------G----TT-------- 2676
29_BF.BR.02.BREPM119            --------------------A--------------------------GT-------------------G----------------------A-------------------------------------------------A---------------------------- 2484
31_BC.BR.02.110PA               --------------------A-------------------A-------T-C--T-----------G-----------------C-----T-----------G-----------------C-----------C---------CA------------G--T----------- 2744
32_06A1.EE.01.EE0369            ------GT-------------T-A---GG----T-A----A-------T-C-----------T--------------------C-----T-----------G-----------------C-----------G----------AA--------C--G--A----------- 2945
33_01B.MY.05.05MYKL007_1        --C----T-------------T-----GCT-----A----AC------T----T--------------G--------G-----------------------A-----------------C-G---------C----------AA-------------------------- 2689
34_01B.TH.99.OUR2478P           -----------------------A---G-T-----A------------T----T--------------G--------G-----------T-----------A-----------------C-----------C----------AA-------------------------- 2512
35_AD.AF.05.05AF095             -------T---------A---T-----GCT-----A----A-------T----------------G--G--------------------T----------------------------------------------------A---------------A----------- 2503
36_cpx.CM.00.00CMNYU830         ------G--------------T-----A-------AC---AT------T----------------------------------------T-----------A-----------------C-----------C---------CAA----------------A--------G 2515
37_cpx.CM.00.00CMNYU926         ------G--------------------G-C--C--AC---AA------T-------------T-----------------------------C--------A-----------------------------C--A------CA-T-------T-------T--------- 2500
38_BF1.UY.05.UY05_4752          ------GAC---G------GA--A------------------------T-------------------G--------------------------------------------------------------------------------------G-------------- 2721
39_BF.BR.03.03BRRJ103           ------G-----G-------A----------------G--A------GT--TA----------G-C-------------------------AG--------C-----------------------------G---------AAA--A--------G-------------- 2799
40_BF.BR.04.04BRRJ115           ------G--------A-----T-A---G-CT--T-A-G--AA------T--TAT--------GG----G-----A--G-----C-------A---------G-----------------------------------C---AA---------T--G--T----------- 2862
42_BF.LU.03.luBF_05_03          -------T------------A-----G--G------C----------C--C--------------------------------------------------A-----------------------------G---------CAA-----------G-------------- 2848
43_02G.SA.03.J11223             ------G--------------T-A---G----C--A----RA------T-------------T--------------------------T-----------A-----G-----------C--------------A------CA---------C--G--A----------- 2835
U.CD.83.83CD003                 ------G--------------T-----GCT-----A-----A------T----------------------------------------T-----------A-----------------C---------------------CA-------A-C-----TGAT-------- 2805
U.CD.90.90CD121E12              ------G--------------T-----GCT-----A------------T----------------------------------------T-----------A-----------------C---------------------CA---------C-----TGAT-------- 2509
U.GR.99.GR303                   ------G-------------A--A---G-------A--A-AA------T--TT------------G--G--------------C-----T-----------------------------C-----------G---------CAAT-------G--G-TA-A--------- 2584
U.NL.95.U_NL_95_H10986_D1       ------G----------A-GAT-A---G-T-----A--A-A-------T----------------------------------C-----T-----------G-----------------C-----------G---------CAA--------C-----G----------- 3195
CPZ.CD.90.ANT                   ------GA--TG-----AA----T---------T-ACA-GTC--------AGAG--T--T--------G--------------------T-AA--------A-----GT-A-----A--C--------T------AAA---AA---A-----GCC---TGAT-----A-- 2749
CPZ.CM.01.SIVcpzCAM13           --C---CAG--GG----A-----T-A-G-----T--C-TCAC-----GT----A--C-----T-----G-----A--------A--C--T--A--------G-----G--T--C--A--------------G-----A---AA---A--------G--GGTA-----A-- 2890
CPZ.CM.05.SIVcpzEK505           ------GAGGC-G-G--AAGA--T--GAGT-----T----CT------T----A--C--T--------G--------------A-----TT-G--------A--------------A--------------C---A-A---CAAT-A-----G-------CA-------- 2882
CPZ.CM.05.SIVcpzLB7             --C--G-AC--GG----A---------GCT--CT--C-A-A------GT-------C--T-----------------G-----A-----T--C--------A--------------A--C-----------------C---CA-T-A--------G--A--T-------- 2865
CPZ.CM.05.SIVcpzMB66            ------GA----G-G--A--AT----GGCT--C--A----AA-----TT-C-----T--T--------G-----A--G--G--A-----T--C--------A-----------C--A--C-----------T---------CA-T-A-----------A----------- 2856
CPZ.CM.05.SIVcpzMT145           -------AG-G-G----A-----C---GC---------A-AA------T----A-----------G--G--------------C-----T--C--------A-----GT-A-----A-----G--------C----AA---AAAT--------C-G--AGTG-------- 2852
CPZ.GA.88.GAB1                  -------A---GG-G--A--A--A---------T--C-T-AA-----GT-C--A--C-----------------A--G--G--A--C--TT-A--------A-----G-----C--A--C-----------C--------TCAAT-A-----------GGTA-------- 3371
CPZ.TZ.00.TAN1                  -----G-A-CTG---A--A----------G--------A-CA------T-AGAG--C--T--------GT-----G--A-----------A-G--------C-----G--A-----AA-------------T--AAA---CACAT-A------CC---GCAG-----A-- 2954
CPZ.US.85.CPZUS                 -----GGA---GG----ACT---C------T--T--C-T-CT------T-C--T--C--T--T--G--G------G-G-----A------T-A--------C--------------A-----G--------G-----C---CA-T-A---C-------A-TT-----A-- 3370
CPZ.CM.x.SIVcpzDP943            -------A---GG-----TTA--T-----------AC-C-CT-----TT----T--T--T--T--G--G-----A-----G--A--T--TT-G--------C-----G--T--C--A--------------G--A--A---CA-T-A---C-------G-C--------- 2896
CPZ.CM.x.SIVcpzMB897            ------GA---GG-G--A--AT-A---GCT-----A--A-A------CT----T-----T--------G-----A--------A-----------------A-----G--------A--C-----------C-----C---CA---------M----------------- 2876
N.CM.02.DJO0131                 ------GAG-C-G-G--A--A--T---AGT-----T----A------CT----G--C--T--------------A--G--G--A--G-----C--------A-----G--------A-----------G--C---A-A---AA-T-A------C-G--TGTA-------- 2816
N.CM.04.04CM_1015_04            ------GAG-C-G----A--A--T---GGT-----T----A------CT----A--C--T--------------A--G--G--A--G-----C--------A-----G--------A--------------C-----A---AA-T-A--------G--T-CA-------- 2815
N.CM.04.04CM_1131_03            ------GAG-C-G----A--A--T---GGC----------A------CT----A--C--T--------------A--G--G--A--------C--------A-----G--------A--------------C-G---A---AA-T-A--------G--T-CA-------- 2814
N.CM.95.YBF30                   ------GAGGC-G----A--C--T---G-T-----T----A------CT----G--C--T-----------------G-----C--G-----C--------A--------------A--C-----------C-----A---AA-T-A--------G--TGTA-------- 2904
N.CM.97.YBF106                  ------GAG-C-G----A--A--T---GGT-----T----A------CT----G--C--T--------------A--G--G--A--G-----C--------A-----G--------A-----------G--C-----A---AA-T-A--------G--GGTA-------- 2902
O.BE.87.ANT70                   -------A--GG--T--ATT---T---G----C--A-ATCA-------T-C--T--C--T--------G-----A--G--------C--T--A--------G-----G--------A--C---------------T-C--CCAAT----TA-C--G--TGTG-----A-- 3364
O.CM.91.MVP5180                 -------AG-GGG-C--ATT---T--GG-----T-A-ATCA-------T----T--C--T--T-----G--------G--------C--TT-A--------A-----G-----C--A--C--G--------------C--CCAAT----T--C-----AGTG-----A-- 3339
O.CM.96.96CMABB637              -----G-A--GGG-T--ACTA--T---G-----T-A-ATC--------T-C--T--C--T--T-----G-----A--G--------C--T--G--------G-----G-----C--A--C--G--------------C--CCAAT----T--C--G--AGTA-----A-- 2794
O.CM.98.98CMA104                -----G-A--GGG-T--ATT---T--GG-----T-A-AT-A-------T-C--T--C--T--T-----G--------G--------C--T--G--------G-----G--------A--C--G-----------A--C--CCA-T----T--T--G--AGTA-----A-- 2796
O.CM.99.99CMU4122               -------A--GGG----ATT-------G----CT-A-ATCA-------T-C--T--C--T--T-----G-----A--G--------C--T--G--------G-----G-----C--A--C-----------------C--CCAAT----TA-C--G--G-TA-----A-- 2795
O.FR.92.VAU                     -----G-A--GGG-T--ACTA--C---A-----T-A-AC-A-------T--TAT--C--T--T-----------A--G-----A--C--T-AG--------G-----G-----C--A--C--G--------------C--CCAAT----T--C--G--AGTG-----A-- 2876
O.SN.99.SEMP1299                -------AG-G-G----ATT---T---G----CT-A-ATCA-------T-C--T--C--T--T-----G-----A--G--------C--TT-G--------G-----G--------A--C-----------G-----C--CCAAT----TA-C--G--GGAA-----A-- 3363
O.US.99.99USTWLA                -------A--G-G-T--ATT---C---G-----T-A-AT-AA--R---T-C--T--C--T--T-----------A--G--------C---M-G--------G-----G-----C--A--C--G--------------G--CCAAT----YA-C-----AGTG-----A-- 2793
```

```
B.FR.83.HXB2          CAATGACATACAGAAGTTAGTGGGGAAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAAACTCCTTAGAGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGAGAGATTCTAA 3480
Pol                   __N__D__I__Q__K__L__V__G__K__L__N__W__A__S__Q__I__Y__P__G__I__K__V__R__Q__L__C__K__L__L__R__G__T__K__A__L__T__E__V__I__P__L__T__E__E__A__E__L__E__L__A__E__N__R__E__I__L__
A1.GE.99.99GEMZ011    ------T--------A--------A---C-A-----------------C--T----------------A------G-----------C--G---G----------G-----TA--G-GA----G--T--G--------AT-----T-------G---------------- 2664
A1.KE.00.KER2008      ------T--------A--------A---C-A--------C-----A-----T-----A----------A---------------T-----G---G----------------TA--G--A----G--T--G--------AT-----T-A-----G-----G---------- 2679
A1.KE.00.KNH1144      ------T--------A------------C-A--------T-----A-----TG----A----------A---C--G----G---T--C--G---G----------------T---G--A--T-G--T--G--------AT----GT------------AG---------- 2685
A1.KE.00.KSM4024      ----------------------------C-A--------G-----A-----T----------------A------G--------A--C------G----------------T---G--A--T-G--T--G--------AT-----T-------G-----G---------- 2676
A1.KE.00.MSA4069      ------T--------A--------A---C-A--------G-----A-----TG---------------A------G--------------G---G----------------TA--G-----T-G--T--G--------AT-----T-------G-----G---------- 2679
A1.KE.00.NKU3005      ------T--------A-----A------C-A--------------A-----TG----A-----------A-----G-----G-----C--G---G----------------TAC-G--A--T----T--G--------AT-----T-------G-----G---------- 2679
A1.RU.00.RU00051      ------T--------A--------A---C-A--------C-----------T---------------------C-G-----------C------G-------G-----T--TA--G--A-------T--G--------AT-----T-------G-----G---------- 2781
A1.RW.93.93RW_024     ------T--------A-----A------C-A--------------A-----TG----A----------A------G-----------C--G---G----------------GA--G--A--T-G--T--G--------AT-----T-------G-----G---------- 2682
A1.SE.95.SE8891       ------T--------A--------A---C-A--------------A-----TGT---A----------A----C-G-----------C------G----------------TA--G--A----G--T--G--------AT-----T-------------G---------- 2669
A1.SE.95.UGSE8131     ------T--------A------------C-A--------------A-----TG---------------AA-----G-----------C--G---G--------T-------TA--G--A--T-G--T--G--------A------T-A-----G-----G---------- 2852
A1.TZ.01.A173         ------T--------A--------A---C-A--------------A-----TG---------------A------G-----G-----C--G---G----------------TA--G--A--A-G--T--G--------AT-----T-------------G--A------- 2679
A1.UA.01.01UADN139    ------T--------A--------A---C-A--------------------T-----------------------G-----------C--G---G----------G-----TA--G-GA----G--T--G--------AT-----T-------G---------C------ 2679
A1.UG.92.92UG037      ------T--------A-----A------C-A--------------A-----TG----A----------AA-----G-----------C--G---G--------T-------TA--G--A--T-G--T--G--------AT-----T-------G-----G------T--- 3496
A1.UG.99.99UGA07072   ------T--------AC-------A---C-A--------------A-----T----------------A----C-G-----------C------G-----A----------TA--G-G-----G--T-----------AT-----T-------G-----G---------- 2679
A2.CD.97.97CDKTB48    ------T--------A-----A--A--G--A--------------------TG--------------------C-G---------------------------T-------CA--G-------G--TAG---------AT-----T---A---G-----G---------- 2818
A2.CY.94.94CY017_41   ------T--------A-----A---------A--------------------TG---------------A----C-G------------------G----------------CA--G--A----G--TA-----------T-----T-A-A---G-----G--A---T--- 2837
A.SN.01.DDI579        ------T--------A-----A--A---C-A--------------------TG--------------------C-G-----------C--G---G----------------CA--G--A--T-G--T--G--------AT-----T-------G-----G---------- 2678
A.SN.01.DDJ369        ------T-----------------A---C-A--------------A-----T----------------A----C-G-----------C--G---G----------------T---G-GT--T-G--T--G--------AT-----T-------G-----G---------- 2681
A.SN.96.DDJ360        ------T--------A--------A---C-A--------------------TG---------------A----C-G-----G-----C--G---G----------------CA--G--A--T-G--T--G--------AT-----T-------G-----G---------- 2678
A.CD.02.02CD_KTB035   ------T--------A--------A------A--------------------T-----A--------G-A----C-G-----------C--G---G----------------CA--G--A----G--T-----------AT-----T-------G-----G--A-----G- 3474
A.CD.97.97CD_KCC2     ------T--------A--------A--GC-A--------G-----------T----------------A----C-G-----G---A-C--G---G-------CT-------CA--G--A----G--T--G--G-----AT-----T-------G-----G--------G- 3476
A.CD.97.97CD_KTB13    ---------------A--------A---C-A--------------------T-----------G----A----C-G-----------C--G---G----G---T-------G---G-GA----G--T--G--------AT-----T-------G-----G--------G- 2684
B.AR.04.04AR151516    ------------------------A--------------C-----------TGG---A----------A---------------------G---G-----T--------------G----------A--------------------------------G------T--- 2702
B.AU.87.MBC925        ------------------------A--------------------------TG---------------A----------------A----G---------------------A--G----------A--------------------------------G--A------- 3491
B.BO.99.BOL0122       ------------------------A---C----------------------T-----------G----A---------------------G-----------G--------------------------------------------------------G---------- 2700
B.BR.03.BREPM2012     T-----------------------A--------------------------TG----------G----A---------------------G------------------------G----------------------A--------------------G------T--- 2905
B.CA.97.CANB3FULL     ------------A-----------A---------------------------G---------------A---------------------G--------------------------------------------------------------------G--A------- 2769
B.CN.05.05CNHB_hp3    ------------------------A-----A--------------------TG----------G----A-G-------------A-----G------------T----------------------------------------------G--------G--A------- 3479
B.CO.01.PCM001        ------------------------A-----A--------------A-----T-----------G----A---------------------G-----------C------------G----C---T---C---G--G---------T-------------G---------- 2685
B.GB.83.CAM1          ------------------------A-----------------------C--TG---------------A---------------------G------------------------G----------------------A------------------------------- 3481
B.GB.86.GB8           T-----------------------A--------------------------T--------------------------------------G---G-----------------------A----G-----------------------------------G---------- 3500
B.GE.03.03GEMZ010     ------------A-----------A--------------------------TG-------------------------------------G------------------------G--T-------------------------------------T--G---------- 2685
B.IT.05.SG1           ------------------------A-----A--N----------------------------------AA--------------------G---------------------A--------------------------------T-------------G---------- 3282
B.JP.05.DR6538        ------------------------A--------------------------TG---------------A---------------------G---G-----------------A-----GTGT----------------------------------------A---T--- 3484
B.KR.05.05CSR3        ---------------A--------A------G-------------------------------G-C--A------------G--A-----G---G-------------------------C------A-------------------------------G--A------- 3505
B.NL.00.671_00T36     ------------------------A---N----------------------T--------------------------------------G-----T---------------A--G-----------C-------------------A---------------------- 3045
B.RU.04.04RU128005    ------------------------A-----A--------------------T-----------------A--G--G--------------G---G--------------------------------C---G-------------------------------------- 2971
B.TH.00.00TH_C3198    ---------------A--------A--G-----------------------TGG---A----GG----A-G-------------------G------------------------G----------------------A--------------------G--A-----G- 2694
B.UA.01.01UAKV167     ------------------------A--------------------------TG---------------AA-----G--------------G------------------------G-----------------------T-------------------G---------- 2706
B.US.04.ES10_53       ------------------------A---C-A---------------------G---------------A-----------GG--------G------------------------G-------------------------------------------G---------- 3489
B.US.99.PRB959_03     ------------------------A-----A-------------------------------------A---------------------G---------T----G--------------------------------A--------------------G---------- 2700
C.AR.01.ARG4006       ------T-----------------A-----A--C----------------------------------AA---C-G--------------G--------------------CA--G-------G--T-C---------AT-----T-------G-----G--A------- 2664
C.BR.04.04BR013       ------T-----------------A-----A--C-----------------------------------A---C-G--------------G--------------------CA--G-------G--T-----------AT-------------G-----G--A------- 2944
C.BW.00.00BW07621     ------T-----------------A--G--A--C-----C-----A---------------------------C-G------------------G----------------CA--G-G-----G--T-----------AT-----T----T--G-----G--A------- 2833
C.CN.98.YNRL9840      ------------------------A-----A--C-----------------------A---------------C-T--------------G--GG----------------CA--G----------T-----------AT-----T-------------G--A---T--- 2664
C.ET.02.02ET_288      ------T-----------------A-----A--C--------C--------------A----------A----C-G--------------G---G-T--------------CA--G----------T-----------AT-----T-------------G--A------- 2682
C.GE.03.03GEMZ033     ------T-----------------A-----A--C--------------------------C--------A--GC-T--------------G--GG----------------CA--G----------T-----------AT-----T-------G-----G--A------- 2661
C.IL.99.99ET7         ------T--------A--------A-----A--C-----------------------------------------G--------------G---G----------------TA--G-------G--T-----------AT-----T-------G-----G--A------- 2661
C.IN.99.01IN565_10    ------T-----------G-----A-----A--C--------C--------T----------------A----C-T-----G--------G--G-----G-----------CA--G----------T-----------AT-----T-------G--------A------- 2861
C.KE.00.KER2010       ------T-----------------A-----A--C-----------------------A----G---------CC-T------------------G----------------CA--G----------T-----------AT-----T-------G-----G-CA------- 2661
C.MM.99.mIDU101_3     ------T-------G---------A-----A--C-----------------------A---------------C-C--------------G--GG----------------CA--G----------T-----------AT-----T-------------G--A---T--- 2831
C.MW.93.93MW_965      ------T--C--------------A-----A--C-----C---------------------------------C-G--C-----------G---G----------------CA--G-------G--T-----------AT-----T-------------G--A------- 2667
C.SN.90.90SE_364      ------T-----A---A-------A-----A--C---------------------------------------C-G--------------G---G-------C--------C---G-------G--T-----------AT----GT-------G-----G--A------- 2649
C.SO.89.89SM_145      ------T-----------------A---C-A--C-----------------------------G---------C-T-----------------GG----------------CA--G----------T-----------AT-------------G-----G--A------- 2701
C.TZ.02.CO178         ------T-----------------A-----A--C---------------------------------------C-G--------------G--G-----------------CA--G----------T-----------AT-----T-A-----G---------------- 2661
C.UY.01.TRA3011       ------T-----------------A-----A--C--------------------------A------------C-G--------------G---G----------------CA--G-------G--T-----------AT-----T-------G-----G--A---T--- 2652
C.YE.02.02YE511       ------T--------A--------A-----A--C---------------------------------------C-G--------------G---G----------------TA--G-------G--T-----------AT-----T-------------G--A------- 2655
C.ZA.04.04ZASK164B1   ------T-----------------A-----A--C-----------------------------G---------C-T--------------G-----T--------------CA--G----------T-----------AT-----T-------G---------------- 2894
C.ZA.05.05ZAPSK240B1  ------T-----------------A-----A--C---------------------------------------C-T--------------G--GG----------------CA--G----------C-----------AT-----T-A-----G-----G--A------- 2884
C.ZM.02.02ZM108       ------T-----------------A-----A--C-----------A-----TG--------------------C-G--------------G---G----GT----------CA--G-------G--T-----------AT-----T----T--G-----G--A------- 3449
D.CD.83.ELI           ------T--------C-------A--G---A--C--------C--------T-----A-----------A--------------------G-----------------------------------------------AT-------------------G--A---T--- 3026
D.CM.01.01CM_0009BBY  ------T-----------------A-----A-----------C--------TT----A----------A---------------------G---G-T------T-------G---------------A----------AT-------A-----------G--------C- 2679
D.KE.01.NKU3006       ------T-----------------------A-----------C--A-----T-----A-----------A--------C---TG------G---G----------G--------------C--G---A----------AT-------------------G--A------- 2685
D.KR.04.04KBH8        ------T-----------------A-----A--------------A-----TG----A----------AA--G--G-----------C--G---G----------G-----------------G---A----------AT-------------------G---------- 3439
D.TD.99.MN011         ------T--------AC-------A-----A-----------C--------T-----A-----G----A-----------G---------G---G-T---A--T-------G---G-----------A----------AT-------------------G--------C- 2704
D.TZ.01.A280          ------T-----------------A-----A--------------------TT----A----------AA--------------------G---G----------G-----------------G---A----G-----AT----------------T--G---------- 2681
D.UG.99.99UGK09259    ------T-----------------------A-----------C--A-----TG----A-----------A--------C---TG-A----G---G----------G---------G-------G----C----------T----G--------------G--A---T--- 2685
D.YE.01.01YE386       ------------------------A-----A--------------------------A-----------A--------C---TG------G---G----------G---------G----------------G-----AT-------------------G--A------- 2679
D.YE.02.02YE516       ------T--------AC-------A-----A-----------C--------------A----------A-----------G---------G---G-T------T-------T---------T-----AG---------AT-------------------G--------C- 2691
D.ZA.90.R1            ------T--------------A--------A--C--------C--------T-----A--------------------------------G------------T-----------G----------------------AT-------------------G--A--CT--- 2826
F1.AR.02.ARE933       ------T--------------A--A---C-A--A-----------------T----------------A---------------------G---G----G-----------CA--G-G-----G--T-C----------T-----T----------T--G--A------- 2774
F1.BE.93.VI850        ---------------------A--A---C-A--------------------T-------------------C------------------G---G----G-----------CA--G-G-----G--T-C---G------T-----T----A-----T--G---------- 2817
F1.BR.01.01BR125      ------T--------------A--A---C-A--------------------T----------G-----A-A-C-----------------G---G----G-----------CA--G-G-----G--T-C----------T-----T-------G--T--G--A------- 2925
F1.ES.x.P1146         ---------------AC----A--A-----A--------------------T-----------------A--------------------G---G----G-----------CA--G-G-----G--TAC----------T-----T-------G-----G---------- 2757
F1.FI.93.FIN9363      -G----T--------A-----A--A---C-A--C-----------N-----T------------------------------------------G----G-----------CA-GG-G-----G--T-C-------A-TT-----T----------T--G---------- 2809
F2.CM.02.02CM_0016BBY ------T-----------------A-----A--------------------T----------G-----A----C----------------G---G----------------C---G-G-----G--T-C------------------------G-----G--A------- 2670
F2.CM.95.MP257        ------T-----------------A---C-A--------------------T----------G-----A----C----------------G--------------------T---G-G--T-----T-C----------T-------------G-----G--A------- 2689
F2.CM.97.CM53657      ------T-----------------A---C-A--------------------T----------G-----A---CC----C---------------G-------G--------C---G-G----------T---------AT-----A-------G-----G--A------- 2670
G.BE.96.DRCBL         T-----T-----A-----------A---C-A--------------------T-----A----------A---------C------A----G--GG--------T-------CA--G--T--A-G----C---------AA-G--GT-------G-----G---------- 3437
G.CM.01.01CM_4049HAN  ------T-----A---------------C-A--------------------T-----A----------A---C-------GG--T-----G--GG----------------CA--G----C--G--G-C---------AA-G---T-------G----AG--A------- 2682
G.CU.99.Cu74          ------T-----A-----------A--GC-A--------------------T----------------AA------------------C-G--GG--------T-------CA--G-------G--T-----------AT-G---T-------G-----G--A------- 3084
G.ES.99.X138          ---C--------A-----------A---C-A-----------C--------T----------------AA---C----------------G---G----------------CA--G----------T-C---------AT-G---T-------G--------------G- 2921
G.GH.03.03GH175G      ------T-----A-----------A---C-A--------------------T-----A----------A---CC------G---------G--GG-----------------A--G--T-C--G----C---------AA-G--GT-------G-----G--A------- 3520
G.KE.93.HH8793_12_1   ------T-----A-----------A---C-A--------------------T---------------------C----------------G--GG----------------CA--G-------G--T-CG--------AT-G---T-------G-----G--.------- 2880
G.NG.01.01NGPL0669    ------T-----A--A--------A---C-A--------------------TG---------------A---------------T--C-----G-----------------CA--G-T-----G--T-----G-----AT-G---T-------------G--A------- 2685
G.PT.x.PT2695         ------T-----A-----------A---C-A-----------C--------T------------A---AA---C----------------G---G----------------CA--G----------T-C---------AT-G---T----------------------G- 3422
G.SE.93.SE6165        ------T-----A-----------A---C-A--------------------T----------------C---C-----------------G--GG----------------CA--G--T----G----C---------AA-G--GT-------G-----G--A------- 2877
H.BE.93.VI991         ------T-----------------A---C-A-------------------------------------A----C-------W--------G--GG-----------------A--G-------G--TA----G-----AT-G---T-------------G---------- 2869
H.BE.93.VI997         ------T-----------------A---C-A--C-----------------T---------------------C----------------G--GG----------------C---G-------G---A----G-----AT-G---T-------------G---------- 2804
H.CF.90.056           ------T-----------------A---C-A--------------------T----AAT---------A----C----------------G--GG--------T-------CA----------G---A----G-----AT-G---T-------------G--------G- 2827
J.CD.97.J_97DC_KTB147 ------T-----A--A--------A---C-A--------------------T----------------A----C----------------G---G-------------G--CA--G--G--T----T-C---------AT-----T-------G-----G---------- 2674
J.SE.93.SE7887        ------T-----A-----------A---C-A--------------------T-----A----------A----C---------------A---GG-T--------------CA--G-----T-G--TAG---------AT-G----------------A----------- 2794
K.CD.97.EQTB11C       ------T-----A-----------A-----A--------------A--A-T------A----------AA--------------------G---GT-------T-------CA--G----------T-C----------T-----T-A-----G--------A------- 2676
K.CM.96.MP535         ------T-----------------A-----A--------------A--A--T----------------AA-----G--------------G---GT---------------CA--G-----T----T-C----------T-----T-A-----G---------------- 2676
```

```
B.FR.83.HXB2                 CAATGACATACAGAAGTTAGTGGGGAAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAAACTCCTTAGAGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGAGAGATTCTAA 3480
Pol                          __N__D__I__Q__K__L__V__G__K__L__N__W__A__S__Q__I__Y__P__G__I__K__V__R__Q__L__C__K__L__L__R__G__T__K__A__L__T__E__V__I__P__L__T__E__E__A__E__L__E__L__A__E__N__R__E__I__L__
01_AE.CF.90.90CF11697        ------T-----A--A--------A---C-A--------------A-----TG----A----------AA---C-G-----------C--G---G-T--G---------------G-------G--T--G--------AT-----T-------G-----G--------G- 3429
01_AE.CN.05.FJ051            ------T-----------------A---C-A--------------A-----TG---------G-A---A-----GC-G-----------C-AG-----T-----------------CA--G------G--T------G-----AT-----T-------G-----G--------C- 3490
01_AE.CN.06.FJ054            ------T--------A--------A---C-A--------------A--------------G----G----A----C-G--------T--C--G-----T------T---------CA--G----------T----------A-----GT-------G-----G---------- 3505
01_AE.HK.x.HK001             ------T--------A--------A---C-A--------------A-----TG---------GA---A-------G------G------C--G---G-T------------------C---G------G--T----------AT-----T-------G-----G---------- 2842
01_AE.TH.01.01TH_R2184       ------T--------A--------A---C-A--------------A-----TG---------G----A----C-G-------------C--G---G-T--------------CA--G------G--T----------AT----GT-------G--T--G---------- 2679
01_AE.TH.02.OUR769I          ------T--------A--------A---C-A--------------A-----T------G---G----A----C-G-------------C--G---G-T----G-T-------CA--G------G--T----------AT-----T-------------G---------- 2679
01_AE.TH.90.CM240            ------T--------A--------A---C-A--------------A-----TG---------G----A----C-G-------------C--G---G-T--------------CA--G------G--T----------AT----GT-------G-----G---------- 3054
01_AE.US.00.00US_MSC1164     ------T--------A--------A---C-A--------------A-----TG---------G----A----C-G--C----------C--G---G----------------CA---------G--T----------AT----GT-------G-----G---------- 2685
02_AG.CM.02.02CM_4082STN     ---------------A-----A--A---C-A--------------A-----TG----A----G----A----C-G-------------C--G---G--------T-------TA--G--A---G--T--G--------AT-----T-------G--T--G--A------- 2679
02_AG.EC.x.ECU41             ------T--------A--------A---C-A--------------------TG----A---------A----C---C-----------C--G---G----------------TA-GG-AT---G--T--G--------AT-----T-------G-----G--A------- 2576
02_AG.GH.03.GHNJ196          ------T--------A--------A---C-A--------G-----A-----T--T--A--------------C-G-------------C--G---G----------------TA--G--A---G--T--G--------AT-----T-------G-----G--A------- 3509
02_AG.NG.01.PL0710           ------T--------A--------A---C-A--------------------TG----A--C-G-A---A----C-G-------------C--G---G----------------CA--G--A---G--T--G--------AT-----T-------G-----G--A---T--- 2664
02_AG.NG.x.IBNG              ------T--------A--------A---C-A--------------------TG----A------A---A----C-G------G------C--G---G----------------TA--G--G---G--T--G--------AT-----T-------G-----G--A------- 3005
02_AG.SE.94.SE7812           ------T--------A--------A---G-A------CA-GTCAG------TG----A----------A------G-------------C------G-----A----------TA--G--A---G--T--G--------AT-----T-------G-----G--A------- 2852
02_AG.SN.98.MP1211           ------T--------A--------A---C-A--------------------T-----A----------A----C-------------C--G------------T-------TA--G------G--T--G--------AT-----T-------G-----G--A------- 2677
03_AB.RU.97.KAL153_2         ----------------C-------A--------------------------TG---------------------------------------G---G---------------------------------------C-------------------------------G---------- 2706
04_cpx.CY.94.CY032           ---C--T--C--------------A---C-A--------------------T----------------A--------------T-----G---G-T------C--------CA--G----------TAC---G------T-----TA------G-----G---------- 2846
05_DF.BE.x.VI1310            ------T---------------A--A---C-A--------------------T----------------A--------------------G----------------------G-----------------------------AT-------------------G--A------- 2862
06_cpx.AU.96.BFP90           ------T--------A--G-----A---C-A--------------------T---------------------------------C--G---G-A-------------CA--G----------T-C--------A------T-------G-----G--A---T--- 3508
08_BC.CN.97.97CNGX_6F        ------T-----------------A-----A--C-----------------------A----------------C-T-----------------G--GG---------------CA--G----------T----------AT-----T-A--G--------G--A---T--- 2660
09_cpx.GH.96.96GH2911        ------T--------A-----A--A-----A--C-----------------TG---------------A----C-G-------------C------G-T-G--------G--TA--G--A---G--T--G--G-----AT-----T-------G-----G---------- 2682
10_CD.TZ.96.96TZ_BF061       ------T-----------------A--G--A--C----------------------------G---------C-G------------------G---G-----T-T-------CA--G------G--T------G-----AT-----T-------------G--A------- 2859
11_cpx.GR.x.GR17             ------T-----A-----------A---C----------------------TG---------------A----C-----------------G--GG--------------------CA--G------G--T-C--------AT-G--GT-------G--T--G--A------- 2782
12_BF.AR.99.ARMA159          ---C--------------------A-----A---------------------------------G-A-----------G---------G--------G-----------------------------------A---------------------------------G--A------- 3485
13_cpx.CM.96.1849            ------T-----A-----------A---C---------------A--C--T-----A----------A----C-----------------G--GG---------------CA--G--T---G--TAC--------AT-----T-------G---------A---T--- 2885
14_BG.DE.01.9196_01          ------T-----A-----------A--GC-A-----------C--------T-----R----------AA---C-----------A-----G---G-----------------CA--G----------T-C--------AT-G--GT-------G-----G--------G- 2997
15_01B.TH.99.99TH_MU2079     ------T--------A--------A---C-A--------------A-----TG-G---------G----A----C-G-------------C--G---G-T-------G-----CA--G----T-G--T----------AT-----T-------G-----G---------- 2879
16_A2D.KR.97.97KR004         ------T--------A--------A--G--A--------------------TG---A-----------A----C-G-----G--------G---G-T-------G-----CA--G------G--TA---------AT-----T-A-A---G-----G---------- 2843
17_BF.PE.02.PE02_PCR0155     ------------------------A--------------------------------------------A-----------G---------G--------G---------------------------A------------------------------G--A------- 2670
18_cpx.CU.99.CU76            ------T--------------A--A---C----------------A-----T----------------A---GC--------------A----G--G-----------------CA--G------G--T-T--------AT-G---T-A-----------G--A------- 2785
19_cpx.CU.99.CU7             ------T-------------------GC-------------C--------T-----A----------A--GC------------------G---G---------------------------------------------AT-------------------G--A------- 2706
20_BG.CU.99.Cu103            ------------------------A---------------------------T-------------------------C--------------G---------T-----------G--------AC-----------------T-------G-----G--A------- 2946
21_A2D.KE.91.KNH1254         ------T--------A--------------A--C--------C--------T-----A----------A---G----------T-----G---------------G---------G-------G---------------AT-------A-----G-----G--A------- 2685
22_01A1.CM.01.01CM_0001BBY   ------T--------A--------A-----A--------------------TG---------------------GC-G--------A-----G---G-------------------G----T-G--T----------AT-----T-------G--T--G---------- 2688
23_BG.CU.03.CB118            ------------------------A--------------------------T-----------------A----C-----------------G-------------------------G------G--T-C--------AT-G---T------------G--A------- 2923
24_BG.CU.03.CB378            ------T-----A-----------A--GC-A--------------------T-------------------A---C-----------------G--------T-------------G-----------AC------------T-------------G--A------- 2905
24_BG.CU.03.CB471            ------T-----A-----------A--GC-A--------------------TG-----------------AA---C-----------------G-------------------------------------C------------T-------------G--A------- 2920
25_cpx.CM.01.101BA           ------T--------A--------A---C-A--------------------T-----A----------AA-----------------------GG-A-------------TA--G------G--T-C--------AT-G---T-------G-----G--A------- 2688
27_cpx.CD.97.97CDKTB49       ------T-----A--------A--A---C-A--------------------T----------------A----C-----------------G--GG--------------CA--G----------T-C--------AT-G---T-A-----G-----G--A------- 2688
28_BF.BR.99.BREPM12609       ------T-----------------A-----A-------------------------------------A----------------------G--------G----------------G--------------------------------------------------G--A------- 2846
29_BF.BR.02.BREPM119         ------------------------A-------------------------------A-------G----A----------------------G---G----G----------------G---------------A--------------------------------G--A------- 2654
31_BC.BR.02.110PA            ------T-----------G-----A-----A--C------------------------A-----G--A---C-G----G---------G---G--------------CA--G------G--T----------AT-----T-------G-----G--A------- 2914
32_06A1.EE.01.EE0369         ------T--------A--G--A--A---C-A--------------A-----T--------------------G--------------C--G---G-A-G-------------TA--G----G-----T-C--------AT-----T-------G---------A---T--- 3115
33_01B.MY.05.05MYKL007_1     ------T--------AC-------A---C-A--------------A-----TG-----------------AA---C-G-------------C-AG---G-T---------------CA--G------G--T----------AT-----T-------G-----G---------- 2859
34_01B.TH.99.OUR2478P        ------T--------A--------A---C-A--------------A-----TG---------G----A----C-G-------------C--G---G-T------T-------CA--G-G-----G--T-------R--AT-----T-------G---------A------- 2682
35_AD.AF.05.05AF095          ------T-----------------A-----A-----------C--------T-----AG------G--A------G-------------C--G-----T-----G--------TA--G---------G-------------AT-------------------G--AC------ 2673
36_cpx.CM.00.00CMNYU830      ------T--------A--G--A--A---C-A--------------------TG---------------A---GC-G--------A--C--G---G-A-------------------G----T-G--T--G--------AT-----T-------G--T--G---------- 2685
37_cpx.CM.00.00CMNYU926      ------T--------A--------A---C-A--------------------TG-------------------GC-G-----G---T--C--G-----T-----GT--------T---G--A--AAG--TA-G--------AT-----T-------G----TTTT----G--- 2670
38_BF1.UY.05.UY05_4752       ------------------------A-------------------------------------------A----------------------G--GG----G-------------------------------A--------------------------------G--A------- 2891
39_BF.BR.03.03BRRJ103        ------------------------A--------------------------T----------------A------------G--------G--------------------------G-------------------------------------------G--T---A------- 2969
40_BF.BR.04.04BRRJ115        ------T--------------A--A---C-A--------------------T--------------G-A---G-----------AA-----G---G----G-----------CA--G-G-----G--TAC------------T-----T-------G--T--G---------- 3032
42_BF.LU.03.luBF_05_03       ------------------------A--------C-----------A-----TG-G--A----------AA---------------------G-------------------------G-------------------------------------------------G------T--- 3018
43_02G.SA.03.J11223          ------T-----A--A--------A---C-A--------------------T-----A----------A----C-G----G------C--G---G-----------------T---G--A---G--T--G--------AT-----T-------G-----R--A------- 3005
U.CD.83.83CD003              ------T--------------A--A---C-A--------------------T--------CC------------C-G-------A-----G-----------------------CA--G-G---A-G--TA----G-----AT-----T-------G-------------- 2975
U.CD.90.90CD121E12           ------T--------------A--A---C-A--------------------T--------CC-------A---C----------A-----G---G-------------------CA--G-G---A-G--TA----G-----AT-----T-------G---------------G 2679
U.GR.99.GR303                ------T--CACAG-AG-TAGT--A-----A--------C-A---------TG---------G--AC--------G---------------G---G--------------------A--G----------T-C--------AT-----T-------G-----G--T------- 2754
U.NL.95.U_NL_95_H10986_D1    ------T-----------------A---C-A-------------------TT----------G---G-A-A-C-----------------G---G-----------------CA--G----G-------C----------T-----T-------------G---------- 3365
CPZ.CD.90.ANT                T--------C-----AC----A--A-----A-----------------C--------A--C---ACT-A---G--G-----G----A-C------GT---GAGT---------TAG-G---A-A-G--TAGG--------AT-----T-A-A------T-A-C-A------C 2919
CPZ.CM.01.SIVcpzCAM13        ------TC-T-----AC-CA-A--A-----A--C-----------------T-----A--A---A---AA--G--G--C---T-AA-A--------A---AGGT-------T---G-T--T--C---CC---------AT-----T-A------------------A--AG-T- 3060
CPZ.CM.05.SIVcpzEK505        ------T-----AC-------A--A-----A--C-----T--C-----C-------------------AA---C-T--C---T-AA----------G-A-----T------------G-CA----T---AG---------AT----GT-A--T--G--T--G-----AT--- 3052
CPZ.CM.05.SIVcpzLB7          ---------------A--G-----A-----A--------------A--A--------C--C--GA---AA--GA----C-GG--T-----G---GTA--GAATTC----------G-CA-CT-T--TAG---------AT-----T-A--------T--G--A--AT--- 3035
CPZ.CM.05.SIVcpzMB66         ------T--C-----A---A-A--A-----A---------T-------A--T-----A--C-G-----AA---C-G--C--G---A----------A-----TT--------TG--A-CT-T---AC---------AT-----T---------T-----A--AT--- 3026
CPZ.CM.05.SIVcpzMT145        T-----T--T-----------A--C-----A--------------A--A--T-----A------A-T-A---GC-T--------T-AA----G---G-G--G-GT--G---------G--G-T--C--CCC------T--A-----------------------TCA------A---- 3022
CPZ.GA.88.GAB1               ------T-----A--ACGA--A--A--G--A--------------------T-----A--A---A---A-------------G-GA--A--------A--GAA---G--------T---G-T--T--C---CC---------AT-----T-A--------T--G-----AG--- 3541
CPZ.TZ.00.TAN1               G---C-T--C--------G--A--C---C-T--------C-----A-----TAT---A--A---ACT-AAG--C----C---TGA--------GTA----G-T----T--GCC-G--GA-A-G--CAGG--------AT-G--GT-A-A------T-AGC--------- 3124
CPZ.US.85.CPZUS              ------T--T-----A-----A------C---C-------------A--T-----A--A---A---A---G-----C---T-AA-A-A----G-T-----T-----T-----TG--AATT-T---C-T--G-----AA-G--GT-A-A------------A------- 3540
CPZ.CM.x.SIVcpzDP943         ------T--C--A---C-G--A--A-----A--------------A--A--T--------A-------AA--GC----C--G---A-A-A----G-A-----TT----T------G-GG-CA-G---C-G--------AA-G---T---A------T---------A---T--- 3066
CPZ.CM.x.SIVcpzMB897         ---------------A-----A--A--G--A--------------A--A--T-----A-----G----AA-----G------G--TA----------A-----TT-------G---G-TR-C--C---C-G--------AT----GT-A----------G-----AT--- 3046
N.CM.02.DJO0131              ------T-----A--A-----A--A--G--A-----------------C--T-----A--C-------AA--GC-T----G-T-AA-C------------TT-G---------G-CAACT-T---A----------AT-----T-A----------G-----AT--- 2986
N.CM.04.04CM_1015_04         ------T-----A--A-----A--A--G--A-----------------C--T-----A--C--------A--GC-T------T-AA-C----------G--TT-G---------G-TA-CT-T------------AT-------A------T--G-----AT--- 2985
N.CM.04.04CM_1131_03         ------T--------A-----A--------A--------------------T-----A-----------A--GC-T--C---T-AA-C----------G--TT-G---------G-TA-CT-T-------------GAT-------A----------G-----AT--- 2984
N.CM.95.YBF30                ------T--------A-----A--A--G--A-----------------C--T-----A--C-G-----AA--GC-C------T-AA-C------------TT-G---------G-CAACT-T------------AT-------A----------G-----AT--- 3074
N.CM.97.YBF106               ------T-----A--A-----A--A--G--A-----------------C--T-----A--C-------AA--GC-T------T-AA-C----------G--TT-G---------G-CA-TT-T---C--------AT-------A----------G-----AT--- 3072
O.BE.87.ANT70                A-----T-----A--AC--A-A--A--GC-A--------------A--C--T-A---A----G---G-AG-----G-----GT-AA-------C-----GT--T-----------G----TT---GTAG---G---------G---T-A-AG-----------A-GGT--- 3534
O.CM.91.MVP5180              A-----T-----A--A-----A--A-----A--------------A--C--T-A---A----G-----AAG----G--C--GT-AA-C-----------T--T-G-----G---G----TT---GTA---G-----A------T-A-A--------------A-AG---- 3509
O.CM.96.96CMABB637           A-----T--C--A--A-----A--A--G--A-----------------A--T-AG--A----G-----AAG-------C--GT-AA-C----------GT--T-G--------G-----TT-G-GTA---G-----A--------A-A---------G--A-GG---- 2964
O.CM.98.98CMA104             T-----T-----A--A---A-A--A--G--A--------------A-----T-A---A----G------AG----------GT-AA-------C-----GT--T-------G---G----TT---GTA---G--------------T-A-A---------G--A-GG---- 2966
O.CM.99.99CMU4122            A-----T-----A--AC----A--A--G--A--------------A--C--T-----A----G------AG----G-----GT-AA-------C-----GT--T-G-----A--G----TT---GTA--------------G---T-A-AG--G---------A-GG---- 2965
O.FR.92.VAU                  A-----T-----A--A--G--A--A---C-A--------C--------C--T-AT--A----G-----AAG----G-----GT-AA-------------T--GT--T-G--------GG-----T---GCA---G-----A------T-A-A---G-----G--A-AG--G- 3046
O.SN.99.SEMP1299             A-----------A--AC----A--A--G--A--------------A--C--T-A---A----G-----AAG-G--G-----GT-AA-------CG----GT--T-G-----A--G----TT---GTA-----------------T-A-AG--G--T-----A-AGT--- 3533
O.US.99.99USTWLA             A-----T--------A-----A--A-----A--------------A--C--T-A---A----G-----ARG----G------T-AA-C-----------T--T-G---------G----TT---GTC---G-----A------T-A-A---------G--A-GA---- 2963
```

```
B.FR.83.HXB2                AAGAACCAGTACATGGAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAGAGCCATTT...AAAAATCTGAAAACAGGAAAATATGCAAGAATGAGGGGTGCCCACACTAATGATGTAAAA 3647
Pol                         K__E__P__V__H__G__V__Y__Y__D__P__S__K__D__L__I__A__E__I__Q__K__Q__G__Q__G__Q__W__T__Y__Q__I__Y__Q__E__P__F__.__K__N__L__K__T__G__K__Y__A__R__M__R__G__A__H__T__N__D__V__K_
A1.GE.99.99GEMZ011          G------T--G--------A--------------------T---G----------------A--A--A-----------------------------G---------...--G-----A-------------------A--A-G--TC---T------------------ 2831
A1.KE.00.KER2008            ----C--T-----------A-------------------------C------------------A-------A------------C-----------G---------...------T-A--G-------------------AA---TC--------------------G- 2846
A1.KE.00.KNH1144            -------T--G--------A--------------------AC-----------G-------A--A-------A------------C-----C---------------...--------A--------G------------CAA---TC---T------------------ 2852
A1.KE.00.KSM4024            ----C--T--------G--------------------------------------------A--A-------A----------------------------------...--------A--------G-------------GA---TCC--T------------------ 2843
A1.KE.00.MSA4069            ----C--T-----------A--C--------------------------------------A--A-------A------------C---------------------...--G--------------G-------------AA---TC---T------------------ 2846
A1.KE.00.NKU3005            ----C--T--G-----G--A--------------------------------G-----------A--A----A----------------------------------...--------A--------G-------------A----TC---T----------------G- 2846
A1.RU.00.RU00051            -------T--G--------A--------------------T---G----------------A--A-------A------------------------G---------...-----CT-A-------------------A--A-G--TC---T------------------ 2948
A1.RW.93.93RW_024           -------T--G--------A--------T--------------------------------A--A-------A------------C---------------------...--------A----------------------AA---TC---A----------------G- 2849
A1.SE.95.SE8891             ----C--T--G--------A-----------------------------------------A--A-------A----------------------------------...--G-----A----------------------AA---TC---T----------------G- 2836
A1.SE.95.UGSE8131           ----C--T--G--------A-----------------------------------------A--A-------A------------------------G---------...-----------------G-------------AA-A-TC---T----------------G- 3019
A1.TZ.01.A173               -------T--G-----------C-----T-------------------T------------A--A-------A------------C---------------------...--------A----------------------A----TC---T------------------ 2846
A1.UA.01.01UADN139          -------T--G--------A--------------------T---G----------------A--A--A----A------------------------G---------...--G-------------------------A--A-G--TC-A-T------------------ 2846
A1.UG.92.92UG037            ----C--T--G-------CA--------------------------------G--------A--A-------A----------------------------------...--------A----------------------AA---TC---T------------------ 3663
A1.UG.99.99UGA07072         ----C--T--G--------A--------T--------------------------------A--A-------A---G------------------------------...-----------------G-------------AA-A-TC---T------------------ 2846
A2.CD.97.97CDKTB48          ----C--T--------G--A-----------------------------------------A--A---------------T-------------------A------...--------A--------G----------A--G--A-TCCA-------C------------ 2985
A2.CY.94.94CY017_41         --ACC--T--------G--A--C--------------------------------------A--A-------A------------------------G--A--C---...--G--------------G----------A--G----TCCA--------------A----- 3004
A.SN.01.DDI579              ----C--T--G-----G-CA-----------------------------------------A--A------AA------------------C---------------...--G----------------------------A----TC-T-T----------------G- 2845
A.SN.01.DDJ369              ----C--T--G-----G--A------------------------G----------------A--A-------A------------------C---------------...--G----------------------------AA---TC---T--T-------------G- 2848
A.SN.96.DDJ360              ----C--T--G-----G-CA-----------------------------------------A--A-------A------------------C---------------...--G---T-A----------------------A--A-TC-T-T----------------G- 2845
A.CD.02.02CD_KTB035         -------T--G-----G-A---------------------T---G----------------A--A-----T-A---------T------------------------...------T-A-------------------A--A----TC--------------------G- 3641
A.CD.97.97CD_KCC2           -G-----T--G--------A-----------------------------------------A--A-----G------------------------------------...------T-A--------G----------A--A----TCA--T----------------GG 3643
A.CD.97.97CD_KTB13          -------T--G--------A--------------------A--------------------A--A-------A---------------------------A------...------T-A--------G----------A--A----TC---------C----------G- 2851
B.AR.04.04AR151516          -------------------A--------------------------------------A-------AC-----------------------C---------------...--------------------G-------R------------------------------- 2869
B.AU.87.MBC925              -----------------------------------------------------------------------------------------------------------...--------A----------------------C---A------------------------ 3658
B.BO.99.BOL0122             -----------------------------------------------------C----------A---G----T-----------C-------T------C------...--------------------G------C--GC--A------T--T---------A----- 2867
B.BR.03.BREPM2012           ----G--------------A--------------------A------------------------------------------------------------------...-----C--------------G--------------------------------------- 3072
B.CA.97.CANB3FULL           -------------------------------------------------------------A---------C-----------------------------------...------------------------------------------------------------ 2936
B.CN.05.05CNHB_hp3          -----A-------------------------------------------------------------A-T---T-----------C----------------A--A-...-----------------------------------------T------------------ 3646
B.CO.01.PCM001              ----G--------------------------------------G---------C----------A--------A------T--------------------------...---------------------------------------A-------------------- 2852
B.GB.83.CAM1                -----------------------------------------------------C-------A---------------------------------------------...--------------------G-------A----------A-------------------- 3648
B.GB.86.GB8                 ----G---------------------------------------G-------G-----------A----T-------------------------------------...--------------------G-------A------------------------------- 3667
B.GE.03.03GEMZ010           G----------------------------------------------------G----------A-A------T---------------------------------...--------------------G--------------------------------------G 2852
B.IT.05.SG1                 -----T----------G-C---C------------------N----------------------------------------------N----T--------M----...-----------N------------------------------------------A----- 3449
B.JP.05.DR6538              ----------------G-CA------------------------G--------G-------A-----------T---------------------------------...--------------------G-------------A------------------------- 3651
B.KR.05.05CSR3              -------------------------C--T-----------T---G-------------A--------A-T------G------------------------------...------------------------------G----------------------------- 3672
B.NL.00.671_00T36           --------------------------------A----------------------------------A---------------------------------------...--------------------G-------A-G----------------A------------ 3212
B.RU.04.04RU128005          ----------------------C---------W----------------------------------------------------------A---------------...-----------G----------------A----------A-------------------G 3138
B.TH.00.00TH_C3198          -------------------------------G-------------------------------------T---------------C-------------------A-...--------------------G--------------------------------------- 2861
B.UA.01.01UAKV167           -----------------------------------------------------G----------A-A------T---------------------------------...--------------------G-------A------------------------------- 2873
B.US.04.ES10_53             ----------------G-----------------------T------------------------------------------------------------------...-----------G----------------------------------------------G- 3656
B.US.99.PRB959_03           G------------------------------C--------------------GC-------A----A-T-T--------------C---------------------...--------------------G------C--------------------------A----- 2867
C.AR.01.ARG4006             -------------------A--------T-----G--------------T-----------A----AA--G-----------------------------------C...----------------------------A-------AC--------------------G- 2831
C.BR.04.04BR013             -------------------A--------T--------------------T------------------G-G-A---------------------------A-----C...----------------------------A-------AC------------------C-G- 3111
C.BW.00.00BW07621           -------------------A--------------------T--------T--C-----A--A------A-T-A---------------------C-----A-----C...--------------------G-------A-------AC---------------------- 3000
C.CN.98.YNRL9840            -------------------A---------T-------------G-----T-----------A--------G-A---------------------C-----AT----C...-----------------G--G-------A-------AC---------------------- 2831
C.ET.02.02ET_288            ----G--------------A--------T--------------------T-----------A------G-T-A---------------------C-----A------...-----------------G--G-------A--G----AC---------------------- 2849
C.GE.03.03GEMZ033           -------------------A-----------------------------T-----------A-----A--T-A---------------------C-----A-----C...-----------------G--G-------A--------C---------------------- 2828
C.IL.99.99ET7               -------------------A-----------------------------------------A------G-T-A---------------------C-----A------...------T----G----------------A------AAC---------------------- 2828
C.IN.99.01IN565_10          ----------------------------T--------------G-----T---T-------A-----A--T-A---G-----------------C-----A-----C...--G-----A--------G--G-------A-------AC---------------------- 3028
C.KE.00.KER2010             -------------------A-----------------------G-----T-----------A--------T-A---------------------C-----A-----C...--------A-----------G-------A-------AC---------------------- 2828
C.MM.99.mIDU101_3           ----------G-------CA-----------------------G-----T-----------A----------A---------------------C-----A-----C...--------------------G-------A-------AC---------------------G 2998
C.MW.93.93MW_965            -------------------A-----------C--------G--------T-----------A--------T-A------------C--------C-----A-----C...-----------------G--G-------A-----A-TC---------------------- 2834
C.SN.90.90SE_364            -------------------A--------------------A--------T-----------A--------T-A---------------------C-----A-----C...-----C--------------G-------A-------AC---------------------- 2816
C.SO.89.89SM_145            -------------------A--C--------------------G-----T-----------A--------T-A---------------------C-----A-----C...-----------------G--G-------A-------ACA--------------------- 2868
C.TZ.02.CO178               G---------------G--A-----------------------G-----T-----------A--------T-A---G-----------------C--G--A-----C...-----------------G--G-------A-------AC---------------------- 2828
C.UY.01.TRA3011             ------------------CA-----------------------------T-----------A--A-AC--G-----------------------C-----A--T--C...------T-----------------A---A-------AC--------------------G- 2819
C.YE.02.02YE511             -------------------A-----------------------------------------A------T-T-A---------------------C-----A-----C...------T----G----------------A---A---AC---------------------- 2822
C.ZA.04.04ZASK164B1         ----------------G--T-----------------------G-----T-----------A-----A--T-AT--------------------C--G--A-----C...--G-----A-----------G-------A------AAC---------------------- 3061
C.ZA.05.05ZAPSK240B1        ----------------G--A-----------------------G-----T-----------A--A---G-G-----------------------C-----A-----C...-----------------G--G-------A-------AC---------------------- 3051
C.ZM.02.02ZM108             ------------------CA-----------------------------T--------A--A--A-----T-A---G-----------------C-----A-----C...--------------------G-------A-------AC---------------------- 3616
D.CD.83.ELI                 -------------------------------------------------------------A--A-----C--------------C--------------A------...--------------------G--------------------------------------G 3193
D.CM.01.01CM_0009BBY        G----------------------------------------------------G-------A--A---G--A--------T-T-GG--------------A------...--------------------G-------A-----A-A-------T--------------- 2846
D.KE.01.NKU3006             -------------------------------------------------------------A--A--A--G-A---------------------------A-A--A-...--------------------G-------A-T----------------------------- 2852
D.KR.04.04KBH8              ----------------------------T------------------------T-------A--A--------------------------A---------------...------T----G--------G-------A----------A-------------------- 3606
D.TD.99.MN011               G---------G------------------------------------------G----A--A--A-----------------T-GG--------------A------...--------------------G-----------A-A------T------------------ 2871
D.TZ.01.A280                G----A------------CA-----------------------------------------A--A-A---------------------------------A------...--------------------G-------A----------A-------------------- 2848
D.UG.99.99UGK09259          -------------------------------------------------------------A--A-------A---------------------------A-A--A-...----------------------------A--G---------------------------- 2852
D.YE.01.01YE386             -------------------------------------------------------------A--A-------AT--G-----------G-----------A-A--A-...--------------------G-------A------------------------------- 2846
D.YE.02.02YE516             G----------------------------------------------------G-------A--A-----------------G-GG-----------G--A------...--------------------G--------------------T------------------ 2858
D.ZA.90.R1                  -------------------------------------------------------------A--A---G-T--------------C--------------A------...--------------------G----------CA--------------------------- 2993
F1.AR.02.ARE933             ----------------G--A-----------------------------------------A--A--------G------------------------------C--...--------A-----------G---------------TC---------------------- 2941
F1.BE.93.VI850              G---------------G--A-----------------------------------------A--A--AG-C--G-----------------------GA-C------...--------A-----------G-------A-G-----TCG--------------------- 2984
F1.BR.01.01BR125            ----------------G--A-----------------------------------------A--A--------G---------------------------------...--------A-----------G-------A-------TC------------------G--- 3092
F1.ES.x.P1146               ----------------G-CA--C-------------------------T------------A--A--------G---------------------------A-----...--------A-----------G-------A-------TC-A-------------------- 2924
F1.FI.93.FIN9363            ----------------G--A---------------------------C--A--T-------A--A--------A--------------------C-G----------...--------A-----------G-------A-------TC---------------------- 2976
F2.CM.02.02CM_0016BBY       ----------------G--A--------------------T--------------------A--A-----T-AT--------------------------A---CA-...-G------------------G-------A--G----TC---------------------- 2837
F2.CM.95.MP257              ----------------G--A--------------------T--------------------A--A--A--C-A------------------C------------CA-...--------------------G----------G--A-TC---------------------G 2856
F2.CM.97.CM53657            ----------------G--A--------------------TC-----------G-------A--A-----C-A-------------------------------CA-...--------------------G----------G----TC---------------------- 2837
G.BE.96.DRCBL               -------T-----------C---------------------C-----------G-------A--A----T-------------------G---------------A-...----------------------------A-GG-G--TC---------------------G 3604
G.CM.01.01CM_4049HAN        -------T-----------C--------------------A------------G----------A----T--A--------------------------------AC...--------A-------------------A--A-G--TC---------------------- 2849
G.CU.99.Cu74                -------T-----------C--------------------AA-----------G-------A--A---A-C-A---------G----------------------AC...--G-------------------------A--G-G--TC---------------------- 3251
G.ES.99.X138                -------T-----------C--------A-----------A------------G-------A--A----T--A---G--------C-----C-------------AC...----------------------------A--G-G--TC---------------------- 3088
G.GH.03.03GH175G            ----G--T-----------C--------------------A------------G-G-----A--A----T--A--------------------------------AC...--G-----A------------------CA--G-G--TC-A----------------T--G 3687
G.KE.93.HH8793_12_1         -------T-----------C--------------------G------------G----A--A--A-AA-----A--------------------C----------AC...----------------------------A--G-G--AC---------------------- 3047
G.NG.01.01NGPL0669          G------T-----------C--C------------------------------G-------A--A----CC-A--------------------------------AC...--------------------G-------A--G-G-ATC--------------------GG 2852
G.PT.x.PT2695               -------T-----------C--------A-----------A------------G-------A--A----C------G--------C-----C--C----------AC...----------------------------A--G-G--TC---------------------- 3589
G.SE.93.SE6165              G------T-----------C--------------------A------------G----A--A--A----T--A--------------------------------AC...--------------------G-------A--G-G--TC---------------------- 3044
H.BE.93.VI991               ------------------CA--------T-----------A-----------------------A----C--A---G-----------------------------C...-----------G----------------A-------TC---------------------G 3036
H.BE.93.VI997               G------------------A--------T-----------------------------------A----C--A---------------------------------C...--------A--G----------------A------AAA-----------G---------- 2971
H.CF.90.056                 G-------A----------A--------T----------------------------G------A--------------------------------G---------...-----------G----------------A------AAC----------------A----- 2994
J.CD.97.J_97DC_KTB147       -------------------A---------T--G-------A------------G-------A--A----T--A----------------------------------...-----C-----------G----------A--G----AC--------------------G- 2841
J.SE.93.SE7887              ----------------G--A---------T--G-------A------------G-G-----A--A----TG-A------------------------G---------...-----C-----------G----------A--G----A----------------------- 2961
K.CD.97.EQTB11C             ----------G-----G--A--C-----------------------------------------A-----C--------------------A-------------A-...------T----------G--G-----------A---TC---------------------- 2843
K.CM.96.MP535               ----G-----G-----G--A-----------------------------------------A--A---A-T-A------------------A------------CA-...-----------------G--G--------------ATC---------------------G 2843
```

```
B.FR.83.HXB2                AAGAACCAGTACATGGAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAGAGCCATTT...AAAAATCTGAAAACAGGAAAATATGCAAGAATGAGGGGTGCCCACACTAATGATGTAAAA 3647
Pol                         K__E__P__V__H__G__V__Y__Y__D__P__S__K__D__L__I__A__E__I__Q__K__Q__G__Q__G__Q__W__T__Y__Q__I__Y__Q__E__P__F__.__K__N__L__K__T__G__K__Y__A__R__M__R__G__A__H__T__N__D__V__K_
01_AE.CF.90.90CF11697       --A-C--T--G--------A-------------------A--G-----------------A--A------A------------------------------------...--------A-------------------AA---TC---T---------------G- 3596
01_AE.CN.05.FJ051           --ACC--T--G--------A----------------------G-------GG-------A--A--A--G-A------------------------------------...--------A--------G--GG----C---AA---TC---T---------------G- 3657
01_AE.CN.06.FJ054           --ACC--T--G--------A-------------------T------------G-------A--A------A--------------------------A---------...--------A-----------G-----C---AA---TC--------------------G- 3672
01_AE.HK.x.HK001            --ACC--T--G--------A----------------------G--------G-----------A--A---A------------------------------------...--------A-------------------AA---TC---T---------------G- 3009
01_AE.TH.01.01TH_R2184      --ACC--T--G--------A----------------------G--------G-------A--A----G-A------------------------------------...--------A----------------C---AA---TC---T--T------------G- 2846
01_AE.TH.02.OUR769I         --ACC--T--G--------A----------------------G--------G-------A--------A------------------------------------...--------A-G-----------------AA---TC---T---------------G- 2846
01_AE.TH.90.CM240           --ACC--T--G--------A----------------------G--------G-------A--A----G-A------------------------------------...--------A----------------C---GAG--TC---T---------------G- 3221
01_AE.US.00.00US_MSC1164    --A-C--T--G-----G--A-------------------------------G-------A--A----G-A---------------------------G--A------...-------------------------C---AAG--TC---T---------------G- 2852
02_AG.CM.02.02CM_4082STN    -------T-----------A-----------GA-----------------------A--A--A----A---------------G-------------------...--------A------------------A--A----TC------------------G- 2846
02_AG.EC.x.ECU41            --C----T--------G--A--------T--A-------T--G--------------A--A--A----A------------------------------------...--------A------------------A--A----TC-------------------- 2743
02_AG.GH.03.GHNJ196         -------T-----------C--C---------G-----------G---------------A--A------A---------------------------G--A------...------T-A------------------A--A----TC-------------------- 3676
02_AG.NG.01.PL0710          -------T-----------A------------A-----------G---------------A--A------A---------------G--A--------A------...--G----A------------------A--A----TC-------------------- 2831
02_AG.NG.x.IBNG             -------T--------G--A------------A-----------G-------T-------A--A------A------------------------------------...------T-A------------------A--A----TC-------------------G 3172
02_AG.SE.94.SE7812          -------T-----------A------------A-----------G---------------A--A---G-A--------------------C------------...--------A------------------A--A----TC-------------------- 3019
02_AG.SN.98.MP1211          -------T-----------A------------A------------------G--------A--A---G-A---------------G-------------------...--------A------------------A--A----TC-------------------- 2844
03_AB.RU.97.KAL153_2        ------------------------------------------G--------------------A-----------------------------------------...--------------------------C---------------------C------ 2873
04_cpx.CY.94.CY032          ----------------G-CA--------------------------------------------A-------T-----------------A----------CA-...-------------G--G----------CC--ATC----------------T-G- 3013
05_DF.BE.x.VI1310           ----------------G------------------------------------G-------A--A---T--G-------------------------A-----...--G-----A----------G--------GGC-------T------------------ 3029
06_cpx.AU.96.BFP90          ----------------G--A------------G-----------------------G------------T-C-------------C----A----G-----CA-...-----C-----------G--G------------A-A-TC-------------------- 3675
08_BC.CN.97.97CNGX_6F       -----------------CA-------------------A--G-----T----------A-------G-A----------------------C----A----C...----------------G--G-------A------AC-------------------- 2827
09_cpx.GH.96.96GH2911       ----G--T--G-----G--A----------------------G----------------A--A------A------------------------------------...--------------------------A----TC---T-----G---------G- 2849
10_CD.TZ.96.96TZ_BF061      -------------------A------------G------------T------------A-------G-A-------------------------A--CAC...-------------------G--------A--G---AAC-----------C--------- 3026
11_cpx.GR.x.GR17            -------------------A----------------------------------G-------A--A--TT--A--G-------------A----------T--C...-----C-----------G---------A--G---AC---------C--------C-G- 2949
12_BF.AR.99.ARMA159         ------------------------------------------------------G------------------T----------------C-T-------------...---------------------------------G----------------------- 3652
13_cpx.CM.96.1849           -------T-----------C-----------GAG-----A-------TG----------A--A--A----AA-------------------------------AC...-----C--------------------A--G-G--TC---------A------------ 3052
14_BG.DE.01.9196_01         -------T-----------C--------A----------A-----------G-------A--A----T------G--------C-----C------------AC...-------------G-----G----------A--G-G--TC-------------------- 3164
15_01B.TH.99.99TH_MU2079    ---CC--T--G-------CA----------------------G--------G-------A--A----G-A------------------------------------...-------------------------C----AA---TC---T---------------G- 3046
16_A2D.KR.97.97KR004        ----C--T--------G--A--C----------------------------G-------A--A--A-C--A--------------------------A-----...-----C--A-------G----------A--G----TCCA------------------- 3010
17_BF.PE.02.PE02_PCR0155    -------------------A----------------------------------------------------------------G------------------...-------------------------------------------------A----- 2837
18_cpx.CU.99.CU76           ----G--T--G-----G--A----------------------GG---------------A--A--AT-T-A------------------------------------...--G--C--A------------------A--G---ATC---T----------------G 2952
19_cpx.CU.99.CU7            ----------G----------C------------------------------G---A--A--A------A---------------------------A------...-------------------------A-------------------------------- 2873
20_BG.CU.99.Cu103           -------------------------------------------------------------------------------------------A-----...-------------------G------------AAA--------------------- 3113
21_A2D.KE.91.KNH1254        ----------------G--A--C------------------A-------------------A--A------A--------------------------A------...--G-----A--------G-----------A--AC---TC-------------------- 2852
22_01A1.CM.01.01CM_0001BBY  G---G--T--G-----G--C-------------------------------------A--A--A----A------------------------------------...------T-A-------------------A----TC---T-----G------------ 2855
23_BG.CU.03.CB118           ---------------------------------------------------------------T--------------------------A-----...-------------------G------------AAA--A------------------ 3090
24_BG.CU.03.CB378           -----------------------------------------G------------------------------------------------A-----...-------------------G------------AAA-------------------- 3072
24_BG.CU.03.CB471           ---------------------------------------A--------------------------G--------------------------A-----...-------------------G------------AAAA------------------- 3087
25_cpx.CM.01.101BA          ----C--T--G--------C-------------------T--------------------A--------A----------------------G------A-...-------------------------A--G-G--TC-------------------- 2855
27_cpx.CD.97.97CDKTB49      -------------------A------A------------A-----------G-------A--A----T--A-----------------A----------G--C...-----C--------------G-------A--G-----C-------------------GG 2855
28_BF.BR.99.BREPM12609      ----------------------------------------------T---------------------A---------------G-------------------...-------------C-------------C-------------A------------ 3013
29_BF.BR.02.BREPM119        ---------------------------------------T------------------A-----------------------------------------------...----------------------------C-------------------------- 2821
31_BC.BR.02.110PA           -------------------A--------------------------T------------A--A----G-A-------------------C----A----C...----------------------------A------AC-------------------G- 3081
32_06A1.EE.01.EE0369        ----G-----G-----G--A--C----------------------------G-------A------A-T-A-----------------A----G-----CA-...-----C-----------G--G-------A---A-A-TC-------------------- 3282
33_01B.MY.05.05MYKL007_1    --ACC--C--G--------A--------T-------------G--------G-------A--A----G-A--------------------------A------...--------A----------------C---GA---TC---T---------------G- 3026
34_01B.TH.99.OUR2478P       --ACC--T--G--------A----------------------G--------G-------A--A----G-A------------------------------------...------T-A----------------C--G-AA---TC---T---------------G- 2849
35_AD.AF.05.05AF095         ---------------------------------------------------G-------A--A------AT-------------------------------------...--------A--G-----G----------A--AA---TC---G------------------ 2840
36_cpx.CM.00.00CMNYU830     ----C--T--G--------C-----------------------------------------A--A--AG---A------------------------------------...--------A---------------------A----TC---T--T---------A----- 2852
37_cpx.CM.00.00CMNYU926     -------T-----------C-TA----T----GA---G--A-----------.C-------A--A------A-----------------T------AT-----A..-----C--A-----------G-------A--A----TC---T-----C---------G- 2837
38_BF1.UY.05.UY05_4752      G-----------------------------------------G-----------A-----A---A-T-A--G------------------------------------...------------------------A-G------------------------------ 3058
39_BF.BR.03.03BRRJ103       --A-T----------------------------------A----------------------A-----A------------------------------------...----T---------------------A-----A-------------------------G- 3136
40_BF.BR.04.04BRRJ115       -G-------------G--A-----------G--------------------------A----------G----------------------------------...--------A-----------G---------A----------------C------------ 3199
42_BF.LU.03.luBF_05_03      -------------------A-------------------------------------------A--A-------------------------------------...--G--------------------G-------------------------------- 3185
43_02G.SA.03.J11223         -------------Y-----A-----------A-------A-------------------A--A------A---------------------------A-----...--------A-------------------A--G-G--TC-------------------- 3172
U.CD.83.83CD003             -------------------A----------------------G-------------------A-------A------T-----------C-------------...-----C-----------------------A--G---TC-------------------- 3142
U.CD.90.90CD121E12          -------------------A-------------------T-------------------A--A------A------T-----------C-------------...-----C--A-------------------A--G---TC-------------------- 2846
U.GR.99.GR303               ----------G-----G--A-----------------------------------------A-----C-----------------A-----------A-...-------------G-----C------G-A---TC---------------A----- 2921
U.NL.95.U_NL_95_H10986_D1   ----------------G--A-------------------T-------------------A--A------A------------------A-----------A-...--------A-------G--G-----------G--A-TC-------------------- 3532
CPZ.CD.90.ANT               -GC-GAA-A--G-G---TAC--C---C-G--TGGTCT-CCAC---A----ACC--------A--A---TC---A--------C--------A------A-TGA-GGA...---CTGT-A---G-------------T--GCCT-CA--AA-T-----------G--T-GG 3086
CPZ.CM.01.SIVcpzCAM13       GTAC-----------G--T------A-T---GAT--G--GC-C----------G---A--A----AG---G--G-----------G--C-T---G--A-A-CA-...-G----T-A--G---------------C---CA---ATCAA-A--T--A-------A-T-G- 3227
CPZ.CM.05.SIVcpzEK505       -------T--------G-CC------A-----GAC-----A-----------G--C----------A----T-------------G----------CTT-CA-...------T-A--G-----------------A-------TC-A-T--T--------CA---G- 3219
CPZ.CM.05.SIVcpzLB7         --A-C-----------G--C-----------GG-T------C-G---------G---A--A--A--A--GA-T--------C--------C-----G--A-A---...-----------------------C---CAAG--TC---T--T--C------------- 3202
CPZ.CM.05.SIVcpzMB66        ----------------G-CC-----------C--------A----------------------A--T-----A--G-----T----------TC--G--A-AG-A-...--G--CT----------G--------T--G----ATC---A--T---------------- 3193
CPZ.CM.05.SIVcpzMT145       -----------------CC--C---------TGA---G--A---------------------A--A--G--T--G--------------A-T---G-----CA-...------C--A--G-------------GC-A-CAA--A-C---A--T---------A-T-GG 3189
CPZ.GA.88.GAB1              GCAC------------G--A--C------T---GAC-----GC-T------------------A-----CA-CT----G-----T-----G--A-T---G--A---CA-...--G---T----G-----G--------C---CAA---TCA--A-----A-----CA-C-GG 3708
CPZ.TZ.00.TAN1              -----AAG--T--A----CA--C------T--TAA-TT-CCTC-GCA-----C----------------------A--G-----------G--A-----G--AGA-GGG...------T-A----------------------A-TCACCA---A-------C-----GA---G- 3291
CPZ.US.85.CPZUS             -G--------------G--C-------------GA------A---G---------G---A--A--A--AAGGA-T--G--------C-------T------A-GGCA-...--G-----A--G---------------C---CAA--ATCA--A--T---------A-C-G- 3707
CPZ.CM.x.SIVcpzDP943        ----C-----G-----------C---------GAG--G--G------------G---A--A-----AA-CA-T--G----T----------T--G--A-AGCA-...--G--CT----------G--G-----T--GCAA--ATCA--A--T--------CA---G- 3233
CPZ.CM.x.SIVcpzMB897        G---C--T--------G-CC--C---------C--------A--G-----------------A--A--T---------------T-----G--------G--A-AGCA-...----------------------------A------A-C------TT-A------------ 3213
N.CM.02.DJO0131             G------CC----------C--C---------GG------AC-------------T--A--A--A--A----AT--G-----------G--------G------A-...------T-A---------------------A--G---ATC------T---------A----- 3153
N.CM.04.04CM_1015_04        -------CC----------C------------GA------A--------------T--A--A--A--A----A-T--G-----------G--------G--AT--YA-...------T-A--G-------G--------A-----AATC------T---------A----- 3152
N.CM.04.04CM_1131_03        -------CC----------C------------GA---G--A--------G-----T--A--A--A--A----A-T--G-----------G--------G---T--CA-...------T-A--G-------G--------A------ATC------T---------A----G 3151
N.CM.95.YBF30               -------CC-G--------C------------GG------A---G----------T--A-----A--A-----T--G-----------G--------G---TT-CA-...------T-A-----------G--------A------ATC------T---------A----- 3241
N.CM.97.YBF106              -------CC----------C------------GG------A--------------T--A--A--A--A-----T--G-----------G--------G------A-...------T-A-----------G--------A--Y---ATC---T--T---------A----- 3239
O.BE.87.ANT70               --C-------G-----G--A--C---C-A--TGAT--G--TC--TGG-TTA-T--T--------A--AGGG-AG--------T--C--G--A-----G--AGA-CA-...--G--C--C--------G------A-T--GCAA-A--CCT-------A------A---G- 3701
O.CM.91.MVP5180             ----G--------------A-----CC-G--TGAC--------GTGG-TTAGT--T--------T--AG----G--------T--C--GG-A-----G--TGA-CA-...--G--C--T------------------T--GCAA-A--CCT-------A------A---G- 3676
O.CM.96.96CMABB637          ----------------T--A------C-A--TAAC---------TGG-TTA-T--T--------T--AA----G--------T-----G--A-----G--TGA-CA-...-----C--C------------------T--GCAA-A--CCT-------A-----GA---G- 3131
O.CM.98.98CMA104            ----------G-----G--A--C---C-A---GAT--------GTGG-TTAGTR-T--------A--AG----G--G-----T--C--G--A-----G--TGA-CA-...--G--C--C--------G-----CA-T--GCA--AA-CAT-------A------A---G- 3133
O.CM.99.99CMU4122           ----------------T--A------C-G--TGAC---------TGG-TTA-T--T--A-----A--AG----G--------T-----GG-A-----G---GA-CA-...-----C--C-----------------A-T--GCAA----CCT-------A------A---G- 3132
O.FR.92.VAU                 ----G-----------G--A--C---C-A--TGAT---------TGG-TTA-C--T--A-----A--AG-G--A--------T--C--G--A-----G--TGA-CA-...--G--C--C--------G--------T--GCAA-AA-CCT-------A------A---G- 3213
O.SN.99.SEMP1299            ----------G-----T--A--C---C-A--TGAC---------TGG-TTA-T--T-----------ATGG--A--------T-----G--A-----G--TGA-CA-...--G--C--C--------G------A-T---CAA-A--CCT-T-----A------A---G- 3700
O.US.99.99USTWLA            ----G-----------T--A--C---C-A--TGAC---------TGG-TTA-C--W--------A--AG----G------T-T--C--GG-A--C--G--TGA-CA-...--G-----C---------------A-T---CA--A--CCT-------A------A---G- 3130
```

```
B.FR.83.HXB2               CAATTAACAGAGGCAGTGCAAAAAATAACCACAGAAAGCATAGTAATATGGGGAAAG...ACTCCTAAATTTAAACTGCCCATACAAAAGGAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCCTGAGTGGGAGTTTGTTAATACCCCTCCCTT 3814
Pol                        _Q__L__T__E__A__V__Q__K__I__T__T__E__S__I__V__I__W__G__K__.__T__P__K__F__K__L__P__I__Q__K__E__T__W__E__T__W__W__T__E__Y__W__Q__A__T__W__I__P__E__W__E__F__V__N__T__P__P__L
A1.GE.99.99GEMZ011         --G-------CA-TG---------G-GG-----------------------------...-------------G---A-----------A------------G---------TG-----C-----G--T--------------A-----------C------------C- 2998
A1.KE.00.KER2008           ------G----A-TG-----G---G-GGT--TG------------------------...-----------------A-----------A--G--------G----------TG--C--------G--T--------------A-----------C------------C- 3013
A1.KE.00.KNH1144           ------G----A-TG--A---------G---TG---G--------------------...-------------G---A-----------A--G-------------------GG--C--------G--T-----------C--A-----A-----C-----------T-- 3019
A1.KE.00.KSM4024           --G---G----T-TG--------GG--GT--TG------------------------...---------------T-A-----------A----------------------TG--C--------G--T--------------A-----A-----C--C--------TC- 3010
A1.KE.00.MSA4069           -----------A-TG--------GG-GGT--TG------------------------...-----A---------T-A-----------A--G-------------------TG--C--------G--T--------------A-----A-----C-----------TC- 3013
A1.KE.00.NKU3005           ------G----A-TG-------G-G-GGT--TG------------------------...-----------------A-----------A-----------G--G-------TG--C--------G--T--------------A-----A-----C-----------TC- 3013
A1.RU.00.RU00051           ----------CA-TG---------G-GG-----------------------------...-------------G---A----------GA------------G---------TG-----C-----G--T--------------A-----------C-----------TC- 3115
A1.RW.93.93RW_024          ------G----A-TG--------GG-GGT--TG------------------------...-----------------A-----------A--G-------------------TG--C--------G--T--------------A-----------C-----------TC- 3016
A1.SE.95.SE8891            ------GT---A-TG--------GG--GT--TG--------------T---------...---------------T-A-----------A-----------G----------TG--C-----------T--------------A-----A-----C-----------TC- 3003
A1.SE.95.UGSE8131          ------G----A-T----------G-GGT--TG------------------------...-----------------A-----------A--C--------T----------TG--C--------G--T--------------A-----A-----C-----------TC- 3186
A1.TZ.01.A173              ------G----A-TG--------GG-GGT----------------------------...-----------------A----------------------------------TG--C--------G--T--------C-----A-----A-----C-----------TC- 3013
A1.UA.01.01UADN139         ----------CA-TG---------G-GG----------------------------A...-------------G---A-----------A------------G---------TG-----C-----G--T--------------A-----------C-----------TC- 3013
A1.UG.92.92UG037           -----GG----A-TG--------GG-GGT--TG------------------------...-----------------A-----------A----------------------TG--C--------G--T--------------A-----------C-----------TC- 3830
A1.UG.99.99UGA07072        ---C-GG----A-TG---------G-GGT-CTG------------------------...---------------T-A-----------A----------------------TG--C--------G--T--------------C-----------C-----------TC- 3013
A2.CD.97.97CDKTB48         -----------A-----A---------G-----------------------------...-T-----------G---A-----------A-----------G-----------G----GG-----G--T--------------------------C-----------TC- 3152
A2.CY.94.94CY017_41        --G--------A-----A-------------TG-----------G------------...---------------T-A-----------------------G---------G-G-----------G--T--------------------------C-----------TC- 3171
A.SN.01.DDI579             ------GT---A-TG---------G-GGT------------------------G---...--C----------G---A-----------A----------------------TG-----------G--T----AT--C-----A-----------C-----------TC- 3012
A.SN.01.DDJ369             ------GT---A-T---------G---GT----------------------------...--C--------------A-----------A-----------C-----------G-----------G--T-----T--------A-----------C-----------TC- 3015
A.SN.96.DDJ360             ------GT---A-TG---------G-GGT------------------------G---...--C--------------A-----------A-----------C----------TG-----------G--T-----T--------A-----------C-----------TC- 3012
A.CD.02.02CD_KTB035        --G--------A-TC---------G-GG-TGT---G--------------------A...--------G----G------T-----G--A-----------GG---------TG--T--------G--A--------------A-----------C------------C- 3808
A.CD.97.97CD_KCC2          ---C--T----A-T----------G-GG-----------------------------...-------------G---A--------G--A-----------GG---------TG--T-----------T--------------A-----------C-----------TC- 3810
A.CD.97.97CD_KTB13         -----------A------------G-GG--TT---G--------G-----------A...-------------GGT-A--------G--A-----------G----------TG--T--------G--T--------------A-----------C-----------TC- 3018

B.AR.04.04AR151516         ------------------A-C------G--------G----------------C---...-T-----------G-T-A-----------A-----------TG---------TG-----------------T-----------------------C-------------- 3036
B.AU.87.MBC925             --------------------G------------------------------------...-T---------------A-----------A-----------------------------------------------------------------C-------------- 3825
B.BO.99.BOL0122            --------G------------------G-----------------------------...-----------------A--------------------------------------------------------------C--------------C-------------- 3034
B.BR.03.BREPM2012          -----------------------G-----------------------------C---...-T-----------G-T-A-----------A--G---------G---------T---------------T--------------------------C-------------- 3239
B.CA.97.CANB3FULL          ------------------A--C-----------------------------------...-----------------A------------------------G----------G--------------T--------------------------C-------------- 3103
B.CN.05.05CNHB_hp3         ---------------------------G-----------------------------...-----------------A------------------------G----------------------------A-----------------------C-------------- 3813
B.CO.01.PCM001             -----------A---------------G-----------------------------...-----------------A-----------A------------G----------------------------------C-----------------C-------------- 3019
B.GB.83.CAM1               ---------------------------G-----------------------------...---------------T-A-----------------------TG----------T--C--------------------------------------C-------------- 3815
B.GB.86.GB8                --G------------------------G-----------------------------...--------------G--A-----------A------------G---------TG-----------G-----------------------------C-------------- 3834
B.GE.03.03GEMZ010          ---------------------------G-----------------------------...-T-----------GGT-A-----------A---------------------G----A--------------A-----------A-----A-----C-------------- 3019
B.IT.05.SG1                -----G----------------------G--TN-----------G------------...--C----------G---A-----------A------------G----------------------------------------------------C-------------- 3616
B.JP.05.DR6538             ---------------------------G-----------------------------...-----------------------------A------------G---------TG-----------------------------A-----------C-------------- 3818
B.KR.05.05CSR3             ---------------------------G-TCA-----------------------G-...-T---------------A-----------A-------------T--------TG--A--------G-----------C-----------------C-------------- 3839
B.NL.00.671_00T36          ---------------------------G-T---------------------------...---------------T-A-----------A------------G----------------------------------------------------C-------------- 3379
B.RU.04.04RU128005         ---------------------------G-------------------C---------...-------------G---A-----------A------------G------------------------------------------------------------------- 3305
B.TH.00.00TH_C3198         ---------------------------G-------------------------G---...-----------------A------------------------G----------------------------A---G-------------------C-------------- 3028
B.UA.01.01UAKV167          --------G------------------G---T-------------------------...-------------G-T-A----------------------------------TG--A--------------------------------------C-----T-------- 3040
B.US.04.ES10_53            -----------A---------------G----------------G------------...-T---------------A-----------A------------G----------------------------------------------------C-------------- 3823
B.US.99.PRB959_03          --G------------------------G-----------------------------...-------------G---A-----------A-----------------------------------------------------------------C-------------- 3034

C.AR.01.ARG4006            --G------------------------G--CAG---------A--------------...--C----------G-T-A-----C--G--A-----------GT-------------C---------------------------------------------------C- 2998
C.BR.04.04BR013            --G------------------------G--CAG-----------------------A...--------G----G---A-----C--G--A------------G-------------C--------------------------A------------------------C- 3278
C.BW.00.00BW07621          --G--------------------G---G--G----------------C---------...-----------C-G---A-----C-----A-----------G--------------C-----------T-----------------------------C---------C- 3167
C.CN.98.YNRL9840           --G-----------T-----G------G---TG------------------------...-------------G-T-A-----C-----------------G--------------C--------------------------------A-------------------- 2998
C.ET.02.02ET_288           -TG--------------------G---G--CT-------------------------...-T-----------G-T-A--T--C--G--A--------------------------C--------G-----------------A-----A------------------C- 3016
C.GE.03.03GEMZ033          --G------------------------G---TG-----------G------------...-------------G---A-----C-----A-----------G--------------T---------------------------------------------------C- 2995
C.IL.99.99ET7              --G---GT--G----------------G-T-TG--GG--------------------...-T-----------G---A-----C-----A-----------G--------------C--------------------------T------------------------C- 2995
C.IN.99.01IN565_10         --G---G--------------------G--CTG--------------------G--A...-------------G-T-A-----C-----A--------------------------C--------------------------------A------------------C- 3195
C.KE.00.KER2010            --G------------------------G---------------------------G-...-------------G-T----A--C-----A-----------G--------------C----------------------------------------------------- 2995
C.MM.99.mIDU101_3          --GC----------T-----G------G---TG------------------------...-------------G-T-A-----C-----------------G--------------C--------------------------------A-------------------- 3165
C.MW.93.93MW_965           --G---------A--------------G-T-TG------------------------...--C----------G---A-----C-----A-----------G--G-----------C----------------------------------------------------- 3001
C.SN.90.90SE_364           --G------------------------G-------G---------------------...-------------G-T-A-----C-----A-----------G--------------C--------------------------A------------------------C- 2983
C.SO.89.89SM_145           --G------------------------G--CAG------------------------...-------------GG--A-----C-----A-----------G--------------A---------------------------------------------------C- 3035
C.TZ.02.CO178              --G------------------------G---TG------------------------...-------------GG--A-----C-----------------GG-------------C---------------------------------------------------C- 2995
C.UY.01.TRA3011            --G------------------------G--CT---G------A--------------...-------------G-T-A-----C--G--A------------G-------------C--------------------------------------------T------C- 2986
C.YE.02.02YE511            --G------------------------G---TG------------------------...-T-----------G---A-----C-----------------G--------------C--------------------C-----A-----------G------------C- 2989
C.ZA.04.04ZASK164B1        --G-----------------G------G--TT-----------------------G-...-------------G---------C-----A-----------G-----------G--C---------------------------------------------------C- 3228
C.ZA.05.05ZAPSK240B1       ---------------------------G-A-T-------------------------...-------------G---A-----C-----------------G--------------A---------------------------------------------------C- 3218
C.ZM.02.02ZM108            --G------------------------G--CTG--G---------------------...GT-----------G---A-----C-----A-----------G--------------C--------------------------------A--------C---------C- 3783

D.CD.83.ELI                ------G---------------G----T----------------G----------G-...-------------G---A---------------------------------G-------------------T-----------------A-----C-----------T-- 3360
D.CM.01.01CM_0009BBY       ---------------------------G--------------A--------------...-T-----------G---A-----------A------------G---------TG-----------G-----T-----------------------C-----------T-- 3013
D.KE.01.NKU3006            ---------------------------G--CA----T-T-----G-----------A...-------------G---A-----------A-----------------------------C-----G-----------------------------C-----------T-- 3019
D.KR.04.04KBH8             ---------------------------G--CA-----------------------G-...-------------G-T-A------------------------G-------------A--------G-----------------------------C--C--------T-- 3773
D.TD.99.MN011              ---------------------------T--CT------------G------------...-------------G-T-A------------------------G---------TG-----------G-----T-----------------------C-----------T-- 3038
D.TZ.01.A280               ---------------------------G--CT----G-T-----G------------...-T-----------G---------------------------------------------------------------------------------C-----------T-- 3015
D.UG.99.99UGK09259         ---------------------------G--CA----T-T------------------...-------------G---A---------------------------------G-------------G-----------------------------C-----------T-- 3019
D.YE.01.01YE386            -----------------A---------G--CA----T-T-----G------------...-------------G---A-----------------------------------------------G--------------A--------------C--C--------T-- 3013
D.YE.02.02YE516            -------------T-------------G--CA------------G------------...-------------G---A------------------------G---------TG-----------------T-----------------------C-----------T-- 3025
D.ZA.90.R1                 ------G----------------TT--T----------------G------------...-------------G---------------------------------------------------G-----------------------A-----C-----------T-- 3160

F1.AR.02.ARE933            -----------A------------G--G-TCA-------------------------...--C-----G--------A-------T---A--G--G-----T----------TCAGTAT------------------------A-----------C--------C--TC- 3108
F1.BE.93.VI850             -----------A-----A-----G---G-TTT---------------------G---AGAT-------G--------A-------TG--A--G--------T--------------T--------------------------A-----------C--------C--TC- 3154
F1.BR.01.01BR125           --------G--A-----------G---G-TCA------------------------A...--------G----G---A-----------A--G--------T-----------------------------------------A-----------C--------C--TC- 3259
F1.ES.x.P1146              -----------------A-----G---G-TCAG---T-------G--------G---...--------G--------A-----------------------GG-------------C--------------T-----------A-----------C--C-----C--TC- 3091
F1.FI.93.FIN9363           -----------A-----------G---G-TCT-------------------------...--------G--------A--T--T-T---A--G--------T-----------------------------------------------------C--------C--TC- 3143
F2.CM.02.02CM_0016BBY      -----------A-TG--A---------G--------G--------------------...GTC----------G---A-----------A-----------------------------------------------------A-----------C-----------TC- 3004
F2.CM.95.MP257             -----------A-T---------G---G--------G-------------------A...GT-----------G---A-------------------------T---------G--------------T--------------A-----------C-----------TC- 3023
F2.CM.97.CM53657           -----------A-T-------------G--------G-------------------A...GT-----------GG--A-----------A-----------------------------------G-----------------A-----------C-----------TC- 3004

G.BE.96.DRCBL              -----------A-T-------------G-------GG-----A--------------...-T-------------T-A--T---A----A------------GTG--------------------G-----T-----------------------C-----------T-- 3771
G.CM.01.01CM_4049HAN       --G--------A-T-------------G-------G---------------------...-TC--------------A--T-----G--A------------GT---------------------G-----------------------------C-----------TC- 3016
G.CU.99.Cu74               ------GT---A-T-------------T-------G--------------------A...-----------------A--T-----G--A------------GT------------C--------G-----------------T-----------C--------A--TC- 3418
G.ES.99.X138               -----------A-T-------------G-------G------A-------------A...-----------------A--T-----G--A------------GT------------A--------G-----------------C-----A-----C-----------TC- 3255
G.GH.03.03GH175G           --G--------A-TG------------G-------GG-------------------A...GT---------------A--T----AG--A------------GTG-----------------------T--------------------------C-----------TC- 3854
G.KE.93.HH8793_12_1        -----------A-T-------------G-------G---------------------...-----------------A--T-----G--A------------GT------------A--------G-----------------T-----------C--C--------TC- 3214
G.NG.01.01NGPL0669         -----------A-----A---------G-------G--------------------A...GT------------G--A--T---AGG--A------------GT------------C--C-----G-----------------T-----------C--------C--TC- 3019
G.PT.x.PT2695              -----------A-T-------------G-------G--------------------A...-----------------A--T-----G--A------------GT------------A--------G-----------------C-----A-----C-----------TC- 3756
G.SE.93.SE6165             --G--------A-T-------------G-------G---------------------...-----C-----------A--------G--A-------------T------------C--------G-----------------------------C-----------TC- 3211

H.BE.93.VI991              -----------A-T---A---------G-T--------------------------A...-T-----------G-T-A--T--------A-----------G---------------C-------------A--------------------------C-------ATC- 3203
H.BE.93.VI997              -----------A---------------G-T--------------------------A...-T-------------T-A--T--------A-----------G---------------C-------------A--------------A-----------C-------ATC- 3138
H.CF.90.056                -----------A-----------G---T-T--------------------------A...-T-----------G---A--T--------A-----------G--C--------------------------A-----------A--------------C-------ATC- 3161

J.CD.97.J_97DC_KTB147      -------------T-------------T-ATGC-----------G------------...-T-----------G---A--------------G-----------------------C--C-----G---------------------------A-C-T----------C- 3008
J.SE.93.SE7887             ------G----A-TG------------G--TTG---GC------------------A...-------------G---A----------GA-----------G--------------C--C-----G-----------------------------C--------C--TC- 3128

K.CD.97.EQTB11C            -----------A-T---------G---G---TG-----------------------A...-------------G-T-A-----------A----------GG--------------A--------G-----------C-----A-----------C------------C- 3010
K.CM.96.MP535              -----------A---------------G--------G-------------------A...-------------G---A-----------A-----------G--------------A--------------------C-----A-----------C-----T------C- 3010
```

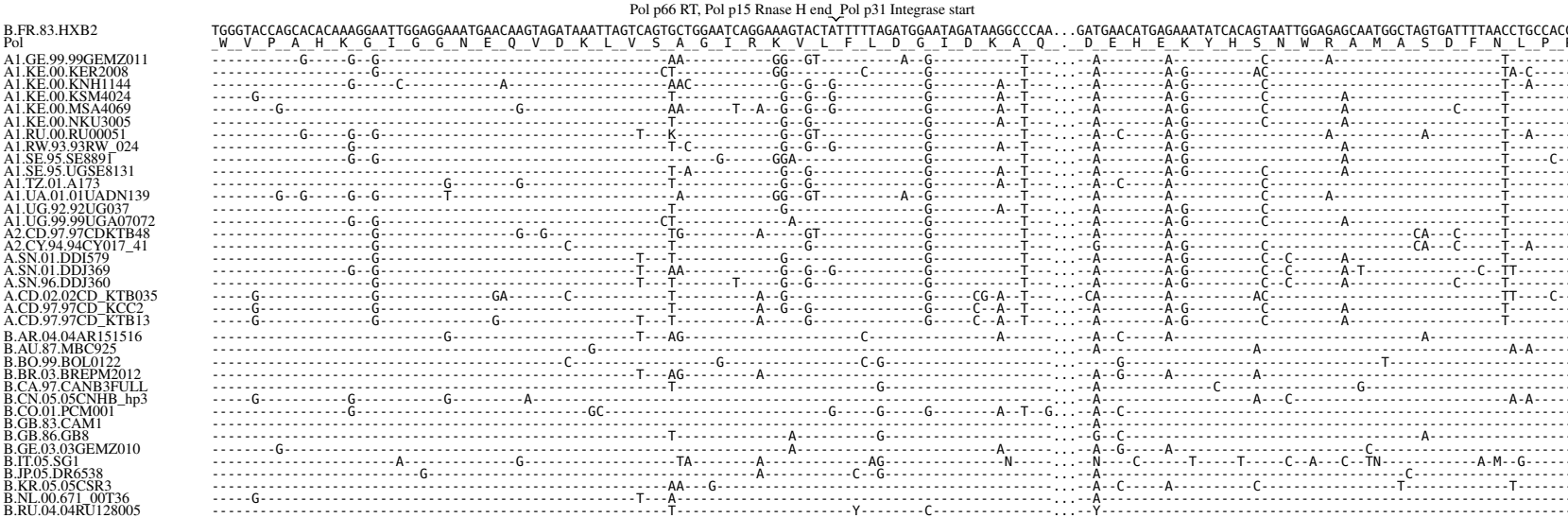
```
B.FR.83.HXB2                 CAATTAACAGAGGCAGTGCAAAAAATAACCACAGAAAGCATAGTAATATGGGGAAAG...ACTCCTAAATTTAAACTGCCCATACAAAAGGAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCCTGAGTGGGAGTTTGTTAATACCCCTCCCTT 3814
Pol                          _Q__L__T__E__A__V__Q__K__I__T__T__E__S__I__V__I__W__G__K__.__T__P__K__F__K__L__P__I__Q__K__E__T__W__E__T__W__W__T__E__Y__W__Q__A__T__W__I__P__E__W__E__F__V__N__T__P__P__L
01_AE.CF.90.90CF11697        ------G----A-TG--------GG-GG-----------------------------...--C----------G---A----------GA---------------------G-G-----------G--T--------------A-----------C--------C--TC- 3763
01_AE.CN.05.FJ051            ------G----A-TG--------G---G-----------------------------...--C--------------A-----------A----------------------TG--------------T--------------A-----------------------T-- 3824
01_AE.CN.06.FJ054            -----------A-TG------------G-----------------------------...--C----------G---A----------GA-----------------------G-----------G--T--------------A-----------------------TC- 3839
01_AE.HK.x.HK001             ------G----A-TG------------GT----------------------------...G-C--------------A----------GA----------------------TG-----------G--T--------------A-----------C-----------TC- 3176
01_AE.TH.01.01TH_R2184       -----------A-TG------------G-----------------------------...-TC----------G---A----------GA----------------------TG-----------G--T--------------A-----------------------TC- 3013
01_AE.TH.02.OUR769I          -----------A-TG------------G-----------------C-----------...--C----------G---A-----------A----------------------TG--A--------G--T--------------A-----------C-----------TC- 3013
01_AE.TH.90.CM240            -----------A-TG---------G--G-----------------------------...--C----------G---A----------GA----------------------TG-----------G--T--------------A-----------------------TC- 3388
01_AE.US.00.00US_MSC1164     -----------A-TG------------G--------G--------------------...--C----------G---A-----------A----------------------TG-----------G--G--------------A-----------------------TC- 3019
02_AG.CM.02.02CM_4082STN     ------G----A-----------GG-GG-T---------------------------...--C----------G---A--T-------GA------------G----------G-----------G--T--------------A-----------C-----------TC- 3013
02_AG.EC.x.ECU41             ------G------T---A------G-GG-T-T-------------------------...--C----------G---A--T-------GA------------G----------G--------------G--------------A-----------C-----------TC- 2910
02_AG.GH.03.GHNJ196          -------------T---A------G-GG-T-----G------A--------------...--C----------G---A-----------A------------G---------TG--T-----------T--------------A-----------C-----------TC- 3843
02_AG.NG.01.PL0710           -------------T---------GG-G--TGT-------------------------...--C----------G---A-----------A---------------------GTG-----------G--T--------------A-----A-----C-----------TC- 2998
02_AG.NG.x.IBNG              -------------T----------G--G-T-TG------------------------...--C----------G---A----------GA--G-------------------TG-----------G--T--------------A-----------C-----------TC- 3339
02_AG.SE.94.SE7812           -------------T---------GG-GG-T---------------------------...--C----------G---A----------GA------------G---------TG-----------G--T--------------A-----------C-----------TC- 3186
02_AG.SN.98.MP1211           -------------T---------GG-GG-T---------------G-----------...--C--C-------G---A----------GA------------G---------TG-----------G--T--------------A-----------C------------C- 3011
03_AB.RU.97.KAL153_2         --G------------------------G----T------------------------...-----------------A-----------A-----------------------------------------------------------A-----C-------------- 3040
04_cpx.CY.94.CY032           -----------A-----------G---G---TG---T--------------------...--------G----G-T-A-----------------------C--------------A--------G-----------C-----A-----A-----C-----------TC- 3180
05_DF.BE.x.VI1310            ------G--------------------G-A--------------G------------...-------------G---A-------------------------T---------------------------------------------A-----------------T-- 3196
06_cpx.AU.96.BFP90           -----------A---------------G--CT---------------------T---...-------------G---A-----------A-----------G--------------A--------------------C-----A-----A-----C------------C- 3842
08_BC.CN.97.97CNGX_6F        --G-----------T-----G------G---TG------------------------...-T-----------G-T-A--A--C-----A-----------G--------------C--------------------------------A-------------------- 2994
09_cpx.GH.96.96GH2911        --G--------A-TG--------GG-G------------------------------...--C------------T-A-----------A---------------------G-------------G--T--------------A-----A-----C-----------TC- 3016
10_CD.TZ.96.96TZ_BF061       ---C----C------------------G--CA----T-T-----G------------...--------G--------A-------------------------T---------G--C--------------------------T------------------------C- 3193
11_cpx.GR.x.GR17             ------G-G--A-TG--A---------G--TTG---G-------------------A...-T-----------G-T-A-----------------------G--------------C--------------------A-----A-----------C--------C--TC- 3116
12_BF.AR.99.ARMA159          -----------A---------------G-------G---------------------...--------G--------A-------T---A--G--------T-----------------------------------------A-----A-----C--------C--TC- 3819
13_cpx.CM.96.1849            ---C-------A-T---------GG-GG----------T------------------...CT---------------C-------G---A------------GT------------C--------G-----------------T-----A-----C-----------TC- 3219
14_BG.DE.01.9196_01          -----------A-T-------------G-------G--------------------A...-----------------A--T----G---A------------GT---------------------G-----------------C-----A-----C-----------TC- 3331
15_01B.TH.99.99TH_MU2079     -----------A-TG------------G-----------------------------...--C----------G---A-----------A----------------------TG-----------G--T--------------A--------------------A--T-- 3213
16_A2D.KR.97.97KR004         -----------A-----A---------GT--TG-----------------------A...-T-----------G-T-A-----------A-----------------------G-----------G--T--------------------------C-----------TC- 3177
17_BF.PE.02.PE02_PCR0155     -----G---------------------------------------------------...--------G--------A-----------A--G--------T-----------------------------------------A--------------------C--TC- 3004
18_cpx.CU.99.CU76            --G---GT---A-TG---------G-GG-----------------------------...---------------T-A--------G--A-----------G--------------A----------------------------------------G------GTTTC- 3119
19_cpx.CU.99.CU7             ---------------------------T--TT------------G------------...-T-----------G---A------------------------G---------GG-----------------T-----------------------C-----------T-- 3040
20_BG.CU.99.Cu103            --G---G--------------------G--CA-------------------------...---------------T-A--T-----G--A----------------------TG--------------T--------------A-----------C-------------- 3280
21_A2D.KE.91.KNH1254         -----------A-----A---------G---TG---GC------------------A...-T-------------T-A-----------A---------A-G----------TG--------------T--------------T-----------C-----------T-- 3019
22_01A1.CM.01.01CM_0001BBY   -----------A-TG---------G-GG--------G--------------------...-------------G---A------------------------G---------TG-----------G--T--------------A-----------C-----------TC- 3022
23_BG.CU.03.CB118            --GC---------T-------------G-ACA-------------------------...-------------G-T-A-----------A------------G---------TG-----------------------------A-----------C-------------- 3257
24_BG.CU.03.CB378            --G------------------------G-ACA-------------------------...---------------T-A-----------A------------G---------TG-----------G--T--------------A-----------C-------------- 3239
24_BG.CU.03.CB471            --G------------------------G-ACA-------------------------...---------------T-A-----------A------------G---------TG-----------G--T--------------A-----------C-------------- 3254
25_cpx.CM.01.101BA           -----------A---------------TGT-----G---------------------...--C----------G---A-----------A--------------------------C--------------------C-----T-----A-----C--C--------TC- 3022
27_cpx.CD.97.97CDKTB49       -----------A-TG--A---------G--T-G---T-------G-----------A...-------------G---A-----------------------G--------------A--------G-----------------A-----------C-----------TC- 3022
28_BF.BR.99.BREPM12609       ------------------------G--G-------------------------C---...-T-----------C-T----------G--A----------------------T------------------------------------------C-------------- 3180
29_BF.BR.02.BREPM119         ---------------------------G--------G--------------------...-T-----------G---A-----------A-----------T-----------------------------------------A-----------C--------C--TC- 2988
31_BC.BR.02.110PA            --G------------------------G--CTG---------A--------------...-------------G-T-A-----C--G-GA------------G-------------C--------------------------------------------T------C- 3248
32_06A1.EE.01.EE0369         -----------A---------------G---TG------------G-----------...-------------G---A-----------------G-----G--------------A--------------------C-----C-----A-----C------------C- 3449
33_01B.MY.05.05MYKL007_1     -----------A-TG--A---------G-------G---------------------...--C----------G---A--T-------GA------C---------------TG-----------G--T--------------A--------------------G--TC- 3193
34_01B.TH.99.OUR2478P        -----------A-T-------------G-----------------------------...--C----------G---A-----------A----------------------TG-----------G--T--------------A-----------C-----------T-- 3016
35_AD.AF.05.05AF095          -------------TG--------GG-G---CTG--G---------------------...-T-----------G-T-A-----------A------------G---------TG--T--------G--T--------------A-----A-----C-----------TC- 3007
36_cpx.CM.00.00CMNYU830      -----------A--G-----------TG-----------------------------...-T---------------A-----------A-------------T--------TG-----------G-----------------A-----------C-----------TC- 3019
37_cpx.CM.00.00CMNYU926      ------G----A--------G---G--G----------T---A--------------...--C----------G---A----------GA------------G---------GG--T--------G--T--------------A-----------C-----------TC- 3004
38_BF1.UY.05.UY05_4752       ---------------------------------------------------------...-T------G--------A------TT------G-----------------------------------------------C--A-----A-----C--------C--TC- 3225
39_BF.BR.03.03BRRJ103        -----G-----A-----------G---G-TTT------T------------------...-------------G---A-----------A--G--------T-----------------------------------------A-----------C--------C--TC- 3303
40_BF.BR.04.04BRRJ115        -----------------------G---G-----------------------------...-T------G----G---------------A--G--------T---------T----------------------------C--A-----------C-----T--C--TC- 3366
42_BF.LU.03.luBF_05_03       ---------------------------G--------G-------C--------G-G-...-T------R------T-A-----------------G----------------T------------------------------------------C-------------- 3352
43_02G.SA.03.J11223          G--------C---TG--------GG-GG-T--G-----------------------A...--C----------G---A-----------A-----------------------------------G--T--------------A-----------C-----------TC- 3339
U.CD.83.83CD003              --GC-------A---------------T-T-TG------------------------...-T---------C-----A-----------------------GT-------------T--------------------C-----------------C-----------TC- 3309
U.CD.90.90CD121E12           --GC-------A---------------T-T-TG------------------------...--------------G--A-----------------------G---------G----C--------------------C-----------------C-----------TC- 3013
U.GR.99.GR303                -----------A-T-------------------------------------------...--C----------G---A--A----------G---------G--------------C--------G---C-------------A---------...--------C--TC- 3085
U.NL.95.U_NL_95_H10986_D1    -----------A---------------G---TG-----------------------A...--------G----G---A--T--------A-----------G--------------A--------------------C-----A-----------C------------C- 3699
CPZ.CD.90.ANT                ------G-T-GA-T-------------GGATT------T-----T--C------G--...GTA--A------C--T-A------ACT-G------T-----TG-C------T----C--------G-----------A--A--A-----------------A--C--A-- 3253
CPZ.CM.01.SIVcpzCAM13        -----GG----A-TCA-A--------TG-TGTG-----T-----C--T--------A...--A----G-----G-T-A--AG-------A----GT-----G---------G-------C-----G-----------C--C-----------CA----C-----C--A-- 3394
CPZ.CM.05.SIVcpzEK505        --------T--A-TG--A-----GG-GG-ACT------T-----T--C--------A...-----------C-G-T-A--AG-G-----A----GT-----G--C--------------------G--A--T-----------T-----------C--C--------AC- 3386
CPZ.CM.05.SIVcpzLB7          --G---G------T-A-------G---G---T------------C-----------A...-TA--A--G----GT--------T-----A---G--------G-------------T--------------T-----C-----A-----A-A-------------T--CA 3369
CPZ.CM.05.SIVcpzMB66         --G-----T----T---A------G--G--CT------------C--C--------A...GTA----G---CGG---A--T--------------------GG-G-----------C-----------A--A-----C-----A-------A------C----------- 3360
CPZ.CM.05.SIVcpzMT145        --GC-TG-T--AA----A---------G-T-----------T--------------A...--A--A-------G---C--TG-T-----A----TG------G-C------G----T--C-----G--T--------------------A--CA-A--C-----A---C- 3356
CPZ.GA.88.GAB1               ---C-GG-G--A-----A--------TG-T--T--------------T--------A...--A----------GGT-A--AG----G--A----GT------G--------G----A--C-----G--A--------C-----A-----A---A-C--------A----- 3875
CPZ.TZ.00.TAN1               ------G---GACTGA-A--G------GG-AT---G------A----T-----T-TT...GTG---------TT-T-A--TG--TCC--A--G------AGCCAG--------T--T--C-------TT------G-A-----------A---A----C-----A--AC- 3458
CPZ.US.85.CPZUS              --GC--GTT--A-TG------------G-T--T--G-----T--C--T---------...GTA--A---------T---TAG-G--G-------GTC--------------T----A--------G--------------A--T-----A-----C------------C- 3874
CPZ.CM.x.SIVcpzDP943         ------G------T---A------G--G-A--T-----T--------C--------A...--A--A-----C-G-T-A--TG-T--------GGTC--------C------T----A--C-----------------C--A--C-----------A------------C- 3400
CPZ.CM.x.SIVcpzMB897         --G--G--T--A-T---A--G---G--G--CTG-----T-----C--C--------A...--A--C--G----G------------G--A-----------GG-T-----------T--C--------A--T-----C-----A-----A-A------C---------C- 3380
N.CM.02.DJO0131              --G---GTT--A-TG--A------G--G-A------G-T--------T---------...-T------G------T-A--AG-----------GTG-----GG----------T--CA----------A--------------------A--C--C--C--T------C- 3320
N.CM.04.04CM_1015_04         --G-----T--A-T---A-----GG--G-A--------T--------T---------...--------G------T-A--MG-----------GTG-----GG----------T---A----------A--T--------A--------A--C-----C--T------C- 3319
N.CM.04.04CM_1131_03         --G-----T--A-T---A-----G---G-A--------T--------T---------...--------G------T-A--AG-----------GTG-----GG----------T---A----------A--T-----------------A--C--C--C---------CC 3318
N.CM.95.YBF30                --G---GTT--A-TG--AAGG---G-GG-A--------T--------T---------...-------------G-T-A--AG-----------GTG-----GG----------C--TC----------A--T-----------------A-----C--C--T------C- 3408
N.CM.97.YBF106               G-G---G-T-CA-TG--A-----GG-GG-A--------T--------T---------...--------G------T-A--AG-----------GTG-----G-----------T---C----------A--T-----------------A--C--C--C--T------C- 3406
O.BE.87.ANT70                ------G----A-T-A-C--G--GG-GT-TCA----TCT---A-T--C--------A...TTG-----------G-----AG-CACT-GA-----T---------------G-G--C-----------------------A--A-----A-----C-GC--A--C--A-- 3868
O.CM.91.MVP5180              -----GG----A-T---C--G--GG-GT-TCA----GCT-----T--------G--A...TTA--------C-GG-----AG-TACT-GA-----T--------T------G----A--------G-----------------A-----A-----C-GC--A--C--A-- 3843
O.CM.96.96CMABB637           ---C--G----A-T---C--G---G--G-TCA----TCT-----T--C--------A...TTG-----------G-----AG-CACT-GA-----T---------------G----C--------------------------A-----A-----C-GC--G--C--A-- 3298
O.CM.98.98CMA104             ------G----A-T-C-A--G---G-GT--CAG---TCT-----T--C--------A...TTG-----G-----G-----AG-CACT-G------T---------------G-G--C--------G--------------A--A-----R-----C-GC--A--C--A-- 3300
O.CM.99.99CMU4122            AT----G----A-T-A-C--G--GG-GT-TCA----GCT-----T--T-----G--A...TTG-----------G-----A--CACT-G---G--T---------------G-G--C-----------------------A--A-----A-----C-GC--A--C--A-- 3299
O.FR.92.VAU                  ------G--A-T-T-A-C--G--GG-GT-TCA----GCT---A-C-----------A...TTG--C--------G-----AG-CACT-GA--GT-----------------G----T--------------------------A-----A-----C-GC--A-----A-- 3380
O.SN.99.SEMP1299             ------G----A-T-C-C--G-GGG-GT-TCA---GGCT---A-T--C--------A...TTG-----------G-----A--CACT-GA-----T---------------G-G--C-----------------------A--A-----A-----C-G---A--C--A-- 3867
O.US.99.99USTWLA             ---C--G----A-T---C--R--GGC-T-TCA----GCT-----T--T-R------A...TTA-----------G-----AG-CACT-GA-----T---------------G----T--------------------------A-----A-----C-GC--A--C--A-- 3297
```

```
                     p51 end and p66 RT continue Pol p15 RNAse H start
B.FR.83.HXB2 AGTGAAATTATGGTACCAGTTAGAGAAA...GAACCCATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCCTAACTGACACAACAAATCAGAAGACTG 3981
Pol __V__K__L__W__Y__Q__L__E__K__.__E__P__I__V__G__A__E__T__F__Y__V__D__G__A__A__N__R__E__T__K__L__G__K__A__G__Y__V__T__N__R__G__R__Q__K__V__V__T__L__T__D__T__T__N__Q__K__T__
A1.GE.99.99GEMZ011 ---A---C-------------G------...--------------------G-------------------------------------GA-------------G---------G-C-------------G-----TC-T--------G-----------A------- 3165
A1.KE.00.KER2008 ---A-----------------G------...--C-------C---------G--T--------------------CC-T-----------GC-------G-----------C---G-C--G-----------G-----TT----G-----G-----------A--A---- 3180
A1.KE.00.KNH1144 ---A-----G--------A-----------...--C-----------------G--T--------------A-----C--T-----------GC-------------G-----C---G-C--G-----------G-----TT----G-----G-----------A--A---- 3186
A1.KE.00.KSM4024 ---A--------------A-----A---...--C--A----C---------G--T---------------------T-----A-----GA-------------G---A-C--AG-C--G-----------G--A--TT----------G-----------A------- 3177
A1.KE.00.MSA4069 ---A-----G--------AC--------...--C-----------------G--T--------------------C--T-------A---GC-------------G-----C---G-C--G-----------G-----TT----------G-----------A------- 3180
A1.KE.00.NKU3005 ---A------------------------...--C-------C---------G--T--------------------C--T-----A-----G--------------G-----C---G-C--G-----------G---A-TT--Y-------G------------------- 3180
A1.RU.00.RU00051 ---A--GC--------------------...--------------------G--T-----------------------------------GC-------------G---------G-C-AG-----------G------T-T--------G-----------A------- 3282
A1.RW.93.93RW_024 ---A---C-------------------G...--C-------A---------G--T-A--------------------C--T-----------GC-------------G-----C---G-C-------------G---A-TT-----G----G-----------A------- 3183
A1.SE.95.SE8891 ---A------------------------...--C-------------------T-------------------------T-----A-----GC-----------------------C---G-C-------------G-----T------G----G-----------A------- 3170
A1.SE.95.UGSE8131 ---A------------------------...--C-------C-----T---G--T--------------------C--T-----------GC-------------G-----C---G-----------------G-----TT----------G-----------A------- 3353
A1.TZ.01.A173 ---A------------------------...--C------A----------G--T--------------------C--T-----------GC------G------G-----C---G-----------------G-----TT----------G------------------- 3180
A1.UA.01.01UADN139 ---A---C-------------G------...--------------------G--T-------------------------A--------GA-------------G---------G-C---------A----G--G--TC-T--------G------------------- 3180
A1.UG.92.92UG037 ---A------------------------...--C-------C---------G--T--------------------C--T-----------GC-------------G-----C---G-C--------G--------G-----TT----------G-----------A------- 3997
A1.UG.99.99UGA07072 ---A--------------------A---...--C------TC-----T---G--T---------------------T-----A-----G---------------------C---G-C-------------G-----TT-----GA---G-----------A---G--- 3180
A2.CD.97.97CDKTB48 ---A---C----------------A-C-...----------C------------T---------------------T-----A---------C-------G-----G-----C---G-----------------A-----C--T-G--A--G-----------A------- 3319
A2.CY.94.94CY017_41 ---A---C-G--------------A---...----------C------------T---------------------T-----A---------C-------G-----G-----C---G-C-------------A-----T----G--G--G-----------A------- 3338
A.SN.01.DDI579 ---A-----------------------G...--C------A----------G--T---------------------T-----------G--C----G---------G-----C---G-C-------------G-----TT--T---TG--G-----------A------- 3179
A.SN.01.DDJ369 ---A-----------------------G...--C-----------------G-----------------A------T-----------G--C----G---------G-----C---G-C-------------G-----TT--T-------G-----------A------- 3182
A.SN.96.DDJ360 ---A--------------------A--G...--C-----------------G--T---------------------T-----------G--C----G---------G-----A---G-C-------------G-----TT--T-------G-----------A------- 3179
A.CD.02.02CD_KTB035 ---A--------------------A---...--C-------------------T---------------A------T---------------C----G---------G-----C---G-C-------------G-----TT----------G-----------A------- 3975
A.CD.97.97CD_KCC2 ---A--------------------A---...--C-------------------T-----------------------T-----------GC-------------G-----C---G-C-----------G---G-----TT----------G-----C-----A--A---- 3977
A.CD.97.97CD_KTB13 ---A--------------------A---...--C-------C-----T------T--T------------------T-----A------C-------------G-----C---G-C-------------G-----TT----------G--C-----------A------- 3185
B.AR.04.04AR151516 ---A------------------------...----------C------------T---------------------T-----A-------------G-------------G-C-------------G------T-------G---------C-------------- 3203
B.AU.87.MBC925 ---A------------------------...---------A-------------T------------------------------G---------------------------C-A------------G------T------------------------------- 3992
B.BO.99.BOL0122 ------------------A---------...--C-------A-----T------T---------------------T--------C-----------------------------------------------------T------------------------------ 3201
B.BR.03.BREPM2012 ------G-----------------A---...-----------------------T---------------------T--A--A--------------------------G-C-------------G-----TC-------A------------------------ 3406
B.CA.97.CANB3FULL ---A------------------------...-----------------------T---------------------T----------------------------------G----G------------------T--A--C------------------------- 3270
B.CN.05.05CNHB_hp3 ---------------T-----------G...----------A------------T---------------------T------------------------------------C-A----------G--------------A-------------------------- 3980
B.CO.01.PCM001 ---A--------------------A---...-------------------------------------------T-----T------C-----------------------G-C-----------------G---T----------T-------------------- 3186
B.GB.83.CAM1 ------------------------------...-----------------------T---------------------T-----A----G---------------------G---G-C----------------------TC-------G------------------------ 3982
B.GB.86.GB8 ---------------T------------...---------A-------------T--------------A------T-------------G--------------------G-C------------------------C------------------------------ 4001
B.GE.03.03GEMZ010 ---A----------------------C-...-----------------------T--T------------------T-----A--C---------C------------------G-C-------------G------T-------A------------------------ 3186
B.IT.05.SG1 ----------------------------...----------A------------T---------------------------------G--------T---------------------CC----------G---G------T-----C------------T----------- 3783
B.JP.05.DR6538 ------GC----------AC----A---...-----------------------T--------------------GT------------------------------G-C-------------G------CA----G------------------------------ 3985
B.KR.05.05CSR3 ---A------------------------...-----------------------T--T---------------A------G---A--A---------GC---G------------A-C---G-C-------------G------------------------------------ 4006
B.NL.00.671_00T36 ------------------------A--G...-----------------------T---------------------T-----A-----G--------------------G-C-----G-------------A--C--A---AG---------------------------- 3546
B.RU.04.04RU128005 ---------------T--------A---...-----A----A------------T----------------------------A----------------------G-----------C------------------------T------------------G----------A---- 3472
B.TH.00.00TH_C3198 ---A------------------------...----------A------------T---------G------------T--A--A------------------------A-----C-A------------------------------------------------------------ 3195
B.UA.01.01UAKV167 ------------------------A-C-...-----------G-----------T--T------------------T-----A-----------C------------------G---------------------G-----TT-------G------------------------ 3207
B.US.04.ES10_53 ---------------T------------...-----------------------T---------------------------A------------------------------------C-A------------G---------T-------------------------------- 3990
B.US.99.PRB959_03 -------------------C----A---...---------T-------------T---------------------T------------------------------C---G-C-------------G------T-----C------------------------ 3201
C.AR.01.ARG4006 ---A--G------------C-G------...---------AC------------T--------------A------T-----A------A-G--G--------G---------G-C--------G--G----A----TG-T------------------------------- 3165
C.BR.04.04BR013 ---A--G------------C-G--A---...---------CC------------T--------------A---------------A------C-G-----------G-----------C--------G--G----A----TT-T--------A------------------- 3445
C.BW.00.00BW07621 ---A---------------C-G--A---...--C---T-G-------T------T--------------A------T-----T------A-------------------------G-C-A------G-GG-----CA-T--T--------A--------------A---- 3334
C.CN.98.YNRL9840 ---A---------------C--------...--T-------C-----T------T--------------A------T-----A------A-------------G---------G--------------GA-G---A----TT-T--------A------------------- 3165
C.ET.02.02ET_288 ---A-----G---------C-------G...----------C------------T--------------A------T-----A------A-------------G---------G-C-------G-----G---A----TT-T--------A------------------- 3183
C.GE.03.03GEMZ033 ---A---------------C-G-----G...----------C-----T------T--------------A-----GT-----A--A------G-------------------G-C--------G--G----A--A-----TA-------A---------------G--- 3162
C.IL.99.99ET7 ---A--------------AC-G------...----------------T------T--------------A------T-----A------A-------------G---------G-C--------G--G----A----TT-T--------A---------------G--- 3162
C.IN.99.01IN565_10 ---A---------------C-G------...----------------T------T--------------A----------------C------A-------G-----G---------G-C------C-G--G----A----TT-T--------G------------------- 3362
C.KE.00.KER2010 ---A---C-----------C-G------...--T-------C------------T--------------A------T-----A------G----G---------G---------G-C--------G--G----------TT-TA---G---A------------------- 3162
C.MM.99.mIDU101_3 ---A---------------C-G------...--T-------C-----T------T--------------A------T-----A------A-------------G---------G---------------A-G----------TT-T--------A-----G------------- 3332
C.MW.93.93MW_965 ---A--------------AC-G--A---...----------C------------T--T-----------A--T----T-----A------A-------------G---------G-C-A------G--G----------TT-T------A-A---------------G--- 3168
C.SN.90.90SE_364 ---A---------------C-G------...--C------A-------------T--------------A------T-----A------A-------------G---------G-C--------G--G----------TT-TA-------A------------------- 3150
C.SO.89.89SM_145 ---A---------------C-G------...----------C------------T--------------A------T-----A------A-G--G--------G---A-----G-C--G--------G---A----TT-T--------A------------------- 3202
C.TZ.02.CO178 ---A---------------C-G------...----------C------------T--------------A------T-----A------------------G-----G---G---A---------G---A--A-TT-T--------A------------------- 3162
C.UY.01.TRA3011 ---A---------------C--------...----------C------------T-----------C--A------T-----A------A-G-----------G---------G------------G--G----A----TT-T--------A------------------- 3153
C.YE.02.02YE511 ---A--------------AC-G------...--------G-------T------T--------------A------T-----A------A-G-------------------A---G-C--------G--G----A----TT-T----A---A---------------G--- 3156
C.ZA.04.04ZASK164B1 ---A---------------C-G--A--G...--T-----GA-C----T------T--------------A------T-----A-----------------G---------G-C-------------------A-TT-T--------A------------------- 3395
C.ZA.05.05ZAPSK240B1 ---A-----------------G------...----------C---G--------T--------------A-----CT-A---C------A-------------G---------G-C--------G--G----A--A-TT-T--------A--------------A---- 3385
C.ZM.02.02ZM108 ---A-----------T---C-G--A---...---------TC------------T--------------A------T-----A------A----G--------G---------G-C-----------G----------TT--A----G--A------------------- 3950
D.CD.83.ELI ---A-----------------------G...---------A-------------T---------------------T--A------------------------------G-C-----------G---------C-TT-G--------G-------------------- 3527
D.CM.01.01CM_0009BBY ---------G-----T-----------G...---------A-------------T---------------------T--A------------------------------G-C-A---------G---------C-GT-T--------T--T------------------- 3180
D.KE.01.NKU3006 ---C--G-----------------A--G...---------A-------------T------------------T----T--A--A---------------------------G-C-------------------------TA-------T------------------- 3186
D.KR.04.04KBH8 ---A----------------------GG...---------A-------------T---------------------T--A-------G--------G------------------G-C------------------A--C-TT--------------------------- 3940
D.TD.99.MN011 ---A-----------T--------A--G...---------A-G-----------T---------------------T--A------------------------------G-C------------------------C-GT-T--------C-------------------- 3205
D.TZ.01.A280 ---T-----------------------G...---------A-------------T-----------------G-----T--A-------G--------------------G-C-A---------G---------T-T------------------------------ 3182
D.UG.99.99UGK09259 ---T-----------------------G...--------------------G-----T---------------------T--A------------------------------G-C-----------G-----G---T-T------------------------------ 3186
D.YE.01.01YE386 ---T-----------T-----------G...-----A----C------------T--------------------GT--A--A------A-----------------------G-C-----------G---------T-T----------------------------A---- 3180
D.YE.02.02YE516 ---A-----------T------------...---------A-------------T---------------------T--A------------------------------G-C-----------G---------C-AT-T--------T-------------------- 3192
D.ZA.90.R1 ---A--------------A--------G...-----------------------T---------------------T--A---------G--------------------G-C-----------G---------C-TT-T------------------------------ 3327
F1.AR.02.ARE933 ---A---C-------T--------A-CG...--G-------C-------------------------------T----T--A--A--C---AA------------------------G-C--------------G-C----C----------G--C------------G--- 3275
F1.BE.93.VI850 -------C----------------A-C-...--G-------C---------C---------------------T----T--A-----C---AA------------------------G---A----A--------------T-------G--G--C------------G--- 3321
F1.BR.01.01BR125 ---A---C-------T--------A-C-...--G-------C-------------------------------T-------A---T-C---AA------------------------G-C-A-----------------C----T----------G--C------------G--- 3426
F1.ES.x.P1146 ---A---C-------T--------A-C-...--G------A-G------------------------------T----T--A-----C---CA------------------------G-C-------------G--G---T--------AGA--C------------G--- 3258
F1.FI.93.FIN9363 ---A---C-------T--A-----A-C-...--G-------C------------A----------G-------T----T--A-----C---AA------------------------G-----------------G------T----------G--C------------G--- 3310
F2.CM.02.02CM_0016BBY ---A--------------A-----A-C-...--G------A-------------T--------------------GT--A---------A----------------------------G-C-------------G------T----------G------------------- 3171
F2.CM.95.MP257 ---A--------------------A-C-...-----T---A-------------T---------------------T--A---------C-----------------A-----G-C-------------G------T----------G------------------- 3190
F2.CM.97.CM53657 ---A-----------T--A-----A-C-...--G------T-------------T---------------------T--A-----C---A-------------------------G-C----------------------C----------G------------------- 3171
G.BE.96.DRCBL ---C-----------T--------A-C-...---------CC-----T------T-A-------------------T----------A-----------G-----------G---G-C-A-----------TTA--A-T-------------A--------C--A---G--- 3938
G.CM.01.01CM_4049HAN ---A-----G-----T--------A-C-...---------CC------------T-A-------------------T--A-----A-----------G---------------G-C-A----A---------A--A-TG----------A--------C--A---G--- 3183
G.CU.99.Cu74 ---A-----------T-G--------C-...--G------C-G-----------T-AT----------------------------A--A------------------------------G-C-A----A---------G---A-T-------------T--------C--A--A---- 3585
G.ES.99.X138 ---A-----------T-G------A-C-...----------C------------T-AT------------------T-----A--A------------------------G-C-A----A-----------A--A-T-------------A--------C--A--AG--- 3422
G.GH.03.03GH175G ---A-----G-----T--------A-C-...---------CC------------T-A-------------------T--------A-----------G-----G---------G-C-A----A-----------A--A-T---T-------A--T-----C--A---G--- 4021
G.KE.93.HH8793_12_1 ---A-----G-----T-G------A-C-...---------CC------------T-AT------------------T-----A--A-----------G-------------------G---A----A-----------A--A-T-------------A--------C--A------- 3381
G.NG.01.01NGPL0669 ---------------T-G------A-C-...---------CC------------T---------------------T-----A--A-----------G-------------------G-C-A----A---------------CA-----A--CAG--A--------C--A------- 3186
G.PT.x.PT2695 ---A-----------T-G------A-C-...---------A-------------T-AT------------------T-----A--A--G--------G-------------------G-C-A----A----------------A-T-------------A--------C--A---G--- 3923
G.SE.93.SE6165 ---A--G--------T-G------A-C-...---------CC------------T-A-------------------T--------A-----------G-------------------G-C-A----AG------------A--A-T-------------A--------C--A---G--- 3378
H.BE.93.VI991 ---------------T----------C-...--G-------A------------T-A-------------------T-----A------A-G--G--------------C---G-C-------A---------A----TT-------G--A--------------A---- 3370
H.BE.93.VI997 ---A-----------T--------AGC-...--G-------C------------T-A-------------------------A------A--------------------C---G-C-A-----A--------G-----TG-------A--A--T------------------ 3305
H.CF.90.056 ---A-----------T--------A-C-...--G-------C------------T-A-----A--------------T-----A-------------------------C---G---------AG--------------T-------G--A------------------- 3328
J.CD.97.J_97DC_KTB147 ---A-----------T--A-----A--G...---------A-G-----------T--------------A--------------------A-------------G-----C---G-C-A---------G-----G--T-------A---------------------------- 3175
J.SE.93.SE7887 ---A--G-----------A-----A--G...---------A-G-----------T-----------------T-----------------AC------G-----G-----C---G-C-A---------G-----A--T------------------------------------ 3295
K.CD.97.EQTB11C ---A---C----------------A-C-...-----------------------T--------------------C--T-----A------CAG--------------------------G-C-A------------------A--T-AA-------A--------------A---- 3177
K.CM.96.MP535 ---A---C----------------A-C-...--------------G--------T-----C---------------C-T-----A------AAG----G----------------G-C-------------G------T--A-------A--------------AG--- 3177
```

```
                                                       p51 end and p66 RT continue  Pol p15 RNAse H start
B.FR.83.HXB2               AGTGAAATTATGGTACCAGTTAGAGAAA...GAACCCATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCCTAACTGACACAACAAATCAGAAGACTG 3981
Pol                        __V__K__L__W__Y__Q__L__E__K__.__E__P__I__V__G__A__E__T__F__Y__V__D__G__A__A__N__R__E__T__K__L__G__K__A__G__Y__V__T__N__R__G__R__Q__K__V__V__T__L__T__D__T__T__N__Q__K__T__
01_AE.CF.90.90CF11697      ---A-------------A-----A---...--C-------C---------G--TC-------------------GT------A---GC------------------C---G-C--------------G--A--TT----------G-----G-----A------- 3930
01_AE.CN.05.FJ051          ---A-------------A-----A--G...--C------T----------G--T-----------------C--C-GT-----A-----GC------------G-----C---G-C--G-----------G--A--TT--T-------G-----------A------- 3991
01_AE.CN.06.FJ054          ---A-------------A-----A---...--C-----------------G--T-------------------GT----------GC------------------C-----C--------------G--A--TT----------A-----------A------- 4006
01_AE.HK.x.HK001           ---A-------------A-----A---...--C--------G--------------------------------GT-----A-----GC------------G-----C---G-C--------------G--A--TT----------A-----------A------- 3343
01_AE.TH.01.01TH_R2184     ---A-------------AC----A---A..--C-----------------G--T--------------------T-----------GC------------------C-----C-A------------G--A--TT----------G-----C-----A------- 3181
01_AE.TH.02.OUR769I        ---A-------------A-----A---...--------------------G--T-----------------C----GT----------GC------------G-----C---G-C-A------------G--A--TT--T-------G-----------A------- 3180
01_AE.TH.90.CM240          ---A-------------A-----A---...--C-----------------G--T-------------------GT----------GC------------G-----C---G-C--------------G--A--TT----------G-----------A------- 3555
01_AE.US.00.00US_MSC1164   ---A-------------A-----A---...--C-----------------G--T--------------------T----------GC------------G-----C-----C--------------G--A--TT----------G-----------A------- 3186
02_AG.CM.02.02CM_4082STN   ---A-----------------------...--C-------C-------------T------------------------------GC---G--------G---A-C---G-C-A------------G-----TT----------A-----------A------- 3180
02_AG.EC.x.ECU41           ---A-----------------------...--C------------T-----T------------------C--T----------GC------------G-----C---G-C-----G-------GG-----TT----------G-----------A------- 3077
02_AG.GH.03.GHNJ196        ---A-------------A-----A---...--C------------------T------------------A--T-----------C------------------C---G-C--------------G--G--TT----------G-----------A------- 4010
02_AG.NG.01.PL0710         ---A--------------C----A---...--C------CC----AT-----T---------------------T----------G----G-------------C---G-C--------------G-----T-----------A-----G-----A--A---- 3165
02_AG.NG.x.IBNG            ---A-------------------A---...--C------------------T-----------A----------T--A-------GA------------G-----C---G-C--------------G-----TT----------G-----------A------- 3506
02_AG.SE.94.SE7812         ---A--------------C----A---...--C------------------T----------------------T----------GA------------G---------G-C--------------G-----TT----G-----G-----------A------- 3353
02_AG.SN.98.MP1211         ---A-----------------------...--C------A----------G--T------------------C--T-----A-----GC-----G-----G-----C---G-C--------------T-----TT----------G-----------A------- 3178
03_AB.RU.97.KAL153_2       ---A-----------------------...---------------------T-----------A----------T------------C---------------------G-C--------------G-----T------------------------------ 3207
04_cpx.CY.94.CY032         ---A--------------------A-C...--C-------C----------T-----------------------T--A--A--A---CAG-------------------G-------C---------------T-----T----A-------------------- 3347
05_DF.BE.x.VI1310          ---A----------------------G...---------T-----------T-----------------------T--A-------------------------A-----G---A--------------C----T----------G--G-----------AG--- 3363
06_cpx.AU.96.BFP90         -------C----------------A-C...--G------------------T-----------------G-----T--A--A-----GCA--------------------G-C--------------G-----T-----------A-----------------G- 4009
08_BC.CN.97.97CNGX_6F      ---A-------------AC-G------...--T-------C-----T-----T------------A---------T------------A---G--------G--------G-C--------GA-G---A----TT------------------------------- 3161
09_cpx.GH.96.96GH2911      ---A--------------------A---...--C------A----------G--T--------------------T----------GC---G-----G--G-----A-----CC-------------GA----AT----------A--C--------A------- 3183
10_CD.TZ.96.96TZ_BF061     ---A--G-----------C-G------...-----------------------------------------T--A-----A-------------------------------G-C-----------G-G--C-A--T-TA-------T------------------ 3360
11_cpx.GR.x.GR17           ---A--------------------A--G...--G------A----G-----------------------------T-----A------C-------------G-----C-----C-A------G---------GA-T----C--T---A-----------A------- 3283
12_BF.AR.99.ARMA159        ---A---C-------T--------A-C...--G-------C--A-----------------------T----T-AA----C---AA--------------------------G-C-A--------------------T----------A--C-----------G--- 3986
13_cpx.CM.96.1849          ---A---C-G-----T-G------A-C...---------C-------------T-A---------------------T----A--A----------G----------------G-C-------A-------A--A-T----------A--------C--A------- 3386
14_BG.DE.01.9196_01        ---A---Y-------T-GA-----A-C...---------C-------------T-A---------------------T----A--A----------G----------------G-C-A------------A--A-T----------A--------C--A--AG--- 3498
15_01B.TH.99.99TH_MU2079   ---A-------------A-----A---...--C-----------------G-------------------------GT----------GC------------G-----C---G-C-----------G--G--A--TT----------G-----------A------- 3380
16_A2D.KR.97.97KR004       ---A--GC-------T--------A-C...--G------AC--------------T------------------C--T--A--------C------G-----G-----C---G-C-A------------A-----T----------T------------------ 3344
17_BF.PE.02.PE02_PCR0155   ---A---C-------T--------A-C...--G-------C--A------------T----------------C--T--A----C---AA---G-----------Y-----G-C-----------------C----T----------G--C-----------G--- 3171
18_cpx.CU.99.CU76          -------C-------T--------A-C...--G-------C-------------T----------------------T-----A------A---G--------G-----C---G-C-A------------G--A--T--A-------A-------------A---- 3286
19_cpx.CU.99.CU7           ---A--------------------A--G...---------A-G-----------T----------------------T--A------------------------------------G-----------G--------T-T---------------------------- 3207
20_BG.CU.99.Cu103          --------------------------G...------------------------T----------------------T--A-------------------------------G-C-------A------G-----T------------------------------ 3447
21_A2D.KE.91.KNH1254       ---A---C----------------A--G...---------C------------T----------------------T--A--------C------G-----G-----C--GG-C--------------A-----T---G--G--G-----------A------- 3186
22_01A1.CM.01.01CM_0001BBY ---A--G----------A-----A---...--C------A----------G--T----C--------------GT----------GC---G-------G-----C---G-C--------G-----G--A---T----------A----T------A------- 3189
23_BG.CU.03.CB118          ---A-----------------------...------------------------T----------------------T--A-------------------------------G-C-------A------G-----T------------------------------ 3424
24_BG.CU.03.CB378          ---A-----------------------...------------------------T----------------------T--A-------------------------------G-C-------A------G-----C------------------------------ 3406
24_BG.CU.03.CB471          ---------------------------...----------A------------T----------------------T--A-------------------------------G-C-------A------GA----T------------------------------ 3421
25_cpx.CM.01.101BA         ---A-----------------------...--C------A----------G--T--------------------T-----A-----GC------------G-----C---G-C--------------G-----GT----------G-----------A------- 3189
27_cpx.CD.97.97CDKTB49     ---A-----------T--------AC-G...--G------------------T--------------------T-----A-----AA------------G--------G-C-A----------G-----GA-T----------A-----------A------- 3189
28_BF.BR.99.BREPM12609     ------------------------A---...---------------------T--------------------T-----A-----CA--------------------G-C--------------------TC-TT----A-------------------- 3347
29_BF.BR.02.BREPM119       ---A--GC-------T--------A-C...--G-------C-------------------------T----T--A--A--C---AA--------------------G-C--------------G-C----C----------G--C-----------G--- 3155
31_BC.BR.02.110PA          ---A--------------C-G------...---------A----G-------T--------------------T----A--A----G---------G--------G-C--------G--G---AG---TT-T--------A-------------------- 3415
32_06A1.EE.01.EE0369       ---A---C----------------A-C...--G-------------T-----A----------------G-----T--A--A-----AA--------------------G-C--------G-----G-------------T---A-----------------A- 3616
33_01B.MY.05.05MYKL007_1   ---A-------------A-----A---...--C------A-----------T--------------------GT----------GC-----------G-----C---G-C-A------------G--A--TT-------A--G-----------A------- 3360
34_01B.TH.99.OUR2478P      ---A-------------A-----A---...--C-----------------G--T--------------------GT----------GCG---------------A-------C--------------G--A--GT----------G-----------A------- 3183
35_AD.AF.05.05AF095        ---A-------------A---------...--C-----------G-----G--T--------G----------C--T----------GC------------G-----C---G-C--G----------G-----TT----G-----G-----------A------- 3174
36_cpx.CM.00.00CMNYU830    -----------------A-----A--G...--C-----------------G--T--------------------T----------GC------------G-----C---G------------------------TT----------G-----------A------- 3186
37_cpx.CM.00.00CMNYU926    ---A--------------------A--G...------------------------T--T----------------G---T--------C--GCA-----------G--C--C--GG---A-----------GA----TT--T-C-----G-----------A------- 3171
38_BF1.UY.05.UY05_4752     ---A---C-------T--------A-C...--G-------C--A-------------------------T----T--A----C---AA--------------------G-C-----------------C----T--T------G--C-----------G--- 3392
39_BF.BR.03.03BRRJ103      ---A---C-------T--------A-C...--G------A-------------------------------C-T--A--A-----GC--------------------G-C---------------------T--T-C-TG----------------A---- 3470
40_BF.BR.04.04BRRJ115      ---A---C-------T--------A-C...--G-------C-----TG-------------------------T--A--A----------------------G---G-C-----------------C----T-------A-------------------- 3533
42_BF.LU.03.luBF_05_03     ---------------------------...--C------------------T----------------------T----A--G-------R-----G-----------G-C-----------------C----T----------G--C-----------G--- 3519
43_02G.SA.03.J11223        ---C-G---------T-----------...--C--------G----TG----T----------------------T----AT----GCA-----------------C---G-C--------------G-----TT----------G-----------A------- 3506
U.CD.83.83CD003            ---A--G-----------------A---...--C------A-G--------G--T-------------------------A---AAG--------------------G-C-------------------A-TT----------A-------------AG-A- 3476
U.CD.90.90CD121E12         ---A--G-----------------A---...--G-------C------------T------------------G-----A--A---AAG--------------------G-CCA------------AC-A-TT----------A-------------AG-G- 3180
U.GR.99.GR303              ---A--------------------T---...-G--------------------T-----------A--T----T--A--A--A---AGG--------------------G-C--G--------W--G-----TG----------A-------------------- 3252
U.NL.95.U_NL_95_H10986_D1  ---A--GC-------T-----G--A--G...--C------A-----------T------G-------------T----A-----A-G-----------G--------G-C--------------G-----T-----------A-------------------- 3866
CPZ.CD.90.ANT              -A-T-G-C-C-----TA-TC--TT-GC-...--C--T--TCC--AG--T-------T-----T-------------T--AA-CT--C----G--G--G-----C-----G--AG-C---------AGT-GG--AAAACA----CAAA-G--C--C-----AC-AG-A- 3420
CPZ.CM.01.SIVcpzCAM13      ---A-----------TAGT-------C-...--C--T---CC-ACG-----C--A-A------------A-----A--T-----A--A--GAA---------------CA-A--AG-C-AG----A-------A-AA---GT---GAGA----C--T-----AC--G-A- 3561
CPZ.CM.05.SIVcpzEK505      ---A-----------T----------C-...-----A--TAGC--G---------AT-----------A--------T-----A-----------G---------T---G--AG----G--T--G-----G--GAC-T-TA-CT-A--A--T--C--C--AC-AG--- 3553
CPZ.CM.05.SIVcpzLB7        T--C--G--T--------A------C--...-----T---CC---G--------A--T-------------------------A--A-----G-----------------------A--AG----G--G-AG--G---A-AA--C--T----A--A-----C--------G--- 3536
CPZ.CM.05.SIVcpzMB66       ---A---C-T--------A------C--...--C------CC------------A--T-------------------------A--A--C------G-------------A--AG-C-A------------GA-CA--T-TT----A--A--C--T---------AG-A- 3527
CPZ.CM.05.SIVcpzMT145      ---A--G-------------------C-...--G--T----A------------A-A----------------C-----A--A--A---AG------------C-----G--A--C--G--G---------GA--CAT-GT--TGAA--T--C--T-----A--AG-A- 3523
CPZ.GA.88.GAB1             ---C-----------AGT--------C-...-----T---CC-ACCA----T--T-AT------------A-----A--T-----A--A---ACT--G---------------A--AG-C-A---G-A-------A-CA-T-G-T--GAAA-T--C--T------C-AG-A- 4042
CPZ.TZ.00.TAN1             -A-C-GGC--------A-TC-GTT-TCT...--C-----CCC--A------------T-----------------A------A--C-G----AAG----G------------A--A--C-----C---T-C-GGTCAAAGGA-T--GAGA----C--T-----AC-AG-A- 3625
CPZ.US.85.CPZUS            C--A---C-T-----TA---------C-...--GG------AG--G--------A--T-----G-----A-----CC-AC-A--A--C--GAA------G-----------------G---G--T---------A--A-A--TT--GAAA-T--T--T-----A---G-A- 4041
CPZ.CM.x.SIVcpzDP943       G--A---------------R-----A-C...---------CC---G--------A---------------------C-----A--C--A--GAA------------T-----C--GG-------T------G---A-A----AT---GAA--T--C--T---------AG-A- 3567
CPZ.CM.x.SIVcpzMB897       G--A--GC-G---------------A-C...-----------G--G------------T-----------------A--T--A--A--C--------------------------A---G-------G-------G--------GA--A--T-------A--A-----T-----A-GAG-A- 3547
N.CM.02.DJO0131            T--A-----------T--------A-C-...--G--A--CACT--G--------A-A-----------A--------T-AA--A--A-----G----------T-T---G--AG---------A---G-----G---T-TA-TGAA-----C--C-----A------- 3487
N.CM.04.04CM_1015_04       T--A-----------T--------A-C-...--G--A--TAGT--G--------A-A-----------A--------T-AA--A--A-----G----------T-T---G--AG---------------G-----G---T-T--TGAAA----C--C-----A------- 3486
N.CM.04.04CM_1131_03       T--A-----------T--------A-C-...--G--A--TAGT--G--------A-A-----------A--------T--A--A--A-----G----------T-T---G--AG---------------G-----G---T-TA-T-AA--T--C--C-----A--A---- 3485
N.CM.95.YBF30              T--A-----------T--------A-C-...--G--A--CAGT--G--------T-------------A--------T-----A--A-----G----------T-T---G--AG----G----------G-----G---T-TA-TG-A-----C--C-----A---G--- 3575
N.CM.97.YBF106             T--A-----------T--------A-C-...--G--A--TAGT--G--------A-A-----------A--------T-AA--A--A-----------------T-T---G--AG---------------G-----G---T-TA-TGAGA----C--C-----A---G--- 3573
O.BE.87.ANT70              GA-C-----------AG-C----A-GT...------T--TA-G--G---------AT-----------A--------T--A----A--C-------G--------------AG-ACA---G-A---G--GA-AA-A-AAT--GA---G--C--C-----A--AG--- 4035
O.CM.91.MVP5180            GA-C-------------------A-C-...------T--T-----G----------T-----------A--------T---A-T--A--C-------G--G-----------AG-ACA-----A---G--CA-AA-A-AGT--GAA--G-----C-----A---G--- 4010
O.CM.96.96CMABB637         -A-C----------------G----GT...------T--T---------------AT-----G-----A--------T---------A--CG--G--G-------------AG-ACA-----A---G---A-AA-A-ATT--GAG--G--C--C-----A---G--- 3465
O.CM.98.98CMA104           GA-C-----------AG------A-GT...------T--TA-G--G-----C-----T-----------A--------T--A--T--A-GC-------G--------------AG-GCA---G-A---G---A-AA-A-AAT--GA---G--C--C-----A---G--- 3467
O.CM.99.99CMU4122          GA-C-------------GA----A-GT...------T--TA-G--G--------A-AT-----------A--------T--AA-T--A---C-------G--------------AG-ACA---G-A---------GA-A-AAT--GAG--G--C--C-----A---G--- 3466
O.FR.92.VAU                GA-C-----------A-------A--T...--G--T--CA----G-----------T-----------A--------T--A--C--A---C-------G--------A----AG-ACA---G-A---G---A-AA-A-AAT--GA---G--C--C-----A---G--- 3547
O.SN.99.SEMP1299           GA-C-------------------A-GT...------T--TA-G--G---------AT-----------A-----------A--T--A---C----------C-----------AG-A-AG--G-----G---A-AA-T-AAT--GAG--G--C--C-----A-G-G--- 4034
O.US.99.99USTWLA           GA-C-----------T--A-----A-CT...---------T-----G---------AT-----------A--------T--A--T--A--------G--G--------------AG-ACA-----A---G--GA-AA-ACAAT--GAG--A--T--C-----A---G--- 3464
```

```
B.FR.83.HXB2              AGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGATTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAACCAGATCAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAGGAAAAGGTCTATCTGGCA 4151
Pol                       E__L__Q__A__I__Y__L__A__L__Q__D__S__G__L__E__V__N__I__V__T__D__S__Q__Y__A__L__G__I__I__Q__A__Q__P__D__Q__S__E__S__E__L__V__N__Q__I__I__E__Q__L__I__K__K__E__K__V__Y__L__A_
A1.GE.99.99GEMZ011        -A----T-----CC-----------------A----C---------T-----------------------------TG---G----------CAGG-----------A-------A----------A-AC----G----A---GA-----C---T-- 3335
A1.KE.00.KER2008          -AC----T-------------C----------A----C---------T----------------G------------------G----------CAG------------A----T--------------A--C-----GG------C--A-----C---T-- 3350
A1.KE.00.KNH1144          -AC----T-----C--------C-----A-----A----C--------------------------G------------------G----------CA-G------------------------------A--C-----GG------C--A-----C---T-- 3356
A1.KE.00.KSM4024          -AC----T-----C--------C-----------A----CG-----G-------------------G------------------G----------CAGG------------------------------A--C-----GG------C--A-----C---T-- 3347
A1.KE.00.MSA4069          -A-----T-----C--------C-----A-----A----C--------------------------G------------------G----------CAG----------------T-------------G---------------T--A-----C---T-- 3350
A1.KE.00.NKU3005          -A-----T-----CC-------C-----------A----C--------------------------G------------T-----G--C-------CAGG----------A---------------------A--C-----GG------C--A-----C---T-- 3350
A1.RU.00.RU00051          -A-----T------C-------------------A----C-A-------T----------T-----G------------T-----G----------CAGG------G-----A--------A----------AA-AY-----G----A------A---------T-- 3452
A1.RW.93.93RW_024         -AC----T-----CC-------C-----------A----C------G-------------------G------------------G----------CAGG---------G-----------G----------A-A------G-------C--A-----C---T-- 3353
A1.SE.95.SE8891           -AC-G--T-----C--------C-----------A----C---------T----------------G--------G-----T---G----------CAG----------------T--------------A--C-----GG------C--A-----C---T-- 3340
A1.SE.95.UGSE8131         -A-----T-----CC-------CC----------A----C--------------------------G------------------G----------CA-G------------------------------A--C-----G-------C--A-----C---T-- 3523
A1.TZ.01.A173             -AC----T-----CC-G-----C-----------A---CC---------------------------------------------G----------CAG-----C-----A-------------------A--C-----G----A--C--A--A--C---T-- 3350
A1.UA.01.01UADN139        -A-----T-----CC-------------------A----C---------T-----------------------------T-----G----------C--G----------A-------A----------A-AC----G----A---GA-----C---T-- 3350
A1.UG.92.92UG037          -A---T-T-----CC-------C-----------A----C--------------------------G------------------G--C--G----CAGG------------------------------A--C-----G-------C--A-----C---T-- 4167
A1.UG.99.99UGA07072       -AC----T-----CCT-T----C-----------A----C---------T----------------G------------------G----------CAG-------------------------------A--C-----GG------C--A-----C---T-- 3350
A2.CD.97.97CDKTB48        -A-----T-----C--T-G----A--------A--G-C--------------T-----G----------------------G-------AGC-------------------------------AA--------G---------G------C---T-- 3489
A2.CY.94.94CY017_41       -A-----T-----C--T-G----A--------A-----------G-------G----T-----G--------------------------AAGG----------A-------------------AA-A-----G---------G------C--AT-- 3508
A.SN.01.DDI579            -A-----T-----C--------C-----------A----C--------------------------G---------------------G----C--G----C------A---------G---------AA--C-----G----------A--A------T-- 3349
A.SN.01.DDJ369            -A-----T-----C-C------C-----------A----C---------T----------------G------------------G----G----CAGG----------A---------G---------AA--C-----G-G--------A--A--C---T-- 3352
A.SN.96.DDJ360            -A-----T-----C--------C-----------A----C-A------------------------G---------------------G----CA-G----------A---------G---------AA--C-----G-G--------A--A--C---T-- 3349
A.CD.02.02CD_KTB035       -------T----C--C-----C--A--A-----A-------------T-------G--------------G---------G----------CAGG----G------A-------------------A--C-----G----A------A----C---T-- 4145
A.CD.97.97CD_KCC2         -A-----T-----C--------C-----------A----C--------------T----------T----------C--G----------CAT-------------A---------G---------A--C-G--G----A--G--A-----C---T-- 4147
A.CD.97.97CD_KTB13        -A-----T-----C--------C-----------A--G-C---------T------------C-------T-------------G----------CAGG-----------G-C---------G---------A--C-----G-------G--A-----C---T-- 3355
B.AR.04.04AR151516        ----R---------C--------------------------------------T----------------------------------------A-G-----------A-----G------------------------------G-----C------ 3373
B.AU.87.MBC925            ---------------------------A--------------------------------------------------------------------------------------G---------------------------------------C------ 4162
B.BO.99.BOL0122           ----------------T--------------A---CC--------------------------------------------------------AGG-----------A-----G-----------A-------------------------C------ 3371
B.BR.03.BREPM2012         --------------C----------------A--------------------------------------------------G-----A-G-----------------------G-------------G-------------------A-----C------ 3576
B.CA.97.CANB3FULL         --------------A----------------------C-----------------------------------------------------A-----------------------G----------------------C-------G-----C------ 3440
B.CN.05.05CNHB_hp3        -------T------C--T---------------A---GC-------------------------G--G--------------------A-G-------T--A-------G-----------A------------------AA----C------ 4150
B.CO.01.PCM001            -------G------C-----------A-----------C-------------------------------------------------A-G------------A-------G------------------------------------------ 3356
B.GB.83.CAM1              -A-----------------------------------------G---------------------------------------------A-G-------------------G-------------G-------------------------C------ 4152
B.GB.86.GB8               --------------C--------------------------------------------------------------------------AGG------------G-----G-------------------------------------C--A--- 4171
B.GE.03.03GEMZ010         --C-----------C----------------A-------------------------------------------G------------A-G-----------------------G-----------A-----------C--------------C------ 3356
B.IT.05.SG1               --------------C-CT------A-------N-T---------------G--------------------------C----------------A---AA-N-----------G-------------A--------------------CT----- 3953
B.JP.05.DR6538            --------------C---------------------------------------------------T----------------------A-G-----------------G-----------A---------------------T--C--A--- 4155
B.KR.05.05CSR3            -------------------------------A-------G-------------------------------------------------------A------------------------------------C--------------G------C---A-- 4176
B.NL.00.671_00T36         --------------C------------A-----A---C---------------------------------------------------A-G-----------------G-------------------------------------C------ 3716
B.RU.04.04RU128005        --------------C-------------------A-------------T-------------------C---------------A-------A---------------G---G-----------A----------------------------C------ 3642
B.TH.00.00TH_C3198        --------------C--T----------------C--------------------------------------------------A-G--------------------G--------------------------------A--A--C------ 3365
B.UA.01.01UAKV167         --C-----------C-----------------C-----------------------------------------------------A-G-------------------G--------------------------------------C------ 3377
B.US.04.ES10_53           --C-----------C-----------------------------------------------------------------------A-----------------------------T-------------C----------A--C--A--- 4160
B.US.99.PRB959_03         -------G------CG----------------A-A------------------------T----------------------A-G------------------------------------C-------------------C------ 3371
C.AR.01.ARG4006           --------------C-G--------A--A-----A-----------------------G------------T------------------A-G-----------------------A--A------C------------------C---T-- 3335
C.BR.04.04BR013           -A------------A-G--------A--------A--C-C-----------------G-------G----T------------------A----------------------------A-----------G-------G------C---T-- 3615
C.BW.00.00BW07621         -A------------C-G-----------------A---CC----------------T----G------C--------------------A-G--------------------------A--AC---------------------C---T-- 3504
C.CN.98.YNRL9840          -A--G-----------A-----------A-----A----C----------------T----G-----------------------------A-G-----------------T--C---------A--A-------------G------C---T-- 3335
C.ET.02.02ET_288          -A------------C-G-----AC-A--------A----C---------------------G-----------G----------------CA-G-------C---G---------------AG-A---------------G------C---T-- 3353
C.GE.03.03GEMZ033         -A------------C-G-----------------A----C----------------T----G----------------------------A-G---------------T--------------A-----------C-------GAA---C---T-- 3332
C.IL.99.99ET7             -A------------CC-GA---------A-----A--G-C------------------C--G----------------------------A-G--------------------------A-------------------G-A---C---A-- 3332
C.IN.99.01IN565_10        -A--G---------C--------C----A-----A----C---------------------G----------------------------A-G-----------------T--C---------A--A-------------G------C---T-- 3532
C.KE.00.KER2010           -A--G---------C-G-----------A-----A----C---------------------G----------------------------A-G-------A----------C---------A--A-------G-------------C---T-- 3332
C.MM.99.mIDU101_3         -A--G-----------A-----------A-----A----C----------------T----G-------G--------------------A-G-----------------T--C---------A--A-------------G------C---T-- 3502
C.MW.93.93MW_965          --------------C-GT----------------A----C---------------------G----------------------------A-G-------A---T--C---------A------------C--------AA----C---T-- 3338
C.SN.90.90SE_364          -A------------C-G--G-----A--------A----C---------------------G-----------------C----C-----A-G----------------------------A----------------------G------C---T-- 3320
C.SO.89.89SM_145          -A------------C-GA----------A-----A----C---------------------G------------------G---------A-G------------------C---------A--A-------------G------C---T-- 3372
C.TZ.02.CO178             -A--G---------C-G-----------A-----A--G-C--------------T---------G------------------C------A-G----------------------------A--A-------------G------C---T-- 3332
C.UY.01.TRA3011           --------------C-G--------A-----C--A-----G-------------------G--------T--------------------A-------------------------------A--A-------------G------C---T-- 3323
C.YE.02.02YE511           -A-----G------C-------------A-----A--G-C-----T--------------G-------------C---------------A-G--------------------------A--A---------------------C---T-- 3326
C.ZA.04.04ZASK164B1       -A------------CTG-----C-----A-----A----C-----------------------------G--------------------A-G----G-----AC------C---------A--AC------C--------A--T--C---T-- 3565
C.ZA.05.05ZAPSK240B1      -A------------C-G-----------A-----A----C---------------------G-------G--------------------A-G-------------A------C---------A---------C--A---G------C---T-- 3555
C.ZM.02.02ZM108           -A------------C-GT-----C-------C--A----C---------------------G----------------------------A-------------A----T----------A----------------C-G------C---T-- 4120
D.CD.83.ELI               --------------A-------C----------------------------T----------------------------------A-G-------------------------------------------------T--C------ 3697
D.CM.01.01CM_0009BBY      --------------A-----------------------G-C-----G-----------T------------------------------CA-G-----------A-------G-----------AG-----------------------C--A--- 3350
D.KE.01.NKU3006           -A--------C---A--------------------------C-------------T------------G--T------------------A-G------------------G------C---------------C-----------A-----C--AT-- 3356
D.KR.04.04KBH8            --------------A--------------------------------------------T---------C-----------C--------A-G------------------G------------A---------------------------C----C 4110
D.TD.99.MN011             --------------A-----------------------C-----G-----------T----G------------------------------A----------AA-----G-----------A---------------------A--C------ 3375
D.TZ.01.A280              ----------C--AA-C----------------------C-----------A--------------------G-------------------A-----------------G-----------A---------G-------------T--C--A--- 3352
D.UG.99.99UGK09259        ----------C---A---------------------------------G--------------------------------------G-----AGG-------A------C-------------A------C-------------------C--AT-- 3356
D.YE.01.01YE386           -A--------C---A-----------------A---------------------------------------------T-----------A-G----------C-----------G------------------------A-----C--AT-- 3350
D.YE.02.02YE516           ----G---------CA---G------------A----C-----G--T---------T--------------------------------A-G---------AG----G-----------A-------------------------A-----C------ 3362
D.ZA.90.R1                --------------A---G-----------------------------------T-----------------------------------G------A-------------------G-----------------C--------------G---T--C------ 3497
F1.AR.02.ARE933           -----A-----G--C-GT-------A--------A---GC------G--------------G---------------------------A-G-----------A-----A--------------------------------------------------- 3445
F1.BE.93.VI850            ----------------T-----------------A----C------G---------------------------G--G----------A-G-----------AA-------------------A--------C---------G------C---T-- 3491
F1.BR.01.01BR125          --------------C-GT-------A--------A----C------G--------------G----------------------------A-G-----------A-----------------------A-------------------------C---T-- 3596
F1.ES.x.P1146             -------------AC--T-------A--------A----C------G--------------G-------------------------------G-----------------------------A------C---------G-A---C---T-- 3428
F1.FI.93.FIN9363          --------------C--T-------A--------A----C------G----------T--------------------------------A-G-----------------------------------C-----------A----C---T-- 3480
F2.CM.02.02CM_0016BBY     -A------------C--T-------------C--A----C---------------------G-------------------C---------A-G------------------C-----------A------C-------------------C---T-- 3341
F2.CM.95.MP257            -A------------C-GT---C------------A----C---------------------G-------------------C---------A-G-----------A------C-----------A------C---------G------C---T-- 3360
F2.CM.97.CM53657          -A------------C--T-------------C--A----C---------------------G-------------------C---------A-G------------------C-----------A------C----A---G------C---T-- 3341
G.BE.96.DRCBL             -A-----T------C-G------C-------C--ACA--C---------------------G-------------------G-----AGG---------AA--------------------A---C---------------------C---T-- 4108
G.CM.01.01CM_4049HAN      -A-----T------C----------------C--AA--CC---------------------G---------------------------AGG---------A-----T-----G--------A---C---------------------C---T-- 3353
G.CU.99.Cu74              -A------------C-G--------------C--A----C---------------------G---------------------------AGG---------A-------------------A--AC--------------------C---T-- 3755
G.ES.99.X138              -A------------C-AA-------A--A--C--A----C---------------------G-----------T---------------AGG---------A-------------------A---C---------------------C---T-- 3592
G.GH.03.03GH175G          -A-----G------CC-AT------------C--AAA--C------G--------------G---------------------------AGG---------A-----T-------------A---C---------------------C---T-- 4191
G.KE.93.HH8793_12_1       -A------------C-G-----G--------C--A----C---------------------G---------------------------AG-----------AC------------------A---C-----------------------T-- 3551
G.NG.01.01NGPL0669        -A-----T------C-G--------A-----C--A----C---------------------G---------------------------AGG---------A-------------------A---C---------------G---T--CT-AA-- 3356
G.PT.x.PT2695             -A------------C-G--------------C--AA--CC---------T-----------G-----------T-----------A--A-G---------A-------------------A---C---------------------C---T-- 4093
G.SE.93.SE6165            -A-----G------C-G-----------A--C--AA---C-----------------G---------G-----C-------------------AGG-----G---A-------------------A---C---------------------C---T-- 3548
H.BE.93.VI991             -A--------------------------A--G--A--GCC-------------G----------------------G-------T--A-G---------AC----T-------------G-A-------------------T----CT--T-- 3540
H.BE.93.VI997             -A--------------------------A-----A--G---------------G----T-----G-------------------T--A-G----G---------T------------AG-A---------G---------------C---T-- 3475
H.CF.90.056               -A--------------------------A-----A--G-------G-------G----T-----G-----C-------------C--A-G-----------------T-------------G-A---------G---------------C---T-- 3498
J.CD.97.J_97DC_KTB147     -AC------C-----T----------------A--GCA--------------------T-----G--------G--T--------------A-------------A-----T------------G-A---------G-------GA---------T-- 3345
J.SE.93.SE7887            -AC----C--C-----T-------A-G-----A--GC---------------------T------------G--G--T--A-----------A-G-----------A---------------------G-----------G---------------C---T-G 3465
K.CD.97.EQTB11C           -A-----------CC--T-------A--------A----C---------------------G----------------------------A-------------A-----T--------------A--------------C-G------C--AT-- 3347
K.CM.96.MP535             -A-----------C-G-T-------A--------A----C---------------------G-------------------G------A-------------T-----T--------------A---------------G-A----C--AT-- 3347
```

```
B.FR.83.HXB2                  AGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGATTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAACCAGATCAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAGGAAAAGGTCTATCTGGCA 4151
Pol                           E__L__Q__A__I__Y__L__A__L__Q__D__S__G__L__E__V__N__I__V__T__D__S__Q__Y__A__L__G__I__I__Q__A__Q__P__D__Q__S__E__S__E__L__V__N__Q__I__I__E__Q__L__I__K__K__E__K__V__Y__L__A_
01_AE.CF.90.90CF11697         -A-----T--G--CC--T----C-----------A----C---T-----T-------------C---C-----------------------G-----------CAGG-----------AA-------C-----------AA--C-----G---------GA-----C---T-- 4100
01_AE.CN.05.FJ051             -A-----T--G--CC-------C-----------A-A--C---------T---------------------------------G--------G-----------CAGG-----------AG-------C-----------G-------------------------------C---T-- 4161
01_AE.CN.06.FJ054             -A-----C-----C---T----C-----------A----C------C--T-----------------------------------------G-----------CAG--A---------AA-------C-----------G-AA-----------------A--GC-C---T-- 4176
01_AE.HK.x.HK001              -------T-----C---T----C-----A-----A----C---------T-----------------------------------------G-----------CAGG-----------AA----T-GC-----------G--C-----------------A-----C---T-- 3513
01_AE.TH.01.01TH_R2184        -A-----T--G--CC--T----C-----------A----C---------T-----------------------------G-----------G-----------CAGG-----------AG-------C-----------G--C----C-----------A-----C---T-- 3351
01_AE.TH.02.OUR769I           -A-----T--G--CC--T----C-----------A----C---------T-----------T-----------------------------G-----------CAGG-----------AG-------C-----------G--C-------------------------C---T-- 3350
01_AE.TH.90.CM240             -A-----T--G--CC--T----C-----------A----C---------T-----------------------------------------G-----------CAGG-----------AG-------C-----------G--C-------------G--A-----C---T-- 3725
01_AE.US.00.00US_MSC1164      -AC----T-----CC--T----C-----------A----C---------T-----------------------------------------G-----------CAGG-----------AG-------C-----------G--C-------------------------C---T-C 3356
02_AG.CM.02.02CM_4082STN      -------T----------G--C------------A----C---------T-----------------G-----------------------G-----------CAGG-----C-----A-------------------AA--C-----G-------G--A-----C---T-- 3350
02_AG.EC.x.ECU41              -A--G--T--------------C-----------A----C---------T-----------------------------------------G-----------CAGG-----------------T--C---------------C-----GG------C---------C---T-- 3247
02_AG.GH.03.GHNJ196           -A-----T------C-------C-----A-----A----C---------T-----------------G-------------T---------G-----------CA-G-----C-----AA----------T---------AA-AC-----G-------C--A-----C---T-- 4180
02_AG.NG.01.PL0710            -A-----T------C--T----C--------------C-C---------T-----------------------------------------G-----------CAGG------------A----T-------------AA--C-----G-------C--A--A--C---T-- 3335
02_AG.NG.x.IBNG               -A-----T------C--T----C--A--------A----C---------T-----------------G-----------------------G-----------CAGG-----------------------------------A--C-------------C--A-----C---T-- 3676
02_AG.SE.94.SE7812            -A-----T------C--T----C-----------A----C---------T-----------------G-----------------------G-----------CAGG------------C------C------------G--C-------------T-GA-----C---T-- 3523
02_AG.SN.98.MP1211            -A-----T------CT------C--A-----A--A----C---------T-----------------G-----------------------G-----------CAGG-----------------------------------A--C-----G-------C--A-----C---T-- 3348
03_AB.RU.97.KAL153_2          --------------C------------------------------------------------------------------------------------------A-G-------------------G---------G-----------------G-------------------C------ 3377
04_cpx.CY.94.CY032            -A--------------CT----------------A----C-------------------------G------A-------------------------------AG-----------T-----T-----------------------CGG-----C---------C---T-- 3517
05_DF.BE.x.VI1310             --------------C-C-----------A-----A-----------G--T---------------G--C---------------------------------A-G-------------------------------------A------G-----A--------------T-- 3533
06_cpx.AU.96.BFP90            -A-------------A-------A--A-------A---CC-------------------------G-------------------------------------A-------------AC-------C-----------A---C----------------AA----C---T-- 4179
08_BC.CN.97.97CNGX_6F         -A--G-----------A-----------A-----A----C-------------------T-----G------------G--------------------A-G-------------------T--C-----------A--A------------------G------C---T-- 3331
09_cpx.GH.96.96GH2911         --C----T--C--AC-------C--A--------A----C---------T-----------------------------------------G-----------CAGG-----------------------------------A--C-----G---A-----------C---T-- 3353
10_CD.TZ.96.96TZ_BF061        -A--------C---A-------------------A----C---------T-----------------G----------C--T-----------C--A-G-----------------------------------------------------A-----C--AC-- 3530
11_cpx.GR.x.GR17              -AC--G--------C--T---------A------A--------------T-----------T-------------G-----T----------------A-G------------------GC-------------A----------------------------C---T-- 3453
12_BF.AR.99.ARMA159           -------T--------CT-------A--------A----C-------------------------G-----------------T---------------A-G----G------------G-------------A----------------------------C---T-- 4156
13_cpx.CM.96.1849             -A--G---------C-G-----------------A----C-------------------------G-------------------------------AGG-------------------------------------A---C-------G--------------C---T-T 3556
14_BG.DE.01.9196_01           -A--G-----G---C-A-------------C--A----C----A--------------------G---------------T-----------------AGG-----------A------------------------A---C--------------------------C---T-- 3668
15_01B.TH.99.99TH_MU2079      -A-----------CC-GT----C----G------A----C---------T-----------------------------------------G-----------CAGG-----------AG-------C-----------G--C-------------------------C---T-- 3550
16_A2D.KR.97.97KR004          -A-----T-----C---T-G-----A--------A--------------------------T-----G-----------------------G-----------AGG-------------A---------T---------AA-------G----A----GA--------AT-- 3514
17_BF.PE.02.PE02_PCR0155      -------C------C-GT-------A--------A----C-----G-------------------G-------------------------------A-G----------------------------------------A------G------------------C---T-- 3341
18_cpx.CU.99.CU76             -A---G--------C-CT----A--------------GCAG-------------G-----T----G-------------------------T---A-------------AA------------------------G----------G-------------C---T-- 3456
19_cpx.CU.99.CU7              ---------------A-------------------------------------------T---------------------T-----------------A-G--------------------G-----------------------G-------G---A--C------ 3377
20_BG.CU.99.Cu103             ---------------C--------------------------------------G--------------------------------------------A---------------------G-------------------------------------T--C--A--- 3617
21_A2D.KE.91.KNH1254          -A-----------C---T-G-----A--------A-A-------G------------T-----G-----------------------G-----------AGG-------------------T---------------A---------GG-------G------C---T-- 3356
22_01A1.CM.01.01CM_0001BBY    -A-----T-----CC--T----C-----------A----C---------T-----------------G-----------------------G-----------CAGG-----------A----------T---------AA-AC-----G---------A-----C---T-- 3359
23_BG.CU.03.CB118             ---------------C----------------C--A----C-------------G-----------------------------------------------A--------------------G-----------A-------------------------T--C--A--- 3594
24_BG.CU.03.CB378             ---------------C--------------------------------------G-----------------------------------------------A--------------------G---------------------------------------T--C--A--- 3576
24_BG.CU.03.CB471             ---------------C------------------A-------------------G-----------------------------------------------A--------------------G---------------------------------------T--C--A--- 3591
25_cpx.CM.01.101BA            -A--G--T-----C--------C-----------A----C---------T----G-----T-----------G-------T-------------------A-C--Y---------A---T------------------------------------------C---T-- 3359
27_cpx.CD.97.97CDKTB49        -A-------------C-AT---------------A----C-------------------------C--G-----------------A-------GT-----A-G-----------A-----------G-------------------------------C---T-- 3359
28_BF.BR.99.BREPM12609        -------------CC-------------------A--------G--------------------------------C-----C---------A--------------------G--------------------------------------------------- 3517
29_BF.BR.02.BREPM119          --C----T------C-GT-------A--------A----C-------------------------G-----------------------G-----------CA-------------------------------------C------------AC--------C---T-- 3325
31_BC.BR.02.110PA             ----G--------CC-G--------A--------A----C-------------------------G-------------T---------G-----A-------------------------------------A-----------------G------C---T-- 3585
32_06A1.EE.01.EE0369          -A-------------A---------A--------A----C-------------------------G-----------------------------------A-------------A---------------------------C------------G-----------T-- 3786
33_01B.MY.05.05MYKL007_1      -A-----T--G--CC--T----C-----------A----C---------T---------------------T-------------G-----------CAGG-----------AG-------C-----------G-----------------------------C---T-- 3530
34_01B.TH.99.OUR2478P         -A-----T--G--CC--T----C-----------A----C---------T--------------------------G-------------G-----------CA------------AG-------C-----------G-AC-----------------A-----C---T-- 3353
35_AD.AF.05.05AF095           -A-----T-----CC-------C-----------A----C---------T-----------------G-----------------------G-----------CA------G----A---------------------A-AC--G--G-------C--A-----C--AT-- 3344
36_cpx.CM.00.00CMNYU830       -A-----T------C--T----C-----------A----C-A-------T-----------------G-----------------------G-----------CAGG------------------T--------------A--C-----G-------T--A-----C---T-- 3356
37_cpx.CM.00.00CMNYU926       -A-----T------C-------CC----------A-A--C------G--T-----------------G-----------------------G-----------CAG------C-----AA-------C--G---------A-AA-T---G-------G-GA-----C---T-- 3341
38_BF1.UY.05.UY05_4752        ---------------CTAA-------A-------A---CC-------------------------G-----------------------------------A-------------------------------------A--------T---------------C---T-- 3562
39_BF.BR.03.03BRRJ103         -A------------------------------------C---------------------T---------------------------------------A-G-------------------G-A-----------G-------------------A-----C------ 3640
40_BF.BR.04.04BRRJ115         ----------------C------------------G----------------T-----------------------------------TG-----A-G-----------A----------------------------------C-------G---T--C--A--- 3703
42_BF.LU.03.luBF_05_03        ---------------C-RT-----C---------A-----------------------T-----G----------G-------------G-----A-G-----------A----------------------------A--------W--------------C---T-- 3689
43_02G.SA.03.J11223           -A-----T--------C-----C---------C--A----C---------T-------------C--G-----------------------G-----------CAG-----------A-----------------------A--C-----G---A--C--A-----C---T-- 3676
U.CD.83.83CD003               -AC---------C-------C--A----------A-----------C-------------T-----G-------------------------------AGC-----------------------------------------A-------------------A-----C---T-- 3646
U.CD.90.90CD121E12            -AC---------C-------T--A----------A-----------C-------------T-----G-------------------------------AGC-----------------------------------------A-------------------A-----C---T-- 3350
U.GR.99.GR303                 -A-----------CC-CT-------A--------A----C---------------T-----------G-----------------------G-----------CACG-----G-----A-------------------------A--C-----GG---A----G---T--C---T-- 3422
U.NL.95.U_NL_95_H10986_D1     -A-------------C-AT---------------A----CT-C--------------------G-----------------------G-----A-C----------T-----T------------------A--------C-G-----CC--------C---T-- 4036
CPZ.CD.90.ANT                 -A-----G--G---CT-A-G---C-AG-------AAC-GGCCC---C--T-----C-----T--------------G---G--T-G----GTACC------------------CCCC----GG-AG-------CC--A-A----------GA---C--A-T--C--CT-C 3590
CPZ.CM.01.SIVcpzCAM13         -----GGG--CT-G-----------A--A--C--T-A-CA-C----G-----------T-----T------TG-----G--T----GT--------------C----------A-----A--------------AG-A-----T----------C-----T--C--TT-C 3731
CPZ.CM.05.SIVcpzEK505         -A-----G--TG-CCTCA-G--C--A--A---G-A---CA---------T-----C--T-----T--G---TCC-------A-----CT-------------A--------------A-----A-----G--------AG-AC-T------------GAA-A-----TT-C 3723
CPZ.CM.05.SIVcpzLB7           ----------C---C-AT----C--A--A--C--A---CCT--G-----T-----T--T--T--------------------C-------C-----G-----G-----T-----A--G--G--------C-----AA-TC-----------------A--T--CT-AA-- 3706
CPZ.CM.05.SIVcpzMB66          -A--------T--AC-AT-G-----A--A-----A-A-G----------T-----C--T--T------------C-G-----A-------GT-----T--AC-------------AA----A-----G--------G-AC-G--C-----A--G--A--G-----CT-- 3697
CPZ.CM.05.SIVcpzMT145         -A---------G---TAA-G--C--A-----C--A---CAG-----C--TG-G--C--T-----C------TGC-G-----T-A---T----C-----C--C------------TA----A-----G--------AG------------------AA-T-----AT-- 3693
CPZ.GA.88.GAB1                -A---A-G--TT-GCT---G--C-----------A-ATCA-C-G--T--------G--T-------------TG-----G--T----GT----G---------C-----------A------------G--------AG-------T-----A-----AA----C--CT-C 4212
CPZ.TZ.00.TAN1                -A---TGG---G-AG-------C--AA-A--C--A---GC-C-G-----T-----C-----T--C-------TTA-G---G-TT-A--G-G-TT--------------C--CCCCA---AG-G-----T--TC-AA------C-C-----AC-GCAA-T------A--- 3795
CPZ.US.85.CPZUS               --C-CAC---GG-A--CT----A--AA-A-----A-AGAATAC---T--TG----C--T-G--C------TC-----------C--CT-T--G-----C---------------A--G---------G--------AG-A---------G--A-----AAGT---A-CT-- 4211
CPZ.CM.x.SIVcpzDP943          --C-GAC---TG----------CC-A--A---G-A-A-CATAT---T--------C--T--T--C------TG---------A----CT-T--G-------------------G----A--G--T---TT----------AG----------------------AAG---C--CT-- 3737
CPZ.CM.x.SIVcpzMB897          -AC-G-----T--A--C-----CC-A--------A-A-G----G-----------T--T--------------------G-------T----G--T--CATT------------CA----A-----G---------------C--G-G-------------A-T-----TT-- 3717
N.CM.02.DJO0131               ----------T--CCT---G--C--A--A-----A---CAG--------------C--T--T--T--G-----TC-G-----A-----CT----C------A----------------A--G--G-GC-------------AG--C-C--------------GA--T-----CT-T 3657
N.CM.04.04CM_1015_04          ----------T--CCT-A-G--C--A--A---K-A---CAG--------------C--T--T--T--G-----TA-G-----A-----CT------------A-G--------------A--G--G-GC-------------AG-AC-C--------------GA---------CT-T 3656
N.CM.04.04CM_1131_03          -------T--T--CCTCA-G--C--A--A--G--A---CAG--------------C--T-----T--G-----TA-G-----A-----CT----G-----A----------------A--G--G-GC-------------AG-AC-C--------------GA---------CT-T 3655
N.CM.95.YBF30                 ----------T--CCT-A-G--C--A--A--G--A---CGG--T-----------C--T-----T--G-----TA-G-----A-----TT----G-----A----------------A--G--G-GC-------------AG--C-C--------------GA--T-----CT-T 3745
N.CM.97.YBF106                ----------T--CCT-T-G--C--A--A--G--A---CAG---C----------C--T-----T--G-----TA-G-----A-----CT-----------A----------------C--G--GGGC-------------AG-AC-C--------------GA---------CT-T 3743
O.BE.87.ANT70                 -A---ATG--G--A-TA-----C--A--------CAA-GA-AC------T-----------T--------------G--CG----CTCCT-C-----TACA--G--------CCCTA----TC-G-------------G-AC---C-----------C----G-----TA-- 4205
O.CM.91.MVP5180               -A---ATG--TG-A-TAA----C-----------CAAGGAGC-----------------------------T---G--C-----ATCCT-C------ACA--G-----C--CCCTA----TC-G--G---------G-AC---C----------CGA--G-----TA-- 4180
O.CM.96.96CMABB637            -A---ATG---G-A-TAA----A--A--------CAAGGA-ACG-----------------T--C-------T---G--C-----CTCCT-C-----TACA--------C--CCCTA----TC-G--G--------AG-AC-G-C---------GCG-ACG---A-TA-- 3635
O.CM.98.98CMA104              -A---ATG---G-ACTAT----C--A--------CAAGGA-A-------T-----------T----------T---G--C-----CTCCT-T-----TACA--G----------CCTA-----C-A--G--------AG-AC---C---G------C-A--G--C--TA-- 3637
O.CM.99.99CMU4122             -A---ATG-----A-TRT----C--AA-------CAAGGA-AC------------------T-------T-----G--C------TCCT-C-----TACG--G-----T--CCCTA----TC-G--G---------G-AC---C-----------C----A-----T--- 3636
O.FR.92.VAU                   -A---ATG---G-A-TGT-------A--------TAA-GA-A-------------------T----------T-C-G--C-----ATCTT-C-----TACA--G-----C--CCCTA----TC-G--G--------AG-AC-G-C-----G----C-A--G-----TA-- 3717
O.SN.99.SEMP1299              -A---ATG---G-A-TAT----C--A--------CAAGGA-AC------T------CA---T--C-------T---G--C-----CTCCT----G--TACA--G--------CTCTC----TC-G--G---------G-AC---C-----------C----G-----TA-- 4204
O.US.99.99USTWLA              -A---ATG---R-A-TGT----C--A-----C--TAA-GAGAG------------------T----------T---R--C-----ATCCT-C-----TACA--G-----C--CCCTA----TC-A--G--------AG-AC-G-C------R--GC-AR-G-----CA-- 3634
```


Pol p66 RT, Pol p15 Rnase H end, Pol p31 Integrase start
B.FR.83.HXB2
Pol
A1.GE.99.99GEMZ011
A1.KE.00.KER2008
A1.KE.00.KNH1144
A1.KE.00.KSM4024
A1.KE.00.MSA4069
A1.KE.00.NKU3005
A1.RU.00.RU00051
A1.RW.93.93RW_024
A1.SE.95.SE889I
A1.SE.95.UGSE8131
A1.TZ.01.A173
A1.UA.01.01UADN139
A1.UG.92.92UG037
A1.UG.99.99UGA07072
A2.CD.97.97CDKTB48
A2.CY.94.94CY017_41
A.SN.01.DDI579
A.SN.01.DDI369
A.SN.96.DDI360
A.CD.02.02CD_KTB035
A.CD.97.97CD_KCC2
A.CD.97.97CD_KTB13
B.AR.04.04AR151516
B.AU.87.MBC925
B.BO.99.BOL0122
B.BR.03.BREPM2012
B.CA.97.CANB3FULL
B.CN.05.05CNHB_hp3
B.CO.01.PCM001
B.GB.83.CAM1
B.GB.86.GB8
B.GE.03.03GEMZ010
B.IT.05.SG1
B.JP.05.DR6538
B.KR.05.05CSR3
B.NL.00.671_00T36
B.RU.04.04RU128005
B.TH.00.00TH_C3198
B.UA.01.01UAKV167
B.US.04.ES10_53
B.US.99.PRB959_03
C.AR.01.ARG4006
C.BR.04.04BR013
C.BW.00.00BW07621
C.CN.98.YNRL9840
C.ET.02.02ET_288
C.GE.03.03GEMZ033
C.IL.99.99ET7
C.IN.99.01IN565_10
C.KE.00.KER20f0
C.MM.99.mIDU101_3
C.MW.93.93MW_965
C.SN.90.90SE_364
C.SO.89.89SM_145
C.TZ.02.CO178
C.UY.01.TRA3011
C.YE.02.02YE511
C.ZA.04.04ZASK164B1
C.ZA.05.05ZAPSK240B1
C.ZM.02.02ZM108
D.CD.83.ELI
D.CM.01.01CM_0009BBY
D.KE.01.NKU3006
D.KR.04.04KBH8
D.TD.99.MN011
D.TZ.01.A280
D.UG.99.99UGK09259
D.YE.01.01YE386
D.YE.02.02YE516
D.ZA.90.R1
F1.AR.02.ARE933
F1.BE.93.VI850
F1.BR.01.01BR125
F1.ES.x.P1146
F1.FI.93.FIN9363
F2.CM.02.02CM_0016BBY
F2.CM.95.MP257
F2.CM.97.CM53657
G.BE.96.DRCBL
G.CM.01.01CM_4049HAN
G.CU.99.Cu74
G.ES.99.X138
G.GH.03.03GH175G
G.KE.93.HH8793_12_1
G.NG.01.01NGPL0669
G.PT.x.PT2695
G.SE.93.SE6165
H.BE.93.VI991
H.BE.93.VI997
H.CF.90.056
J.CD.97.J_97DC_KTB147
J.SE.93.SE7887
K.CD.97.EQTB11C
K.CM.96.MP535

Pol p66 RT, Pol p15 Rnase H end Pol p31 Integrase start

```
B.FR.83.HXB2                TGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTCAGTGCTGGAATCAGGAAAGTACTATTTTTAGATGGAATAGATAAGGCCCAA...GATGAACATGAGAAATATCACAGTAATTGGAGAGCAATGGCTAGTGATTTTAACCTGCCACC 4318
Pol                         _W__V__P__A__H__K__G__I__G__G__N__E__Q__V__D__K__L__V__S__A__G__I__R__K__V__L__F__L__D__G__I__D__K__A__Q__.__D__E__H__E__K__Y__H__S__N__W__R__A__M__A__S__D__F__N__L__P__P
01_AE.CF.90.90CF11697       -----------------G--G------------------------------------T----------A--G--G--------------G--------A--T---...--A--------A-G-------------------A-------------------TT------- 4267
01_AE.CN.05.FJ051           -----------G--------G------------------------------------T-A-----------G--G--------------G-----------T---...--A--------A-G---------C---------A-------------------TT------- 4328
01_AE.CN.06.FJ054           --------------------G------------------A-----------------T-A-----T--A--G--G--------------G-----------T---...--G--------A-G---------C---------A-------------------TT------- 4343
01_AE.HK.x.HK001            --------------T-----G--------T---A-CA--------------------T-A--------------G---------A----G-----------T---...--A--------A-G---------C---------A-------------------TT------- 3680
01_AE.TH.01.01TH_R2184      -----------------G--G--------T---------------------------T-A-----------G-----------------------------T---...--A--------A-G---------C---------A-------------------TT------- 3518
01_AE.TH.02.OUR769I         --------G--------G--G------------------------------------T-A-----------G--G--------------G-----------T---...--A--------A-G---------C---------A-------------------TT-A----- 3517
01_AE.TH.90.CM240           -----------------G--G------------------------------------T-A-----------G--G--------------G-----------T---...--A--------A-G---------C---------A-------------------TT------- 3892
01_AE.US.00.00US_MSC1164    -----------------G--G---------------------------C--------T-A-----------G--G--------------G-----C-----T---...--A--------A-G------T--C---------A-------------------TT------- 3523
02_AG.CM.02.02CM_4082STN    --------------------G-----G------------------------------AG------------------------------T---------------...--A--G-----A-G---------C-----------------------------T-------- 3517
02_AG.EC.x.ECU41            --------------T-----G--------------------------------T---AA---------A-G------------------C--------A------...--A--T-----A-G---------C---------A-------------------T-------- 3414
02_AG.GH.03.GHNJ196         --------------------G------------------------------------AA-----------G------------------C--------A------...--A--------A-G------T--C-------------------A---------T-------- 4347
02_AG.NG.01.PL0710          --------------------G---------------------------C--------AA----G-T--A--G-----------------C--------A------...--G--T-----A-G---------C-----------------------------T-------- 3502
02_AG.NG.x.IBNG             --------------------G------------------------------------AA------------G-----------------C--------A------...--A--G-----A-G---------C-------AG--------------------T-------- 3843
02_AG.SE.94.SE7812          --------------------G-----------------G--G---------------AG------------G-----------------C---------------...--A--G-----A-G---------C-----------------------C-----T-------- 3690
02_AG.SN.98.MP1211          --------------------G------------------------------------AAG-----------------------------C--------A------...--A--G-----A-G---------C-----------------------A--C--T-------- 3515
03_AB.RU.97.KAL153_2        -----------------------------------------T---------------------------G-------------------------------A---...--A------------------G---------------------------------------- 3544
04_cpx.CY.94.CY032          --------------------G-----------------------------------CAA---------A--G--G--------------------------T---...--A-------------------A---C--------------------------T------T- 3684
05_DF.BE.x.VI1310           --------------------G--------------------------------------------T--A--------------G--------------A--T--G...--------------------T--C-----------------------------T-------- 3700
06_cpx.AU.96.BFP90          --------------------G------------------------------------AA------------G-----------------C--------A-----G...--A--T-----A-G---------C-----------C-------A---------T-------- 4346
08_BC.CN.97.97CNGX_6F       --------------T--------------------------------------A---AA---C-----------G------C----------------A--T---...--A--G-----A--G--------C-----------------------C-----T-------- 3498
09_cpx.GH.96.96GH2911       --------------T------------------------------------------AG---------A--------------------C--------A--T---...--A--------A-G------T--C-----------------------------T-------- 3520
10_CD.TZ.96.96TZ_BF061      -----------------------------------------G-----------A---AGG--------A--------G---C----------------A--T---...--A--G-----------C----AC-------------------------------------- 3697
11_cpx.GR.x.GR17            ---------------------------------------------------------T----------------G------C----------------A--T-C-...--A-G------A-G---------C--C-----------------------C--T-------- 3620
12_BF.AR.99.ARMA159         -----------------G--G-----------------------------------------GG----A---A----G-----------G-----------A--G...--G--------A----------AC-----------------------------T-------- 4323
13_cpx.CM.96.1849           --------------------G--------------------------------T---AA---------------G--------G-----C-----C--A------...--A--T-----A-G---------C---------A-------------------T--A----- 3723
14_BG.DE.01.9196_01         --------------------G-----------C------------------------AG------------G---T----C--------T--------A------...--A--C-----A-----------------------------------------T-------- 3835
15_01B.TH.99.99TH_MU2079    ---------------------------------------------------------T-A-----------G--G--------------G-----------T---...--A--------A-G---------C---------A-------------------TT------- 3717
16_A2D.KR.97.97KR004        --------------------G-------A---------G-----C------------T----C-----------GT-------------G-----------T---...-----------A-----C-G----------------------CA---C-----T-------- 3681
17_BF.PE.02.PE02_PCR0155    --------------------G-------------------------G------T--------------A---A----G-----------G-----------A--G...--G--------A-G--------AC-----------------------------T-------- 3508
18_cpx.CU.99.CU76           --------------------G-----------------------------------CAG---------A--GG--T-------------------A-----T---...--A--------T----------AC-----------------------------TT------- 3623
19_cpx.CU.99.CU7            --------------------G------------------------------------T-------T-----G--G-----C--------G-----------T---...--A--------A-----------C-----------------------------T-------- 3544
20_BG.CU.99.Cu103           --------------------G------------------------------------AG------------------------------C--------A------...--A--------A-G---------C-------------------A---------T-------- 3784
21_A2D.KE.91.KNH1254        --------------------G-----------------------C------------T----------A-----GT----C--------------------T---...--A--------A------------------------------CA---C-----T-------- 3523
22_01A1.CM.01.01CM_0001BBY  -----------------------------------------------------T---T-A--------A--GA--T-----C-------G-----------T---...--A--C-----ACG---------C--------GA------T------------T-------- 3526
23_BG.CU.03.CB118           --------------------G------------------------------------AG------------------------------C--------A------...--A--G-----A-----------C-------A---------------C-----T-------- 3761
24_BG.CU.03.CB378           -----G---------------------------------------------------AG------------------------------C--------A------...--A--G-----A-G---------C-----------------------------T-----T-- 3743
24_BG.CU.03.CB471           -----G--------------G------------------------------------AA---------------G--------------C--------A------...--A--G-----A-G---------C-----------------------------T-----T-- 3758
25_cpx.CM.01.101BA          --------------------R------------------------------------AG---------------G------C-------C--------A-----G...--A--G-----A-G---C--T--C-----------------------------T-------- 3526
27_cpx.CD.97.97CDKTB49      --------------------G-----G--T-----G---------------------AG-------------A----------------T--------A------...--A--T-----A-G--------CC-----------------------------TT------- 3526
28_BF.BR.99.BREPM12609      -----------------G--------------------------------------------------A--GA--------C-----------------------...---------------------------------T------A---------------A----- 3684
29_BF.BR.02.BREPM119        --------------------G------------------------------------------G----A---A----G-----------G-----------A--G...--G--C-----A----------AC---------T-------------C-----T-------- 3492
31_BC.BR.02.110PA           --------------T--G--------G--------------------------A---AAG--------------G--G---C-------------------T---...--A--------A-G---------C-----------------------------T-------- 3752
32_06A1.EE.01.EE0369        --------------------G------------------------------------AG------------G---------C-------C--------A------...--A--G-----A-G---------C-----------C-------A---------T-------- 3953
33_01B.MY.05.05MYKL007_1    --------------------G---------------A--------------------T-A--------A--G--G--------------------------T--G...--A--------A-G---------C---------A-G-----------------TT------- 3697
34_01B.TH.99.OUR2478P       --------------------G------------------------------------T-A-----------G--G--------------G--------A--T---...--A--------A-G---------C---------A-------------------TT------- 3520
35_AD.AF.05.05AF095         --------------T--G--G-----------------------------------CA-------------G--------C--------------G-----T---...--A--------A-G---------------------------G-----------TT------- 3511
36_cpx.CM.00.00CMNYU830     ---------------------------------------------------------AAG------------A----------------T--------A-----G...--A--T-----A-G---------C-----------------------------T-------- 3523
37_cpx.CM.00.00CMNYU926     -----------------G--T------------------A-----------------T-A-----------G-GGG-------------G-----------T---...--A--T---A-A-G---------C-------------------A---------T-------- 3508
38_BF1.UY.05.UY05_4752      --------------------G-----------------------------------------------A---A----G-----------------------A--G...-----G-----A-----------C-----------------------------T-------- 3729
39_BF.BR.03.03BRRJ103       ------------------------------------------------C--------T-A-----T--------------------------------A------...--A--G----------------A------------------C-------------------- 3807
40_BF.BR.04.04BRRJ115       ---------------------------------------------------------T-A-----------------G---C-----------------------...--G--G----------------A---------------------------------A----- 3870
42_BF.LU.03.luBF_05_03      --------------------G------------------------------------A-------T------A--T-----C----------------A------...--A--T-----A--G----------------A----------------------A-A----- 3856
43_02G.SA.03.J11223         --------------------G------------------A-----------------AG---G--------------------------C--------A------...--A--G--Y--A-G---------C---------------------------G-T-----C-- 3843
U.CD.83.83CD003             -----G--------------G-------------------C------------------C--G--------------------G-----------------T---...--G--G-----A-G------T--C-----------------------------T-------- 3813
U.CD.90.90CD121E12          -----G--------------G--------------------------------------A-----------------------G-----------------T---...--A--------A--------T--C-----------------------------T-------- 3517
U.GR.99.GR303               --------------------G------------------A-----------------AA---------A--------------------C-----C--A------...--A--T----RA-G------T----------------------A---------T-------- 3589
U.NL.95.U_NL_95_H10986_D1   --------------------G-----------------------------------CAA------T--A--GA----------------------------T---...--A--C-----A----------CC-----------------C-----C-----T-------- 4203
CPZ.CD.90.ANT               --------------T-----C--A--------------G-----C-----------CAA---G-----AC----C--C---C-G--A--C--------A--T---...--G--C-----T--------T--C---------T--T-A--AGA---A-AC--T--T--C-- 3757
CPZ.CM.01.SIVcpzCAM13       --------T-----T--G--C--A-----------------------G---------T----------------------CC-G--------------A--A--G...--A--------A-G------T--C-------A---TT----GG----A------T-A----- 3898
CPZ.CM.05.SIVcpzEK505       --------T-----T-----T-----------------GA----C-----------CA-A--------------CT----CC-------T-----------T---...--A--------A-G------T----------AG--------C-----------T--A--C-- 3890
CPZ.CM.05.SIVcpzLB7         --------T-----T--G-----A--------------------C------------T-A-----T--------C--T--C--G------------CGA-----G...--A--G-----A-----C--T-A------------T-----C--------C--T-------- 3873
CPZ.CM.05.SIVcpzMB66        --------T------------------------------A----------------CT-A-----T--------C--C--CC-G-----G--------------G...--A--G-----A-----C--T-A---------G--C-----C-----------T-------- 3864
CPZ.CM.05.SIVcpzMT145       ----------------G---T--A--------------G---------C----T---T-A--G-----A-----T--T--C--------------A--A--T---...--A--------A-G---C---TC--------AG---------CAG--C------T-A----- 3860
CPZ.GA.88.GAB1              --------------------T--A-----------G-----------------------A--------------G-----CC-------------C-G-------...--A--------A-GG-----T----------A---T-----------------TT-A----- 4379
CPZ.TZ.00.TAN1              -----------C--T-----T--A--G--T------G-------C-----G--T---AAAAAT--T--A---A--T----CC-G--------TA--G-A--A--G...--A--C-----T-------------------A---TT-A---GA---A-A---T-----C-- 3962
CPZ.US.85.CPZUS             --------------T------------------------A--------C--------T-A--G-----A-----T--T--CC-------C--------A--A--G...--A--C-----C--------T----------C---T-----C-------------------- 4378
CPZ.CM.x.SIVcpzDP943        -----G--------------------------------G-----------------CT-------T----G---T--C--C--------T-----------A---...--A--------A--------T-A------------T-----A-----C--C-----C----- 3904
CPZ.CM.x.SIVcpzMB897        --------------------G------------------A----C--------T-----A--------------C--T--CC-G---------------------...--A--------------C----AC-------A---T-----AGA---A-----TT------- 3884
N.CM.02.DJO0131             --------T-----T-----T--------------G--G-----C--------T--CT-A-----T--A------T----CC-------T-----A--A------...--A--T-----R-G------------C--------------C------------T-A--CT- 3824
N.CM.04.04CM_1015_04        -----R--T-----T-----T--------------G--G-----C--------T--CT-A-----T--A------T-----C-------T-----A--A------...--A--G-----A-G------------C----A---------C------------T----C-- 3823
N.CM.04.04CM_1131_03        --------T-----T-----T--------------G--G-----C--------T--CT-A-----T--A------T----CC-------T-----A--A------...--A--G-----A-G------------C--------------C------------T-R--C-- 3822
N.CM.95.YBF30               --------T-----T-----T--------------G--G-----C--------T--CT-A-----T--A---A--T----CC-------T-----A--A------...--A--T-----C-G---------C-------A---------C------------T-A--C-- 3912
N.CM.97.YBF106              --------T-----T-----T--------------G--G--------W-----T--CT-A-----T--A-R----T----CC-------T-----A--A------...--A--------A-G-----------------A---------C------------T-A--C-- 3910
O.BE.87.ANT70               -----T--T--C--T-----C--A------------A--A-------------A--CAAG-AT--T--A-G---C--G--CC----A--------CC----A---...--A--T-----A--------T----------A----C-A--------A---GGA--A----- 4372
O.CM.91.MVP5180             -----T--T--T--------C--A------------A--A-------------A--CAAA-AC--T--A-G---C--G--CC-G--A---------C----A---...--A--T-----A--------T---------------T-A--------C---GGAT-A----- 4347
O.CM.96.96CMABB637          -----T--T--T-----G--C--A--GR--------A--A-------------A--CAAA-AC--T--A-GG--C-----CC----A---------C-R--A--G...--A--T-----A-----------C------------T-A--------C---GGA--A----- 3802
O.CM.98.98CMA104            -----T--T--C--T--G--T--A--G------------A-------------A--CAAG-AT--T--A-G---CT----CC----A--G------C----A--G...--A--T-----A--------T--CM-----------T-A--------C---GGA-------- 3804
O.CM.99.99CMU4122           -----T--T--T--T-----C--A------------A--A-------------A--CAAG-AT--T--A-G---C-----CC----A--K-----CC----A---...--A--T-----A--G-----T--C------------T-A--------C---GGA--A----- 3803
O.FR.92.VAU                 -----T--T--T--T-----T--A--G---------A--A-------------A---AAA-AT--T--A-G---C-----CC----A---------C----A---...--A--T-----A--------T---------------T-A--------C---GGA--A----- 3884
O.SN.99.SEMP1299            -----T--T--C--T-----C--A------------A--A-------------A--CAAG-AT--T--A-G---C-----CC----A--------CC----A---...--A--T-----A--G-----T--C------------T-A------------GGAT-A----- 4371
O.US.99.99USTWLA            -R---T--T--T--------C--A--G---------A--A-------------A--YAAA-AT--T--A-G---C-----CC----A---------C----G--G...--A--T-----A--G-----T--C----R-------T-A--C---------GGAT-A----- 3801
```

```
B.FR.83.HXB2 TGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAA...TGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAGGAAAAGTTATCCTGGTAGCAGTTCATGTAGCCAGTGGATATATAGAAG 4485
Pol __V__V__A__K__E__I__V__A__S__C__D__K__.__C__Q__L__K__G__E__A__M__H__G__Q__V__D__C__S__P__G__I__W__Q__L__D__C__T__H__L__E__G__K__V__I__L__V__A__V__H__V__A__S__G__Y__I__E__
A1.GE.99.99GEMZ011 -A----------G------------------------...-----A--------G-----T-----------G-----------------G---------T-------C------C-------------A---A-A--------C--------------C---------- 3669
A1.KE.00.KER2008 -A--A-------G------------------------...-----A--------G-----------------------------------G--G------T-------C------C-------------A--T-----------C--------------C---------- 3684
A1.KE.00.KNH1144 -A----------G------------T-----------...-----------------------A-----------------C--------G---------T--------------C-----------A-A--TT----------C--------------C---------- 3690
A1.KE.00.KSM4024 -A----------G------------------A-----...--------------G--------------------------C--------G---------T-------C------C-----------A-A---A-A-----------------T-----C---------- 3681
A1.KE.00.MSA4069 ----A-------G------------T-----------...--------------G--------A--------------------------G-----------------C------C-------------A--T--------------------------C--C------- 3684
A1.KE.00.NKU3005 -A----------G------------------------...--------------G--------------------------C--------G---------T-------C------C-------------A--T-----------C--------------C---------- 3684
A1.RU.00.RU00051 -A----------G------------------------...-----A--------G--------A--C--------------------------C-----GT-------C------C-------------A--------------C--------------C---------- 3786
A1.RW.93.93RW_024 -------------------------------------...--------------G--------------------------C--------G---------T-------C------C-------------A--T--A--------C--------------C---------- 3687
A1.SE.95.SE8891 -A----------G------------------------...-----A--------G-----------------------------------G--G------T-------C------C-------------A--T-----------C--------------C---------- 3674
A1.SE.95.UGSE8131 -A----------G------------------------...--------------G--------------------------C--C---------------T-------C------C-------------A--T-----------C--------------C---------- 3857
A1.TZ.01.A173 -A-----------------------------------...-----A--------G-----AG----------------T-----------G--G------T--------------C-------------A--T-----------C------------------------- 3684
A1.UA.01.01UADN139 -A----------G------------------------...-----A--------G-----T-----------G-----------------G---------T-------C-----CC-------------A---A-A--------C--------------C---------- 3684
A1.UG.92.92UG037 -A----------G------------------------...--------------G-----------------------------C-----G---------T--------------C-------------A--T-----------C-----G--T-----C--C------- 4501
A1.UG.99.99UGA07072 -A----------G------------------------...-----A--------G-----------------------------C-----G--G------T-------C------C-------------G--T-----------C--------------C---------- 3684
A2.CD.97.97CDKTB48 -A-------C---------------------------TG.GTCAG---------G--------------------------C--------------------------C------C-----------A-----G----------C--------------C---------- 3825
A2.CY.94.94CY017_41 ------------------------T------------...--------------G-----------------------------------------------------C------C-T--------------------G-----C-----G--------C---------- 3842
A.SN.01.DDI579 -A----------G------------------------...-----A--------G---------------------------------------------T-------C------C-------------C--TA-A--------C--------------C---------- 3683
A.SN.01.DDJ369 -A----------G------------------------...-----A--------G-------A------------------------------------------C--C------C-------------C--TA-A--------C--------------C---------- 3686
A.SN.96.DDJ360 -A-------G--G------------------------...-----A--------G---------------------------------------------T-------C------C-------------C--TA-A--------C--------------C---------- 3683
A.CD.02.02CD_KTB035 -A----------G-----------T------------...--------------G-----------------------------------G---------T-------C------C----------------------------C--------------C-----G---- 4479
A.CD.97.97CD_KCC2 ------------------------T------------...--------------G---------------------------------------------T----C--C------C----------------T-----------C--------------C---------- 4481
A.CD.97.97CD_KTB13 ------G-----G-----------T--------C---...--C-----------G-----------------------------------G--------GT--------------C----------G-----T-----------C--------------C---------- 3689

B.AR.04.04AR151516 ---------------------------T---------...-----------------------A--C--G-----------------------------G--------------C--G---------------------------------------------------- 3707
B.AU.87.MBC925 -------------------------T-----------...---------------------------------------------------------------------------------------------------------------------------------- 4496
B.BO.99.BOL0122 -------------C----------T-A----------...------G----G---------------------------------------------------------------------------A------------------------------------------ 3705
B.BR.03.BREPM2012 -A-----------------------------------...-----------------------------------------------------------G--------C-----C-----------------------G-----C--------------G---------- 3910
B.CA.97.CANB3FULL ---------------G---------------------...---------------------------------------------------------------------------------------A----------------C------------------------- 3774
B.CN.05.05CNHB_hp3 ---G---------------------T-----------...--------G-----------------------------------C------------------------------C-----------A-------------G---------------------------- 4484
B.CO.01.PCM001 -A-----------------------------------...-----A---------------------------------------------G-----------------------------------A------------------------------------------ 3690
B.GB.83.CAM1 -------------------------------------...-----------------------------------------------------------G--------------------------------------------G------------------------- 4486
B.GB.86.GB8 -------------------------------------...---------------------------------------------------------------------------------------A-------A---------------------------------- 4505
B.GE.03.03GEMZ010 ---G---------------A-----T-----------...------A----------------------------------------------------G---------------------------A-----T------------------------------------ 3690
B.IT.05.SG1 -A-----C-----A--T-----C-----G-------T...--------------------AT---T---------G---N-----------T-----------A--G---------A-A-------G--A----NN---C----C-------------N--T-------- 4287
B.JP.05.DR6538 ---G---------------------------------...--C---A-------------------------------------C--------C-----------------------------G---A------------------------------------------ 4489
B.KR.05.05CSR3 ---G---------------------------------...--------------------------------G-----------------------------------------C-----G------A-----------------------------------G------ 4510
B.NL.00.671_00T36 ---------------------------T---A-----...-----------------------------------------------------------G---------------------------A------------------------------------------ 4050
B.RU.04.04RU128005 -------------------------------------...--C---------------------------------------------------------T--------------------------A----------------A--------------G---------- 3976
B.TH.00.00TH_C3198 ------G------------------------------...--------------------------------------T-----C------------------------------------------A------------------------------------------ 3699
B.UA.01.01UAKV167 A------------------------------------...-----------------------------------------------------------G---------------------------A------------------------------------------ 3711
B.US.04.ES10_53 -----------G---------------T---------...-----A-----------------------------------------------------------------------G---------A----T------------------------------------- 4494
B.US.99.PRB959_03 ---G---------------------------------...-----------------------------------------------------------------C---------------------------------------------T------------------ 3705

C.AR.01.ARG4006 CA----------------------T------------...--------------G-----T-CA------------------------------------T--------------------------A-C--------------C-----------------C------- 3669
C.BR.04.04BR013 CA-------------------------T---------...--------------G--------------------------------C------------T----------------------G---A-C--------------C-----------------C------- 3949
C.BW.00.00BW07621 CA----------------------------------G...--------------G--------A--------------------------G---------T-------------------------GACC--------------C--------------C--C--G---- 3838
C.CN.98.YNRL9840 CA----------------------T---G-----C--...--------------G-----------------------------------G---------T--------------------------A-C--------------C--------------C--C------- 3669
C.ET.02.02ET_288 CA----------------------T------------...--------------G-----------------------T-----------G-----------G-----------C------------A-C--------------C--------------C--C------- 3687
C.GE.03.03GEMZ033 CA----------------------T-----------G...-----A--------G-----------------------------------G---------T--------------C-----------A-C---T----MS----C--------------C--C------- 3666
C.IL.99.99ET7 -------------------------------------...--C--AT-------G--------A--------------------------G---------T----------------------------C--------------C--------------C--C------- 3666
C.IN.99.01IN565_10 CA----------------------T---------C-G...--------------G---------------------------------------------T-------------------G------A-C--------------C--------------C--C------- 3866
C.KE.00.KER2010 CA----------------------T------------...--------------G-------CA-----------------------G--G---------T-G------------------------A-C--------------C-----------------C------- 3666
C.MM.99.mIDU101_3 CA----------------------T---------C--...--------------G-----------------------------------G---------T--------------------------A-C--------------C--------------C--C--G---- 3836
C.MW.93.93MW_965 CA-C---------------------------------...--------------------T--A--------G-----------------G---------T--------------------------A-C-----------------------T-----C--C------- 3672
C.SN.90.90SE_364 CA----------------------------------G...--------------G-----A--A--------------------------G---------T--------------------------A-C--------------C--------------C--C------- 3654
C.SO.89.89SM_145 C----------G------------T------------...--------------G-----------------------------------G---------T--------------------------A-C--------------C--------------T--C------- 3706
C.TZ.02.CO178 CA-----------------------------------...-------A------G--------A-----------------C--------G---------T--------------------------A-C--T-----------C-----------C--C--C------- 3666
C.UY.01.TRA3011 C---A-------------------T------------...----------------------CA-----------G-----------------------GT-------C-----------------GA-C--------------A-----------------C------- 3657
C.YE.02.02YE511 CA-------------------------T---------...------T----G--G--------A--------------------------G---------T----------------------G---A-C--------------C--------------C--C------- 3660
C.ZA.04.04ZASK164B1 CA-------------------------------C---...--------------G-------CA--------G-----T-----C-----G---------T----------------------------C--T-----------C--C-----------C---------- 3899
C.ZA.05.05ZAPSK240B1 CA----------------------T------------...--------------G---------------------------------------------T----------------------------C--T-----------C-----------------C--G---- 3889
C.ZM.02.02ZM108 C----------G-------------------------...--------------G---------------------------------------------T--------------C-------------C-----A--------C--------------C--C--G---- 4454

D.CD.83.ELI C--G--------------------T------------...------------------------------------------------------------T-------------C--------------------------------------------C---------- 4031
D.CM.01.01CM_0009BBY ---G--------------------T------------...-----A------------------------------------------------------T--------------C-------------------------------------------C---------- 3684
D.KE.01.NKU3006 ---G--------------------T------------...---------------------T--------------------------------------T-------------C-----------------------------C--------------T---------- 3690
D.KR.04.04KBH8 ---G--------------------TT-----------...-------A----------------------------------------------------T----------C---C-------------------------------------------C---------- 4444
D.TD.99.MN011 ---G--------------------T------------...-----AA-----------------------------------------------------T-------------C---------C--A-------------------------------C---------- 3709
D.TZ.01.A280 ---G---------------------T-----------...------------------------------------------------------------T-G-----C--------------------------------------------------C---------- 3686
D.UG.99.99UGK09259 ---G--------------------T--------C---...--C------------------T--------------------------------------T--------------------------A----------------C--------------C---------- 3690
D.YE.01.01YE386 ---G--------------------T------------...--------------G------T--------------------------------------T-------------------------GA----A-----------C--------------C---------- 3684
D.YE.02.02YE516 ---G-----------G--------T--------C---...--C--A-A----------------------------------------------------T--------------C-----------A-------------------------------C---------- 3696
D.ZA.90.R1 ---G--------------------T------------...------------------------------------------------------------T--------------------------------A-A-----------------------C---------- 3831

F1.AR.02.ARE933 ------------------------T------------...------T-------G-----------------------T---------------------T-------------------------------------------C--------T--------CC------ 3779
F1.BE.93.VI850 -A----------------------T-----------G...--------------G-----------C-----------------------G---------T-------C--------------C--G-----------------C--------T----AG--C------- 3825
F1.BR.01.01BR125 ------------------------T-----------G...-----A--------G-----------------------T-----------G---------T----------------------G--------------------C--------T-----G--CC------ 3930
F1.ES.x.P1146 ------------G-----------T------------...--------------G--G--------------------------------G---------T--------------------------A-------A--------C--------T--------C------- 3762
F1.FI.93.FIN9363 ------------------------T-----------G...--------------G-----------------------------------G---------T--------------------------A----------------C--------T--------C------- 3814
F2.CM.02.02CM_0016BBY -A-----------------------------------...------T-------G--------A-----------------C--------G---------T--------------------------A----------------C--------T-----C---------- 3675
F2.CM.95.MP257 -------------------------------------...--------------G--------------------------C--------G---------T--------------------------A-------A--------C--------T-----C---------- 3694
F2.CM.97.CM53657 ---------C---------------------------...--------------G--------------------------C--------G------------------------------------A-----T----------C-----G--T-----T---------- 3675

G.BE.96.DRCBL -A-------------------G---------------...--C---T-------G---------------------------------------------T-G------------------------A----------------C--------------C---------- 4442
G.CM.01.01CM_4049HAN ---------------------G---------------...--------------G---------------------------------------------T-------------C------------A-------A--------C--------------C---------- 3687
G.CU.99.Cu74 -A--A----------G-----G---------------...--------------G--------------T------------------------------T-------------------------GA----------------C-----G--------C---------- 4089
G.ES.99.X138 ----A----------------G---------------...-----A--------G--------------C-----------------------------GT--------------------------A----------------------G--------C---------- 3926
G.GH.03.03GH175G -A-------------------G---------------...--------------G------T-A--------------------------G---------T--------------------------A-----A-A--------C--------------C---------- 4525
G.KE.93.HH8793_12_1 -A-------------------G---------C-----...--------------G--------------T------------------------------T--------------------------A----------------A--------------G---------- 3885
G.NG.01.01NGPL0669 ---------------------G---------------...--C-----------------------------G---------------------------T-------------------------GA----------------C--------------C---------- 3690
G.PT.x.PT2695 ---G-----------------G---------------...-----A--------G--------------C-----------------------------GT----C--------------------GA----------------C-----G--------C---------- 4427
G.SE.93.SE6165 -A-------------------G---------------...--------------G---------------------------------------------T--------------------------A-----A-A-----------------------C---------- 3882

H.BE.93.VI991 -A----------------------T------------...------T-------G-----------------------------C-----C---------T-------------C------------A----T-----------C--------------C---------- 3874
H.BE.93.VI997 -A----------------------T------------...-----------------------------------G--------C---------------T-------------------------------T-----------C--C-----------C------S--C 3809
H.CF.90.056 -A----------------------T------------...--------------G-----------------------------C---------------T-------C--------G------C-------T-----------C--------------C---------- 3832

J.CD.97.J_97DC_KTB147 -A----------------------T------------...-----------------------------------G--------------G--G------T----C---------------------A----------------C---A----------C---------- 3679
J.SE.93.SE7887 ---------------G--------T------------...-----------G--G-----------------------------------G---------T--------------C-------------------------------------------C---------- 3799

K.CD.97.EQTB11C ------------------------T------------...--------------G--------A--------------------------G--------GT--------------------------A-------A-----------------T-----C---------- 3682
K.CM.96.MP535 -A-------C--------------T--------C---...--------------G--------A------------------------------------T--------------------------A----------------C--------T--C--C---------- 3681
```

```
B.FR.83.HXB2               TGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAA...TGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAGGAAAAGTTATCCTGGTAGCAGTTCATGTAGCCAGTGGATATATAGAAG 4485
Pol                        __V__V__A__K__E__I__V__A__S__C__D__K__.__C__Q__L__K__G__E__A__M__H__G__Q__V__D__C__S__P__G__I__W__Q__L__D__C__T__H__L__E__G__K__V__I__L__V__A__V__H__V__A__S__G__Y__I__E__
01_AE.CF.90.90CF11697      -A-------G----------------A----------...-----A--------------T--------------G--------------G---------T-------C--------------------C------------------C--C--G-------------------- 4434
01_AE.CN.05.FJ051          -A----------G-------------A----------...-----A--------G-----T--------------G--------------G---------T-------C-------C----G--------C------------------C--C--G-----------C------- 4495
01_AE.CN.06.FJ054          --------------------------AT------C---...-----A--------G-----A-----------------------C-----G--------GT----C--C-------C-------------A-C------------------C--C--G-------------------- 4510
01_AE.HK.x.HK001           -A----------G-------------A----------...-----A--------G-----T-----------------T-----------G---------T-------C-------C----G--------C------------------C--C--G-------------------- 3847
01_AE.TH.01.01TH_R2184     ----A-------G------------------------...-----A--------G-----T--A-----------G--------------G---------T-------C-------C-------------C------------------C--C--G-------------------- 3685
01_AE.TH.02.OUR769I        -A----------G-------------A----------...--------------G-----T--------------G--------------G---------T-------C-------C-------------C------------------C--C--G--T----------------- 3684
01_AE.TH.90.CM240          -A----------G--------A---A----------...-----A--------G-----T-----------------------------G---------T-------C-------C-------------C------------------C--C--G-------------------- 4059
01_AE.US.00.00US_MSC1164   -A----------G-------------A----------...-----A--------G-----T-----------------------------G---------T-------C-------C-----------A-C------------------C--C--G-----------C------- 3690
02_AG.CM.02.02CM_4082STN   -A-------------------G---------------...--------------G----------------------------------------------T-------------------------G------------------C--------------C---------- 3684
02_AG.EC.x.ECU41           ---------------------G---------------...-----A--------G--------A------------------G-----------------T---------T--------------A-----------------C--------------C---------- 3581
02_AG.GH.03.GHNJ196        -A-------------------G---------------...--------------G----------------------------------------------T---------------------------A-----------------C--------------C--C------- 4514
02_AG.NG.01.PL0710         AA---------G---------G---------------...--------------G--------A------------------G-----------------T-------------------------GA----T-A------------------------C---------- 3669
02_AG.NG.x.IBNG            -A-------------------G---------------...------A-G-----G---------------------------G-----------------T---------------------G---A----T-A-----C--------------C---------- 4010
02_AG.SE.94.SE7812         ---------------------G---------------...--------------G--------------C-------------------------------T---------------------------A-----------------C--------------C---------- 3857
02_AG.SN.98.MP1211         -A-------------------G---------------...--------------G----------------------------------------------A---------------------------A----A-A----------C--------------C---------- 3682
03_AB.RU.97.KAL153_2       ---G---------------------------------...-----AT------------------C-----------------------------------------------------------A------A------------------------------------ 3711
04_cpx.CY.94.CY032         A--G-----------G--------T------A-----...--------------G--------------------G--------------G--------GT----------------------T------A----------------G--T--------C------- 3851
05_DF.BE.x.VI1310          ------------------------T-----------G...------T-------------------------------------------------------T---------------------G-----------A-----------------------C---------- 3867
06_cpx.AU.96.BFP90         -A-------------------G-----------C---...--------------G--------A--------------T-----------G---------T-------C-----CC------------A-A----------------C--------------C---------- 4513
08_BC.CN.97.97CNGX_6F      CA----------------------T---------C--...------------------------------------------------G---------T--------------------------A-C--T-------------C--------------C--C------- 3665
09_cpx.GH.96.96GH2911      ---------------------G---------------...--------------G----------------------------------------------T---------------------------A----T-----------C--------------T---------- 3687
10_CD.TZ.96.96TZ_BF061     ---G--------------------T------------...----------G---------T----------------------C---------------T---------------------G-------------------------C--------------C---------- 3864
11_cpx.GR.x.GR17           ----A-------G-------------A----------...--C--A--------G----------------------------A-----------T--------------C-----------A-A---------------C--------------C---------- 3787
12_BF.AR.99.ARMA159        ------------------------T-----------G...--------------G-----------------T-----------G---------T----C----------------------AC---------------C--------------G--CC------- 4490
13_cpx.CM.96.1849          ---------------------G-----T---------...--C-----------G-------C----------------------------------------T---------------------------A-----------------C--------------C---------- 3890
14_BG.DE.01.9196_01        ---------------------G---------------...-----A--------G--------------C---------------------------GT---------------------------A-----M----------C----G--------C---------- 4002
15_01B.TH.99.99TH_MU2079   -A----------G--------G---------------...-----A--------G--G--T-----------------------------G---------T-------C-------C-------------C------------------C--C--G-------------G------ 3884
16_A2D.KR.97.97KR004       ---G--------------------TT-----------...------G-------G----------------------------------------------T-------C-------C-----------A-------A----------C--------------C---------- 3848
17_BF.PE.02.PE02_PCR0155   ------------------------T-----------G...-----A--G-----G--------------G---A----T-----C-----G---------T-------------------------------------------C--------T-----G--CC------- 3675
18_cpx.CU.99.CU76          A--------------G--------T------------...--------------G-----------------------------C-----G---------T----CC---------------G---A----A-A----------C--------------C---------- 3790
19_cpx.CU.99.CU7           -A----------G--------G---------------...--------------G-----T-----------------------------------------T-------C-------C-------------A--------------C----G--------C--C------- 3711
20_BG.CU.99.Cu103          -A-------------G-----G---------------...--------------G--------------C-------------------------------T----C----------------------A----------G------C----G--------C---------- 3951
21_A2D.KE.91.KNH1254       -A--A-------------------TT-----------...--------------G----------------------------------------------------C------------------------------------C--------------C---------- 3690
22_01A1.CM.01.01CM_0001BBY -A----------G--------G---------A-----...-----A--------G--------A-----------------C--------G---------T-------C-------C-------------G---T-------------C--------------C---------- 3693
23_BG.CU.03.CB118          ---------------G-----G---------------...--------------G--------------C-------------------------------T---------------------------A-----------------C----G--------C---------- 3928
24_BG.CU.03.CB378          -A-------------G-----G---------------...--------------G--------------C-------------------------------T---------------------------A-----------------C----G--------C---------- 3910
24_BG.CU.03.CB471          -A-------------G-----G---------------...-----------------------------C-------------A-----------------T---------------------------A-----------------C----G--------C---------- 3925
25_cpx.CM.01.101BA         -A-------------------G---------------...--------------G------C---------------------------------------T---------------------------A----T--A--G------C--------------C---------- 3693
27_cpx.CD.97.97CDKTB49     -A-------------------G---------------...--------------G--------------G--G----------------------------T---------------------------A----------G------C--------------C---------- 3693
28_BF.BR.99.BREPM12609     -A-----------------------------------...-----------------------------------------------------------G--------------C-----------------A------------------------------------------ 3851
29_BF.BR.02.BREPM119       -A----------------------T--T--------G...--------------GA----------------------T-----------G---------T--------------G------------------A-A----------C--------T-----G--C------- 3659
31_BC.BR.02.110PA          CA----------------------T------------...--------------G-------CA----------------------------------GT--------------------------A-C--------------C-----------------C------- 3919
32_06A1.EE.01.EE0369       -A-------------------G-----------C---...--C-----------G------------------------------------G---------A-------C-----CC---G------A-A---A-A----------C--------------T---------- 4120
33_01B.MY.05.05MYKL007_1   -A----------G-------C-----AT---------...-----A--------G-----T------------------------------------G-G-------C-------C-------------C------------------C--C--G-------------------- 3864
34_01B.TH.99.OUR2478P      -A----------G-------------A----------...-----A--------G-----T--------------G--------------G---------T-------C-------C-------------C------------------C--C--G------------------A 3687
35_AD.AF.05.05AF095        -A----------GC------------A----------...-----A--------G-----------------------------------G--G------T-------C-------C------------A---TA-A----------C--------------C---------- 3678
36_cpx.CM.00.00CMNYU830    -A----G--------------G---------------...--C-----------G----------------------------------------------T---------C-----------------A----A-A----------C--------------C---------- 3690
37_cpx.CM.00.00CMNYU926    -A-------------------G---------------...------T-------G-----T--------------G-----C--------------------G-----C-------C-C-G-----------A--------------C--C--G--------C--------G- 3675
38_BF1.UY.05.UY05_4752     ------------------------T-----------G...--------------G-----------------T-----------G---------A---------------------------A----A-A----------C-----------------CC------- 3896
39_BF.BR.03.03BRRJ103      -------------------------------------...--------G-------------------------------------------G-----C--------CC--------------------------------------------------------------- 3974
40_BF.BR.04.04BRRJ115      -------------------------------------...-----------------G-----------------------C------------------G--------C-----CC------------------------------C-----------------------G- 4037
42_BF.LU.03.luBF_05_03     ---G-----R---------A-----T-----------...-----------------------------------------------------------------------------------------------R-------------------------T-----G--CC------- 4023
43_02G.SA.03.J11223        ---------------------G---------------...-------A------------------------------------------------------T-------------C----------------A--------------C--------------C---------- 4010
U.CD.83.83CD003            -A-----------------------------------...--------------G------------------------C--------G---------T--------------CC----G-----------T----------------C--------------C---C------ 3980
U.CD.90.90CD121E12         -A----------------------T------------...----------------------A--------------------C--------G--------------------CC----G------A-----------------C--------------C---------- 3684
U.GR.99.GR303              ----A----------------G----------C-GG-...AT--G---------G---------------------------------------------GT-------------------------------T-C---G------C--------T-A---C---------- 3756
U.NL.95.U_NL_95_H10986_D1  A-----------------------T--T---------...-----A--------G-----T-----------------------------G--G------T---------------------------A-----------------CC--------T-----C---------- 4370
CPZ.CD.90.ANT              -A-T--G--T--------TA----ACAG---------...-----CG----G----------G-------------------C------------------G-------C--C-----------T-----A---A-A----------G------T-T-----C-TC------- 3924
CPZ.CM.01.SIVcpzCAM13      CA-------------G--------------T-A----...------G----------G--TC-------G--G--T--T-----C-------C------GA-G-----C-----------------C---A-C---A-A------------------A-------------C------- 4065
CPZ.CM.05.SIVcpzEK505      CA-------T-----------------T--------G...-----AT-G-------------A-----G---A-CA-----------GG--------T-------C-----------G--G------A-----T--------C----------G--------C--CT------ 4057
CPZ.CM.05.SIVcpzLB7        CA-T-----------G--------T--T---------...----------G-----G--A-------G-----------C--------------G-T-------------CC-------------C--T-A---------G--------T-----C-----G---- 4040
CPZ.CM.05.SIVcpzMB66       CA-------------G--------A-AT--------G...-----AT----------G-----A-------G---------------------------T----C--C-----C--G--------GA-C-----A--G-----A-----------------G-----G---- 4031
CPZ.CM.05.SIVcpzMT145      AA----G--T---------------CAA---------...--------------------A--------------T-----C----------------GA-T-----C--C------------------A-A--TA-A--G-----C-----G-----------C------- 4027
CPZ.GA.88.GAB1             CA-------------------G---CAT--------G...--C---G----------------G--G-----------C-----G--T------G-----------C--CC-------C-----G---A-A--G--------C----------------C---------- 4546
CPZ.TZ.00.TAN1             A--T--G--T--------TA-T--TCAG----CCA---...--C--TA----------G-----T--A--------G--G----AC-------A--C-------A----C-----C--CC-------C---C--G--C---A-A----------G--------T-----T-TC------- 4129
CPZ.US.85.CPZUS            A--G--C--T--G--G------G--------------...-------C----------G-----A-----G--G-----T--C---------T--C-----G-----------------------G---A-C--TA-A--G-----A-----G------------CC------- 4545
CPZ.CM.x.SIVcpzDP943       ------G--C-----------------T-----C--G...--C--AT-----------------------G----------C--C--------G---------T----C--C-------C-------G-----C-----A--G-----A--------------------CC-----G- 4071
CPZ.CM.x.SIVcpzMB897       CA-------C---------------------------...--C--AT---------------------G----------------------------------C--C-----------------GA-C---T-A-----T--A------------------C-----G- 4051
N.CM.02.DJO0131            ------G------------------------C-----...--C-----------G--------A---------G--CA-T------------G-G-----G---------------------------A-C--T--T-----G-C-------G--------C--CT------ 3991
N.CM.04.04CM_1015_04       CA----G------------------------C-----...--C-----------G-------------------G--CA-T------------G-G-----GT-------------C----------------A-C--T--T-----G-C-------G--------C--CT------ 3990
N.CM.04.04CM_1131_03       CA----G------------------------C-----...--C-----------G-------------------G--CA-T------------G-G-----GT-------------C----------------A-C--T--T---------C-----G--------C--CT------ 3989
N.CM.95.YBF30              CA----G------------------------C-----...--C-----------G-------------------G--CA-T------------G-G------T-------------C---------G------A-C-----T--G--G--C-----G--------C--CT------ 4079
N.CM.97.YBF106             CA-----------------------------C-----...--C-----------G--------------G--GA-CA-T------------G-G-----GT-------------C----------------A-C--T--T---------C-----G--------C--CT------ 4077
O.BE.87.ANT70              A--G--G--C--G-----CA-T--T------CC----...-----TA-------G-----A--T-----T--G---------C-------A-G--------A-------C------A-G-----C---A-C--AA-A--T--T--C-----G--A-----G-TC------- 4539
O.CM.91.MVP5180            AA-------C--G-----CA-T--T--T---CC----...--C--TA-------G-----A-C------T----------AC--C----AG---------A-G-----------------------C---A-C--AA-A--T--T--C--------A---AC-T--------- 4514
O.CM.96.96CMABB637         A--------C--------TA-T--T--T---CC----...-----TA-T-----G-----A--T-----T--G-----T-AC-------A----------A-G-----C---------------C-----C--AA-A--T--T--C-----G--A-------TC------- 3969
O.CM.98.98CMA104           A-----G--C--------TA-T--T-AT---CC----...-----TA-------G-----M--T-----G----------C-------A-G--------A-G-----C------AC---------C-----C--AA-A-----T--C-----G--A-----G-TC------- 3971
O.CM.99.99CMU4122          A--G-----C--G-----CA-T--T-AT---CC----...-----TA-------G-----A--T-----T--G---------C-------A-G--------A-------C------AC---------C-----C--AA-A--T--T--------G--A-----G-T-------- 3970
O.FR.92.VAU                A--------C--------CA-T--A--T---CC---G...-----TA---G-A-G-----A--------C--G---------C-------A----------A-G-----C---------------C--GA-C--AA-A--T--T--C-----G--A-------TC------- 4051
O.SN.99.SEMP1299           A--G--G--C--------CA-T--T-AT---CC-C--...-----TA-------G-----A--T-----T--G---------C-------A-G--------A-G-----C------C----------C---A-C--AA-A--T--T--C-----G--A-------TC------- 4538
O.US.99.99USTWLA           A--------C--G-----CA-T--T--T---CC----...-----TA---G--AG---C-A------R-T--G------------------A----------A-G--C--C---------------C--GA-C--AA-A--T--T--C--------A-------TC------- 3968
```

```
B.FR.83.HXB2             CAGAAGTTATTCCAGCAGAAACAGGGCAGGAAACAGCATATTTTCTTTTAAAATTAGCA...GGAAGATGG...CCAGTAAAAACAATACATACTGACAATGGCAGCAATTTCACCGGTGCTACGGTTAGGGCCGCCTGTTGGTGGGCGGGAATCAAG...CAGGAATTT 4646
Pol                      A__E__V__I__P__A__E__T__G__Q__E__T__A__Y__F__L__L__K__L__A__.__G__R__W__.__P__V__K__T__I__H__T__D__N__G__S__N__F__T__G__A__T__V__R__A__A__C__W__W__A__G__I__K__.__Q__E__F_
A1.GE.99.99GEMZ011       ----------C--------------A--A--G--------C-----GC-----------...---------...---------GT-G---C--A--T-----CC-----------A-CAG-G-A---A--T-------------A-AT---C-A...--------- 3830
A1.KE.00.KER2008         ----------C--------------A-----G--------C-----GC-----------...---------...---------GT-G---C--A-------------C------A-C---G-AT---AA--A-------------AAATG--C-A...--A----A- 3845
A1.KE.00.KNH1144         ----------C--------------A--A--G--------C---A-AC-----------...---------...---------GT-----C--A--------------------A-C---G-A---AA--A-------------A-AT---C--...--------- 3851
A1.KE.00.KSM4024         ----------C--------------A--A--GG-------C-----AC-----------...---------...---------GT-G-G----A--T-----------------A-----G-AT---AA--A-------------AAAT---C-A...--------- 3842
A1.KE.00.MSA4069         -------A--C--------------A--A--G--------C---A-A------------...---------...---------GT-G---C--A-------------C------A-C---G-A---AA--A-------------AAAT---C-A...--------- 3845
A1.KE.00.NKU3005         ----------C--------------A--A---G-------C-----AC---G-------...---------...---------GT-G---C--G--------------------A-----G-AT---AA--A-------------AAATG----A...--------- 3845
A1.RU.00.RU00051         ----------C--------------A-----G--------C-----GC-----------...-----G---...---------GT-----C--A--T-----------------A-CAG-G-A---AA--T-------------AAAT---C-A...--------- 3947
A1.RW.93.93RW_024        ----------C--------------A--A--G-----C--C-----AC-----------...--C------...---------GT-G---C--A------T------------A-C---G-A---AA--A-------------AAAT---C-A...--------- 3848
A1.SE.95.SE8891          -------G--C--------------A--------------C-----GC-----------...---------...---------GT-G---C--A--------------------A-C---G-AT---AA--A-------------AAATG--C-A...--------- 3835
A1.SE.95.UGSE8131        ----------C--------------A-----GG-------C-----AC-----------...---------...---------GT-G---C--A--------------------A-C---G-AT---AA--A-------------AAATG--C-A...--A------ 4018
A1.TZ.01.A173            ----------C--------------A-C---G--------C-----GC-----------...---------...---------GT-G---C--A--------------------A-C---G-AT---AA--A-------------AA-TG--C-A...-------A- 3845
A1.UA.01.01UADN139       ----------C--------------A-----G-----T--C-----GC-----------...---------...---------GT-----C--A--T-----CC---------TA-CAG-G-A---A--T-----C--------AAAT---C-A...--------- 3845
A1.UG.92.92UG037         -------------------------A-----G--------C-----AC---GC------...---------...---------GT-G---C--A--------------------A-C---G-A---AA--A-------------AAATG-T--A...--------- 4662
A1.UG.99.99UGA07072      ----------C--------------A-----G--------C-----AC-----------...-----G---...---------GT-G---C--A-------CC----------A-C---G-AT---AA--A-------------AAATG--C-A...--------- 3845
A2.CD.97.97CDKTB48       -------C--C--------G-----A-----G--------C-----A------------...---------...---------GT--------A------GCC----------A---AG-A---A--A-------------A-ATG---A...--------- 3986
A2.CY.94.94CY017_41      -------C--C---A----------A-----T--------C---A-A-----C------...---------...---------GT--------A------GCC-------T-A---A--A---A--A-------------A--T---C-A...--A------ 4003
A.SN.01.DDI579           ----------C--------------A-----GG-------C--C--GC-----------...---------...---------GT-----C--A-------------C------A----G-AA-G-AA--A-------------AAAT----A...--------- 3844
A.SN.01.DDJ369           ----------C--------------A--------------C--C--G------------...---------...---------GT-----C--A-------------C------A----G-A--G-AA--A-------------AAAT---C-A...--------- 3847
A.SN.96.DDJ360           ----------C--------------A--------------C--C--GC-----------...---------...---------GT-----C--A-------------C------A----G-AA-G-AA--A-------------A-AT----A...--------- 3844
A.CD.02.02CD_KTB035      -------C--C--------------A-----GG-------C---A-AC-----------...---------...---------GT--------A--------------------A-CA--G-A---AAA-A----C--------A--TG--C-A...--------- 4640
A.CD.97.97CD_KCC2        ----------C--------------A-----G--------C---A-AC-----------...---------...---------GT--------A---------T----------A-CA--G-A---AA--A-------------AA-T---C-A...--------- 4642
A.CD.97.97CD_KTB13       ----------C--------------A-----GG-------C---A-AC-----------...---------...---------GT--------A--------------------A-C---G-A---AA--A-------------AAAT---C-A...--A------ 3850
B.AR.04.04AR151516       -------------------G--------A-----------C---A-C------------...---------...--------------C----A--------------------A--A--------A--T-------------A--G----T...--------- 3868
B.AU.87.MBC925           -------------------G--------------------C-----C------------...---------...-------------------A-------CC----------A--AA---A---A-----------------G-----...--------- 4657
B.BO.99.BOL0122          ----------C-----------------------------C---A-C------------...-----G---...-------------------A--------------------A--AA-G---C-A--------C--------A--GG----...--------- 3866
B.BR.03.BREPM2012        ----------------------------A-----------C-----C------------...---------...-------------------A--T-----T-----------A----G-A---A-----------------A--GG----...--------- 4071
B.CA.97.CANB3FULL        -------------------G--------------------C-----C--G---C-----...---------...-------------------A--------------------A------A---A-----------------A--G-----...--------- 3935
B.CN.05.05CNHB_hp3       -----A--------A----G--------A-----------C---A-CC-----------...---------...-----G-------------A----------A-----T--A--A--T-A---A-----------------A--G-----...--------- 4645
B.CO.01.PCM001           -------------------G-------C------------C---A-C------------...---------...-------------------A--T-----------------A-CA----A---A-----------------A--GG----...--A------ 3851
B.GB.83.CAM1             -------------------G--------A-----------C-----C------------...---------...-----------C-------A-------TG---------T-A--A---A---A-----------------A--G-----...--------- 4647
B.GB.86.GB8              -------------------G-----------C-----C--C---A-C------------...---------...-------------------A--T-----C-----------A--A--------A-----------------A--G-----...--------- 4666
B.GE.03.03GEMZ010        -------------------G--------A-----------C-----C-----C------...-----G---...---------GT--------A--------------------A--AA-G--C---AA-------------------G-----...--------- 3851
B.IT.05.SG1              -------------------G--------------------C-----C--G---A-----...---------...-------------------A-------------C------A--AG-------A-------------------G-----...--------- 4448
B.JP.05.DR6538           -------C--------------------------------C-----C------------...---------...------------G------A-------------C------A-CA--------A--T-------------A--G-----...--------C 4650
B.KR.05.05CSR3           -------C--------------------------------C---A-C------------...---------...-------------------A--------------------A--A---A---A--T-----------------GG----...--------- 4671
B.NL.00.671_00T36        ----------------------------A-----------C-----C------------...---------...----------C--------A-------CC---------T-A--A--G-A---A-------------------G-----...--------- 4211
B.RU.04.04RU128005       -------------------G-----A--------------C-----C--G---------...---------...-C-----------------A--------------------A--AA-----C-A-------------------G-----...--A------ 4137
B.TH.00.00TH_C3198       ----------C---A----G--------------------C-----CC-----------...---------...-----G------C-------A----------A--------A--AA-T-----A-----------------A--G-----...--------- 3860
B.UA.01.01UAKV167        -------C-----------------A--A-----------C-----C-----C------...-----G---...-----------T-G------A--------------------A-CAACG-----A-----------------A--G-----...--------- 3872
B.US.04.ES10_53          -------------------G--------------------C---G-C------------...---------...--------C----------A--------------------A--A--G-A---A-------------------G-----...--A------ 4655
B.US.99.PRB959_03        -------------------G--------------------C---A-C------------...---------...--------G---G------A-------CC-----------A----G-----A-----T-----------A--G-----...--------- 3866
C.AR.01.ARG4006          ----G-----C--------G-----A--A-----T-----C---A-AC-G---------...---------...------C--GGT--------A----------T---------A--AA-G-A---AA--A-------------A--T---C--...--------- 3830
C.BR.04.04BR013          ----G-----C--------------A--A-----------C---A-AC-----------...---------...------C--GT---------A----------T---------A--AA---A---AA--A-------------A------C-A...--------- 4110
C.BW.00.00BW07621        ----G-----C--------------A--A-------------ACA-AC-----------...---------...------C---T---------A----------T---------A----G-A---A--A-------------A--T---C--...--------- 3999
C.CN.98.YNRL9840         ----G-----C--------------A--A--G--------C---A-AC-----------...---------...------C--GT---------A-------T--T---------A----G-A---A--A-------------A--T---C-A...--------- 3830
C.ET.02.02ET_288         ----G--------------------A--------------C-A-A-AC-----------...---------...------C--GGT--------A--T-------T--C------A--AG-G-A---AA--A-------------A--T--T-CA...--------- 3848
C.GE.03.03GEMZ033        -------------------------A--A-----------C-A-A-AC-----------...---------...------C--GT---------A----------T---------A--AG-G-A---A--A-------------A--T---C-A...--------- 3827
C.IL.99.99ET7            ----G-----C--------------A--A--G--------C-A-A-AC-G---------...---------...------C--GT---------A----------T---------A----G-A---A--A-------------A--T---C-A...--------- 3827
C.IN.99.01IN565_10       ----G-----C--------------A--A-----------C-A-A-AC-----------...---------...------C--GT---------A-------T--T---------A----G-A---A--A-------------A--T---C-A...--------- 4027
C.KE.00.KER2010          ----G-----C--------------A--A-----------C-ACA-AC-----------...---------...------T--GT---------A----------T---------A----G-A---A--A-------------A--T---C-A...--------- 3827
C.MM.99.mIDU101_3        ----G-----C--------------A--A--G--------C---A-AC-----------...---------...------C--GT---------A-------T--T---------A----G-A-C--A--A-------------A--T---C-C...--------- 3997
C.MW.93.93MW_965         ----------C--------------A--A-----------C-A-A-A------------...---------...------C--GT---------A----------T---------A----G-A---A--A-------------A--T---C-A...--------- 3833
C.SN.90.90SE_364         ----G-----C--------------A--A-----------C---A-AC-----------...---------...------C--GT---------A----------T---------A----G-A---A--A-------------A--T---C-A...--------- 3815
C.SO.89.89SM_145         ----G--------------------A--A--G--------C-A-A-AC-----------...---------...------T--GT---------A----------T---------A--AA-G-A---A--A-------------A--T---C-C...--------- 3867
C.TZ.02.CO178            ----G-----C--------------A--A-----------C-A-A-AC-----------...---------...------C--GT---------A---------------T---A-CA--G-A---A--A-------------A--T---C-A...--------- 3827
C.UY.01.TRA3011          ----G-----C-----------G--A-----------------A-AC---G--------...---------...------CC-GT---------A----C--CCT-------TA--AA---A---A--A-------------A--T---C-A...--------- 3818
C.YE.02.02YE511          ----G-----C--------------A--A-----------C-A-A-AC-G---------...--G------...------C--GT---------A----------T---------A----G-A-------A-------------A--T---C-A...--------- 3821
C.ZA.04.04ZASK164B1      ----G-----C--------------A--A-----------C-A-A-AC-----------...---------...------C--GT---------A----------T---------A--A--G-A---A--A----C--------A--T---C-A...--------- 4060
C.ZA.05.05ZAPSK240B1     ----G-----C--------------A--A-----------C-A-A-AC-----------...---------...------C--GT-C-------A----------T--------TA----G-A---A--A-------------A--T----A...--------- 4050
C.ZM.02.02ZM108          -------------------------A--A-----------------A-AC---------...---------...------C--GT---------A-------T-CT---------A-------A---A--A-------------A--T---C-A...--------- 4615
D.CD.83.ELI              -----------------------------------------------------------...---------...---------GT-G-------A--------------------A----G-A---A-----------------A--T----A...--------- 4192
D.CM.01.01CM_0009BBY     ----------------------------A-----------C-A--C-------------...---------...---------GT-G-------A----------T-----T---A----G-T---A-----------------A--T-----...--------C 3845
D.KE.01.NKU3006          -------------------------------------C--C--C--C--G----G----...--G------...---------GT-G-------A-------------C------A-C---G-A---A-----------------A--C--T---...--------- 3851
D.KR.04.04KBH8           -------------------------A--------------C-----C------C-----...---------...---------GTG--------A--------------------A----GTA---A-----------------A--C-----...--------- 4605
D.TD.99.MN011            ----------------------------------------C------------------...---------...----------GG-G------A----------A-----T--A------T---A-----------------A--T--T---...--------- 3870
D.TZ.01.A280             -------------------------------------C--C--C---A-C---------...---------...---------GT--------A-------TC-----------A------T---AA--T-------------A--C-----...--------- 3847
D.UG.99.99UGK09259       --------------------------------GG------C-----C--G--------C...---------...--------G-GT--------A--------------------A-CAA-G-AC---A-----T-----------A--C-----...--------- 3851
D.YE.01.01YE386          ----------C-----------------A-------T------------G---------...---------...--------GGT---------A-------TG---------T-A-C---G-----A-----------------A--T-----...--------- 3845
D.YE.02.02YE516          ----------------------------------------C------------------...---------...---------GT-G-G------A------------------T--A---G-G-T---A--A----C--------A--T--T---...--------C 3857
D.ZA.90.R1               ----------------------------------------C-----C------------...---------...---------GT-G-------A--------------------A----G-A---A--------C--------A--T----A...--------- 3992
F1.AR.02.ARE933          ----------C--------------A--A--------C--C--C--CA-AC---G----...---------...-----C-------------A-------TCC-----------A---CG-----AA--AT------------A--T--TC--...--------- 3940
F1.BE.93.VI850           -------------------------A-----------C--C--C--CA-A----G----...---------...-----------T-------A--------------------A---CG-----A--AT------------A--T---C-A...--------- 3986
F1.BR.01.01BR125         -------A--C--------------A--A--C-----C--C--C--CA-AC---G----...---------...-----------T-------A--T-----CC-----------A---CG-----A--A-------------A--T---C--...--------- 4091
F1.ES.x.P1146            -------------------------A-----------C--C--C--CA-AC---G----...---------...--T------G---------A----------T---------A--AG-G-A---A--A-------------A-AT---C--...--A------ 3923
F1.FI.93.FIN9363         ----------C--------------A-----C-----C--C--C--CA-AC---G----...---------...-----------TG------A--------------------A---CG-----A--A-------------A--T---C--...--------- 3975
F2.CM.02.02CM_0016BBY    ----------------------G--A-----------C-----C--CA-A----G----...---------...-----------T-------A--------------------A--A--GT----A--A-------------A--T---C-A...--------- 3836
F2.CM.95.MP257           -------------------------A-----G-----C--C--C--CA-AC---G----...---------...-----------T-------A-----------------T--A----GT----A--A-------------A--T---C--...--------- 3855
F2.CM.97.CM53657         ----------C-----------G--A-----------C--C--C--C--AC---G----...---------...-----------T-------A--T-----------------A--A--GT----A--A-------------A--T---C--...--------- 3836
G.BE.96.DRCBL            ----------C--------------A--------------C---A-A------------...-----G---...-----------T-------A-----------------T--A----G-A--A-A--A--A----------AA-C----CA...--------- 4603
G.CM.01.01CM_4049HAN     ----------C--------------A---------G---C---A-A-------------...---------...---------GGT-----C--A-------------------TA----G-A--A-A---A-------------AAAT----CA...--------- 3848
G.CU.99.Cu74             -------------------------A--------------C---A-AC-----------...---------...--------G--GTG-------A-------------------A----G-A--A-A--A--A----------AAAT----CA...--A------ 4250
G.ES.99.X138             ----------C--------------A--------------C---A-AC------G----...-----G---...--------G---T-------A--T-----------------TA--A-G-A--A-A--A--A----------AAAT----CA...--A------ 4087
G.GH.03.03GH175G         ----------C--------------A-----------C--C--C---A-A---------...-----G---...-----------GT-C------A----------T---------A-C---G-A--A-A--A--A----------AAAT----CA...--------- 4686
G.KE.93.HH8793_12_1      ----------C--------------A--------------C--CA-AC-----------...-----G---...--------G--GT--------A--------------------A-----G-A--A-A--A--G---C-------A-AT--T-CA...--A------ 4046
G.NG.01.01NGPL0669       ----G-----C--------------A--------------C---A-A------------...-----G---...-----------GT--------A------------------AA--AG-G-A--A-A--A--T----------AA-T----CA...--A------ 3851
G.PT.x.PT2695            ----------C--------------A--------------C---A-AC------G----...-----G---...--------G-G--T-------A--------------------A-----G-A--A-A--A--A----------A-AT----CA...--A------ 4588
G.SE.93.SE6165           -------------------------A--------------C---A-A------------...-----G---...----------C-GT-------A--T-----------------A-----G-A--A-A--A--A----------AAAT----CA...--------- 4043
H.BE.93.VI991            ----------C--------------A--------------C---A-A--G--C------...-C-------...-----------TG-------A-------------C----AA-----G-----A--A-------------A-AT---C-C...--------- 4035
H.BE.93.VI997            -------C--C--------------A------------------A-A--G--C------...-C-------...-----------TG-------A----------C--------AA--A--G-----A--A-------------A-AT---C-A...-----C--- 3970
H.CF.90.056              -------C--C--------------AA-------------C--C--G--G--C------...A-C------...---------GT--------A-------------------GA-----G-----A--A-------------A-AT---C-A...--------- 3993
J.CD.97.J_97DC_KTB147    ----------C--------------A-----------T----A-A-------G------...---------...--------G-GT--------A-------------C------A----G-T---A--A-------------A--GG----...--------- 3840
J.SE.93.SE7887           ----------C--------------A-----G-----T----A-A--------------...--C------...---------GT--------A-------------C------A--G-G-T--G-A--A-------------A-AT-----...--------- 3960
K.CD.97.EQTB11C          ----------C-----------G--A-----------C--C--CA-AC-----------...---------...--------G-GT--------A--------------------A----GTA---A--A-------------A-AT-----...--------- 3843
K.CM.96.MP535            ----------C-----------G--A--A--------C--C---A-AC-----------...---------...---------GT--------A---------C----------A-CA--GT----A--A-------------A--TG----...--------- 3842
```

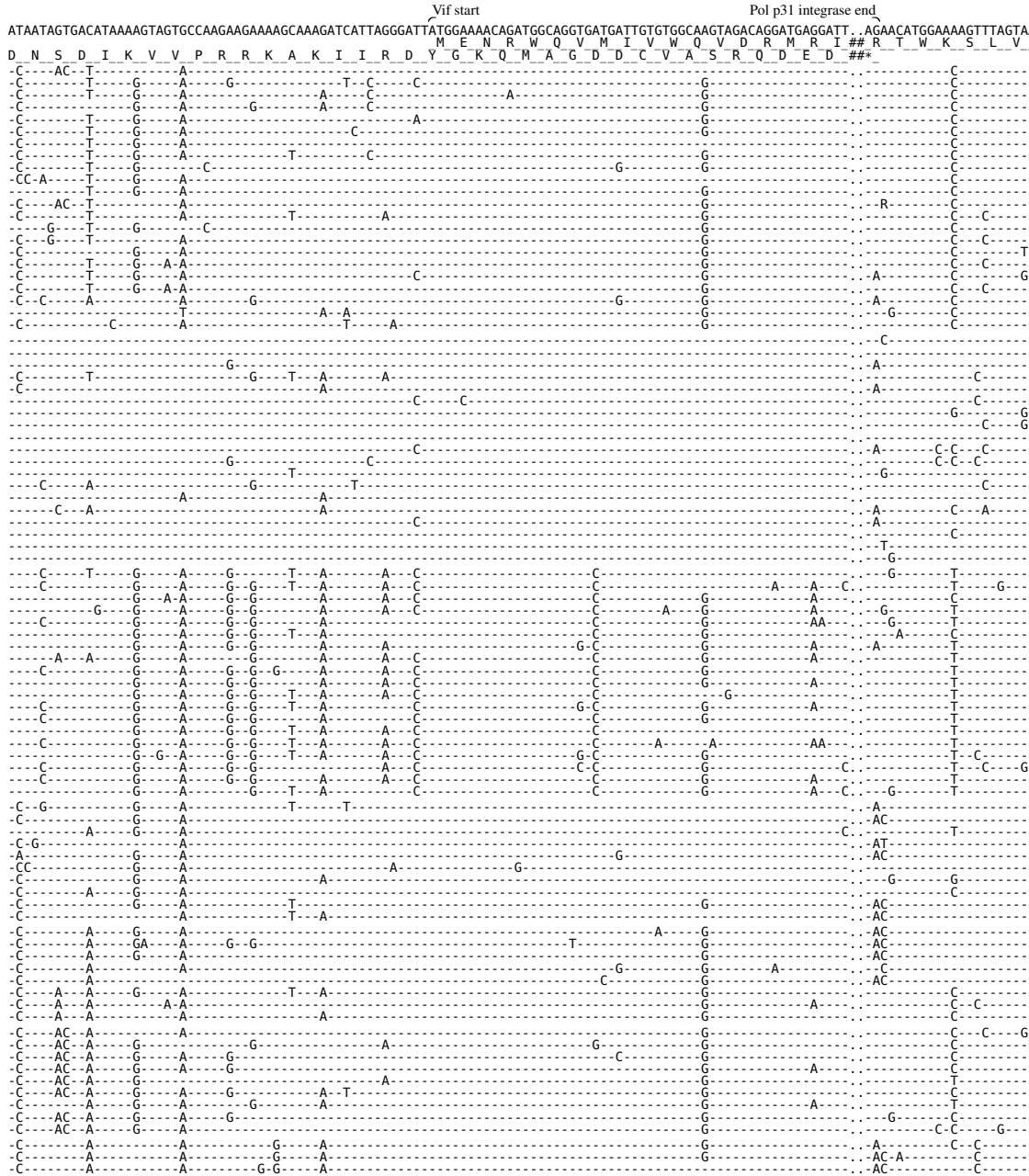
```
B.FR.83.HXB2                 CAGAAGTTATTCCAGCAGAAACAGGGCAGGAAACAGCATATTTTCTTTTAAAATTAGCA...GGAAGATGG...CCAGTAAAAACAATACATACTGACAATGGCAGCAATTTCACCGGTGCTACGGTTAGGGCCGCCTGTTGGTGGGCGGGAATCAAG...CAGGAATTT 4646
Pol                          A__E__V__I__P__A__E__T__G__Q__E__T__A__Y__F__L__L__K__L__A__.__G__R__W__.__P__V__K__T__I__H__T__D__N__G__S__N__F__T__G__A__T__V__R__A__A__C__W__W__A__G__I__K__.__Q__E__F_
01_AE.CF.90.90CF11697        ----------C-------------A----G--------C-----GC-----------...-------A-...---------GGTC-----C--A--------T-----------A-C---G-TA-G-AA--A-------------CAATG--C-A...--------- 4595
01_AE.CN.05.FJ051            ----G-----C-------------A----G--------C-----GC-----------...---------...----------GT------C--A-----C--G-----------A-C---G-A---AA--A-------------C-ATG--CGA...--------- 4656
01_AE.CN.06.FJ054            ----------C-------------A----G--------C-----GC-----------...---------...-----C----GT------C--A--------T-----------A-C---G-A---AA--G-------------CAAC---CGA...--A------ 4671
01_AE.HK.x.HK001             ----------C-------------A----G--------C--C--GC-----------...---------...----------GTC-----C--A--------T-----------A-C---G-A---AA--A-------------CAATG--CGA...--------- 4008
01_AE.TH.01.01TH_R2184       ------------------------A----G--------C--C--GC-----------...---------...----------GT------C--A--T-----T-----------A-C---G-A---AA--A-------------CAATG--CGA...--------- 3846
01_AE.TH.02.OUR769I          ----------C-------------A----G--------C-----GC-----------...---------...----------GT------C--A--------T-----------A-C---G-A---AA--A-------------CAATG---A...--------- 3845
01_AE.TH.90.CM240            ----------C-------------A----G--------C-----GC-----C-----...---------...----------GT------C--A-----C--T-----------A-C---G-A---AA--A-------------CAATG--C-A...--------- 4220
01_AE.US.00.00US_MSC1164     ----------C-------------A--A-G--------C-----GC-----------...---------...----------GT------C--A-----C--T-----------A-C---G-A---AA--A-------------CAATG---A...--------- 3851
02_AG.CM.02.02CM_4082STN     ----------C-------------A----G--------C---A-A------------...---------...------GC--GT------C--A-----------------T-A-----G-A--A-A---A--A-----------A-AT-----A...--A------ 3845
02_AG.EC.x.ECU41             ----------C------------------G--------C---A-A------------...---------...-----------T------C--A-------------------A-----G-A--A-A---A-------------AAATG---CA...--A------ 3742
02_AG.GH.03.GHNJ196          ----------C-------------A-------------C---A-A----G-------...---------...------G-G-GT------C--A-------------------A-----G-A--A-A---A--A-----------A-ATG----A...--A------ 4675
02_AG.NG.01.PL0710           -------G--C---T-C-------A----G--------C---A-A------------...---------...------G-G-GT------C--A-------------------A-----G-A--A-A---A--A-----------AAAT----CA...--A------ 3830
02_AG.NG.x.IBNG              ----------C-------------A----G--------C---A-A------------...---------...------G---GT------C--A-------------------A-----G-A--A-A---G--A-----------AAATG---CA...--A------ 4171
02_AG.SE.94.SE7812           ----------C-------------A--A-G--------C---A-A------------...---------...------G---GT------C--A-------------------A-----G-A--A-A---A--A-----------AAATG---CA...--A------ 4018
02_AG.SN.98.MP1211           ----------C-------------A----G-----C--C---A-A-----------GGGAAG-TGGCCAGT......---GT-G---C--A-------------------A-----G-A--A-A---A--A-----------CAA-G-T-CACCAGGAATTGGG 3846
03_AB.RU.97.KAL153_2         ------------------------A-------------C---G-C------------...---------...-----------T--------A-------------------A--A--G------A---T-------------A--G------...--------- 3872
04_cpx.CY.94.CY032           ----------C-------------A----------C--C--CA-AC-----------...---------...------G----TG------G-A-----C---CC---------A-----G------A---A-------------A-AT-----C...--------- 4012
05_DF.BE.x.VI1310            ------------------G-----A-----C-----C--C--CA-A-----GC-G---...---------...-----------TG--------A--T------CC---------A---G-G------A---A-------------A--T---C--...--A------ 4028
06_cpx.AU.96.BFP90           ----------C-------------A----G--------C---A-A------------...---------...----------GTG-----C--A--------T-----------A-----G-A---AA--A-------------AAAT----CA...--A------ 4674
08_BC.CN.97.97CNGX_6F        ----G-----C-------------A--A----------C---A-AC-----------...---------...------C---GT---------A--------T--T---------A-----G-A---A---A-------------A--T---C-A...--------- 3826
09_cpx.GH.96.96GH2911        ----------C-------------A--A----------C---A-A------------...---------...----------GTG-----C--G-----------T---------A-----G-A--A-A---A--A-----------AAATG---CA...--A------ 3848
10_CD.TZ.96.96TZ_BF061       -------------------------------------C-----C--C--G--------...---------...----------GT-G--------A-------------------A-C---G-A---A-----------------A--C------...--------- 4025
11_cpx.GR.x.GR17             ----------C-------------A-------------C---A-A------------...---------...----------GT---------A--T-----------------A-CAA-G-A---AA--A-------------A--T---C-A...--------- 3948
12_BF.AR.99.ARMA159          ----------C-------------A--A--------C--C--CA-AC----G------...---------...-------------------------A---------CC--------T-A----C-T-----A---A-------------A--T---C-A...--------- 4651
13_cpx.CM.96.1849            ----------C-------------A-------------C---A-A------------...---------...----------GT------C--A-----------------T---A-----G-A--A-AA--A--A-----------AAAT----CA...--A------ 4051
14_BG.DE.01.9196_01          ----------C-------------A------G-------C---A-AC-------G---...-----G---...-----G----T---------A--T-----------------T-A--AA---T----AA--------C---------A--G------...--------- 4163
15_01B.TH.99.99TH_MU2079     ----------C-------------A-------------C-----GC-----------...---------...----------GT------C--A-----C--T-----------A-C---G-AA-G-AA--A-------------CAATG--CGA...--------- 4045
16_A2D.KR.97.97KR004         -------C--C-------------A----G-----T--C---A-A------C-----...---------...----------GT---------A-------GCC--------T-A----AC-A----A---A-----------------TG--C-A...--------- 4009
17_BF.PE.02.PE02_PCR0155     ----G-----C-------------A--A--------C--C--CA-AC----G------...---------...-------------------------A----------C----------A----CG------A---A-------------A--T---C-A...--------- 3836
18_cpx.CU.99.CU76            ----------C-------------A----------C--CC----AC-----------...---------...----------GT------C--A---------CC---------A-----G-A--A-A---A--T-----------AAAT--T-CA...--A------ 3951
19_cpx.CU.99.CU7             ----------C-------------A----G--------C-----GC-----------...-C-------...---------GGT------C--A--------TCC---------A-C---G-A---AA--A-------------AAATG--C-A...--------- 3872
20_BG.CU.99.Cu103            ----------C---T---------A-------------C---A-AC----------T...-----G---...-----G---GT---------A---------CC--------T-A-----G-A--A-A---A--A-----------AAAT----CA...--A------ 4112
21_A2D.KE.91.KNH1254         -------C--C-------------A--A-G--------C-GGA-A------C-----...---------...----------GT---------A-------GCC--------T-A--A-AC-A----A-----------------A--T-----A...--------- 3851
22_01A1.CM.01.01CM_0001BBY-  ----------C-------------A----G--------C---A-AC-----------...--G------...----------GT------C--A--------T-----C------A-C---G-A--G-AA--A-------------A-A----C-A...--------- 3854
23_BG.CU.03.CB118            ----------C-------------A-------------C---A-AC----------T...-----G---...-----G---GT---------A--Y-----CC--------T-A-----G-A--A-A---A--A-----------A-AT----CA...--A------ 4089
24_BG.CU.03.CB378            ----------C-------------A-------------C---A-AC----------T...-----G---...-----G---GT---------A---------CC--------T-A-----G-A--A-A---A--A-----------A-AT----CA...--A------ 4071
24_BG.CU.03.CB471            ----------C-------------A-------------C---A-AC----------T...-----G---...----------GT---------A---------CC---------A-----G-A--A-A---A--A-----------AA-T----CA...--A------ 4086
25_cpx.CM.01.101BA           ----------C-------------A--A-G--------C-----A------------...---------...----------GTG--------A---------CC--------T-A-----G------A---A-------------A--T------...--------- 3854
27_cpx.CD.97.97CDKTB49       ----------C---A---------A-------------C---A-A------C-----...---------...----------GT------C--A-----------A---------A--A--G-A--A-A---A--A--C--------AAATG-T-CA...--A------ 3854
28_BF.BR.99.BREPM12609       ------------------G--------A----------C-----C------------...---------...------G-----G-------A--T-------------------A--AA-G------A-------------------AG------...--------- 4012
29_BF.BR.02.BREPM119         ----------C-------------A--A-G----GT--C--CA-AC----G------...---------...-------G-------------A---------CC--------T-A--A--G------A-----------------A--G------...--------- 3820
31_BC.BR.02.110PA            ----G--C--C-------------A-------------C---A-AC-----------...---------...------C-G-GT---------A-----------T---------A--AA---A--A-AA--A-------------A--T---C-A...--------- 4080
32_06A1.EE.01.EE0369         ----------C-------------A----G--------C---A-AC-----------...---------...----------GT------C--A--------T-----------A-----G-A---AA--A-------------AAAT--T-CA...--A------ 4281
33_01B.MY.05.05MYKL007_1     ----------C-------------A----G-T------C---T-GC-----C-----...--G------...----------GT------C--A-----C--T-----------A-C---G------AA--A-------------CAATG--C--...--------- 4025
34_01B.TH.99.OUR2478P        ----------C-------------A----G--------C-----GC-----------...---------...----------GT------C--A-----C--T-----------A-C---G-A---AA--A--A-----------CAATG--CGA...--------- 3848
35_AD.AF.05.05AF095          -------G--C------------------G--------C-----GC-----------...---------...----------GT-G---C--A-------------------A-C---G-AT---AA--A-------------AA-T---C-A...--------- 3839
36_cpx.CM.00.00CMNYU830      ----------CT------------A----G--------C-----A------------...---C------...----------GT------C--A--------T-----C------A-C---G-A--G-AA--A-------------A-AT---C-A...--------- 3851
37_cpx.CM.00.00CMNYU926      ----------C-------------A-------------C-----GC---GC------...---------...--------GGGT------C--A--------T-----------A-C---G-A---AA--A-------------AAATG---A...--------- 3836
38_BF1.UY.05.UY05_4752       ----------C-------------A--A-----T--C--C--CA-AC----G------...---------...--------G------------A---------CC---------A----CG------A---G-------------A--T---C--...--------- 4057
39_BF.BR.03.03BRRJ103        ------------------G-----A--A----------C---A-C------------...---------...----------------------A-------------------A--A----A---A-------------------G------...--------- 4135
40_BF.BR.04.04BRRJ115        ------------------G-----A--A----------C------------------...-----G---...-----------T-G----------------CC---------A--AAC--A--C-A---A-------------A--T---C--...--------- 4198
42_BF.LU.03.luBF_05_03       ----------C-------------A--A--------C--C--CA-AC-----------...---------...-------------T-------A---------R-T-------Y-A--A--GT-----A---------C-----------G------...--------- 4184
43_02G.SA.03.J11223          ----G--C--C-------------A-----C-------C---A-A------------...---------...----------GT------C--A---------------C------A-----G-A--A-----A--A-----------AA-TG---CA...--A------ 4171
U.CD.83.83CD003              ----------C--G----------A-------------C---G-AC-----C-----...---------...--------C-GT---------A-------------------A-----G-A--C-AA--A-----C---------A--G---C-A...--------- 4141
U.CD.90.90CD121E12           ----------C-------------A-----C-------C---G-AC-----C-----...---------...--------C---T--------A--T----------------A-----G-A--C-AA--A-----C---------A--G---C--...--------- 3845
U.GR.99.GR303                ----------C-------------A--A----------C-----AC---G-------...---------...----------GT------C--A------------------TA-----G-TA-G-A---A--A-----------AAATG---CA...--A------ 3917
U.NL.95.U_NL_95_H10986_D1    -------C--C--------------------------C-A-A-C-------------...-G-------...--G-----GTG---------A---------CCT---------A-----G---C-A---A-------------A--T------...--------- 4531
CPZ.CD.90.ANT                -T-----A--GG---AT--G------A-AAGT--------C--C--G-----------C...A-C------...-----------------C-----------AGCT--------AA----AG-A--A-AA--G--A-----------TAAT---C-A...--------- 4085
CPZ.CM.01.SIVcpzCAM13        ----------C-------------A--A-----T-----C--CA-A-------------...A-C------...------G---GTCT---C--A--T----G-CA----T--AA--AA-G-A---AA--A--A-----------CAAC-----A...--AA-G--- 4226
CPZ.CM.05.SIVcpzEK505        -------C-----T-----------A--------T---------A-CC----G-----T...---------...----------GTC----------T----G----C-----A---C--T--A-AA--A-------------AAAT---C-A...-----G--- 4218
CPZ.CM.05.SIVcpzLB7          -------C--A--T-----------A-----------T------A-C------------...TCG------...-----------TT----------T----T-------T--TA-CA--G-A--C-A---T--A-----------CAG---C--...--------- 4201
CPZ.CM.05.SIVcpzMB66         -------A--A--T--T--------A--A-----------C---A-C--------G---...--C------...----------GTC----------T----A--T--C--T---A--AG---T--C-A---A--T-------------T---C-A...--A------ 4192
CPZ.CM.05.SIVcpzMT145        -------C--C--T--T--------A--------------G-CC-T--G---------...---------...--------G--T-----C--A-------GCCT--------TA--AA-G-A--A-A---A--A-----------A--T--T-GA...--------- 4188
CPZ.GA.88.GAB1               ----------C-----T--G-----A--A--------T-----C--G-------------...--T------...------------T--T--C--A--T----GCCA-----T--AA-----G-A--C-A---T-------------A-AC------...--------- 4707
CPZ.TZ.00.TAN1               -------C--A----A----------AAGA-----C--T--C--CA-CC-------G---...---------...--T-----G-A---------A--T----ACCA-----T--TA--A-AG-A--G-A---A-----C--------ACA---TC-A...--T------ 4290
CPZ.US.85.CPZUS              ----G--C-----T-----------A--A-G-----C---------A-C-----------...---------...--T-----GTG--T--------T----ATCT--C--T--AA--AG---A-----A--A--T--C--------A--C--AC-A...--A--G--- 4706
CPZ.CM.x.SIVcpzDP943         ----G--C--C--T-------T--A--------T-------A--C-----------...---------...----------GT---T--C-----T----ATCT-----T---A--AA---A---A---A-------------A--C---C-A...--------- 4232
CPZ.CM.x.SIVcpzMB897         -------C--A--T--T------------A--G-----T------G-GC-G--------...---------...--------GGTT-----C-----T-----------C--T--TA-----G-T--C-A---T--A-----------AAAT---C-A...--------- 4212
N.CM.02.DJO0131              -------------T----------A--------------------A-----------T...---------...----------GTT-----C-----T----ACC---------TA----C--T--A-AA--A-------------AAAT----CA...---------C 4152
N.CM.04.04CM_1015_04         -------------T----------A--K-----------------A------G-----T...---------...----------GTT-----C-----T----ACC---------TA----C--T--A-AA--A-------------AA-T-----A...--------- 4151
N.CM.04.04CM_1131_03         -------------T----------A--------------------A------G-----T...---------...----------GTT-----C-----T----ACC---------TA----C--T--A-AA--A-------------AA-T-----A...---------C 4150
N.CM.95.YBF30                -------------T-----G-----A--------------------A------G-----T...---------...----------GTT-----C-----T----ATC---------TA----C--T--A-AA--A-------------AAAT-----A...--------- 4240
N.CM.97.YBF106               -------------T----------A--------------------A------G-----T...---------...----------GTT-----C-----T----ACC------T-TTA----C--T--A-AA--A-------------A--T-----A...---------C 4238
O.BE.87.ANT70                -------G--A-------------A--A-----T--C--C--C--G------C-G--T...-C-------...--T--T---GT---------A-----C--GCCT-----T--AA--A-A--TA-G-A---T--A-----------CAAC--AC-A...--T--G--- 4700
O.CM.91.MVP5180              ----G--G--A-------------A--------T--C-----C--G-------------...-C-------...--T--C---GT---------A--------ACCT-----T--AA----AG-CA-G-AA--T--A---------A-A--C--AC-A...--T--G--- 4675
O.CM.96.96CMABB637           -------G--A-------------A--A-----T--C-----C--G-------------...-C-------...--TA-T-G-GT-C-------A--------GCCT-----T--AA----A--CA-G-A---T--A-----------CAAC--A-GA...--T--G--- 4130
O.CM.98.98CMA104             -------A--A-------------A--A-----T--C-----C--G------C------...-C-------...--T--T---GT---------A--------GCCT--------AA----A--CA-G-A---T--A-----------CAAC--AW-A...--T--G--- 4132
O.CM.99.99CMU4122            -------G--A-------------A--A-----T--C--C--C--G-----------T...-C-------...--T--T---GT---------A-----C--GCCT--------AA----A--CA-G-A---T--A---------A-CAAC--A--A...--T--G--- 4131
O.FR.92.VAU                  -------C--A--------G-----A--------T--C--C--C--G-------------...-C-------...--TA-T----T-C-------A--------GCCT--C--T--AA----AG-CA-G-AA--T--A-----------AAAC--AC-A...--T--G--- 4212
O.SN.99.SEMP1299             -------A--A-------------A--------T--C--C--C--G-----------T...-C-------...--T--T--GGT---------A--------GCCT-----T--AA----AG-CA-G-A---T--A-----------CAAC--A--A...--T--G--- 4699
O.US.99.99USTWLA             -------G--A-------------A--A-----T--C-----C--G-------------...-C-------...--T--T----TC--------A--T---A-GCCT-----T--AA----A--CA-G-AA--T--A---------A-A--C--AC-A...--T--G--- 4129
```

```
B.FR.83.HXB2              GGAATTCCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAAAGAATTAAAGAAAATTATAGGACAGGTAAGAGATCAGGCTGAACATCTTAAGACAGCAGTACAAATGGCAGTATTCATCCACAATTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGC 4816
Pol                       _G__I__P__Y__N__P__Q__S__Q__G__V__V__E__S__M__N__K__E__L__K__K__I__I__G__Q__V__R__D__Q__A__E__H__L__K__T__A__V__Q__M__A__V__F__I__H__N__F__K__R__K__G__G__I__G__G__Y__S__A
A1.GE.99.99GEMZ011        --G--------------------------------G--G-----------G--------------C-----G-----------G--A--------C--------------------------------T----------------------------------------- 4000
A1.KE.00.KER2008          -----------------------------------G--------------G--------------C-----G--------G--C--A--------C--C-----------------------------T----------------------------------------- 4015
A1.KE.00.KNH1144          --------------C-----------------------------------G--------------C-----------------G--A--------C-----A--------------------------T----------------------------------------- 4021
A1.KE.00.KSM4024          --G--------------------------------G--G-----------GC--------C----C--------A--------G--A--------C-----A--------------------------T----------------------------------------- 4012
A1.KE.00.MSA4069          --G--------------------------------G--G-----------G--------------C--------------------A--A-----C-----A--------------------C-----T----------------------------------------- 4015
A1.KE.00.NKU3005          --------------------------------------------------G--------------C-----G-----------G--A--------C-----------G--------------G-----T----------------------------------------- 4015
A1.RU.00.RU00051          --G--------------------C--------G--G--------------G--------------C-----G-----------G--A--------C--------------------------------T----------------------------------------- 4117
A1.RW.93.93RW_024         --G--C-----------------------G--------------------G--------------C-----------------G--A--------C-----A--------------------------T----------------------------------------- 4018
A1.SE.95.SE8891           --G--C-----------------------------G--------------G--------------C-----G-----------A--A--------C--------------------------------T----------------------------------------- 4005
A1.SE.95.UGSE8131         -----------------T-----------------G-----------------------------C-----G-----------G--A--------C----G---------------------------T----------------------------------------- 4188
A1.TZ.01.A173             --G-----A--------------------------G--------------G--------------C--------A--------G--A--------C-----A--------------------------T----------------------------------------- 4015
A1.UA.01.01UADN139        --G--------------------------------G--G--C--------G--------------C-----G-----------G--A--------C--------------------------------T----------------------------------------- 4015
A1.UG.92.92UG037          --T-----------------------------------------------G--------------C-----G--------G--G--A-----------------------------------G-----T----------------------------------------- 4832
A1.UG.99.99UGA07072       --G--------------------------------G--------------G--------------C-----G-----------G--A--------C--------------------------------T--------------------------------A-------- 4015
A2.CD.97.97CDKTB48        --G--------------------------------G-----------------------------C-----G--------------A-----------------------G-----------G--------------------------------------A-------- 4156
A2.CY.94.94CY017_41       --G--------------------------------G-----------------------------C-----G--------------A--------C--------------------------------T--------------------------------A-------- 4173
A.SN.01.DDI579            -----------------------------------G--------------G--------------C-----G-----------G--A--------C--------------------------------T----------------------------------------- 4014
A.SN.01.DDJ369            -----------------------------------G--------------G--------------C-----G-----------G--A--------C--------------------------------T----------------------------------------- 4017
A.SN.96.DDJ360            -----------------------------------G--------------G--------------C-----G-----------G--A--------C--------------------------------T----------------------------------------- 4014
A.CD.02.02CD_KTB035       --G--------------------------------G-----------------------------C--------------------A--A-----C-----A-------------------------------------------------------------------- 4810
A.CD.97.97CD_KCC2         --G--------------------------------G-----------------------------C-----G--------G--G--A--G-----------A--------------------------T----------------------------------------- 4812
A.CD.97.97CD_KTB13        --G--------------------------------G-----------------------------C-----G--------------A--G-----C-----A--------------------------T----------------------------------------- 4020
B.AR.04.04AR151516        --C----------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4038
B.AU.87.MBC925            --C-----------------------------------------A----------------------------------------------------------------------------------------------------------------------------- 4827
B.BO.99.BOL0122           --C-----------------------------------------------T-----------------------A--------------------------A-------------------------------------------------------------------- 4036
B.BR.03.BREPM2012         --C-----------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 4241
B.CA.97.CANB3FULL         --C-----------------------------------------------------------------------A-----------------------------------------------------------------G--------------------A-------- 4105
B.CN.05.05CNHB_hp3        --C-----------------------------------------------T--G--------------------------------A--------------------------------------T--------------G----------------------------- 4815
B.CO.01.PCM001            --C-----------------------------------------------T--------------------G--------------------------------------------------G-----T------------------------------A---------- 4021
B.GB.83.CAM1              --C-----------------------------------------------T-----------------------------------A----------------------------------------------------------------------------------- 4817
B.GB.86.GB8               --C--------------------------------------------------------------------------------------------------A-----------------------------------------------------------A-------- 4836
B.GE.03.03GEMZ010         --T--------------------------G-----------C----------------------------C------------------------------A-------------------------------------------------------------------- 4021
B.IT.05.SG1               --C----------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4618
B.JP.05.DR6538            --C--------------------------------------------------------------------------------------------------------------------------------------------------------------A-------- 4820
B.KR.05.05CSR3            --C--C----------------------------------------------------------------C---A-----------------------------------------------------T----------------------------------------- 4841
B.NL.00.671_00T36         --C--------------------------------------------------------------------------------G--------------C--------------------------T----------------------------------A---T----- 4381
B.RU.04.04RU128005        --C--------------------------------------------C-----------------------------------------------------------------------------------------------------------------A-------- 4307
B.TH.00.00TH_C3198        --C-----------------------------------------------------------------------------------A-------------------------------------------------A--------------------------------- 4030
B.UA.01.01UAKV167         --T--------------------------------------------------------------------------------------------------------------------------T-------------------------------------------- 4042
B.US.04.ES10_53           --C--C-----------------------------------------------------------------------------G-------------------------------------------------------------------------------------- 4825
B.US.99.PRB959_03         --C----------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4036
C.AR.01.ARG4006           -----C--------------------G--------------C-----C-----------------C-----G--------------A-----G--C--------------------------------T----------------------------------------- 4000
C.BR.04.04BR013           --------------------------G--------------C--------G---C----------C--------------------A-----G--C-----------------------G--------T----------------------------------------- 4280
C.BW.00.00BW07621         --------------------------G--------------C-----C-----------------C--------------------A-----G--C--------G-----------------------T----------------G------------------------ 4169
C.CN.98.YNRL9840          --------------------------G--------------C-----------------------------G--------------A-----G--C--------------------------------T----------------------------------------- 4000
C.ET.02.02ET_288          --G--------------------C-----------------C-----------------------C----T------------C--A-----G--C--------------------------------T----------------G--------------AC---T---- 4018
C.GE.03.03GEMZ033         --------------------------G--------------C-----------------------C-----G--A-----------A-----G--C--------------------------------T----------------------------------------- 3997
C.IL.99.99ET7             --------------------------G--------------C-----------G-----------C-----G-----------A--A-----G--C--------------------------------T----------------------------------------- 3997
C.IN.99.01IN565_10        --------------------------G--------------C-----------------------------G--------------A-----G--C--------------------------------T----------------------------------------- 4197
C.KE.00.KER2010           --------------------------G--------------C--------G--------------C--------------------A--A--G--C--------------------------------T----------------------------------------- 3997
C.MM.99.mIDU101_3         --------------------------G--------------C-------------------G---C-----G--------------A-----G--C--------------------------------T----------------------------------------- 4167
C.MW.93.93MW_965          -----------------------------------------C-----------------------C-----G--A-----------A-----G--C---------------------------C----T----------------------------------------- 4003
C.SN.90.90SE_364          --------------------------G--------------C-----------------------C-----G-----------A--A-----G--C--------------G-----------------T----------------G------------A----------- 3985
C.SO.89.89SM_145          --------------------------G-----------G--C-----------------------C--------------------A-----G--C--------------------------------T----------------G------------------------ 4037
C.TZ.02.CO178             --------------------------G--------------C-----------------------C-----G-----------G--A-----G--C--------------------------------T----------------G------------------------ 3997
C.UY.01.TRA3011           --------------------------G--------------A-----------------------C--------------------A-----G--C--------------------------------T----------------------------------------- 3988
C.YE.02.02YE511           --------------------------G--------------C--------------------G--C-----------------A--A-----G--C--------------------------------T----------------------------------------- 3991
C.ZA.04.04ZASK164B1       --------------------------G--------------C-----------------------C-----G-----------G--A--A--G--C--------------------------------T----------------------------------------- 4230
C.ZA.05.05ZAPSK240B1      -----C-----------T--------G-----T--------C-----------------------C-----G--------------A-----G--C--------------------------------T----------------------------------------- 4220
C.ZM.02.02ZM108           --------------------------G--------------C-----------------------C-----G--------------A-----G--C--------------------------------T----------------------------------------- 4785
D.CD.83.ELI               --------------------------------------------------------------------------------------A----------------------------------------------------------G-A-------------A-------- 4362
D.CM.01.01CM_0009BBY      -----------------------C--------------------------------------------------------------A--------------A--------------------------T----------------------------------------- 4015
D.KE.01.NKU3006           --------------------------------------------------------------------------A-----------A----------------------------------------------------------------------------------- 4021
D.KR.04.04KBH8            -----------------------------------------------------C--------------T--------------G--------G----------------------------------------------------------------------------- 4775
D.TD.99.MN011             ----------------------------------------------------------------------A---------------A-----------------------------------------T----------------------------------------- 4040
D.TZ.01.A280              -------------------------------------------------G---G--------------------------------A--------C-----------------------------T-------------------------------------------- 4017
D.UG.99.99UGK09259        -----------T-----------------T--G------------------------------------AA---------------A--------------------------------------------------------------------------A-------- 4021
D.YE.01.01YE386           --------------------------------------------------------------------------AA----------A-----------------------------------------T-----------------------------A-------R--- 4015
D.YE.02.02YE516           --------T--------------------------------------------------------------G--A-----------A-----------------------------------------T----------------------------------------- 4027
D.ZA.90.R1                -----------------------------------G--------------------------------------------------A---------------------T----------------------------------------------------A-------- 4162
F1.AR.02.ARE933           --------------C-----------------------G--------------GC-------G--C--------------------A---------------------------------------------------------------------------A-------- 4110
F1.BE.93.VI850            -----C--------------------------------------A--------G-----A--G--C--------------------A----------------------------------------------------------------------------------- 4156
F1.BR.01.01BR125          --------------C-----------------G--------------------GC----A--G--C---AA---A-----------A---------------------------------------------------------------------------A-------- 4261
F1.ES.x.P1146             ---G-------------------C-----------------------------------------C--------------------A-----------------------------------------T-----------G-------------------A--------- 4093
F1.FI.93.FIN9363          -----------T--C--------------------------A-----------GC-------G--C--------------------A----------------------------------------------------------------------------------- 4145
F2.CM.02.02CM_0016BBY     --------------------------------------------------------------------------A-----------A-----------------------G-----------------T----------------------------------------- 4006
F2.CM.95.MP257            -----------------------------------------------------------------C--------A-----------A-----------------------G-----------------T----------------------------------------- 4025
F2.CM.97.CM53657          -----------------T--------------------------------------------------------------------A-----------------------------------------T----------------------------------------- 4006
G.BE.96.DRCBL             -----------------------C-----------G--------------G--------------C-----------C--G-----A--------------------------G--------------T----------------------------------------- 4773
G.CM.01.01CM_4049HAN      -----------------------C-----------G--------------G--------------C--C-CG-----T--G-----A-----------C--------------G--------------T----------------------------------------- 4018
G.CU.99.Cu74              -----C-----------------C-----------G-----------------------------C--T--G-----C--G-----A--------C-----------------G--------------T----------------------------------------- 4420
G.ES.99.X138              -----------------------Y-----------G--G--------------------------C--C--G-----C--G--C--A--------C-----------------G--------------T----------------------------------------- 4257
G.GH.03.03GH175G          -----------------------C-----------G--------------G--------------C--T--G-----T--G-----A--------------------------G--------------T----------------------------------------- 4856
G.KE.93.HH8793_12_1       -----------------------C-----------G-----------------------------C--C--G-----C--G-----A--------C--------------------------------T---------------------------------------T- 4216
G.NG.01.01NGPL0669        -----------------------C-----------G--------------------------------T--G-----C--------A--------C-----------------G--------------T----------------------------------------- 4021
G.PT.x.PT2695             -----------------------C-----------G--G--------------------------C--C--G-----C--G-----A--------C-----------------G--------------T----------------------------------------- 4758
G.SE.93.SE6165            -----------------------C--------------------------G--------------C--C--G-----C--G-----A--------------------------G--------------T----------------------------------------- 4213
H.BE.93.VI991             -----------------------C--G---------------------------C-------G--C-----G-----------C--A--------C----G---------------------------T----------------------------------------- 4205
H.BE.93.VI997             --------------------------------------------------G--------A--G--C-----G-----------C--A--------C----G---------------------------T----------------------------------------- 4140
H.CF.90.056               --G-----------------------G-----------------------------------G--C-----G-----------C--A--A-----C--------------------------------T----------------------------------------- 4163
J.CD.97.J_97DC_KTB147     -----------------------------------------------------------------C--------A--------A--A--------C--------------------------------T--------------------------------A-------- 4010
J.SE.93.SE7887            -----------------------------------------------------------------C-----------------A--A--------C--------------------------------A--------------------------------A-------- 4130
K.CD.97.EQTB11C           -----------------------C-----------------------------------------C--------------G--G--------------------------------------------T----------------------------------------- 4013
K.CM.96.MP535             --G--------------------C-----------------------------------------C--------------G-----------------C--A-------------------------------------------------------------------- 4012
```

```
B.FR.83.HXB2                    GGAATTCCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAAAGAATTAAAGAAAATTATAGGACAGGTAAGAGATCAGGCTGAACATCTTAAGACAGCAGTACAAATGGCAGTATTCATCCACAATTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGC 4816
Pol                             _G__I__P__Y__N__P__Q__S__Q__G__V__V__E__S__M__N__K__E__L__K__K__I__I__G__Q__V__R__D__Q__A__E__H__L__K__T__A__V__Q__M__A__V__F__I__H__N__F__K__R__K__G__G__I__G__G__Y__S__A
01_AE.CF.90.90CF11697           --G--------------------C--------------------------G------------G-C-----G-----------G--A--------C--------------------------------T----------------------------------------- 4765
01_AE.CN.05.FJ051               --G-----------------------------------------------G--G-----------C-----G-----------A--A--------C-C------------------------------T----------------------------------------- 4826
01_AE.CN.06.FJ054               --G--C-----------T--------------------------------G--------------C-----G-----------G--A--------C------------C-------------------T-----------G----------------------------- 4841
01_AE.HK.x.HK001                --G--C--------------------------------------------G--C-----------C-----G-----------A--A-----------------------------------------T----------------------------------------- 4178
01_AE.TH.01.01TH_R2184          --G--C---------------------------------AG---------G--------------C-----G-----------G--A--------C--------------------------------T-----------G----------------------------- 4016
01_AE.TH.02.OUR769I             --G--C--T-----------------------------------------G--------------C-----G-----------G--A--------C--------------------------------T----------------------------------------- 4015
01_AE.TH.90.CM240               --G--C--------------------------------------------G--------------C-----G-----------G--A--------C-----A--------------------------T----------------------------------------- 4390
01_AE.US.00.00US_MSC1164        --G--C-----------------C-----------------C--------G--------------C----AG-----------A--A--------C--------------------------------T----------------------------------------- 4021
02_AG.CM.02.02CM_4082STN        -----------------------C-----------G-----------------------------------G-----C--G-----A--------C-----A-----------G--------------T----------------------------------------- 4015
02_AG.EC.x.ECU41                -----------------------C-----------------------------------------------------C--G-----A--------C-----------------G--------------T----------------------------------------- 3912
02_AG.GH.03.GHNJ196             -----------------------C-----------G-----------------------------------------C--G-----A-----G--C-----------------G--------------T----------------------------------------- 4845
02_AG.NG.01.PL0710              -----------------------------------G--------------G--------------------G-----C--G-----A--------C--------------------------------T----------------------------------------- 4000
02_AG.NG.x.IBNG                 -----------------------C-----------G-----------------------------------------C--------A--------C-----------------G--------------T----------------------------------------- 4341
02_AG.SE.94.SE7812              -----------------------C-----------G-----------------------------------G-----C--G-----A--------C-----------------G--------GC----T----------------G------------------------ 4188
02_AG.SN.98.MP1211              AATTCC...--------------------------G-----------------------------------G-----C--G-----A--A-----C-----------------G--------------T----------------G------------------------ 4013
03_AB.RU.97.KAL153_2            --C------------------------------------------------C--------C---C---------------------A----------------------------------------------------------------------------------- 4042
04_cpx.CY.94.CY032              -----------------------C-----------G-----------------------------C-----G-----C--G-----A--------C-----------------G--------------T----------------------------------------- 4182
05_DF.BE.x.VI1310               --------------------------------------G--------------G-----------C--------------------A------------------------------A-----------------------------------------------G---- 4198
06_cpx.AU.96.BFP90              --------T--------------------------G-----------------------------C-----G-----------A--A--------C--------------------------------T----------------------------------------- 4844
08_BC.CN.97.97CNGX_6F           --------------------------G--------------C--------G-----------C--------G--------------A-----G--C--------------------------------T----------------------------------------- 3996
09_cpx.GH.96.96GH2911           -----------------------C-----------G-----------------------------C-----G-----C--G-----A--------C-----------------G--------------T------------------A---------------------- 4018
10_CD.TZ.96.96TZ_BF061          --------------------------G--------------C------------C----------C--------------------A--------G--------------------------------T----------------------------------------- 4195
11_cpx.GR.x.GR17                --G-----------------------------G--G--------------G-G-------G----C-----G-----------A--A-----------------------------------------T------------G----G----------------------- 4118
12_BF.AR.99.ARMA159             --------------C-----------G--------------------------GC----A--G--C---A-C--------------A-----------------------G--------------------------------------------------A-------- 4821
13_cpx.CM.96.1849               -----------T-----------------------G-----------------------------C-----G-----C--G--G--A--------C--------------------------------T----------------------------------------- 4221
14_BG.DE.01.9196_01             --C-----------------------------------------------------G-----------C---------M----------------------A----------------------------------------------------------------C--- 4333
15_01B.TH.99.99TH_MU2079        T-GGA---------------------------------------T-----G--------------C-----G--AA-------G--A--------C----------------C---------------T----------------------------------------- 4215
16_A2D.KR.97.97KR004            --G--------------------------------G--------------------------G--C-----G---A----------A--------C----------------------------------------C---G--------------------A-------- 4179
17_BF.PE.02.PE02_PCR0155        --------------C--------------------------------------G--------G--C--------------------A---------------------------------------------------------------------------A-------- 4006
18_cpx.CU.99.CU76               -----------------------C-----------------------------------------C--------A--------C--A-----------G-----------C-----------------T----------------------------------------- 4121
19_cpx.CU.99.CU7                --G-----------------G--------------G-----------------------------C--------A--T--G-----A-----------------------------------------T----------------------------------------- 4042
20_BG.CU.99.Cu103               -----------------------C-----------G-----------------------------C--T--G-----G--G-----A--------C-----------------G--------------T----------------------------------------- 4282
21_A2D.KE.91.KNH1254            --G-----T--------------------------G-----------------------------C-----G--------------A--------C----------------------------------------------------------------AC--T----- 4021
22_01A1.CM.01.01CM_0001BBY      --------T--------------------------G--------------G-----------G--C-----G--AA----G--A--A-----------------------------------------T----------------------------------------- 4024
23_BG.CU.03.CB118               -----------------------C-----------G--------------G--------------C--C--G-----C--G-----A--------C-----------------G--------------T----------------------------------------- 4259
24_BG.CU.03.CB378               -----------------------C-----------G-----------------------------C--C--G-----T--G-----A--------C-----------------G--------------A----------------------------------------- 4241
24_BG.CU.03.CB471               -----------------------C-----------G-----------------------------C--C--G-----C--G-----A--------C-----------------G--------------T----------------------------------------- 4256
25_cpx.CM.01.101BA              -----------------------C-----------G-----------------------------C-----G-----T-----G--A--------C--G--------------G-------------------------------------------------------- 4024
27_cpx.CD.97.97CDKTB49          -----------------------C-----------G-----------------------------C-----G-----C--G-----A--A-----C-----------------G--------------T----------------------------------------- 4024
28_BF.BR.99.BREPM12609          --C------------------------------A-------C--------G-----------------------------------A-----------------------------------------------------G--------------------A-------- 4182
29_BF.BR.02.BREPM119            --C-----------------------------------------------G-------------------A----A-------------------------------------------------------------------------------------A-------- 3990
31_BC.BR.02.110PA               --------------------------G--------------A-----------------------C--------------------A-----G--C--------------------------------T----------------------------------------- 4250
32_06A1.EE.01.EE0369            -----------------------------------G--------------G--------------------G--------------A--------C--------G-----------------------T----------------------------------------- 4451
33_01B.MY.05.05MYKL007_1        --G--C--T--------------------------------C--------G--------------C-----G-----------G--A--------C--------------------------------T----------------------------------------- 4195
34_01B.TH.99.OUR2478P           --G--C--------------------------------------------G--------------C-----G--------------A----------------------------------------------------------------------------------- 4018
35_AD.AF.05.05AF095             -----------------------------------G--------------G-----------G--C-----G-----------G--A--------C--------------------------------T--------------------------------A-------- 4009
36_cpx.CM.00.00CMNYU830         -----------------------C-----------G-----------------------------------G-----C--------A--------C-----------------G-----G--------T----------------------------------------- 4021
37_cpx.CM.00.00CMNYU926         -----------T--C--------------------G-----------------------------C-----------C--G-----A--------C-----------------G--------------T----------------------------------------- 4006
38_BF1.UY.05.UY05_4752          -----------------------------------------------------G--------G--C---CA---------------A-----------------------------------------T----------------------------------------- 4227
39_BF.BR.03.03BRRJ103           --C-----------------------------------------------T--------------------------------------------------A-------------------------------------------------------------------- 4305
40_BF.BR.04.04BRRJ115           --G--------------T-----------------------------------GC-------------------------------------------------------------------------T--------------------------------A-------- 4368
42_BF.LU.03.luBF_05_03          --C-------------------------------------------------------------------A---A--------------------C--C----------------------------------------------------------------------- 4354
43_02G.SA.03.J11223             -----------------------C-----------G-----------------R-----------C--------A--T--G-----A--------C-----------------G--------------T----------------------------------------- 4341
U.CD.83.83CD003                 -----------------------C-----------------C-----------------------C-----G--------------A--------C--------------------------------T----------------------------------------- 4311
U.CD.90.90CD121E12              -----------------T-----------------------C-----------------------C-----G-----------G--A--------C--C-----------G-----------------T----------------------------------------- 4015
U.GR.99.GR303                   ---------------T-------C-----------G--------------G--G-----------C-----G-----C--G-----A--------------------------G--------------A----------------------------------------- 4087
U.NL.95.U_NL_95_H10986_D1       -----------------------------------------------------------------C--------------------A-----------G-----------------------------T----------------------------------------- 4701
CPZ.CD.90.ANT                   -----A--A--------A-----------------G-----C---------C----G---C-------------AA-T-----C--A--A-----AT-A------------GT---------GCA---T-------------------------------------CAC- 4255
CPZ.CM.01.SIVcpzCAM13           --------------C--A--------------G--------C-----C-----------------C--T-A---A--T-----------A--------A--A--------------------G----GT-----------G-------------------------C--- 4396
CPZ.CM.05.SIVcpzEK505           -----A-----------T--------G--T-C---------C--------------------------------AA-T--------A--A------T-A--A--------------------C-----T-----------G--------------------A----C--- 4388
CPZ.CM.05.SIVcpzLB7             --------------------G-----------G--------A-----C------C-T--------A-----G-----------------A--G--C--A--A--------------------------T--------------------------------A-------- 4371
CPZ.CM.05.SIVcpzMB66            -----------------T-----------G-----------------------------------A-----T---A----------A--A-----------A--------C-----------------T----------------------------------------- 4362
CPZ.CM.05.SIVcpzMT145           --------A-----C--A-----------------G-----------------------A--G--C----AG--A--C------------------T-A--A--------------------C-----T----------------------------------------- 4358
CPZ.GA.88.GAB1                  -----A-----T-----A--G-----------G--------CT-A--------GC----------A-----------T--G-----A--A------T-A--A--------------------G-----T-------------------------------------C--- 4877
CPZ.TZ.00.TAN1                  --G-----A--T-----T---------------------------------C--------C---------AG--A--C--G--C--A--A--G--A--G-G----------ATC--------G-AT----------------------------------A-----C--- 4460
CPZ.US.85.CPZUS                 --------A--------A--------------G--------C-----------------------C--------AA-C--G-----A--A--G---T-A--------T-----G--------------T-----------------------------------T-C--- 4876
CPZ.CM.x.SIVcpzDP943            --------T--T--C--A-----C-----------------A-----C--------------G--C--------AA-T-----A--A--A--G---T-A-----------T-----------G-----T-------------------------------------C--- 4402
CPZ.CM.x.SIVcpzMB897            -----------------T--G--------------------A-----C-----G-----------C-----G--------------A--A------T-A-----------G-----------C-----T----------------------------------------- 4382
N.CM.02.DJO0131                 --G--A-----------T--------G--------------C--------------------------------AA-C--------A--A--------A--A--------G--------G--G-----T-------------------------------A-----C--- 4322
N.CM.04.04CM_1015_04            --G--A-----------T--------G----C------G--C--------------------------------AA-C--------A--A--------G--A-----------------G--C-----T-------------------------------------C--- 4321
N.CM.04.04CM_1131_03            --G--A-----------T--------G----C------G--C--------------------------------AA-C--------A--A--------A--A--------G-----------C-----T-------------------------------------C--- 4320
N.CM.95.YBF30                   --G--A-----------T--------G----C------G--C--------------------------------AA-C--------A--A--------A-----------G--------G--T-----T-------------------------------------C--- 4410
N.CM.97.YBF106                  --G--A-----------T--------G----C------G--C--------------------------------AA-C--G-----A--A--------A--A--------G--------G--T-----T-----------------------------------M-CG-- 4408
O.BE.87.ANT70                   -----A--A--T-----A---------------------G-C--------G--------ATC-------CAG-----G--G--C--A--A-----CT-A-GA-----------------------TG-T-------------------------------------C--- 4870
O.CM.91.MVP5180                 --G--A--A--T-----A---------------------G-C-----------------ATCT------CAG-----G--G--C--A--A--G---T-A--A--------------------C--TG-T-------------------------------------C--- 4845
O.CM.96.96CMABB637              -----A--A--T--C--A---------------------G-C--------G---------TC---C---CAG-----G--G--C--A--A--G---T-A--A-----------------------TG-T----------------------------------G--C--- 4300
O.CM.98.98CMA104                -----A--A--T-----A---------------------G-C--------G---------TC---C----A---A--G--G--C--A--A--G--CT-A-GA--------------------------T-------------------------------------C--- 4302
O.CM.99.99CMU4122               -----A--A--T-----A---------------------G-C-----C--G--------ATC---A---CAG-----G--G--C--A--A-----CT-A-GA-----------------------TG-T-------------------------------------C--- 4301
O.FR.92.VAU                     -----A--A--------A--------G-----------G--C-----C--G-----GG-ATC---C--GCAG-----G--G--C--A--A-----CT-A--------------------------TG-T-------------------------------------C--- 4382
O.SN.99.SEMP1299                -----A--A--T-----A--------------------GG-C-----C--G--------ATC-------CAG-----G--G--C--A--A-----CT-A--A-----------------------TG-G-------A-----------------------------C--- 4869
O.US.99.99USTWLA                R----A--A--------A----------R----------G-C--------G--------ATC-------CA------G--G--C--A--A--G--CT-A--A-----------------------TG-T-------------------------------------C--- 4299
```

```
B.FR.83.HXB2              AGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAAATCCACTTTGGAAAGGACCAGCAAAGCTCCTCTGGAAAGGTGAAGGGGCAGTAGTAATACAAG 4986
Pol                       __G__E__R__I__V__D__I__I__A__T__D__I__Q__T__K__E__L__Q__K__Q__I__T__K__I__Q__N__F__R__V__Y__Y__R__D__S__R__N__P__L__W__K__G__P__A__K__L__L__W__K__G__E__G__A__V__V__I__Q__
A1.GE.99.99GEMZ011        -------------A---------------------------------------------------T------------------------------A---------G-----A-------------------A--A-----------------------G--------G- 4170
A1.KE.00.KER2008          -------------A--------------T-------------C------C--------------------------------------------------------G-------------------------A--A--------------------------------G- 4185
A1.KE.00.KNH1144          -------------A----T-----------------------------------------C---------------------------------------------G-----G-------------------A--A--T-----------------------------G- 4191
A1.KE.00.KSM4024          -------------A----------------------------------------------T-------------------------------T-------------G-----G-------------------A--A--------------------------------G- 4182
A1.KE.00.MSA4069          -------------A--------------------------------------------------------------------------------------------G-------------------------A--A--------------------------------G- 4185
A1.KE.00.NKU3005          -------------A----T-----------------------C-G---------------------------------A---------------------------G-----G-------------------A--A--------------------------------G- 4185
A1.RU.00.RU00051          -------------A-------------------T----------G---------------------------C-----------------------A---------G-C---G-------------------A--A-----------------------G--------G- 4287
A1.RW.93.93RW_024         -------------A-------------------T------------------------------------------------------------------------G-----A-------------------A--A--T-----------------------------G- 4188
A1.SE.95.SE8891           -------------A-----------------------------------C--------------------------------------------------------G-------------------------A--A------------------------C-------G- 4175
A1.SE.95.UGSE8131         -------------A--------------------------------------------------------.-----------------------------------G-----A-------------------A--G--T--------------------G--------G- 4357
A1.TZ.01.A173             -------------A--------------------------------------------------------------------------------------------G-----A-------------------A--A---------------------------------- 4185
A1.UA.01.01UADN139        -------------A----------------------------------------------------------------------------------A---------G-----A-------------------A--A--------------------G--G--------G- 4185
A1.UG.92.92UG037          -------------A--------------T-----T-------------------------------------------A---------------------------G-----A-------------------A--A-----------------------G--------G- 5002
A1.UG.99.99UGA07072       -------------A------------------------------G----C----------C-----------G-------------------T-------------G-------------------------A--A--------------------------------G- 4185
A2.CD.97.97CDKTB48        -------------A-------------------T-----------------------------C-T----------------------------------------G-----A-------------------A-----T------------------------------- 4326
A2.CY.94.94CY017_41       -------------A-------------------T---------------C------G-------------------------------------------------G-----A-------------------A-----T------------------------------- 4343
A.SN.01.DDI579            -------------A--------------------------------------------------------------------------------------------G-C---A-------------------A--A--------------------------------G- 4184
A.SN.01.DDJ369            -------------A-----------------------------C---------------------T----------------------------------------G-C---A-------------------A--A--------------------------------G- 4187
A.SN.96.DDJ360            -------------A--------------------------------------------------------------------------------------------G-C---A-------------------A--A--------------------------------G- 4184
A.CD.02.02CD_KTB035       -------------A-------------T---------------------------------G--TT----------------------------------------G----TA-C-----------T--------A---------------------------------- 4980
A.CD.97.97CD_KCC2         ----------------------------------T-----------------------------------------------------------------------G-C---A-------------------A--A---------------------------------- 4982
A.CD.97.97CD_KTB13        -------------------------------------------------C--------------------------------------------------------G-C---A----------------C--A--A---------------------------------- 4190
B.AR.04.04AR151516        ----------------------------------------------------------------------------------------------------------G--------------A-------------T---------------------------------- 4208
B.AU.87.MBC925            -------------A--------------------------------------------------------------------------------------------G----------------------------T---------------------------------- 4997
B.BO.99.BOL0122           -------------------------------------------------C--------------------------------------------------------G--------------------------------------------------------------- 4206
B.BR.03.BREPM2012         -------------A-----------------------------------C--------------------------------------------------------G----------------------------T---------------------------------- 4411
B.CA.97.CANB3FULL         ----------------------T-----------------------------------------------------------------------------------G----------------------------T--G------------------------------- 4275
B.CN.05.05CNHB_hp3        ---------G----------------------------------G----C--------------------------------------------------------G----------------------------T--G-----------------------G------- 4985
B.CO.01.PCM001            ------------------------------------------------------------G---------------------------------------------G----------------------------T---------------------------------- 4191
B.GB.83.CAM1              -------------A-------------------T-----------------------------------G------------------------------------G----------------------------T---------------------------------- 4987
B.GB.86.GB8               ---------------------------------T------------------------------------------------------------------------G----------------------------T---------------------------------- 5006
B.GE.03.03GEMZ010         ------------------T--------------T---------------C--------------------C---------------------T-------------G----------------------------G---------------------------------- 4191
B.IT.05.SG1               -------------A-------------------------------------------C-------------------------------------------A----G-----T----------------------T---------------------------------- 4788
B.JP.05.DR6538            ------G---------------------------------------------------------------------------------------------------G-G--------------------------T---------------------------------- 4990
B.KR.05.05CSR3            ----------------------------------------------------------------------------------------------------------G-----A----------------------T--------------------------G------- 5011
B.NL.00.671_00T36         -------------A-------------------------------------------G------------------------------------------------G----------------------------T---------------------------------- 4551
B.RU.04.04RU128005        ---A-----------------------------T------------------------------------------------------------------------G----------------------------T--T--------------------R---------- 4477
B.TH.00.00TH_C3198        --------------------------------------------G-------------------------------------------------------------G-------------------------------T------------------------------- 4200
B.UA.01.01UAKV167         ---------------------G-----------T-----G---------C--------------------------------------------------------G-------------------------A--G---------------------------------- 4212
B.US.04.ES10_53           -------------------------------------------------C---------------------------------------------------A----G----C--------------------A--T--G------------------------------- 4995
B.US.99.PRB959_03         ------------------------------------------------------------------------------------------------A---------G----------------------------T---------------------------------- 4206
C.AR.01.ARG4006           -------------A-----------------------------------C---------------T------------------------------A---------G----TA----------------C--A--A---------------------------C------ 4170
C.BR.04.04BR013           -------------A-----------------------------------C---------------T------------------------------A---------G----TA----------------C--A--A---------------------------C------ 4450
C.BW.00.00BW07621         -------------A------------------------------G--------------G-----T-C----------------------------A---------G-C--TA----------------C--A--A---------------------------------- 4339
C.CN.98.YNRL9840          -------------A--------------T------------------------------------T----------------------------------------G-C--CA----------------C--A--A---------------------------------- 4170
C.ET.02.02ET_288          -------------A----T--------------------G---------C-------C------TT------------------------------A---------G-C--TA----------------C--A--A---------------------------------- 4188
C.GE.03.03GEMZ033         -------------A---------------------------------------------------T------------------------------A---------G-C--TA----------------C--A--A--------------------G------------- 4167
C.IL.99.99ET7             -------------A-------------------T-------------------------------T------------------------------A---------G-C--TA----------------C--A--A--------------------------G------- 4167
C.IN.99.01IN565_10        -------------A-----------------------------------C---------------T------------------------------A---------G-C--TA----------------C--A--A---------------------------------- 4367
C.KE.00.KER2010           -------------A------------------------------G------------G-------T-C----------------------------A---------G-C--TA----------------C-----A---------------------------------- 4167
C.MM.99.mIDU101_3         -------------A---------------------------------------------------T------------------------------A---------G-C--CA----------------C--A--A---------------------------------- 4337
C.MW.93.93MW_965          ------G------A----T-------------------------G--------------------T------------------------------A---------G-C--TA----------------C--A--A---------------------------------- 4173
C.SN.90.90SE_364          -------------A---------------------------------------------------T------------------------------A---------G-C--TA----------------C--A--A--------------------------G------- 4155
C.SO.89.89SM_145          -------------A---------------------------------------------------T------------------------------A---------G-C--TA----------------C--A--A--------------------------C------- 4207
C.TZ.02.CO178             -------------A-----------------------------C--------------G--------------------------------------A---------G-C--TA----------------C--A--A---------------------------------- 4167
C.UY.01.TRA3011           -------------A------------------------------.----C---------------T------------------------------A---------G----TA----------------C--A--A-----------------------G---C------ 4157
C.YE.02.02YE511           -------------A-------------------T-------------------------------T------------------------------A---------G-C--TA----------------C--A--A---------------------------------- 4161
C.ZA.04.04ZASK164B1       -------------A----T--------------------------------------------------------------C--------------A---------G-C--TA----------------C--A--A---------------------------------- 4400
C.ZA.05.05ZAPSK240B1      -------------A--------------T------------------------------------T------C-----------------------A---------G-C--TA----------------C--A--A---------------------------------- 4390
C.ZM.02.02ZM108           -------------A-----------------------------C----------------------T------------------------------A----A----G-C--TA----------------T--A--A---------------------------C---G-- 4955
D.CD.83.ELI               -------------A---------------------------------------------------T------------------------------A---------G-----A--------------------------------------------------------- 4532
D.CM.01.01CM_0009BBY      -------------A------------------------------G----C----------C--C-----C-------------------------------A----G-----A----------------------T--T-----------------------G------- 4185
D.KE.01.NKU3006           -------------A---------------------------------------------------T----------------------------------------G-----A----------------------T---------------------------------- 4191
D.KR.04.04KBH8            -------------A------------------------------G---------------------------A---------------------------------G-C---A-------------------A--G---------------------------------- 4945
D.TD.99.MN011             -------------A------------------------------G------------------C------------------------------------------G-----A----------------------T-----------------------G---------- 4210
D.TZ.01.A280              ---A---------A--------------T-----T---------C------------------C--------------C---------------------------G-C---A----------------------T---------------------------------- 4187
D.UG.99.99UGK09259        -------------A-------------------------------------------------C------------------------------------------G-----A-------------------A--T---------------------------------- 4191
D.YE.01.01YE386           -------------A-------------------------------------------------C------------------------------------------G-----A-------------------A--T---------------------------------- 4185
D.YE.02.02YE516           -------------A-------------------------------------------------C------------------------------------------G-----A----------------------T---------------------------------- 4197
D.ZA.90.R1                -------------A-------------------------------------------------C------------------------------------------G-----A-------------------A--T---------------------------------- 4332
F1.AR.02.ARE933           -------------------------T-------------G----G--------------------T---------------C------------------------G-C---G----------------C-----A--------------------------T-----G- 4280
F1.BE.93.VI850            -------------A-----------T------------------G---------------------G--------------C--------------------------C---G----------------------A--------------------------C------- 4326
F1.BR.01.01BR125          -------------A-----------T------------------G--------------------T---------------C------------------------G-C---G----------------------A--------------------------C------- 4431
F1.ES.x.P1146             -------------A----T----------------------------------------------T---------------C------------------------G-C---G-------------------A--A--------------------------C------- 4263
F1.FI.93.FIN9363          -------------A-----------------------------------------------G-------------------C------------------------G-C---G----------------------A--------------------------C------- 4315
F2.CM.02.02CM_0016BBY     -------------A----------------------------------------------G--CT-------C-------------------T---A---------G-C---A----------------------A--------------------------C------- 4176
F2.CM.95.MP257            -------------A--------------------------------------------------T---------------------------T-------------G-C---G-------------------A--A--------------------------C------- 4195
F2.CM.97.CM53657          -------------A-----------------------------------------------------------------C-------------T-------------G-C---A----------------------A--------------------------C------- 4176
G.BE.96.DRCBL             -------------A--------------T----T---------------C--------------------------------------------------------G-----A-------------------A--A---------------------------------- 4943
G.CM.01.01CM_4049HAN      -------------A--------------T----T---------------C--------------------------------------------------------G-----A-C-----------------------------------------------------G- 4188
G.CU.99.Cu74              -------------A----T---------T----T---------------C------------------------------------------T-------------G-C---A-------------------A--A--------------------------G------- 4590
G.ES.99.X138              -------------A--------------T------------------C-C------------------------------------------T-------------G-C---A----------------------A--------------------G------------- 4427
G.GH.03.03GH175G          -------------A--------------T----T---------------C------------------------------------------T-------------G-C---A-------------------A--A---------------------------------- 5026
G.KE.93.HH8793_12_1       ------G--G---A--------------T----T---------------C----------------MC------------------------T-------------G-C---G-------------------A--G---------------------------------- 4386
G.NG.01.01NGPL0669        ---------G---A--------------T--------------------C--------------------------------------------------------G-C--CA-------------------A--A--------------------G------------- 4191
G.PT.x.PT2695             -------------A--------------T----T---------------C-------G----------------------------------T----------C--G-C---A-------------------A--A---------------------------------- 4928
G.SE.93.SE6165            ---A---------A--------------T----T---------------C--------------------------------------------------------G-C---G----------------C--A--A---------------------------------- 4383
H.BE.93.VI991             -C-----------A------------------------C----------C--------------T--C----------A---------------------------G-C---A-------------------A--T---------------------------------- 4375
H.BE.93.VI997             -------------A--------------------------------------------------T----T--------A---------------------------G-C---A-------------------A------------------------------------- 4310
H.CF.90.056               -------------A--------------------------------------------------T----C--------A---------------------------G-C---A-------------------A------------------------------------- 4333
J.CD.97.J_97DC_KTB147     -------------A--------------T----------------------------------C-T----------------------------------------G-C---A-------------------A--A---------------------------------- 4180
J.SE.93.SE7887            -------------A------------------------------G-------------------------------------------------------------G-C---A-------------------A--A-C------------------------------G- 4300
K.CD.97.EQTB11C           ---A--G------A----T---------------------------------------------------------------------------------------G-A---A--------------------------------------------------------. 4182
K.CM.96.MP535             ------------------T---------------------------------------------TT---C--------A---------------------------G-A---A----------------------A---------------------------------- 4182
```

```
B.FR.83.HXB2                  AGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAAATCCACTTTGGAAAGGACCAGCAAAGCTCCTCTGGAAAGGTGAAGGGGCAGTAGTAATACAAG 4986
Pol                           __G__E__R__I__V__D__I__I__A__T__D__I__Q__T__K__E__L__Q__K__Q__I__T__K__I__Q__N__F__R__V__Y__Y__R__D__S__R__N__P__L__W__K__G__P__A__K__L__L__W__K__G__E__G__A__V__V__I__Q__
01_AE.CF.90.90CF11697         -------------A-------------------A---------------Y--------------------------------------------------------G-C---A-------------------A--A---------------------------------- 4935
01_AE.CN.05.FJ051             ---A---------A--------------------------------------------------------------------------------------------G-C---A-------------------A--A---------------------------------- 4996
01_AE.CN.06.FJ054             -------------A--------------------------------------------------------------------------------------------G-----A-------------------A--A--T--------------A---------------- 5011
01_AE.HK.x.HK001              -------------A-------------------------------------------------------------------------C------------------G-C---A-------------------A--A---------------------------------- 4348
01_AE.TH.01.01TH_R2184        -------------A------------------------------G-------------------------------------------------------------G-C---A-------------------A--A---------------------------------- 4186
01_AE.TH.02.OUR769I           -------------A--------------------------------------------------------------------------------------------G-C---A-------------------A--A---------------------------------- 4185
01_AE.TH.90.CM240             -------------A--------------------------------------------------------------------------------------------G-C---A-------------------A--A---------------------------------- 4560
01_AE.US.00.00US_MSC1164      -------------A--------------------------------------------------------------------------------------------G-C---A-------------------A--A---------------------------------- 4191
02_AG.CM.02.02CM_4082STN      ---A---------A--------------T----T---------------C--------------------------------------------------------G-C--CA-------------------A--A--T-----------------------------G- 4185
02_AG.EC.x.ECU41              -------------A--------------T----T---------------C---------------T---------------------C------------------G-C--CA-------------------A--A-----------------------T---G----G- 4082
02_AG.GH.03.GHNJ196           -------------A--------------T--------------------C---------------T------------------------------A---------G-C--CA-------------------A--A--------------------------------G- 5015
02_AG.NG.01.PL0710            -------------A----T---------T----T---------------C--------------------------------------------------------G-C--CA-------------------A--A--T-----------------------------G- 4170
02_AG.NG.x.IBNG               -------------A----T---------T----T---------------C----------G----T----------------------------------------G-C--CA-------------------A--A--------------------------------G- 4511
02_AG.SE.94.SE7812            ---A--G------A--------------T----T---------------C----------G---------------------------------------------G-C--CA-------------------A--A--------------------------------G- 4358
02_AG.SN.98.MP1211            -------------A----T---------T----T---------------C-------------------T--------A---------------------------G-C--CA-------------------A--A---------------------------------- 4183
03_AB.RU.97.KAL153_2          -------------A---------------------------------------------------T------------------------------A---------G-----A-------------------A--A-----------------------G--------G- 4212
04_cpx.CY.94.CY032            -------------A--------------T----T---------------C--------------------------------------------------------G-A---A-------G-----------A--A--------------------------------G- 4352
05_DF.BE.x.VI1310             -------------A-----------A-------------------------------------C-T----------------------------------------G-C---G----------------------A--------------G--A--------C------- 4368
06_cpx.AU.96.BFP90            -------------A--------------T----T---------------C---------------T----------------------------------------G-C---A-------------------A--A---------------------------------- 5014
08_BC.CN.97.97CNGX_6F         --------------------------------------------G--------------------T------------------------------A---------G-C--CA----------------C--A--A---------------------------------- 4166
09_cpx.GH.96.96GH2911         -------------A--------------T----T---------------C-------------C------------------------------------------G-C---A-------------------A--A--------------------------------G- 4188
10_CD.TZ.96.96TZ_BF061        -------------A-----------------------------------C---------------T------------------------------A---------GGC--TA----------------C--A--A---------------------------------- 4365
11_cpx.GR.x.GR17              -------------A--------------------C--------------C----------------------------A---------------------------G-C---A-C-----------------A--A---------------------A----------G- 4288
12_BF.AR.99.ARMA159           -------------A-----------T------------------G--------------------T---------------C------------------------G-C---G----------------------A--------------------------C------- 4991
13_cpx.CM.96.1849             -------------A--------------------T-------A------C---------------T----------------------------------------G-C---A-------------------A--G---------------------------------- 4391
14_BG.DE.01.9196_01           ---A------------------------T----T---------------C------------------------------------------T-------------G-C---A-------------------A--A---------------------------------- 4503
15_01B.TH.99.99TH_MU2079      -------------A-----T----------------T-C--------------C---C---...T-C---------------------------------------G-C---A-------------------A--A---------------------------------- 4382
16_A2D.KR.97.97KR004          -------------A-------------------T---------------C------G-------------------------------------------------G-----A-------------------A------------------------------------- 4349
17_BF.PE.02.PE02_PCR0155      -------------A-------G---T------------------G--------------------T---------------C------------------------G-C---G----------------------A--------------------------C------- 4176
18_cpx.CU.99.CU76             ----------------------------T----A-----------------C------------------------------------------------------G-C---A-------------------A--A--------------------------------G- 4291
19_cpx.CU.99.CU7              -------------A--------------T--------------------C-----------------------TC---C---------------------------G-C--CA-------------------A--A---------------------------------- 4212
20_BG.CU.99.Cu103             -------------A----T---------T----T---------------C--------------------------------------------------------G-C---A-------------------A--A---------------------------------- 4452
21_A2D.KE.91.KNH1254          -------------A-------------------T----------------------G---------------------------------------A---------G-----A-------------------A-----------------------T------------- 4191
22_01A1.CM.01.01CM_0001BBY    -------------A----------------------------------------------------G---------------------------------------G-C---A-------------------A--A---------------------------------- 4194
23_BG.CU.03.CB118             ---A---------A-----------------------------------C--------------------------------------------------------G-C---A-------------------A--A-----------------------G---------- 4429
24_BG.CU.03.CB378             -------------A----T---------T----T---------------C------------------------------------------T---A---------G-C---A-------------------A--A---------------------------------- 4411
24_BG.CU.03.CB471             -------------A----T------T--T----T---------------C----------------------------------------------A---------G-C---A-------------------A--A-----------------A---------------- 4426
25_cpx.CM.01.101BA            -------------A--------------T--------------------C---------------T----------------------------------------G-C---A-------------------A--G--------------------------------G- 4194
27_cpx.CD.97.97CDKTB49        -------------A--------------T----T---------------C--------------------------------------------------------G-C---A----------------T--A--A---------------------------------- 4194
28_BF.BR.99.BREPM12609        -------------A-----------T------------------G---------------------------C---------------------------------G-C---G----------------------A--------------------------C------- 4352
29_BF.BR.02.BREPM119          -------------A-----------T------------------G------------------------------------C-------------------A----G-----A----------------------A-----------------------G--C------- 4160
31_BC.BR.02.110PA             -------------A-------------------------------G---C---------------T----------------------------------------G----TA----------------C--A--A---------------------------C------ 4420
32_06A1.EE.01.EE0369          -------------A--------------T----T---------------C----------------T---------------------------------------G-C---A----------------T--A--A--T-----------------------------G- 4621
33_01B.MY.05.05MYKL007_1      -------------A--------------------C----------------------G------------------------------------------------G-C---A-------------------A--A---------------------------------- 4365
34_01B.TH.99.OUR2478P         -------------A--------------------------------------------------------------------------------------------G-C---A-------------------A--A---------------------------------- 4188
35_AD.AF.05.05AF095           -------------A----T------------------------------C--------------------------------------------------------G-------------------------A--G--------------------------------G- 4179
36_cpx.CM.00.00CMNYU830       -------------A--------------T----T---------------C--------------------------------------------------------G-C---A-------------------A--A---------------------------------- 4191
37_cpx.CM.00.00CMNYU926       -------------A----T---------T----T--------C------C--------------------------------------------------------G-C--CA-------------------A--A---------------------------------- 4176
38_BF1.UY.05.UY05_4752        ---A---------A-----------T------------------G----C---------------T---------------Y------------------------G-C---G-C--------------------A--------------------------C------- 4397
39_BF.BR.03.03BRRJ103         ------------------------------------------A-G------------------C------------------------------------------G----------------------------T---------------------------------- 4475
40_BF.BR.04.04BRRJ115         -------------A----T------T-----------------------C----G---------T-----------------------------------------G-----G----------------------A--------------------------C------- 4538
42_BF.LU.03.luBF_05_03        -------------A--------------------------------------------------------------------------------------------G-C---G----------------------A--------------------------C------- 4524
43_02G.SA.03.J11223           ---------R---A--------------T----TC--------------C--------------------------------------------------------G-C---A-------------------A--A--T-----------------G-----G------- 4511
U.CD.83.83CD003               -------------A----------------------------------------------------------------------------------A---------G-C--GA-------------------A--A-----------------A--------C------- 4481
U.CD.90.90CD121E12            -------------A---------------------------------------------------T-C----------------------------A---------G-C--CA-------------------A--A-----------------A--------C------- 4185
U.GR.99.GR303                 -------------A----T-------------------------G--------------------T----------------------------------------G-C---A-------------------A--A--------------------------------G- 4257
U.NL.95.U_NL_95_H10986_D1     -------------A-------G----------------------C-------------------------------------------------------------G-C---A-------------------A--A--------------------------C------- 4871
CPZ.CD.90.ANT                 T--AC-G------T--------C----------------G--A-CTC----------T------TT----------C-A---------C----------T------G-C--TG-G---------------C-A--T--G-----------------------C--CA--- 4425
CPZ.CM.01.SIVcpzCAM13         -------------A----------------------------A------C--------------TT------------A-----------------A---------G-C--GA-------------G--C-CC--G--G--------------------------T--G- 4566
CPZ.CM.05.SIVcpzEK505         -------------A-------------------T--------A-C-A-GC------C-------TT----G-------------------------A---------G-G--CAC---------------C--A--A--G--------A-----------G-----T---- 4558
CPZ.CM.05.SIVcpzLB7           ---A---------A-------------------T----------C-------------------T-----G-------A---------------------C-----G----TA----------------C--A--A--G-----G--------A--------GC-G---- 4541
CPZ.CM.05.SIVcpzMB66          ---A---------A--------C----------T----------C-A--------C--------T----T--------A---------------------------G-C--TA----------------T--A--A--------------G--A--------GC-C---- 4532
CPZ.CM.05.SIVcpzMT145         ---A--G------A-------C--------------------A------C--------------TT----------C-A--------C--------A-------A-G-----A----------------G--A-----G--------------------G-----T---- 4528
CPZ.GA.88.GAB1                -------------A----------------------------A-GC------------------TT----G-------A-----------------A---------G-C---A-------------G----CCT-A--G-----------------------G--C--G- 5047
CPZ.TZ.00.TAN1                ----------C--T--------C--A-T---A-T-----G--A---C-----------------TT----G----------------------T------GC----G-----A-------G--------GCGA--A--G---------------------------A--- 4630
CPZ.US.85.CPZUS               ---A--------CA-------C------T----AC-------AG-CTT----------------TT----G----------------C------------------G-----A----------------C--A--T--G--------------------------CA-G- 5046
CPZ.CM.x.SIVcpzDP943          ---A-----G--TA-------T--------------------A-CTA-T---------------TT----G-----------------------------------G-----A-------------T--C-GA--T--G--------------------------CA--- 4572
CPZ.CM.x.SIVcpzMB897          ------------CA----T--------------T----------G----C--------------T-----G-------A---------------------------G-A--TA-C--------------C-----T--------------G--A--G-----GC-T---- 4552
N.CM.02.DJO0131               -------------A----------------------------A-C-A---------C-------TT---G--------------------------A---------G----CA----------------C--A--T--G--------A-----------G-----C---- 4492
N.CM.04.04CM_1015_04          -------------A-------------------------G--A-C-A---------C-------TT----G-------------------------A---------G----CA----------------C--A--G--G--------A-----------G-----C---- 4491
N.CM.04.04CM_1131_03          ------G------A-------------------------G--A-C-A---------C-------TT----G-------------------------A---------G----CA----------------C--A--T--G--------A-----------G-----C---- 4490
N.CM.95.YBF30                 -------------A-------------------------G--A-C-A-T-------C-------TT----G-------------------------A---------G----CA----------------C--A--T--G--------A-----------G-----T---- 4580
N.CM.97.YBF106                -------------A----------------------------A-C-A---------C-------TT----G------------R------------A---------G----TA----------------C--A--T--G--------A-----------G-----C---- 4578
O.BE.87.ANT70                 ---A-----G---A--------T-----T--C-A--------A-C-------------------TT----N----C--A--------C--------A---------G-C--TA-C-----------G---C-------G-----------G--A--------C------- 5040
O.CM.91.MVP5180               ------G---C--A--------C-----T--C-A--------A-C----C--------------TT------CA-C-----------C--------A--T------G-C--TA-------------G---C-A-----G-----------G-----------C------- 5015
O.CM.96.96CMABB637            ---A--G---T--A----T--GC-----T--C-A--------A-C----C----------C---TT------------A----A---C--------A---------G-C--TA-------------G---C-A-----G-----------G--A--------C------- 4470
O.CM.98.98CMA104              ---A-----GT--A--------T-----T--C-A--------A-C----C--------------TT---------------------C--------A---------G-C--TA-------------G---C-A-----G-----------G--A--------C------- 4472
O.CM.99.99CMU4122             ---A-----G---A--------T-----T--C-A--------A-C-------------------TTC-----------A----A---C--------A---------G----TAC------------G---C-------G-----G--------A--------C------- 4471
O.FR.92.VAU                   ---A--G---T--A--------T-----T--C-A--------A-C----C--------------TT---------------------C--------A---------G-C--TA-------------G---C-------G-----------G--A--------C------- 4552
O.SN.99.SEMP1299              ---A-----G---A--------T-----T--C-A--------A-C-------------------TTT-----------A----A---C--------A---------G----TA-------------G---C-------G-----------G--A--------C------- 5039
O.US.99.99USTWLA              ---A--G---T--A--------T-----T--C-A--------A-C----C--------------TT---------------------C--------A---------G-C--TA-------------G---C-------G-----------G--A--------C------- 4469
```


Vif start
Pol p31 integrase end
B.FR.83.HXB2 ATAATAGTGACATAAAAGTAGTGCCAAGAAGAAAAGCAAAGATCATTAGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATT..AGAACATGGAAAAGTTTAGTAAAACACCATATGTATGTTTCAGGGAAAGCTAGGGGATGG 5154
Vif M E N R W Q V M I V W Q V D R M R I ## R T W K S L V K H H M Y V S G K A R G W
Pol D N S D I K V V P R R K A K I I R D Y G K Q M A G D D C V A S R Q D E D ##*
A1.GE.99.99GEMZ011 4338
A1.KE.00.KER2008 4353
A1.KE.00.KNH1144 4359
A1.KE.00.KSM4024 4350
A1.KE.00.MSA4069 4353
A1.KE.00.NKU3005 4353
A1.RU.00.RU00051 4455
A1.RW.93.93RW_024 4356
A1.SE.95.SE889I 4343
A1.SE.95.UGSE8131 4525
A1.TZ.01.A173 4353
A1.UA.01.01UADN139 4353
A1.UG.92.92UG037 5170
A1.UG.99.99UGA07072 4353
A2.CD.97.97CDKTB48 4494
A2.CY.94.94CY017_41 4511
A.SN.01.DDI579 4352
A.SN.01.DDI369 4355
A.SN.96.DDJ360 4352
A.CD.02.02CD_KTB035 5148
A.CD.97.97CD_KCC2 5150
A.CD.97.97CD_KTB13 4358
B.AR.04.04AR151516 4376
B.AU.87.MBC925 5165
B.BO.99.BOL0122 4374
B.BR.03.BREPM2012 4579
B.CA.97.CANB3FULL 4443
B.CN.05.05CNHB_hp3 5153
B.CO.01.PCM001 4359
B.GB.83.CAM1 5155
B.GB.86.GB8 5174
B.GE.03.03GEMZ010 4359
B.IT.05.SG1 4956
B.JP.05.DR6538 5158
B.KR.05.05CSR3 5179
B.NL.00.671_00T36 4719
B.RU.04.04RU128005 4645
B.TH.00.00TH_C3198 4368
B.UA.01.01UAKV167 4380
B.US.04.ES10_53 5163
B.US.99.PRB959_03 4374
C.AR.01.ARG4006 4338
C.BR.04.04BR013 4618
C.BW.00.00BW07621 4507
C.CN.98.YNRL9840 4338
C.ET.02.02ET_288 4356
C.GE.03.03GEMZ033 4335
C.IL.99.99ET7 4335
C.IN.99.01IN565_10 4535
C.KE.00.KER2010 4335
C.MM.99.mIDU101_3 4505
C.MW.93.93MW_965 4341
C.SN.90.90SE_364 4323
C.SO.89.89SM_145 4375
C.TZ.02.CO178 4335
C.UY.01.TRA3011 4325
C.YE.02.02YE511 4329
C.ZA.04.04ZASK164B1 4568
C.ZA.05.05ZAPSK240B1 4558
C.ZM.02.02ZM108 5123
D.CD.83.ELI 4700
D.CM.01.01CM_0009BBY 4353
D.KE.01.NKU3006 4359
D.KR.04.04KBH8 5113
D.TD.99.MN011 4378
D.TZ.01.A280 4355
D.UG.99.99UGK09259 4359
D.YE.01.01YE386 4353
D.YE.02.02YE516 4365
D.ZA.90.R1 4500
F1.AR.02.ARE933 4448
F1.BE.93.VI850 4494
F1.BR.01.01BR125 4599
F1.ES.x.P1146 4431
F1.FI.93.FIN9363 4483
F2.CM.02.02CM_0016BBY 4344
F2.CM.95.MP257 4363
F2.CM.97.CM53657 4344
G.BE.96.DRCBL 5111
G.CM.01.01CM_4049HAN 4356
G.CU.99.Cu74 4758
G.ES.99.X138 4595
G.GH.03.03GH175G 5194
G.KE.93.HH8793_12_1 4554
G.NG.01.01NGPL0669 4359
G.PT.x.PT2695 5096
G.SE.93.SE6165 4551
H.BE.93.VI991 4543
H.BE.93.VI997 4478
H.CF.90.056 4501
J.CD.97.J_97DC_KTB147 4348
J.SE.93.SE7887 4468
K.CD.97.EQTB11C 4346
K.CM.96.MP535 4350

```
                                                                                   Vif start                                          Pol p31 integrase end
B.FR.83.HXB2                 ATAATAGTGACATAAAAGTAGTGCCAAGAAGAAAAGCAAAGATCATTAGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATT..AGAACATGGAAAAGTTTAGTAAAACACCATATGTATGTTTCAGGGAAAGCTAGGGGATGG 5154
Vif                                                                                  M__E__N__R__W__Q__V__M__I__V__W__Q__V__D__R__M__R__I_##_R__T__W__K__S__L__V__K__H__H__M__Y__V__S__G__K__A__R__G__W_
Pol                          D__N__S__D__I__K__V__V__P__R__R__K__A__K__I__I__R__D__Y__G__K__Q__M__A__G__D__D__C__V__A__S__R__Q__D__E__D__##*_
01_AE.CF.90.90CF11697        -C--------T-----------A-----------------------T----------------------------------------G-------------A---..---------T-C--------------T---------A-C---AA--------AAAAC--- 5103
01_AE.CN.05.FJ051            -C--------T-----------A-----------------------T----------------------------------------G-----------------..-----------T--------------T---------A-C---AA--------AAAAG--- 5164
01_AE.CN.06.FJ054            -CG-------T-----------A----------------------------------------------------------------G-----------------..-----------C--------------T-----------C---AA---C----AA--G--- 5179
01_AE.HK.x.HK001             -C--------T--------------------------A-------------------------------------------------G-----------------..-----------C--------------T-----------C---AA--------AAAAG--- 4516
01_AE.TH.01.01TH_R2184       -C--------T-----------A----------------------------------------------------------------G-----------------..-----------C--------------T-----------C---AA--------AAAAG--- 4354
01_AE.TH.02.OUR769I          -C--------T--C--------A---------------------------C------------------------------------G-----------------..-----------C-----------G--T---------A-C---AA--------AAAAG--- 4353
01_AE.TH.90.CM240            -C--------T-----------A----------------------------------------------------------------G-----------------..-----------C--------------T---------A-C---AA--------AAAAG--- 4728
01_AE.US.00.00US_MSC1164     -C----A---T-----------------------------------T----------------------------------------G-----------------..-----------C--------------T--C--------C---AA-----T--ACAAG--- 4359
02_AG.CM.02.02CM_4082STN     -C--C-----T-----G-----A-----------------A----------------------------------------------G-----------------..-----------C-------------T-T--C------AGA--TA---------A--AT--- 4353
02_AG.EC.x.ECU41             -C--------T-----G-----A-------------T---A----------------------------------------------G-------------A---..-----------C-------------T-T------------C--TAA--------A---T--- 4250
02_AG.GH.03.GHNJ196          -C--------T-----------A-----G-----------A-------A-----------------------------C--------G------------A----..--G--------C-----------G--T------------C--TAA--------A--AT--- 5183
02_AG.NG.01.PL0710           -C--------A-----G-----A-----------------A----------------------------------------------G-----------------..-----------C--------------T---------A-C--TAAA------CA---T--- 4338
02_AG.NG.x.IBNG              -C--------T-----G-----A-----------------A----------------------------------------------G-----------------..-----------T--------------T-----------C---AA--------A---T--- 4679
02_AG.SE.94.SE7812           -C--------T-----G-----A----------G-T---A---G-------------------------------------------G-----------------..--G--------C--C----------T-T---------AAA--TA---------A--AT--- 4526
02_AG.SN.98.MP1211           -C--------T-----G-----A------------A----A---C------------------------------------------G--G-----A--------..-----------C--------------T-----------C--TAA-------CACC-T--- 4351
03_AB.RU.97.KAL153_2         -C----AC--T-----------A---------------------------------------------------------------------------------..-----------C-------------------A---A-----AA--------------- 4380
04_cpx.CY.94.CY032           -C--C-----T--C--------A-----------------------T-------C-----C---------------A----------G------------A----..-----------C---C----G-----T---------------AA---------AA-----A 4520
05_DF.BE.x.VI1310            -C--------A-----G-----A----------------------------------------------------------------G-----------------..-AC-------------------T--------C--------AA---------ACA----- 4536
06_cpx.AU.96.BFP90           -C--------A-----G-----A-----------------------------A----------------------------------G-----------------..-AT-------------------T--------C--A-----AA---------AAA----- 5182
08_BC.CN.97.97CNGX_6F        ----------------G-----A-----G--G--------A-------A---C------------------C---------------G------------A----..-----------T---------------------------C---A---G-----AT------ 4334
09_cpx.GH.96.96GH2911        -C----A---T-----G-----A-------------------------------C--------------------------------G-----------------..-AT-------------------T--------C--AC----AA---------AAA----- 4356
10_CD.TZ.96.96TZ_BF061       -A--C-----------G-----A-----G--G----T---A-------A---C------------------C-----------------------------C..-----------T----------G-------------------A---------AT------ 4533
11_cpx.GR.x.GR17             -C--------T-----------A----------------------------------------------------------------G---C-------------..-----------T--------------T-----------C---AA---------AA--G--- 4456
12_BF.AR.99.ARMA159          -CG-------A-----G-----A--C--------------------------C----------------------------------G-----------------..-AC-------------------GT--------C--------AA------C-AAA----- 5159
13_cpx.CM.96.1849            -C--------A-----------A-----------------------------A----------------------------------G-----------------..-AT--------------------T--------C--------AA-------G-AA----- 4559
14_BG.DE.01.9196_01          -C----AC--A-----G-----A-----G----------------------------------------------------------G-----------------..-----------C--------------T-----------C---A----------A---C--- 4671
15_01B.TH.99.99TH_MU2079     -C--------T-----------A---------G------------------------------------------------------G-----------------..-----------C---C----------T-----------C---AA---------AAAAG--- 4550
16_A2D.KR.97.97KR004         -C--------T-----G-----A----------G-----------------------------------------------------G-----------------..-----------T----------G--T---------------AA---------AA-AT--- 4517
17_BF.PE.02.PE02_PCR0155     -C--------A-----G-----A----------------------------------------------------------------G-----------------..-AC-----------------T---T-------AC----G---AA----A-C-AAAAT--- 4344
18_cpx.CU.99.CU76            -C--C-----A--T--G-----A----------------------------------------------------------------G-----------------..-----------C---A----------T-----------C---AA---------AAAAT--- 4459
19_cpx.CU.99.CU7             -C--------------------T-----------------------------A----------------------------------G-----------------..-----------T--------------T---------A-C---AA---------ACAA---- 4380
20_BG.CU.99.Cu103            -C----AC--A-----G-----A-----G----------------------------------------------------------G-----------------..---------C-C--------------T-----------C---AA---------AA--C--- 4620
21_A2D.KE.91.KNH1254         -G--------T-----G-----A-----------------A-------A-----------------------G-------------------A------------..-----------C--------------T---------A----A---------AA-AT--- 4359
22_01A1.CM.01.01CM_0001BBY   -C--------T-----------A-----------------A-----------G----------------------------------G-------------A---..-----------C--------------T-----------C---AA----------A--T--- 4362
23_BG.CU.03.CB118            -C-G--AC--A-----------A-----G-----G------------------------------------------------------G-----------------..---------C-C--------------T--Y-A----A-C---AA---------AA--C--- 4597
24_BG.CU.03.CB378            -C----AC--A-----G-----A-----G----------------------------------------------------------G-------------A---..-----------C------A-------T---------A-C---AA---------AA-AC--- 4579
24_BG.CU.03.CB471            -C----AC--A-----G-----A-----G----------------------------------------------------------G-------------------..-A---------C--------------T---------A-C---AA---------AA--C--- 4594
25_cpx.CM.01.101BA           -C--------A-----------A-----------------------------A------------------------A-----G-----------------..-AT-A-------------------T--------C----C---AA---------AAAA---- 4362
27_cpx.CD.97.97CDKTB49       -C--------A-----G-----C---------------------------AA-----------------------------------G-----------------..-AT--T-------------------T-----------A-----AA----------AA-G--- 4362
28_BF.BR.99.BREPM12609       ----------------------------------------A---------------------------------------------------------------..-----------C-----------------------A-----A---------A-AA---- 4520
29_BF.BR.02.BREPM119         -C--------A-----G-----A---------------------------------------------------------------------------------..-AC-------------------T------AC--A-----AA------C--AAA---- 4328
31_BC.BR.02.110PA            ----C-----------G-----A---------G----T---A-------A---C-----------------C-------A------------------A----..----Y------T----------G-----C------R-----A---G------T------ 4588
32_06A1.EE.01.EE0369         -C--------A-----G-----A-----------------A---------A--G-------------------------A------T--G-----------------..-AT--------------G------T--------C--A-----AA---------AAAAT--- 4789
33_01B.MY.05.05MYKL007_1     -C--------T------------------------------------------------------------------G---------------------------..-----------C--------------T----A----A-C---AA---------AA------ 4533
34_01B.TH.99.OUR2478P        -C--------T-----------A----------------------------------------------------------------G-----------------..-----------T--CA----------T-----------C---AA---------AAAA---- 4356
35_AD.AF.05.05AF095          -C--------T-----G--------T-------------------------------------------------------------G-----------------..-----------T--------------T---------A-A---AA----A---AA--T--- 4347
36_cpx.CM.00.00CMNYU830      -C--------T-----------A------------------------------A---------------------------------G-----------------..-----------T-----------G--T------C--A----TAAA-C-----AA-AT--- 4359
37_cpx.CM.00.00CMNYU926      ----------T-----------A-------------------------C--------------------------------A-----G-----------------..-----------C--------------T------C--A-A---AA---------AA--C--- 4344
38_BF1.UY.05.UY05_4752       -C--------A------A----A----------------------------------------------------------------G-----------------..-AC-------------C---------T-------AC--------AA---------AA------ 4565
39_BF.BR.03.03BRRJ103        ---------------------------------------------------------------------------------------G-----------------..-AC-------------------T------A-C--------AA--------A--A---- 4643
40_BF.BR.04.04BRRJ115        -C--------A-----G-----------------------------T------------------------------------------------------------..-A-----------------------------C--A-----A---G----G-TA----- 4706
42_BF.LU.03.luBF_05_03       ----------------------------G-------------------------------------------------------------------------------..-----------C-------------------------------AA---------A------- 4692
43_02G.SA.03.J11223          -C--------A-----G-----A-----------------A----------------------------------A-------------G-----------------..-----------C---Y----------T-----------C---AA---------AA--C--- 4679
U.CD.83.83CD003              -C--------A-----G-------------G------------------C-----------------------------------------------------A---..-A---------C--------------T--C------A-----AA---------A------- 4649
U.CD.90.90CD121E12           -C--------A-----G-------------G------------------C-----------------------------------------------------A---..-----------C-----G--------T--C---------------AA---------ACAA---- 4353
U.GR.99.GR303                -C--------A-----G-----A-----------------T-------C-----------------------------------------G-----------------..-AT-----------------G------T--------C--------AA---------AAA----- 4425
U.NL.95.U_NL_95_H10986_D1    -C--------G-----------A-----------------T-------C-------G-------------------------------------------------..--T---------------------C---T--------C--------AA---------AA------ 5039
CPZ.CD.90.ANT                -CC-AGAG--A--T--G-----C--TC----G-----------T--A-A---G---A-----A----A-A--A-AGA---GA-T-----G--------A----------..-AT-TC------TCC-----G--------C--A-GG-AGA-TAAAGT-CT--AACC---- 4593
CPZ.CM.01.SIVcpzCAM13        --C-AG-A--A-----------A--T-----G--------A--A-----A---------------A----------------------------------------..-A-G-------T--CC-----T---T-T---------AAAA--AA------A-AA-AT--- 4734
CPZ.CM.05.SIVcpzEK505        ----CG-G--T-----------C---C-T-----G--C--A--A-----A--------------------------G-------------G-----------AG-C..--G--------T-----G-----G--T--C--------G---AA---G--AC-A--T--- 4726
CPZ.CM.05.SIVcpzLB7          --C-GGAA--A-----G-----C---C-T--------T--A--T---------------------A-----C-A-----C-----GA--------------A---..-A-G--------C-----G------T-T-----C------A--CA--------A--A--C--- 4709
CPZ.CM.05.SIVcpzMB66         -CC-AGAA--A-----G-----C--TC----G--------A-----------------------------------------------------------A--C..-A--------T--CC------------T--C--------G--TA-------C--AAAT--- 4700
CPZ.CM.05.SIVcpzMT145        --C-GGAC--A--------G-----T-----G-----T-----A-------------------------CA---A--------------------------AG--..-A---C-----T--------G---T-T--------CAGA---AA------A-AAAAT--- 4696
CPZ.GA.88.GAB1               --C-AG-G--AC----G-----A---------G-----------------A----------------------------------------A----------..-AG--T-----T--------T---T-T-----T---AGG---AAA--G--A--A--T--- 5215
CPZ.TZ.00.TAN1               -AGGAGAA-----T--------A--C--G------------A-----A-AA--G---------------------CA-G-G--A---ATGA-------A-----AC--..-----------C---CC----T-----T-----C-T-ACAA-CAAATGCTG--AA-AT--- 4798
CPZ.US.85.CPZUS              -A--CGAG--GG-T--------A--C---------------A--T--A-AA--C---------------------C----A--A-----G---------------AG--..-AGG-T-----T---C-G------T-T--C------AGA---AA---------AA------ 5214
CPZ.CM.x.SIVcpzDP943         ----AGAG--AG-T-----------C--G--G--------A--A--C--A---------------------------G-A--A--------G-----------AG-C..-A---C-----T--CC--------T----------CAGAA--AA------A--AAA---- 4740
CPZ.CM.x.SIVcpzMB897         ----GGAA--T-----------C---C-C-----G--T--A--A--C-----C-----------G---A---A--------C-----GA--------------AG--..-A-G--------C-----G--------T--C--------G--TAAA--G--AC-A--T--- 4720
N.CM.02.DJO0131              ----CG-A--T-----------C---C-T--G--------A--A--------------------------------G-----------G------CG---A--C..-AG------C-T--C--------G--T------------G---A-A--G--AGAA--G--- 4660
N.CM.04.04CM_1015_04         ----CG-G--T-----------C---C-T--G--------A--A-----A--------------------------G-----------G----------A--C..--G-T------T--------------T------C----G--AAA--G--A-AA------ 4659
N.CM.04.04CM_1131_03         ----CG-G--T-----------C---C-T--G--------A--A-----A--------------------------G-----------G---------AA--C..--G-T------T--------------T------C----G--AAA--G--A-AA------ 4658
N.CM.95.YBF30                ----CG-G--T-----------C---C-T--G--------A--A--------------------------------G-----------G---------AA--C..--G-A------T--C-----------T------------G--AAA--G--A-AA------ 4748
N.CM.97.YBF106               ----CG-G--T-----G-----C---C-T--G--------A--A-----A--------------------------G-----------G---R-----AA--C..--G--------T--C-------------------------G--AAA--G--A-AA--G--- 4746
O.BE.87.ANT70                ----GG-A-----T--G-----A---------G--G----A--A--C--A--G-------------------AC----A--A--------G----ACA--AAG-G..-A-G-G-----C--CC----G---T------A---CAGG--TA-A--GA-CGA-AAC--- 5208
O.CM.91.MVP5180              ----AG-A-----T-----G--A------------G----A--A--C--A----------------------AC----A--A------A------ACA--AAG-G..-A-G-------C--CC-G--G---T------AA--CA-G--TAA---G--CGC-AAC--- 5183
O.CM.96.96CMABB637           ----AG-A--T--C--G-----A-----------------A--AC-C--A-----------------------C----A--A--------G----ACA--RAG-G..-A-G-------T--CC-G--R---T------A---CARG--CAAA--G--CCA-AA---- 4638
O.CM.98.98CMA104             ----AG-A-----T--------A-----------G-----A--A--C---C----------------------C----A--A--------G----ACA--AAG-G..-A-GTG-----C--CC-G--G---T-T----AA--CAGG--TA----G--C-AA-AC--- 4640
O.CM.99.99CMU4122            ----AG-A-----T--G-----A---------G--G----A--A----AA----------------------AC----A--A--------G----RCA--AAG-G..-A---G-----C--C--G--G---T------AAC-CAGA--TA-A--GA-C-A--A---- 4639
O.FR.92.VAU                  ----AG-A-----T--------A-----------G-----A--A--C--------------------------C----A-----------G----ACA--A-G-G..-A-G-G-----C--CC----G---T------AA--CAGG--TA-A--G--C-C--AT--- 4720
O.SN.99.SEMP1299             ----AG-A-----T--G-----A-----------G-----A--A--C--AC---------------------AC----A--A--------G----ACA---AG-G..-A---G-----C--CC-G--G---T------AA--CAGG--------GA-C--A-AC--- 5207
O.US.99.99USTWLA             ----AG-A--A--T--G-----A-----------G-----A--A----------------------------AC-------A--------G----ACA--AAG-G..-A---------C--YC----G---T-T----AA--CAGG--TAA---G--C-A--AC--- 4637
```

```
B.FR.83.HXB2            TTTTATAGACATCACTATGAAAGCCCTCATCCAAGAATAAGTTCAGAAGTACACATCCCACTAGGGGAT......GCTAGATTGGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGAGACTGGCATTTGGGTCAGGGAGTCTCCATAGAATGGAGGAAAAAGAG 5318
Vif                     _F__Y__R__H__H__Y__E__S__P__H__P__R__I__S__S__E__V__H__I__P__L__G__D__.__.__A__R__L__V__I__T__T__Y__W__G__L__H__T__G__E__R__D__W__H__L__G__Q__G__V__S__I__E__W__R__K__K__R
A1.GE.99.99GEMZ011      ------------------------AGG-------A----------------------------------......-----GC-A---G---A-------------------G---------A---------C--------T--G-----------------AC-G--A-- 4502
A1.KE.00.KER2008        ------------------------AGG-------A-G--------------------------------......----A---A---G---GG------------------C---------A---T-----C--------T--G------------------CT---A-- 4517
A1.KE.00.KNH1144        G-C-----G-------T-------AGAA--------G--------------------------------......----A-C-A---G---G-------------------C---------AT-----------------T--G------------------CT---A-- 4523
A1.KE.00.KSM4024        ----------------T-------AGG-------A-G----------------------------A---......----A---A---G---T------------C----------------A---------CC-------A--G------------------CG-GGA-- 4514
A1.KE.00.MSA4069        -GG---------------------AGA-------A-GC----------------------G-------A......--------A-------G----------------A------------A--C------A--------T--G------------------CT--GA-A 4517
A1.KE.00.NKU3005        --------------T---------AGG---------G--------------------------------......------A-A-------G-----------------------------A---------A--------T--G------------------CT---A-- 4517
A1.RU.00.RU00051        C----------C------------AGG---------G-------------------------------C......------C-A---G---G-----------------------------A---------A--------T--G-----------------AC-G--A-- 4619
A1.RW.93.93RW_024       ------------------------AGA---------G-----C--------------------------......------C-AA--G---G-----------------------------A---------C-----G--T--G------------------CT---A-- 4520
A1.SE.95.SE8891         C---------------T-------AGG-------A-G----------------------------A--G......--------A---G---G------------------G----------A---------C--------T--G------------------CT--GA-- 4507
A1.SE.95.UGSE8131       --C---------------------AGG-------A-G-------------------------------C......------C-AA--G---G-------------------C-------------------A--------T--G------------------CT--GA-- 4689
A1.TZ.01.A173           ------------------------AGA-------A-G--------------------------------......--------A--GG--CG------------------G----------A---T-----C--------T--G------------------CT---A-- 4517
A1.UA.01.01UADN139      G------A----------------ATAA------A-G-----C--------------------------......-----GC-A---G---A-----------------------------A---------C--------T--G-----------------AC-G--A-- 4517
A1.UG.92.92UG037        ------------------------AGG-------A-G-----C-----------------A--------......------A-A---G---G------------------G----------A---------C--------T--G------------------CT---A-- 5334
A1.UG.99.99UGA07072     ----------------T-------AGA-------A-G----------------------------A---......------C-A---G----------------------G----------A---------A--------T--G------------------CT---A-- 4517
A2.CD.97.97CDKTB48      --C---------------------AGA---------G----------------------G---AA---G......--------AA--G---G------------------CC---------A---------C--------T--G------------------C-GGGA-- 4658
A2.CY.94.94CY017_41     G-C----A----------------AGAA-------------------------------G--------G......------A-AA--G---G------------------C-T--------A---------C--------T---------------------C----C-- 4675
A.SN.01.DDI579          --------------------CT--AGG---------GC-------------------------------......----TG--A---G---G------------C-----CC--------CA------------A-----T--G-----------------AC-G--A-- 4516
A.SN.01.DDJ369          G------A----------------AGG-------A-G--------------T-T--------------G......----T---A---G---G------C-----C-----CC--------C----------A--A-----T--G-----------------AC-G--A-- 4519
A.SN.96.DDJ360          ------------------------AGG---------G--------------------------------......----T---A---G---G------------C-----CC--------CA------------A-----T--G-----------------AC-G--A-- 4516
A.CD.02.02CD_KTB035     G---------------T--------GG-------A-GC------------G--------G---------......------C-AA--G---G------------C-----A--------------------C--------T--G---------------CAACTG--A-- 5312
A.CD.97.97CD_KCC2       --------------T---------AGG-------A------------------------G-----A---......--------A---G---A------------C-------------------G------C--------T--G----------------A-CTG-GA-- 5314
A.CD.97.97CD_KTB13      --------------T---------AGG-C-----A------------------------G---------......--------A---G---G------------C-----GC-----------------T-C--------T--G----------------A--TG-GA-- 4522
B.AR.04.04AR151516      A-----------------------A--A-------G---------------------------------......--------AA----------------------------------------A---------------------------------------G---A 4540
B.AU.87.MBC925          -G----------------------A--A--------------C-------------------------A......----A----------GT---------------------------------------------------------T-----------A-------- 5329
B.BO.99.BOL0122         -G--------------------A--T----------------GG----------------T----ACC-......----A------G-----------------C--------------------------------------------------------A--GG---- 4538
B.BR.03.BREPM2012       --------------T---------A-------------------------------T------------......--------A--------------------------------------G--T----------------------------------------G--- 4743
B.CA.97.CANB3FULL       A------A--------------A--A-------------------------------------------......----A-G-----G-----------------------------------------------------------------------------G---- 4607
B.CN.05.05CNHB_hp3      -G-----A---C--------T---A---------A----------------------------------......----A---------------------------------------------------------C-------------------------G-G-C-- 5317
B.CO.01.PCM001          ------------------------A---------A---------------------------------A......AA------A---G------------------------------------------------------------------------------G--- 4523
B.GB.83.CAM1            -C---------C------------A-------------------------------------------G......--------A---G---------------------------------------------------------------------------C----G- 5319
B.GB.86.GB8             G------A--------------A-A-------------------------------------------A......--------------------------------------------------------------C-----------------------A----G--- 5338
B.GE.03.03GEMZ010       CG----------------------A----C----A------------------T---------------......---GA---A---G---T-------------------T-------------------------A-----------------------A-------- 4523
B.IT.05.SG1             G---------------------A----------------G-----------------------------......-G--A----------------------------------------CA---------AC-------T------------------------G---- 5120
B.JP.05.DR6538          G-------------T--------TA---------A----------------------------------......-A--A------------------------------C-----------G--------------C-----------------------A----G--- 5322
B.KR.05.05CSR3          G---------------------A-AT----------G-------------G------------------......----A-------G-------------------------------------T-----------------------------------A--GG---- 5343
B.NL.00.671_00T36       -----------------------TA----------G--------------------------------G......----A---A------------------------------------------------------------------------------G---G--- 4883
B.RU.04.04RU128005      G-------------------G---A--A------A---------------C---------T--------......----A-----------A----------------------------CA---------C--------T-------------------------G-G- 4809
B.TH.00.00TH_C3198      A------A------------T---T---------A----------------------------------......----C---------------------------------------------------------C-------------------------------A 4532
B.UA.01.01UAKV167       -G-----A----------------A--A-C-------C---------------T---------------......----A---A---G---T---------------------------------A-----------A--A----------------------------A 4544
B.US.04.ES10_53         ----------------------------------A---------------------------------G......--------------------------------------------------------------A-----------------------A----G--- 5327
B.US.99.PRB959_03       ------------------------AGG---------------C--------------------------......----A--------------------------------------------------------------------------------------G--- 4538
C.AR.01.ARG4006         -A------------T---------AGA-------A-G----------------T------T----A---......--------A-------T-------------T----A--------------T--------------T--------------------ACTG-RA-A 4502
C.BR.04.04BR013         -A------------T---------AGA-------A---------------------T---T-------A......----A---A---------------------T-R--A-----G------A-T--------------T--G-----------G-----ACTG-GA-- 4782
C.BW.00.00BW07621       -----C--------T---------AGA-------A-T-------------------T---T--------......--------A-------A-------------T-------------------T--------------T-------------------AATTG-GA-- 4671
C.CN.98.YNRL9840        C----C--------T-----C---AGA-------A-G-----------------------T-------G......--------A-------A------------AT----A--------------T--------------T--------------------ATTG-GA-- 4502
C.ET.02.02ET_288        --G----A------T---------AGAA---------C----------------------T-------A......--------AA--G---A-------------T-A--A--------------T--------------T--------------------ATTG-GAG- 4520
C.GE.03.03GEMZ033       --C--C--------T-T-------AGA-------A--C----------------------T--------......--------A-------A-------------T----A--------------G--------------T--------------------ATTG-GA-A 4499
C.IL.99.99ET7           --C-----------T---------AGA-------A-G------G----------------T--------......--------A-------A-------------T-------------------T--------------T--------------------ATTG-GA-- 4499
C.IN.99.01IN565_10      -----C--G-----T-----C--TGGA-------A-G-----------------------T--------......--------A-------A------------GT----A-----------G--T--------------T-----A--------------ATTG-GA-- 4699
C.KE.00.KER2010         -----C--------T-T-------AGA-------A-G-------------------A---T-------G......--------A-------A-------------T----A--------------T--------------T-----T-------------AATTG--A-- 4499
C.MM.99.mIDU101_3       -----C--------T-----C---AGA-------A-G-----------------------T-------G......-----G--AA------A------------GT----A--------------A--------------T--------------------ATTG-GA-- 4669
C.MW.93.93MW_965        --C--C-----C--T---------AGA-------A-G-----------------------T--------......--------A-------A-------------T----G--------------T--------------T--------------------ATTG-GA-- 4505
C.SN.90.90SE_364        --------------T---------AGA-------A-G-----------------------T--------......--------A-------A-------------T--------------C----T--------------T--------------------ATTG-GA-- 4487
C.SO.89.89SM_145        -----C--------T---------AGA--C----A-G-------------------A---T-------G......--------A-------A-------------T----A--------------T--------------T--------------------ATT--GA-- 4539
C.TZ.02.CO178           --C--C------------------AGG-------A-G-----------------------T------GG......--------A---------------------T----A--------------T--------------T--------------------ATTG-GA-- 4499
C.UY.01.TRA3011         -A------------T---------AGA-------A-------C-------G-----T---T-------G......----A---A---------------------T----A--------------A--------------T--------T-----G---C-ACTG-GA-- 4489
C.YE.02.02YE511         --------G-----T---------AGA-------A-G-----------------------T-------G......--------A-------A-------------T-------------------T--------------T--------------------ATTG-GA-- 4493
C.ZA.04.04ZASK164B1     -A---C-----C--T---------AGA-------A-G-----------------------T----A--G......--------A---------------------T-A--A--------------A--------------T--------------------ATTG-GA-- 4732
C.ZA.05.05ZAPSK240B1    -----C--------T---------AGAA-C----A-G-----------------------T--------......--A--G--A-------A-------------T----G--------------T--------------T--------------------ACTG-GA-- 4722
C.ZM.02.02ZM108         -----C-A------T---------AGAA------A-------------------------T--------......--------A--------------------------------------------------------T--------------------ATTG-GA-- 5287
D.CD.83.ELI             --------------------------C--C----A------------------------------A--A......------C---------A---------------------------------A-----C----------------------------------G--- 4864
D.CM.01.01CM_0009BBY    -----------G--------G--------C----A-------------------------T--------......----A-C----GG-------------------------------------------C-----------------------------A-------- 4517
D.KE.01.NKU3006         --G-----------T------T--AA---C----A------------------------------A--A......------C-----G---A---------------------------------A-----C----------------------------------G--- 4523
D.KR.04.04KBH8          G-------G-----------C--AA-----------------C----------------------A---......-A--A---A---G-G------------------A--------------A-G-----C-----------G--------T--------ACTG--A-- 5277
D.TD.99.MN011           --------------------T--------C----A---------------G---------T----A--A......------C-----G---A----------------A-------G--------T-----C----------------------------------GA-- 4542
D.TZ.01.A280            --G-----------------C----A---C----A---------------------T--------A---......------C----GG-------------------------------------A-----C-----------------------------A----G--- 4519
D.UG.99.99UGK09259      --G-----------------CT---A--------A------------------------------A--A......--C---C-----G-G-G---------------------------------A-----C-----------G--------------------G-G--- 4523
D.YE.01.01YE386         --G-----------------CT--TA---C----A------------G--------T---T----A---......------C-----G---A---------------------------------T-----C----------------------------------G--- 4517
D.YE.02.02YE516         --------------------T--------C--C-AG--------------G--T------T----A--G......----A-C-----G---T----------------A----------------T-----C-----------------------------AC---G--- 4529
D.ZA.90.R1              --------------------C-----C-------A------------------------------A--A......-AG---C-----G---A-------------------------------A-A-----C-------CT--------------------A----G--- 4664
F1.AR.02.ARE933         ----------------T-------AGG---------G----------------------------A--A......----A---A-------T------C-----C--------------------T-----C-----G------------------------C-NGG--A 4612
F1.BE.93.VI850          -C--------------T-C-----AGG---------G----------------------------A--A......-T--A---A--------------------G-----CC-------------A-----C------------------------------C-GGG--A 4658
F1.BR.01.01BR125        C-G-------------T-------AGG--------GG----------------------------A--A......----A---A--------------C-----G--------------------T-----C------------------------------C-GGG--- 4763
F1.ES.x.P1146           -G------------T-T-------AGGA-C------G--------------------------AAA--A......----A-A-A---G-------------------------------------G-----C------------------------------C-GGG--A 4595
F1.FI.93.FIN9363        ----------------T-------AGA-------A--------------------------C---A--A......---GA---A--------------------G---A----------------T-----C--------A---------------------C-GGG--- 4647
F2.CM.02.02CM_0016BBY   ------------------------AGG---------G-----------------------T---A-A--...GATT---A---A------GT---C-----------A-------------A---T--------------T---------------------C-G--A-- 4511
F2.CM.95.MP257          ------------------------AGG---------G-------------------T--------A--A...GATT---A---A--C---T----C-----------A--------------------------------A--------------------AC-G--A-- 4530
F2.CM.97.CM53657        ------------------------AGG---------G-----C--------------------AA----...GATT---A---A------GT--GC-----------A-----------------T--------------A---------------------C-G--A-- 4511
G.BE.96.DRCBL           -----------------------TAGG--C----A-G----------------------------A---......----A-C-A---G---G---------------------------C-A---A-----C-----G--T--G------------------C-GGGA-- 5275
G.CM.01.01CM_4049HAN    ------------------------AGG-------A------------------------------A---......----TGC-AA--G---G-----------------------------AG--------C--------T-----T--T------------C-G-GA-- 4520
G.CU.99.Cu74            -GG---------------------AGA-------A------------------------------A---......------C-AA--G---G-----------------------------A---T-----C--------T--G----------------A-C-G-GA-A 4922
G.ES.99.X138            --C-----------T---------AGG-------A-G-----------------------T----A---......----C-C-A---G---G-----------------------------A---T--------------T--G------------------C-G-GA-- 4759
G.GH.03.03GH175G        ----------G-------------AGA-------A-G----------------------------A---......--C-CCC-A---G---A------------C---------------CA--C------C-----CA-T--GACT---------------C-G-GA-- 5358
G.KE.93.HH8793_12_1     ----------------T-------AGG-------A-G----------------------------A---......----A-C-A---G---GC---------------A------------A---T-----A--------T--G------------------C-G-GA-- 4718
G.NG.01.01NGPL0669      ------G-----------------AGG----------C------------------T---A----A---......----AGA-A---G---T-----------------------------A--C------A--------T--G------------------C-G-GA-- 4523
G.PT.x.PT2695           -C--------------T-------ATG-------A-G----------C------------T----A---......----A-C-A---G---G-----------------------------A---T--------------C--G------------------C-G-GA-- 5260
G.SE.93.SE6165          -----------C--------C---AGG---------G----------------------------A---......----C-C-A---G----------------G----------------A---------A--------T--G-----------G------C-G-GA-- 4715
H.BE.93.VI991           G-------------T--------TA--A-C----AG-C----------------------G--------......-----G--A--C--------------------------C--------------------------T--------------------AC-GG-A-- 4707
H.BE.93.VI997           G-------------T-----C--------------G------------------------T--------......-----G--A--T-----------C-----------------------------------A--C------------------------C-G--A-- 4642
H.CF.90.056             --------------T-T-------A----------G------------------------T----A--A......-----G--A--C-----C-----C---------A----------------A--------A--C-----------------------ACTG--A-- 4665
J.CD.97.J_97DC_KTB147   C-------------------C--AA---------A-G-----C-----------------T----ATG-......---GA-C-A---G-------------------A-----------------T------------------------------------CTG--A-- 4512
J.SE.93.SE7887          C-G-----G-----------C--AA---------A------------------------------A--G......----T---A--------------------------G--------------T-----C------------------------------C-G-GA-- 4632
K.CD.97.EQTB11C         CG------------------C--AA---------A----------------------------------......--AGA---A--GG----------------------------------G--A-----C------------------------------CT-----A 4510
K.CM.96.MP535           -A----------------------AGG-------A----------------------------------......--AGAG--A---G---------------------T------------G---------------------------------------CTG--A-- 4514
```

```
B.FR.83.HXB2             TTTTATAGACATCACTATGAAAGCCCTCATCCAAGAATAAGTTCAGAAGTACACATCCCACTAGGGGAT......GCTAGATTGGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGAGACTGGCATTTGGGTCAGGGAGTCTCCATAGAATGGAGGAAAAAGAG 5318
Vif                      _F__Y__R__H__H__Y__E__S__P__H__P__R__I__S__S__E__V__H__I__P__L__G__D__.__.__A__R__L__V__I__T__T__Y__W__G__L__H__T__G__E__R__D__W__H__L__G__Q__G__V__S__I__E__W__R__K__K__R
01_AE.CF.90.90CF11697    -------------T---------AG------A-G--------------------A-----------......--------A------G-------------------G-------A---------C-------T------------------C-G-GA-A 5267
01_AE.CN.05.FJ051        -C---------C--T---------AG------AGG-------------T-T-------------G......-G--A---A------G-----C--------A--G---------A---------C-------T-----------------A-C-G-G-CA 5328
01_AE.CN.06.FJ054        --------G-----T---------AG------AGG----------------T----------A--......----G--A------G-------------------G---------AG--T-----A-------T------------------C-G-GA-A 5343
01_AE.HK.x.HK001         -A------------T---------AG------AG-----C--------------------A--G......--------A------A-------------------G---------AG-------A-------T-----------------AC-G-GA-A 4680
01_AE.TH.01.01TH_R2184   --------------T---------AGA-----AGG----------------T---------A--G......---A--A------G-------------------G---------AG-------C-------T------------------C-G-G-GA 4518
01_AE.TH.02.OUR769I      -C------------T---------AG------A-G------------------------A--G......---A--A------G-------------------G--------CAG-------C-------T------------------C-G-GA-A 4517
01_AE.TH.90.CM240        --------------T---------AG------AGG---C-------------T--------A--G......-------A------G-------------------A---------AG-------C-------T------------------C-G-GA-A 4892
01_AE.US.00.00US_MSC1164 GC------------T---------AG------AGG----------A----T----------A--G......---T-C-A------G-------------------G---------AG-----------------T------------------C-G-GA-A 4523
02_AG.CM.02.02CM_4082STN -----C----------T----T-TAAG--C----A-GC---------------------------......--------A------G---------------A------------------------C-------T--GT--------------A-C-G--A-A 4517
02_AG.EC.x.ECU41         ----------------------TAGG-------A-G-----------------------------......--------A---G--GG--------------A------------------------CC------T--G---------------A-C-G-GA-- 4414
02_AG.GH.03.GHNJ196      --------------T-T-----TAGA-------A-GC----------------------------......--------A---G--G---C---------T--A------------------------C-------T--G----------------C-G-GA-- 5347
02_AG.NG.01.PL0710       -Y--------------------TAGG---------G-----------------------------......--------A---G--G-----------------------------------C-----C-------T--G---------------AC-G--A-- 4502
02_AG.NG.x.IBNG          ----------------T-----TAGG-------A-G--T--------------------------......----G--A---G--G--------------------------C------C-------C-------T--G---------------A-C-G--A-- 4843
02_AG.SE.94.SE7812       --C-------------------AG---------GG------------------------------......--------A---G--G---G--------------------------------------C-------T--G--T-------------A-C-G-GA-- 4690
02_AG.SN.98.MP1211       ----------------------TAGG-------A-G----------------T------------......-----C---AA--G--G-------------------GC--------------------------T--G---------T----CA-C-G--A-- 4515
03_AB.RU.97.KAL153_2     G------A---------------AGGA------------------------------T--------......---A-----------A--------------------------------------------------C-------C---------------A---G---- 4544
04_cpx.CY.94.CY032       --C-------------------TAGG--C----A-G---------------T-------------G......--------A---G--G-------------------GC----G---CA---------C-------T-------------------CTC-GA-- 4684
05_DF.BE.x.VI1310        -G--------------T-----TAGAA---------------------------------A--C......---A-C-----G------------------A------------------------------------------------G---------G--- 4700
06_cpx.AU.96.BFP90       AA-------------------C---AA--------A-------------------A---T----AAG-......---GA---A------------------------AA-------------A-A-----C-----C------------------------TTG-GA-- 5346
08_BC.CN.97.97CNGX_6F    -----C--------T-----C---AGA-------A-G-------------------T---A--G......--------A------A-----------GT----A--------------T--------------T----C---------------ATTG-GA-- 4498
09_cpx.GH.96.96GH2911    --------------------T---AA--------A-----C----------------G-----A-C-......---GA---A---G-G--------------------------G--------A-A------------------G-------------------CTG--A-- 4520
10_CD.TZ.96.96TZ_BF061   -C------------T---------AGA-------A------------------------------A--A......-----TC--A------A-----C----------------------------------A-------------T--G------------------ACTG--A-- 4697
11_cpx.GR.x.GR17         --G--C--G-----------G---AGG---------G----C--------------------A---G......-A--TG--A---G-G---------------------------------------------A---------------------G------------------CGG--A-- 4620
12_BF.AR.99.ARMA159      ----------------T-------AGG---------G----------------------------A-A-A......---GA---A---------------C-----G-----GC-------------------T------C------------------------------C-GGG--- 5323
13_cpx.CM.96.1849        --------------------C---AA--------A------------------------------......---A---A---G-G-----------------------------------------A-A-------------A--G--------T---------CTG--A-- 4723
14_BG.DE.01.9196_01      ----------------T-------GGG-------A-G-------------------T----A---......---A-C-A--GG---G---------------------------A---T-------------T--G-C------------------C-G-GA-- 4835
15_01B.TH.99.99TH_MU2079 --------------T---------AG-------AGG---C---C------T----------AA--G......--------A------A-------------------G---------AG--T--------------T------------------C-G-GAG- 4714
16_A2D.KR.97.97KR004     -GC---------------------ATAA--------G---------G-----------G-------GG......---A-A-AA--G--A-----------------CC---------A---------C----------G------------------C-GG-A-- 4681
17_BF.PE.02.PE02_PCR0155 -C--------------T-------AGGA-----------------------------------C......---A---A------G-------C----G-------------------------T------C----------C--------------C-GGGAG- 4508
18_cpx.CU.99.CU76        ----------------------TAGA-------A-G---C------------------G------......---C---A---G----------------------------------A---A-------------T--G------------------CTG-GA-- 4623
19_cpx.CU.99.CU7         G-----------------------AGGA------AGG----------------------------......---A---A---G---A--------------------------------A---------C-------T--G------------------C-G-GA-- 4544
20_BG.CU.99.Cu103        -GG-------------T-------AGA-------A-G---C--------G-------------A--C......---A-C-AA--G--G--------------------------------A---T-----------Y--T--G--T-------------A-C-G-GAGA 4784
21_A2D.KE.91.KNH1254     G-C---------------C-----AGA---------G--------------------G--G-A---G......---A-C-AA--G--G-----------------CT---------A---------C-------T--G------------------C----AG- 4523
22_01A1.CM.01.01CM_0001BBY-----------------------AGG-------A------------------------------......--------A---G--G-----------G---------------A---A-----C-------T--G------------------C-G--A-- 4526
23_BG.CU.03.CB118        -G-----A-------------G---A--A-------G-----C--------------A----------A......---A-----A----AG------------------------------------------T----------C-------------------------------- 4761
24_BG.CU.03.CB378        -G-----A--------T-------AGAA------A-G---C--------G-------------A--A......---A---AA--G--G------------------------------A---T-------------T--G--------------A-C-G-GAGA 4743
24_BG.CU.03.CB471        GG--------------T-------AGAA--------G---C--------G-------------AA-C......---A---AA--G-T-A----------------------------A---T-----G-------T--G--------------A-C-G-GAGA 4758
25_cpx.CM.01.101BA       --C---------------------AA---------------C-----------------------......---GA---A-C-G-------------------C-----------------A--------A--------G------------------CTG--A-- 4526
27_cpx.CD.97.97CDKTB49   --------G-----------C---AA---------------C-----------------------......---GT---A---G-G-------------------------------------A--------A--------G------------------CT--GA-- 4526
28_BF.BR.99.BREPM12609   G-G----A----------------A-----------G-----------------------------......---A---A---G---T---T-------------------------------------------------------------------------G---- 4684
29_BF.BR.02.BREPM119     --G-------------------A-AGA-------A--------------------------A---A......---A---A------T-----C-----G-----G-----------T------C----------------------------------C-GGG--- 4492
31_BC.BR.02.110PA        -R------------T---------AGA-------A-R--------------------T------G......--------A------A-----------T---A-------------T-------------T---------------G-----ACTG-GA-- 4752
32_06A1.EE.01.EE0369     --------------------C----AC-------A-----------------------------AAG-......---GA---A------------------------AA-------------A-----C-----C-------------------------TTG--A-- 4953
33_01B.MY.05.05MYKL007_1 --G-----------T-----------AA-------AGG-----------------T------A----A--G......---A-A-A-C-----G-------------------G---------AG-----------------T------------------C-G--A-A 4697
34_01B.TH.99.OUR2478P    --------------T---------T-AG-------A-G----------------------------A--G......--C-----A-----CG-------------------G---------AG--T-----AC-------T------------------C-G-GA-A 4520
35_AD.AF.05.05AF095      G------A----------------AGG-------A-G---------G-----------T----------......---A-A-AA--G--G-----------------------------A-----------------T--G------------------CT---A-- 4511
36_cpx.CM.00.00CMNYU830  ------------------------AGA-------A-G-------------------T---------......--A-----------G------------------------------------------------T--G-----T-----------AAC-G-GA-A 4523
37_cpx.CM.00.00CMNYU926  ------------------------AGGG--------G------------------------------A......---T-G--A---G---G-----------A-----GC----G--------------C-------T--G------------------CTG-GA-A 4508
38_BF1.UY.05.UY05_4752   -G-----A--------T-------AGGA--------G------------------T--AA---A......---C---A-C---C--------C-----G--------------------T------C----------------------------------CTGGG-G- 4729
39_BF.BR.03.03BRRJ103    G-C---------------------AA-A--------------------------------------A......---A---A---G-------------------------------------A---------A-------------------------------------G--- 4807
40_BF.BR.04.04BRRJ115    ------------------------A---------AG-----------------------------A......---GA---A---G------------------C--------------T--------------A--------------------------------G--- 4870
42_BF.LU.03.luBF_05_03   -A----------------------A----C-----A------------------------------......RG------A------A-------------A------------------------------------------------------------------------- 4856
43_02G.SA.03.J11223      ------------------------AGGA------A-G--------------------M--R-----......---C---AR--R---R--------------R--G-----------G--Y-----A-----C--T--G----------------R-C-G--A-A 4843
U.CD.83.83CD003          G------A---C------------A--A-----------------------------T------A--......-----A-A---G-----G--------------------------------------A--------------------------------C-GGGA-- 4813
U.CD.90.90CD121E12       G------A---C------------AT-A-C---------------------------T---A-A-C......-----A-A----------G-------------------------------T-------A--------------------------------C-GGGA-- 4517
U.GR.99.GR303            G--------------------C---AA-----------------C-----G--------------A-A-A......---A---A---G-G-------------------------------------G---------------A--G-----------------TG--A-- 4589
U.NL.95.U_NL_95_H10986_D1 -----C------------------AGA---------G-----------------------------AA--C......---A---A-------------C-------------------G-------T----------------------------------G--C-G-GA-A 5203
CPZ.CD.90.ANT            AAG-----------T-------ATGA---C----AG-A-G-AGA---------T--A------CCAACATTGGATAAA-A-------G-G---GTG-TC-----AT-A--ATGT------C-CCCA-----C--A--A--T--T--GAG-------------ATGTGGA-A 4763
CPZ.CM.01.SIVcpzCAM13    --C--C-----C----------CAT---A--------G-TTC------G---------T--TT-TAAA---...AATT-A-A---AA---C-----C-------C-T-AACC-----------------CG---------A-----T-------------C-G-------TGGGATC 4901
CPZ.CM.05.SIVcpzEK505    --C-----------T-------CAGA-A------A-------------A-----------T----A--A......-A-A---A------GTT--C--------AT-AATGC----------GCCA--------A--A--T-----A---------------AC--GGA-T 4890
CPZ.CM.05.SIVcpzLB7      --------------T-----G--TGAAA------A-G-----------A----T--------------G......-C-A---A------GT---T--------A--TACAC--------C-G--------C--A-----T--G--A--T--C--G----AAC-G----A 4873
CPZ.CM.05.SIVcpzMB66     --C--------C------------AA-A------A-----------------------------G--A--A......---A---A-----------------------A--CACA-------------C-------C-A--C--T--G--AAGT-----------AC-G-GA-- 4864
CPZ.CM.05.SIVcpzMT145    --C-------------------TG--A-CA-C--T--G--GGC----C-GA--T-T--T---A--AATC-G......----AGC-AA-TG-C----------------A--AATGC-------------T------C-----A--T-----AG-A------------AG---GAG- 4860
CPZ.GA.88.GAB1           -----C--------T-----TCA----A------A-G-TGC---T---A-------A--TT-CA-A----...TATT---A---AA--G-G-----C-------CCT-AAGCC----------G-C----------A--C--T-------------TC---------CT-GGATC 5382
CPZ.TZ.00.TAN1           AAG-----------T-------CTGA-ACA----A--G-GCAGGG---A-------A--T---ACA--A...AGAT-A-A---A--GG-TTT-CAT-----------AGCCTGT---------CCA------C-A-----T--CA-AGGAT------------AC--GGA-A 4965
CPZ.US.85.CPZUS          --C--C--G-----T-----G---A--A----C---G---A-------------T-----TA---ATA-A......-A-A-A-A--G--T-T---G----------GT--A-A---------CA--C------C--A--G--T-----A--A------------AC--GGA-C 5378
CPZ.CM.x.SIVcpzDP943     --------G--C--T--------TT--A-C--C---G---------A----T-----TT---AAA-A......-A-A---A--GG-C--C--C---------AT-AATGC------GCAT--A-----CC-A--A--T-----G-------------AT-TGG-TC 4904
CPZ.CM.x.SIVcpzMB897     -----C--------T--------TGAGA------A-G-T---------A-T-----T---G----A---......-C-A---A------GT---C--------A--CATGC---------------------A--A--T--G-----T------------AC-G--A-- 4884
N.CM.02.DJO0131          -A------------T-------CAAA--------A------------------T-----G----TCTG......-A--G--A--G-----C--T--------A--AACA-----------GTCT------C-A--A--T-----A-------------AACT-GGA-A 4824
N.CM.04.04CM_1015_04     -A------------T------TCAAGG-------A--C---------------T-----G----TC-G......-A-C---A--G---GTC--T--------AT-AACA---------CAGCCT------C-A--A--T-----A--------------ACT--GA-A 4823
N.CM.04.04CM_1131_03     -A------------T-------CAAGG-------A------------------T-----G----TC-G......-A-----A--G---GTC--T--------A--AACA---------CAGCCT------C-A--A--T-----A--------------ACT--GA-A 4822
N.CM.95.YBF30            -A------------T-------CA-A---C----A------------------T-----G----TC-G......-A-----A--G-C-GTC--T--------G--AACA---------CAGTCT------C-A--A--T-----A--------------ACT--GA-A 4912
N.CM.97.YBF106           -A------------T-------CAAAG-------AG-C-----------R--T-----G----TCMG......-A-R---A--G---GTC--T--------A--AACA---------CAGCCT------C-A--A--TR----A--------------AC--GGA-A 4910
O.BE.87.ANT70            -GG--C--------T------TC-AGAA--------G-T------AGT---T-T--T---G------TG......--CCATG-A---G----C-----------AT--ATGC----G--------TGAA--------A--T--G--TAGT-----------C-TT-T----A 5372
O.CM.91.MVP5180          CG------G-----T------TC-AGGA------A-G-C-----G-CG--GT-T--T---G---CA--A......---GATA-A--GG-C--C-----------AT-AATGC----G--------GGAA--C-----A--T--G--TAGT-----------CAAT-C---GA 5347
O.CM.96.96CMABB637       -A------------T-------C-AGAA----------C--C-----GT---T-T--T---G----TACA......----ATA-A--GG-G--T-----------AT-AATGC----G--------TGAA-----A--A--T-----AG-------------CAAT-T--A-A 4802
O.CM.98.98CMA104         -AC--C-------------MTCTAGAA--------G-C-------GT---T-T--T---G-----C-G......---T-TA-A---G-G--C-----------AT-AATGC----G--------GGAT--G--A--A--T-----TAGT-----------CA-T-C----A 4804
O.CM.99.99CMU4122        CA---C-A---C--T-T----TT-AGAA--------GCC-------GT---T-T--T---G----ATG......---T-TG-A---G-G--C--C----------AT--ATGC----G-----G--TGAA--------A--T--G--TAGT-----------CA-T-C----A 4803
O.FR.92.VAU              -A-----A------T------TCTATAA-------GG-C------AGG---T-T--T---G---TCCA......---T-TA-AA-GG-G--C-----------AT-AATGC-------------GGA---G-----A--T-----AGT-----------CA-TGC-G--A 4884
O.SN.99.SEMP1299         -A---C-----C--T-T----TCTAGAA--------G-C-------GT---T-T--T---G------GG.....C--T-GA-A---G-G--C-----------AT--ATGC----G--------TGAA--------A--T--G--TAGT--------------T-C----A 5371
O.US.99.99USTWLA         -A------------T-------CTAGAA----------C-------GT--CT-T--T---G----A-GA......--C-TGA-A--GG----T-----------AT-AATGC----G--------TGA----C-A--A--T--G--TAGT-----------GA-T-C----A 4801
```

```
B.FR.83.HXB2               ATATAGCACACAAGTAGACCCTGAACTAGCAGACCAACTAATTCATCTGTATTACTTTGACTGTTTTTCAGACTCTGCTATAAGAAAGGCCTTATTAGGACACATAGTTAGCCCTAGGTGTGAATATCAAGCAGGACATAAC...AAGGTAGGATCTCTACAATACTTGG 5485
Vif                        __Y__S__T__Q__V__D__P__E__L__A__D__Q__L__I__H__L__Y__Y__F__D__C__F__S__D__S__A__I__R__K__A__L__L__G__H__I__V__S__P__R__C__E__Y__Q__A__G__H__N__.__K__V__G__S__L__Q__Y__L__
A1.GE.99.99GEMZ011         G------------A----------C--G--G------------------C----T-----------------A-----C-----G--A---A--G---------------------A----A-C----C-C-------C-G-...------------------------- 4669
A1.KE.00.KER2008           ------------------T--------G---------------------C----T---A-------------A-----C-----G--A--------------AG-------------T------------------------...--------------------T---- 4684
A1.KE.00.KNH1144           T--C--------------T-----C--G--------------------------T----------------G------C-----G--A---A--------G-AT-------------A------------A-------C-G-...--------------------T--A- 4690
A1.KE.00.KSM4024           -------------A----T-----TT-G--------------A------C-C--T----CT------A----------C--C--G--A---A----------AG-------A-----A----------C-A-----------...--------------------T--A- 4681
A1.KE.00.MSA4069           ------T-----------T-----C-AG--------------------------------------------A-----C-----G--A---A----------AG-------------------C--C---A-----------...--------------------T--A- 4684
A1.KE.00.NKU3005           -------------C--ACT-----CA-G---------------------C----T---A-------------------C-----G--A---A-------G--AG--------------------------A-------C---...--------------------T--A- 4684
A1.RU.00.RU00051           G--Y---------A----------C--G-----------G--------------T-----------------------C-----G--A---C--------G-A--------A--------------------------C---...-----------------R------- 4786
A1.RW.93.93RW_024          -------------A----T-----C--G---------------------C----T-----------------A-----C-----G-GA---A----------AG-------------T----------CA------------...C----------C--------T--A- 4687
A1.SE.95.SE8891            -------------A----T-----C--G---------------------C----T---A-A-----------------C-----G--A---A--------G-A---------------------------------------...-----------C--------T---- 4674
A1.SE.95.UGSE8131          -------------A----T-----C--G---------------------C----T---A-------------------C-----G--A---A-------G-GAG-------------T------------A-----------...--------------------T--A- 4856
A1.TZ.01.A173              ------------GA----T-----C--G---------------------C----------------------------C-----G--A---A----------AG-------A-----------C------------------...--------------------T---- 4684
A1.UA.01.01UADN139         G------------A----------C--G--G-----------------------T-----------------A-----C-----G--A---A--G-------A------------------A-C----C---------C---...---------C--------------- 4684
A1.UG.92.92UG037           -------------A----------C--G---------------------C----T---A-------------------C-----G--A---A-------G--AG-------------------T------A-------C---...------------T-------T--A- 5501
A1.UG.99.99UGA07072        -------------A----T-----C--G--------------C------C----T---A-------------------C-----G-GA---A----------AG-------------T------------------------...--------------------T---- 4684
A2.CD.97.97CDKTB48         G--C--T------A----------T--G--------T-----C-----------T-----------------A-----C-----G-GA---A--------G-A--------G-----A-----------------------T...--------G-----G-----T---- 4825
A2.CY.94.94CY017_41        G---CAT------A----------T--G--------T-----C-----------T-----------------A-----C-----G--A---A--A-----G-A--------T------------------------------...--------G-----G-----T---- 4842
A.SN.01.DDI579             -------------A----T-----C-----------------------AC----T-----------------------C-----G-GA---A----------A--------A------------------A-------C---...------------T-------T---- 4683
A.SN.01.DDJ369             -------------A----T-----C------------------------C----T---A-------------------C-----G--A--------------A--------A-----T------------------------...------------T-------T---- 4686
A.SN.96.DDJ360             -------------A----T-----C-----------------------AC----T---A-------------------C-----G--A---A----------A--------A----------------C-A-----------...------------T-------T---- 4683
A.CD.02.02CD_KTB035        G-----T------A----------C--G---------------------C----------A-----------------C-----G--A---A--------A----------------A-----C-----------------T...--------------------T---- 5479
A.CD.97.97CD_KCC2          G-----T-----GA----------C--G--------------------AC----T---AC------------A-----C-----G--A--TA------------G-----------A-------------------------...--------G-----------T---- 5481
A.CD.97.97CD_KTB13         G-----T------A----------C--G-----T---------------C----------------------------C-----G--A---A--------------------------------------------------...--------------------T---- 4689

B.AR.04.04AR151516         ------T-----------------CA----------------------------T-----T-----------------------G--A--------------T--------A-----T------------------------...-----------C--------T---- 4707
B.AU.87.MBC925             -----------------G------C-----------------------------------T-----------A------------------C----------T-G------T------------------------------...--------G--------G------- 5496
B.BO.99.BOL0122            ----GCA---------AC------CA----------------------------T-----T------G-----------------------A---------GT--------------T-----C-T--G-------------...-----------------G------- 4705
B.BR.03.BREPM2012          ------------G-----------T--------------------C--------T---A-T-----------A--------------T---A---------G---------T----A-------------------------...-G---------C-----G------- 4910
B.CA.97.CANB3FULL          ------------------------C--------T--------------------------T-----------A--------------C---C---------GT--------T-----T-----------------------T...--------G--------G------- 4774
B.CN.05.05CNHB_hp3         --------------C---A---AGC-----------------------------T-----T-----------A-----C--------A---A-T-------GTG--------T----T--CA-T----C--A--------G-...-----------C-----G------- 5484
B.CO.01.PCM001             ------------------------C-----------------------------T-----T-----------A--------------T---A--------A--T-------G------------------------------...-------A---------G------- 4690
B.GB.83.CAM1               -----A------------------C--------------------------C--T-----T-----------A------------------A--G------GAT--------------------------------------...-----------------G------- 5486
B.GB.86.GB8                ------A----------------GC----------------------A------T-----T-----------A--------------T---C---------GT-C------T-----T---A--------------------...------------------------- 5505
B.GE.03.03GEMZ010          ------------------------CTC---------------------A-----T-----T-----------G-----------G-GA---A--A------GTT-------T------------------------------...--------G--------G------- 4690
B.IT.05.SG1                ------------------------C-----G----G------------------------T-----------A--------------T--AA---------GT-G------T---TTA-----G------------------...-----------------G------- 5287
B.JP.05.DR6538             ----------------------A-C-----------------------------T-----A-----------AG-----------C-TA--A---------GT--------T----------------------------C-...------------------------- 5489
B.KR.05.05CSR3             -----AT--------------AAC-----------------------G------------T-----------------C------------A----------T--------T-----T------------------------...-----------------G------- 5510
B.NL.00.671_00T36          -----C------------------C---------------------G-------T-----T-----------A--------------T---A---------GTT--------------------------------------...--------G--------G------- 5050
B.RU.04.04RU128005         -----A------------------C-----------------------------------T-----------A--K--Y--------T---A-A--------T--------T----------------C-A-----------...-----------------G-----A- 4976
B.TH.00.00TH_C3198         C----A-----------------GC--------------------C--------T-----T-----------A--------------T---A-TG------GTG-------------T-----T------------------...--------------G--G------- 4699
B.UA.01.01UAKV167          -------------A----------CT----------------------A-----T-----T----------GG-----------G-GT---A---------GTT-------T-----T------------A-------C-G-...------------T----G------- 4711
B.US.04.ES10_53            ------------------------------------------------------------T-----------A--------------T---A----------G--------------------G------------------...-----------C-----G------- 5494
B.US.99.PRB959_03          ---------------------CA-C---------------------A-------------T-----------A------------C-T---A----------T-G------T-----T------------------------...-----------------G------- 4705

C.AR.01.ARG4006            -----------------------GC--G------------------A--C----T-----T------A----------C-----G--A---C----------T-G------------------------------------T...------------T--------C--- 4669
C.BR.04.04BR013            ---C--T-----------------C--G------------------ACAC----T-----T------G----------C-----G--A--------------T-G------------------C--C--------------T...---------------------C--- 4949
C.BW.00.00BW07621          ---C-------------------GC--G--------G---------A--C----T-----T------G----------C--------A--------------AG------TT-----------C-----------------T...--------------------T---- 4838
C.CN.98.YNRL9840           -----A-------A----A----GC--G--------G-----------------T-----T------G----------C-----G--A---A-C----------------TT-----------T-----------------TAAT---------------------C--- 4672
C.ET.02.02ET_288           -----C----------------TGC--G------------------A--C----T-----T------G----A-----C-----G--A---A----------G-G------------------C-----------------T...------------------------- 4687
C.GE.03.03GEMZ033          -----------------------GC--G--------G---------A--C----T-----T------G----------C--C--G--A---A----------T-------TT-----------C-----------------T...------------------------- 4666
C.IL.99.99ET7              -----------------------GT--G--------G-----C--CA--C----T-----T------G--A-A-----C--------A---A----------T-------TT-----------C-----------------T...------------------------- 4666
C.IN.99.01IN565_10         C-----T-----------A----GCT-G--------G---------A--C----T-----T------G----------C--------A---A------------------TT-----------C-----------------T...-----------------G------- 4866
C.KE.00.KER2010            G----A-----------------GC--G--------G---------A--C----T-----T------G----------C--------A---A------------------TT-----------------------------T...---------------------C--- 4666
C.MM.99.mIDU101_3          ------------------A----GC--G--------G-----------------T-----T------G----------C-----G--A---A-C----------------TT-----------T-----------------TAAT---------------------C--- 4839
C.MW.93.93MW_965           -----------------------G---G------------------A--C----T-----T------G----------C--------A---A----------T--------------------C------------------...------------------------- 4672
C.SN.90.90SE_364           G---------------------CGC--G--------G-----C---A--C----T-----T------A----------C--------A---A----------T-------TT-----------------------------T...------------------------- 4654
C.SO.89.89SM_145           -----------------------GC--G--------G---------A--C----T-----T------G----------C--------A---A------------------TT-----------C-----------------T...------------------------- 4706
C.TZ.02.CO178              -----------------------GC--G--------G---------A--C----T-----T------G----------C--------A---A------------------TT-----------T----C------------T...--------------G---------- 4666
C.UY.01.TRA3011            -------------A---------GC--G------------------A--C----T-----T------G----------C-----GG-A---A----------T-G------------------C--------G--------T...C------------------------ 4656
C.YE.02.02YE511            -----------------------GT--G--------G-----C--CA--C-C--T-----T------A----------C--------A---A----------T-------T------------T----------------CT...------------------------- 4660
C.ZA.04.04ZASK164B1        -----------------------GC--G--------G---------A----C--T-----T------G----------C--------A---A------------------TT-----------C-----------------T...---------------------C--- 4899
C.ZA.05.05ZAPSK240B1       ------------------T----GC--G-----T------------A--C----T-----T------G----------C------C-A---A----------T--------------------C--C---------------...--------------G---------- 4889
C.ZM.02.02ZM108            -----------------------GC--G--------G---------A--C----T-----T--C---G----------C--------A---A----------T-----C--T-----------C--C--------------T...------------------------A 5454

D.CD.83.ELI                -----------------------GC--G------------------A-------T-----T-----------A--------------A---A--------G-T--------T-----------G------------------...-----------C-----G--T---- 5031
D.CM.01.01CM_0009BBY       -----------------------GTT-G------------------A-------------T------G-------------------A---A----------T-G------T-----------T------------------...-----------CT-------T---- 4684
D.KE.01.NKU3006            -----------------------GC--G------------------A-A-----T-----T------G----A--------------A--------------T-------CT---------A-T------------------...--------------------T---- 4690
D.KR.04.04KBH8             G--C--A------A---------GC--G-----T------------A--C----------T-----------------C-----G--A---A----------T-G------------------T------------------...-----------------G--T---- 5444
D.TD.99.MN011              -----------------------GC--G------------------AC-C----------T--------------------------A---A----------G-G------T-----------T------------------...-----------CT----G--T---- 4709
D.TZ.01.A280               -----------------------G---G------------------A-------T-----T------A----------C-----G--A--AA-------T--T-----G--T-----------T------------------...------------T----G--T---- 4686
D.UG.99.99UGK09259         -----------------------GC--G------------------T-A-----T-----T------G----A-----C-----G------A----------G-------CT---------A-T--------------C---...------------T-------T---- 4690
D.YE.01.01YE386            -----------------------GC---------------------A-A-----T------------G----A-----------G--A---A-----------G------CT---------A-T------------------...------------T----G--T---- 4684
D.YE.02.02YE516            ----------------------AGC--G------------------ACA--C--------T-----------A--------------A---A---------TT-G------T-----------T------------------...-----------CT----G--T---- 4696
D.ZA.90.R1                 -----T-----------------GC--G------------------A-A-----T-----T-----------A--------------A---A----------T-G------T------------------------------AAC-----------CT----G--T---- 4834

F1.AR.02.ARE933            C-----G------A---------GC--G-----------G--C---A-A--C-CA-----------------A-----C-----GGGA---A----------T-G-A--------------A-C------------------...-G-T----------G-----T---- 4779
F1.BE.93.VI850             G-----G------A---------GC--G--G--------G--C---A-A-----T-----T-----------A-----C-----G--A---A----------T-G-A--------------A-C-----------------T...-----------C--------T---- 4825
F1.BR.01.01BR125           ------G------A---------GC--G-----------G--C---A-A-----T-----T-----------A-----C-----G--A---A----------T-A-A--------------A-C------------------...C-----------T------------ 4930
F1.ES.x.P1146              G------------A----------C--G-----------G--C---A-A-----T-----T-----------A-----C-----G--A---A--------T-T-G----------------A-C------------------...------AA---C------------- 4762
F1.FI.93.FIN9363           G-----G------A---------GC--G-----------G--C---A-A-----------T-----------A-----C-----G--A---A----------T-G-A----------------C------------------...-----------C--------T---- 4814
F2.CM.02.02CM_0016BBY      G-----G-----------T----GCT-G------------------AAA-----T-----T-----------A-----C-----G--A---A----------T-G------T---------A-C------------------...-----------C--------T---- 4678
F2.CM.95.MP257             G-----G----------------GCT-G---------------------C----T-----T-----------------C-----G--A---A----------G-G----------------A-T------------------...-----------C--------T---- 4697
F2.CM.97.CM53657           ------G----------------GCT-G---------T--------T-------------T-----------A-----C-----G--A---A----------G-G------T---------A-C------------------...-----------C--------T---- 4678

G.BE.96.DRCBL              ------T-----------T-----C-----------C------------C----T---A-T-----------A-----C-----G--A---A----------A-C------G-----T----------C------G-----T...------------T-------T---- 5442
G.CM.01.01CM_4049HAN       ---C--T------A----T-----C-----------T--G--------------T-----T-----------A-----C-----G--A---A--------G-G--------T-----------------------------T...-----------C--------T--A- 4687
G.CU.99.Cu74               -------------A----T-----C--------------G---------C----T---A-A-----------------C------C-A---A--------G-AG-------A--------------------------C--G...--------------------T---- 5089
G.ES.99.X138               -------------A----T-----C------------------------C-C--T-----------------------A-----G-----AA--------AGA--------T----------------------------CT...--------------------T---- 4926
G.GH.03.03GH175G           -----AT------A----T-----C-----------C--G---------C----T-----T-----------A-----C-----G--A---A--------G-G-----------------------------------C--T...--------------------T---- 5525
G.KE.93.HH8793_12_1        -------------A----T-----C--------------G---------C----T-----------------A-----C-----G--T---A--------AGA--------A-----------------------------T...--------------------T---- 4885
G.NG.01.01NGPL0669         G------------A----T-----T--------------G---------C----T-----T------A----------C-----G--A---A--------G-AT-------A-----------------------------T...--------C--C--------T---- 4690
G.PT.x.PT2695              -------------A----T-----C-----------------------AC----T-----------------------A-----G-----A---------GGA--------T----------------------------CT...--------------------T---- 5427
G.SE.93.SE6165             ------A-----------G-----C-----------T--G---------C----T-----T-----------------C-----G--A---A----------A--------------------------------------T...C-------------------T---- 4882

H.BE.93.VI991              G------------A----------C--G-----T------------T--C----T-----T-----------------C-----G--A--TA----------T-G------T--C-TT-----C----------------GA...------------T-------T---- 4874
H.BE.93.VI997              G----------------------GCT-G------------------ACCC----T-----T-----------A-----C-----GGGA---A---------GAG-------T------------------------------...C----------------G--T---- 4809
H.CF.90.056                G-----------------G----GC--G------------------A--C----T-----T-----------A-----C-----G--A---A---------GTG-------A---------A-C----C------------A...C--------A--------------- 4832

J.CD.97.J_97DC_KTB147      G-----A----------------GC--G------------------A--C--C-------T-----------AG----C-----G--A---A----------T---------------------------------------...------------T-G------C--- 4679
J.SE.93.SE7887             G--C--A----------------GC--G------------------A---G---------------------------C-----G------A----------A--------------------C--C---------------...--------------------TC--- 4799

K.CD.97.EQTB11C            G--C--A-----------T----GC--G------------------A-A-----T-----T-----------A-----------G--A---C----------T-G------T----------------------------C-...C----------CT----G--T---- 4677
K.CM.96.MP535              G--C--A-----------A-----C---------------------A-A-----T-----T-----------A-----AG----G------A----------T-G------T-------------G----------------...-----------CT-------T---- 4681
```

```
B.FR.83.HXB2                ATATAGCACACAAGTAGACCCTGAACTAGCAGACCAACTAATTCATCTGTATTACTTTGACTGTTTTTCAGACTCTGCTATAAGAAAGGCCTTATTAGGACACATAGTTAGCCCTAGGTGTGAATATCAAGCAGGACATAAC...AAGGTAGGATCTCTACAATACTTGG 5485
Vif                         __Y__S__T__Q__V__D__P__E__L__A__D__Q__L__I__H__L__Y__Y__F__D__C__F__S__D__S__A__I__R__K__A__L__L__G__H__I__V__S__P__R__C__E__Y__Q__A__G__H__N__.__K__V__G__S__L__Q__Y__L__
01_AE.CF.90.90CF11697       -----------------T-----C----------G--AG---------C-A--T--------------------C-----G--A---A-G--------AG-------A---------------C-A----------...-----------G--------T---- 5434
01_AE.CN.05.FJ051           -----A-------A---T-----C-------------G-------AAC-A--T--------------------C-----G-GA---A----------AG-------AAG---------T----C-T-----------...-----------C--------T---- 5495
01_AE.CN.06.FJ054           -------------A---T-----C-----G-------G--------AC-A--T-------------------------G-GA---A----------AG-------A-A--------------C-T-----------T...-----------C--------T---- 5510
01_AE.HK.x.HK001            C------------A---T------G---------G--G--------AC-A--T-----------G----------C-----G-GA---A--C-------AG-------ATA--A-----------C-T-----------...---------------------------- 4847
01_AE.TH.01.01TH_R2184      -------------A---T-----C----------G--G--------AC----T--------------------C-----G--A---A----------AG---A---A-G--A----------C-T----------G-...C----------CT-------T---- 4685
01_AE.TH.02.OUR769I         C------------A---T-----C-------------G------A-AC-A--T--------------------C-----G-GT---A----------AG-------AAG------------C-C-T-----------...-----------C--------T---- 4684
01_AE.TH.90.CM240           -------------A---T-----C---------A---G--------AC-A--T---G-----------------C-----G--A---A----------AG-------A-G--------------C-T-----------...-----------C--------T---- 5059
01_AE.US.00.00US_MSC1164    C----A-------A---T-----C-------------G--------AC-A--T--------------------C-----G--A---A----------AG-------A-G--------------C-T-----------...-----------C--------T---- 4690
02_AG.CM.02.02CM_4082STN    -------------A---T-----C-----------T--G-----C---C-C--T--------------A-----C-----G--A---A----------AG------------A----------------------T...C----------GT-------T---- 4684
02_AG.EC.x.ECU41            -------------A---T--C--C-------------T-G-----CT--C---T--------------A-----C-----G--A---A-----------G-------A---A----------------------T...-----------G--------T---- 4581
02_AG.GH.03.GHNJ196         G------------A---T-----C-----T-------G-----C--------T--------------A-----C-----G--A-T-C----------AG-------A-----T-------------------C-GT...-----------G--------T---- 5514
02_AG.NG.01.PL0710          -------------A---T-----C-------------G-----C---C----T--------------------C-----G--A---A--G-----G-AC-------A-------------------------T...-----------G-----G--T---- 4669
02_AG.NG.x.IBNG             -------------A---------CT------------G-----C--------T---A-T--------------C---------A---A--------G-A--------A-------------------------T...-----------G--------T---- 5010
02_AG.SE.94.SE7812          ------------GA---T-----C-------------G-----C---C----T--------------------C-----G--A---A-----------A-------A----A---------------------CT...-----------G--------T---- 4857
02_AG.SN.98.MP1211          -------------A---------T-------------G-----C---C----T--------------A-----C-----G--A-----------AG-------A---A--------------------T...-----------G--------T---- 4682
03_AB.RU.97.KAL153_2        ----------------------A-C----------------------------T-----T-----------A------------T---A----------T-G-----------T----------G-------------...-----------------G-------- 4711
04_cpx.CY.94.CY032          ---------------G--T-----C--G-----------------A--C----------T-----------A-----C-----G--A---A---------T-G------T----------------------------T...-----------CT--------C--- 4851
05_DF.BE.x.VI1310           ----------------------AGC--G-----------------G-------T-----T-----------A-----C---------T---A---------GT--------T----------G-T-----------------...-----------CT-----G--T---- 4867
06_cpx.AU.96.BFP90          ------A---------------AGTA-G-----------------A-AC----------T-----------A--------C--G--A--------------T-G-------------------T-------------------...---------------------TC--- 5513
08_BC.CN.97.97CNGX_6F       C----A-------A----A---GC--G--------G---------A-------T-----T-----G-----------C-----G--A---A-----------G-----TT----------C--------------TAAT------------------------ 4668
09_cpx.GH.96.96GH2911       G--C--G-------------CAGC--G------------------A-AC----------T-----------A-----C-----G--A---A--A-----G-T-G----------------------------------T...------------T--------TC--- 4687
10_CD.TZ.96.96TZ_BF061      ------T------A--------GC--G------------------A--C---T-----T-----G----A-----------G--A---A---------AG------CT---------A-T---------------A...C-----------T--------T---- 4864
11_cpx.GR.x.GR17            G--C--A-----------------G--------------------A--C-C--------T-----G----A-----C-----G--A---A---------T-G------T------------------C---------G-----...------------T--------T---- 4787
12_BF.AR.99.ARMA159         G-----G------A--------GC--G-------G---G--C---A-A-----T-----T-----------A-----C---------A---A---------T-A-A--------------A-C------------------...---------------------T---- 5490
13_cpx.CM.96.1849           ---C--A---------------GC--G------------------AC-C----------T-------------------C-----G--A---A---------T-G-------------------C----------------T...---------------------TC--- 4890
14_BG.DE.01.9196_01         C---------------C-------C--------------------------C--------T-----------AA----------G--T---A-----------T--------A-----T----------G------------...---------------------------- 5002
15_01B.TH.99.99TH_MU2079    -------------A---T-----------------G---G--------AC-A--T-----------------A-----C-----G-GA---A----------AG-------ATA---T-----T-----C-T----------...-----------C--------T---- 4881
16_A2D.KR.97.97KR004        ---C--T------A----------T-----G-G---T-----C-------C----T--------------------C-----G-----A--------A-A---------A-------------------------------...-------------T-G---------- 4848
17_BF.PE.02.PE02_PCR0155    ------G------A--------GC--G-----------G--C---ACA-----T-----T-----G----A-----C-----G--T---A----------T-G-A--------------A-C------------C--G...---------------------------- 4675
18_cpx.CU.99.CU76           G--C--A-----------T-----C--G-------------------------T-----------------A-----C-----G--T---A----------T-G------T-----------------------------T...------------T-------TC-T- 4790
19_cpx.CU.99.CU7            -------------A---T-----C--------T-----G--------C----T-----T--------------------C-----G--A-------------AG-------A---------------------------T...---------------------T---- 4711
20_BG.CU.99.Cu103           -------------A---T-----C-------------CA-G-------------T---A-T------------------C---------A---A--------G-A--------A-----------T---C-----------T...---------------------T---- 4951
21_A2D.KE.91.KNH1254        G-----T------A----------T--G-----T--T-----C-------------T---------------A-----C-----G--A---A--------AGAT-------T----A---------------------C-G-...--------------G------T---- 4690
22_01A1.CM.01.01CM_0001BBY------T-----------T-----C--G------------------A-----------T---------------------C-----G--A---A----------A---------------T------------A-----------T...-----------C--------T---- 4693
23_BG.CU.03.CB118           ----------------------GC--------------------AA-----T-----T-----------A------------GT---A----------T-G-----T-----T--------------------------T...---------------------T---- 4928
24_BG.CU.03.CB378           -------------A---T-----C-------------G-------A-C----T--CA-A-----------T--C--C--------A---C--------G-A-A------A-----T--C--G-----C-----------GT...---------------------T---- 4910
24_BG.CU.03.CB471           -------------A---T-----C-------------G-------A-C----T--CA-A--------------------C---------A---A--------G-ACA-A----A-----T-----G--------------T...---------------------T---- 4925
25_cpx.CM.01.101BA          G--C--G---------------GTT-G-------G---------A--C----------T-----A----------C-----G--A---A---------GT--------T-----T----------------------G...--------C-----------T---- 4693
27_cpx.CD.97.97CDKTB49      G--C--A---------------GC--G-------G---------AC-C----------T-----------A-----C-----G--A---A--------T-T-G------T----A---------C----------------T...---------------------TC-A- 4693
28_BF.BR.99.BREPM12609      ------------------------C-------------------------T--C--T-----------A-----------G--T---A---------GT--------A----------T------A-----------------...-----------------G-------- 4851
29_BF.BR.02.BREPM119        G-----T---------------GC--G-----------G--C---TCAC----T---A-T-----------A---------------T---A---------GT--------AT---------G-----------------CA...------------C-----G-------- 4659
31_BC.BR.02.110PA           C-----------------------C--G-----------------A--C----T-----T-----G----------C-----G--A---A----------T-G--------------------C-----------------T...-------------------------YC--- 4919
32_06A1.EE.01.EE0369        -----AA--C---A--------AGT--G-----------------A--C----------T-----------A-----C-----G--A---A----------T--------A-A------------------G------------...-----------------G--TC--- 5120
33_01B.MY.05.05MYKL007_1    -------------A---T-----C-------------G--------AA----T--------------------C-----G-GA---A----------AG-------ATA---A----------C-T-------C---...-----------C--------T---- 4864
34_01B.TH.99.OUR2478P       C------------A---T-----C----------G---G--------AC-A--T--------------A-----C-----G-GA---A----------AG-------A-G--------------CGT-----------...-----------C--------T---- 4687
35_AD.AF.05.05AF095         -------------A---T-----TG-G---------------------C----T---A-T--------------------C-----G--A-T-A----------AG-------------C---C-----------------T...---------------------T--A- 4678
36_cpx.CM.00.00CMNYU830     ---C---------A---T-----C-------------G-----C---C----T--------------------C-----G--A---A----------A------------------T----C--------------T...---------------------T---- 4690
37_cpx.CM.00.00CMNYU926     ------A------A---T-----C--------T-----G--------C----T--------------------C-----G-CA---A--------GA--------A-------------------C-------------GT...---------------------T---- 4675
38_BF1.UY.05.UY05_4752      C-----G------A--------GC--G-----------G--C---A-A-----T-----T-----------A-----C-----G--T---A-T--------T-G---------------A-C----C--------------------...---------------------T--A- 4896
39_BF.BR.03.03BRRJ103       ------------------------C----------------------------T-----T-----------A-------------C-T---A---------G---------A--------A-T--------------------...--A--------------G-------- 4974
40_BF.BR.04.04BRRJ115       ------------------------C-------------------C-----C-T-----T-----------A-------------T---C----G----G---------------T---A-T--------------------...------------------G-------- 5037
42_BF.LU.03.luBF_05_03      ----------------------A-C--------------G-------------------T-----------A---------------TA--A----------T-G------T-----------------------------...C-----------------G-------- 5023
43_02G.SA.03.J11223         -----------------T-----C-------------G-----C---C----T--------------A-----C-----G--A---A----------A--------A-------------------------------T...-----------A--G-----T---- 5010
U.CD.83.83CD003             G--C---------A----------C--G------------C--C---C----T-----T--------------------C-----G--A---A------------------------------------C---A-------------...---------------------T---- 4980
U.CD.90.90CD121E12          -------------A----------C--G------------CA--C---T-----T--------------------C-----G--A--AA--------------------------------------C--C---A-------------...---------------------T---- 4684
U.GR.99.GR303               T--C--A---------------AGC--G-----------------A--C----------T-----------A-----C-----G--A---A------------A------------------------C-----------------T...------------T-------T---- 4756
U.NL.95.U_NL_95_H10986_D1   G--C--------------------GC--G-----------G-----A--C----T-----T--------------------C-----G--A---C-----------T-G------T------------------------T...------------CT-------TC--- 5370
CPZ.CD.90.ANT               G----T---------------A---AC---T--T---A-G--A--C-AA-----T--CCC------------TCAA--AG-C----C-A--AA-CC-----G-A-G-A-ATTGA-ATAC--CC-C--CA-GAAG-------GT...C--------GA-A--G--G--T--A- 4930
CPZ.CM.01.SIVcpzCAM13       ----GTT-----G-----T---T-TG-------TA--T-G--C--CTCTC-A--T-----T-----G----------C--C----GA--TA------G--T--T--AGAA--CC------CT--CAG--A------C-GG...C--------CA-C--------T-C-A- 5068
CPZ.CM.05.SIVcpzEK505       ---C--G------A----T-------T-G------A-G-----A--C--T-----T-----T-----A---C------C--C--GC-A--GG-C-----GAGACC---A-TA----A----------C-CT-----G-----A...C--------C--C--------TC--- 5057
CPZ.CM.05.SIVcpzLB7         ------T------A-----------GT-G-----T--TA-G-----T-A-----T---A-T-----G---A------C------G------A---------GAG-------------A-------T--C-----------C--G...----------A-C---------T---- 5040
CPZ.CM.05.SIVcpzMB66        T-----G------A----T------T--G-----------T-----C---T---C--T-----T-----G---A------C------G------A----------A----------------------C-A---------------...--------CAG-T--------T---- 5031
CPZ.CM.05.SIVcpzMT145       T--C--T--C--G-------------AC------TAGG-----A--C--TC----T-----T-----A---CAA----C--C--G--C--TA-C----A-A-GGC---CCT-----A---------TC-C--A---G--C-G-...--------G---T----------G-- 5027
CPZ.GA.88.GAB1              T---GT------G--G--T--CTTCAC------TAG-T-G--C--CTCTC-A--T-----T--C---G----AA----C--C----GA--TA--------G--GC-T--AGCG---C-T---------CA---A-------C-GG...C-----------CT-----G-T----- 5549
CPZ.TZ.00.TAN1              ---C--T------A------------AC-----------T-G-----CACTAGG--T---AC-------G-T-CAGGA--AG-TC-GC-A--AA---------G-A-G-A-ATTGA-ATTC--CC-C-T----T-------C-GA...C--------GA----G-----T---A- 5132
CPZ.US.85.CPZUS             T--C-AG------A----T-----GA-G------A-G------------------T-----T-----A---C------C--C------A--TG-T-----GAGACC---A-TT--C-AA---------C-CTA----G--C--T...---------C---T-G-----T----- 5545
CPZ.CM.x.SIVcpzDP943        ---C--A------A----T------A-G-----TAG---------C--TC----T-----T-----G---CT-----C--C------A--TG-T-----GACACA---AGTT----A-----G--C-C------G--CC-A...C--------G--C---------T--A- 5071
CPZ.CM.x.SIVcpzMB897        -------------A----T------T-G--------C---------T--C----T-----T-----G---G------C-----G--A---A--C--------A---------------A-------TC-C-----------C-CA...---------CA-----------T---- 5051
N.CM.02.DJO0131             ---C-AG--------T--T------A-G------A-G-----A------T-----T---A-T-----A---C------C-----GC-A--GG-C-----GAGACC---ATT---------------------CG-----G--C--T...C--------CA--T-G---C-TC-A- 4991
N.CM.04.04CM_1015_04        ---C-AG-----G--T--T------A-G------A-G-----A------T-----T-----T-----A----T-----C-----GC-A--GG-C-----GAGACA---ATTT--------------------CT-----G--C-----...C------ATA-CT--------TC-A- 4990
N.CM.04.04CM_1131_03        ---C-AG-----G--T--T------A-G------A-G-----A------T-----T-----T-----A---CG-----C-------C-A--AA-C-----GAGACA---ATTT--------------------CT-----G--C--T...C--------CA-CT--------TC-A- 4989
N.CM.95.YBF30               ---C-AG--------T--T------A-G------A-G-----A-----TC-----T-----T-----A---C------C-----GC-A--GG-C-----GAGACC---ATTA---------------------C------G--C--A...C--------CA-C---------TC-A- 5079
N.CM.97.YBF106              ---C-AG--------T--T------A-G------A-G-----A----TGT-----T---A-T-----A---CT-----C-----GC-A--GG-C-----GAGACC---GTTA-----------M----CG-----G--C--M...C--------CA-----------TC-A- 5077
O.BE.87.ANT70               G----AA-----GA-T----------AC-------AGGA-G--A------C----T---AC-------A---CA--A--AG-C--G--A---A-CC-------G-G---ACTGA-C-A----------C-CTA----------GT...C--------GA-A------CTGC-A- 5539
O.CM.91.MVP5180             G----AA-----GA-T--T-------AC-------AGGA-G--A------C----T--CACA------A----A--A--A--C--G-------A-TC----G--G-G---GCTGA-C-A----------C-TG-----------GT...C--------GA-A-------T---A- 5514
O.CM.96.96CMABB637          ------T-----GA-T----------AC------TAG-A----A------C----T--CAC-------A---GG--A--A--C--G-GA---A-TC----G--G-G---GCTGA-C-----------C-CT------------GT...C--------GA-A--G---CT-C-A- 4969
O.CM.98.98CMA104            G-----A-----GA-T----------AC------TAG-A-G--A-----------T--CAC-------A----G--A--A--C--G--A---ACC-----G--G-G---ACTGA-C-A------C--C-CT-A----------GT...---------GA-A--G---CTAC-A- 4971
O.CM.99.99CMU4122           C----CA-----GA-T----------AC-------AGGA----A----AAC----T---AC-------A----G--A--A--C--G--A---A-CA----G--G-G-A-AYTGA-C-A----------CTCT-----------GT...C--------GA-AT-G---CTAC-A- 4970
O.FR.92.VAU                 G----TA-----GA-T----------AC-------A-G-----A------C---------AC------A----G--A--A--T--G-GA---ACTC----G--G-G---GCTA--CCA----------C-CT-A----------GT...C--------GA-A-----GCT-C-A- 5051
O.SN.99.SEMP1299            G----AA-----GA-T----------AC-------AG-A-G--A--A-A-----T----C-------A----G--A--A--C--G--A---A-CC----G--G-G---ACTGA-C------------C-CT-----------GT...C--------GA-A--G---CTAC-A- 5538
O.US.99.99USTWLA            T----AA-----GA-T----------AC-------AGGA-G--A------C-C-----CAC-------A----T--A--A--C--G--A----GTC----G--G-A---GCTGA-C----------TC--TAA---------GT...C--------GA-A------CT-C-A- 4968
```

```
                                                                                            Vpr start                                                  Vif end
B.FR.83.HXB2               CACTAGCAGCATTAATAACACCAAAA...............AAGATAAAGCCACCTTTGCCTAGTGTTACGAAACTGACAGAGGATAGATGGAACAAGCCCCAGAAGACCAAGGGCCACAGAGGGAGCCACACAATGAATGGACACTA.GAGCTTTTAGAGGAGCTTAAG 5639
Vif                        A__L__A__A__L__I__T__P__K__.__.__.__.__._ _K__I__K__P__P__L__P__S__V__T__K__L__T__E__D__R__W__N__K__P__Q__K__T__K__G__H__R__G__S__H__T__M__N__G__H__*#
Vpr                                                                                         _M__E__Q__A__P__E__D__Q__G__P__Q__R__E__P__H__N__E__W__T__L_#_E__L__L__E__E__L__K_
A1.GE.99.99GEMZ011         ----GAA-------G---------C-...............-GAGA--G--------------------GA--G--A----------------------------------G-----G--------A----------------TGT--.--A--G-----A--T------ 4823
A1.KE.00.KER2008           ----GAA-------G----------G...............----C-----------------------A------AG------------------------------T--G-----------A------T--------------T--.--AT-G-----A--------A 4838
A1.KE.00.KNH1144           ----GAA-------G--------C--...............----C-----------------------G--T-T-AG---------------------A-----------G-----C-----A------T------------TGT--.-----G-----A--T------ 4844
A1.KE.00.KSM4024           -G--GAA-------G---------C-...............----C--G------G-------------G--T-T-AG---------------------------------G-----T-A----------T------------TGT--.-----G-----A--------- 4835
A1.KE.00.MSA4069           ----GAA-------G-G---------...............----C-----------------------G--T-T-A-------------------------G--------------C-----A-------------------TGT--.-----G--------T------ 4838
A1.KE.00.NKU3005           -----AA-------G--------CC-...............-G-----G------C-------------G----T-AG---------------------------------G-----C------------T------------TGT--.-----G--G--A--------- 4838
A1.RU.00.RU00051           --T-GAA-------G---------C-...............-G-GG--G---------A----------G------A----------------------------------G-----G--------A---T------------TGT--.--A--G-----A--------- 4940
A1.RW.93.93RW_024          ----GAA-------G---------C-...............----C--G--------------------G----T-AG------------------------G-----------A---------------T------------TGT--.-----G-----A--T------ 4841
A1.SE.95.SE8891            ----GAA-------G-----------...............-------G------------------------GT-AG-------C--------------------------------------------T-----------G--T--.--A--GG----A--------- 4828
A1.SE.95.UGSE8131          ----GAA-------G-----------...............-G--C-----------------------G---GT-A-----------------------------------------------------T------------TGT--.-----G-----A--T--C--- 5010
A1.TZ.01.A173              ----GAA-------G-----------...............----------------------------G----T-AG------------------------------T--G-----C------------T----------------- .--A--G-----A--------- 4838
A1.UA.01.01UADN139         ----GAA-------G---------T-...............-G-GA--G---------A----------G------A---------------------T------------G-----G--------A----------------TGT--.--A--G-----A--A------ 4838
A1.UG.92.92UG037           -----AA-------G--------TC-...............-G--------------------------A----T-AG---------------------------------G-----------A------T------------TGT--.--T--G-----A--T------ 5655
A1.UG.99.99UGA07072        ----GAG-------G---------C-...............-G--C--G--------------------G---GT-AG-------C-------------------------G------------------T--------------T--.--A--G-----A--A------ 4838
A2.CD.97.97CDKTB48         --T-GAG-------G--G-TT---C-...............-G--C-----------------------G----T-AGT--------------------------------G-------C----------------------GTGT--.--A--G-----------A--- 4979
A2.CY.94.94CY017_41        --T-GAA----G--G--G-TT---C-...............-G--C-----------------------G----T-AGT---------------------------------------------------T------------TGT--.--A--G-----------C--- 4996
A.SN.01.DDI579             ----GAG-------G--G--T-----...............----C--G------------------C-G----T-A----------------------------------G------------------T------------TGT--.--A--G--------A--G--- 4837
A.SN.01.DDJ369             ----GAA-------G--------G--...............-G--C-----------------------GA---T-A----------------------------------G------------------T-------G----TGT--.--A--G--------T------ 4840
A.SN.96.DDJ360             ----GAG-------G-----------...............--------------------------C-G----T-AG---------------------------------G------------------T------------TGT--.--A--G--------AA----- 4837
A.CD.02.02CD_KTB035        ----GAA----C--G--G-------G...............GG--C-----------------------G----T-A----------------------------------G------------------T------------TGT--.--A--G--------------- 5633
A.CD.97.97CD_KCC2          ----GAA-------G---------C-...............-GAGC-----------------------G----T-A----------------------------------G------------------T--------------T--.--A--A--------------A 5635
A.CD.97.97CD_KTB13         ----GAA-------G--CAG------...............-GA-G-----------------------G----T-AG----------------------------A-------------------A---T-----------GTGT--.--A--A--------------- 4843
B.AR.04.04AR151516         --------------G---A-------...............-G--A--G--------------------AA-----------------------------------A-------------------A---T-------G------T--.--------------A------ 4861
B.AU.87.MBC925             --------------------------...............-------------------------------------------------------------------T---------------------T-T---------------.--------------------- 5650
B.BO.99.BOL0122            --------------G-----------...............----C-------------------C--CA--T--------------------------------------------------A------------------------.--A--------A--------- 4859
B.BR.03.BREPM2012          ------------------------G-...............-G-G------------------------GA--------------------------------C----------------------A---TT----------------.--------------A------ 5064
B.CA.97.CANB3FULL          --------------------------...............C--------------------------G------------------------------------------------------A------T-----------------.--A------------------ 4928
B.CN.05.05CNHB_hp3         --------------G-----------...............-G--G----------------------G------------------------------------G--T-----------------A---T-------G---------.---------------A----A 5638
B.CO.01.PCM001             -----A----------C---------...............--------------G-------------A--------------------------------G-----T---------------------T-----------------.--------G------------ 4844
B.GB.83.CAM1               -----A-----------G--------...............----------------------------G---------------------------A--------------------------C-----T-----------------.--------------A------ 5640
B.GB.86.GB8                -----A-------G------------...............-----------------------------------------------------------------------------------------T-----------------.--------------------- 5659
B.GE.03.03GEMZ010          -----A--------G---AG------...............----A---A---------------A---G---------------C-----------A-------G------------------------TC----------G-----.--A-----G-----A------ 4844
B.IT.05.SG1                ------------------------G-...............-G--G----------------------CA------------------------------------------------------------T-----AG----------.--A--------A--------- 5441
B.JP.05.DR6538             -------------G----AG------...............---------------------------G----------------C-----------------------------A--------------T-----------------.--A------------------ 5643
B.KR.05.05CSR3             -----A--------------------...............----C----------------------G-A---T---------------------------------------------A---------T-----------T-----.--------------------- 5664
B.NL.00.671_00T36          ---------T----------------...............----------------------------G--------------------------------G--------------------A------T------C----------.------C-------A------ 5204
B.RU.04.04RU128005         -----------A---C---------C...............----G----------------------G-A-----AG-------Y--------------------A----G-----------A--A---T-----------------.--AA----------------A 5130
B.TH.00.00TH_C3198         ---------T----------------...............----G----------C-----------G------------------------------------G--------------------------------G---------.--------------------- 4853
B.UA.01.01UAKV167          --T-----------G---A-------...............-G-CAG----------------------G-------------------------------------------------C----------T------C----G-----.--------G-----A------ 4865
B.US.04.ES10_53            --------------------------...............-G--G---A-----------C-A----G-A-----A-------------------------G-----------------------A---T-----------------.--------------------- 5648
B.US.99.PRB959_03          -----A--------G-----------...............----C---------A--A-----------------A------------------G----------A-----------------------T-----------------.--------------------- 4859
C.AR.01.ARG4006            ----GA-------GT---A-------...............----G----------C------------G----T-AGT---A----------------------------G--A--G--------A---T-T---------------.--A--C--------A--C--A 4823
C.BR.04.04BR013            ----GA-------G----A-------...............-------G-------C------------A---GT-AGT--------------------------------G--A--G------------T-----------------.--A--------A--A--C--- 5103
C.BW.00.00BW07621          ----GA-------G----A-------...............----G----------C------------G----T-AGT--------------------------------G-----G--------A---T-----------------.-----CC-------A--C--A 4992
C.CN.98.YNRL9840           ----GA-----C-G----A-------...............---------------C------------A----T-AGT---A--------------C----------T--G-----G--------A---------------------.------C-------A--C--- 4826
C.ET.02.02ET_288           --T-GA-------GC---AG------...............----G----------C------------GT---T-AGT--------------------------------G-----G--------A---T---G-------------.--------------C--C--- 4841
C.GE.03.03GEMZ033          ----GA-------G----A-------...............C---G----------C------------GA---T-AGTG-------------------------------G-----G--------A---T-----------------.---A--C-------A--C--- 4820
C.IL.99.99ET7              ----GA-------G----A-------...............----G---A------C------------G---G--AGT--------------------------------G-----G--------A---------------------.--------------A--C--- 4820
C.IN.99.01IN565_10         ----GA-----C-G----A-------...............---------------C--------A---G----T-AGT--------------------------------G-----G--------A---T-----------------.------C----------C--- 5020
C.KE.00.KER2010            ----GA-------G----A------C...............---------------C-----------CA----T-AGT------------------C-------------G-----G--------A---T-----------------.---A--C-------A--A--- 4820
C.MM.99.mIDU101_3          ----GA-----C-G----A-------...............---------------C--------A---A----T-AGT------------------C----------T--G-----G--------A---------------------.------C-------A--C--- 4993
C.MW.93.93MW_965           -G--GA-------G----A-------...............---------------C------------A----T-AGT--------------------------------G-----G--------A---T-----------G-----.-----------A--A--C--- 4826
C.SN.90.90SE_364           ----GA-------G----A------G...............---------------C------------G----T-AGT---A----------------------------G-----G--------A---T-----------------.--------------A--C--- 4808
C.SO.89.89SM_145           ----GA-------G----A-------...............---------------C-------CA---G----T-AGT------------------T-------G--T--G-----G--------A---T-----------------.---A--C-------A--C--- 4860
C.TZ.02.CO178              ----GA-------G----AG------...............---------------C--------A---G----T-AGT------------------C-------------G-----G--------A---T-------G---G-----.---A--C-------T--C--A 4820
C.UY.01.TRA3011            ----GA-C----------GG------...............---------------C------------A----T-AGT--------------------------G-----G--A--G--------A---T-----------------.--A-----------A--C--- 4810
C.YE.02.02YE511            ----GA-------G----A-------...............--------A------C------------G---GT-AGT--------------------------------T-----G--------A---T-----------------.--------------A--C--- 4814
C.ZA.04.04ZASK164B1        ----GA--A----G----A-------...............-G-------------C-------CAC--G----T-AGT--------------------------------G--A--G--------A---T-----------------.---A--C-------A--C--- 5053
C.ZA.05.05ZAPSK240B1       ----GA-------G----A-------...............-G--C----------C------------G----T-AGT--------------------------------G-----G--------A---------------------.-----------A--A--C--- 5043
C.ZM.02.02ZM108            ----GA-------G----A-------...............----G----------C------------G----T-AGT--------------------------------G-----G--------A---T-----------------.--------------A--C--A 5608
D.CD.83.ELI                -----A-----------G--------...............C---------------------------G---G--A-----A---------------------C------G------------------T-----------G--T--.--------------------- 5185
D.CM.01.01CM_0009BBY       -----A-----C-----GAG------...............--------A-------------------G---GT-A-----A--C-------------------------G--------------A---T--------------T--.---A-------------C--- 4838
D.KE.01.NKU3006            -----A------------A-------...............-G------A-------------------G---G--A-----A--C-------------------G------------------------T--G---C-------T--.--------------------- 4844
D.KR.04.04KBH8             --T--A--------------------...............-G--------------------------G---GT-A-----A---------------------G------------G-----------CT-----------G--T-G.-----------A--------- 5598
D.TD.99.MN011              -----A------------AG----G-...............----G---A-------------------G---GT-A-----A---------------------C-------------------------T-----------------.--A--------------C--- 4863
D.TZ.01.A280               -----AT-----------A-------...............----------------------------A---G--A----------------------------G-----------------------CT--------------T--.--------------------- 4840
D.UG.99.99UGK09259         -----A-----------G---A----...............--------A-----G-------------G---G--------A--C-------------------G------------------------T--------------T--.--------------------- 4844
D.YE.01.01YE386            -----A------------A-------...............GG--C---A--------A----------G---G--------A--C-------------------G------------------------T--------------T--.-----------------C--- 4838
D.YE.02.02YE516            -----A-----------G-G---CG-...............-G--G---A-----G-------------A---GT-A-----A---------------------C-------------------------T--------------T--.--A--------A-----C--- 4850
D.ZA.90.R1                 -----AA----------G--T-----...............----C-----------------------G---G--A-----A----------------------------G------------------T--------------T--.--A-----G------------ 4988
F1.AR.02.ARE933            -----A-----------G-----C-G...............----C------C--------------CCA-------GT---A--C-------------------------G------------------T-----------G-----.--T------------------ 4933
F1.BE.93.VI850             -----A-----------G-T---G-G...............----C---------GC----------CCA------AGT------C------------------G------G------------------T-----------G-----.--AA----------------- 4979
F1.BR.01.01BR125           -----A-----------G-T-----G...............-----------G--------------CCA------AGT---A--C-------------------------G--A---------------T-----------------.--------G------------ 5084
F1.ES.x.P1146              -----A-----------G-T-----G...............----C---------------------C-A-------GT---A------------G------G-----T--G-----G--------A---T-----------------.--------------------- 4916
F1.FI.93.FIN9363           -----A--------G--T-T-----G...............---GC------G--------------C-A------AGT---A--C------------------G---T--G------------------T-----------G-----.--A------------------ 4968
F2.CM.02.02CM_0016BBY      -----A-----C--G----C-----G...............---G-------G--------------C-A------AGT--------A---------C--------------------------------T-----------------.--------------------- 4832
F2.CM.95.MP257             -----A-------------T------...............-----------G-----A--------C-G------AGT------------------C-------------G------------------T-----------------.------C-------------- 4851
F2.CM.97.CM53657           -----A-------------C-----G...............-----------G--------------C-G------AGT------------------C-------------G------------------------------------.--------------------- 4832
G.BE.96.DRCBL              --T--AA--T----G-GG------C-...............-G--G--G--------------------G----T-A------------------G---------------G-----------A--A--C------------G--T--.--A--G-----A--------A 5596
G.CM.01.01CM_4049HAN       --T-GAA-------G--C------C-...............-G--A--G--------------------GT---T-A---------------------T------------G-----------A--A--CT--------------T--.-----G-----A--------A 4841
G.CU.99.Cu74               ----GAA-------G---------C-...............----------------------------A----T-A-----A----------------------G-----G-----------A--A--C---------------T--.--AT-G--------------A 5243
G.ES.99.X138               -----AA-----------------C-...............---GC-----------------------G------A-----A----------------------G-----G--------------A--CT-----------G--T--.--A--G-----A--------A 5080
G.GH.03.03GH175G           --T-GAA-------G---AG---GC-...............-G--A-----------------------G------A-----AA--C-------------------A----------------A--A--CT-----------G--T-G.--A--G-----A--------A 5679
G.KE.93.HH8793_12_1        ----GAA-------G---------C-...............----C--G--------------------G----T-A-----A---------------T------------G-----------A-----CT-----------G--T--.--A--G-----A-G------- 5039
G.NG.01.01NGPL0669         ----GAA-------G---A-------...............-----------------A--C-------G----T-A-----A-------------------G---A----------------A--A--CT-------G------T-G.--A--G-----A--A-----A 4844
G.PT.x.PT2695              ----GAA-------GC--------C-...............---GC-----------------------G----T-A-----A----------------------G-----G--------------A--CT-----------G--T--.--A--G-----A--------A 5581
G.SE.93.SE6165             --T-GAA--T----G-----T-----...............-G-GT--G------------C---------G--T-AG----A----------------------------G-----------A--A--CT-----------G--T--.--A--G-----A--------A 5036
H.BE.93.VI991              ----GA-----------T--------...............-G--C-----------------------GA-----AGT--------------------------------G------------------------------------.--A-----G-----A-----A 5028
H.BE.93.VI997              ----GA--------G--TG-------...............----C-----------------------AA--G--AGT-------G------------------------G----------------------------A-------.--------------------- 4963
H.CF.90.056                --T--A--------G-GG--------...............----------------------------GA--G--AGT--------------------------------G------------------------------------.---------------A----- 4986
J.CD.97.J_97DC_KTB147      -----A----T-------A------G...............----G----------C-----------CA---G--AGT--------------------------------------------A------T-----------------.--------G-----------A 4833
J.SE.93.SE7887             -----A------------G-------...............-G--G----------------------CA------AGT--------------------------------C-----------A------T--C--------------.--------G-----------A 4953
K.CD.97.EQTB11C            -T---A-----------G-T------...............----C---A-----GG----------CCA------AGT---A----------------------------G------------------T-------G------A--.---A----------------- 4831
K.CM.96.MP535              ----CA--------G--G-T----G-...............-G-CC----------G------------A------AGT--------------------------------G------------------A--------------T--.---A----G------------ 4835
```

```
                                                                                                                     Vpr start                                                        Vif end
B.FR.83.HXB2                  CACTAGCAGCATTAATAACACCAAAA..............AAGATAAAGCCACCTTTGCCTAGTGTTACGAAACTGACAGAGGATAGATGGAACAAGCCCCAGAAGACCAAGGGCCACAGAGGGAGCCACACAATGAATGGACACTA.GAGCTTTTAGAGGAGCTTAAG 5639
Vif                           A__L__A__A__L__I__T__P__K__.__.__.__.__.__K__I__K__P__P__L__P__S__V__T__K__L__T__E__D__R__W__N__K__P__Q__K__T__K__G__H__R__G__S__H__T__M__N__G__H__*#
Vpr                                                                                                          _M__E__Q__A__P__E__D__Q__G__P__Q__R__E__P__H__N__E__W__T__L_#_E__L__L__E__E__L__K_
01_AE.CF.90.90CF11697         ----GAA-------GC----------..............----C--G-------C------------G----T-A-----A-----------------------------G-----------A--A--CT-------------T--.--A--G--------------A 5588
01_AE.CN.05.FJ051             ----GAA--------C----------..............-------G---G---C-------C----A----T-A-----A--------------------------T--G-----------A-----CT-----C----G--T--.--A--G--------T-----A 5649
01_AE.CN.06.FJ054             ----GAA--------C---------G..............-------G-------C------------A---GT-A-----A--------------------------T--G-----------A-T---CTC------------T--.--A--A--------------A 5664
01_AE.HK.x.HK001              -----AGG-------C-GGG------..............GG--GC-G---G---C----C-----A-A----T-A-----A-----------G-G---------------G-----------A--A--CT-------------T--.--A--A--------------A 5001
01_AE.TH.01.01TH_R2184        ----GAA------G-C----------..............-G-----G-------C------------A----T-A-----A--------------------------T-GG-----------A--A--CT-------------T--.--A--A--------------A 4839
01_AE.TH.02.OUR769I           ----GAA------G-C----------..............-G-----G-------C------------A----T-A-----A--------------------------T--G-----------A--A--CT---C---------T--.--A--G--------------A 4838
01_AE.TH.90.CM240             ----GAA--------C----------..............-G-----G-------C----------AGAA-T-...-----A-------------------------...-G-----------A--A--CT-------------T--.--A--G--------------A 5207
01_AE.US.00.00US_MSC1164      --T-GAA--------C----------..............-G-----G-------C------------G----T-A-----A--------------------------T--G-----------A--A--CT-------------T--.--A--G--G-----T-----A 4844
02_AG.CM.02.02CM_4082STN      -----AA-------G---------G-..............----C-----------------------GA--GT-AG----A--C----------G---------------G-----------------GTC---------------.--A--G-----A--------A 4838
02_AG.EC.x.ECU41              -----AA----C--G--G------C-..............-G--C-----------------------A---GT-AG----A--C--------------------G--T--G--------------A--GTT---------------.--A--G-----A-----C--A 4735
02_AG.GH.03.GHNJ196           ----GAA-------G--G------C-..............-G--G------------A----------A---GT-A-----A---C-------------------------G--------------A--GTC----C----------.--A--G--------------A 5668
02_AG.NG.01.PL0710            --T-GAA---C-----GA--------..............----C-----------------------GA--GT-A-----A--C---------AG---------------G-----------------GTT------------T--.--A--G-----A--------A 4823
02_AG.NG.x.IBNG               ----GAAT--T---G--G------C-..............----C-----------------------G---GT-AG----A--C---------AG---------------G-----------------GTC---------------.--A--G-----A--------A 5164
02_AG.SE.94.SE7812            ----GAA-------G---------C-..............-G--C-----------------------A---GT-AG----A--C----------G---------------G-----------------GTT---------------.--A--A-----A--------A 5011
02_AG.SN.98.MP1211            ----GAA-------G---------C-..............-G--G-----------------------A---GT-AG----A--C----------G---------------G--------------A--GTT---------------.--A--G-----A--A--G--A 4836
03_AB.RU.97.KAL153_2          ---------------G----------..............-----------------------------------------------------------------G--------A--------------T-------G---------.--------------------- 4865
04_cpx.CY.94.CY032            ----------------T-C-------..............----C-----------------------A------AGTG--------------------------------G-----G-----A--A---A-T--------------.--------G-----------A 5005
05_DF.BE.x.VI1310             -----A--------------------..............----C---------------C-------G---G--A-----A--C--------------------------------G----------A---T-----------T----.--------------------A 5021
06_cpx.AU.96.BFP90            -----A----T-------A---G-G-..............----G---A------------------CAA--G--AGT---------------------------------G-----------A------T----------------.--A-----G--------G--A 5667
08_BC.CN.97.97CNGX_6F         ----GA-----C-G---A--------..............-G-------------C--------A---A----T-AGT-------------------C----------T--G-----G----------A--------------------.------C-------A--C--- 4822
09_cpx.GH.96.96GH2911         -----A------------AC----G-..............--A-C-----------------------A------AGT---------------------------G-----G-----------A--A---A----------------.--------G------------ 4841
10_CD.TZ.96.96TZ_BF061        -----A-T------------------..............----C---A-------------------G---G--A-----A-----------------------G------------------------T-------------T--.-----C--------------- 5018
11_cpx.GR.x.GR17              --T-GAA-------G-----------..............-G--C-----------------------G----T-A-----------------------------------G-----G----------A---T---G-------TGT--.--A--G-----A--------- 4941
12_BF.AR.99.ARMA159           -----A----T------G-T-----G..............--A-C------G--------------C-A---G--AGTG--A--C------------C-------------G------------------T----------------.--------------------- 5644
13_cpx.CM.96.1849             --T--A-----------CAC------..............-----------G--------------G-AA---T-AGT---------------------------------G--A--------A--A---A----------------.--------G-----------A 5044
14_BG.DE.01.9196_01           --T-----------C--------C--..............---GC---------C-------------G----T-A-----A-----------------------G-----G----------------A--CT-------------T--.--A--G-----A--------A 5156
15_01B.TH.99.99TH_MU2079      ----GAA--------C------C---..............-G-----GA------C-------G----A----T-------A----------G---------------T--G-----------A--A--CT-----G----C--T--.--A--G--------------A 5035
16_A2D.KR.97.97KR004          --T-GAA-------G--GGTGA--C-..............-GA-C---------CC------------G---G--A-----A-------------------------------------------------T-------------T--.-----C-----A--T------ 5002
17_BF.PE.02.PE02_PCR0155      -----A-----------G-C-----G..............----C---------------------C-A------AGT---A--C-------------------G------G-----------------T----------------.--------------------- 4829
18_cpx.CU.99.CU76             ------------------T-------..............----C------------A----------G----T-AG----A-----------------------------G--------------A---A----------------.--------G-----------A 4944
19_cpx.CU.99.CU7              ----GAA-------G---------C-..............-G--GT-G--------------------G----T-A------------------------------A----G-----------A--A--CT----------G--T--.--A--G-----A--C------ 4865
20_BG.CU.99.Cu103             ----GAA-------G---------C-..............----------------------------G----T-A-----A-----------------------G-----G-----------A--A--C--------------T--.--A--G--------------A 5105
21_A2D.KE.91.KNH1254          --T--AA-------G--G-TG-----..............-G--C----------C----------A-G---G--A-----A-----------------------------------------------TCT-----------T--.--------------------- 4844
22_01A1.CM.01.01CM_0001BBY    --T-GAG-------G--G--------..............--------A-----CC------------G----T-A-----A-----------------------------G-----------A-TA--CT-------------T-G.--AG-G--------------A 4847
23_BG.CU.03.CB118             ----GAA-------G---------C-..............----------------------------G----T-A-----A-----------------------G-----G-----------A--A-----------------T--.--A--G--------------A 5082
24_BG.CU.03.CB378             ----GAA-------G---------C-..............----------------------------G----T-A-----A-----------------------G-----G-----------A--A-----------------T--.--A--G--------A-----A 5064
24_BG.CU.03.CB471             ----GAA-------G---T-----C-..............----------------------------GA---T-A--G--A-----------------------G-----G--A--------A--A---T-------------T--.--A--G--------------A 5079
25_cpx.CM.01.101BA            -T---AA-------G--G-C------..............--------------C-------------G----T-A-----A---------------A-------------G-----------A--A---A-------------T--.--A--G-----A--------A 4847
27_cpx.CD.97.97CDKTB49        --T---------------C----G--..............----C-----------------------A------AGTG--------------------------------G-----T-----A--A---A----------------.--------G-----------A 4847
28_BF.BR.99.BREPM12609        -----A---------C--A-------..............----C-----------------------G---------------------------------------T------------------A---T----------------.--------------A------ 5005
29_BF.BR.02.BREPM119          -----A--------C---------G-..............----C-----------------------A---------------C-----------------------------AT----------------T----------G--T--.-----A--------A------ 4813
31_BC.BR.02.110PA             ----GA-----C-G---A--------..............----G---A------C------------G----T-AGT---A------------GC---------------G--A--G----------A---T----------------.--A-----------A-----A 5073
32_06A1.EE.01.EE0369          -----AA---C-------A------G..............----G----------C-----------CAA-----AGT-----------------G---------------G-----------A-------------------G---.--A-----G------------ 5274
33_01B.MY.05.05MYKL007_1      ----GAA--T-----C-----A----..............-G-----G-------C------------A----T-A-----A--------------------------T--G-----------A--A--C-G---C--G--G--T--.--A-----------------A 5018
34_01B.TH.99.OUR2478P         ----GAA--------C----------..............-G-----G-------C------------A----T-A-----A----------------------C---T--G-----------A--A--C-----C--------T--.--A--G--------------A 4841
35_AD.AF.05.05AF095           --T-GAA-------G---------C-..............------------------------A---G----T-AG----A-------------G---------------G-----C-----A------T-------------T--.-----GC----A--A------ 4832
36_cpx.CM.00.00CMNYU830       -G--GAA-------G---------C-..............----G-----------------------A----T-A-----A--C--------------------------G-----------------GTT---C--------T--.--A--G-----A--------- 4844
37_cpx.CM.00.00CMNYU926       ----GAA-------G--C------C-..............-G--A--G--------------------A---GT-A-----------------------------------G-----------A--A--CT----------G--T--.--A--G-----A--------- 4829
38_BF1.UY.05.UY05_4752        -----A-----------G-C-----G..............----C---------------------C-AA-----AGTG--A--C-----------------C-G------G------------------T----------------.-----C--------------- 5050
39_BF.BR.03.03BRRJ103         -----A----------C---------..............-G----------------------A---A----T--------------------------T-G--------------------A------T----C-----------.------C----A--A------ 5128
40_BF.BR.04.04BRRJ115         -----A-----------AG-------..............------------------------A---A------------------------------------GA-------------------------T----------------.------C----A--A------ 5191
42_BF.LU.03.luBF_05_03        -----A--------------------..............----Y-----------------------GA---T-A-----A-------------------------------------------------------T----------------.--------------------- 5177
43_02G.SA.03.J11223           ----GAA-------G---A-----C-..............-G--A----------C-----------G-A--GT-A-----A--Y-------------T------------G-----------------GTY---------------.--A--G-----A--R-----A 5164
U.CD.83.83CD003               ----GA--A-----G--G------C-..............----G---------------------C-G------AGT---A--------------------------------------A-----------T-----C----------.--A-----------A------ 5134
U.CD.90.90CD121E12            ----GA--------G--C------T-..............----G-----------------------G----T-AGT---A--C-----------------------T-----------A----------T----------------.--A-----------A-----A 4838
U.GR.99.GR303                 -----A------------AC------..............-G--C-----------------------A------AGT---------------------------------------------A------A-------------T--.--------G------------ 4910
U.NL.95.U_NL_95_H10986_D1     -----A---T----G----T------..............CG--------------------------G------AGT---A-----------------------------G-------C---A------T----C-----G-----.---A----------------- 5524
CPZ.CD.90.ANT                 -TT-TTGTAAGA-CC--GA-TTC-G-GGATAC...CCCAAGGGTCC--GAAG--AG--C------C-GT-T-TCT-A-----A--------------------GC-G--TG-G-----------A--A---ATG----G---TTGT--.--AAC------A--AA-A--A 5096
CPZ.CM.01.SIVcpzCAM13         -TT--AA---T---G---A-GA----..............-G-C-T-GA-----C--A--C-----AG-T---T-A--------------------C--------G-------------CC--A--A---T-------G---G-T--.--AA-A-----A--CT-A--- 5222
CPZ.CM.05.SIVcpzEK505         --T--AT---C-GGG-GGG-GT-C-G..............----G--GA-----C--A--------G--C--GT-A--------------------G--A--------T-----A----------------T----C-----G-TT--.--AT-G--G--A--C--A--- 5211
CPZ.CM.05.SIVcpzLB7           --A-CT-T--C---T--CA-AG-C-G..............CCA-C--G---G--------------C-GC---T-AG----A----------------T-------------------------------T-T----G---G--T--.---G-A-----A--A--C--A 5194
CPZ.CM.05.SIVcpzMB66          --T-GT--A---------A-AGC---..............CCA-G--GA-----C-----C-------GA---T-AG--------------------CT------------G--A-----A-----A---T-T--------------.---A-A-----A--A-----A 5185
CPZ.CM.05.SIVcpzMT145         -T---AAGCAG--GG-T-A-A-C--G..............--A-AT-GA------C----CTCA---AG-A-GT-AG----A-----------------------G-----G--A---TCA--A-----CT-TG-GC-G--CTGT--.--CACA--G--A--AA----A 5181
CPZ.GA.88.GAB1                -C---AA---T------T--GAG-G-..............-G-CAT-GA--------A--------CG-C---T-A-----------------------A-----G--------T---CA---A--A--TT--C-G-----G--T--.---ACC--G--A--AT-A--A 5703
CPZ.TZ.00.TAN1                -TT-CAG-AAGG--G-TGAGAGCC--GATAAACAGCCAAAGGACC--G-AGG--C-----ATC------A-----A--------C--------------A--GA-C---A-C-----G-----A--A---T---C---G--CTGT--.--CA-CC----A--AA-A--A 5301
CPZ.US.85.CPZUS               -TT--AAG--C---G--GG--AG-GC..............--A-G--GA-----C--A--------G--T---T-------A----------------TG------------------GCC--A--A---T------G---G--T--.--TG-C-----A--A--A--A 5699
CPZ.CM.x.SIVcpzDP943          -CT--AAG--C-GGG--GG-GT-G--..............----A--GA-----C-----------CG-C--GT-AG----A-----------------A--------T--G--A---GCA--A------T----------G--T--.--A--A--G-----A--G--A 5225
CPZ.CM.x.SIVcpzMB897          --A-CT-----C-GC---A-AGG--G..............CCA-AG--A------C-A----------GT---T-AG----A--C-----------------G-----------------A--A------T----------------.--AT-A-----T--A--C--A 5205
N.CM.02.DJO0131               ----CA----C-GGG-GGG-GT---G..............----G---------C--A--------G---T-GT-A-----A-----------------A---------GC-----------------A--CT-T--------------.--AT-A-----A--AT-A--A 5145
N.CM.04.04CM_1015_04          --T--A----T-GGG-GGG-AT---G..............----G---------C--A--------GG-C--GT-A-----A--------------G--A---------GC-----------------A--CT-T--------G-----.--AT-A-----A--CT-A--A 5144
N.CM.04.04CM_1131_03          --T--A----T-GGG-GGG-AT---G..............----G---A-----C--A--------GG-T--GT-A-----A-----------------A---------GC-----------------A--CT----------G-----.-GAT-A-----A--CT-A--A 5143
N.CM.95.YBF30                 -----A----C-GGG-GGG-G----G..............----G---------C--A--------G--T--G--A-----A-------------G---A--------TGC-----------------A--CT-T--------G-----.--AT-A-----A--AT-A--A 5233
N.CM.97.YBF106                --M--A----C-RGG-GGG-GT---G..............----G--G------C--A--------G--T--GT-A-----A-------------G---G---------GC-----------------A--CT-T--------G-C---.--AT-A-----A--AT-A--A 5231
O.BE.87.ANT70                 -T---AG----G--G---A-G---G-..............-GC-GG-----T--CC-A--C-----CCA----T-A-----A-----------------A--T-GA-T--G---A---GCT-AA-----CTT---------G-----.-----C-----A-----A--A 5693
O.CM.91.MVP5180               -CT-GAA----G--G-G-A-GT----..............-GA-AT-----T--CC-A--C-----CCA--G-T-A-----A-------------------TG--A-T--G---A---GCT--------TT----------------.-----CC-G--A-----G--A 5668
O.CM.96.96CMABB637            --T--AG---G-GG---A-GAC----..............-GA-GT-----T--CC-A--C-----CCA---GT-A----GA--C-----------G--A--T--G--T--G--A---GCA--A-----TT----------W--T--.-----CC-G--A--A--A--A 5123
O.CM.98.98CMA104              -TT--AG--T-G-TG---A-A-----..............-GA-AT-----T--C--A--C-----CCA---GT-A-----A--------------GA-A--T-----T--G--A---GCT--A-----TT----------G-----.-----CC----A--T--G--A 5125
O.CM.99.99CMU4122             -T---AG--T-G--G---A-GA----..............-G--AT-----T--CC-A--C-----CCA------A-----A-----------------A--T--G--T--T--A---GCT-----A---TT---------G-----.---ACCC----A--T--A--A 5124
O.FR.92.VAU                   -CT--AG----G--G---A-G----G..............-GA-AT-G---T---C-A--C-----CCA---GT-A--------------------G--A--T--G--T-----A---GCC---------GTT---C----G-----.-----CC----A-----A--- 5205
O.SN.99.SEMP1299              -T---AG--TGG--G---A-GA----..............-GA-AT-----T--CC-A--C-----CCA---GT-A-----A-----------G--G--A--T--G--T--G--A---GCT--A-----TT----------T-----.-----CC----A--A--A--A 5692
O.US.99.99USTWLA              -CT--AG--T-G--G--GG-GA----..............-CA-GT-GA--T--AG-A--C-----CCA---GT-A----GA------G-------G-TGT-------T-----A---GCT--A-----TT-------------A--.C----C-----A-----A--A 5122
```

```
                                                                                                                                                                      frameshift insert in HXB2
                                                                                                                                                                       Vpr premature end (HXB2 only)
B.FR.83.HXB2           AATGAAGCTGTTAGACATTTTCCTAGGATTTGG...CTCCATGGCTTAGGGCAACATATCTATGAAACTTATGGGGATACTTGGGCAGGAGTGGAAGCCATAATAAGAATTCTGCAACAACTGCTGTTTATCCAT.T.TTCAGAATTGGGTGTCGACATAGCAGAATAGG 5804
Vpr (frameshifted)                                                                                                                                                     F  R  I  G  C  R  H  S  R  I  G
Vpr                    _N__E__A__V__R__H__F__P__R__I__W__.__L__H__G__L__G__Q__H__I__Y__E__T__Y__G__D__T__W__A__G__V__E__A__I__I__R__I__L__Q__Q__L__L__F__I__H_#_#F__Q__N__W__V__S__T__*_
A1.GE.99.99GEMZ011     C----------C-----------A---CCG---...--------AC----A------------A-C--C----------------A---G--T-----T----------C-T-----------------G-T---...--------------C-A--------------- 4987
A1.KE.00.KER2008       ------------------------G--CCA---...--T-----A-----A---T--------C-C--C----------------A------T-----T--GC--------T----------A------G-T---...--------------CAC--------------- 5002
A1.KE.00.KNH1144       C----------------C---------CCA---...--T-----A-----A-----C--------C--C----------------A------T-----T----------C------------A------G-T---...----------------A--------------- 5008
A1.KE.00.KSM4024       C--------------------------CCA---...--T-----A-----A------------A-C--C-----A----------A------T-----T-----------------------A------G-T---...--------------C-A--------------- 4999
A1.KE.00.MSA4069       C--------------------------CCA---...--T-----A-----A-----C--------C--C----------------A---G--T-----T---------C-G-----------A-------CT---...--------------C-A---------G----- 5002
A1.KE.00.NKU3005       C--------------------------CCA---...--------A-----A-----C-----------C----------------A------T-----T------------T----G-----A------G-T---...--------C-----C-A---C----------- 5002
A1.RU.00.RU00051       C----------C--------C------CCG---...--------A-----A------------A-C--C----------------A---G--T-----T------------T----------A------G-T---...--------------C-A--------------- 5104
A1.RW.93.93RW_024      C-------C-----------------ACCA---...--T-----A-----A-----------------C------------CA--A------T--G--T---G-------------------A------G-T---...--------C-----C-A--------------- 5005
A1.SE.95.SE8891        C--------------------------CCA---...--T-----A-----A------------A-C--A--C-------------A------T-----T----------C-T----------A------G-T---...--------------C-A--------------- 4992
A1.SE.95.UGSE8131      C--------------------------CCA---...--------A-----A---T-------CT-C--C----------------A------T-----T--G-------C-T----G-----A------G-T---...--------C.....................-- 5154
A1.TZ.01.A173          C--------------------------CCA---...--A-----A-----A------G----------C--------C-------A------T-----T-----------AT----------A------G-G---...--------------C-A--------------- 5002
A1.UA.01.01UADN139     C----------C-----------A---CCG---...--------A------------------A-C--C----------------A---G--T-----T------------T----------A------G-T---...--------------C-A--------------- 5002
A1.UG.92.92UG037       C--------------------------CCA---...--------A------------------C-C--C----------------A------T-----T------------T----G-----A------G-----...--------C-----C-A---C----------- 5819
A1.UG.99.99UGA07072    C--------------------------CCA---...--------A-----A------------A-C--C----------------A------T-----T----------C-T----------A------G-----...--------------CTC--------------- 5002
A2.CD.97.97CDKTB48     C-A-----A-----------C------CCG---...--A------C-G--A------------A-T--C-----------A---AG------T-----T--G---------T----------A------G-T---...--------------C-A--------------- 5143
A2.CY.94.94CY017_41    C-G-----------------C------CAC---...--A------C----A---T-C------A-T--C------------C---A------T----TT--C------TA------------A------G-----...--------------C-A---------G----- 5160
A.SN.01.DDI579         C-G-----A--C---------------CCA---...--A-----A-----A------------A-C--C----------------A------------T-----------GT----------A--T---G-T---...----------------A--------------- 5001
A.SN.01.DDJ369         C-------------------C------CCA---...--T-----G-----A---T--G-----A-C--C----------------A------T-----T------CA---CT----------A--T---G-T---...----------------A--------------- 5004
A.SN.96.DDJ360         C-G-----A------------------CCA---...--A-----A-----A---T--------A-C--C----------------A------------T-----------GT----------A--T---G-T---...----------------A--------------- 5001
A.CD.02.02CD_KTB035    C-------------G-----------ACCA---...--------A-----A-GT-----A---A-C--C--------C-------A------T-----T----C------CT----C--G--A--A---------...-A------------C-A--------------- 5797
A.CD.97.97CD_KCC2      ---------------------------G-G---...--------------A--------A---A-C--C----------------A------T-----T------------T----------A--------T---...-A------------C-A--------------- 5799
A.CD.97.97CD_KTB13     C----------------------------G---...--------------A--------A---A-C--C----------------A------T-----T--G--------GT----------A-----C--T---...-A------------C-A--------------- 5007
B.AR.04.04AR151516     -G---------C-R--------------YA---...-----GA-------A---T----T-----------------C--------------A-----------------A--------------------T---...-----------A-------------------- 5025
B.AU.87.MBC925         ---------------------------CCA---...-----------G-----------------------------------------------G---------------------------T-------T---...----------------A--------------- 5814
B.BO.99.BOL0122        -G--------------------------CG---...-----G-----G--A-----------------------------------------------------------A---------------C----T---...-------------------------------- 5023
B.BR.03.BREPM2012      -G---------------C-----------A---...------AA------A------------------------------------------------C-GC-----TC---------------------T---...-----------A----A--------------- 5228
B.CA.97.CANB3FULL      -G---------G---------------G-G---...---------C----A---------------------------------------------------------------------T--A-------T---...----------------A--------------- 5092
B.CN.05.05CNHB_hp3     -GG-------------------------GA---...-------AG--G--A---T-C-----------A----------------T--------------------------------------------CT---...-----------A----A--------------- 5802
B.CO.01.PCM001         -GA-T---------------------CTCA---...--A------C----A------------------------A---------------------T-T---------CC--------------------T---...--------C-----C----T------------ 5008
B.GB.83.CAM1           -G------------G------------CCA---...----T--CA-----A---------------------------------C--------------T---------G---------------------T---...-----------A--C----------------- 5804
B.GB.86.GB8            -G------------------------T--A---...--------A---------T-------------------A--------------------------------------------------------T---...------G------------------------- 5823
B.GE.03.03GEMZ010      -G-------------------------CCA---...------AA------A---T-------------------A---------A------------------------C---------------------T---...-----------A-------------------- 5008
B.IT.05.SG1            -----------------C--------A-CA---...---------C--------T----A------------A---------A------------G-------C-------A-----------------------.-.--------C-------AT-------------- 5606
B.JP.05.DR6538         -GA------------------------CCA---...------A--C----A--------T---------------------------------------------------------------------------.-.--------C-------A--------------- 5807
B.KR.05.05CSR3         -GG--------------C---------CCA---...------A-------A---T-------------C---------------A-------------TT----------G--------------------T---...----------------A---------G----- 5828
B.NL.00.671_00T36      -G-------------------------GGA---...--------------A------------------------------------------------C-------------------------------T---...-----------A----A--------------- 5368
B.RU.04.04RU128005     -----------G--------C------G-G---...------------------T-------------------A---------R-----------------------------------T----A---G-T---...--------------C--G-------------- 5294
B.TH.00.00TH_C3198     -G---------------------------A---...--------G--G--A--------------C--A------------------------------T--T-G----C----------T----------T---...-------------------------------- 5017
B.UA.01.01UAKV167      -G------------------------ACCA---...------A-------A---T-C--------------------------------------------------------------------------T---...-----------A----A--------------- 5029
B.US.04.ES10_53        -----------------C-----C-C-G-G---...-----------------------------------------------------------G--------------------------A--------T---...----------------A--------------- 5812
B.US.99.PRB959_03      -----------------------C---CCA---...--T---A-----------T--------------------------------------------------------------------------G-T---...----------------A--------------- 5023
C.AR.01.ARG4006        C-G--------C-----C--------ACCA---...--T---A-------A---T-------------C-----A---------ATG-----------T-----------------------A------G-----...--------------C--T------C------- 4987
C.BR.04.04BR013        C-G--------C-----C--------ACCA---...--T---A-------A---T----------G--C---------------A-G-----------T-----------------------A------G-----...--------------CAAT------C------- 5267
C.BW.00.00BW07621      C-G--------C-----C--C-----ACCA---...------A----------G---------A-T--A-----A---------A-------A--G--T------------T----------A--------T---...-----------A--C-AG-------------- 5156
C.CN.98.YNRL9840       C-G--------C-----C--------ACCA---...--T-----------A---T--G----------C---------------A-------C---A-T---------C-A--A-----------------T---...--------------C-AG-------------- 4990
C.ET.02.02ET_288       C-G--------C--------------ACCA---...------A-T-----A---T--------A-C--C---------------AGG-----A-----TC------C-T-C---------TCA------------...--------------C-AC-------------- 5005
C.GE.03.03GEMZ033      C-G--------C-----C--------ACCA---...------A--A----A---T----T--------C---------------A-G-----T----TG---------G-----------T-A-----C--T---...--------------C-AG-------------- 4984
C.IL.99.99ET7          C-G--G-----C-----C--------ACCA---...------A-------A---T----T--------C-----------C---A-G-----T-----TC---------C------------AT------CT---...--------------C-AG--C----------- 4984
C.IN.99.01IN565_10     C-G--------C-----C---------CCA---...--T-----------A---T--G----------C-----A---------A-------T-----T---------C-A-----------A--------T---...--------------C-AG-------------- 5184
C.KE.00.KER2010        C-A--------C-----C--------ACCA---...------A--A----A---T----------G--C---------------A-------C-----T-----C----C------------A--------T---...--------------C-AC-------------- 4984
C.MM.99.mIDU101_3      C-G--------C-----C---------CCA---...--T-----------A---T--G----------C---------------A-------T---ATT-----------A-----------A--------T---...--A---G-------C-AG-------------- 5157
C.MW.93.93MW_965       C-G--------C-----C---------CCA---...------AA------A-----------------C---------------A-G-----C----TG---------------------T-AT-------T---...--------------C-AT-----T-------- 4990
C.SN.90.90SE_364       C-G--------C-----C--------ACCA---...------A-------A-----------------C---------------A-G-----C-----T-----------------------A--------T---...--------------C-AG--C----------- 4972
C.SO.89.89SM_145       C-G--------C-----C--------ACCA---...------AA------A-----------------C---------------A-------C-----T-----T-----A-----------A-------CT---...--------------C-A--------------- 5024
C.TZ.02.CO178          C-A--------C-----C--------ACCA---...------A-------A-------------C---CC--------------AAT-----C-----TC----------G-----------A------G-T---...--------------C-AG-------------- 4984
C.UY.01.TRA3011        C-G--------C-----C--------ACCA---...--T---A-------A-----------------C-----A---------ATG-----------T---------------------T-A------G-----...--------------C-AG------C------- 4974
C.YE.02.02YE511        C-G--------C-----C-----C--ACCA---...------A----G------T-------------C-----A-----A---A-G-----T-----T--C--------------G-----AT-----G-T---...--------------C-AG--C----------- 4978
C.ZA.04.04ZASK164B1    C-G--------C--------------ACCA---...------A-------A---T----T--------C-----A-----------------C-----TC-GC-------A--------GA----------T---...--------------C-AT-------------- 5217
C.ZA.05.05ZAPSK240B1   C-G--------C-----C--------ACCA---...------AA------A---T-------------C-----A---------A-------AC---TG---------------------T-A--------T---...-A----G-------C-AG-------------- 5207
C.ZM.02.02ZM108        C-G--------C-----C--------ACCA---...--T---------------T-------------A-----A---------A-G-----T-----T---------------------A----------T---...--------------CTCG-------------- 5772
D.CD.83.ELI            -G---------------------------A---...------A-------A--------T--------------------C----T------T-----T-----------A---------T-A--------T---...----------------A--------------- 5349
D.CM.01.01CM_0009BBY   -G------C-----------C------G-A---...--T-----------A--------T--------------------C-----------T-----T----C----G-C-----------A-----C--T---...--------------C-A--------------- 5002
D.KE.01.NKU3006        -G---------------C-----------G---...--T---A--C----A-------------C---------------C---A-------T-----TC----------C--T--------A------------...------G---------TC--------G----- 5008
D.KR.04.04KBH8         -G-------------------------CCA---...--T---A-------A-----------------------------C----T------T-----TC--C-----G-------------AT-----G-----...----------------A--------------- 5762
D.TD.99.MN011          -G------------------C--------A---...--T-----------A--------T--------------------C-----------T-----TC---------C-------------A----C--T---...-------------------------------- 5027
D.TZ.01.A280           -G----------C----C---------T-G---...------AA------A---TT------------------A-----A-----------T-----T-----------C--A--------A--------T---...----------------A--------------- 5004
D.UG.99.99UGK09259     -G---------------C-----------G---...---------C----A-----------------------A-----C---A-------T-----TT--C-----T-C-----------AT-----G-----...----------------A--------------- 5008
D.YE.01.01YE386        -G---------------C---------CCA---...--T---A--C-G--A---T----------T--------------C---A-------T-----TC---------CC--T--------A------------...----------------A--------------- 5002
D.YE.02.02YE516        -G---G--------------C------GCA---...--T-----------A--------T--------------------C-----------T-----T----------CC-----------A---C-C--T---...----------------AG-------------- 5014
D.ZA.90.R1             -G---------------C-------TA-CA---...------TC------------------------------------C-----------T-----T-----------------------A--------T---...----------------A--------------- 5152
F1.AR.02.ARE933        ---------------------------CCA---...------A-T-----A-----C------A-C--C-----------C-----G-----T-----TC--C-----G-AT----------A------------...--------------C-AT--------G----- 5097
F1.BE.93.VI850         --------------G------------CCA---...--T--C--------C------------A-C--C--------C-------AG-----T-----T---------G-AT----------A------------...--------------C--T-------------- 5143
F1.BR.01.01BR125       ---------------------------CCA---...------A-------A-----C------A-C--C----------------AG-----T-----T-----------GT----G-----A------------...--------------C-AT-------------- 5248
F1.ES.x.P1146          ---------------------------CC----...-----------G--A---T-C-----AGC---C----------------AG-----T--G--T---------G-A--------T--AT-----------...----------------A--------------- 5080
F1.FI.93.FIN9363       --C------------------------CCA-TC...---------A----A---T--------A-C--C----------------AG-----T-------------A---AT----G-----A------------...--------------C-AT-------------- 5132
F2.CM.02.02CM_0016BBY  C-------------------------AGAG---...-----------G--A--G---------A-T--C----------------AG-----T-----T---------G-CA----------A--A---------...--------------C-AT-------------- 4996
F2.CM.95.MP257         C--------------------------GAG---...--------------A--G---------A-C--C----------------AG-----T-----T---------G--A----------A--A---------...--------------C-AT-------------- 5015
F2.CM.97.CM53657       C-------------------C------GAG---...--------------A--G--------CA-C--C----------------AG-----TC----T----------CA-----------A--A---------...----------------AT-------------- 4996
G.BE.96.DRCBL          C--------------------------CCC---...--------T-----A--G-T-------A-C-------------------AG-----T-----------------A--A--------A------------...--------------C-A--------------- 5760
G.CM.01.01CM_4049HAN   ---------------------------CCC---...--------------A---T--------A-C--------A----------A------T-----------------A--A---------------G-----...--------C-----C-A--------------- 5005
G.CU.99.Cu74           ---------------------------CCC---...--------------C--GT--G-----A-C-------------------A------T-----------------A--A--------A------------...--------------C----------------- 5407
G.ES.99.X138           ---------------------------CC----...--------------A--GT--------A-C-----------C-------A------T-----------------A--A-----G--AT-----------...----------------A--------------- 5244
G.GH.03.03GH175G       ---------------------------CCC---...--T-----------A--GT--------A-C-------------------A------T-----T-----------G--A--------A------------...-A------C-----C-A--------------- 5843
G.KE.93.HH8793_12_1    ---------------------------CCC---...--T-----------A--G---------A-C-------------------A---G--T------T----------A-----------A------------...--------------C-A--------------- 5203
G.NG.01.01NGPL0669     GT------------------------ACCC---...--------------A------------A-C-------------------A------T-----------------A--A------T-A------------...--------------C-A--------------- 5008
G.PT.x.PT2695          ---------------------------CCC---...--T-----------A--GT--------A-C-------------------A------T-----------------A--A------T-AT-A---------...--------------C-A--------------- 5745
G.SE.93.SE6165         ---------------------------C-----...--T-----------A--G---------A-C-------------------A------T-----------------A--A--------A------------...--------------C-A--------------- 5200
H.BE.93.VI991          -G------------G----A-------CCA---...--T---A-------A--G---------A-T--C----------------AG-----T-----T----------C----------T-A-----CGCT---...----------------A---C----------- 5192
H.BE.93.VI997          -G---------C--G-----------AGAC---...------CAA-----A--G---------A-T--C-----A----------G------C--G-TT----------C------------A--------T---...--------------C-AT-------------- 5127
H.CF.90.056            -----G--------G-----------AG-A---...------CAA-----A--G---------A-C--C-----A----------T------T----TT----------CG-----------A--------T---...--------------C-A--------------- 5150
J.CD.97.J_97DC_KTB147  -GA------------------------CCA---...------AA------A--G---------A-C--------A----------A------T-----------------AT-------------------T---...--------C-----C-AC-------------- 4997
J.SE.93.SE7887         -----------------C---------CCA---...--T---A-T-----A--GT-------AGC---------A-----A----A------C-----------------A-----------AT-------T---...--------C-----C-AT-------------- 5117
K.CD.97.EQTB11C        -GA-----------------C------CCA---...-------A---G--A--G---------A-C-A-----------------A------T-----T-----------AT-A--------A------------...----------------AC-------------- 4995
K.CM.96.MP535          -GA------------------------CCA---...------AATC-G--A--G---------AC--------------------A----C-T-----T-----------AT----------A--------T---...--------------C-AC-------------- 4999
```

```
                                                                                                                                                      frameshift insert in HXB2
                                                                                                                                                          Vpr premature end (HXB2 only)
B.FR.83.HXB2               AATGAAGCTGTTAGACATTTTCCTAGGATTTGG...CTCCATGGCTTAGGGCAACATATCTATGAAACTTATGGGGATACTTGGGCAGGAGTGGAAGCCATAATAAGAATTCTGCAACAACTGCTGTTTATCCAT.T.TTCAGAATTGGGTGTCGACATAGCAGAATAGG 5804
Vpr (frameshifted)                                                                                                                                                 F__R__I__G__C__R__H__S__R__I__G
Vpr                        _N__E__A__V__R__H__F__P__R__I__W__.__L__H__G__L__G__Q__H__I__Y__E__T__Y__G__D__T__W__A__G__V__E__A__I__I__R__I__L__Q__Q__L__L__F__I__H_#_#F__Q__N__W__V__S__T__*_
01_AE.CF.90.90CF11697      -T-------------------------CCC---...--------------A--GT-------A-C------------------A--G--T---------------C-T---------A------G-T---...-------------C-A--------------- 5752
01_AE.CN.05.FJ051          ---------------------------CCC---...--------------A--G--C-------C------------------A--G--T---------------C-T------G--A------G-T---...----------------A--------------- 5813
01_AE.CN.06.FJ054          ---------------------------CC----...--T----------A--GT-C-----A-C-A----------------A--G--T----T----------C-T---------A------G-T---...----------------A---------G----- 5828
01_AE.HK.x.HK001           ---------Y-----------R--CCC---...-----A------A--GT-C-------T----------------A--G--T----T---C----C-T---G-----A------G-T---...----------------A--------------- 5165
01_AE.TH.01.01TH_R2184     -C----------A-------------CCC---...--------C---A--G--C-----A-C-----------------A--G--T----T-----------T---------A------G-T---...----------------A--------------- 5003
01_AE.TH.02.OUR769I        ---------------------------CCC---...--------------A--GT-C-----A-C-A----------------A--G--T----T-----------T-A--------A------G-T---...----------------A--------------- 5002
01_AE.TH.90.CM240          ---------------------------CCC---...-----A--------A--GT-C-----A-C-A----------------A--G--T----T----------GT---------A------G-T---...----------------A--------------- 5371
01_AE.US.00.00US_MSC1164   ---------------------------CCC---...--------------A---T-C-----T-C-A----------------A--G--T----TC--G---T-T--T---------A------G-T---...-----G---------A--------------- 5008
02_AG.CM.02.02CM_4082STN   -----------C--G-----------CCA---...--T-----A-----A-------------C--A----------------A--G--T----T-----------T------T-------G-A---...----------------A--------------- 5002
02_AG.EC.x.ECU41           C-----------G------------CAG---...-------A-----A---T-------A-C--A----------------A--G--T-T---T-----------T-------T-A------G-T---...----------------AC-------------- 4899
02_AG.GH.03.GHNJ196        C-A-----------------------CCG---...--T-----A-----A---T-------A-C--A-------C-------A--G--T----T----------CT---------A------G-T---...----------------A--------------- 5832
02_AG.NG.01.PL0710         C---------------C---------CCA---...--A-----AC----A--------------G--A----------------A--G--TAG---T-------C----T---------A--------T---...----------------A--------------- 4987
02_AG.NG.x.IBNG            C---------------C---------CCG---...-------A-----A-----------------A----------------A--G--TA----T-----------T---------AT-----G-T---...----------------A--------------- 5328
02_AG.SE.94.SE7812         C-----------C-------------CCG---...-------A-----A-----------------A----------------A--G--T-G---T-----------T-------------------G-T---...----------------A---------G----- 5175
02_AG.SN.98.MP1211         C----------C--------------CC----...-------A-----A-----------------A----------------A--G--T-G---------------T-A-----T-A------G-T--C...----------------A--------------- 5000
03_AB.RU.97.KAL153_2       -G--------------C---------G-G---...--T---A------A---T-C---------------A-----------------------------------------------------------T---...-----G---------A--------------- 5029
04_cpx.CY.94.CY032         -------------G-----------ACCC---...---------C---A--G---------A-C--------A----C---A--G--T----T-----------T---------A--------T---...-------------C-A-----T--------- 5169
05_DF.BE.x.VI1310          C-------------------------CCA---...------A---------GT-------A-C--C----------------AG----T----T--G---G--AT-A-----G--A-------------...-------------C--C--------G----- 5185
06_cpx.AU.96.BFP90         -G-----------------------ACCA---...------AA-A-------GT----T---AGC------------------A------T----TC--T------CA-------T---A------G-----...--------C-----C-AC--------G----- 5831
08_BC.CN.97.97CNGX_6F      C-G--------C-----C--------ACCA---...--T----------A---T--G---------C--------------A------T---A-T---C-------A---------T-A-----C--T---...-------------C-AG-------------- 4986
09_cpx.GH.96.96GH2911      ---------------------------CC----...--------------A-------------T--C----------------A------T----T-----------T-------T-A------G-T---...-------------C-AT------------- 5005
10_CD.TZ.96.96TZ_BF061     ----------------C-------------------...--T---AA-C----A---------------------A-----C---A------T----T----------C--T------------------------...----------------CC------C------- 5182
11_cpx.GR.x.GR17           C-------A-----G--------A---CCA---...-------A-----A-----------A-C--C-----A---------A------T----T-----------T---------A------G-T---...-----G---------C-A--------------- 5105
12_BF.AR.99.ARMA159        -------------------C------CCA---...--------------A--G---------A-C--C----------------AG----T--G--T-------G--AT----------A-------------...-------------C-AG-------------- 5808
13_cpx.CM.96.1849          C------------A-------------CC----...--T-----A-----A--G---------A-C-----C-------------A------T----T---C------G--A---------A--------T---...-------------C-A--------------- 5208
14_BG.DE.01.9196_01       ---------------------------CCC---...--------------A--GT-------A-C------------------A------T---------------A--A-----G--AT-----------------...-------------C-A--------------- 5320
15_01B.TH.99.99TH_MU2079   -----------A-A-TT--C-...----CCC---...--------------A--GTCC------ACC-A----------------A--G--T----T--G--------T----C-----A------G-T---...--T-----------AC------------C 5196
16_A2D.KR.97.97KR004       -GA-----------------------CCA--A...------A-------A---T----T---C------------C--A------T----T-------------------A-----------T---...-----------A---A--------------- 5166
17_BF.PE.02.PE02_PCR0155   ---------------------------CCA---...------A-------A-------------A-C--C----------------AG----T----------G--AT-----------A-------------...A-----------C-AT------------- 4993
18_cpx.CU.99.CU76         ----------------C------C-ACCC---...---------C---A--GT-------A-C------------------A------T----T----------A------------------------T---...-----G-------C-A--------------- 5108
19_cpx.CU.99.CU7          ----------------C---------CCA---...--T-----A-----A---T--G----AC---C------------------A------T----T----------A-----------A--A-----T---...-------------C-A--------------- 5029
20_BG.CU.99.Cu103         ---------------------------CCC---...--T-----------AGTG---------A-C------------------A------T-----C---------A--A--------A-------------...-------------C-A--------------- 5269
21_A2D.KE.91.KNH1254       -G--------------------------CC----...------A--C---A-------T-------C----------C----------T----T---C------------G-T-A--------T---...----------------A--------------- 5008
22_01A1.CM.01.01CM_0001BBY--------------------G--CCC---...---------------AG-----------A-C--------A---------A--G--T----T-----------T-------T-A------G-T---...----------------A---C----------- 5011
23_BG.CU.03.CB118          ---------------------------CCC---...--T-----------AGTG---------A-C------------------A------T----------------A--A-------AT------C-----...-------------C-A--------------- 5246
24_BG.CU.03.CB378          ---------------------------CCC---...--------------A--G---------A-C------------------A------T-------C--------A--A---------A-------------...-------------C-A--------------- 5228
24_BG.CU.03.CB471          ---------------------------CCC---...--T-----------A-----------A-C------------------A------T-------C--------A--A-----G--AA-------------...-------------C----------------- 5243
25_cpx.CM.01.101BA         C------------G------------CCC---...--T-----A-----A--GY---------A-T------------------A------T-T--------------A--A--------A--------A---...-------------C-A--------------- 5011
27_cpx.CD.97.97CDKTB49     -C-------------------------CCC---...-------A-C----A--G---G-T---A-C------------------AG----T----T-----------AT--------------------T---...----------------AC------------- 5011
28_BF.BR.99.BREPM12609     -G---G-----------------------A---...--------------A------------------------------A------T-----------------------A--------A---...-----------A------------------- 5169
29_BF.BR.02.BREPM119       -G---------G-------------AG-G---...------A----------T--------------------------A-------------C---------A--------------------T---...----------------A--------------- 4977
31_BC.BR.02.110PA          C-G--------C-----C--------ACCA---...--T----------A---------------C-----A----------G--G------T-------------------------A------G-----...-------------C-AG------C------- 5237
32_06A1.EE.01.EE0369       -G-----------A------------CCA---...------A-------A--GT-------A-C--C----------------A------T--G----------------A---------A--------T---...--------C------AAC------------- 5438
33_01B.MY.05.05MYKL007_1   ---------------------------CAA---...--------------T-----C------C-C----------------A--G--T----T--G-------C-T---------A------G-T---...----------------A--------------- 5182
34_01B.TH.99.OUR2478P      ---------------------------CCC---...--------------A--GT-C--------C-A--------------CA--G--T----T-----------T---------A------G-T---...----------------A--------------- 5005
35_AD.AF.05.05AF095        C---------------C---------CCA---...--T-----A-----A---T---T--CA-C----------------------T----T----------CT---G-----------------T---...--------C-----CA---C----------- 4996
36_cpx.CM.00.00CMNYU830    C-------------------------CCA---G..--------A-----A---T-------------A----------------A--G--T----T----G-----CT---------A------G-T---...-------------C-A--------------- 5009
37_cpx.CM.00.00CMNYU926    -------------G-----C------CCA---...--T------C--------T---------------C----------------CA------T------------A---AT-A--G---T-A--------------.-.-------------C-A--------------- 4994
38_BF1.UY.05.UY05_4752     C-------------------------CCA---...--T---A-------A---T------CA-C--C--------C--A---A------T----T---------A-----------AA-------------...-------------C-AT--------G----- 5214
39_BF.BR.03.03BRRJ103      GG---------------------------A---...--------------A-----C---------------------------------A-G----------------------------------------T---...-------------C----------------- 5292
40_BF.BR.04.04BRRJ115      -G-------------------------CCA---...--T---AA-----A---T---------------------------A-------------------------------------A---------T---...-----------A----AC------------- 5355
42_BF.LU.03.luBF_05_03     ---------------------------CCA---...--T-----T--G--A-----------------------------------A----------T------------------------T-------T---...------------------------------- 5341
43_02G.SA.03.J11223        C---------------------C---CCG---...--------AY---A-----------A-C--A----------------A--G-------T---------Y-T-A----R--A------G-Y---...----------------A--------------- 5328
U.CD.83.83CD003            C-------------------------CCG---...--T---AA------A--GT-------A-T--C--------------CA------T----T---------G-----------------A-------------...-------------C-A--------G----- 5298
U.CD.90.90CD121E12         C-G-----------------------CCG---...--T---AA------A--G--------A-T--C-------C-----CA------T----TC-----------------------A-------------...-------------C-A--------------- 5002
U.GR.99.GR303              C------------------C------CCA--C...--T-----C---A-------------A-C--------A----------A------T----T--C-------T-----------A--------T---...-------------C-A--------G----- 5074
U.NL.95.U_NL_95_H10986_D1  -GA-----------------------CCA---...--T--------------------A-C--A----------------A--G--T----T-----------A-----------------CT---...-------------C-A--------------- 5688
CPZ.CD.90.ANT              -----G--A--A--------C--ACA-CC-ACA...T-A--ACA-C-T--AA-TTGGG-A----C--A---C--A---T-A---AG-------C-G-TTT-G---TTC-C--T---A--GCT-------CA---...-AT---CAC--A---GCT--------G----- 5260
CPZ.CM.01.SIVcpzCAM13      -----------A-A---C-A---C--A-G-----...--T--CCAA-----C--G-----T---ACC--C--------C--G----A--G-----G--A--T--------C--A---------A------G-T---...-T---G-------C-AC-------------- 5386
CPZ.CM.05.SIVcpzEK505      -----G---C-GC-C--C------C--CC-----...--A-----AC--------T-CT-----A-T--A-----A-----C----AG-----A--G-----C--T--G-CA--A---------T--------A---...-AT--G-----C----A--------G----- 5375
CPZ.CM.05.SIVcpzLB7        C-G-----A--A-----C--------ACCA---...-----CCAA--G--A---T----A-----C--C-----A--C--C----AG-----TC----T--C---------C-----------C--------T---...-A---------C----A--------------- 5358
CPZ.CM.05.SIVcpzMB66       -GG-----A--A-----C--C---C-AGA-----...------CAGC----A--G----------ACC---------A----------------G-----G--T----------C------------T--------T---...-A-------------AT--------------- 5349
CPZ.CM.05.SIVcpzMT145      C-A-----A--------C-----C--ACCC---...T-A---CAC----AG-------T--------A----A--CT-C------G-----AC----T-------------T--------T-A------G-----...-----------C--C-AT-------------- 5345
CPZ.GA.88.GAB1             ---------A--A--------C-----ACCA---...T-G---CAA-----A---TTC--T-----C---------A-----A----T------A--G-----C--------CT-A-----CT----------------...-T---C------C-AG-------------- 5867
CPZ.TZ.00.TAN1             C-A-----A--G-A---C-----A--ACCAATA...T-A--G--GG----AA-TTGGG---TCACC-T------A--CT-C----AG-----AC-G-AAT----C-AG--CT----GAG-GCTT------C---C...-ATC-CCA---T---ATC--C----------- 5465
CPZ.US.85.CPZUS            G-A--G---T-GC-C-----------ACCC---...--A--A--------A--GT-C-----CAGT---------A--C--------AG-----A-----A--T-----G--CT-A---------TT----------------...-AT--G------C---AAT--------G----- 5863
CPZ.CM.x.SIVcpzDP943       -----G-----AC----C-------C-AGCA---...T-A-----A------A-----C--A-----C--A-----A-----A---------T--A-----T--C--T------A---------G--CT-------T---...-AT--G------C----A--------------- 5389
CPZ.CM.x.SIVcpzMB897       C-G-----A--A--------------ACAG---...--T----AT------A--G--C--T---A-C--A-----A--C---------G--G--T--G--T--C-----G--C----------T-----------T---...-A---------C--C-A--------------- 5369
N.CM.02.DJO0131            -----------GC-C--------A----C----...--A-----G------------C------A-C--A-----A--C--C----AG--G--A--G--A--T--T--G--A--A------T-A--------------...-AT--G-----C--C-AG-------------- 5309
N.CM.04.04CM_1015_04       -----------GC-C--------A-------------...--A-----G------------C--------C--------A--C--C----AG-----A--G--A--T--C--G--GY-A------T-A--------------...-AT--G-----T--C-AG--C----------- 5308
N.CM.04.04CM_1131_03       -----------GC-C--------A-------------...--A-----A------------C--------C--A-----A--C--A----A------A--G--A--T--C--G--G--A------T-A--------------...-AT--G-----T--C-AG--C----------- 5307
N.CM.95.YBF30              -----------GC-C--------A-------------...--A-----G-----A-----C------A-C--A-----A--C--C----AG--G--A--G--A--T--C--G--A--A------T-A--------------...-AT--G-----C--C-AG--C----------- 5397
N.CM.97.YBF106             -----------GC-C--------A-------------...--A-----G-----A-----C------A-C--A-----A--C--C----AG--G--A--G--A--T--C--G--A--A------T-A--------------...-AT--G-----C--C-AG--C----------- 5395
O.BE.87.ANT70              GCA-----A--A--------C-----CC-----...--A--C-C---G--A---T-C--T-----G-----------C-------T------TATG--A--T--------CT-A------T--------T--------C-----...-AT--------A--C-A------T-------- 5857
O.CM.91.MVP5180            G-A-----A--A--------C-----CC-----...T-A--A-C--GT-----GT-C--T-----G--------A--C-------A------TATG--A--T--------CT-A---------A--------A--------C-----...-AT--------A--C-A------T-------- 5832
O.CM.96.96CMABB637         GCA-----A--A-----C---------GC-----...--A--A---------A---T-C---------T--------A--C--C----A---G---ATG--A--C--------CT-A-----G--AA-A---GC-----...--T--------A----A------T-------- 5287
O.CM.98.98CMA104           -CA-----A--A---------------CC-----...--AY-C-C---G--A---T-C---------G--------A--C-----TT------TATG--A--T-C---G--CT-A-------A-A---A---GC-----...--T--------A--C-AG-----T-------- 5289
O.CM.99.99CMU4122          GCA-----A--A--------C-----ACC-----...--A--C-C---G-----T-C--T-----G------------C--G----T------CAT---A--T--G-----AT-A-------------A---C-----...--T--------A----A------T-------- 5288
O.FR.92.VAU                GCA-----A--A---------------A-AC-----...--G--A-C---G--A---T-C--T-----G--------A--C-------A------AAT---A--T-C------CT-A-------A-AA-A----C-----...-AT--------A--C-A------T-------- 5369
O.SN.99.SEMP1299           GCA-----A--A--------C-----ACC-----...--A--G-C---G--A---T-C--T-----C------------C-------T------TATG-----T--------CT-A---------------A---GC-----...--T--------A--C-A------T--G----- 5856
O.US.99.99USTWLA           -CA-----A--A--G--------C---CC-----...T-A-----A--G--A---T-C--T--------------A--C--C----A-------ATG--A--T--------CT-A------A-A--A----C---C...-AT--------A--C-AG-----T-------- 5286
```

```
                                                                    Tat exon 1 start  Vpr end
B.FR.83.HXB2             CGTTACTCGACAG............AGGAGAGCAAGAAATGGAGCCAGTAGATCCT..AGACTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAAAAGTGTTGCTTTCATTGCCAAGTTTGTTTCATAACAAAAGCCTTAG 5960
Vpr (frameshifted)       __V__T__R__Q__.__.__.__.__R__R__A__R__N__G__A__S__R__S__##*
Tat exon 1                                                   _M__E__P__V__D__P__##_R__L__E__P__W__K__H__P__G__S__Q__P__K__T__A__C__T__N__C__Y__C__K__K__C__C__F__H__C__Q__V__C__F__I__T__K__A__L__
A1.GE.99.99GEMZ011       -A--GAG-A-AG-.............--A----GC--G-----T-------------..-AC-------------T----G----------------------G-----T--C---------------GG--------C----C--TC---A-----G----- 5143
A1.KE.00.KER2008         -A---T--CCAG-.............--A-----C--GG----T--G----------..-AC----C-------C-A--C--G------------C----C-----A---G--T-TC---------------A-----T-C-----C--GC-G-AC----G----- 5158
A1.KE.00.KNH1144         -A---T---GGGA.............--A----GC--G-----T-------------..-AC-----C------C--C--G------------C----C-----A---G--T--C---------------A---------GTA---C--TT-G-AC----G----- 5164
A1.KE.00.KSM4024         -A---T----GG-.............--A----GC-T------T--G----------..-AC--------A----C-----------C------C--------------G---A--T--C----------------------T-----C--C--TT-G-AC----G----- 5155
A1.KE.00.MSA4069         -A---T--C-GG-AGAAGAGGG....--A----TC--G-------------------..GAC-------------C--C--G------------C---AC-----G---A--T--C--------------------------AG---C--TT-G-AC----G----- 5167
A1.KE.00.NKU3005         -A---T----GG-.............--A----TC--GG----T--G----------..-AC-------------T----G------------CG---C-----G---G--T--C---------------A-------TT------C---CAG--C----G----- 5158
A1.RU.00.RU00051         -A---A----AG-.............--A----TC--G-----T-------------..-AT-------------T----G------------------------G---A--T------------------GG-------------C--TY-G-RC----G----- 5260
A1.RW.93.93RW_024        -A---T--C-GG-.............--A----GC--G-----T-------------..-AC-------------C----G------------C-----------A---G--T--C---------------A------T-C-AG---C--TT-G-AC----G----- 5161
A1.SE.95.SE8891          -A---T----AG-.............--A----TC--GG----T-------------..-AC-------------T----G----------GC-----------G-GCG--T--C---------------A------T-C-------C--TC-G-AC----G----- 5148
A1.SE.95.UGSE8131        -A---T----GG-.............--A-----C--GG----T-------------..-AC-------------C--C--------------C----C-----A---G-AT-TC-------G--------A-------TT------C--TCAGCAC----G----- 5310
A1.TZ.01.A173            -A---T----GGAGGG..........--A----TC--G-----T-------------..-AC-------------T--C--------------GC----------G-CCG--T--C--C------------A------T--------C--TC-G-AC----G----- 5161
A1.UA.01.01UADN139       -A--GTG-A-AGA.............--A----GC--G-----T-------------..-ACA------------T-------------------------------G-----T--C-------G--------GG--------A----C--TC-G-A-----G----- 5158
A1.UG.92.92UG037         -A---A-ATT-GAGGG..........--A----TC--GG----T--G----------..--C-------------C----G----------------C-----A---G--T--C------GT--------A----------TGC--C--TC-G-AC--G-G----- 5978
A1.UG.99.99UGA07072      -A----G---AG-.............--A-----T--GG----T-------------..-AC-------------C----G----------GC-----------G--CG--TC-C---------------A------T-C-------C--TC-G-AC----G----- 5158
A2.CD.97.97CDKTB48       -A---T--C-AGA.............--A----T---GG-------------C---..-ACT------------C----G----C--------------------A--------------------------A----------AGG--C--TC---AC----G----- 5299
A2.CY.94.94CY017_41      -A---T----AGA.............--A----T---GG-------------C---..-A---------------C----G--------------------------A--T--------CGC-------A---------GT-G--C--T----AC----G----- 5316
A.SN.01.DDI579           -A---T----AGAGGG..........--A----TC--GG-------------C---..-AC-----------C-C----G-------A--------------G------T---------GCA-------GG----T--------C--TC-G-AC----G----- 5160
A.SN.01.DDJ369           -A--------AGAGGG..........--A----GC--GG-----------------..-AC-----------C-A----G-------A--------------G------T---------GCA-------GG----T----C---C--TC-G-AC----G----- 5163
A.SN.96.DDJ360           -A---T----AGAGGG..........--A----TC--GG-------------C---..-AC-----------C-C----G-------A--------------AT--G--T---------GCA-------GG----T--------C--TC-G-AC----G----- 5160
A.CD.02.02CD_KTB035      -A---T----AGC.............--A---AGG--GG----A-TG---------..-AC------------------G------------------C---C-G---A---------------------A-------T-CAG---C--TC---A-----G----- 5953
A.CD.97.97CD_KCC2        -A------C-AGC.............--AG---T---G------------------..-AC-----------G------G------------------C-----A-CGG--T-----------A-------A---C---C-----C--TG---AC----G----- 5955
A.CD.97.97CD_KTB13       -A---T--C-AGC.............--A---AT---G------------------..-AC------------------G------------G----------A-CGG--T--C-----GTT-------A-------T-CA-G--C--TC---A-----G----- 5163
B.AR.04.04AR151516       -A------TG--A.............------------G-----------------..--------------------------------------G---C---------A------------------------C-------------TG-G------G----- 5181
B.AU.87.MBC925           -A-----------.............--A-----------------------C---..--------------------------A---GG-----------CC--------------------------------------------------------- 5970
B.BO.99.BOL0122          -A-C---TAC---.............-AA---------------------------..----------------------------------------------------------C----------------------------T----A-----G----- 5179
B.BR.03.BREPM2012        -A---T-------.............------------------------------..-------------------------------------C---------A---------------------C---------------C-------A-----G----- 5384
B.CA.97.CANB3FULL        -A-------G---AGG..........--AGC-AG----------------------..-------------------------------GG----------A--------------------------------------------------GC----G----- 5251
B.CN.05.05CNHB_hp3       -A-----------.............------------------------------..---T----------------------------G----------A---------------A--CA-------------------------C--A-----G----- 5958
B.CO.01.PCM001           -A-C...-A----.............--AG--------------------------..--------------A-----------------G---------------------------C--------------------------C--A-----G----- 5161
B.GB.83.CAM1             -A---T--A----.............-------T----G-----------------..--------------------------C-----GG---------------------------C-------------------------------------G----- 5960
B.GB.86.GB8              -A-----------.............---------C--------------------..---A----------------------------GG----------T----------------------------------------C---T-G-G-----G----- 5979
B.GE.03.03GEMZ010        -A-----------.............--------------T---------------..--------------------------C-----G-----------------T-----------A-------------------A-----------------G----- 5164
B.IT.05.SG1              -A-G--------T.............------------------------------..-AC--------------------------A--C--G--------C-------------------------A---C-----TA--------C-T-----G----- 5762
B.JP.05.DR6538           -A---T--C----AGG..........--AGC-AG----------------------..-------------G-----------C-----GG----------AT------------------C-------------------------C--A-----G----- 5966
B.KR.05.05CSR3           -A-A-T-------.............------------------------------..---------------T-----------A----G-G---C-----A------------------C----A--------------------C--G-----G----- 5984
B.NL.00.671_00T36        -A---------GA.............G------T---G-------------C---..-------------------------------G----------A------------------------------------------------C--G-----G----- 5524
B.RU.04.04RU128005       -A-----------A............------------------------------..---------------G-----------------G-----------GCC--T---------------------------------------------A-----G----- 5450
B.TH.00.00TH_C3198       -A------A----.............------------------------------..---T------T--------------------G------------------------------C---------------------------------A-----G----- 5173
B.UA.01.01UAKV167        AA---T-------.............------------G-----------------..--------------------------------GG-------C----A-------------------------------------------G-----G-G----- 5185
B.US.04.ES10_53          AA-----------.............------------------------------..---------------A--------------------------A-TCA--------------------A--------C---G--C---------------G----- 5968
B.US.99.PRB959_03        -A---T-------.............------------------------------..---------------------------C-----GG----------------------C-----------------------------C---T---G-----G----- 5179
C.AR.01.ARG4006          -A---TAA-----.............------------------------------..-A---------------C------------------------C---C--T--A-----------A--CA--A-----T--------C--TC-G-------G----- 5143
C.BR.04.04BR013          -A---TGA-----.............--A---------------------------..-AC--------------C----------------------CC---C-AT-G---T---------GC--CA-T-A-----T--------C--TC-G-------G----- 5423
C.BW.00.00BW07621        TA---TG------.............--A---A-------------A---------..-AC--------------C------------------------C-----AT--G----T---C----GT---A---A-----T-T-------C--TCAG-------G----- 5312
C.CN.98.YNRL9840         -A--TGGA-----.............--A---------------------------..-AC--------------A----------------------------C-AT--C--T---------T-C---A---AC----T-T-------C--TCAG-------G----- 5146
C.ET.02.02ET_288         -A-C--AA-----.............--A---------------------------..-AC--------------T-----------C------------C---C--T-----T----------------A---A-----T-C-------C--TC-G-------G----- 5161
C.GE.03.03GEMZ033        -A--TTG------.............--A---------------------------..-ACT-------------C-----------------------------T--G---T---C-----A--C----A-----T-C--C---C--TC-G-------G----- 5140
C.IL.99.99ET7            -A---TG-C----.............--A---------------C-----------..-AC--------------C--------------------GC---------AT-CA--T--C------GCC--CA---A-----T-T-------C--TCAG-A-----G----- 5140
C.IN.99.01IN565_10       A---TTG------.............--A---------------------------..-AC--------------C------------------------------C-AT--C--T-----------T---A---AC----T-T-------C--TCAG-A-----G----- 5340
C.KE.00.KER2010          -A--CT------A.............--A-------------T-------------..-A---------------T-----------C------------------C-AT-CG--T------------A---A---A-----T-T-------C--TCAG-A-----G----- 5140
C.MM.99.mIDU101_3        -A--GTGA-----.............--A---------------------------..-AC--------------C----------------G----------C-AT---------------CGC---A---AC----T-T-------C--TCAG-A-----G----- 5313
C.MW.93.93MW_965         -A--TTG------.............--A---A-----------C---A-------..-AC----A---------C------------------------------AT-GG--T--C--C---CGC--C----AC----T-T-------C--GCAG-------G----- 5146
C.SN.90.90SE_364         -A---TA-A----.............--A---------------------------..-AT--------------C------------------------------AT--G--T---------CGC---A---A-----T-T-------C--TCAG-------G----- 5128
C.SO.89.89SM_145         -A--TTG------.............--A---------------------------..GA---------------C-------T---A------------------AT-----T---------CTC---A---AC----T-T-------C--TCAG-------G----- 5180
C.TZ.02.CO178            -A--TTG-A----.............--A---------------------------..-AC--------------C----------C----A----CC-----ATC-G----------------C-C---A---A-----T-T-------C---CAG-------G----- 5140
C.UY.01.TRA3011          -A-CCTAA-----.............--A---------------------------..-AC--------------C--------------------------C-AT--C--T-T-------------A--CA---A-----T-T-------C--TCAG-------G----- 5130
C.YE.02.02YE511          -AC--TA-----T.............--A-----------------A---------..-A-T-------------T-------------------G---C----C-AT--A--T--C------GCC--C----AC----T-T-------C--TCAG-A-----G----- 5134
C.ZA.04.04ZASK164B1      -A--TTG-A---C.............--A---------------------------..-AC--------------C----G------------------C-C----A-CCA-----------------A--CA---A-----TTT-------C--TCAG-------G----- 5373
C.ZA.05.05ZAPSK240B1     -A--TGG------.............--A-----------------A---------..-ACT---A---------A-----------------------C------ATGGG--T----C-------A--CA---A-----T-T-------C--TCAG-A-----G----- 5363
C.ZM.02.02ZM108          -A---TA--T---.............--A---------------------------..-AC--------------T-----------------C-G---C--------A---T--C---T-T---A---A---C--T-T-------C--TCAG-------G----- 5928
D.CD.83.ELI              -A---T-------.............--A------------T--------------..-AC--------------C-----------------GG---C-----A---G--TC-----------------A--------C--------C---T---AC----G----- 5505
D.CM.01.01CM_0009BBY     -A-----------.............--A------------T--------------..-AC--------------A---------------A-GT----------A-TCA--T---------C---------A---------C----C-------AG----G----- 5158
D.KE.01.NKU3006          AA---T-------.............--A---A--G--------------------..-AT----A------G--------------------GG---C-----A---A--------------------C----A-----T-C--G--C--------------G----- 5164
D.KR.04.04KBH8           -A---T---G---.............--A------------T--------------..-AC--------------A-------------------GG---------A----------------------C-------------------C---T---AG----G----- 5918
D.TD.99.MN011            AA--CAGA--ACA.............GCA------------C--------------..-ACT-------------C---------------A-GT---T-----G-TCA--T----------TT------AC-------GC----C--------------G----- 5183
D.TZ.01.A280             -A------T----.............--A---A--------T-------C------..-AC--------------T--------------------GG---C-----A-----T------------------A-----T---AA---C-----------G---G----- 5160
D.UG.99.99UGK09259       -A-----------.............--A--GA-----------------------..-AC--------------T------------------G-G--CC-----A---A--T-------------------C----A----------CA---C---G-----G---G----- 5164
D.YE.01.01YE386          -A--CT-------.............--A---AT----------------------..--C--------------C--------G---------GG---C------T-G----------TG---C---A---------C-------C-----G-------G----- 5158
D.YE.02.02YE516          -A---T---C---.............--A------------T-------C------..-AC-----------------------------G---G-GT---C-----G-CC---T----------------A-------C----T---------------G---G----- 5170
D.ZA.90.R1               TA---T-------.............--A------------T--------------..-AC--------------C--------------------G---C-----A---G--TC------------------A-----------GC----C-------T-C----G----- 5308
F1.AR.02.ARE933          -A-----------.............--A--------------T------------..-ACT---T-------C----------C------C---C-------------T--------CG---------------TT-CCGG--C--AGC-CGG--G-G----- 5253
F1.BE.93.VI850           -A--GT--C----.............--A---T-----------------------..--CT---T-------T----------C------C---C--------A--T--------------CGA------------------TGG--C--T-C---G--G-G----- 5299
F1.BR.01.01BR125         -A---TG------.............--A--AT-----------------------..-ACT---T-------C----------C--A---CC--C-----A---A--T--------CG---------------TT-CTGG--C--TGC---G--G-G----- 5404
F1.ES.x.P1146            -A-----------.............--A---T-----------------------..-ACT---T-------T----------C------C---C-C---A---A--T---------------T---------------T-CCAC--C--TGC---G--G-G----- 5236
F1.FI.93.FIN9363         -A---T-------.............--A---T------------T----------..-ATT---T-------C----------C-----CC---C-----A---A---------CG----------------T-CTGG--C--TGC---G--G-G----- 5288
F2.CM.02.02CM_0016BBY    -A---T-----A..............--A--AT---------AATG----------..-A-----T-------C-------------------G-A-C-----AT--A--T---------G----------------------C-G--C--T-C--GC--G-G----- 5152
F2.CM.95.MP257           -A---T-----A..............--A--AT---------AGTG----------..-AC----T-------A-------G---------G-G--C-G---AT--A---------------------------------C-G--C--T-C--GG--G-G----- 5171
F2.CM.97.CM53657         -A---T-----A..............--A--CT---------AATG----------..-A-------------T-------------------G-G--C-----AT--A--T--C----------------------------T-G--C--T-C--GG--G-G----- 5152
G.BE.96.DRCBL            -A---T--C---AGG...........--A---T--GG----T--G----------..-A------------------G--G------------C-C---A-----------GTGT-GCTGCATTG...--------C--TC-G-AC----G----- 5916
G.CM.01.01CM_4049HAN     TA------C----.............--A---T--G-----C--G----------..-AT------------T----G--G---------T---C-C---AT--C-----------------GG--------TA---C--TT-G-AC----G----- 5161
G.CU.99.Cu74             -A---G--T--G-.............--A---T--GG----C--G----------..-AC------------T-------G------------------------A---------------GA-------GG----T---AG---C--TC--CAC----G----- 5563
G.ES.99.X138             -A--GTG-C----.............--A---T--GG----C--G----------..-AC------------T-------G---------GC---------------C-------------G-------GG--------------C--TC-G-AC----G----- 5400
G.GH.03.03GH175G         -A---TC-C----.............--A---T--GG----C--------C---..-AT------------T-------G---------C---C-C---A---A--T----C---GT-------GG----T--------C--TT-G-AC----G----- 5999
G.KE.93.HH8793_12_1      -A------T--G-.............--A---T--G-----C-------------..-AC------------C----G--G----------G---------A---G---------------TA-------GG----T--------C--TC-G-AC--G-G----- 5359
G.NG.01.01NGPL0669       -A------C--G-.............--A---T--GG----T--G----------..-AT------------T----G--G---------G----------A---C---------------------------GG-----------------C--TC-G-AC----G----- 5164
G.PT.x.PT2695            -A------C--G-.............--A---T--GG----C--G----------..-AC------------T-------G--------------------------A----------------GT-------GG-----------------C--TC-G-AC----G----- 5901
G.SE.93.SE6165           -A----A-C--G-.............--A--G-T--GG----C--G----------..-AC------------T----G--G--------------C-C---A---A-T-T--------GT-------GG-----------------C--TC-G-AC----G----- 5356
H.BE.93.VI991            -A-------G---.............--A---T--G-----C--------------..-AC-A-----------C----------C-----G----------A-----T-----------------C----A----------T-A--C--TT---AG----G----- 5348
H.BE.93.VI997            -A---T-------.............--A---T--------C--------------..-AC-A-----------C--------A-------------C-----A-----T------------C----CA-------------T-G--C--TT---AG----G----- 5283
H.CF.90.056              AA-----------.............--A---T--------C--------------..-A--------------C------------------C-----------A-----T-----------------C----A----------A-G--C--TT---AG----G----- 5306
J.CD.97.J_97DC_KTB147    -A---TA--C---.............--A----GG--G-----T------------..-ACA------------T--G----G--C-------G----------A-C-G--T-----C----G--------A---------C-T-G--C--TT---AG--G-G----- 5153
J.SE.93.SE7887           -A---TA-C---A.............--A----GG--G------------------..-ACAG-----------T----------------------------------T----T------------------A------T------G--C--TC--CAG--G-G----- 5273
K.CD.97.EQTB11C          -A-CTTA------.............--A----G-G--------------------..-ACA------------T--G----------A----------C---A-C-G--T--------------------A------T---C-A--C--TT-GCAG--G-G----- 5151
K.CM.96.MP535            -A---TA-C----.............--A----GG---------T-----------..-ACA------------C--A-------------------------A-C-G---------------CG---------A----------A-A--C--TT---AG--G-G----- 5155
```

```
                                                         Tat exon 1 start   Vpr end
B.FR.83.HXB2               CGTTACTCGACAG............AGGAGAGCAAGAAATGGAGCCAGTAGATCCT..AGACTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAAAAGTGTTGCTTTCATTGCCAAGTTTGTTTCATAACAAAAGCCTTAG 5960
Vpr (frameshifted)         __V__T__R__Q__.__.__.__.__R__R__A__R__N__G__A__S__R__S__##*
Tat exon 1                                                     _M__E__P__V__D__P_##_R__L__E__P__W__K__H__P__G__S__Q__P__K__T__A__C__T__N__C__Y__C__K__K__C__C__F__H__C__Q__V__C__F__I__T__K__A__L__
01_AE.CF.90.90CF11697      -A---TA---GG-............--A----GC--G---------G---------..--C-------------C-----G-----------C-----------G---G-T----------A-------GG---------C-A--C--TC-G-A---G-G----- 5908
01_AE.CN.05.FJ051          -A--TTG-C-GG-............--A----GC--G---------G---------..-AC-------------T-----G-----------C---G------A---------C-------A-------GG----T---T-A--C--TC-G-A----G----- 5969
01_AE.CN.06.FJ054          -A-A-TA-C-GG-............--A----GC--G---------G---------..-AC-------------T-----G-----A-----C-----------G---G-T--C-------TT-------GG---------C-A--C--TC-GCA----G----- 5984
01_AE.HK.x.HK001           -A---TA-C-GG-............--A----GC--G-------------------..-AC-------------T-----G-----A-----C---C-----------G-T--C-------TA----C--GG---------T-A--C--TC-G-A----G----- 5321
01_AE.TH.01.01TH_R2184     -A---TG-C-GG-............--A----GC--G---------G---------..-AC-------------T-----G-----------C-----------G---G-T--C-------TA-------GG---------C-A--C--TC-G-A----G----- 5159
01_AE.TH.02.OUR769I        -A---TA-C-GG-............--A----GC--G-------------------..-AC-------------T-----G-----A-A---CG----------G---A-T--C-------TA-------GG---------T-A--C--TC-G-A----G----- 5158
01_AE.TH.90.CM240          -A---TG-C-GG-............--A----GC--G------A-TG---------..-AC-------------T-----G-----------C-----------G---G-T--C--------A-------GG---------C-A--C--TC-G-A----G----- 5527
01_AE.US.00.00US_MSC1164   -A---TA-C-GG-............--A----GC--G---------G---------..-AC-------------T-----G--G--------C-----------G-----T--C--------A-------GG----T---A-A--C--TC-G-A----G----- 5164
02_AG.CM.02.02CM_4082STN   -A---T----GG-............--A----GC--G--------TG---------..-AC-------------T--C--G-----------C-----------G--CA-T--C--------A--C----GG---------T-A--C--TC-G-AC--G-G----- 5158
02_AG.EC.x.ECU41           A----T----GG-............--A----GC--G---------G---------..--C-----T-------T--C-------------GC----------A---G------C-------T---C----GG---------A-A--C--TC-G-A---G-G----- 5055
02_AG.GH.03.GHNJ196        -A---T----GG-............--A----GC--G------T-TG---------..-AC-----T--A----C--C--G-----------C-----------G---G-T----------TA--C----GG---------T-GG-C--TC-G-AC--G-G----- 5988
02_AG.NG.01.PL0710         -A---T--C-GG-............--A----GC--G--------TG---------..--C-------------C-----G-----------C-----------G---G-T--C-------TA--C----GG---------T-G--C--TC-G-A---G-G----- 5143
02_AG.NG.x.IBNG            -A---T--A-GG-............--A----GC--G--------TG---------..--C-------------C--C--G-----------C-----------G---G-T--C-------T---C----G.---------T-G--C--TC-G-AC--G-G----- 5483
02_AG.SE.94.SE7812         -A---T----GG-............--A----GC-----------TG---------..--C-------------C--C--G-----------C-----------G---A-T--C-------TC-------GG---------T-A--C--TC-G-AC--G-G----- 5331
02_AG.SN.98.MP1211         -A---T--C-GG-............--A----GC--G--------TG---------..-A--------------C--C--------------C-----------G---G-T--C------TT---C----GG---------T-A--C--TC-G-AC--G-G----- 5156
03_AB.RU.97.KAL153_2       -A---T--A--G-..........................................-..---------------------------C------------------------------------------TC--------------------G-A----G----- 5185
04_cpx.CY.94.CY032         -A------CT--AAGGAGA.......--AG-CAGG..GG-----------------..GAC-------------C-----G-----------C----A-----A--G-T-TC---------------GG---------------C--TC-G-A----G----- 5329
05_DF.BE.x.VI1310          -A--CT-------............--A---AT---------T-------------..-AC-------------T-----------------GG---------A-C-G-----------T--------A---------AA---C--------G----G----- 5341
06_cpx.AU.96.BFP90         -A---TA-C----............--A----GG----------------------..-A-A-----------T--G---------G-----G-------------A--------------------A--------C---G--C--TC---AC--G-G----- 5987
08_BC.CN.97.97CNGX_6F      -A--GTGA-----............--A----------------------------..-AC-------------C------------------G---------C-AT--C-----------CGC---A---AC-----T-T------C--TCAG-A----G----- 5142
09_cpx.GH.96.96GH2911      -A---T----GG-AGG.........--A----GC--GG-------TG---------..-AC-------------C------------------GC----------A---A-T--C-------GA-------A----------CAC--C---T-G-AC----G----- 5164
10_CD.TZ.96.96TZ_BF061     -A--GTG------............--A---------------A------C-----..-AC-------------C--C--------C------------------AT--A---C--------GCA---A---A------T-T------C--TCAG------AG----- 5338
11_cpx.GR.x.GR17           -A---TC---GG-............--A----GC--------------G----AT-..-ACA----T-------T--G--G--------------G----------A-C-G-T----------G--------A------T---CA---C---T---AG----G----- 5261
12_BF.AR.99.ARMA159        -A---TG------............--A----T-C----------T----------..-ACT------------C--C--------CA----TC----C----------A-T----------CG----------------TT-CTGG--C---G---------G----- 5964
13_cpx.CM.96.1849          T---GTG-----A............--A----GT---------T------------..-A-A------------C--G--------------------C----------G-T--------------------------T--------C-----------G-G----- 5364
14_BG.DE.01.9196_01        -A---T--C----............--A----T---G------C--G---------..-AC-------------T--------G-------------------------C--------------------GG---------------C--TC-G-AC----G----- 5476
15_01B.TH.99.99TH_MU2079   -A---T--C-GG-AGA.........-AGCCAGG--G------C--G-------T-..-AC-------------T-----G-----G----C----------A-----T--T--C-------TA-------GG---------A-A--C--TC-G-A----G----- 5355
16_A2D.KR.97.97KR004       TA---TG-C-...............--AG-----------------T---------..-AC-------------C--------C---------GG----------A---G-T----------G--------A------T-C---G--C---C---AC----G----- 5319
17_BF.PE.02.PE02_PCR0155   -A-----------............--A----T----------T------------..-ACT----T--------C-----------C-------C--KC---------GA------------CG----------------TT-CTGG--C-------------G----- 5149
18_cpx.CU.99.CU76          TA--GTG-----A............--A----GG----------------------..-ACA------------C--G------------------------------A------T----------GA-------A------T---C-A--C--TT---AG----G----- 5264
19_cpx.CU.99.CU7           -A------C--GA............--A----T---GG-----C--G---------..-AT-------------T-----G--G--------------------A--CA-----------------------------T---C-G--C--T-C--AG--G-G----- 5185
20_BG.CU.99.Cu103          -A---G--T--G-............--A----T---G------C------------..-AC-------------T-----G--G--------------------A--G-----------GT--------GG----T--------C--TC-G-AC--G-G----- 5425
21_A2D.KE.91.KNH1254       T---GT-------............--A---------------T------------..-AC-------------T-----------------GG--C------A--G-TC-------------------A-------AC------C---T----------G----- 5164
22_01A1.CM.01.01CM_0001BBY -A---TA---GG-............--A----GC-----------TG---------..--C-------------T-----G-----------C---G------AT----T-T-------GT--------GG---------------C--TC---A-----G----- 5167
23_BG.CU.03.CB118          -A---G--C--G-............--A----G---G------C--G---------..-AT----C--------T-----G--G--------------------A--G------------T--------GG----T--------C--TC-G-A---G-G----- 5402
24_BG.CU.03.CB378          -A---G--T--GAAGAAGA......CTA----G---GG-----T--G---------..-AC-------------T--C--G--G--------------------A--G-----------GT--------GG----T--GC----C--TC-G-A----G----- 5390
24_BG.CU.03.CB471          -A---G--T----............--A----G---GG-----T--G---------..-AC-------------T-----G--G--------------------A--C------------TA--C----GG----T--------C--TC-G-AC--G-G----- 5399
25_cpx.CM.01.101BA         TA------C----............--A----GG---------C------------..-ACA---A--------C--A--------------C----------A-----T--C-------TA--C----GG---------T-A--C--TC-G-AC--G-G----- 5167
27_cpx.CD.97.97CDKTB49     TA---TA------............--A----GG---------T------------..-ATA------------C--G--------------------A-------A-C-G-T----------GC-------AC-----T---C-G--C--TT---AG----G----- 5167
28_BF.BR.99.BREPM12609     -A---A-------..........................................-..--------------C-A--C-----G---------GG---------A---A--------------------C--------------------G-A----G----- 5325
29_BF.BR.02.BREPM119       -A---T-------............------C------------------------..-------------------------------------C-----A-----------------C----GA--C--T-----------AAG-----------------G----- 5133
31_BC.BR.02.110PA          -A--CTAA-----............--A----------------------------..-ACT----Y-------C--------------------------C-AT----T--T-----------A--CA---A------T-----C--C--TC-G-------G----- 5393
32_06A1.EE.01.EE0369       -A---TA-C---A............--A----GG--------A----------A-..-A-A---A--------T--G----------G-----G---C-G-----CCA--T-------------------A-------T-CA-G--C---T---AC--G-G----- 5594
33_01B.MY.05.05MYKL007_1   -A---TA-C-GG-............--A----GC--G--------TG---------..-AC-------------T-----G-----A-----C-----------G-----T--C-------GA-------GG---------T-A--C--TC-G-A----G----- 5338
34_01B.TH.99.OUR2478P      -A--TTG-C-GG-............--A----GC--G-------------------..-AC-------------T-----G-----------C-----------G---G-T--C--------T-------GG---------C-A--C--TC-G-A----G----- 5161
35_AD.AF.05.05AF095        -A---T----GG-............--A--G-TC--GG-----T-TG----C---..GAC----T-------C--C--G-----------C--GC------A---G-T-TC-----GT----A---A------T--------C--TC-G-AC----G----- 5152
36_cpx.CM.00.00CMNYU830    -A---TA---GG-............--A----GC--G--------TG---------..-ACA------------T-----G----C------G----------A---C-T-T------------C----GG---------------C--TC-G-AC--G-G----- 5165
37_cpx.CM.00.00CMNYU926    -A------C--GC............--A----TG--G------C------------..-AT-------------------G--G--------------------A---C-------C---------C-----------T--------C--T-C--A--G-G----- 5150
38_BF1.UY.05.UY05_4752     -A-----------............--A----T-----------AAT---------..-ACT----T--------C--C--------C------GC---GC-----A-----T----------CGA--------------TT-CTAC--C--TGA--A---G-G----- 5370
39_BF.BR.03.03BRRJ103      -A---A--C----......................-----G---------------..------------------------------------G----------AT--A------------CT---C----A---------------C------------G----- 5448
40_BF.BR.04.04BRRJ115      -A---TG-----------------------T-----------T-------------..-----------------------------------G------------T--A-----T-------------------------------------G------G----- 5511
42_BF.LU.03.luBF_05_03     TA---ACAT--G-..........................---------A-------..-AGT-----------G--------------------------C----------------------G-----------------------C-------G-----G----- 5497
43_02G.SA.03.J11223        -A---T----GG-............--A----GC--GC-------TG---------..--C--------T-----C--C--G-----------C-----------G---G-T--C-----------C----GG---------T-A--C--TC-G-AC--G-G----- 5484
U.CD.83.83CD003            -A-------G---............--A---------GC-----C-----------..--G--------A-----T--C-------------C-----------C-A--G------------G---------A-----T--------C--T--G-------G----- 5454
U.CD.90.90CD121E12         -A---T---G---............--A---------GC-----C-------C---..-AG--------A-----T-----------------C-----------C-A-----T-------------------A-----T--------C--TG-G-A----G----- 5158
U.GR.99.GR303              -A--GTG-----A............--A----GG---------T------------..-ACA------------C--G--G-----------------------------C--T-------------------------T----G--C--T---CGC---G----- 5230
U.NL.95.U_NL_95_H10986_D1  -A--TTG-A----............--A----GG--G------------A------..-A-A------------T--G-----------A-------------C-A-C-G----------------C---A----------C-A--C--TT---AG--G-G----- 5844
CPZ.CD.90.ANT              A...................................C-----C--C-----CG--..GA-AC-CCT--A---TTA--C--TCCTGC-ACA--AGC---GC-A--C-AT--C-T--C------TGC-----T-AC-------CTC-C--C----C--A-----G---G- 5392
CPZ.CM.01.SIVcpzCAM13      TA-A-TA-C---A...............-----AG-GC-----------A----C---..GAGT-G-----A--------------C--C-----A-----A------AT--------------GCT-----T--C-----TATGTA---C----C--AG---G----- 5542
CPZ.CM.05.SIVcpzEK505      AA-C----CT..............CAA--GAG---G-------------------..GA-T-------------T--C-------------A-----G--A-----C-AT-----C-------GTT-------A---C--TGTGTA---C----C--A----G----- 5528
CPZ.CM.05.SIVcpzLB7        -A----A-C---A............--A---AGG---G--------A--------..--C-----T--------T------------------------C-C----A---------T----------------A---C--T-C-T----C----C--AG---G----- 5514
CPZ.CM.05.SIVcpzMB66       -A-C-A-ATTA-TCAGCAG.......--A---AG------------A--------..-ATT-----------T-------------------------------AT-----------------C-A--C---AC--C--T---C----C----CG-AG---G----- 5511
CPZ.CM.05.SIVcpzMT145      -A-A-AC-CT---............--A---AGG---G-----A---A---G---..GACT---A--A----C-------G--CA-A---G---A-AA--C-AT--G-----T---C----------T-A-----TAT-C----C--T-C--AG---G----- 5501
CPZ.GA.88.GAB1             -A-CCTC-C---A............--A---AG-TCC-----T---A--------..GAC--G-----------A--C-----C--------A-G---A-T----AT--------------C---GCT-----T-A---C--TAT-TA---C----C--A----G----- 6023
CPZ.TZ.00.TAN1             ATCATGAATC-CA............TA--.......................-----.CAGG---CA--A---G-A---------GC-GCA---G----AC-------A--C--T--C--------A--C---------------C-----C--T-CG-A------A---- 5598
CPZ.US.85.CPZUS            GA--T-CTTGGCAAGGAGAACTCCTCAAG-GAG---G-----T------------..-ACA---------------G--------C--------G---------C-AT-----------C-------------------C--TGTG---------T-CG-AG--G-G----- 6031
CPZ.CM.x.SIVcpzDP943       GA---AC-CT..............G----GAGG---------------------..GAT------------C-A-----------C--A----G---------C-AT-----------------TGT--C-----------ATG--C-----TGAG-A----G----- 5542
CPZ.CM.x.SIVcpzMB897       -A--CTG-C---A............G-A---AGG---------T---A-------..-ACT----A--------T-----------C-------------CC-C----A-----------------T-------A--C--T---T---C----C--G----G----- 5525
N.CM.02.DJO0131            GA------CT..............C--G-GAG---G-------------------..---T-------------T----------C--A--C--G--A----C-ATGGA-------------T-T--C----A---C--TATGTG---C----C--AG---G----- 5462
N.CM.04.04CM_1015_04       GA------CT..............CAA--GAG-----------------------..---T-------------T----------------A--C------A----C-AT--G--------------T--C----A-----TGTG----C----C--AG---G----- 5461
N.CM.04.04CM_1131_03       GA------CT..............CAA--GAG-----------------------..---T-------------T----------C--A--C------A----C-AT-----------------A--C----A-----TGTG----C----C--AG---G----- 5460
N.CM.95.YBF30              GA-C----CT..............CAA--GAG---G------A------------..---T-------------T----------C--A-------A----C-AT----------------GA-------A---C---TT-TA---C----C--AG---G----- 5550
N.CM.97.YBF106             GA------CC..............CAA--GAG---G-------------------..---T-------------T----------C--A--C--G--A----C-AT--G--------------A--C----A---C--TATGTG---C----C--AG---G----- 5548
O.BE.87.ANT70              AA---AC-C-AGAGGA.........--AG--AG---------T------------..GAGG-GCCC--T---C-C--C--T----------CC-G-TCC-----AT---------------C----GA--C----A-------T-T---------G---G---G-GT--G- 6016
O.CM.91.MVP5180            AA--CTC-C-TCTAACACAAGAGGA--AG--AG---------T------------..GAGA-GCCC--T---C-T--C--T--G--CA---CC----CC-----AT---------------C----GA--C----A-------T-T-----------C--AG--G-GT--G- 6000
O.CM.96.96CMABB637         AA----C-C-TCTAACGCAAGAGGA--AG--AG---------T------------..GA-G-GCCC-----C-T--C--T-----------CCC---CC----C-A-GC-------------C-----A--C--T-A-------T-TC----------C--AG--G-GT--G- 5455
O.CM.98.98CMA104           AA---AC-C-TCTAACACAAGAGGA--AG--AG---------T------------..GAGA-GCCC--T---C-T--C--T--------A--CC-G--CC-----AT---------------C----GC--C----A-------T-T------C-TG-G-G---G-GT--G- 5457
O.CM.99.99CMU4122          AA---AC-C-TCTAACACAAGAGGA--AG--AG---------T------------..GAGA-GCCC--T---C-T--C--T--------A--CC-G--CC-----AT--G------------C-----A--C----A-------T-T-----------C--GC--G-GT--G- 5456
O.FR.92.VAU                AA---T--C-TCTAACGCAAGAAGA--AG--AG---G-----T--CA--------..GA-A-GCCC--T---C-T--C--T-----C--A--CC-G--CC-----A-GC-------------C---C-A--C----A-------T-TC----------C--AG--G-GT--G- 5537
O.SN.99.SEMP1299           AA----C-C-TCTAACACAAGAGGA--AG--AG---------T------------..GAGA-GCCC--T---C-T--C--T------------CC-G--CC-----AT--G------------C---GCA--C----AC------T-T--------TGC--GC--G-GT--G- 6024
O.US.99.99USTWLA           AA---TC-C-TCTAACACAAGAGGA--AG--AG---G-----T---------C-A-..GAGA-GCCC------C-T--C--T------AGC--CC-G--CC-----AT-CC-----------C----GT--C----A-------ATT-----------C--AG--G-GT--G- 5454
```

Rev exon 1 start
Tat Rev exon 1 end_intron start
Vpu start (ACG in HXB2)
B.FR.83.HXB2
Tat exon 1
Vpu
Rev exon 1
A1.GE.99.99GEMZ011
A1.KE.00.KER2008
A1.KE.00.KNH1144
A1.KE.00.KSM4024
A1.KE.00.MSA4069
A1.KE.00.NKU3005
A1.RU.00.RU00051
A1.RW.93.93RW_024
A1.SE.95.SE889I
A1.SE.95.UGSE8131
A1.TZ.01.A173
A1.UA.01.01UADN139
A1.UG.92.92UG037
A1.UG.99.99UGA07072
A2.CD.97.97CDKTB48
A2.CY.94.94CY017_41
A.SN.01.DDI579
A.SN.01.DDI369
A.SN.96.DDJ360
A.CD.02.02CD_KTB035
A.CD.97.97CD_KCC2
A.CD.97.97CD_KTB13
B.AR.04.04AR151516
B.AU.87.MBC925
B.BO.99.BOL0122
B.BR.03.BREPM2012
B.CA.97.CANB3FULL
B.CN.05.05CNHB_hp3
B.CO.01.PCM001
B.GB.83.CAM1
B.GB.86.GB8
B.GE.03.03GEMZ010
B.IT.05.SG1
B.JP.05.DR6538
B.KR.05.05CSR3
B.NL.00.671_00T36
B.RU.04.04RU12005
B.TH.00.00TH_C3198
B.UA.01.01UAKV167
B.US.04.ES10_53
B.US.99.PRB959_03
C.AR.01.ARG4006
C.BR.04.04BR013
C.BW.00.00BW07621
C.CN.98.YNRL9840
C.ET.02.02ET_288
C.GE.03.03GEMZ033
C.IL.99.99ET7
C.IN.99.01IN565_10
C.KE.00.KER2010
C.MM.99.mIDU101_3
C.MW.93.93MW_965
C.SN.90.90SE_364
C.SO.89.89SM_145
C.TZ.02.CO178
C.UY.01.TRA3011
C.YE.02.02YE511
C.ZA.04.04ZASK164B1
C.ZA.05.05ZAPSK240B1
C.ZM.02.02ZM108
D.CD.83.ELI
D.CM.01.01CM_0009BBY
D.KE.01.NKU3006
D.KR.04.04KBH8
D.TD.99.MN011
D.TZ.01.A280
D.UG.99.99UGK09259
D.YE.01.01YE386
D.YE.02.02YE516
D.ZA.90.R1
F1.AR.02.ARE933
F1.BE.93.VI850
F1.BR.01.01BR125
F1.ES.x.P1146
F1.FI.93.FIN9363
F2.CM.02.02CM_0016BBY
F2.CM.95.MP257
F2.CM.97.CM53657
G.BE.96.DRCBL
G.CM.01.01CM_4049HAN
G.CU.99.Cu74
G.ES.99.X138
G.GH.03.03GH175G
G.KE.93.HH8793_12_1
G.NG.01.01NGPL0669
G.PT.x.PT2695
G.SE.93.SE6165
H.BE.93.VI991
H.BE.93.VI997
H.CF.90.056
J.CD.97.J_97DC_KTB147
J.SE.93.SE7887
K.CD.97.EQTB11C
K.CM.96.MP535

```
                          Rev exon 1 start                                           Tat Rev exon 1 end_intron start                                       Vpu start (ACG in HXB2)
B.FR.83.HXB2              GCATCTCCTATGGCAGGAAGAAGCGGAGACAG......CGACGAAGAGCTCATCAGAACAGTCAGACTCATCAAGCTTCTCTATCAAAGCAGTA........................AGTAGTACATGTAACGCAACCTATACCAATAGTAGCAATAGTAGCATTAGTA 6100
Tat exon 1                G__I__S__Y__G__R__K__K__R__R__Q__.__.__R__R__R__A__H__Q__N__S__Q__T__H__Q__A__S__L__S__K__Q
Vpu                                                                                                                                     _T__Q__P__I__P__I__V__A__I__V__A__L__V_
Rev exon 1                          _M__A__G__R__S__G__D__S__.__.__D__E__E__L__I__R__T__V__R__L__I__K__L__L__Y__Q__S__
01_AE.CF.90.90CF11697     ---------------------------A---C......------G--C---C--CAGG---A--GA------AA-C--A--C----------......................AGTATTA----T------T-TC-G--T-G-A---CTAG----------G-C-GA-- 6052
01_AE.CN.05.FJ051         --------C--------------------A......------G--A---C-----G---CA--GAC------AA-C--G--C----------..............................----T----C-T-AC----T--GA---TAGT----------G-C--A-- 6106
01_AE.CN.06.FJ054         ---------------------------AG--C......------G--A---C-----G---A--GA------AA-CT-A--CG-G-------..................AGT-AATA--T------T-AC----T-GGA---TAGT----------G-C-GA-- 6127
01_AE.HK.x.HK001          ---------------------------AG--C......------G--A---C--G--G---G--GGC------AA-C--G--C----------..........AGTATTTGTAAA-TA----A------T-AC----C-T-A---TAGT----------G-C-GA-- 5473
01_AE.TH.01.01TH_R2184    ---------------------------AG--T......------G--A---C-----G---A--GA------AA-C-----C--G-------..................AGTA-TA-A--T------T-AC----T-GGA---TAGT-------C--G-C-GA-- 5303
01_AE.TH.02.OUR769I       --------C------------------AG--C......--------------C-----G---G-AGA-----G-A-C--A--C----------..................AGTA-TA-A--T------T-TC----T-GGA---TAGT----------G-C-G--- 5302
01_AE.TH.90.CM240         ---------------------------AG--C......------G--A---C-----G---A--GA------AA-C--A--C----------..................AGTA-TA----T------T-AC----T-GGA---TAGT----------G-C-GA-- 5671
01_AE.US.00.00US_MSC1164  ---------------------------AG--C......------G--A---C-----G---A---GC-----GAA-C--A--CG-G-------..................AGTA-TA----T------T-AC----T-G-T---TAGT----------GGC-GAC- 5308
02_AG.CM.02.02CM_4082STN  ------------------------------GC......------G--A---C-----G-CA---GA------AG-C--G--C---------G............AGTAGTAATA-CC----T----G-T-------T--G------CT----------G------- 5308
02_AG.EC.x.ECU41          -T----------------------------GC......------G--A---C-----G-C---C-GA------AA-C--G--C---------G.............AGTACAATA-T-----T-G--G-T----T--T--GA-----CT----------G-C----- 5204
02_AG.GH.03.GHNJ196       ------------------------------CC......------G--A---C-------C-----GA-T----AA-C--GC-C---------G............AGTAGTGCTA-T-----T--A-G-T----T-CT--GT-----CT----------G-C----- 6138
02_AG.NG.01.PL0710        ------------------------------GC......------G--A---C-----G-CA---GA------AA-C--G-------------............AGTAGTAAGA-T-----T------T----T--T--AA-----CT----------G------- 5293
02_AG.NG.x.IBNG           ------------------------------GC......------G--A---C-----G-C-----GA------AA-C--G--C---------G............AGTAGTAATA-T---C-T----G-T-------T--A-----ACT----------G-C----- 5633
02_AG.SE.94.SE7812        ---------------------------A--GC......------G--A---C-----G-C-----GACA------A-C--G--C---------G............AGTAGTAATA-T---C-T----G-T-A-CT--T--GAC-----T----------G-C----- 5481
02_AG.SN.98.MP1211        ------------------------------GC......------G-GA---C-----G-C-----GA-----GAA-C--G--C----------........AGTAGTAAGAATTAA-T-----T----G-T-------T--G--G---CT----------G-C----- 5311
03_AB.RU.97.KAL153_2      --------------------------------......--------------C---AG---A-------G------T-------C----------..............................--------------T----T-CT--G------C------------------- 5325
04_cpx.CY.94.CY032        ---------------------------A---T......------G--T---T---AGG---CA--GGC------AA-CT-A--C----------..................AGTATTA----T------T-TT-TTCTGGGA---CTGG----------G-C-G--- 5473
05_DF.BE.x.VI1310         ------------------------------C-......--------C---C---AGG-------G---------A-C--G--C--G-------............AGTAGTGATA------CT------T-TC-GA-C-G-T--C-A-----G-----------A-- 5491
06_cpx.AU.96.BFP90        ----T---------------------------......------CAG----C--CTGG---A---A--------A-C--G--------------..........AGTAATAATAATTAGCTA--TG----T---GG-AT--GAG----CT----------G-C----- 6139
08_BC.CN.97.97CNGX_6F     ----T---------------------------......--------C----C-----G---G--GA------AA-CT-A-------------..........................-GTATCTGTAATGTTAGATTT-GA-T-----G---G------GA-- 5276
09_cpx.GH.96.96GH2911     -------------------------------A......------GC----TC-TGCGG---A---G-A------A-C--A--C---------G............AGTATAATTA-A--AGCTC-----T----T--T-G-A-----C-AGC-------G-A--A-- 5314
10_CD.TZ.96.96TZ_BF061    ----T---------------------------......--------C----C--CA-G---G--GA------AA-CT-A-------------..AGTGCTAAATAGTAAATGTAATGT--AGTTTAC--G-AAG-GTATATTAT-A-T--CG-G---G-------A-- 5500
11_cpx.GR.x.GR17          ----T-----------------T---------......-----C-C----TC--GC-G---CA---A--------A-C--A--C--G------G..........AGTATAATTAGTT---GC-GT-ATG-TAA-TTTGCGA--AGT--AGCAT-G---C-TT-G-TAGC 5414
12_BF.AR.99.ARMA159       --------------------------------......--------------C---AGG---C------A-----------------------..............................------G-------T----T--T--GT--C-T-------------------- 6104
13_cpx.CM.96.1849         --------------------------------......--------T---TC---C----------A--------A-CT-A--C----A----G.............AGTATAATT-AATAGTAT-----T-A-TG--T---------A--AGC-------------A-- 5513
14_BG.DE.01.9196_01       ---------------------------AG-GC......------G--C---C---A-G---G--GA------AA-C--G--C----------............AGTAGTAATA-T-----T------T-A----AT--GA----TCT----------G-C--A-- 5626
15_01B.TH.99.99TH_MU2079  ---------------------------AG--C......------G--A---C-----G---A--GA------AA-C--GC-C----------..................AGTA-TA----T------T-TC----T--GA---TAGT----------G-C-GA-- 5499
16_A2D.KR.97.97KR004      ------------------------------CC......----------C---C----GG-G--AC-GGC-----GAA-C--A--C--G-------............AGTAGTAGTTCT--A--G------T-A-TT--T---A----T-GT----------GTC-GA-- 5469
17_BF.PE.02.PE02_PCR0155  --------------------------------......----------T---C---AGG---A----------AGC------C----------..............................--------------T-------T--G-G---T------------------- 5289
18_cpx.CU.99.CU76         ----T---------------------------......---------C-A---C-T-CG-----A--CA--------A-CT-A--C--G-------........................AGTAG-T---G-T-KGGK--T-TGA--------GT---A--------A-- 5402
19_cpx.CU.99.CU7          --------------------------------......------G--C---C-----GTCA----GA-------AA-C--G--C----------..............................-------T----G-T---GG--T---A-----CT----------G-C--A-- 5325
20_BG.CU.99.Cu103         ---------------------------AG--C......------G--C---C--C-GGT--CA--GA-------AA-C--G--C---------G..........AGTAGTGACAATTA---RT------T----T--C--GA-G---CT----------G----A-- 5576
21_A2D.KE.91.KNH1254      ----------------------------C---......----------C---C---A-GTG---------C-----TCCA-A--C----------..............................---------------T-------T---AG-------------C---------- 5304
22_01A1.CM.01.01CM_0001BBY---------------------------AG--C......------G--A---C----GG----A--GA-------AA-C--A--C----------.................AGTATTAA----T------T------CT-GGA---CTGTA-T-------G-C--A-- 5312
23_BG.CU.03.CB118         -A-------------------------AG--C......------G--A---C----G-----A--GA-------AT-C--G--C----------............AGTATTAACA-T-----T-G----T---CT-AC--GA-----CT---G------G----A-- 5552
24_BG.CU.03.CB378         ---------------------------AG--C......------G--A---C--TT-G--A-A--GA-------AA-C--G--C----------............AGTAGTGACA-T-----TC-----T----T-AC--GA-----CT----------G----A-- 5540
24_BG.CU.03.CB471         ----------------------------G--C......------G--A---C--CA-G---A--GA-----GAA-C--G--C----------............AGTAGTGACA-T-----T------T----T-GC--GA-----CT----------G----A-- 5549
25_cpx.CM.01.101BA        -------------------------------A......----------A---C--C-GG---A--GAC------TA-C--G--C-G--------............AGTAATAATA--C----T------T------CT--GA----TGT----------GGC--A-- 5317
27_cpx.CD.97.97CDKTB49    ----T-T-C---------------A--G--A......------GCGA---C-T-CGG---A---A--------A-C--A--C----------..............AGTATACTT-AATAGT-C-----T----G--T--GA-------AGC-------------A-- 5316
28_BF.BR.99.BREPM12609    --------------------------------......----------C-------AG---A-----A-A------T-------------------..............................----T---------T----G-CT--AT-----C------------------- 5465
29_BF.BR.02.BREPM119      --------------------------------......----------T---C-AG-C------C-G-----------C-------------G..............................----A--T----G-T-TT----T---AG--------------T-------- 5273
31_BC.BR.02.110PA         -----------------------A--------......--------C----C--CA-R---G--GA-------AA-CT-A--------------...AGTATTCAATAGTAGATGTAAT--TAAG-T-AAT-GAT--AA--G-TT-TAGA-----G---S------GA-- 5554
32_06A1.EE.01.EE0369      ----T-A------------------------A......-------C-----C--CTGG---A---A--------A-C--GA----------............AGTAACAATAGT--T-GTG-----T---GG-AT--GA----TCT----------G-C--A-- 5744
33_01B.MY.05.05MYKL007_1  ---------------------------AG--CAAGCAC------G--A---C---A-G---G--GA-------AA-C-----C----------..................AGTA-TA----T------A-AC----T-GGA---TAGT----------G-C-GA-- 5488
34_01B.TH.99.OUR2478P     ---------------------------AG--C......------G--A---C-----G---A--GAC------AA-C--A--C--G-------..................AGTA-TA----T------T-TC----T-GGA---TAGT----------GGC--A-- 5305
35_AD.AF.05.05AF095       -------A------------------------......----------A---C-----G-------GA-------AA-C-----C----------............AGTAGTAGCA-T------------T--CT---T-G-AT--CT-G----------G-C-G--- 5302
36_cpx.CM.00.00CMNYU830   ------------------------------GC......------G------C-----GTC-----GAC-------A-C--G-------------..........AGTAGTAGTAGTA-T-----T----G-T----T--T---A-----CT------C--G-C----- 5318
37_cpx.CM.00.00CMNYU926   -A------------------------------......-------C---GC-----G-C-----GA-A-----CA----G--C---------G............AGTAGTAATA-C-----T----G-T-------T--GA----TATT---------G------- 5300
38_BF1.UY.05.UY05_4752    --------------------------------......-------A-A--TC-TCA-G---C-CA-TA-------A-C--G--C---------G.......AGTATTGTTAAGTGATTTAT---ATGTC-GAATT-AACT-GTT-GC-T---GC----C-------A-- 5526
39_BF.BR.03.03BRRJ103     --------------------------------......----C----A---C---A-G--A----TTA-------A-CT-G--C--G-----CT............AAGTATTGATA-A-----T------T-TC-TA-T-GTT-GC-A--AGCT--AC-------A-- 5599
40_BF.BR.04.04BRRJ115     --------------------------------......----------T---C---AGG-C--------TG-----TA------------------..............................-------T------T------CT--GT-----C-----------G------- 5651
42_BF.LU.03.luBF_05_03    ------------------------------C-......-------ACC---C--CAG-----G-------------G---T-------------..............................--------C-----T----T--T--GTG---T------------------T 5637
43_02G.SA.03.J11223       ------------------------------GC......------G--A---C-----GYC-----GA-------AA-C--G-------------............AGTAGTTATA-T-----T------T----Y-CT--GA-----CY----------G-C--A-- 5634
U.CD.83.83CD003           -A--T--------------------G--A--.......----------A---C-----G-G---C--A-------AA-AT-G-------------..........................-----TACC--T-TCCAT-T-G-A-------------------A--A-- 5591
U.CD.90.90CD121E12        -A--T--------------------G--A--.......----------A---C-----G-GA--C-GA-------AA-AT-G--C----------..............................-GTAT--C-AATGTGG--TTTGCC-----G----------A--A-- 5292
U.GR.99.GR303             --------------------------------......------GC-----C---CG---A-A---A--------A-CT-A--C----------..........AGTATAATTTA---AGGTC-----T----T--T--GA-----C-AGC-------G-C--A-- 5381
U.NL.95.U_NL_95_H10986_D1 ----T---A--------------A------CA......A------C-A---C-T-CG-----A---A--------A-C--A--C--G-------.............AGTATTGTTA-A--GTATATGTAATGAA---T-T-G------G-----C---C------ 5991
CPZ.CD.90.ANT             -G--T----------------GAA-AGCT-GAAGGAACA------C-A--GC-G-A-G----G-A-A-A------A-C--G--------T----..............................----C-CTTAA-G--TA-TATAT-TGAGTAT-CTTTTC-T-CCTTTAGTA-- 5538
CPZ.CM.01.SIVcpzCAM13     -A--T-------------------AGC--GA......-----C-CG---TT-GCA-G----A-A-A-A------A-C--G--CG----------...............................AGT-TATT-T--GAC--T--GA--TT-----CTTT-GGTTGC 5672
CPZ.CM.05.SIVcpzEK505     -------A--------------------T--A......----------A---C-----G--A-A---G--------A-C-----C----------..........................CGTAA--CTA--T--A-GTT-TTG-T-----AAGT---G-TTC----G-TT-GC-A-- 5673
CPZ.CM.05.SIVcpzLB7       -A--T-----------------T-----TC-......----------A--TC-A---G-----C-CA-T------A-C--A-------------...............AGTACTCCTA--AGTTG-AGTCAAT-TATGGAC-T-----A-TT-GG-T--AT-GGCT-- 5662
CPZ.CM.05.SIVcpzMB66      -T--T------------G--------A---......----------A--TC-G---G--------A--------A-C--G--C---------G.....AGTAATTCATGCAGATATAT-CTTAGGC-T---T-G-TATAG-G-AGCA-----GCT--T---T-G-TC-- 5670
CPZ.CM.05.SIVcpzMT145     ------------------------AG-GAGA......---AA-CCT---TCGAGC--TCAGG-T...A--A-G-A-C-----CG---------..................AGTAAAAG-CTCACATGCAGCTAG--AT-G-G-T----A--TT-T-CA----T---A-G 5645
CPZ.GA.88.GAB1            -A----------------------A-CCACAAGA..-----C-CG----C-GCAGG---A-A-A-A------A----A--C-----------.....................AGTATA-GACT-TGT--GTCGGCTTAG-T-TC---T---T-GGCT--ATTGC-TGG 6170
CPZ.TZ.00.TAN1            -A-------------------GAA-AG--..........--CAA-...T--GC-GTAC-----AC--A-A------A-C--G--CG-C-------.............................----CCCATGA---AAAT-GTAG-GGG--GT--GT---CTAAT-TCA---GC 5732
CPZ.US.85.CPZUS           -A--T--------------------G-G-GACGC...A------C-A--GC-A---G----GC-GA-A------A-C-----CGC---------.................AGTGCCCT-TTGTGC-A-GCTTA-TTGGT-TGA-----GTTT-----C-TT-GG-A-- 6180
CPZ.CM.x.SIVcpzDP943      -A----------------------A--GGGACGC...A---ACCCGC--GC-ACA-GT---AC-CA-A------A-C--T--CG-C--------.........AGTACTTTACATGC-A-GC--AC-TGGGAA---ATAGGCTTG--T-CTTT-GG-A---A-GGCA-T 5699
CPZ.CM.x.SIVcpzMB897      -A--T-A-------------------AG.........------G-----AC-G---G----A--GA-------A-A--A--C----------.............AGTAATAAGT---A--T-ATAT-T-G--ATAT-CAT---CT---G-C--A---G-A----T 5671
N.CM.02.DJO0131           ----A--A--------------------T--A......----------A---C-A---G----A---A--------A-C--A--C--G-------..............................----A-A-CTGTAT-TATGC-GGAGTGG-G-TTCGC--CGTTGGG---G 5599
N.CM.04.04CM_1015_04      -------A--------------------T--A......-------A-A---C-A-C-G----A--GA-T------A-C--A--C--G-------..............................----A-AACTGTAT-TATGC-GT--T-G-G-TTCTC--CGCTCGGG--G 5598
N.CM.04.04CM_1131_03      -------A--------------------T--A......----------A---C-A-C-G----A---A-T------A-C--A--C--G-------..............................----A-AACTGTAT-TATGC-TGA-T-G-G-TTC----CGTTTGGG--G 5597
N.CM.95.YBF30             -------A--------------------T--A......----------A---C-----G----A---G--------A-CT-A--C--G-------..............................----A-ACCTGTAT-TATGC-GT--T-G-G-TTC----CGTT-GG--C- 5687
N.CM.97.YBF106            -------A--------------------T--A......----------C---C-A-A-G----A-AGA--------A-C--A--C--G-------..........................AGTA-AACTGTAT-TATGC-GTGGT-G-G-TTC----CGTT-GG---G 5684
O.BE.87.ANT70             -A----------------------AG--..........----C-...---GC-GCA-G-CA--CAGA---A--A-C--G--C---------..............................----ACGC---T---T-A--GGGACC-GC--------A--ATTA-TAGT 6147
O.CM.91.MVP5180           -A-----C----------------A--.........A---C-GC---GC-GCA-G-TA--CAGA-A--A---A-C--G--C--G-------..............................----ACGC-G-T---T-AAGAGAACC-GC----CT--A----T---AGT 6134
O.CM.96.96CMABB637        -A-----C-----------------C--.........----C-GC---GC-GCA-G-T--TC--A-A--A---A-CT-G--C--G-------..............................----ACGC---T-A-TTA--AGGA-T-GC--AGCT--A---TTG-TAGT 5589
O.CM.98.98CMA104          -A----------------------AC--..........---CC---AC-GCA-G-CA--C-GA-A--A---A-C--G--C--G-------..............................----ACGC---T---TTA--GGGACC-GT-----T--A--ATTG-TAGT 5588
O.CM.99.99CMU4122         -A----------------------AC--..........--TCCC---GC-GCAC--CA--C-GA-A--A---A-CT-G--C--G-------..............................----ACGC---T-----A--AGGATC-GC--AT-T--A--ATTACTAGT 5587
O.FR.92.VAU               -G-----C----------------A--.........A---C-GC---GC-GCA-G-TA--CAGA-A--A---A-C--G--C--G-------..............................----ACGC-G-T---T-A-CAGGGCC-GC---T-T--A---TCA-TA-T 5671
O.SN.99.SEMP1299          -A----------------------AC-G.........A---C-GCC---GC-GC-CG-CA--CAGA-A-----A-AT-G--C--G-------..............................----ACGC---T---T-A--GGGACC-GC-----T--A--A-TACTAGT 6158
O.US.99.99USTWLA          -A-----------------------C--.........A---C--C-A--GC-G-ACG-CA-ACAGA-A--A---A-C--G--C--G-------..............................----TCGCC--T---GTAC-AGGACT-GC--TTCT--C----T-ATAGT 5588
```

```
                                                                                                                                                                  Env start
B.FR.83.HXB2            GTAGCAATAATAATAGCAATAGTTGTGTGGTCCATAGTAATCATAGAATATAGG...AAAATATTAAGACAA.................AGAAAAATAGACAGGTTAATTGATAGACTA...ATAGAAAGAGCAGAAGACAGTGGC..AATGAGAGTGAAGGAG..... 6239
Vpu                     _V__A__I__I__I__A__I__V__V__W__S__I__V__I__I__E__Y__R__.__K__I__L__R__Q__.__.__.__.__.__.__R__K__I__D__R__L__I__D__R__L__.__I__E__R__A__E__D__S__G_##_N__E__S__E__G__.__.
Env                                                                                                                                                                    _M__R__V__K__E__.__.
A1.GE.99.99GEMZ011      ----TGT-TG-C-----------------A-T------GGT-------------...---T-GC---A----.................--------------------A-G---A--...-G---------------------..-----A--C--G--G-ATGCA 5431
A1.KE.00.KER2008        A----TT----C-----------------A-T------GGT-------------...-G-T-G---A----.................---------------------G---A--...-G---------------------..-----------T--G-ACACA 5456
A1.KE.00.KNH1144        -----GGC---CC---------A--------A-T------GGT-------------...---C-G---A----.................--G------------------G---A--...-G-----------G---------..-----------T--G-ACACA 5461
A1.KE.00.KSM4024        -----GC----CT---------A----A---A-T------GGA-------G--A-...-G-G-GC--GAC---.....AGAAAAATACAA---------------------G---A--...-G---------------------..-----------T--G-ATACA 5464
A1.KE.00.MSA4069        -----GC----C--------------A---A-T------GGT----------A-...-G-T-GC---A----.................--G------------------A---A--...-G---------------------..-----------T--G-ATACA 5464
A1.KE.00.NKU3005        -----GC----TC------------------A-T----GGGT----------A-...---T-GC---AG---.................---------------------G---A--...-G---------------------..-----------G--G-ATACA 5449
A1.RU.00.RU00051        -----GC-T--C-------------------A-T------GGT-------------...---T-GC---A----.................--G------------------G---A--...-G---------------------..-----A--C--G--G-ATGCA 5548
A1.RW.93.93RW_024       -----GC----CC------------------A-T------GGA----------A-...---G-GC-----------.................-------------A------G---A--...-G---------------------..-----------T--G-ATACA 5449
A1.SE.95.SE8891         --G-TGTC---C-------------------A--------GGT----------A-...---T-GC---AG---.................-----G-----T--------AGG---A--...-GT--------------------..-----------T--G-ACACA 5442
A1.SE.95.UGSE8131       ------C----CC------------------A-T------GGGC---------A-...---T--AA---G---.................--------------C------A---A--...-G---------------------..-----------T--G-ACACA 5598
A1.TZ.01.A173           --G--GC----C----T--------------A-T------GG-----------A-...-G-T-G----A----.................--G------------G-----G---A--...-G---------------------..-------------G-ATACA 5447
A1.UA.01.01UADN139      ----TGTCTG-CT--------T---------A-TGC----GGT-------------...---T-GCG--A----.................---------------------A---A--...-G---------------------..-----A--C--G--G-ATGCA 5446
A1.UG.92.92UG037        -----GC----C-------------------A-T------GGT----------A-...---T-GC---AG---.................--------------G--------A---A--...-G---------------------..-----------T--G-ATAGA 6272
A1.UG.99.99UGA07072     --G--GC----CC------------------A-T------GGT----------A-...-G-T--AG--AG---.................--G----------A----A-G---A--...-GT--------------------..-----------T--G-ACGCA 5453
A2.CD.97.97CDKTB48      ------TC---TT------------------A-TG-----T-T-----------A...--G--TAAG-AG---.................--G--------T----C---G---A--...-GT--------------------..--C--------T--G-ACACA 5593
A2.CY.94.94CY017_41     -----TT----CT--------------A---A-T------T------------A-...-----TAAG-AG---.................--G--------T------CA-A---A--...-GT--G-----------------..-----------T--G-ACACA 5610
A.SN.01.DDI579          -----GC----CT----------A------A--------GGG----------A-...---G-GC---AG---.................-A--G------------------A---A--...-G---------------------..-----------T--G-ATACA 5454
A.SN.01.DDJ369          ------G----CT----------A------A-T------GGG----------A-...---T-GAG--AG--G.................---------------------G-A-A--...-G---------------------..--C--------T--G-ATACA 5457
A.SN.96.DDJ360          -----GC----CT----------A------A--T-----GGG----------AA...---G-GC---AG---.................-A--G------------------A---A--...-G---------------------..-----------T--G-ATACA 5454
A.CD.02.02CD_KTB035     -----TT-C--C----T--------------A-T------T-T-------------...---C-GAG--AG---.................--------------C-G----A---A--...-G---------------------..-----A--C--G--G-ATACA 6234
A.CD.97.97CD_KCC2       -----TT-T--CT---C------G---ATT----C-T-T----------A-...---T-GAG--AG---.................--G--------TT---------A---A--...-G---------------------..-----A-----G--G-ATACA 6246
A.CD.97.97CD_KTB13      -----TT-C--CT------C-----------A-T------T-T-------------...--GT-GAG--A----.................-AG------------------A---A--...-G---------------------..-----A-----G--G-ATACA 5457

B.AR.04.04AR151516      A---T-GG-----------------------G-T-----GC-T-------------...-----------------.................-----------T------C--------A--...C---------------------..--------C-----G-ACCAG 5465
B.AU.87.MBC925          ------------------------------------T------T-------------...-----------------.................--G------------------A--...-GT--------------------..-----------------G-ATCAG 6254
B.BO.99.BOL0122         ------GC----C------------------A-T------T-------------...-------C---------.................-----------T------------A--...-G---------------------..-----------T--G-ATCAA 5463
B.BR.03.BREPM2012       ----T--C-----------------------A--T----GC-------------...-----------------.................--G-------------------A--...GC--------------------..--------C-----G-ACCAG 5677
B.CA.97.CANB3FULL       ------GC-----------------------A--------GG------------...-----------------.................---------------------A--GA--...--T--------------------..------------------CCCAG 5535
B.CN.05.05CNHB_hp3      ----T-GC-----------------------A--------T-AC-----------...-------------G.................-----------T------------A--...-G---------------------..--------C--C--G-ATCAG 6251
B.CO.01.PCM001          ---A--GC-----------------------A--------GG-T-----AT----...-GG-------------.................---------------------A---A--...GC-----------------C---..--------C-----G-ATCAG 5445
B.GB.83.CAM1            ----T-GC-----------------------A-T-----GT--------------...-------AG-------.................-A-C--G----------------A--...----------------------..--------C-----G-ATCAG 6244
B.GB.86.GB8             ------GG----G------------------AT-------T-------------...-----------A-A--..............AA.-A--------------------A--...---------------------A-..-----A--C-----G-ACCAG 6256
B.GE.03.03GEMZ010       ----T-GG-----------------------A-------GT-T-------------...-----------A----.................----G----------------A--...GC--------------------..-----------G--G-ATCAG 5448
B.IT.05.SG1             -------------------------------------GT-A-----------A...-GG---------------.................---G---A-----------------...----------------------..-----------------G-ATCAA 6046
B.JP.05.DR6538          ----T--C-------------A---------A--------T-------------...------C----------.................--G------------------A--...-C---------------------..-----A--C-----G-ACCAG 6253
B.KR.05.05CSR3          ------GC----T---------A-A------A-------TT-T------------...-------AG---G--.................-AG-------A--A----------GA--...-G---------------------..-----------T--G-ATCAG 6268
B.NL.00.671_00T36       ------GC-----------------------A-T-----GT--------------...-----------------.................--G--------T-------------A--...-G---------------------..-----A-------G-ATCAG 5808
B.RU.04.04RU128005      K-----GY-----------------------------T-A-----------AA...-MT--------------.................-----------A-----------A--...-G---------------------..--------C--G--G-ATCAG 5734
B.TH.00.00TH_C3198      ------GC--------------A--------A-T------C---------------...-------------G.................---------------------A--...-G---------------------..--------C--T--G-ATCAG 5457
B.UA.01.01UAKV167       ----T-GC-G------------------A---A--T----GT-A------------...-----------------.................---G-----------------A--...----------------------..-----------------G-ATCAG 5481
B.US.04.ES10_53         ----T-GC---------C-----G----------------------AT----...-----------------.................-AG---------G-----------A--...-G---------------------..-----------T---ATCTG 6252
B.US.99.PRB959_03       T-------------------------------------T-------------...-----------------.................--G--------T--------------A--...-G---------------------..-----------T--G-ATCAA 5463

C.AR.01.ARG4006         ------C----C---------A--------A------CTTAT-------------...---T-GG---------.................-A--G-------T----------A---A-T...-GG--------------------..--------C-----G-ATACA 5435
C.BR.04.04BR013         A-----C--G-CT---------A--------AT-----CTTAT-------------...---T-GG---------.................-A--G-------T----------A---A-T...-GG--------------------..-----------------G-ACACA 5729
C.BW.00.00BW07621       A-----C----C-------------------A------C-TAT-----G-------...---T-G---A----.................-AG-----------C-----A-G---A-T...--------G--------------..-----------T--G-ATCAT 5618
C.CN.98.YNRL9840        ------C----C-------------------A--------TAT-------------...-G-T-GG---------.................--------------------A-A---A-T...-GG--------------------..-----------G--G-ACACG 5430
C.ET.02.02ET_288        ------C-G--CC---------A--------A------C-TAT-----------A...---T-G----------.................C-G-------Y----------A---A-T...-GG--------------------..-----------G--G-ATACA 5467
C.GE.03.03GEMZ033       ------C----CT---------A-A-C----A------C-TAT-----------A...---T-G--------G.................-A--G-T-----T----------A---A-T...-AG---T----------------..----G---------G-ATACT 5445
C.IL.99.99ET7           ------C----C----------A--------AT-------TAT-------------...---T-G----------.................----G-------T-----G--A-C---A-T...-AG--------------------..-----------C--G-ATGCA 5440
C.IN.99.01IN565_10      A-----T----C-------------A------A--------TAT-------------...-G-T-GG---------.................-AG---------T---------A-A---A-T...-GG--------------------..-----------G--G-ATACT 5627
C.KE.00.KER2010         ------C----C----------A--------A-T------TAT-------------...---T-G----------.................--G---------T----------A---A-T...-GG--------------------..-----------G--G-ATACA 5436
C.MM.99.mIDU101_3       ------C----C----------A--------A--------TAT-------------...---T--G---A----.................------------T----------A---A-T...-GG--------------------..-----------G--G-ACACG 5597
C.MW.93.93MW_965        ------C----C----------A-A------A--------TAT-------------...---T-GG---------.................------------T-----G--A-A---A-T...-GG--------------------..-----------T--G-ATACG 5452
C.SN.90.90SE_364        ------C----C----------A--------AT-----C-TAT-------------...---T-G----------.................-AG--T----TT-----G--A-G---A-T...-GG--------------------..-----------C--G-ATGTT 5434
C.SO.89.89SM_145        ------C----C-------------------A--------TAT-------------...---T-G----------.................------------T-----G--A-A---A-T...-GG--------------------..-----------G--G-ATACT 5476
C.TZ.02.CO178           ------C----C-------------------A------C-TATT------------...---T-G----------.................----G-------CA-A-----A-A---A-T...-GG--------------------..-----------T--G-ATACT 5430
C.UY.01.TRA3011         ------TC---C----T-----A-A------AG------TTAT-----------A...---T-GG---------.................----G-------T-----G---G---A-T...-GG--------------------..-----------------G-ATACA 5436
C.YE.02.02YE511         ------C----C----------A--------A-T----C-TAT-------------...---T-GG--T-----.................------------T-----G--A-A---A-T...-GG--------------------..-----------C--G-ATACT 5439
C.ZA.04.04ZASK164B1     T-----C----CC------------------A-T------TATC------------...-------G--A----.................-----------TT-----------A---A-T...-GG--------------------..-----------G--G-ACACT 5670
C.ZA.05.05ZAPSK240B1    ------T----C----------A-A------A--------TAT-------------...---T--G---------.................-----------TT------A-A---A-T...-GG-------------T------..-----------------G-ATACA 5669
C.ZM.02.02ZM108         A----------CT---------A----A---A------CCTATC----------A...---T-GG---------.................--G--------T--------A-A---A-T...GA-----G-------T------..-----------T--G-ATCAT 6233

D.CD.83.ELI             ------------C------------------A--------T-------------A...-GG---AA--AG---.................--G-G-------T-T---C---------A--...-C---------------------..--------C--G--G-ATAGA 5789
D.CM.01.01CM_0009BBY    -----GC-----T------------------A------C-TA------------...-----G--AAG--G---.................-----------A-TCA------------A--...--------------A--------..-----------G--G-ATAGA 5442
D.KE.01.NKU3006         -----GC-----C-----G------------A-T--------A-----G-G--AA...-G-G--CA---G--G.................-AG------------------------A--...-G--------------T------..-----------C--G-ACAGA 5448
D.KR.04.04KBH8          ---A-GC---------------A--------A-------TT-------G-GCT-T...-G-T--AG--AG---.................-----G--------------A-----A--...-G--------------T------..-----------G--G-ATGAA 6199
D.TD.99.MN011           -C---GC---------------A--------A--------TA------------A...-------AAG--G---.................------------TATC----------A--...----------------------..-----------T---ATAAA 5467
D.TZ.01.A280            A----TTC----C-------------G-----A-T------TA-----G------...-G----AA--AG---.................-AG----------T--------------A--...-G--------------T------..-----------T---ACACA 5444
D.UG.99.99UGK09259      -----GC------------------------A-T------T-------G-G---A...-G-G--AG---G---.................------------T--------------A--...-GT-------------T------..-----A-----T--G-ATAAA 5448
D.YE.01.01YE386         -----GC-----C---------A--------A-T-----TT-------G-G---A...-G-T--AG---G---.................-----------TT--C-----------A--...GC--------------------..-----------T--G-ATAGA 5442
D.YE.02.02YE516         A----GGC--------------A--------A--------TA------------A...-------AGG--G---.................------------CA-C----------A--...-G---------------------..-----------G--G-ATGAA 5454
D.ZA.90.R1              ------T-----T------------------A-T------T-------------...-G----AG--AG---.................------------TAT-----------A--...-G---------------------..--------C--G--G-ATGGA 5592

F1.AR.02.ARE933         ------C------------------------A-T----C-TAT-----------AA...---C-GG---G---.................--G-G---A-T-----TA-A-A---A--...-G--------------T------..-----------G--G-ATGCA 5547
F1.BE.93.VI850          ------C---------C--------------A-T------TAT-----------A-...---C-GG---G---.................---------A-T-AA---TA-A-A---A--...-G---------------------..-----------G--G-ATGCA 5593
F1.BR.01.01BR125        ------C------------------------ATT------TAT-------------...---C-GG---G---.................-AG-----A-T-----TA---A---A--...-G---------------------..-----------G--G-ATGCA 5698
F1.ES.x.P1146           ------C-----T---------A--------A-T------TAT-------------...------G---G---.................-A--G---A-T-----TA---A---T--...-----------------T------..-----A-----T--G-ATGCA 5536
F1.FI.93.FIN9363        -----CC----------T-------------A--------T-T-----------A-...--GC--G---G---.................---------A-T----GTA-ATA---A--...-G---------------------..-----------G--G-ATGCA 5582
F2.CM.02.02CM_0016BBY   -----GC-----GC--------A--------A-T------TAT-----------A-...---CAG-----G---.................-A--G---A-T-----TA---A---A--...-G---------------------..-----------G--G-ATGCA 5436
F2.CM.95.MP257          ----TGC-----C---------A--------A-T------TAT-----------A-...------------G---.................-A--G---A-T-----TA---A---A--...-----G------------------A..-----------G---ATGCA 5455
F2.CM.97.CM53657        ---------C------------A--------A-T----GTAT-----------A-...---C-G----G---.................-A--G---A-T-A---TA-A-A---A--...-G---------------------..-----------G---ATGCA 5436

G.BE.96.DRCBL           ------TCC---GC---C-------------A-T------T-T-----------A...------AG--A---G.................-A--G-----A-A---C-------GA--...-G--------------------A..-------------G-ATACA 6210
G.CM.01.01CM_4049HAN    ------T-C---GC---C---A----A---AGT------T-T-----G-------...------AG--A---G.................-A-----------A---C---G---A--...-G--------------------A..-----------G--G-ATACA 5454
G.CU.99.Cu74            A-----GGC---GC---C-------------A-------TT-----C-------A...G----AG--A---G.................-A--------T--A-C---A---A--...-G--------------------A..-----------G--G-ATACA 5857
G.ES.99.X138            ------GGC---GC---C---A---------ATT-----CT-T---C-------A-...G----AG--A---G.................-AG-----C-A--A-C-------A--...-G--------------------A.-----A--C--G--G-ACACA 5694
G.GH.03.03GH175G        -------GC---GC---C---A---------AG-------T-T-------------...-G---AG--A---G.................-A---G----A-AA-----------A--...-G--------------------A..-------------G-ATACA 6293
G.KE.93.HH8793_12_1     ------T-C---GC---C---A---------A--------C-TG--C---------...G----AG--A---G.................------G---A--A-----------GA--...-G--------------------A..-------------G-ATAGA 5653
G.NG.01.01NGPL0669      ------GCC---GC---C-------------A-T-------T--------------...------AG--A---G.................-A--G-----T-----C--------A--...-G--------------------A.-----A--------G-ATACA 5458
G.PT.x.PT2695           ------GGC---GC---C---A---------AT------CT-T---C-------A-...G----AG--A---G.................-AC-----C-A-A-A-C--------A--...-G--------------------A..-----------G--G-ACACA 6195
G.SE.93.SE6165          ------T-C---GC---C-------------A--------T-T-------------...G----AG--A---G.................-A--G----GA-A---C--------A--...-G--------------------A..-----------C--G-ATACA 5650

H.BE.93.VI991           -----TT-T--C-----C-------------A------C-TAT-------------...---C-G.---AG--G.................--G------------------A---A--...-G---------------------..--C--------T---ACACA 5625
H.BE.93.VI997           ------T-T--C-----C-----A-------A--------TAT-------------...--GT-GG--AG---.................-AG----------------C-G--A-C...------G----------------..--C--------T-AG-..... 5556
H.CF.90.056             ---A--T-T--C----CG--A----------A--------TAT-----------AA...---T-GG---G---.................-AG----------C-------A---A--...GG--------------------..--C--------T---ACACA 5587

J.CD.97.J_97DC_KTB147   ------G-T--C------G--A---------A--------T-T-------------...-GG-G-----G---.................--G--------T--------CA-----A--...-G--------------------T..--C--A-----T---ACACA 5447
J.SE.93.SE7887          --------TT-CC-T-------G-A-----A-T----GTAT-----------A-...---C-G----------.................--G--------T-A-------------A--...-G--------------------T..--C--------T---ACACA 5566

K.CD.97.EQTB11C         -C---GC-----C---------A--------A------CCTATC----------A...---G-GG---------.................-A--G---A--T-----T--------A--...-G---------------------..--------C--G---ATACA 5442
K.CM.96.MP535           ------C-----T---------A----A---A--------TAT-------------...---T-GG--A----.................-A--G---A--T------A------A--...-G---------------------..-----------G--G-ATGCA 5446
```

```
                                                                                                                                                        Env start
B.FR.83.HXB2                 GTAGCAATAATAATAGCAATAGTTGTGTGGTCCATAGTAATCATAGAATATAGG...AAAATATTAAGACAA.................AGAAAAATAGACAGGTTAATTGATAGACTA...ATAGAAAGAGCAGAAGACAGTGGC..AATGAGAGTGAAGGAG..... 6239
Vpu                          _V__A__I__I__I__A__I__V__V__W__S__I__V__I__I__E__Y__R__.__K__I__L__R__Q__.__.__.__.__.__.__R__K__I__D__R__L__I__D__R__L__.__I__E__R__A__E__D__S__G_##_N__E__S__E__G__.__._
Env                                                                                                                                                     _M__R__V__K__E__.__.
01_AE.CF.90.90CF11697        -----GC----CT---------A------A-T-----GT-T----------A-...------C---G--.................-----------------G----A---A--...-G-------------------A..---G--------G-ACACA 6196
01_AE.CN.05.FJ051            -----GC----CT--------A-A------A-T------GC-------AT--A-...-G----C-T--GG-C.................----------------G--A-G--A--...-G-------------------A..----------------ACACA 6250
01_AE.CN.06.FJ054            -----GC----CT---------A------A-T------GCT------T--A-...------C---G--.................---------------------A-A---A--...-G-------------------A..----------------ACACA 6271
01_AE.HK.x.HK001             -----GC----CT---------A------A-T------GGTT-----GT---...------C---G--.................-----------T--------A-G-A-A--...-G-------A-----------A..----------------ACACA 5617
01_AE.TH.01.01TH_R2184       -----GC----CT---------A------A-T------GGT-------T--A-...------C---G--.................---------------------A-G---A--...-G-------------------A..----------------ACACA 5447
01_AE.TH.02.OUR769I          -----GT----TT---------A------A-T------G-T------CT--A-...------C---G--.................---------------------A-G---A--...CA------------------A..--------------G-ACACA 5446
01_AE.TH.90.CM240            -----GC----CT---------A-----AA-T------GCT-----GT--A-...------C---G--.................-----------------G--A-G---A--...-G-------------------A..----------------ACACA 5815
01_AE.US.00.00US_MSC1164     -----GC----CT---------A------A-TT-----GCT-------T---...------C---G--.................-----------------G--A-G---A--...-G-------A-----------A..----------------ACACA 5452
02_AG.CM.02.02CM_4082STN     ------GGC---GC---C------------A-------GC-T-----------A...------AGG-A---G.................-AG-------T-----C--------A--...-G--------------T------..-----------T--G-ATACA 5452
02_AG.EC.x.ECU41             ------T-C---GC---CT-----------AG-------C-T--T---------A...------AGG-A---G.................-AG-G--------C--C-------A--...-G--------------T------..-----------T--G-ATACA 5348
02_AG.GH.03.GHNJ196          ------T-C---GC---C------------A------AGTAT-----------A...-G---AGG-A----.................-A-C------T-----C--------A--...-G--------------T------..-----------T--G-ACACA 6282
02_AG.NG.01.PL0710           --------C---GC---C----CG------A-T------TAT-----------A...------AGG-A---G.................-AG-GG------AA------------A--...-G--------------T------..---G-----T---ATACA 5437
02_AG.NG.x.IBNG              ------T-C---GC---C------------AG-------TAT-------------...------AGG-A---G.................-AG---------------C-------A--...-G--------------T------..-----------T--G-ATACA 5777
02_AG.SE.94.SE7812           ------T-C---GC---C------------A-------GTAT------------...------AGG-A---G.................-AG---------------C-------A--...-G--------------T------..-----------T--G-ATACA 5625
02_AG.SN.98.MP1211           ------T-C---T----C------------A--------T-T------------...------AG--A----................G------------A---C------GA--...-G--------------T------..-----------T--G-ATACT 5455
03_AB.RU.97.KAL153_2         ----T-GG------------------G----------T-------------...---------------.................--------------------------A--...-G---------------------..----------------ATCAG 5469
04_cpx.CY.94.CY032           -----GC----T----T------A------A-TT-----T-T----------A-...---T-GAG---G---.................--G-G---------C--GTACA-----A--...-G---------------------..-----------T--G-ATGCA 5617
05_DF.BE.x.VI1310            ------C---------------------ATT------TAT-----------A-...---T-GG----G---.................---------A-T------TA-A-A---A--...-G---------------------..-----------G--G-ATGCA 5635
06_cpx.AU.96.BFP90           ------T-CT--GC---C------------A-T------T-T---C----------...G-----AG--A---G.................-A---------A-A---C-------A--...-G-------------------A..--------------G-ATACA 6283
08_BC.CN.97.97CNGX_6F        ------C----C------------------A--------TAT-------------...-G-T-GG---A----.................--G---------T--------A-A---A-T...-GG--------------------..-----------G--G-ACACG 5420
09_cpx.GH.96.96GH2911        ------T-C---GC---C---A---C----A--------T-T------GT----...--G---AG--A----.................-AG-G------T----------------A--...-G---T------------------..----------------ATACA 5458
10_CD.TZ.96.96TZ_BF061       ------C----C------------------AT--G----TAT---C----------...---T--G--GA--CT.................-CG---------T-------CA-A---A-T...-GG--------------C------..C----C------C--G-ATACA 5644
11_cpx.GR.x.GR17             AGCATTTA.T-----------A--------A-T----C-GGT-----------A...-GGC-G---------C.................--G---TAGATAGAT-AT...T-----A--...-G---.-------------------..----------------ACACA 5553
12_BF.AR.99.ARMA159          ------GC----T------------C-------------T-A-------------...---------------.................--G-----------------------A--...-G---------------------..-----------G--G-ATGCA 6248
13_cpx.CM.96.1849            ------T-T--TT-----G-----------A------C-TAT-----------A...--GC-GG--CAT--G.................-AG--------A-A----T-G------A--...-GG--G----------------A..---------C--G--G-ATACA 5657
14_BG.DE.01.9196_01          -------GC---GC---C---A--------AT------CT-T---C------A-...G-----AG--A---G.................-AG------C-A--AA--C--------A--...-G---T----------------A..------A--C------G-ATACA 5770
15_01B.TH.99.99TH_MU2079     -----------CT----------A------A-T------GCT------AT----...--T-----G-AG---.................-----------------G--A-G---A--...-G-------A-----------A..----------------ACACA 5643
16_A2D.KR.97.97KR004         ------T-C--CT-----G-------A---A-T------T------------A-...------TAAG-AG---.................-AC---------T--------A-G---A--...-GT---------------T------..-----------G--G-ATACA 5613
17_BF.PE.02.PE02_PCR0155     ------------C------------------T-----G-----------------...---------------.................-----G---------R---------A-A-A--...-G---------------------..-----------G--G-ATGCA 5433
18_cpx.CU.99.CU76            -----TT-T--T------G--TG-------A--------TAT-----------A...--G--T----A----.................-AG--------TG-----T----G---A--...-G---C----------------..T---------G--G-ACACA 5546
19_cpx.CU.99.CU7             ------T-C---GC---C------------A-------GTAT-----------A...------AG--A-G-G.................-----------A-A---C--------A--...-G---------------------..-----------T--G-ACACG 5469
20_BG.CU.99.Cu103            ------GGC---GC---C---A--------G-------TT-----CT-----A-...G---TAG--A---G.................-A---------T---A-C----A---A--...-G-------------------A..--------------G-ATACA 5720
21_A2D.KE.91.KNH1254         ------GC----------------------A--------G-------------A...-GG---AA---G---.................--G---------GAT-------------A--...-G---------------------..-----------G----ATGAA 5448
22_01A1.CM.01.01CM_0001BBY   -----GC----CT-------------A---A-T------GCT------AT--A-...---G--AG---G---.................-AG--------T-----C----A---A-C...-G-------------------A..-----------G--G-ACACA 5456
23_BG.CU.03.CB118            ------GGC---GC---C---A--------G-------TT-T--TCT-----A-...C-----AG--A---G.................-A---------T---A-C----A---A--...-G-------------------A..--------------G-ATACA 5696
24_BG.CU.03.CB378            ------GGC---GC---C---A--------G-----A-TT-----C------A-...G-----AG--A---G.................------------T---A-C----A---A--...-G-------------------A..--------------G-ATACA 5684
24_BG.CU.03.CB471            ------GGC---GC---C---A--------G-------TT-----C------A-...G-----AG--A---G.................------------T---A-C--C-A---A--...-G-A-----------------A..-----------G--G-ATACC 5693
25_cpx.CM.01.101BA           ------T-C---GC----G-----------A-T------T-T------AT--A-...-------AG--A---G.................-A---G-----ACA----C----C---A--...-GGAC---G--------T------..-----------G----ATACA 5461
27_cpx.CD.97.97CDKTB49       ------T-T--C----T-G-----------A--------TAT----------AA...--GC-G---------.................------------T---A-------G---A--...-G---C-------------------..--------------G-ACACA 5460
28_BF.BR.99.BREPM12609       ----T-GC--------------A-------A-------GT-T-------------...---------------.................--G------------------------A--...-G---------------------..-----------G--G-ATGCA 5609
29_BF.BR.02.BREPM119         ----T-GCC---------------------A-------GT-T-------T-----...---------------.................----G------T---------------A--...-G---C-------------------..--------C--G--G-ACCAG 5417
31_BC.BR.02.110PA            A-----C----C----------A-------A-------TTAT-------------...----T-GG---A---.................------------TT-----G--A-A---A-T...-GG--------------------..--------------G-ATACA 5698
32_06A1.EE.01.EE0369         A----TTC-T--GC---C------------A-T------T-A---C----------...G-----AG--A---G.................------------A-AA---C-------A--...-G-------------------A..-----------G--G-ACACA 5888
33_01B.MY.05.05MYKL007_1     -C---GC----CTC---------A------A-T------GCT--------TGCA-...------C---AG--.................---------------C--G--A-G---A--...-G-------------------A..----------------ACACG 5632
34_01B.TH.99.OUR2478P        -----GC----CT----------A------ATT------GCT-------------...------C-----G--.................------------T------G--A-G---A--...CA------------------A..----------------ACACA 5449
35_AD.AF.05.05AF095          --G--GC----CC-----------------A-T------GGT------CT--A-...-G-T-GC-G--G---.................------------T--AA-----A-G---A--...-GT--------------T-----T..--C--------T--G-ACACA 5446
36_cpx.CM.00.00CMNYU830      ------T-C---GC---C----G--G----A--------C-T----------AA...------A---G---.................-A---------------C-----A---A--...-A---------A----T------A..---------C--G--GAACACA 5462
37_cpx.CM.00.00CMNYU926      ------T-C---GC---C------------A-T------C-T-----------A-...------AG--A---G.................-------------------------A----A--...-G---------------------..-----------T--G-ATACA 5444
38_BF1.UY.05.UY05_4752       ------T---------------------ATT----C-TAT-A---------A-...---C--G----G---.................----G----A-T--C---TA-C-A---A--...-GT---------------------..-----------G--G-ATGCA 5670
39_BF.BR.03.03BRRJ103        ------T-C-------------------A-T------TAT-------------...--GC-GG----G---.................------------A-T-AT---TA-A-A---A--...-G---------------------..-----------G--G-ATGCA 5743
40_BF.BR.04.04BRRJ115        ----T-GC----------------------A-TC----GC-----------C---...---------------.................------------------------A---A--...G---------------------..--------------G-ATCAA 5795
42_BF.LU.03.luBF_05_03       ----Y-GC------------------A-----------C----------------...----------A----.................------------T----------C----A--...-GC--------------------..-----------C--G-ATCAT 5781
43_02G.SA.03.J11223          ------T-CY--GC---C------------A-A------T-T-------------...------AG-----G.................-A--R------T--A---C--R-----A--...-G-------------------A..----------------ATACA 5778
U.CD.83.83CD003              ------T----CT-----------------A--------TA-----------A-...-G-C-G---------.................------------TT--------------A--...-GG---------------------..----------------ATACA 5735
U.CD.90.90CD121E12           ------T--GGCT-----------------A--------TA-C---------A-...---C-G---------.................CA--G------TT--------------A--...-GG---------------------..--------C-------ATACA 5436
U.GR.99.GR303                ------T-T---GC---C------------GA-------C------T------A...---TA-GA--AC---.................----G------T---------T----A--...-G---------------------..----------T-----G-ATACA 5525
U.NL.95.U_NL_95_H10986_D1    --G---T-T--CT---T-G--A--------A--------TATTGT-------A-...--GC-GAA---G---.................GA--G------T-A-A-----A-----A--...-G---------------------..-----------G--G-ATGCA 6135
CPZ.CD.90.ANT                ---CT-TGG--T--CTGT---CC-A-ACTC-AT-AGC-TTAT-A-AT-----A-......CAGCAGCA-AT-GATAATAAGAGAAATCAA----TC-----AGT-----G-AGA-----T...-GTAT-GATAGT-C-ATAGAA-AA..G-----GAA-CC-ATACATAT 5697
CPZ.CM.01.SIVcpzCAM13        ............T----C--T-CCT-AATT-TAC-TAAC--TT-TAT--GG---AATTT-TGGAG-TT-TGTAAACAAAATAAGCTTGAAGT-G----T--G-ACC-TGC-TTA-------...-C---------------T-----A..-----A-----T----GAGAG 5825
CPZ.CM.05.SIVcpzEK505        -AGA-T...C------T------A--T---G-A------TAT-G-AT---C--A...G--G--AA-GT-G--.................GA-------TCTCA-C---GAC-G---A--...-G---C-----------T------..-----A-----C----ATGCA 5814
CPZ.CM.05.SIVcpzLB7          --TAT-GA-C----CAT-G--A-A--AGTA-GGT--.......-A--C----CA-......C-T-AT-A-G-G.....AACATAAGACAA-AGGC----A-T-AA--------G--GA--...-G--------GG-------------..--------------G-ACATG 5809
CPZ.CM.05.SIVcpzMB66         A-TAT-GA-T--G-TAT-G-CA-A--AATA-GGG--.......-A--TT----AA......C-C-GT-A-G-G.....GATAGGAGGCAA-AG--------T--A-------CA---A--...-G---------------------..-----A--C--T----ACACA 5817
CPZ.CM.05.SIVcpzMT145        C-T-T-GCT---T-T--TTGGA-A-CAGCT-AT..........-A---G----AA...--GT--.................CAACAAGTA--GCG------A--AC-TCAG--C---A-T...-G-AGT--------G----------..-----A-----T--G-ATGAA 5780
CPZ.GA.88.GAB1               AACATTTGT---TGG-G-TACA--A-TAAA-GGGG-TAT...-G-AGG----AA......-G-CAT--G...............CTTGAG-C-G----T--G--AC-T-ACCTA-T-T-G...-G--------------T-----A..-----A---A-T-----...AA 6308
CPZ.TZ.00.TAN1               A-TCTTTGT---T--CTG--TT-AA-AG--GGAGGCT-GC-A----GTATAG-T...-T--G-AG-GAGTT-.........GAAAGGGAA--GC--CATC-A--AG--T-A--A--G---GCT-G-AG-TT-AGCAT---------A..GTA--AGAA--T-A--...AA 5885
CPZ.US.85.CPZUS              -A--GT--TT--G---TG---A-CTG-G-ACTAG-T-CT-GAT-GTGGAGACAA...-T--A-A---A-G--.........AATACTCAACA-G-G--TC-A-ATC-TT-A--A---A-C...-G-AT-----A------T-----A..-----A-----T-----...AA 6330
CPZ.CM.x.SIVcpzDP943         A-T--T-CTG--G-CTGGGGGA-A-CT-TTATA-CGTGG-GA-G-AG-A-A-TAGAAG--GACAG-GC----.................-----CATTGAT---C-A--A---A--...-G--C----CAG--C--T-----A..-------------T---AGGAT 5843
CPZ.CM.x.SIVcpzMB897         A---A-T--G------------A--A----TA-A--CTTATGA-TGT----A-...GC-C-TAA-------.................GA--GG-G---TCAA--G-----C---A--...-G---------------T-----A..-----A-----T--G-ACACA 5815
N.CM.02.DJO0131              -C-ATT...---------G--A-AA-C---GTTT--C--TAT-A--------A-...----C--AA-TTG--G.................GA---------ACAAA---GAC-A---A--...-G---T---A-------T------..-G-----A---T--G-ATGCA 5740
N.CM.04.04CM_1015_04         -C-ATT...---------G-CA-AA-C---GTAT--C--TAT-A--------A-...G-----AA-CTG--G.................GA---------ACAAA---AAC-G---G--...-G---C-----------T--------------------T--G-ATGCA 5739
N.CM.04.04CM_1131_03         -C-ATT...---------G--A-AA-C---GTGT--C--TAT-A--------A-...G-----AA-CTG---.................GA---------A-A-A---GAC-G---A--...-G---T-----GA---G------..--A--------T--G-ATGCA 5738
N.CM.95.YBF30                -C--TT...-GC------G--A-A--C---G-AT--C--TAT-G--------A-...------AA-TTG--G.................GA-------A-ACACA---GAC-G---A--...-G--------A------T------..-----A-----T--G-ATGCA 5828
N.CM.97.YBF106               -C-ATT...---------GC-A-AA-C---GTAT-GC--TAT-A--------A-...------AA-TTG--G.................GA---------ACAAA---GAC-G-A-A--...-G---T---A----G-----G---..--G--------T--G-ATGCA 5825
O.BE.87.ANT70                -CTTTGC-GT-T---AATG--A--T-A---GGGT-TA-TC-T-G-A-----TTA...G--CA-AAGGA----......GACAGAAAGGAG---G-----CTTGA-AGGT-AAGA---A--...-G-----TTAGG--T--T----A-..T----A--CA-T----AAGAG 6303
O.CM.91.MVP5180              -CTTTGTGTC-T---AATG--C--A-A----TGT-TAACC-T-G-ATT---TTA...GTGCA-AG--A----......GATAGAAGGGAGCAGG-----CTTGAAAG-T-AAGG---A--...-AG----TCAGG--T-------A-..T----A---A-T-A--AAGAA 6290
O.CM.96.96CMABB637           --TTTGC-TC-TGC---T-----GA-A---ATGT-TA-TC-T-AGA-----TTA...G--CAGAAGGA----......GATAGAAGGGAA---G-GC-GCTT-AAAG---A--A--------...-R----GYCAGG--T--T----A-..T----A---A-T----ATGAA 5745
O.CM.98.98CMA104             -CTTTGC-TC-T---AATG--C--T-A---AG-T-TA-TC-T-G-A-----TTA...G--CA-AAG-A----......GACAGAAGGGAA---G-----CTTGAAAGGT-AAGA-A-A--...-G-----TTAAG--T-------A-..T----A--CA-T-A--AAGAA 5744
O.CM.99.99CMU4122            -CTTTGC-GC-T---AATG--A--C-A---CTGT-T--TC-T-AGC---G-TTA...G--CA-AAG-A----......ACTAAAAGAGAA---G-----ATT--AAGGT-AAGA---A--...-G-----TT-AG--T--T----A-..T--------CA-T----AAGAG 5743
O.FR.92.VAU                  -CTTTGC-TC-T---AATG--C--C-A---ATGT-TAATC-T-G--C-----TTA...G--CA-AAG-A----......GATAGAACAGCA---G-----CTTGAAAG---AAGA---A--...-G----GTT-AG--T-----C-AT..T----A--CA-T--G-AAGAG 5827
O.SN.99.SEMP1299             -CTTTGC-GC-T-C-AATG-----C-A---A-AT-TA-TC-T-G-C-----TTA...---CA-AAG-A----......GATAAAAGGGAA---G-----CTTGAAAGGT-AAGA---A--...-G-C---TT-AG--T-------A-..T----A-----T---ACAGAG 6314
O.US.99.99USTWLA             -CTTTGT-TC-T-C-A-CGCGC--A-A---ATGT-T--GC-T-A-CTT-GCTTA...G--CA-AAG-----......GGAAAAAGGGAA---G-----CTTGATAG-T-AAGA---A--...-GR--R-TCAGG--T-------AT..T----A--CA-T----RGGAA 5744
```


HIV-1/SIVcpz
Complete Genomes

HIV-1/SIVcpz Complete Genomes

```
                                                              Vpu end
                                                  Env signal peptide end Env gp120 start
B.FR.83.HXB2            ....AAATATCAG...CACTTGTGGAGATGGGGGTGG...........AGATGG..............GGCACCATGCTCCTTGGGATGTTGATGATC...TGTAGTGCTACAGAAAAA......TTGTGGGTCACAGTCTATTATGGGGTACCTGTGTGGAAGGAAG 6366
Vpu                     _.__E__I__S__.__A__L__V__E__M__G__V__.__.__.__.__E__M__.__.__.__.__.__G__H__H__A__P__W__D__V__D__D__.__L__*
Env                     __.__K__Y__Q__.__H__L__W__R__W__G__W__.__.__.__.__R__W__.__.__.__.__.__.__G__T__M__L__L__G__M__L__M__I__.__C__S__A__T__E__K__.__.__L__W__V__T__V__Y__Y__G__V__P__V__W__K__E__
A1.GE.99.99GEMZ011      GAGG--T-----A..------A-----G---.....................................--A--T---T--T--T-----A-A--A--G..----AG---G-------C......-----------G--A--C------------------GA--T- 5550
A1.KE.00.KER2008        GAGG--T-GG---...-C---A-T-------.....................................--A--T---A--T-G-----C-A-------..----------G----C--C......--------T--T----------------------A--T- 5575
A1.KE.00.KNH1144        GATG--T-G---...-----A-T-------......................................--A--T---A--T-G--AT--A-A--A---..----A---GT-A-C-GC...AAC--------T--T----------------------A--T- 5583
A1.KE.00.KSM4024        GAGG---G---...-----A-T-------.......................................--A-TT---A-AT-G----AA-A--AT--..-----------A---C--C......--------T--T--T-------------------A---- 5583
A1.KE.00.MSA4069        GAGG--T-----A..-G--CA-T---C---......................................--G-TT---G--T-G-----A-A--A---..----------G-------C......--------T--T----------------------A--T- 5583
A1.KE.00.NKU3005        GACG--T-GG-G-..-------T-------......................................--A--T---A--T-G--A---A-A--A---..-------A--G-------G......--------T--TA---C-----------------A--T- 5568
A1.RU.00.RU00051        GAGG--G-----A..-----A---G-G---......................................--A--G-----TT--T-----A-A--A--G..----AA-------G---C......C---------G--A--C-----------C-------GA--T- 5667
A1.RW.93.93RW_024       GAGG--T------..-------TAGA----......................................--A--T---A--T-G-----A-A--AT--..----------GT------C......--------T--T----C-----------G--------A--T- 5568
A1.SE.95.SE8891         GATG--T-GG--A...A-------------......................................--A--T---A--T-G----AA-A--A--T..----------G-------C......--------T--T-----------A-----------GA--C- 5561
A1.SE.95.UGSE8131       GAGG--T-G---A..-----A-T--AC---......................................--G-TT---A--T-G-----A-A--A---..-------A--G-------C......--------T--T--T--C-----------G--------A--C- 5717
A1.TZ.01.A173           GAGG--C-C----..-----C-T-------......................................--A--T---A--T-G--AT--A-A--A---..----------G---C--C......--A-----T-------C-----------A--------A--T- 5566
A1.UA.01.01UADN139      GAGG--T-----A..-----A---G-G---......................................--A--T---T----T-----A-A--A--G..----AG---------C......----------------A--C-------------------GA--T- 5565
A1.UG.92.92UG037        GAGG--T----CT..TG--G-----C----......................................--G-TT---A--T-G-----A-A--A--T..----A-A--G-------C......--------T--T----C-----------A-A--------T- 6391
A1.UG.99.99UGA07072     GAGG--T-GG--A...A-------------......................................--T--T---A--T-G-----A-A--A--T..-------A--G-------C......--------T--T----C-------------------A--C- 5572
A2.CD.97.97CDKTB48      GAGG--T-G---A...A-A-G----GA---......................................--A-T-T--G-TT-------A-A-----G..----AG---G----T.........--------------A--C----A-------------GA--T- 5709
A2.CY.94.94CY017_41     GAGG--C------..-----------G---......................................----T-T--A-TT-G-------A--A--G..----AA--------T.........--------------A--C----A--------------A--T- 5726
A.SN.01.DDI579          GAGG--T----TA..-C--------G---.......................................--G-TG--ATC---CT-T---A-A--A--T..----AA---------C......-----------G----C----------G--------A--C- 5573
A.SN.01.DDJ369          GAGG--C-G-ACA..-C--G-----G---.......................................--G-TA--AT------AT--A-A--A--T..----AG---------G......------------------C----------G--------A--C- 5576
A.SN.96.DDJ360          GAGG--T----TA..-C--------G---.......................................--G-TG--AT------A--A---A--T..----AG---------C......-----------G----C----------G--------A--C- 5573
A.CD.02.02CD_KTB035     GAGG--C-----A..-----A----A----......................................--G-T-T--A-GT-G-----A-----A--T..G-------G--C--G-G......--------T--G----C--------------------GA--T- 6353
A.CD.97.97CD_KCC2       GAGG--C-C---A..-----A---------......................................--A---T--A-TT--T---CAA-A--A--T..----------G--------......-----------G----C----------C--------A--T- 6365
A.CD.97.97CD_KTB13      GAGG--C-----A..-----A-----G---......................................--A---T--A-TT-------TA-A--A--T..----------G-------C......-----------G----C-------------------A--T- 5576
B.AR.04.04AR151516      GAAG--T-GG---..----G----------......................................--------------------A-------..----------ACC--.........C-------------------------------A---- 5581
B.AU.87.MBC925          GAAG--T-------..---------------.....................................-----------------T-A--A-------..----------G----C---......------------------------------A---- 6373
B.BO.99.BOL0122         GATG--T-G----...T--------G-----A-C......AC--TG---------------A--------..---------------C--......--------------------------A---- 5585
B.BR.03.BREPM2012       GAAG--T------..-----------------....................................----TG-----------A-----A-----GATC-----------A--G--...CAA-------------------------A---- 5802
B.CA.97.CANB3FULL       AAAG--T------..-----------------....................................--------------------------------..----------G----C--......--------------------------GA---- 5654
B.CN.05.05CNHB_hp3      GAAG--T---........--GCAT---------...................................------T-AT----C--A---C-------..----------G--------......--------------------A------A---A 6367
B.CO.01.PCM001          GAAG--T-GG---..-----------A----.....................................-----G-----------------A-----A..-----------------C--......--------------------A------A---- 5564
B.GB.83.CAM1            GAAG--T-G----...-G---A---------.....................................-----A-----------------A-------..----------G----C---......--------------------A------A---- 6363
B.GB.86.GB8             GAAG--T------..-----------A----.....................................----T---------------------------..-----------------------......------------------------A---- 6375
B.GE.03.03GEMZ010       GAGG--C--------...-C---A---------...................................----T-G-------------------------..-----------------------......--A-----------------------A--T- 5567
B.IT.05.SG1             GAAG---------..----------------T-----...............................----TG-----------T-A--A-------..-----------A----G-C......--------A-----------------T------------ 6171
B.JP.05.DR6538          GAAG--T------..-----A---------......................................--------------------A-------..----------G----GC-G......--------T-------------------------A---- 6372
B.KR.05.05CSR3          GAGG--T----A...A----------G---......................................----T-T------------A----------..-----------A---GC-G......------------------------------A---- 6387
B.NL.00.671_00T36       GAAG--T--------...-T-----------.....................................----T-------------C-------------..------------------T......--------------------------A---- 5927
B.RU.04.04RU128005      GAAG--T--------..GG---A-T-------....................................------T-----------A-------T..----------G---G---T......--------------------------A---- 5853
B.TH.00.00TH_C3198      GAGG--T------...--T--A---------.....................................------------------------------..----------G-------C......--------------------A------A---- 5576
B.UA.01.01UAKV167       GAAG--T-G--GA..-----------T-----....................................------T---------A--A---------..----A---------C---......--A-----------------------A---- 5600
B.US.04.ES10_53         GAGG--T-G----..-------------G----CA-C.........TGG------.............----T--------------------------..-----------G--A-C--T......-------------A-----------------A---- 6386
B.US.99.PRB959_03       GAGG--T-------..-----------A-GA-....................................----T-T-------------------------..-----------G-----C-G......------------------------------A---- 5582
C.AR.01.ARG4006         GAGG--T-GGA-A..--A-G-----T----......................................----T-T-AGG-T--T-C-----M-----T..----A--TG-GG-GG--C......--A----------A---------------------A---- 5554
C.BR.04.04BR013         GAGG--T-GGA-A..--A-G-----T----......................................--A-T-T-AGG-T--T-------CC----T..----A--TGGTG-GG--C......--------------A---------------------A---- 5848
C.BW.00.00BW07621       GAGG--T-G---A..--A-G-----T----......................................-T--T-T-AGG-T--T-----C-A---G-T..----A--TA-T--GG--C......-------------------------------GA---- 5737
C.CN.98.YNRL9840        GAGG--TC------..--A-G-----T----.....................................--A-T-T-AGG-T--T-------A-----T..----A--TGGG--G---C......C----------------------------A---- 5549
C.ET.02.02ET_288        GAGG--T-G---A..--A-G-----T----......................................--ACT-T-AGG-T--T-------C-----T..----A--TGGTG-GG--T......--------------------------A-------A--C- 5586
C.GE.03.03GEMZ033       GAGG--T-G---A...-G--G-----T----.....................................--A-T-T-AGG-T--T-----C-A---G-T..-----CCAGGA-A-C.........--A-----------C------------A---GA---- 5561
C.IL.99.99ET7           GAGG--T-G---A...CA-G-----T----......................................----T-T-AGG-T--T-----C-A-----T..-A--A--TGGG--G---C......--------------T-----------G----------A---- 5559
C.IN.99.01IN565_10      GAGG--T-----A..--A-G-----T----......................................--A-T-T-AGG-T--T-------A-----T..----A--TGGT--G---C......--------T-----------------------------A---- 5746
C.KE.00.KER2010         GAGG--T-GG--A..--A-G-----T----......................................----T-T-AGG-TC-T-------A----CT..--------TGGA--GG--C......-----------------C------------------A---- 5555
C.MM.99.mIDU101_3       GAGG--T-----A..--A-G-----T----......................................---GT-T-AGG-T--T-------A----G-..----A--TGGG--G---C......-------------------------------GA---- 5716
C.MW.93.93MW_965        GAAG--T-GG--A..--A-------T----......................................----T-T-AGG-T-------C-A-----T..---GA-ATG-AG-GC.........------------A---C------------------A---- 5568
C.SN.90.90SE_364        GAGG--T-G---A...CA-G-----T----......................................----T-T-AGG-T--T-----C-AT----T..-A--A--TGGGG-GG--C......--------------A---------------------A---- 5553
C.SO.89.89SM_145        GAGG--T-GG--A..--A-G-----T----......................................----T-T-AGG-T--T-------A-----T..--------GGGG--GG--C......--------------------------G------GA---- 5595
C.TZ.02.CO178           GAGG--T-GG--A..--A-G-----T----......................................--A-T-T-AGG-T--T-------A-----T..----A--TGGTG-GG--C......-------------------------------GA---- 5549
C.UY.01.TRA3011         GAGG--T-GGA-A..--A-G-----T----......................................----T-T--GG-T--T--C----A---GGT..-A--A--TA-GG-GG--C......--------------A-----------------C-A---- 5555
C.YE.02.02YE511         GAGG--T-----A..--A-G-----T----......................................----T-T-AGG-T--T----A--A-----T..----A--GGGG--GG--C......--------------T---------------------A---- 5558
C.ZA.04.04ZASK164B1     GAGG--T-GG---..--A-G-----T----......................................----G-T-AGG-T--T-------A---G-G..--------TGGTG-G---C......-------------------------------GA---- 5789
C.ZA.05.05ZAPSK240B1    GAGG--T-----A..--A-G-----T----.....................................A---T-T-AGG-T--T--T--A-A--A--T..A-CG-G-TGGTG-GG--C......-----------------C--------G----A-----A---- 5788
C.ZM.02.02ZM108         GAGG--T-G---A..--A-G-----C----......................................----T-T-AGG-T--T-----C-A------..-A-----TA-TG-GG--T......-----------------------A------------A---- 6352
D.CD.83.ELI             GAGA--T-G---A...A---G-----A----.....................................----T--------------A-------C-..----------G----C--T......C------------T-----------G-----A----------- 5908
D.CM.01.01CM_0009BBY    GAGG--T-----A..-----A----A----......................................-----------------------C----AACA-A------AG--AGT--T......--------------A-----------G-----A-----A---- 5564
D.KE.01.NKU3006         GAGG--T------..-----------A----.....................................----T-----------------------A..----A--T-----G---G......--------------T---------------------A---- 5567
D.KR.04.04KBH8          GAGG--T-----A..-----A----A----......................................-----------------A---G-----T..----------G---G---C......--------------T---------------------A---- 6318
D.TD.99.MN011           GAGG--T-----A..-----A----C----......................................------C----------C----T---GACC-A------AG----GG-T......-C------------A--C---------G-----------A---- 5589
D.TZ.01.A280            GAGG--T-----A..-----------------....................................----T-----------A---G----CA...-A----T-G-----C-G......--------------A-T----------------T---GA---- 5563
D.UG.99.99UGK09259      GAAG--T-----A.....CA----A----.......................................----T-----------A---------CA..----A---G-----C-G......--------------T---------------------A---- 5567
D.YE.01.01YE386         GAGG--T------..-----------------....................................----TG----------A---------CA..--------T-----G-G-G......C--------------T---------------------A---- 5561
D.YE.02.02YE516         GAGG--T-----A..-----A----A----......................................-----GC---------A------C----GACC-A------AG----T--G......--------------G---------------A-----A---- 5576
D.ZA.90.R1              GAGG--T-G---A...A---------A----.....................................--T-TG----------A-------------..----------G-------T......--------------T-----------G-----A----------- 5711
F1.AR.02.ARE933         GAGG--T-GG---..------G---A----......................................---C-TT-AT----A-----A-----A---..----A-----------C......----T---T----------------------------GA---- 5666
F1.BE.93.VI850          GAGG--T-GG---..------G---A----......................................---CTTT-AT----G-----A--A--A---..----A----G----C--C......--------------------------------------A---- 5712
F1.BR.01.01BR125        GAGG--T-GG---..------G---A----......................................---CT-T-AT----G-----A-----A---..----AC---GA---C.........------------------------------A---- 5814
F1.ES.x.P1146           GAGG--C-GG---..------G---AG---......................................---CTTT-AT----G-----A--A--A-C-..------------A-GG-C......--------T----T------------------------- 5655
F1.FI.93.FIN9363        GAGG--T-GG---..------G-A-A----......................................---CTTT-AT----G--------A--A---..----AA---G-G--TG-C......------------A--------------------A---- 5701
F2.CM.02.02CM_0016BBY   GAGG--T-GG---..------G---AG---......................................---TTTT-AT----G--A--A--A--A---..----A----G----T--C......-----------------------A-----A-------A---- 5555
F2.CM.95.MP257          GAGG--T-GG---..------G----G---......................................---CT-T-AT----G-----A--A--A---..----------G----T--G......-----------------------A------------A---- 5574
F2.CM.97.CM53657        GAGG--T-GG---..------G---A----......................................---CTTT-AT----G-----A--A--A---..----A----G----T--C......-----------------------A------------A---- 5555
G.BE.96.DRCBL           GAGG--T-GG--A..-----A----AC---......................................--G-TTT--A-------T--G----A--T..--------AGA-A--.........----------------------A-----------G-A--T- 6326
G.CM.01.01CM_4049HAN    GAGG--T-GG---..------A----A----.....................................--A--TT--A-------T--G----A--T..--------CT--A-T--C......---------------------------C---------G-A---- 5573
G.CU.99.Cu74            GAGG--T-G----..-----------A----.....................................--G--TT--A--A-----T--GC---A--T..--------CT--ART--C......------------A-----------------------G-A--C- 5976
G.ES.99.X138            GAGG--T-GG--A...-C--------A----.....................................--G-TTT--A------AT--G----A--T..----A---CT--A-TG-C......----------------------------------G----T- 5813
G.GH.03.03GH175G        GACG--T-GG---..-----------A----.....................................--G--TT--A-------T--G----A--T..--------CT----C---......----------------------------------G-A--T- 6412
G.KE.93.HH8793_12_1     GAGG--T-GG--A..-----A----A----......................................--G--TT--A-------T--G----A--T..--------CT--A-T--C......----------------------------------G-A--C- 5772
G.NG.01.01NGPL0669      GAGG--T-GG--A..-----------C----....................................T---TT--A-------T--G----A--T..--------CT--AGT----......------------------------A-------G-A--T- 5577
G.PT.x.PT2695           GACG--T-GG--A...-T--------A----.....................................--G--TT--A------AT--G----A--T..--------CT--A-T--C......----------------------------------G----T- 6314
G.SE.93.SE6165          GAGG--T-GGT-A..-----------A----.....................................--G--TT--A-------T--G----A--T..--------CT--A-T--T......--------T-----------------------G-A--T- 5769
H.BE.93.VI991           GAGG--T----CA..AG-------------......................................------T-AA--T-G--T-----C-----..-------T-GT--G---C......--------------A-----------G----------A---- 5744
H.BE.93.VI997           .....--T----CA..--A-G-------G---....................................----T-T-A--TT-G--T-----C----T..-A-------G---G---C......--------T-----A-----------G----------A---- 5671
H.CF.90.056             GAGG--T----CA..AG---A---------......................................------T-AA--T-G--T-----C-----..----------G--C----C......--------T-----A-------------------------A--G- 5706
J.CD.97.J_97DC_KTB147   GATG--T-GGA-A..A---------A----......................................---CT----A--T----A-----A-----T..----A--G-G---------......--------------A----------------------A--T- 5566
J.SE.93.SE7887          GAAG--T-GG---..AC-----------G--.....................................---CT----A--T----A-----A-----T..----AG--A-A----G-C......------------------------------A--T- 5685
K.CD.97.EQTB11C         GAGG--T-GG---..------G---A-C--......................................----T-T-AT--T-G-----A--A--A---..-----C---G--A-C--C......------------------------------A---- 5561
K.CM.96.MP535           GAGG--T-GG---...ACA---G---AC---.....................................----T-T-AT--T-G--A--A-----A---..-----CAA-G-----C---......------------------------------A---- 5565
```

```
                                                                                          Vpu end
                                                                                Env signal peptide end Env gp120 start
B.FR.83.HXB2               ....AAATATCAG...CACTTGTGGAGATGGGGGTGG...........AGATGG..............GGCACCATGCTCCTTGGGATGTTGATGATC...TGTAGTGCTACAGAAAAA......TTGTGGGTCACAGTCTATTATGGGGTACCTGTGTGGAAGGAAG 6366
Vpu                        _._ _E_ _I_ _S_ _._ _A_ _L_ _V_ _E_ _M_ _G_ _V_ _._ _._ _._ _._ _E_ _M_ _._ _._ _._ _._ _._ _G_ _H_ _H_ _A_ _P_ _W_ _D_ _V_ _D_ _D_ _._ _L_ _*
Env                        __.__K__Y__Q__.__H__L__W__R__W__G__W__.__.__.__.__R__W__.__.__.__.__.__G__T__M__L__L__G__M__L__M__I__.__C__S__A__T__E__K__.__.__L__W__V__T__V__Y__Y__G__V__P__V__W__K__E__
01_AE.CF.90.90CF11697      AATG--T-GG-CA...----------AG---...........................................--G--TT--A-------T--G---A--T..---------CT----C-CC......--------T----T----------T---------GA--T- 6315
01_AE.CN.05.FJ051          GATG--T-GG-CA...A---------A---............................................--A--TT--A-------T--G---A--T..----A---CT----C--C......--------T----T----------T----------A--T- 6369
01_AE.CN.06.FJ054          GATG--T-GG-CA...A---------A---............................................--G--TT--A-------T--G---A--T..---------CT--A-C--T......--------T----T----------T----------A--T- 6390
01_AE.HK.x.HK001           GATG--T-GG-CA...A--------------...........................................--G--TT--A-------T--G---A--T..---------CT---C--C......--------T----T----------C----------A--T- 5736
01_AE.TH.01.01TH_R2184     GATG-CT-GG-CA...A---------A---............................................--G--TT--A-------T--G---A--T..---------CT---C--C......--------T----T----------T---------GA--T- 5566
01_AE.TH.02.OUR769I        GATG-GT-GG-CA...----------C-G--...........................................--G-TTT--A--A----T--G---A--T..---------CT-------C......--------T----T-------A--C---------GA--T- 5565
01_AE.TH.90.CM240          GATG--T-GG-CA...A---------A---............................................--G--TT--A-------T--G---A--T..---------CT---C--C......--------T----T----------G---------GA--T- 5934
01_AE.US.00.00US_MSC1164   GATG--T-GG-CA...A---------A---............................................--G--TT--A-------T--G---A--T..---------CT---C---......--------T----T----------T--G--------A--T- 5571
02_AG.CM.02.02CM_4082STN   GAAG--T----CA...-T---A---------...........................................--T-TG--AA-AT--T-T--AA---A--T..----A-----A---T.........----------G----C------------------A--C- 5568
02_AG.EC.x.ECU41           GAAG--T----CA...-T---C---------...........................................--GGTG--AA-AT--T---CAA-------T..----A----G--A-T.........----------------C----------A-------GA--C- 5464
02_AG.GH.03.GHNJ196        GAAG--T----CA...TG--------G---............................................--G-TT--AA--T--T----AA--C----A..----AG-G-GA---T.........C---------G-------------------------GA--C- 6398
02_AG.NG.01.PL0710         GAAG-GT-G--CA...-C---A----A---............................................--T-TG--AA-AT--T---CAA---A--T..----A----GAT--G.........----------G----C------------------A--C- 5553
02_AG.NG.x.IBNG            GAAG--T----CA...-T---A---------...........................................--T--G-ATA-AT--T----AA---A--T..----A----GA-C--.........----------G----C--------------------ACG- 5893
02_AG.SE.94.SE7812         GAAG--T----CA...-T---A---------...........................................--T-TG--AA-AT--T----AA---CA--T..----------GG-A-T.........----------G------------------------GA--C- 5741
02_AG.SN.98.MP1211         GAAG-GT-G--CA...-C---C----------T.........................................ATG-TA-----G---T----A-----A--T..----A----GA-A-T.........----------G----C---------------C---GA--T- 5574
03_AB.RU.97.KAL153_2       GAAG...C-C.........--A--------------......................................------T--T----------C--------..----------------T.......--A----------T-----------------------A---- 5579
04_cpx.CY.94.CY032         GAGG--T----CA...---------GA---............................................--A--TT--A-------T--G---A---..---------CT--A-C--C......--------------T---------------------GA--C- 5736
05_DF.BE.x.VI1310          GAGG--T-GG-C-...------G---AG---...........................................---CTTT-AT----G-----A--A--A---..---------------C---.......--T------------------------------A----GA---- 5754
06_cpx.AU.96.BFP90         GACG-GT-GG--A...-----A----A---............................................--G--TT--A-------T--G---A--T..---------CT-TA-G--C......A-------T-----------------G----C----G-A--T- 6402
08_BC.CN.97.97CNGX_6F      GAGG--T-----A...--A-G-----T---............................................---GT-T-AGG-T--T-------A-----T..----A--TGGA--G---C......-------------------------------------A---- 5539
09_cpx.GH.96.96GH2911      GAGG--T-GG--A...A--------------...........................................---CT----A-T-----A-----------T..----A----G-------C......-------------------------------------A--G- 5577
10_CD.TZ.96.96TZ_BF061     GAGG--T-G--GA...--A-G-----T---............................................--A-T-T-AGG-T--T-------A-----T..----A---A-GG--G--G......--------------A----------------------A---A 5763
11_cpx.GR.x.GR17           GAGG--T-GG--C...A--------------...........................................---CT----A--T----A-----A------..----A---AGA-A--.........A----------------------------------GA--T- 5669
12_BF.AR.99.ARMA159        GAGG--T-GG---...--T---G---AC---...........................................---TT-T-AT----G--A--A-----A---..----A------------T......-------------------------------------A---- 6367
13_cpx.CM.96.1849          GATG-CT-GG--A...A---------A---............................................--A--TT--A-------T--G---A--T..---------CT--A-T--C......--------------T-----------T---------GA--T- 5776
14_BG.DE.01.9196_01        GAGG--T-G---A...-G--------A---............................................---GTG-C----A----------A------..----A----GT----C--......C--------------------A--------A-----A---A 5889
15_01B.TH.99.99TH_MU2079   GATG--T-GG-CA...A---------A---............................................--G---T--A-------T--G---A--T..----------G-------C......------------------------------A-----A--C- 5762
16_A2D.KR.97.97KR004       GAGG--T-----A...-----A----A---............................................--A-T-T--A-TT-G------C-A-----T..A--AA---------G-T......C------------A--C-----A-------------GA--T- 5732
17_BF.PE.02.PE02_PCR0155   GAGG--T-------...----------A---...........................................----T---------------A--A-------..----A---G---A--G-T......-------------------------------------GA---- 5552
18_cpx.CU.99.CU76          GAGG--T----CA...AG---A----A---............................................--A----AA---G--T------C---G..---------------TC-C......--------------A-----------G----------A--T- 5665
19_cpx.CU.99.CU7           GAGG-GT----CA...-C---A----AT--............................................--G--TT--A-------T--G---A--G..---------CTTT--C--C......--------------T-----------T---------GA--T- 5588
20_BG.CU.99.Cu103          GACG--T-GG---...TG--------C---............................................--G--TT--A--A----T--GG---A--T..---------CT--A-T--T......--------T--------------------------G-A--C- 5839
21_A2D.KE.91.KNH1254       GAGG--T-G---A...A---------A---............................................----T-------T---------A-------..----------------T--T......C-------T----T-----------G-----A----G----- 5567
22_01A1.CM.01.01CM_0001BBYGATG-CT-GG--A...A--------------...........................................--A--TT--A-------T--G-A--A--T..---------CCA-A-T.........--------------T-----------C---------GA--G- 5572
23_BG.CU.03.CB118          GAAG--T-GG---...-G--------C---............................................--G--TT--A--A----T--GG---A--T..---------CT----T--C......-----------------------------------G-A--T- 5815
24_BG.CU.03.CB378          GATG--T-GG--A...----------C---............................................--G--TT--A--A----T--GG---A--T..---------CT--A-T--C......-------------------------------------A--C- 5803
24_BG.CU.03.CB471          GATG--T-GG--A...-T--------C---............................................--G--TT--A--A----T--GG---A--T..---------CT--AGT--C.......--A--------------------------------G-A--C- 5812
25_cpx.CM.01.101BA         GAGG--T----CA...-T---A---------...........................................--G-TTT--A--A----T--GG---A--T..---------CT----C--C......-------------------------------------A--T- 5580
27_cpx.CD.97.97CDKTB49     GAGG--T-GG--A...GG-------------...........................................---CT----A--T-G--A---GCA-----T..-----------A----G-C......-----------------------G----------GA--T- 5579
28_BF.BR.99.BREPM12609     GAGG--T-GGA--...------G---C---............................................---CT-T--T--T-G--A--A--A--A--T..----A-------A-CC--.......-CA---------------------------------A---- 5728
29_BF.BR.02.BREPM119       GAAG--T-G----...----------GA---...........................................----T-T---------------A------..---------GA-C--.........----------------------------------A---- 5533
31_BC.BR.02.110PA          GAGG--T-GG--A...--A-G-----T---............................................----T-T-AGG-T--T-------A-----T..-A--A--TA-GG-GG--C......--------------A-------------------GA---- 5817
32_06A1.EE.01.EE0369       GATG--T-GGTCA...-----A----A---............................................--A--TT--A-------T--G---A--T..----A---CT--A-GG-C......C------------T-----------G---C----G-A--T- 6007
33_01B.MY.05.05MYKL007_1   GATG--T-GG-CA...A---------A---............................................--G--TT--A-------T--G-A--A--G..---------CT--A-C--C......--------T----T-----------T---------GA--T- 5751
34_01B.TH.99.OUR2478P      GATG--T-GG-TA...A---------A---............................................--A--TT--A------AT--G---A--T..---------CT---C--C......--------T----T-----------T---------GA--T- 5568
35_AD.AF.05.05AF095        GATG--T-GG--A...A--------------...........................................--A--T---A--T-G--A---A-A--A--T..---------GT------T......--------T--T------------------------A--C- 5565
36_cpx.CM.00.00CMNYU830    GAGG--T-G--TA...----------A---............................................--G--TT--A--A----T--G---A--T..---------CT----CTTG......--------T----T-----------T---------GA--T- 5581
37_cpx.CM.00.00CMNYU926    GAGG--T----CA...-C---A----AT--............................................--G--TT--A-------T--G---A--T..---------A-ACA-CCTC......--------------T--C--------T----------A--T- 5563
38_BF1.UY.05.UY05_4752     GAGG--T-GG---...-C----G---A---............................................A--TTGT-AT----G-----A--A--A--A..----A----GA---C.........-----------A---------------------GA---- 5786
39_BF.BR.03.03BRRJ103      GAGG--T-GG---...------G---A---............................................---CTTT-AT----G-----A--A--A---..----A----G-------C......-----------------------------------G---C- 5862
40_BF.BR.04.04BRRJ115      GAAG--T-G----...-G---A----A---............................................----TG----------T-----------..----A---C-AT--CC--.......-G-----------A---------------------A---- 5914
42_BF.LU.03.luBF_05_03     GAAG--T------...----K-----A---............................................----T-------------A---------..A-----------------......--------------T---------------------A---- 5900
43_02G.SA.03.J11223        GAGG--T-GG--A...T---------C-AT-...........................................T-G-SYT--A-------T--GK---A--T..---------CT--A-T--C......----------------------------A------G-A--T- 5897
U.CD.83.83CD003            GAGG--T-----A...----------A---............................................A--CTG--AA-TT-A--A---A-A-----A..----AA----T-------.......-C------------------------------------A--T- 5854
U.CD.90.90CD121E12         GAGG--T-----C...T---------A---............................................A--CTG--AA-TT-A--A---A-A--A--A..----AA-G-GA----C--.......-C-----T-----------------------------A--T- 5555
U.GR.99.GR303              GAGG--T-G---A...A----A---------...........................................-C-CTG---A-TT----A-----A-----T..----AG-------C--C......-------------------------------------A--T- 5644
U.NL.95.U_NL_95_H10986_D1  GAGG--T-GG-G-...------G--------...........................................--T-T-T--T--T-A-----A---------..---------G----C---......------------A----------------------C-A--T- 6254
CPZ.CD.90.ANT              TATT.....................................................................................TGGGG-CT-GCT---C-A---CAG-T--TA-AG-AG-GG-CGAATGAAGAC-AT--A-----A-TC-----A--C-----C----GAA-T- 5792
CPZ.CM.01.SIVcpzCAM13      GAAG-G-AGCTTC...A--GGC-TCTACAT-ACC.............................-TG-CAA-ATGCAT-GCT---C-A---CCT---TTGA-AT-T--T.........GAA----A----G-------A-------------GA---- 5944
CPZ.CM.05.SIVcpzEK505      GAAG--T-GG..................................................CTAAT-TG-TGC-----GATAGGA-----C..........-AGATA-TT-GGTCT...GAA-------A----G--C-------------C----GA--T- 5915
CPZ.CM.05.SIVcpzLB7        GACG-GC-GCAT-..............................CCATAC-ACGTAGTGGGGATCCTGA--TTGTGCTCAT-GATA--TAGT.........A-C-C-AAA-AT-----T........-A----G----G--C-----A--G-----C-----A---- 5931
CPZ.CM.05.SIVcpzMB66       GAGG--T-GCAG-...AC---A-CACTGAA--A-ATA.........................AT-TTATGCACAT-G-TAT-AGG---A--A..G-A-TAAT--AGTGTG--GATAACA-------G-----T--C-----A--------C----GA--G- 5948
CPZ.CM.05.SIVcpzMT145      ATATT-T-GGTG-...AGGAAC-CAT-CA--TCCATC.........................A---T--A-T-AA-CCT---TGGATG-----..GCATCCTG-TTT-GTG--GAAAAT-G---------T--T-----------C--A--A----G---G- 5911
CPZ.GA.88.GAB1             GAAG--GAGAG-C...TGGAACA-CTT--CCATAATT.........................ACA-T---AACAA-CATTT--C-A-CCCCA..---TTGA-CT-T--G..........-A-----A----A--------A-------T---C-T--T- 6430
CPZ.TZ.00.TAN1             TTTA-TTGGA.........A-AACTTTGATC.....................................C-CA-AA--ACA--CC-AGG---T..G-ATT-AGC--T............-AT-ACAC------G-T------A--------T-----A--G- 5986
CPZ.US.85.CPZUS            GAAG--GAGA.........C-C---CT-A-T.............................TATTGTC-T---TCGA-CT--A-A--CCCA..G-ATTGT---GC............--A----CT----A--------------------A----GA--T- 6437
CPZ.CM.x.SIVcpzDP943       CAGAGGCAGC..........................................................T-ATGG--AT-GGT-----A---CCA---TTGA--GGTAGTG-T...CAG-------CA-----G--C-----A------------------A--T- 5944
CPZ.CM.x.SIVcpzMB897       GAGG-GT-GGA--...--A--C-TATCCCT-A--.........................ATCGTCATGTG-T--TGG--ATTAGG-G-T.........AT---AT--GA-A-T--T.......--A-----A-----G--C-----A--------A-----A--G- 5937
N.CM.02.DJO0131            GATTGGT-GG...................................................CCATT-TT-TGT---A-GATA-GT----CA--T..G-GTC-AAA-A-............-AT----CA-----G--C-----A-----A--A----GA--T- 5838
N.CM.04.04CM_1015_04       GAGTGGT-GG...................................................CCATT-TT-TGT-----GATA-GC-C---A--T..G-GTC--AACA-...........CAT-----A-----G--C---------G--A--A----GA--CA 5837
N.CM.04.04CM_1131_03       GAGTGGT-GGGT-ACAAGGAGAGT-TTGG--AT-AT-AGTGGTTGG...................CCATT-TT-TGT-----GATA-GC-C---A--T..G-GTC--AACA-...........CAT-----A-----G--C-----------A--A----GA---A 5872
N.CM.95.YBF30              GAGTGGT-GG............................ATGGGGATG-AGA-TGGTTGGTTACTCTTCTATCTTC-AG-AAGCTT---CAA-G-A--T..G-GTC--AACA-...........CAT-----A-----G--C-----------A--A----GA---- 5953
N.CM.97.YBF106             GAGTGGT-GG...................................................CCATT-TT-TGT-----GATA-GC----CA--T..G-GTC--ACC--...........CAT-----A-----G--C-----A------A--A----GA--C- 5923
O.BE.87.ANT70              GAAC--GA-G.........--A----CC-T-.......................................TA-T-AGC-A-G-CTT--A-A-CCCCA..---TTGAGCCTTAG-C-G......C-A-AT-CA---------GC-----G-----A---G-A--T- 6413
O.CM.91.MVP5180            CAAC-GGA-G.........-CA-----C-TA.......................................TA---AGC-A-G-CTT--C-A--CCCA..---TTGAG-TATAGT---...CAA----AT-C------T----C-------------A---G-A--G- 6403
O.CM.96.96CMABB637         GAAC--GA-G.........--A--C-TC-T-.......................................TG---A---A-G-CTT--C-A--CCCA..---TTGAACT-T--CC-G......C-A-AT-CA-------T--C------G---------G-A--T- 5855
O.CM.98.98CMA104           GAAC--GAGG.........--A----TC-TA.......................................TA---AG--A-G--TT--A-A-CCCCA..---TTGAGCT-TAGCC-G.......--T-AT-CA---------GC---A--G-----A---G-AA-T- 5854
O.CM.99.99CMU4122          GAAC-G-C-G.........--AG---TC-TA.......................................TA---AG--A-G-CTT--A-A--CCCA..---TTGAGCTAT--CC-G......C---AT-CA----------C-----------A---G-A--G- 5853
O.FR.92.VAU                GAAC-GGA-G.........--AG---TC---.......................................TG-T--A-TT-G-CTT--A-A--CCCA..---TTGAGCTGTA-CC--......C-A-AT-C-----------C-----------A---G-A--T- 5937
O.SN.99.SEMP1299           GAAC-GGA-G.........--AG---TC-T-.......................................TG---AG-TA-G-CTT--A-A-CCCCA..---TTGAGCTATA-CC-G......CAC-AT-CA---------GC-------------A---G-A--G- 6424
O.US.99.99USTWLA           GAAC-GGA-G.........-CA----TC-T-.......................................TG---RA--A-G-CTT--A-A--CCCA..---TTGAACT-T---C-G......C-A-AT-C-----------C------G---------G-A--T- 5854
```

```
B.FR.83.HXB2          CAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTGGTAAATGTGACAGAAAATTTTAACATG...TGGAAAAATGACATGGTAGAACAG 6533
Env                   A__T__T__T__L__F__C__A__S__D__A__K__A__Y__D__T__E__V__H__N__V__W__A__T__H__A__C__V__P__T__D__P__N__P__Q__E__V__V__L__V__N__V__T__E__N__F__N__M__.__W__K__N__D__M__V__E__Q_
A1.GE.99.99GEMZ011    --GAG-----C-----C-------------------------A---A-----C-----C-----T--------------------------TG-----A-----A--A-T----A-------------A---------...-------G-A-------------A 5717
A1.KE.00.KER2008      --GAG-----CT-----------------------G-----AAA---------C-----T----------------------------G----------A--CCT---AA---------------G---------...-----------A-------------- 5742
A1.KE.00.KNH1144      --GAG-----CT---------------------A-A-----AAAG--------C-----T-------------G-------------------------A--CCT----A-------------G-G---------...-----------A-A---------A 5750
A1.KE.00.KSM4024      --GAG------T---------------------A-G-----AA-G--------C-----T---------------------------------------A--CAT--A-A-------------G-A--C------...----------CA-----------A 5750
A1.KE.00.MSA4069      --GAG-----CT----------------G-------G------A-----AAAG--------C-----T---------------------------------------A--CACC--AA-------------GGG--------A...-----------A-A-------G--- 5750
A1.KE.00.NKU3005      --GAT-----C-------------------------G-A---A--G--------C-----T-------------G-----------------------A--C-T---AA-------------G-G---------...-----------A----------G--- 5735
A1.RU.00.RU00051      --GAG-----C-----------------------C----A-----------C-----T--------------------------TG-----------A--ACT----AC-----------G-A-----T---...-----------A-------------- 5834
A1.RW.93.93RW_024     --GAG-----C-----------C--------------------G-----A--G--------C-----T-------------G-----------A-----------A--CCT----A---------------G---------...-----------T-T--------G--- 5735
A1.SE.95.SE8891       --GAG-----CT-------------------A----------------------G--------C---------------------T-----------------------A--AAT----A-------------G-G---------...-----------A----------G--- 5728
A1.SE.95.UGSE8131     --GAG-----C--------------------C-----------G-A---A--G--------C-----T---------------------T-----------------T--TAT----A-------------G-----------...-----------A-T--------G--- 5884
A1.TZ.01.A173         --GAG-----CT-G-----------------G------------G----A--G--------C-----T---------------T--------------------------A--CAT----A-------------G-G---------...-----------A----------G--- 5733
A1.UA.01.01UADN139    --GAG--------------------------------------------A---A-----------C-----T--------------------------T--------------A--CCT----A---------------------T---...-----------A-------------A 5732
A1.UG.92.92UG037      ---AT-----CT-------------------G-----------------A--G--------C-----T--G--------------T--------G----------C--AA-A---A-------------G-G---------...-----------A----------G--- 6558
A1.UG.99.99UGA07072   --GAG-----CT-------------------G-----G-----------A--A-G--------C-----T--G--------------T-------------------------A--AAT---AA-------------G-A---------...-----------A----------G--- 5739
A2.CD.97.97CDKTB48    --GAT-----C-----------------C-------------C------AAAG--------C-----------------------------------------------AACC---C------A---------G-----------...-----------A----------G--- 5876
A2.CY.94.94CY017_41   --GAT----TC---------------------------------------A--G--------A---------------------------------------------A--AACC---A------A---------------T---...-----------A----------G--- 5893
A.SN.01.DDI579        --GAG-----C---------------------------------CA---A--G--------------T-----C----------------------------------GT---AT----A-------------G-G---------...-----------A-------G--G--- 5740
A.SN.01.DDJ369        --GAG-----C--------------------------------GCG---A--G--------------T-----C--------G------------------------A---AT----A-------------G-G---------...-----------A-T-----G--G--- 5743
A.SN.96.DDJ360        --GAG-----C---------------------------------CA---A--G--------------T-----C-----------------------------------T---AT----A-------------G-A---------...-----------A-------G--G--- 5740
A.CD.02.02CD_KTB035   --GAA-----C--------------------------------G-A---AA-G--------C-----T---------------------------G-----------A-G-ACC----C------------------...-----------A-T--------G--- 6520
A.CD.97.97CD_KCC2     --GAA-----C--G------------C-------T--------G-A---ACA-----T---C-----T--C-----------T---------------------TC---A-G-AT--A-AT-----------G-G---------...-----------A-T--------G--- 6532
A.CD.97.97CD_KTB13    --GAA-----C-----C-----------C--C-----------G-G---A--G--------C-----T--------------T-----------------------T-A-GCCTC-A-CCG-----------GTA---------...----TG-----T--------G--- 5743

B.AR.04.04AR151516    -G-----T-----------------------------------C-------C------------------------------------------T--Y-----------C----RAA--------------------------...-------------------------- 5748
B.AU.87.MBC925        --------------------------------------------------------------------------------------------------------------------AA-----------------------------...-----------A-------------A 6540
B.BO.99.BOL0122       -----------------------------------------AG-------AA-------A--------------------------------------------------CC-----A-----------------A--------...-----------A-------------- 5752
B.BR.03.BREPM2012     ----T------T-------------------G-----------A------A-G-----------------------------------------T--T-------------------A-----------------------------...----------------G--C--- 5969
B.CA.97.CANB3FULL     --CA--------------------------------------------------------------------------------------------------G--------------G-------------------------...-----------A-------------- 5821
B.CN.05.05CNHB_hp3    ----T-----------------------G-------------C----------------A----------------------------------------------------C----A------------------------...---G-----A-------------- 6534
B.CO.01.PCM001        ---A---TGT------------C-------------------G-----------------C-----------G--------------------------------------------A-----------------------...------G---------C--- 5731
B.GB.83.CAM1          -------------------------------------------C--------------------------------------------------T----------------------A-----------------------------...-----------A-------------- 6530
B.GB.86.GB8           -------------------------------------------C-----AA-------------------------------------------------------------G----------G-----------------------------...-----------A-------------- 6542
B.GE.03.03GEMZ010     -------------------------------------------C------A-G-----------------------A-----------------TG----------G---AG-----A---------------------T---...-----------A----------T--A 5734
B.IT.05.SG1           -------------------------------AT---G----AG--------C-------------------------------------A-----G---------------------A-----------------------------...-----------A-------------- 6338
B.JP.05.DR6538        --GAA-----------------C-----C------A-C-------------C-----G---------------------------------A-------------------------A--------------------------...G--------A-A-----------A 6539
B.KR.05.05CSR3        -C-A-------T----------------------------CAG------------------------------------------------T---------------A---------A------A---------------T---...-----------A-------------- 6554
B.NL.00.671_00T36     -------------C-----------------------------A--------------------------------------------------T------------C-------A-A---------------------T---...-----------A-------------- 6094
B.RU.04.04RU128005    --GA--------------------------------G-Y--AG---------------------------------A------------------------------A---AC----A------------------------...---C-----A-------------- 6020
B.TH.00.00TH_C3198    ---T-----------------------------------------------C-----------------------------------------------------------C----A----------------C------------...-------------------------- 5743
B.UA.01.01UAKV167     ---AA-----------------------------G------C-C--------------------------------------------------T------------A-----A---A---------------------T---...-----------A----------T--- 5767
B.US.04.ES10_53       ---A------------------C------------------AG---------------------------------------------------------------------C----AA-----------------------------...-----------A-------------A 6553
B.US.99.PRB959_03     ---------------------------------G-------A---A------------------------------------------------------------------------...--------------A---------...----------CA----------T--- 5746

C.AR.01.ARG4006       ---AA--T---T------------------------G------------A--G--------C-----T-----------------------------T---------T-G--T---AA------A---------------------...---G-------------T--- 5721
C.BR.04.04BR013       ---AA--T--C--------------------------------------A--G--------C-----T-----------------------------T---------A-G--T------------A-------------------...---G-------------T--- 6015
C.BW.00.00BW07621     ---AG--T-----------------------------------C-G---A--G--------C-----T---------------------------------------C-G--T----A------A-----------------------...-----------T-------G--C--- 5904
C.CN.98.YNRL9840      ---AA--T--------C--------------------------G-A---A--G--------C-----T---------------------------------------A-G-AT----A------A---------------------...----------------G--T--- 5716
C.ET.02.02ET_288      ---G-C-T-----------------------------------------A--G--C-----C-----T---T-----------------------------------T-G--TA---A---------------------T---...----------------G------ 5753
C.GE.03.03GEMZ033     ---AA--T----------------------------G------A-A---A--G--C-----C-----T-----C-----------------------T-----G---C----T----A------A-----------------------...-------T-----G--T--A 5728
C.IL.99.99ET7         ---AA--T---------------------------A-------G-G---A--G--------C-----T-----------------------------------T---TGG--T----A------A-----------------------...----------------G--C--- 5726
C.IN.99.01IN565_10    ---AA--T--------C--------------------------C-A---A--G--------C---------------------------------------------T-GT-G-----------A---------------T---...----------------G--T--- 5913
C.KE.00.KER2010       ---AAG-TC-------C--------------------------G-A---A--G--------C-----T----------------------------G----------A----T----A-----------------------------...-------T-----G--T--A 5722
C.MM.99.mIDU101_3     ---AA--T--------C---------A----------------G-----A--G--------C-----T---------------------------------------A-G-AT----A------A---------------------...---G-C-----T-----G--T--- 5883
C.MW.93.93MW_965      ---AA--A--------------------G--------------G-A---A-----------------T---------------------------------------A-G--T----A------A---------------T---...-------T-----GA-T--- 5735
C.SN.90.90SE_364      ---AA--T-----------------------------------G-A---A--G--------C-----T---------------------------------------A-G--T----A------A------T---------------...----------------T--- 5720
C.SO.89.89SM_145      ---AA--T-----G---------------T-------------G-G---A--G--------C-----G----------------------------G----------A-G-C-T----G------A-----------------------...-------T-----G--T--- 5762
C.TZ.02.CO178         ---G---T--------C-------------------G------G-A---A--G------A-C-----T--------------------------------------GC-G--T----C------A--------C------------...----------------G--G--- 5716
C.UY.01.TRA3011       ---AAG-T--------------------C--------------------A--G--------C-----T---------------------------------------A-G-CT----A------A---------------------...---G-----A-------G--T--- 5722
C.YE.02.02YE511       ---AA--T-----------------------------------G-GC--A--G--------C-----T---------------------------------------A----T----A------A-------G-------------...-------T-----G--T--- 5725
C.ZA.04.04ZASK164B1   ---AA--T--------C--------------------------G-G---AA----------A-----T--------------------------------G------A----T----A-----------------------------...----------------G--T--A 5956
C.ZA.05.05ZAPSK240B1  ---AA--T-----------------------------------C-----A--G--------------T--------------------------T------------A-G--T----AC-----A-----------------------...-------T-----G--T--- 5955
C.ZM.02.02ZM108       ---AA--T--------------------------G--------G-A---A-----------C---------------------------------------------C---A----AA------A-----G-----------------...-------T-----G--T--- 6519

D.CD.83.ELI           -----------------------------------T-------A-------C-------A-C--------------------------G------------------A---C-C---A--------------C------------...-----------A-------G------ 6075
D.CM.01.01CM_0009BBY  ---A-G-----------------------------------A-AG------C-------A-C-----T--------------------G--------T---------A---A-C--AA------C---------------T---...-----------A-------G-----A 5731
D.KE.01.NKU3006       -------T---------------------------T-------AG------C-------A-C-----T-----------------A------------------G----A--AA-C-A-A------C--------C------------...-G---A-T-----G--G--- 5734
D.KR.04.04KBH8        -------------------------G--------CT--CT-A-AG------C-------A-C--------------T-----------G--T----G------G-G-----------C---C------C-----G-C---A-CATGT...-----------A-------G--C--- 6485
D.TD.99.MN011         ---AAG-----------------------------T--AT-A-A-------C-------A-C-----TT-------------------G-------GT--C------A--AGC----A------C--------C------------...-----------A-------G--G--- 5756
D.TZ.01.A280          ---A-------------------------------T---T---C-------C-------A-C-----T-----------C-----A---------------------A-G-ATC-----------TT-------C------------...-----------A-------G------ 5730
D.UG.99.99UGK09259    -------T---------------------------T-----ATA-A-----C-------A-C-----T-----------------A---------------------A--AATC-A-G------C--------C------------...-G-----------------G------ 5734
D.YE.01.01YE386       -------T----------------------------G--TCA-A-------C-------A-C-----T-----------------A---------------------A--A--C-TAA------C--------C------------...-----------A-------G--G--- 5728
D.YE.02.02YE516       ---AAG-----------------------------T--GT-A-A-------C-------A-C-----TT-------------------G--------T--C------A--AGC---AA------C--------C------------...----------------G--G--- 5743
D.ZA.90.R1            -------T----------------------------G----A-A-A-----C-------A-C-----T--------------------G--------------------G-C-----A----------------C-----T---...---G-----------------A-C--- 5878

F1.AR.02.ARE933       -------T---T----C------------------T-------A-G---------------C-----T-----------------------T---G-T----G---------CC-N-A------A--------------G-T---...-----------A-------------A 5833
F1.BE.93.VI850        -------T--------C--------------------------A-G-----C------C---------------------------------------------------T-TC--AA------A--------------G-T---...-----------A-------------A 5879
F1.BR.01.01BR125      -------T--------C--------------C-----------A-A---------------C-----T--------T--------------T-----T---------A-G--TC-ACC------A--------------G-T---...-----------AG------------A 5981
F1.ES.x.P1146         -------T-----------------------A----G-----TACA-----C---------C-----T-----------------------------------G---A--TATC---A------A---------------------T---...---G-----A-------------A 5822
F1.FI.93.FIN9363      ---A-------T----C------------------T-------A-A---A-----------C-----T-----------------------------T-------------CTC-A...----------------------T---...---G-------------------A 5865
F2.CM.02.02CM_0016BBY -------T-----------------G-----------------A-A-----C---------C-----T-----------------T--------TG-----------A--T-TC---AC-----A-----------------------...-----------A----------C--- 5722
F2.CM.95.MP257        -------T-----------------------------------A-G---------------C-----T---T-------------T--------A-G----------T-G--TC-T-G------AT-------A------------...-----------A----------C--A 5741
F2.CM.97.CM53657      -------T-----------------------------------A-G----A----------C-----T---T-------------T---------------------T----TC---G------A-----------------------...-----------A----------C--- 5722

G.BE.96.DRCBL         ---ATG--C----------------T------------C--AG--T---AAGT------A-C--------------T--------------T----G----------A--AATA--AG------A-----------------------...-----------A-------------- 6493
G.CM.01.01CM_4049HAN  ---ATC----------C--C-----T---------------AG-T-T--AAA---------C-----T---------------------------------------A----TA---A------A-----------------------...-----------A-------------- 5740
G.CU.99.Cu74          --GATG--C----------------T---------------AG--T---AAGC--------C-----T--------A-----------------------------GA--CCTC---A------A--------G-----------------...-----------A-------------- 6143
G.ES.99.X138          ---AT--------------------T---------------AG--T---AAGC--------C-----T-------------------------------------GA--CCTC---A------A-----------------------...-----------A-------------- 5980
G.GH.03.03GH175G      --GAT-----------C--------T---------T-----AG-T-T--AAA---------C-----------------------------------T---------A----TA--AA------A--T---T------------------...-----------A-------------- 6579
G.KE.93.HH8793_12_1   ---AG--------------------T---------------AG--T----AG-------A-C-----T---------------------------G----------GA--CCCC-A-G------A---------------G--...-------------------------- 5939
G.NG.01.01NGPL0669    ---ATG-TC----------------T---------------AG--T---AAG---------C-----------------------------T--------------GA-GA-TC-A-A-------------CC-----------------...-----------A-------------A 5744
G.PT.x.PT2695         --GAT--------------------T---------------AG--T---AAGC--------C--------------------------------------------GA--CCTA--AAG-----A-----------------------...-----------A-------------- 6481
G.SE.93.SE6165        --GAT--------------------C---------T-----AG-G-T--AAGC--------C-----T---------------------------------------A--A-TA---A------A------T---------------...-----------A-------G-----A 5936

H.BE.93.VI991         ---AG--T--------C--------------------------C-----AAGG--------A-----T---------------------------------------A-G--C----A------A-------CC-----T---...---GT------------------- 5911
H.BE.93.VI997         ---AA-----------C-----------------G--------GC----AAAG--------C-----T----------------------------G---------GA-G--CC---C------A-----G-C------T---...---G-C-----T--------G--- 5838
H.CF.90.056           ---AA-----------C-----------------G--------G-----AAAG--------C-----T--------A---------------------------GA-G--CA--AG------A-----G-GC-----T---...---G-----A-------G--G--- 5873

J.CD.97.J_97DC_KTB147 ---AG------T-------------T---------T-----AG--T---AAA---------C-----T--------T-----------------A-------AG--GA--ACTA---A------A---------------T--A...-------------------------- 5733
J.SE.93.SE7887        ---AG------T-------C-----T---------------AG--T---AAA---------C-----T--------T--------------T--A-G---------GA-GAATC--CC------A--------C------------...----------------C--- 5852

K.CD.97.EQTB11C       ----A-----------C--------------------------A-----------------C-----T--------------------G-----------------------T----A------A-----------------------...-----------A----------G--- 5728
K.CM.96.MP535         ----AC----------C--------------------------G-AG--------------C-----T--C-----------------G----------------------AGA---A------A--------C------------...-----------A----------G--- 5732
```

```
B.FR.83.HXB2                  CAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTGGTAAATGTGACAGAAAATTTTAACATG...TGGAAAAATGACATGGTAGAACAG 6533
Env                           A__T__T__T__L__F__C__A__S__D__A__K__A__Y__D__T__E__V__H__N__V__W__A__T__H__A__C__V__P__T__D__P__N__P__Q__E__V__V__L__V__N__V__T__E__N__F__N__M__.__W__K__N__D__M__V__E__Q_
01_AE.CF.90.90CF11697         --GAT-----C-------------------C--C---C-A--G-----A--G--C-----C-----T-------T-----------------------------A--CACC---A------A-----------G-----..---------A---------G--- 6482
01_AE.CN.05.FJ051             --GAT-----C--G----------------C------C--A-G-----A-----C-----C-----------------------------------------A-GCACC---A------A-----------------..---------A---------G--- 6536
01_AE.CN.06.FJ054             --GA------C-----C-------------C-G----C----G-----A-----C---A-C----T------------------------------------A--C-C---A------A-----------------..---------A------------- 6557
01_AE.HK.x.HK001              --GAT-----C-------------------C------C----G-----A--G--C-----C--------G------------------------------A--AA-C---G-G----A-----------------..---R-----A-A----C---G--- 5903
01_AE.TH.01.01TH_R2184        --GAT-----C-------------------C------C----G-----A-CG--C-----C--------G------------------------------A--CACC---A------A-----------------..--------CA---------G--- 5733
01_AE.TH.02.OUR769I           --GAT-----C-------------------C------C----G-----A--G--------C------------A--------------------------A--CACC---A------A-----------------..---------A---------G--- 5732
01_AE.TH.90.CM240             --GAT-----C-------------------C--G---C----G-----A--G--C-----C-----------------------------------------A--CACC---A------A-----------------..---------A---------G--- 6101
01_AE.US.00.00US_MSC1164      --GAG-----C-------------------C-G----C----G-----A--G--C-----C-----------------------------------------A--TACC---A------A-----------------..---------A-G----C---G--- 5738
02_AG.CM.02.02CM_4082STN      --GAG-----C---------------------G------------------------------T----------C-----------T-G---------A-GCCT---A------A-----G-------------..---------A---------G--- 5735
02_AG.EC.x.ECU41              --GAG--T--C----------------------------C-G-----A---------------T--------------------T-G---------A--AAT---G------A-----------------..---------A---------G--- 5631
02_AG.GH.03.GHNJ196           --GAT-----C--G----------------G---T--------------A---------------------------------T--T-G---------A--TAT---A------A-------------T---..---------A------------- 6565
02_AG.NG.01.PL0710            --GAG-------------------------------------------------C---------C-----T--------------------T-----------A--CAT---A------A-------------T---..---------A-A----------- 5720
02_AG.NG.x.IBNG               --GAG-----C--------------------------------------A----------C-----T--------------------T-----------A--CAT---A------A------G-----------..---------A------------- 6060
02_AG.SE.94.SE7812            --GAG-----C--------------------------------------A----------C-----T----------C---------T-----------A--CAT---G------A-----G-------------..---------AG--------G--- 5908
02_AG.SN.98.MP1211            --GAG-----C-------------------C-----------G-----A----------C-----T--------------------T-G---------A-GCAT---A------A-------C---------..---------A------------A 5741
03_AB.RU.97.KAL153_2          ----------------------------------------AG--AG--------------------T-------------G-------G---------A--CC----AA----------------------...G-------A------------- 5746
04_cpx.CY.94.CY032            --GAG-----C-----------------A------------G-A---A--------A-C-----T-----------------------------------------CTC--A------A-----G--C---------..----------------------- 5903
05_DF.BE.x.VI1310             ---AA--T--C--G---------------------G------C-A---------------C-----T-------T------------------------------TC---G------A-------------T---..----------------------A 5921
06_cpx.AU.96.BFP90            --GAT----T---------------T--------------AG-C-T--TAAG--------C-----A-------T---------------------------A--TCTC--AA-----A-------------T---..---------A--------C--- 6569
08_BC.CN.97.97CNGX_6F         ---AA--T--------C-----G-----------------------G-----A--G--------C-----T---------------------G---------A----TA---A------A-----------------..-----T-----T-----GA-T--- 5706
09_cpx.GH.96.96GH2911         --GAG-----CT------------------C-G-----CTC--A---AAA--------------T------------G-----G---------------------A---GTC---A------A-----------------..---------A---------G--- 5744
10_CD.TZ.96.96TZ_BF061        -------T-----------------------------C--CA-AG------C-------A-C-----T----------C-----T---------------G---A----T----A------A-------------T---..----------G------G--T--- 5930
11_cpx.GR.x.GR17              -GGAT----T---------------T-------G-A-----AG---T--AAA---------C-----T-----CT-------------------------G----C--AGTC-A-A------A-----------------..---------A------------- 5836
12_BF.AR.99.ARMA159           -------T--------C-------------------T-------A-GG---------------C-----T--C-------------------T--A--T-------------TC---A------A-----------G-T---..---------A------------A 6534
13_cpx.CM.96.1849             --GAG-----C---------------------G------TC------AAA------C--C-----T-------------------------------------A--AAT--A-GC--------------C---------..---G----CA-T--------G--- 5943
14_BG.DE.01.9196_01           --------------------------------G-----------------------------------------------------------------G---G-T-------------A-----------------..---------A------------- 6056
15_01B.TH.99.99TH_MU2079      --T-------------------------------------------------C-------G-----T-------------G------G-----------A-A------G---------------C---------..---G-------T----------- 5929
16_A2D.KR.97.97KR004          --GAG--T--C--------------------------------------A-C-------------------------------------------------A--AACC---A------A-----------------..----------G------------- 5899
17_BF.PE.02.PE02_PCR0155      ---M---T--------C---------------------G-----G-------------------------------------------T-------------TC---A------A-----------G-T---..---------A------------A 5719
18_cpx.CU.99.CU76             --GTA------------------T---------T-----AG---T--AAGC--------C-----T----------------------G-----------A--ACTC---G------A-------------T---..---------A------------- 5832
19_cpx.CU.99.CU7              --GAA-----C-------------------------------------------G--------C-----T-------------------------G---------A--AAT----A------------G-G---------..---------A------------A 5755
20_BG.CU.99.Cu103             ---AT---C---------------T--------------AG---T---AGC--------C-----T-------A-------T-----------------GA--CCTC---A------A-----------------..---------A------------- 6006
21_A2D.KE.91.KNH1254          ---A-G--T-----------------------------ACA-A-----C-------A-C-------------------------G-----------------A---A-C---A--------------C---------..---------A-------G----- 5734
22_01A1.CM.01.01CM_0001BBY    --GAG-----C--G--C-------------C------------C---A--G--------C-----T-------------------------T-G---------A---AG----CC-----A-----------------..---------A-A-------G--- 5739
23_BG.CU.03.CB118             --TAT---C---------------G--------------AG---T---AGC--------C-----T-------A-------T-----------------GA--CCTC--RA------A-----------------..---------C------------- 5982
24_BG.CU.03.CB378             --GAT---C---------------T---------T-----AG---T--AAGC--------C--------------G-----G---------------------A--CCTC---A------A-----------------..---------A---------G--- 5970
24_BG.CU.03.CB471             --TAT---C---------------T------------------T-G-T---AGC--------------T-------G-----G------------------C--GA--CCTC--CA------A-----------------..---------A---------G--- 5979
25_cpx.CM.01.101BA            --GAA----TY-------------T--------------AGCC-T--A-CC--------C-----T-------T-----------------G---------A---AGC-AAC------A-----------------..---------A-T----------- 5747
27_cpx.CD.97.97CDKTB49        --GAT-----C-----C-----C---------------C--ACG-----A--G-------C-----T----------------------T--------------A--A-TC---A------A-----------------..---------A------------- 5746
28_BF.BR.99.BREPM12609        ----------------------------------C---GT-A-A-------C-------A-------------------------------------T-T-----G-------------AA------C-----------------..---------A------------A 5895
29_BF.BR.02.BREPM119          ------------------------A--------------------------C--------G----------------------------G----------A--CC----AAG--------------------------..----------------------- 5700
31_BC.BR.02.110PA             ---AA--T------------------------------------------------G--------C-----T--------------------------GT---------C-G-AT--A-G------A-----------------..---G------------G--T--- 5984
32_06A1.EE.01.EE0369          --GAT--T-T-T------------T------------C--AG---T--AAGG--------C-----T-------T---------------------------A--CAT---AA------A-T---------------..----------C------------- 6174
33_01B.MY.05.05MYKL007_1      --GA------C-------------------C------C-C--G-----A--G--C-----C-----------------------------------------A--CACC--AA----A-A-----------------..---------A-A----C---G--- 5918
34_01B.TH.99.OUR2478P         --GAT-----C-------------------C------C----G-----A--G--C-----C--------T-------------------------------A--CACC-A-A------A-----------------..--------CA---------G--- 5735
35_AD.AF.05.05AF095           --GAG-----C------------------------------G-----AA-G--------C-----T-------------------------------------A--CCT----G------------G-----------..---------A-A-------G--- 5732
36_cpx.CM.00.00CMNYU830       ---AT-----CT------------------C--------------------A--G--------C-----T--------------------T-----------A---AT----C-----A----G-----------..---------A---------G--- 5748
37_cpx.CM.00.00CMNYU926       --GAT-----C---------------------C--------T-C----A--G--------C-----T---TTGC-TG--ACC-A--G-CC--A-...-----GA--A--AATG---GT---------A--G----------TGGAAA--T--CATGG-A-A-C-GAT- 5730
38_BF1.UY.05.UY05_4752        ----------------------------------------------------G----------------------------------T---G-T---------A----CC---G------A-----------G-T---..----G----CCA----------A 5953
39_BF.BR.03.03BRRJ103         -------T--------C-------------------------A---------------C-----T-----------------T-----T-----T-------------AGC--AA------A-----------G-T---..--------CA---------G--A 6029
40_BF.BR.04.04BRRJ115         ----------------------------------G------A--------G-------------------------------------T--T-----------A----------G-----------------------..----------------------- 6081
42_BF.LU.03.luBF_05_03        ---S------T-------------------------------C--------------A-----------------------------------T------------------------------------------..---------A------------A 6067
43_02G.SA.03.J11223           --RAY---C---------------TA-------------AG---T--AAA---------Y---------------------------------------------A-GACTC---A------A------CC---C---------..---------A------------- 6064
U.CD.83.83CD003               --GAA-----------------------------G-------G-A---ATCG--------C-----T-------------------------G---------GC----T----G------A-------C---------..-----G---A-A--------G--- 6021
U.CD.90.90CD121E12            --GA--------------------------------G------A-A---A-C--------C-----T-------------------------G---------T----T-------------A-------C---------..---------A------G--G--- 5722
U.GR.99.GR303                 --GAG-----C--------C-----------C------C--AG-G----AAA---------C-T---T----------C----------------G---------A-GA-T----G------A-------------T---..---------A---------G--- 5811
U.NL.95.U_NL_95_H10986_D1     ---AT------T------------T---------T-----AG--AT--AAG---------------T---------------------------------------T--CCCC-A---------A-----------------..---------A------------- 6421
CPZ.CD.90.ANT                 -G--AC-T-------------CA--A----CTCCATGACAAG---------G--C-----A-----A--TACCAGT-----G--A-T---T--AG-T--TATT-T---TAGG...C-C---ACCT---TCTGG-----TGCT...-AT------T-T---------AGT 5956
CPZ.CM.01.SIVcpzCAM13         TC-A---AGTA--C--C-----T---------------C--AGC-------CT-----CA-A----------C--A--------T--------T--T--T------------TC-TCAT-----A-----GG-C-----T---..-----T---C-G------------- 6111
CPZ.CM.05.SIVcpzEK505         --GAG--AGT------C-----T-------------G--CC--AG--------CT--C---A----------G--A--A--C--C-----T---------------------C--A-ACC--------------CG----G-T---..--------CA-T--------T--A 6082
CPZ.CM.05.SIVcpzLB7           --GAAG-A-----C--------G---------A------C-AA-GG------C---C---A-A------T-T--A--T-------------T--TG-T--TA---C-A---A-C-AAGC-----A-----------------T---..----C-----CA-----A-T--- 6098
CPZ.CM.05.SIVcpzMB66          --GA---A-----G------------------A------C-GA-CC----A--G--C-----A-----TT-T--A--T------T-----A----T------G-G--GA---A-C-AAC------A--T---T-C--C--TGCA..----G----CA-T--------G--A 6115
CPZ.CM.05.SIVcpzMT145         ---AA-----C--T--C-----C--C--------GAGT---AG------A-CT-----CA-C---------C--A--T--------T--T--T---CT--T-----G---C-C--ACCT-----C--T----G-G-----------..---G-----T----------T--A 6078
CPZ.GA.88.GAB1                -TGA-C-GGTA--C--------C-----C-----G---C--AG--------CT------A--------------G--A--------T-----T---GT--T--G------T-TC-TCC------A-T----TCA-------------..---------A-T-----G--C--A 6597
CPZ.TZ.00.TAN1                -CCAAC-A--CT-G--------C--T------G-TATTAC-AG--G---TAA---C--CA-A-----A------AA------G--TTT---T------T--TT-T------ACCC-A-CC------T--AT--GG-----T---...GAAG------T--------GC--G-- 6153
CPZ.US.85.CPZUS               T-GAG--A--CT----C-----C--T-----A--G-----CA-GCAG----CC--C--CA--------------G--A-----C--T--T---------T--------------CAT---CC----------T-----G---G-----..---G-----A-T-----C------A 6604
CPZ.CM.x.SIVcpzDP943          -GGAG--A--A--C----------G--------TC-----CT-A-C-A---A-CC--C---A---------T-G--A---------T------------T--------------GCATG-ACC----A-A--T----CC---G-----..---G-----A-------------A 6111
CPZ.CM.x.SIVcpzMB897          ---AA--A-----T---------------C-----------C-GATAG-T-----C---C--CA-A-----TT-T--A--T-----G--T-----T--T-----TG------A--ACG--A-A-------A------G-G---G-TGCA..------C---A-----------C--A 6104
N.CM.02.DJO0131               T-GAA--AGT---T--C--C--C----------------CC--AG---------CT--C--CA-C------------A--A-----C-----T--T-----T------------CC-C-A-AC---------T---CC------T---..---G-----A-A----C------A 6005
N.CM.04.04CM_1015_04          --GAG--AGT---T--C-----T----------------CC--AG---------CT--C--CA-C------------A--A-----T-----T--T-----T------------GCC---AA-C---------T---C-----G-T---..---------A-------C------A 6004
N.CM.04.04CM_1131_03          --GAT--AGT---T--C-----T----------------CC--AG---------CT--C--CA-C------------A--A-----T-----T--T-------------------CAT--AA-C---------T------A---G-T---..---------A-------C------A 6039
N.CM.95.YBF30                 --GAG--A-----T--C-----T----------------CC--AG---------CT--C--CA-C------------A--A-----T--T--T--T-----T------------GC----ACCC-----A--T-----A-----T---..---G-----A-A----C---C--A 6120
N.CM.97.YBF106                --GAG--AGT---T--C--C--T----------------CC--AG---------CT--C--CA-C------------A--A-----C-----T--T-----T------------GC----AACC---------T---T-------T---..---G-----A-A----C------A 6090
O.BE.87.ANT70                 ----AC-AGTA-----C-----T-----------CCT-ACAAGC--T--AAAG------A-------AT----A---------T--T----------CT---T-T---TATCC----CAC------------TG-C-----T--A...-----------T-------------A 6580
O.CM.91.MVP5180               --G-AC-AGTA-----C-----T-----------CCT-ACAAGC--T--ACAG------A-------AT----A-----C--T--T-----T-----T-----T---T-TCC-C-A-GC-----------T--C---G-T--A...-----------T-------G--C--A 6570
O.CM.96.96CMABB637            -----C-A--AT----C-----T---------C--CCTTAC-AGC--T---AAG------A-------AT----A-----C--C-------------TCA-----T--GTATCC-C-AA-C-----------T--A--C--T--A...---G-----T-----------T--A 6022
O.CM.98.98CMA104              ---A-C-AGTAT----C-----T-----------CCT-ACAAGCC-T---CA-------A-------AG----A---------T----------------T-----T---TATCC----CAC-----------T-GA---G-T--A...---G-----T-------G------ 6021
O.CM.99.99CMU4122             ---AAC-A--RT----C-----T-------T---CCT-ACAAGC--T---AAG------A-------AT----A---------T--------T----CT-----T---TATCC---AAA------------T--A--C--T--A...---G-----T-------------A 6020
O.FR.92.VAU                   ---AAC-A--AT-G--C-----T-----------CTTGACAAGC--T--ACAG------A-------A-----A---------T-------------GT---A-T---TAT-AGC-AAA------------T--A--C--T--A...---G-----T-T---------C--A 6104
O.SN.99.SEMP1299              -----C-AGTAT----C-----T-----------CCT-ACAAGC--T---CAG------A-------AT----A---------T-------------TCT---T-T---TATCCT--AACC--G--------T-----C--T--A...-----------T-------------A 6591
O.US.99.99USTWLA              ---G-C-T--A-----C-----T-------T---TTT-ACAAGC--T---CAG------A-------AT----A-----C--T--------T----GT---A-T--GTAT-AC--AAA------------TT-------T--A...-----------T----------C--A 6021
```

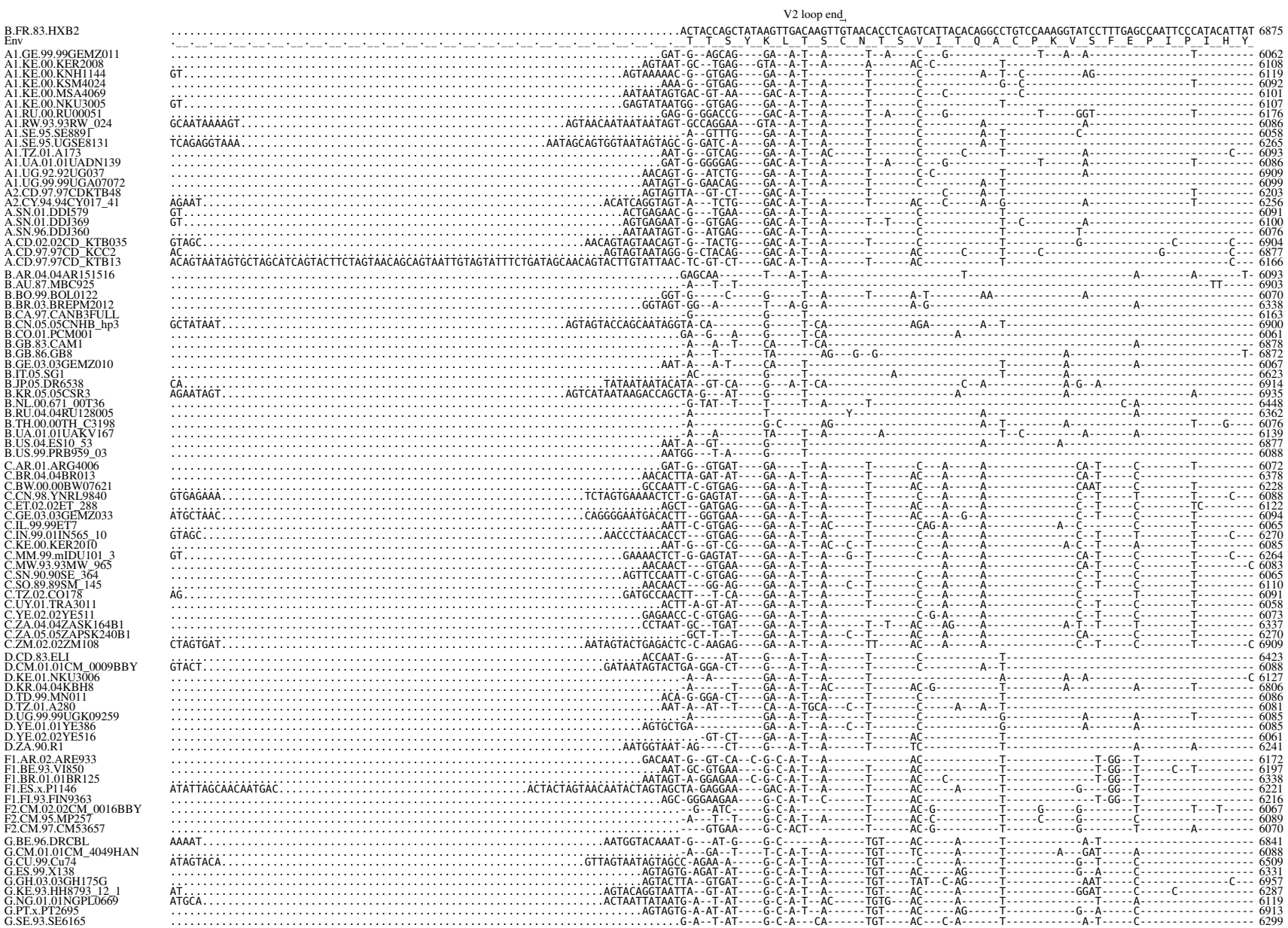
HIV-1/SIVcpz Complete Genomes

```
                                                                                                              V1 loop start
B.FR.83.HXB2               ATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACC...CCACTCTGTGTTAGTTTAAAGTGCACTGATTTGAAGAATGAT................................................................. 6635
Env                        _M__H__E__D__I__I__S__L__W__D__Q__S__L__K__P__C__V__K__L__T__.__P__L__C__V__S__L__K__C__T__D__L__K__N__D__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.
A1.GE.99.99GEMZ011         -----AACA------------C-C-----C-----------------------G------...--T-----C--C-C------T--T---A-CCCC--C-TCAC-................................................................. 5819
A1.KE.00.KER2008           ------ACA------------C-------C---------------------C-G------...--T-----C--C-C-C----T--T-G---CG-C-CC-CCA--...............................................................AA 5846
A1.KE.00.KNH1144           ------ACA------------C-------C---------C-------------G------...--T-----C----C------T--T--A---G-T-CT----T-................................................................. 5852
A1.KE.00.KSM4024           ------ACA------------C-G-----C-----------------------GC-----...--T-----C----C------T--T-GCA--A-AG-C---AGC................................................................. 5852
A1.KE.00.MSA4069           -------CA------------C-------C-----------------------G------...--T----------C---GG-T--TGTCAC--AC--TGGCACA................................................................. 5852
A1.KE.00.NKU3005           ------ACA--------------------C-----T---C-------------G------...--T----------C------C--T--CA--GYC--C--CTG-................................................................. 5837
A1.RU.00.RU00051           -----AACA--------T---C-------C---------C-------------G------...--T-----C----C------T--TCAA--GA-C-GC----TCACTAACTATAACAGT.................................................. 5951
A1.RW.93.93RW_024          ------ACA-----------CC--------G----------------------G------...--T-----C----C------T--T-GC---GCC-TTGGCA-C................................................................. 5837
A1.SE.95.SE8891            ------ACA------------C-------C-----------------------G------...--T-----C----C------T--T-GCTC-G-C-CC---AGC................................................................. 5830
A1.SE.95.UGSE8131          ------ACA------------C------------------------------C-G------...--T-----C----CC-----T--T-GCA-CAATGTC-C--CC................................................................. 5986
A1.TZ.01.A173              ------ACA------------C-------C------T----------------G------...--T-----C----C------T--T-GCA--G-C-GT---A--...............................................................AC 5837
A1.UA.01.01UADN139         -----AACA------------C-C-----C-----------A-----------GC-----...--T-----C----C------T--T---A-ACCC-C-CCCCCA................................................................. 5834
A1.UG.92.92UG037           ------ACA------------C-------C-----------A---------C-G------...--T-----C----C----G-T--T-GCT--AAC-TC-CCA--................................................................. 6660
A1.UG.99.99UGA07072        -------CA------C-T---C-------C-----------------------G------...--T-----C----C----G-A--TCAAA-C-ATG-AGGCACC................................................................. 5841
A2.CD.97.97CDKTB48         -------CA------------C--------------------------------------...--T-----C--C-C------T--T-GCA--GCC--C-CCAC-................................................................. 5978
A2.CY.94.94CY017_41        -----A--A-----------C---------------------------------------...--G-----C--C-T------T--T-GCA--GCC--T-CCAGC................................................................. 5995
A.SN.01.DDI579             ------ACA--C---------C-------C------T--------------C-G------...--T-----C----C-C----C--T--C--CA-CCCC--GA-C................................................................. 5842
A.SN.01.DDJ369             ------ACA------------C-------C------T----------------G------...--T-----C----C------C---CACA-CG-C-CC--GACC................................................................. 5845
A.SN.96.DDJ360             ------ACA------------C-------C------T----------------G------...--T-----C----C------C--T-GCA-CA-C--C-GCACC................................................................. 5842
A.CD.02.02CD_KTB035        --------A---------------------------T----------------G------...--T----------C----G-T---CACA--G-C--TGT--GC...............................................................AA 6624
A.CD.97.97CD_KCC2          --T-----T---------C--C---------GC---T--C-----C--------------...--T-----C----C----G-C---CAGCCAG-T--C-GCACC................................................................. 6634
A.CD.97.97CD_KTB13         --------A-----------CC--------------T----A---------C-G--G---...--T----------C----C-T---CA-T--G-C-TC--CA-C................................................................. 5845
B.AR.04.04AR151516         ------------G----------------------------------------------A...-------------C---G--T----------G---A-G-A--................................................................. 5850
B.AU.87.MBC925             --------------------------------------------------G---------...-------------C------T--------C-ACTCA--GA--.........................................................GCTACTAA 6650
B.BO.99.BOL0122            ------------------------------------------------------C-----...-----T-----C-C------C--T---A--G-A-CC-TAAG-................................................................. 5854
B.BR.03.BREPM2012          --------------------------------------------------G---------...---A---------T----C-T------A---AT-CT---AC-AGTGCTAATACCTTTAGTAGTCCT......................................... 6095
B.CA.97.CANB3FULL          -------------------A----------------------------------------...-------------C------T-------------G-----G-................................................................. 5923
B.CN.05.05CNHB_hp3         --------A------------------------G------------------G-------...-------------C------T-----------C----G-A--AGTACCAGTAGT..................................................... 6648
B.CO.01.PCM001             --------A--------T------------G-------------------C---------...-----A-----C-C------T-----GA--GCT--T-CCAC-................................................................. 5833
B.GB.83.CAM1               ------------------------------------------------------------...-------------C-----TT------A--G-A--T---AC-................................................................. 6632
B.GB.86.GB8                ------------------------------------------------------------...-------------C------T-------------G-------................................................................. 6644
B.GE.03.03GEMZ010          ------------------------------------T-----------------C-----...---------A-C-C-C----T--------------C---AG-................................................................. 5836
B.IT.05.SG1                -----G------------------------------------------------------...------------GC------T-----------A.......................................................................... 6431
B.JP.05.DR6538             -----------------T--CA--------------------------------------...-------------C------T---------GAA-TTCTCA--................................................................. 6641
B.KR.05.05CSR3             --------A--------------------------------------------------T...-------------C-------------A--AATGT-------................................................................. 6656
B.NL.00.671_00T36          -----------C------------------G----------A--------G---------...-------------C------T---------GCC--T-T-AC-................................................................. 6196
B.RU.04.04RU128005         --------A----------------T----------------------------------...-----T-------CC-----T--------------A--AG--................................................................. 6122
B.TH.00.00TH_C3198         --------A---------G-C---------------------------------------...-------------C----C-T------A-------TGC-ACC................................................................. 5845
B.UA.01.01UAKV167          ------------------------------G-----T----------C------A--G--...-----------C-C-C----T------A------G---GA--................................................GTCACTAATACCAATGG 5886
B.US.04.ES10_53            -----G--A--------------------------------------------------T...-------------C------T----------G-G-T---C-ACT............................................................... 6658
B.US.99.PRB959_03          -----------------T-----------C------------------------------...-----------C-C------T------A--GC-GGA----T-................................................................. 5848
C.AR.01.ARG4006            ------C----------------------------------------------G--G---...-------------CC-----T-----G---GCT--C-GAAC-................................................................. 5823
C.BR.04.04BR013            ------C-----G----------------------------------------G--G---...-------------C------C--T---AC-GCT--C-GAAC-................................................................. 6117
C.BW.00.00BW07621          -----------------------------------------A-----------G--G---...-----------C-C------C--T-T-A--G-T-CT----C-................................................................. 6006
C.CN.98.YNRL9840           ------------G----------------------------------------G--G---...-----T-----C-C----G-A--T-GAA--G-T-GA----T-................................................................. 5818
C.ET.02.02ET_288           ------C-----------------------G----------------------G--G---...-------------C----C-C--T-G-TC-AATGTT-CC.........................................................AGTAATTCTAA 5863
C.GE.03.03GEMZ033          --------------------C---------G---------------C-----G--G---...-----------C-C------T-----CA--GCT--A-T-A--................................................................. 5830
C.IL.99.99ET7              -----------------------------------------------------G------...-----------C-C------C--TGG---GG-T--TGT-AC-................................................................. 5828
C.IN.99.01IN565_10         ------------G----T-----------------------------------G------..A-----------C-C----G-A--T-GAA--G-T-CCTTG-T-................................................................. 6015
C.KE.00.KER2010            -----------------------------------------------------G--G---...-----------C-C------T--T-GCA--G-T-CC-G-A--...............................................................GC 5826
C.MM.99.mIDU101_3          ------------G--G-------------------------------------G--G---...-----------C-C----G-A--T-AAA--G-T--A-T--T-.........................................................AACAGTAC 5993
C.MW.93.93MW_965           --------A------------C----------------C--A-----------G--G---...-----------C-C------C--T---A--GCT--TGG-AC-................................................................. 5837
C.SN.90.90SE_364           -----------------------------------------------------G--G---...-----------C-C------C--T-GGA--G-T-CC-CCAGC................................................................. 5822
C.SO.89.89SM_145           -----------------------------------T-----------------G--G---...-----------C-C------T--T--A--AGCT--T-T-A--................................................................. 5864
C.TZ.02.CO178              --------------------------------------T--------------GC-G---...-----------C-C------T--T-G-A-GG-T-GA-T-A--...............................................................GC 5820
C.UY.01.TRA3011            -----------------------------------------------------G--G---...-------------CG-----C----G---GCT---C-G-AC-................................................................. 5824
C.YE.02.02YE511            -----G-----------A------------------T----------------G--G---...-----------C-C------C----A-A--AAT-TT-CCA--................................................................. 5827
C.ZA.04.04ZASK164B1        --------A-------C-------------G--G----------------------G---...-----------C-C-C--G-T--TGT----G-T---GT-A--................................................GTTACCAAGAATAATAC 6075
C.ZA.05.05ZAPSK240B1       ---------------------C-G-----------------A-----------G--G---...-----------C-C------C--T--AA-GG-T-CC---A--................................................................. 6057
C.ZM.02.02ZM108            ------------G----------------------------------------G--G---...-----------C-C-C----A--T-GC---G-T-CCT-CA--............................................................ACCAG 6626
D.CD.83.ELI                -----------------------------------------A------------------...-----------C-C------C--T-G----GAATT--GGA-C................................................................. 6177
D.CM.01.01CM_0009BBY       -----------------------------A-----------A------------------...-----A-----C-C------C---------GCT--AGTGAGC................................................................. 5833
D.KE.01.NKU3006            ------A----------------------------------A------------------...-----------C-C------C--T------G---------AACTATCTTCAAAACTTTCAAT............................GAGACCATGAAAAATGA 5874
D.KR.04.04KBH8             -----------------------------G-----------A------------------...-----------C-C------C--T------A-CCG----A-G................................................................. 6587
D.TD.99.MN011              -----------------------------------------A------------------...--G----------C---G--C---------G-C-GC--CA--................................................................. 5858
D.TZ.01.A280               --------------------C--------GG----------A------------------...-----------C-C----C-C---T----CGC---C-CGAC-................................................................. 5832
D.UG.99.99UGK09259         -----------------------------------T-----A------------------...---T-------C-C------C--------A-G-G-----A-G................................................................. 5836
D.YE.01.01YE386            -----------------------------------T-----A------------------...-----------C-C------C----T---A-G--CTGGATGG................................................................. 5830
D.YE.02.02YE516            ---------------------C-------------------A------------------...--G--------C-C----GC---T------A-T-------T-................................................................. 5845
D.ZA.90.R1                 -----------------------------------------A------------------...--G--------C-C------C---------GCC-TCTT-A-CTCCACCTCCTTCAGCAGC............................................... 5998
F1.AR.02.ARE933            ------ACA-----------------------------------------G--G------...--------C-N--C-----NT--T------GCC-TC-GGA-C................................................................. 5935
F1.BE.93.VI850             ------ACA--------------------------------------------G------...-----------G-C------C--T---A--GCC-CC---A-C................................................................. 5981
F1.BR.01.01BR125           ------ACA-----------------------------------------G--G------...-----T-------C------T--T--CA--GCC--TGG-AC-................................................................. 6083
F1.ES.x.P1146              --------A------------C-------------------------------G------...-----------C-C------T--T---A--GCC---GGCA--TTCACTAGCCAAAGCAGCACT............................................ 5945
F1.FI.93.FIN9363           ------A-A--------------------------T-----------------GC-----...-------------C------T--T---A--GCC-CT-CCACC................................................................. 5967
F2.CM.02.02CM_0016BBY      --------A--------------------------------A-----------GA-----...-------------CC---C-C--T-G----G-T--T-T-ACG................................................................. 5824
F2.CM.95.MP257             --------A---------------------G-G--------------------G------...-------------C------T--T---A-GGCT-TC-T-A--................................................................. 5843
F2.CM.97.CM53657           --------A------------------------------------------C-GA-----...-----------A-CC-----T--T------G-TCCTGT-A-C................................................................. 5824
G.BE.96.DRCBL              ------------------------------G----------------------GC-----...--T----------C------C--T-----GA-A--C---A--................................................................. 6595
G.CM.01.01CM_4049HAN       ------C----------T------------G----------------------GC-----...--T----------CGC----C--T---A--G-C-CC-CCACC................................................................. 5842
G.CU.99.Cu74               -----------------T------------G----T-----------------GC-----...--T----------C------C--TGT-A--G-A-GTTGCACC................................................................. 6245
G.ES.99.X138               ------------------------------G----------------------GC-----...--T----------C-----CC--T---A--G-A-CC--CA--................................................................. 6082
G.GH.03.03GH175G           --------A------------C--------G----------------------GC-----...--T--------C-C------C--T-----CA-TGTCT--TC-......................................................AATTGTACTAG 6692
G.KE.93.HH8793_12_1        ------------------------------G----------------------GC-----...--T----------C------C--T------GCA--TGTAACC................................................................. 6041
G.NG.01.01NGPL0669         -----A------------------------G----------------------GC-----...--T--T-------C------C--TGTG---G--G-TGTCA--................................................................. 5846
G.PT.x.PT2695              ------------------------------G----------------------GC-----...--T--------C-C-----CC--T---A--G-A-CT-T-A--GTAACCAGCAGTAGCAACAGTAGTGTAACCAGCACTAGCACTGTGGGTAGCACTACCAGCACTAG 6648
G.SE.93.SE6165             ------------------------------G----------------------GC-----...--T----------CC-----C--T------G-A-CC--CA-A......................................................GGCAACAAAAG 6049
H.BE.93.VI991              ------ACA-----------------------G---T----------------G------...-------------C-C--G-C--T-G-AG-G-A--TGCAACA................................................................. 6013
H.BE.93.VI997              -----GACA--C------------------------T-----------------------...-------------C-C--G-C--T-G-A--A-A-CT-GAA--................................................................. 5940
H.CF.90.056                ------ACA--C------------------------T-G--A------------------...-------------C-C----C--T---A--G-C-GA--CA--................................................................. 5975
J.CD.97.J_97DC_KTB147      --------T------------G--------G-----T----A------------A-----...--T----------CC-----T--T-----GGCA-CA-TCA-CAATAGTACTGATAACAGTAAT............................................ 5856
J.SE.93.SE7887             -----G--A------------G--------G-----T----------------GA-----...--T--A-----C-C------T--T-G-A-CA-C-CT-G-A-C................................................................. 5954
K.CD.97.EQTB11C            ------ACA---------------------G----------A-----------G------...--T--------C-C-C---CT--T--CA--G-A-CT---A-C................................................................. 5830
K.CM.96.MP535              ------ACA---------------------G----------A--T------G-G------...--T--------C-C------T--T-------AC--AGGAACC................................................................. 5834
```

```
                                                                                                                                                   V1 loop start
B.FR.83.HXB2                 ATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACC...CCACTCTGTGTTAGTTTAAAGTGCACTGATTTGAAGAATGAT................................................................ 6635
Env                          _M__H__E__D__I__I__S__L__W__D__Q__S__L__K__P__C__V__K__L__T__.__P__L__C__V__S__L__K__C__T__D__L__K__N__D__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__._.
01_AE.CF.90.90CF11697        -----A------G-----------------------T----------------G-----T...--T--T--C--C-C----C-T--T--CA-GGCT--ACTGA--................................................................ 6584
01_AE.CN.05.FJ051            -----G-----CG-----------------------T----------------G-----T...--T--T-------C------T--T--AA--GCTGCTGTCA--................................................................ 6638
01_AE.CN.06.FJ054            -----A------G-----------------------T-----A----------G---------T-----C----CCC----T--T--CACGGCT--TTTGCC-................................................................ 6659
01_AE.HK.x.HK001             -----G------G----------------------------------------G-----T...--T----------C------C--T--CC-GGCT--TTGGA-A................................................................ 6005
01_AE.TH.01.01TH_R2184       -----G------G-----------------------T--------T-------G-----T...--T-----C----C------T--T--C---GCT---TGGACC................................................................ 5835
01_AE.TH.02.OUR769I          -----G------G-----------------------T----------------G-----T...--T----------C------T--T--CA--GCC--TTTGACC................................................................ 5834
01_AE.TH.90.CM240            -----G------G-----------------------T----------------G-----T...--T-----C----C------T--T--CA--GCT--TTTGACC................................................................ 6203
01_AE.US.00.00US_MSC1164     -----C------G-----------------------T----------------G-----T...--T----------C------T--T--CAC-GCT--TTTGACC................................................................ 5840
02_AG.CM.02.02CM_4082STN     -------TA--------T---C-------C----G------------------G---------T----------C----C-T------A-C--C--C-GATC-................................................................ 5837
02_AG.EC.x.ECU41             --------A--------T----------------------------------G----------T-----C----CA-----T--TCAAA-CG-C-CC--CATA................................................................ 5733
02_AG.GH.03.GHNJ196          --------A--------T---C-------C-------T----A---------G-G--------T-----C----C----G----TCA-AG-G-C-CC--CAGC................................................................ 6667
02_AG.NG.01.PL0710           --------A--------T---C-------C-----------------------GC--------T-----C----C------T--T--A--CCGCCTC--CA-C................................................................ 5822
02_AG.NG.x.IBNG              --------A--------T---C-------C-----------------------G---------T-----C----C----G-T--TCA-A----C--C--CAGC................................................................ 6162
02_AG.SE.94.SE7812           --------A--------T---C-----------------------------C-G---------T----------C----C-T--TCAG---AACCTC-C-AGC................................................................ 6010
02_AG.SN.98.MP1211           -------TA--------T---CC-------------T----------------G---------T----------C----G---TCA-A-C-AT--CT--ACC................................................................ 5843
03_AB.RU.97.KAL153_2         -----------------------------------------------------G---------------------C------T-----------------G--G................................................................ 5848
04_cpx.CY.94.CY032           ----------------------------A--G--G-------A------C---GC-----...T-T----------C---T-CA--T-T-A--GCA-CT-C-AC-................................................................ 6005
05_DF.BE.x.VI1310            ------ACA--------T-------------------T---------------G------...-------------C-C----C--T--------C---GCCA--................................................................ 6023
06_cpx.AU.96.BFP90           -------------------------------G----T----------------GC--------T-----C----C-----CA--T---A--GCT-CATTA-G-...............................................AACAAAACATTAGGTAA 6688
08_BC.CN.97.97CNGX_6F        ------------G---C------------------------------------G--G---...-----------C-C----G-A--T-GAA--G-T-GC-G-A--................................................................ 5808
09_cpx.GH.96.96GH2911        ------A-A--C---------C---------------T---------------G---------T-----C--C-CG---G-T---CA-AGAG-C--C--CAGC................................................................ 5846
10_CD.TZ.96.96TZ_BF061       ----------------------------------G-------A------A----------...-----------C-C------C----A--GCA-T--TGCCAC-................................................................ 6032
11_cpx.GR.x.GR17             ------------G-------C----------G-----T----A----------GC--------T---------GC------T--T-----GCA.......................................................................... 5929
12_BF.AR.99.ARMA159          ------ACA-----------------------------------------G--G------...-------------C------T--T-G-A--GCT--TGCCAC-................................................................ 6636
13_cpx.CM.96.1849            ------ACA--------T---A------------------------------GT---------T-----C----CC-----T--T-A-A-CA-C-CT-G----................................................................ 6045
14_BG.DE.01.9196_01          ---------------------------------------------------G--------...-------------C---G-T------A-------TG-GAC-................................................................ 6158
15_01B.TH.99.99TH_MU2079     -----G--A------------C--------------------------------------...--------------------T------------G--------................................................................ 6031
16_A2D.KR.97.97KR004         --------A------------C--------------------------------------...--T--------C-C------T--T-GCAGGG-C--A---ACC................................................................ 6001
17_BF.PE.02.PE02_PCR0155     ------ACA---------------------------T--------------G--------...-------------C------T--T---A--GCC--TGCCACC...............................................AAGAGTGCCAGTGCCAA 5838
18_cpx.CU.99.CU76            --------A--C-------------------G----------A----------GC--------T----------C------C--T---A--G-A-MTTR-AG-................................................................ 5934
19_cpx.CU.99.CU7             -------CA------------C-------C-----------------------G---------T-----C----C----G-T---.................................................................................. 5839
20_BG.CU.99.Cu103            ---A-------C-----T-------------G----T----------------GC--------T----------C------C--T---A----ATCAGTAA--................................................................ 6108
21_A2D.KE.91.KNH1254         -------------------------------G----------A-----------------...-----------C-C------T--TG---GGAGC-CC-C-TCA..............................................................TC 5838
22_01A1.CM.01.01CM_0001BBY   -----G-----------T-----------------------------------G-----T...--T-----C----C---------TGAA---GCT-G-TTCA--................................................................ 5841
23_BG.CU.03.CB118            -----------C-----T-------------G----T----------------GC--------------C-C----C--T---A-GG-AG-CTG---G................................................................ 6084
24_BG.CU.03.CB378            -----------C-----T-------------G----T----------------GC--------T----------C------C--T---AC---A--C-G-A--................................................................ 6072
24_BG.CU.03.CB471            ---A-------C-----T-------------G----T----------------GC--------T----------C------C--T---AG-G-AG-CTG-A--................................................................ 6081
25_cpx.CM.01.101BA           -------------------------------G----------A--------C-GC--------T--T-------C------C--TGT-AC---A--TTG-AG-......................................AATAGAATCAACAATAGTAACTGTAG 5875
27_cpx.CD.97.97CDKTB49       -------------------------------G----T----------------GC----T...-----A-------C-C----T--T---A--GCA--TGCAACA................................................................ 5848
28_BF.BR.99.BREPM12609       -------------------------------G-G---A----------------------...-------------C------T------A--AATGT--C--T-................................................................ 5997
29_BF.BR.02.BREPM119         ------C-A--------T----------------------C-------------------...-------------C------T--T-G------AGG----A--................................................................ 5802
31_BC.BR.02.110PA            ------C-------------C----A---------------------------G--G---...-------------CC-----C----G----GCT-TC-G-A--...............................................GCTAACAAAACTAATGC 6103
32_06A1.EE.01.EE0369         ----------------AA-------------G---------------------GC--------T----------C----C-C--T-AGA--GCA-CCTT-A-G..............................................................AA 6278
33_01B.MY.05.05MYKL007_1     -----G------G-----------------------T--------------C-G-----T...--T-----C----T-C--C-T--T--CA--GCT--TTTGACC................................................................ 6020
34_01B.TH.99.OUR2478P        -----G------G-G---------------------T----------------G-----T...--T-----C----C------T--T--CA--GCTGCTTTGACCAATGTAACTATA.................................................... 5849
35_AD.AF.05.05AF095          ------ACA------------C-------C-----------------------G---------T-----C----C------A--T--CA-CGCA-CC-CCA-CATA............................................................. 5837
36_cpx.CM.00.00CMNYU830      -----G-----------T-----------------------------------G---------T-----C--A-C-C--G---TCAGA-C--T--CGGCAGCAACAACCAC........................................................ 5859
37_cpx.CM.00.00CMNYU926      CA-ATA--TATA--C-GTCTA-GGG-ACCAA-GTTA...----------A---G------CCCT-T--GC--CACTT.-----TA-T-GCC--AA-GTC--CCCCCAAGACCAACGGGAG................................................. 5846
38_BF1.UY.05.UY05_4752       ------ACA---------------------------------A-------G--G------...-------------CC---C-T--T---A-CC-T-CC-CCA-G................................................................ 6055
39_BF.BR.03.03BRRJ103        ------ACA---------------------------T-------------------G---...--C----------C------T--T-G-A--G-T--TGT-AC-................................................................ 6131
40_BF.BR.04.04BRRJ115        -----------------------------------------------------G------...-----T-------C----------G---CC--C--TGGCAC-................................................................ 6183
42_BF.LU.03.luBF_05_03       --------------------------------------------------G--G------...-------------CC-----T--T---A--NNNNNNNNNNNNNNNNNNNNNNNNNN.................................................. 6184
43_02G.SA.03.J11223          --------T----------------------G----T---------------TGC--------T----------CC----CC--T---A--GCC-CTTG-A--................................................................ 6166
U.CD.83.83CD003              ---------------------------------------T-------------G------...TTT--T-----C-C------C--T-T---G-A------AG--................................................................ 6123
U.CD.90.90CD121E12           -----------------T---------------------T-------------G--T-T-...T-T--T-----C-C------C--T------G-AGG--C-A-C................................................................ 5824
U.GR.99.GR303                ------ACA------------C-------C------------A----------G---------T-----C--C-C------T--T-AAA-CCCCTCTTG--T-................................................................ 5913
U.NL.95.U_NL_95_H10986_D1    G------------------------------G-------G--A-------T--GC--------T--------C-C------T--T-----CG-A-CA---AGA................................................................ 6523
CPZ.CD.90.ANT                ---ACA--A-----GN-ACAA---TCC-A--------AT---------------C----A...--TA-G---A-A-AAA-G--T--T--A-GA-AC--TGGAACA.....................................................CCTACAACACC 6069
CPZ.CM.01.SIVcpzCAM13        -----A--A--C--C-GT---C-C------C--G--T--T--A----------GC----T...--TT-A----A-CAA-G--A--T-G-A-CG-T-CT-GAC-A...............................................AATAGCACTAGCAACTC 6230
CPZ.CM.05.SIVcpzEK505        -----A--A--C--T--T--C--G------A--G--TT----A--C----T--G------...---T-A-----A-CCC--TCA--T-GCTCA-G--G--G--TG................................................................ 6184
CPZ.CM.05.SIVcpzLB7          -----AC-----G----T--C---------------------A------A---GC-----...--T--T-------C----C-T----G-ATCCCT--ATT----................................................................ 6200
CPZ.CM.05.SIVcpzMB66         --------------G---A-C-T-----C--G--TG------A---------------TT...--TT-G-------C------C--TT-ACTG--T--ATGCAT-...................................AAAGAAAATGGCAACACCACCAACTGCAC 6246
CPZ.CM.05.SIVcpzMT145        -----G--A--C----A---C-C-----A-------T--T--A----------------A...--CT-A-----G-CCC-T-CT--T-A-A--CCC-CT---AC-................................................................ 6180
CPZ.GA.88.GAB1               --------A--C--T-----C-C----------G--T--C-----T-------G--G---...--TT-A-----A-CC--GC------G-A-GGCT--CTT-AGC................................................................ 6699
CPZ.TZ.00.TAN1               ---A-A--A------T-ATCAC-T-TTC--A--G--TT-T-----T-------------A...---T-T--CA-A-AGA-G-CA--T--AATGACT--T-CCACA.....................................................AATAAAACCCT 6266
CPZ.US.85.CPZUS              -----G--A-----T--T--CC--------C--G---T----A--T---A---G------...---T-G-------C-A-G-CT--TCT-A-CCCCG-T-G-A--................................................................ 6706
CPZ.CM.x.SIVcpzDP943         -----GACT-----T--T---C--------------T-----A-------T--G--G--A...-----A-----C-C-A-GC-T--T-T-CCA-AC--C-TAAGC................................................................ 6213
CPZ.CM.x.SIVcpzMB897         -----A--A--C-C--T--CC---------------------------------C-G---...--C--------G-CC-----C--TT-CACAAAT-CA-C-ACA................................................................ 6206
N.CM.02.DJO0131              G----A-----C--T--T-----------A--G---T-------C-----T---------...---T-A-----A-C-A-G--T--T-GCA--AGC--TGGAA--................................................................ 6107
N.CM.04.04CM_1015_04         -----A-----C--T--------------A--G---T-------C-----TM--------...---T-A-----A-C-A-GC-T--TWACA--AGC--TGGGKT-................................................................ 6106
N.CM.04.04CM_1131_03         -----A-----C--T--T-----------A--G---T----A--C-----T---------...---T-A-----A-C-A-GC-T--T-AC---AGC--TGGGTT-................................................................ 6141
N.CM.95.YBF30                -----A--------T------C-G-----A--G---T-------C-----T---------...---T-A-----A-C-A-GCTT--T-AC---AGCT-TGGG--G................................................................ 6222
N.CM.97.YBF106               -----A-----C--T--------------A--G---T-------C-----T---------...---T-A-----A-C-A-GCTT--T-ACA--AGC--TGGGA--................................................................ 6192
O.BE.87.ANT70                -----G--A--C--T--T------------C--G--T--T--A--T----TC--A-G--T...TTC--G-----ACAAA-GG----T--AA-CA-AGCTGGAACA................................................................ 6682
O.CM.91.MVP5180              --------A--C--C--T-----G------A--G--TT-------T----AG--A-G--T...TTCT-A-----ACAAA-G--C--TGTA---C--C-A-CAA--................................................................ 6672
O.CM.96.96CMABB637           -----G--A--C--T--T------------G--G--T-----A--T----TC--A-G--T...TTC--G-----ACAAA-G--T--T--A---G-C--TGCAACC..............................................................AC 6126
O.CM.98.98CMA104             -----A-----C--T-CTGA----------G--G--TT----A--T-----G--A-G--T...TTCA-G-----ACAAA-GG-A--T-A---CA-A--T-CAACA................................................................ 6123
O.CM.99.99CMU4122            -----G--A--C--T--T------------A--G--TT----A--T----TC--A-G--T...TTC--G-----ACAAA-G--T-----AA-CG-A--CG-CACA................................................................ 6122
O.FR.92.VAU                  -----G--A--C--T--AGA---G------C--G--TT----A--T----TC--A-G--T...TTCT-G-----ACAAA-G--T--T--A---A-C--A---AG-................................................................ 6206
O.SN.99.SEMP1299             -----G--A--C--T--T------------A--G---T----A--T----TC--A-G--T...TTC--G-----ACAAA-G--T--T--AA-C-ATGTAC-A-GG................................................................ 6693
O.US.99.99USTWLA             --------A--C--C--T---------------G--TT----A--T----GC--A-G--T...TTCT-A-----GCAAA-G--C--T--A---G-A--AG-CA--................................................................ 6123
```

```
                                                              V1 loop end┐      ┌V2 loop start
B.FR.83.HXB2          .............ACTAATACCAATAGTAGTAGCGGGAGAATGATAATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAGGTGCAGAAAGAATATGCATTTTTTTATAAACTTGATATAATACCAATAGATAATGAT....... 6785
Env                   __.__.__.__.__T__N__T__N__S__S__S__G__R__M__I__M__E__K__G__E__I__K__N__C__S__F__N__I__S__T__S__I__R__G__K__V__Q__K__E__Y__A__F__F__Y__K__L__D__I__I__P__I__D__N__D__.__.__
A1.GE.99.99GEMZ011    ..............-GCTC--A---C-----CGTTAAC----ACGAT-GCAT-GTT-A---A--G-G--------------C--G-C----GAAC------A---AAG-A-A-CT-T----T--C---------G----------G--G---CT-G--T-A--....... 5969
A1.KE.00.KER2008      CACCAAGGTCAAT--C-CC-GTGCCCC--CC-C--CTGCC-CTG-C-ACA-TG-CA-T--C--G--------T----A------G-C----GAAC----G-A----AAA------ATG---T--C-C-------G-------G--GA--T----A----GA--AAT.... 6012
A1.KE.00.KNH1144      ....ACAGATGTT-G-GG---G-GGG-C-AC-T-ACC-TC----A-GA-AT-G-G-----A----------T-----------G-C----GAA-----G-A----AAA-------T----T--C-C-------G-------G--G--------A--C-G-G-AATAGTA 6018
A1.KE.00.KSM4024      ...................--G---GTC-AC-T-ACA---GAAGAGGAA--A-T-AA-A-CT-------T-------------G-C----GAAC----G-A----AAGA-------T-A--T--C---------G----------G--AA---TA--G-AAG-AGT.... 5999
A1.KE.00.MSA4069      ......................-ATGCC-CAG-G--A-T-GAC-AC-GC-CAAT--GG-A---AG----------A---------G-G-C----GAG--T--G-A----AGAA----G---C--T--C---------G-G-----T--G---A---T---G-AATGCAAGGT. 6002
A1.KE.00.NKU3005      .ACCAATGCCAAC-ACTG----C--GCC-AC---A-C-TC-C--ATGACAT-GC------A---------------------G-C----GAAC----G-A----AAA-------T----T--C----C-----------------G---A---TA--GGAA--AATGGTA 6006
A1.RU.00.RU00051      ..................................---AT-TCT-AC-GC-GCAG-T-T-----A--G-------------------G-C----GAAT------A--G-AAAAGA--G-T-C--T--C------------------G--T---CTAG--G-A--....... 6083
A1.RW.93.93RW_024     ...................................T-T-C--ATATC--G-A---A----C----------------G-G-C----GAAC----G-A----ACAA-------T-C--T--C---------G---------G-----C-TA--G-AT--AGTAGTA 5967
A1.SE.95.SE8891       ..........................TC-GTC-CTAAT-TC-CC-GTGACAT-GC------A---------------------G-C----GAA-----G-A----AG--------T-C-----C-C-------G---G---G-G-------G--C--CA--....... 5968
A1.SE.95.UGSE8131     ..............-AC-CC-A--GC-CG---GC-AACCTC-C-GAT-GC-T---------A--G-G-----------------T-C----GAAC----G-A----AAAA------T---CT--C---------------------G--AA---TA----AA--AAAAGCT 6143
A1.TZ.01.A173         TGACAGTGACAAC-G-G-C-ATGCC-CC-A-TC-ACC-TC-CC-GGGAAAT--TT-----A---------------------A-C----GAA-----G-AC---AAA-------T----T--C------------------G--G-GGA---TAG----AG-....... 6000
A1.UA.01.01UADN139    ....ACTTTAAATTG--C-G-ATCC-ACT-C-CTATCTCT-AC-GC-GCTT-T-T-A---A--G-G--------------C--G-CA---GAAC------A---A-A-A-A----T-C--T--C---------------------G--T---CT-G--G-A--....... 5993
A1.UG.92.92UG037      ..............-A--TC--------C-TC-C-AAT--CTCAG-T-ACAT--G--A---A---------------------G-C----GAAT----G-A----AAT-G---G-T----T--C------------------G--G---A---TA------G-AAT.... 6813
A1.UG.99.99UGA07072   .......AATTTA-------GT--C-CC-C-GT-AAC-TC-CCGAGGACAT---------A---------------------G-C----GAAT----G-A----AG-------ATGC--T-GC---------G---G------G---A---TA--GGCA-CAGAAGC. 6003
A2.CD.97.97CDKTB48    ...................................-AC-CC-ATAGC-CT-A---A--------------A------GCC----GAAC---A--A----AACA-------TC---T--C-G------G--------G--G---T-C-TA---GAAG-AAGAAT. 6107
A2.CY.94.94CY017_41   ...............-CCCAT-GC-A---C--TA-C-CCCA--GCCCCATT--T-A---A--------------A-----CT-C-----TAC------A----ACA--A---TT---T--C-G-------G-------G--G---A-C-T---G-AAG-GAAAATA 6149
A.SN.01.DDI579        ...................TT--CC-A--C--A--ATG-C-CC-A--CCT-T-A------G-----------G--------A-C----GAAC------A----AAA-----AT---T--C---------G-------G--G-------T--G----G-AGTAACA 5990
A.SN.01.DDJ369        ...............---T-----TCC-A-TC-A-TGTC-AA-ATGAAAT-CCT-----A-------------G-A----G-A-C----GCA-------A----AG------T-T----T-GC---------G-G-----G--G-------TA--G-AA--AGTAATA 5999
A.SN.96.DDJ360        ..................................A--TATGAC-CC-A---C--G-A------G-----------G---------A-C----GAAC------A----AGA----GT-T----T--C------------------G--G-------TA-----A--AGTAGT. 5977
A.CD.02.02CD_KTB035   TGTTGATATCAGC-GCC---T---C-AC--CTAT-TC-ATG-C-GTG-C--AG-GCAC-CA--G-C----------------G-G-C----GCA-------A---ACAA-------T-C--T-TC-C---------GG-----T--G---A-T---GGG-A---AAGGGTA 6794
A.CD.97.97CD_KCC2    .................................-A-AAT-CT-A-G--GAA-T-CCG-----A--G-C---------------------G-C----GAAC----T-A----AAA-------T-CGGT-TC---------GGA-----T--G---AG--T-G----A--ACAAATG 6773
A.CD.97.97CD_KTB13    ...................--AGCCCAC-AGGTTTCTGTTTCTCC-GAAAT-C---A-A-AG--GC----------------CG-G-C----GAAC------AC-GAAAA-------T----T-TC---------GGGAA---T--G----T--T---G-G---AAAAATA 5996

B.AR.04.04AR151516    .......GCTACT-A--G-----GA-TC-AC-ATA-T--TT--GG--C-AT-G-G---C-A------------------------T-T----G---GA-AA---A----GA------Y-----C-C---A---GC------G--G---M---------G.......... 6003
B.AU.87.MBC925        TACCACTAGTAGT-GCG-G--AGTG-TGGAG-AG--AGA---AAG-G--------G-----A------------------------C-----A-----G-------------------------C-----------------G--G-----G-------G--C....... 6813
B.BO.99.BOL0122       .....................................AT-ATAGT-C------T--G------------------------CTG--G-A----AG-A------AA-------------------C---------T------G--G----C------G-----AAT.... 5977
B.BR.03.BREPM2012     ....................CT--CC-A-GATA-C--GTG-GA-GATAT---G-A---A------------------------C---G------A--A------A-A------C-----C-----A---G-------G--G---------------CAGAGAAT. 6242
B.CA.97.CANB3FULL     ..............---------C--A-------A--GA----GG------G------A-------------------G----C-----A--G-----A---A--T--------------------C------------------G--G-----------------....... 6073
B.CN.05.05CNHB_hp3    ..................................-ATAAC--T-GT-C--CA-----------A------------------------C--------C-G--AC------A...-G---C-C----C-----A----G-----G--T--------------AG-AGTACCA 6784
B.CO.01.PCM001        ........................-G----G-G--A-T--TGCCC-C----A--CC--A---A-----------------G---C---------G--A-------A...C-----AC----C---------G------G--GA------------GA--....... 5971
B.GB.83.CAM1          .......AGGACC-A--G--GTG--T-GGAC--GA--GA-GGAGA--A-AT---------A-----------------G---C--------------AA--------G---------------C------------------G--G--------------G-C-....... 6788
B.GB.86.GB8           ..........................-C--A--C--AT-AT-GT---------GG------A--G-----------------------C-------------G-A----A-------------------C----G---------------G--T------G--G-------....... 6782
B.GE.03.03GEMZ010     ..........................-CA-CC-ATA-T--T-AT-GTTG-A-AG-G-TG--A--G----T--------------G---C-----------G--AA----A----C-------------C----G-----G--C---T--G-----------AG-GA--....... 5974
B.IT.05.SG1           .................................................-A--G------------------G---C--AC-A----G--AAA-C--C-...--C---------C--C--------TA-------G---A-------G--T--....... 6533
B.JP.05.DR6538        .AGCACTGAGACT-A-----GA-C--A--CC--TACC-CT-GTG--CCA--T-G--A-CCA--G--------------G--G---C-------C------A----A-TACA------------------C---------G--C----G--G------------C-GAC-AATAATA 6810
B.KR.05.05CSR3        .......ACTACC------GG--C-TC--AGT-GAATG-TGAT-AT-GCACT-T------A--------A----A-------CT-C----G----G-T--A----A--A--G----------------A--------C---------G-----C--------AG-CTTGCTA 6819
B.NL.00.671_00T36     .AGTTCCAATAAT-T--C-GGTTCC-A--A--ATA-T-ATT--GA-CA-AT-GC-A----A----GT----------------------C-----C-----A-AA----AGA----G-----T----C-G--GAG---------------G---------------------....... 6358
B.RU.04.04RU128005    ..............----CGGA-----C--C--ATAAT--T-CA-------G-----A---A----------------------CT-C-----A-------AA---AA-----------------------C-----------G---------G----------GG----....... 6272
B.TH.00.00TH_C3198    ....................-AT-G-----CG-C-AAT--T-GT--G--A-G-GG-----A--G--G-----------------G---C-----A----G---C------A...-C----------C-----------AC-------G---A-----C---------....... 5986
B.UA.01.01UAKV167     TACCATTACCAGT-A--G--------A-G---CTA-TT-GGAATAC------GG------A--G-----T--------------G---C----------G--AA-G-AA--A--G----T-------C-----------G----------AGT-----A-G---------....... 6049
B.US.04.ES10_53       ...................................C-TGC--ATAT--CG----A-----------A-------------C-----C--G----A-------A----C------C-C-------GCAC----G--G--AA----A-----A--GAT.... 6784
B.US.99.PRB959_03     ....................-------C--C---TAACT-GGAA-AG------CC------A---------------------G---C-----------------A------------------------TC------------------G--G---A-------G--ATAGAT.... 5995

C.AR.01.ARG4006       ..........ATT-GCRGA--T-T---C-A-GATACC-TTC-TGATGACAT---G-----A--------T--T-----------G-C----GAAC------A----CAAA----G-T-C-----C---------G-------R--G-----C-TA-------GAAT.... 5979
C.BR.04.04BR013       ....ATTATTGGCTAC----A-C-A-C--T--AATAC-ATGA--CC------GTG---C-A--------T--T-----------G-C----GAAC------A----AAAAG---G-T-C-----C---------G-----------G-----C-T--GG---GGAGTAAT. 6282
C.BW.00.00BW07621     ...................................-CC-ATG-TACT--T-CT--CTAT-----T---------------A-C----GAAC------A----AGAA-------------------C---------G---G------G---A-C-T-GGG-GA--AAT.... 6132
C.CN.98.YNRL9840      ..............-GC-G--ATGG--CCTAC-ATATT-TCCAC-ATGA-ACCT-CAAG--A--G-----T-----------GCA-C----GTAG--GA--A--G-AA--------TGC-----C---------G-----------G-----C-T---G-GA-CAACTCTA 5975
C.ET.02.02ET_288      TGTTACCAGTAATC-----GTT-CC------TCTACTGATGCC-GT-AA-TT-TT---C-A--G-----T--------------A-C----GAAG------A----AACA------C-------C----C---CG-----------G-----C-TAC-------AAC.... 6029
C.GE.03.03GEMZ033     ...................-GT-CC-A-GC--AAATT-AT-GT-CCTCC-TAGGG-A---A--G-G---T-----------G-G-C----GAAT------A-----AAA-----C-------GC-------------------------C-T-----G-GGAATAGCA 5981
C.IL.99.99ET7         ...................GTT-C-----A-GTTACC-TC-AT-ATGATACC-TCAAT--AC-------T--------------A-C----GAA-------A----AAAA---GT----T----C---------G----------G-----C-TA--G-GC-G....... 5972
C.IN.99.01IN565_10    .......................---GT---CTCGAAT--TGAT-AC-ATACT-GG-----A----GG--T---A----T-----G-C----GAGC---A--A----AAAA--C---TG-C----C---------G----------G-----C-T---GG-A-GAATGAGA 6163
C.KE.00.KER2010       TACCAGCAATAGC------GGG-CAG-C-A-GATACCTACGA--ATGACAT--CGA----A--G-----T-------------CA-C----GAA-------A----CA--------T----T--C------------------G--T--C-T--G----G-AAT.... 5992
C.MM.99.mIDU101_3     TGGTAGAAATGTT-GC-G--GT-G-G------AT--T-CCTAC--TGG-A-CT-CACT-----------T-------------CA-C----GTTT---A--A--G-AA----TC--TG-------C---------G----------G-----C-TA--G-GA-GAAATCTA 6163
C.MW.93.93MW_965      ................----AT---G---C-GTTAATGTC-ATGAT-CCAT-T-T-----A--------T--------------G-C----GAAT------A----AAAA--C---T-------C------------------G--T--C-TA--G-GA--AGT.... 5987
C.SN.90.90SE_364      ........................-A--A-GC-ACC--T-ATGAT-ACCCT--T-----A--------T--------------A-C----GAAC------A----AGAAG---T---------C---------G----------G-----C-TAG-GGGAG-AAAAAC. 5966
C.SO.89.89SM_145      ................-T-----G-GT-G-C-AT-TT-CC-GT-ATGACA-T--G-A---A--------T-----------GCA-C----GAAT----T-A----AAA-------TG-------C---------G----------G-----C-TA--G-G-G-GAA.... 6014
C.TZ.02.CO178         TACCATTAGAGAGTA-GT-G--CT-GAACA--AT-CT--CC-TGG--ATT-T-GT-AT--A--------T-----------G-A-C----GAAC---AG-A----ACAA---G--TGC-----C-C------TT----------GA----C-TA--G-GA-GGAGGACA 5990
C.UY.01.TRA3011       ......................-GA-C--AA-A-AACCAC-CTGAT-ACAT---G-----------G--T--TA-----------G-C----GTAC------A----AAAG----G-T-C-----C------------------G-G---C-T---GG-A-C....... 5965
C.YE.02.02YE511       ...............---GTT-CC-A-GA-GTTACCTTC-ATGAT-CCAT--GG--G--A--------T--------------A-C----GAA-----G-A----AGAA--C-G-----C---C---------GG---------G-G---C-T---GGAA-AGAA.... 5977
C.ZA.04.04ZASK164B1   CAGTAATGGTACT----CCC-T-G--C--C--C-CCT-CCCCC-AT-----C-CG--G-GA--G-----T--------------A-C----GAAC------A-----AAA------TGA------C---------G----------G-T---C-GA-G-C-A--GAG.... 6241
C.ZA.05.05ZAPSK240B1  ...................................-AT-CTGACAT-G-GAA---A--G-G---T-----------GCA-C----GAA-------A----CAA-------CG------C------------C----T--G-----C-T--GG-GA--....... 6180
C.ZM.02.02ZM108       TAGTGTTAAAGAG-A--CC-----C-A-GC--T-TTT-AC--A-C-GAAATA--GAA---A--G-GT--T-------------G-A-C----GAAC------A----AAAA--C-G---------C---------G---------------TC-TA---GG---GAGAATA 6796

D.CD.83.ELI           ...................-ATGGC-C--TGG-GAAC-ATG-C-CT-CA---G-GAA--GA--G-----------------G-A-C----GTAC---A--A----AA----C---T-------C---------G----------G---------C------AGTAGT. 6327
D.CM.01.01CM_0009BBY  .....................-GC-C-GAG-CT-AATTG-AC-AC-GCAGC-CT--------G-----T--------T-----A-C----GAAT----G-A----AAAA--C-----------------C------------------G--GA---------GA-GGATAATA 5981
D.KE.01.NKU3006       GACCATGAAAAATGAG--C-TG--A-ACCTC-AT-AC-CC--A-A--ACCTC--T-AGACC--G-C-----------T-GGG-A-C----GAAT------A-----A-A--C---TGC-----C---------G-------G-GG--AA--C----GGG-G-....... 6037
D.KR.04.04KBH8        ...............................ACT-TC-AC-AC-ACAGC-CT--C--A--G-----------------G-A-CT---GTGG----G-A----AA-A--C------------C---------G----------G-------G--------......... 6716
D.TD.99.MN011         ...........................GTTACAGTG-GT-AC-GCAGCTT---------G---------------------A-C----GTAG----G-A----AA-A--C------------C---------G----------G--GA---G-G-G--A--GAG.... 5993
D.TZ.01.A280          .......AACAGTGGG---GGG-C--AC-CG-CT-ATCCTCGCC-T--T--A-----------G---------------------A-C----GAA-------A----AGAA--C---T-C-A---C-----------------G-GG-------------AA-G....... 5988
D.UG.99.99UGK09259    ....GATGCCACG-A---G-ATGCC-C--A--ATACT--T-ATG-C-CTA-TG--AT--GA--G----TT------------G-A-C----GAAG------A--G-AA-A--C---T--C----C-----------------G-GG--A--------G--ACC....... 5995
D.YE.01.01YE386       .......AATACC----G-GG-----A--CC-CTAAT-AT--C-CT-AT---G-TAT--GA--G-----------------G-G-C----GAAT------A----AA-A--C---T-C-----C-----------------G-GG-----------------A--AAT.... 5989
D.YE.02.02YE516       ...........................-CA-TAGAC-AC-AC-GCACA-G--A------G-----------------G-A-C----GTAG----G-A----CAAA--C----------G--G-----------------G--G---G------......... 5974
D.ZA.90.R1            .........................TCCG-G---ACC-CC-CTGAG-A-CCA-C-CT----G-GC-----------------G-AGCT---GTAG------A----AGAA--G---C--C---GC---------G--C-------G-----------CC-G---AATAAA. 6142

F1.AR.02.ARE933       .........................-CG-CA-A-AAC-CCC--GCGGAT---CC---G-CA---C-------T---------G-G-C----GAAG------A----AA--T-----T-C-C---A---------G----------G-------CAG-GGCAG-AAT.... 6076
F1.BE.93.VI850        ...................................-GTCAGGAAA--CC---G-CA--GC-------T-----------G-C----GAAG------A----AA--T----TT--C---GC--------GG----------G---------GC---A-C....... 6104
F1.BR.01.01BR125      ..........CAA-AC-GC---CTG-AGGAAG-A-ACGCCC---AGGAA---CC---G-CA---C-------T-----T-----G-C----GAAT------A----CA-AT-----T-C-----C---------GG---------G-------CAGC---ATCAGTAAT. 6242
F1.ES.x.P1146         ..........................GAGCAA-A-A-C-CCC---AGGAA---G------CA-----------T-----------G-C----GTAG---AT-AC---AA-TTA----T-C-GT--C---------G---------CG-----C-CA-C---A--AATAAGA 6090
F1.FI.93.FIN9363      ..........AATGAC-C-CT-TC-GACCAA---A-T-CCC---AGGAA--ACC---G-CA---C-------T-----------G-C----GAAG--GA--A----AA--------T-C----GC---------G----------GA------CAGC---A--AAT.... 6123
F2.CM.02.02CM_0016BBY ..........GCT-A----------CAGCC--ATAAC-TC-CCC-GGAA---C----G--A-----------T-------------C----GAG----A--A----AG-A-A--T----T----AC-----------A-------G-------T--A--AA--....... 5977
F2.CM.95.MP257        .......GTCACCT--TC--AT----CC-CCTTG-CTCCT-ATG-T-CCATCTCG-AG--A--G--------T-----------A-C----GAG-------A----CA-A-A-------------C------------------G--G---A---CA----CAG-....... 5999
F2.CM.97.CM53657      .......ATCACT-A-GGG-A--GC-CCCTGGATAAC-TC-CCC-GGAA--AC----G--A-----------T-------------C----GAG----AC-A--TTAA-A-A-------C-----C---A---------GG------G--G-------CA-----AG-....... 5980

G.BE.96.DRCBL         ...........................--TACTC-G-AT--C-CT---G-GTACAGA--G-C------------------G-C----GAAC-G----A----AA-A--GC------C--G------C--GC---C----G-AC----G--G-------TA--G-GATGAATAATG 6734
G.CM.01.01CM_4049HAN  .......ACTAATGG--C--TAG---A---C-ATAAT-TC-C-G-GGAT-GT----A---A----CC--T--T--C---------A-C----GAA-------AC--AA-AAG-C-G-----C--G-----C--GC---C-----------G--G-------TA--------......... 5998
G.CU.99.Cu74          ...................-A--GC-C-G-G--AAATT-CTCTG-GGAAA-AGGT-A---A---------------------A-C----GAG----AG-A---AAA-A-AC----------G---------GC---C---G-----------G-G-----TA--G--A--AGAGATA 6396
G.ES.99.X138          .................-GC--T----AA-CAGA-AAC--T-GC-CTG-A--C-------AC---------------------A-C----GAA-------A----AA-A--C----G----G-----C---G-----C---G-----------G-------TAC-G--A--GGTAAT. 6235
G.GH.03.03GH175G      AAAGCATCCCAATGGC-C-GTGG--TAC-A----ACTGTGGAT-AC-GTAGCTGT-A-ATA---------------------A-CT---GAAC------AC--AAGTA-------G--C--G---G--C--GC-C--C---GG-C----G--G-------T---GG-A--GATAAT. 6861
G.KE.93.HH8793_12_1   ............................-CAGTA-CC-ACGAA-GCG--AGTGCCCA---A---------------------A-C---GGAA-------A--GAAA-AG----------C--G------C--GC---C---G-------G---------TA--G-----AGTAGTA 6183
G.NG.01.01NGPL0669    ....TGTACTAAG-AC--C-A--C-G-G--C-CGAATGT--CTTGT-ATA--G---T---A--G------------------A-C----GAA-------A---ACA-A-A--------C--G------C--GC---C-----G--------GG--T----TA--G---G-GATAGTA 6012
G.PT.x.PT2695         CACTGTGGGTAGC---GCC-G--C---C-C-GTG--T--T-CTGCGGGTT-C-G--A---AC--------------------A-C----GAA----A--A--G-AG-A--C--------C--G------C--GC--------------------G-------TA--G---G-AGGAAT. 6817
G.SE.93.SE6165        AAATAATAGCACTGA-----G--C-GAA-CC-A-AAT--C-CTG-GGATA-CCC------A-----------------G-A-C----GAA-------A----AAAA----------C--G------C--G-----C---G-------G--G-------TA-----.......... 6209

H.BE.93.VI991         .......AATGTA--A--A-GT----AC----CA-AT-TC-AT---GGA----T-CAG--ACA--G---------------G-A-CT---GC---------A----AAC-----G-T-C-----C-C-------G-GC---C---G---AG--T---G-A-G-GAAAGA. 6175
H.BE.93.VI997         .......GATACT-AC-G--G--GC-C-GTC-AT-CC-CT-GC-GTCCTTCAGC-AAT--AT---C---T------------G-A-CT--TGTA-------A----CA-----G--T-C-----C-C-------G-------G--G-G-----T---G-GAC-AGTAATA 6103
H.CF.90.056           ......................-CCTC--AC---ACT--C-GT--GGA--CAGG---G--AC---C---T------------G-A-CT---GTAC------A----CA--------T-C-----C-C-------G-------G--G-------T-----CA--....... 6116

J.CD.97.J_97DC_KTB147 ......................-CC-C----CCTA-TCCT-CT--C-CTA-TCCTCTT-G---G---------------------A-C----GAA---G---A--G-AG-A----------C--G------C--GC---C-------AA---G--G--T----T------AG-....... 5997
J.SE.93.SE7887        ..................-GT----CA-C---TAAC--T-GTG---GTAGTCCT-ATATC--G-C--------------------A-CT---GAA------AA---AAG-A--C----G--C--G---G--C--GC---C----G--AA---G--G-------T----G-A-C....... 6098

K.CD.97.EQTB11C       .CGTACCAATGCA-A---G-A-G---CC-A--T-AATGC--CTG-G-CATCA-CT-AC-----------T--------------A-C----GAAC---A--A----AAAA---G--TG-C----C----C----G---------G-G-A----A-GC-ATC-GAAATT. 5998
K.CM.96.MP535         ..................----G--CC-A--A-GCAACT--C-CTG-GG-AAGCCC--C---A--------T--------------A-C----GAA----A--A----AA-A-A--G----C----C----C----G-------G--C-G---C-TA--GGC--GGGTAAT. 5987
```

```
                                                V1 loop end      V2 loop start
B.FR.83.HXB2      .............ACTAATACCAATAGTAGTAGCGGGAGAATGATAATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAGGTGCAGAAAGAATATGCATTTTTTTATAAACTTGATATAATACCAATAGATAATGAT....... 6785
Env               __ __ __
```



HIV-1/SIVcpz
Complete Genomes

```
                                                                                                                      V2 loop end┐
B.FR.83.HXB2                  ................................................................................ACTACCAGCTATAAGTTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACATTAT 6875
Env                           .__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__._ _T__T__S__Y__K__L__T__S__C__N__T__S__V__I__T__Q__A__C__P__K__V__S__F__E__P__I__P__I__H__Y_
01_AE.CF.90.90CF11697         AT..................................................................AGTAGTGGAAAT-G--GTGAG----TA--A-T-CA---------T----------AG--A--T------------A-------T------------------ 6845
01_AE.CN.05.FJ051             ................................................................................-A---TGAG----G---A-T--A------T--T----------AG-----T---------A----------T--------T--------- 6881
01_AE.CN.06.FJ054             .......................................................................AATGAAAAT-A---TGAG----GA--A-T--A------T--T----------A----C-T---------A-------C--T--------T--------- 6920
01_AE.HK.x.HK001              ...................................................................................-GT-CG----G---A-TT-A------T--T----------AG--A--T--------------------T--------T--------- 6251
01_AE.TH.01.01TH_R2184        ..........................................................................AGTAAT-G---TGAG----G---A-T--A------T--T-----A----AG-----T---------A----------T------------------ 6090
01_AE.TH.02.OUR769I           ................................................................................-G--GTGAG----G---A-T--A------T--T----------AG-----T---------A----------T--------T--------- 6077
01_AE.TH.90.CM240             .............................................................................ACT-G--GTGAG----G---A-T--A------T--T----------AG-----T---------A----------T--------T--------- 6455
01_AE.US.00.00US_MSC1164      .............................................................................AAT-G--GTGAG----G---A-T--A------T--T----------AG-----T-----C---A----------T--------T--------- 6086
02_AG.CM.02.02CM_4082STN      ................................................................................GTAGGGCAG----G---A-T--A------T--T----C------------T--C---------------C-------------------- 6062
02_AG.EC.x.ECU41              .............................................................................GAC-A--G-CAT----G---A-T--A------T-------C---------A--T------------A-------------------------- 5952
02_AG.GH.03.GHNJ196           CAATA............................................................AATACAACTAACAGC---GATCAG----G---A-T--A------T------AC-G----------T--------------------------------------- 6931
02_AG.NG.01.PL0710            .............................................................................AAC-G--GTGAG----G---A-T--A---C--T-------C------------T---------A--A-------------------------- 6053
02_AG.NG.x.IBNG               .............................................................................AATGG--GTCAG----G---A-T--A------T-------C------------T--------------------------------------- 6399
02_AG.SE.94.SE7812            .............................................................................GAT-AC-TTCAA----G---A-T--A------T-------C------------T--------------------------------------- 6241
02_AG.SN.98.MP1211            ................................................................................G----TTAT----G---A-T--A------T-------C------------T-----------------------------T--------C 6068
03_AB.RU.97.KAL153_2          ...................................................................................GAT-------G-----T---------T---------G-------A------------A----------------------------- 6073
04_cpx.CY.94.CY032            TTAATGGT......................................................AGTAATAGGAATAATAGT--AGAAGAG----T---A-T--A-------G-----AC-----A------T--C--------G--T------------------------ 6263
05_DF.BE.x.VI1310             ..........................................................................AGTAGC-A--G---T----G-C-A-T--A------T------AC-----A------T--------------T-GG--T------------------ 6278
06_cpx.AU.96.BFP90            .............................................................................AAT-A--GTGAT----GA--A-T--A------TGT----AC-----A------T-----------G--T-----A-----C------------ 6947
08_BC.CN.97.97CNGX_6F         GT.....................................................................GAGAACTCT-G-GAATAT----GA--A-T--A------T-------C---A-----A--------------CA-T-----T--------T-----C--- 6063
09_cpx.GH.96.96GH2911         ................................................................................-A--ATGAG----GAC-A-T--A------T-------C------------T------------A-------------------------- 6074
10_CD.TZ.96.96TZ_BF061        ................................................................................-A-----------GA--A-T--A------T-------C------------A------------A----------------T--------- 6248
11_cpx.GR.x.GR17              ..........................................................................AGTAAT-TAGGTCAG----G-C-A-T--A------TGT-----CAG-C-A------T--CT----A---A----------------T-----C--- 6148
12_BF.AR.99.ARMA159           ................................................................................-G-GAA-CA--C-G-C-A-T--A------T------AC-------T----T--------------T-GG--T------------------ 6888
13_cpx.CM.96.1849             GT..................................................................GGTAATAGTAGT-G--GTGAA----GA--A-T---------T-------------A---A--T------------A-------------------------- 6297
14_BG.DE.01.9196_01           .............................................................................AAT-G-----AA----G---A-T----------------A----------------C---------A-------------------------- 6398
15_01B.TH.99.99TH_MU2079      ...................................................................................----------G-----G--AG------------A----------A--T--------------------------------------- 6247
16_A2D.KR.97.97KR004          AT.....................................................................GATACACTG-A--GGCAG----GAC-A-T-CA----G-T------AC---C--------T--C--------------------------T--------- 6244
17_BF.PE.02.PE02_PCR0155      ................................................................................GGC-GTGAA--C-G-C-A-T--A------T------AC------------T--------------A-GG--T------------------ 6089
18_cpx.CU.99.CU76             ..........................................................................AGTAATTA--GTGAT----G---A-T--A------TGT----AC-----A---------C-----A--G--T-----C------------------ 6183
19_cpx.CU.99.CU7              ............................................................................................---A-A-T--A------T-------C------------T-----------CA-------------------------- 6019
20_BG.CU.99.Cu103             AT..................................................................AATAGTAATAACTA--GTGA-----G-C-A-T--A------TGT-----C-----A------T-----------G--T-----T------------------ 6345
21_A2D.KE.91.KNH1254          .............................................................................GAT-A--GT--T----G---A-T--A------T-------C------------T---------A----------------------------- 6094
22_01A1.CM.01.01CM_0001BBY.   ................................................................................-A--AT--T-T--GAC-A-T--A------T-------C---------A--T------------A-------------------------- 6066
23_BG.CU.03.CB118             GT..................................................................AGTAATAAAAACTA--GTGA-----G-C-A-T--A------TGT----AC-----A------T-----------G--T-----T------------------ 6345
24_BG.CU.03.CB378             .......................................................................AATAACAACTA--GTGA-----GAC-A-T--A------TGT-----C-----A------T--C--------G--T-----T------------------ 6333
24_BG.CU.03.CB471             GTAGT...............................................................AATAATAATAACTA--GTGA-----G-C-A-T--A-------GT----AC-----A------T--C--------G--T-----T------------------ 6345
25_cpx.CM.01.101BA            ..........................................................................AATAGT-A-GAT--T----G-C-A-T--AC-----T-------C-----A------T------------A-------------------------- 6137
27_cpx.CD.97.97CDKTB49        ..........................................................................ACTACTGGGG-TGAG----G-C---TC-A------T-------------A------T------------A-------------------------- 6091
28_BF.BR.99.BREPM12609        ................................................................................-A-G-A-------C-C---T--A------------------------A-----------------------A------------------ 6249
29_BF.BR.02.BREPM119          ................................................................................-A---T-------GA----T---C--------T------C--------------------AC---------------------------- 6039
31_BC.BR.02.110PA             ..........................................................................AGCAGTT--GGTGAT----GA--A-T--A------T-------C---A-----A--------------C--T-----C--------T--------- 6365
32_06A1.EE.01.EE0369          .......................................................................AATGGTAAT-G--GTGCT----TA--A-T--A------TG--A--AA----GA------T-----------G--T-----A-----C------------ 6546
33_01B.MY.05.05MYKL007_1      ..........................................................................AGTAAT-G--GTGAG--C-G---A-T--A------T-------------AG-----T-----------G--------T--------T--------- 6275
34_01B.TH.99.OUR2478P         ................................................................................GG-GGTGAA----G-----T--A------T--T----------AG-----T---------A----------T--------T--------- 6068
35_AD.AF.05.05AF095           .............................................................................AAT-G--GT--T----GA--A-T--A------T-------C---------A--T------------A-------------------------C 6053
36_cpx.CM.00.00CMNYU830       ..........................................................................AATGAT-G--GT--G----GAC-A-T--A------TG------C---------A--T------------A-------------------------- 6087
37_cpx.CM.00.00CMNYU926       ................................................................................-A-GAG-CT----GA--A-T--A---C--T--------------------T------------A--------------------T-C--- 6069
38_BF1.UY.05.UY05_4752        GTAGT...............................................................AGAAGTAATAGC-G-GGAGAA--C-G-C-A-T--A------T-G----AC------------T-----------G--T-GG--T------------------ 6325
39_BF.BR.03.03BRRJ103         GC.....................................................................AGCAATGGC-G-GGA-AA--C-G-C-A-T-CA------T------AC------------T--------------T-GG--T------------------ 6383
40_BF.BR.04.04BRRJ115         .......................................................................AACAGTAGTGAA-AATA---C-G-C-A-T--A------T------AC------------T--------------T-GG--T------------------ 6444
42_BF.LU.03.luBF_05_03        .............................................................................AAT-G---A-AT----T-----T---Y--------------YC----------------T---A----------A------------------ 6415
43_02G.SA.03.J11223           ..........................................................................AATGAAT-AGGTGAT----G-C-C-T--A------TGT----AC---------A--T-----------GA-T-----C-----------------C 6403
U.CD.83.83CD003               .......................................................................TCTAGTATA-G--GT-C-----G---A-T--A------T------AC------------T-----------------------------T--------- 6357
U.CD.90.90CD121E12            ..........................................................................AGTAGCTA-GGG-C-----G-C-A-T--A------T------AC---C--------T-----------------------------T--------- 6043
U.GR.99.GR303                 .......................................................................AAGAGCAATTA--GTGAG----GA--A-T--A-----GT-------C------------T--------A--GA-------------------------- 6168
U.NL.95.U_NL_95_H10986_D1     .......................................................................AGCAGCAATTA--GTGAA----G---A-T--A------T-T----AC-----A------T-----------GGAT-----A------------------ 6772
CPZ.CD.90.ANT                 .............................................................................GAG----A-T-----T-C--ATGGCAC--C------A--ACT--C-----AT-----GA----TCTA-T-----A--------T--------- 6328
CPZ.CM.01.SIVcpzCAM13         .............................................................................ACTGG-GAT-CA----CAA---TT-A----------A---C---A--------A--------A--G-----------T-----A-----C--- 6486
CPZ.CM.05.SIVcpzEK505         .......................................................................AGTGGTAAT-AC-GTCAA----G---A-T--A----------A---CTG-G--------------T--AA-T-----------------A---T----- 6424
CPZ.CM.05.SIVcpzLB7           ................................................................................--CGAG---------C-T-T--A------T--TA---C---A--C-----T--C------A-T--T-----------A--T--------- 6422
CPZ.CM.05.SIVcpzMB66          CA..................................................................GAAGGAGATAAG-ACG--TTA----GA--ATAT-A----------A---C---G--T--A--T--CT----A--C--TC----A--------T--------- 6520
CPZ.CM.05.SIVcpzMT145         ................................................................................-A--GT-CT--C-G-----TT-A------T---A---C---C--T--A--T--------ACT---------A--------A--------C 6405
CPZ.GA.88.GAB1                ................................................................................-AC-A--CA----G-A-A-TT-A---C--T--TA---C---A-----A--A--------AACG-----------T-----A-----C--- 6936
CPZ.TZ.00.TAN1                ................................................................................-A--AT--T---T-T--ACAT-A---C--T--------------T--A-AA---GAT--ATCTA-T-----A-----------CAGA--C 6519
CPZ.US.85.CPZUS               ...................................................................................-AT-C-----GAC---T--A---C------A-C-CT--C--------T-----T--ACC------------------A-----C--- 6919
CPZ.CM.x.SIVcpzDP943          ................................................................................-M--AKTC-----G-C-A-T--AC-----T---A---CT--------A--A--C--T--AACC-----------------A--------- 6444
CPZ.CM.x.SIVcpzMB897          ..........................................................................GTTAAA-A-GGA-CT----G---A-TT-AC-----T---A---C---A--C-----T--------AAC---A-----A--------T--TT-C--- 6425
N.CM.02.DJO0131               .............................................................................AAT-A--GT-CA--C-G-C-A-T--A------T---A---CTG-A-----A--T-----T--ACTA-----------------A--------C 6359
N.CM.04.04CM_1015_04          ..........................................................................AGTACG-A-CAA-CA----G-C-A-T---------A---A--TCTG-A-----A--------T--ACCA-----------------A--------C 6352
N.CM.04.04CM_1131_03          ..........................................................................AGTGAC-G--AT-CA----CTC-A-T--A------G---A----TG-A-G---A--T-----T--ACCA-----------------A--------C 6387
N.CM.95.YBF30                 ................................................................................-A--AA-CA----G-C-A-T--A------T---A---CTG-G-----A--T-----T--ACT------------------A--------C 6468
N.CM.97.YBF106                .............................................................................GAC-A--AT-CA----G-C-A-T--A------T---A---CTG-G-----A--T-----T--ACTA-----------------A--------C 6441
O.BE.87.ANT70                 CA..................................................................AATAAGACAAAC-GC-AA-TG----CA--A--T-A-------T--A--AC---C--G--A-----------------T-----A--------A-----C--- 6910
O.CM.91.MVP5180               ..........................................................................GTAAATGGA--A-CA----T---A--T-A-------T--A--A-T--C-AG-----------G------AGT--------C-----------C--- 6909
O.CM.96.96CMABB637            GT.....................................................................AATAATACTGAAC-A-TA--C-CA--A-TT-A---C--TT--A--AC---C-AG--A-----C-----A--TAAT--------C--C--A--------- 6397
O.CM.98.98CMA104              .......................................................................AACAAGACA-GAGAA-TA----CA--A-TT-A-------T-TA-TACA-CC--G--A-----------------T--------C-----A---A----- 6366
O.CM.99.99CMU4122             .............................................................................ACC-GC--A-TG----CA--A-TT-A---C--TT--A--AC-----AG--A---------------A-T--------------A-----C--- 6347
O.FR.92.VAU                   .............................................................................AATGAA--A-TG----GA--A-TT-A-------T--A--AC---C-GG--------------------T--------C-----------C--- 6449
O.SN.99.SEMP1299              ..........................................................................GACACA-AGGA--TG----CA--A-TT-A-------T--A--AC-----AG--A-----------------G--------------A-----C--- 6948
O.US.99.99USTWLA              ................................................................................--A-A--TA----CA--A-TT-A------TT-----AC---C-AG--A---------------A-T--------C-----------C--- 6369
```

```
B.FR.83.HXB2               TGTGCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAG...GTAGTAATTAG 7042
Env                        _C__A__P__A__G__F__A__I__L__K__C__N__N__K__T__F__N__G__T__G__P__C__T__N__V__S__T__V__Q__C__T__H__G__I__R__P__V__V__S__T__Q__L__L__L__N__G__S__L__A__E__E__E__.__V__V__I__R_
A1.GE.99.99GEMZ011         --------A-------------------G--C-GGG--CCA-AT--T-----------G-----C-AG-------------------C-------------A-------------------------------A----------------------GA--A-...---A-G----- 6229
A1.KE.00.KER2008           -----T--A-----------------TG-----GG-----GA---------------G-T---C-AG-------------------C-----------C-A---------G--T--------A--A-----------T----G---A--AG-...---A-G----- 6275
A1.KE.00.KNH1144           --------A--------------C--G--G----GGG-----GA---------------GGA---C-AG-------------C-----C-----------C-A-------------------------A--------------------------A--...---AA---C-- 6286
A1.KE.00.KSM4024           --------A--------------CT----G-----GG--G---AA--------------G-----C-A--------------C-----C-----------C-A--------------------------T--------------------G-...---...---AA------ 6256
A1.KE.00.MSA4069           --------A--------------C------------G---T--AT-A---G---CA---G-----CCAG--------T-----------C-----------C-A--------------------------T--------------T-------C-----...A---C------ 6268
A1.KE.00.NKU3005           --------A--------------C-----G-----G--G--GA----------------G-----C-A--------------------C-----------C-A---------------------------------------------A--A-TA-T...---ACT----- 6274
A1.RU.00.RU00051           --------A-------------------G--C--CC-A--AGAT--T-----------G-----C-GG-------------------C-----------C-A-------------------------------------------------G-G-A-T...---ACG----- 6343
A1.RW.93.93RW_024          --------A--------------C-----G-----GG-C-CAGAT-A-----------G-----C-GG-------------------C-----------C-A---------------------------------------C---------AG-A-T...---AC------ 6253
A1.SE.95.SE8891            --------A--------------C-----G--C---G-----GA---T----------G-------AG---------T----------C-----------C--------------------------------------------T-------AC-----...---A-G----- 6225
A1.SE.95.UGSE8131          --------A--------------C-----G------G-AG-TGA-----------T--GA----C-GG-------------------C-----------C--------------------------A--------------------A----------...---AG------ 6432
A1.TZ.01.A173              -----T--A-------------------G-----G-----A-----------------G-----C-AT---------T----------C-----------C--------------------------------------------------------G---...A--CA------ 6260
A1.UA.01.01UADN139         --------A-------------------G--C---G---CA--T--T---------------C-GG-------------------C-----------C-A------------------------------------------------A-T---...---A-G----- 6253
A1.UG.92.92UG037           --------A-------A-------------------G-----GA---------------G-T---C-AG-------------------C-----------C--------------------------------------------T---------G-A--...---A-G----- 7076
A1.UG.99.99UGA07072        --------A-----------------G--G-----GG---GAGAT--------------G-----C-AG-----------------G--C-----------C-A--------C--------T---------------------------------G-----...--TA-G----- 6266
A2.CD.97.97CDKTB48         --------A-------A-----------G-----GG-----GAA---------A--GT---C-GT----T--T----------G----------------C------------------------C-G--------------------G-A--...---A-G----- 6370
A2.CY.94.94CY017_41        --------A-------------------G-----GG--CC--GA---------------GT----C-AG----T--T--------------------AA------C-----------------------G-------------------G--G-AAAA--A-G----- 6426
A.SN.01.DDI579             --------A-------------------G-----GG-G---AA----------------G-----C-GG----------G-------------------C-A----------------------------G-------------G---------...---A-G----- 6258
A.SN.01.DDJ369             --------A-------------------G-----GG-----GAT----------------A----C-GG-------------------C-----------C-A---------GA-------------------G----------------G-A--...---A-G----- 6267
A.SN.96.DDJ360             --------A--------------C-----G-----GG------AA--------------G-----C-GG----------G-------------------C-A-----------------------T---G-------------G---------...---A-G----- 6243
A.CD.02.02CD_KTB035        -----A--A-----C--C-----------G-----AG-----GA---------------G-----C-AG-----------------------------A-----------------------------------------------------G-A-A...--CA-G----- 7071
A.CD.97.97CD_KCC2          --------A------------------G--G-----AG--C--GA---------------G-----C-A---------------------------AA--T-------------------------------------------T---------G--GA...A--AAG----- 7044
A.CD.97.97CD_KTB13         ----T---A-----------A------------CG-G---GA---------------G-----C-AG----T--T------------------C-A---------------------------------C--------------------A-----ACA...---A------- 6333
B.AR.04.04AR151516         ---A----A-------------------G------G------A---T-------------C-------------T-------------C----------A-----G--G--------------------------T---T------------C...A---------- 6260
B.AU.87.MBC925             ----------------------------G-----AG--G-A-A-----------A-----T-------------------------------------------------------------------------------------------G-----...---A------- 7070
B.BO.99.BOL0122            ---A-T--A-----------A---------------G-----AT----------------------A-------------------------------A--------G----------------------------------------G-----A...A---------A 6237
B.BR.03.BREPM2012          ----------------------------G-------------A--A--------T---G------A---------------------------------------------G--C-------------------------C-----------G---...A---------- 6505
B.CA.97.CANB3FULL          ---------------------C------G-----CG------AA-------------------------------T--------------------------------G--------------------------------T---------------...----------- 6330
B.CN.05.05CNHB_hp3         ---A----------------C-C-----G-----C-------A-------------T--------------------T---------------------A--------A-------------G-----T---T--------C-------...----------- 7067
B.CO.01.PCM001             ---A--------------------C-G-------G-----A-A---------------------------------------------------------------------------G---------A---C------------------...A---------- 6228
B.GB.83.CAM1               ---A------------------------G------G-------A--------A-------------------------------------------G--G--C--------AT--------------------------A----...----------- 7045
B.GB.86.GB8                ----------------------------G-----C-------A---G------A-------------------------------------------------------------------------------------------------A--...----------- 7039
B.GE.03.03GEMZ010          -------------A-----A---G----------G--G---A--------------------A------T---------------------------AC---G-----------A------------------A--------A...---A------- 6234
B.IT.05.SG1                --------A-------------------G-----CG----A-G---------T-----------------------------------------------------------------------G--------T-A------------------...----------- 6790
B.JP.05.DR6538             ----------------------------G---------T-GC-------------------------------------------------------------A---------G--------------------------------A----...----------- 7081
B.KR.05.05CSR3             ---A------------------------G-----CG---CA-A-----------------------A--------------------C-----------------------G--------T-A-------------T-------------C...A---------- 7102
B.NL.00.671_00T36          ---A--------------------------------G------A---T--------G--------A-------------------------------------------G-G-----C---T-A-----------------------------T...----------- 6615
B.RU.04.04RU128005         -------------G--------------G-----AG-G--T-A-------------------A-------------G--G----------------------------------------C----G------------------------------------G---------G--- 6529
B.TH.00.00TH_C3198         ----------------------------G-----C---GCA-G-----------------------T-----------------------------------------------------------------T------------------...--A-G---- 6243
B.UA.01.01UAKV167          --------T--------------T----G------G------A---------A---------------T-------------------------------------------G--------T-A-------------------------GG--C...A---------- 6306
B.US.04.ES10_53            --------------------------G--G----GAG---A-A--------------------------T----T----------T----------------------------------------------------------------T------------------...---A------- 7044
B.US.99.PRB959_03          ----------------------------------C-----A--A---------T-----------AT-----------------G-------------C--------------------------A--------------------------...----------- 6255
C.AR.01.ARG4006            -----T--A--G----A-----------G--------------A------------------C--T--G-----T-----------------------A------G------------------A-----------T------------------...A--A------- 6239
C.BR.04.04BR013            -----T--A--G----A-----------G--------------A------------------C-AT--------------------------------A------G------------------A-----------T--------------------A...A---------- 6545
C.BW.00.00BW07621          -----T--A-------A-----------G-------C------A---G--------------C--T----------T--------------------A-----------------C-----A-----------T-----T----------G-...A---------- 6395
C.CN.98.YNRL9840           --CA-T--A-------A-----------G-----G----T--A--------G----------CCAT----T-------------C-------------A------G-----------G--A--------------T-C-----------G----...A--A------- 6255
C.ET.02.02ET_288           -----T--T-------A-----------G-----G-------A-------------------C-A-------------------C-------------AA-----G------------------AT-----------T---A------------T---...AC-A------- 6289
C.GE.03.03GEMZ033          -----T--A-----C-A-----------G--------------A-------------G-----C-AC-----------------------------------A------G------------------A-----------T-C--------------------...A--A------- 6261
C.IL.99.99ET7              -----T--T-------A-----------G-------C-----A-------------------C-AT--------------------------------AA-----G-----------G--A--C--------T-------------A-------...A----G---- 6232
C.IN.99.01IN565_10         -----T--A-------A-----------G--------G----A-------G------------C-GT-------------------------------A-----------------------A-----------T-C--------A----------...A--A------- 6437
C.KE.00.KER2010            -----T--A-------A-----------G--------------A-------G----------CCAG----T---------------------------A------G------------------A-----------A-C--------------------...A--A-G----- 6252
C.MM.99.mIDU101_3          --C--T--A-------A-----------G-----GG------AA-------G--------C-CCGT----T---------------------------A------G--------C------A-----------T-C--------------------...A--A------- 6431
C.MW.93.93MW_965           -----T--A-------A-----------G--------------A----CA----T------------C-AG-----------------T---------A------G------------------A-----------T-C--------------------...A--A--G---- 6250
C.SN.90.90SE_364           -----T--A-------A------T----G--------------A-----------------------C-AT--------------------------A------G------------------A-----------T--------------G---A...A--A------- 6232
C.SO.89.89SM_145           -----T--A-------A-----------G--------G----A-------G------------CCGT-------------------------------A------G------------------A-----------T-C--------------G---...A--A-G----- 6277
C.TZ.02.CO178              -----T--A-------A-----------G--------------A-------G-----------C-A--------------------------------A------G------------------A-----------------C---A------------...A--A------- 6258
C.UY.01.TRA3011            -----T--A--G----A-----------G--------G----A-------------------C-AT------------------G--------------A------G--G---------------A-----------T-----------C-A----A...A--A----C-- 6225
C.YE.02.02YE511            -----T--T-------A-----------G-------C-G---A-------------------C-AT--------------------------------AA------------------------A-----------------------------------...A---------- 6240
C.ZA.04.04ZASK164B1        -----T--A-------A-----------G----G-------A-A-------G------------C-AG----T-------------------------A------G------------------A-----------T-C---------------G-...A--A----C-- 6504
C.ZA.05.05ZAPSK240B1       -----T--A-------A-----------G--------------A-------G------------CCTT-------------------------------A------G------------------C-----------T-C--------------------...A--A------- 6437
C.ZM.02.02ZM108            -----T----------A-----------G--------------A-------------G-----C-AT----------T------------------A------------------------------A----G----T---T------------------...A---------- 7076
D.CD.83.ELI                --------A-------------------G---GAG------A--------------C-----C--------------------------------------------G--G------------------G-------------------------------...--CA------ 6590
D.CM.01.01CM_0009BBY       --------A-----A-----A-----C-----C---G-C----G--------------C-----C--------------------------------C-A--------G--------T-------------------------------C...A-TA-G----- 6255
D.KE.01.NKU3006            --------A-----A-----A---------------C-----------------G--G--T-----C-A----------------G---------G---A--------G--------------G---------T------------------...AC-A------- 6294
D.KR.04.04KBH8             --------A-----------A-------G-----C-----A--A--------------C-----C------------------------------------A-----G-------------------------------------A-G--A...A--A------- 6973
D.TD.99.MN011              --------A----------G--------------G-----A-----------------T--------------------------G-------------C-A-----G--G--------T-T--A-------------G-----------------...A-C--G---- 6253
D.TZ.01.A280               --------A-----------A-------G-----AG------A--------------C-----C-A--A----T-------------------------------G--G--------------------G--------T-----G--------...A--A------- 6248
D.UG.99.99UGK09259         --------A----A-A---------------------------T----G--G--T---CCA-----------------G---------G---A---------------AT--------------------------G-------...A--A------- 6252
D.YE.01.01YE386            --------A----A-A-------------------G------G---T----G--G--C-T--C-GG--C----------------G--C-------G--------------G--C-----G---T----G--------C-------------A...A--A------- 6252
D.YE.02.02YE516            --------A---------GA----------------G-A---AT-A--CA---A---C-----C-AC--------------------------C-A-----G--G--------T-T----G--------------G---------------...A--A------- 6228
D.ZA.90.R1                 -----T--A-----------A---------------GAG------A--------------C-----C--------------------------------------A-----G-----------T--------G-----------------------...A-CA------- 6408
F1.AR.02.ARE933            ---A-T--A-------A-----------G--N--G---A-G--------G--G--G----C-GG-------------------------------CA-----G-----------TC-T----------T-C--------------T...A--A----C-- 6339
F1.BE.93.VI850             -----T--A-------A--A-------------G---A-GA--------G-----G----C-AG-------------------------------AA-----G-----------T--T-------------C--------------GT...A------C-- 6364
F1.BR.01.01BR125           -----T--A-------A-----------G-----GG--G--A--------G--G--G----C-A--------------------------------AA-----G-----------T--T-------------C---------GA---T...A--A----C-- 6505
F1.ES.x.P1146              -----T--A-------A------------------G-GG-AGA--------G--G--G----C-GG-------------------------------A-----T------------------A-----------C--------G-----T...A--A----C-- 6388
F1.FI.93.FIN9363           -----T--A-------A-----------G-----GG---A-G--------G--G--G------GG----------------------------------------G-----------T--T-------------C---T------G--GT...A--A----C-- 6383
F2.CM.02.02CM_0016BBY      -----T--T-------A-----------G-----G---A-G--------G-A--GTT--C--G--------------------------------AA-----G-----------A-------------CT--------A--A-T...A--A------- 6234
F2.CM.95.MP257             -----T--T----C--------------G------------A--A--------G-----TT--C-GG----T------------------------AA-----G-----------A-------------C-------------A--...A-GA------- 6256
F2.CM.97.CM53657           -----T--T-------------------G-----G---AGA--------G----GTT----GG-------------------------------AA-------G---------A-------------C-------------G---T...A---------- 6237
G.BE.96.DRCBL              -----T--A------------C-----G--GTGG------A--------------A-----AC-------T-------------------------A---------------------A---C-G--------T--------A----T...A-CA------- 7008
G.CM.01.01CM_4049HAN       -----T-----------------T----G---GAG------AT--------------T------G-------T-------------------------A-----G------------------A---C-G--------T--------A-A...A--AG------ 6255
G.CU.99.Cu74               -----T--A-----G-------------------GGG-----GA--------------------AT-------T-------------------------A-----G-----------T-A---C-G--------CT--------A----A...A--A-G----- 6676
G.ES.99.X138               -----T--A---------------------C-GGG-----RA---------------T-----A--------T----------------G---A--G--GA----------T----AC---C-----T------------A...A--A------- 6498
G.GH.03.03GH175G           -----T--A--------------T-------GGGG------A-----------------T----AC----T--T-------------------------A-----G--------C-----A---C-G--------G-GA----G-----A...A--A-----CA 7124
G.KE.93.HH8793_12_1        -----T--A-G------------T----G-----GG-A---GA--------------------CA--------T-------------------------A-----G------------------A---C-G--------T--------G---A...A--A------A 6454
G.NG.01.01NGPL0669         -----T--A-----G-----A--T----G------G-----AC---------------A-----A--------TT-----------------G---A------------------T-A--AC-G--------G-------------GA...A---------- 6286
G.PT.x.PT2695              -----T--A-------------------G---GGG-----GA--------------TT-----A--------T-------------------------A-----G------------------T-----C-G--C------T--------GG--A...A--A-G----- 7080
G.SE.93.SE6165             -----T--AA-------------T----G---GGG-----GA-----------------------A--------T-------------------------A-----G------------------A--AC-G--------T--------G-A-A...A--AA-G---- 6466
H.BE.93.VI991              -----T--A-------------------G--C---GGG--A-AA--T------------------C-----------------------C------------------------G------------------T---------A--C----------TG---GAG---A------- 6444
H.BE.93.VI997              -----T--A--------------T----G-----C-----A-----T------------------------------------------------------A-----G------------------A--------A--C----------G-C--...--CA------- 6374
H.CF.90.056                -----T--A----C--------------G-----C-----A--A--------------TT------------------T--------------------------A-----G------------------A--------A--C-------------C--...A-CA------- 6373
J.CD.97.J_97DC_KTB147      --------A-------------------G-----G---T-A--------------G-----C-A--------T-------------------------A-----G--G------------A---C-------T--A----G--A---A...---A------A 6254
J.SE.93.SE7887             --------A-------------------------G-----AT----------------T-----C-A--------T----------------C------A-----G------------------A---C------------A---------G---C...A--A------- 6355
K.CD.97.EQTB11C            -----T--A----C--------------G-----------C--A-G-----------G-----C--------------------------------AA-----G-----------T----A------------C-----------------...A--A-T----- 6264
K.CM.96.MP535              --------A----C--------------G-----GG---A-G--------G--------------A-----------------------------------AA-----G------------------------------C-----------------...A--A------- 6250
```

```
B.FR.83.HXB2                 TGTGCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAG...GTAGTAATTAG 7042
Env                          _C__A__P__A__G__F__A__I__L__K__C__N__N__K__T__F__N__G__T__G__P__C__T__N__V__S__T__V__Q__C__T__H__G__I__R__P__V__V__S__T__Q__L__L__L__N__G__S__L__A__E__E__E__.__V__V__I__R
01_AE.CF.90.90CF11697        ---A-T--A-------A-------T----G------G-G---AT-------------G-----C-A---------T---------C-------------A-----G--G---------T--T-------------------------C...A--A----C-- 7012
01_AE.CN.05.FJ051            ---A-T--A-------A--T----T----G--C---G----A-AT--------G----G-------A---------T---------CG---------------G--G-C------C-----------G------------------------...A--A----C-- 7048
01_AE.CN.06.FJ054            ---A-T----------A-------T-T--G------G-----AT--------G----G-------A---------T---------C-------------A-----G-------------T------------------T------G---A...A--A----C-- 7087
01_AE.HK.x.HK001             ---A-T--A-------A--T----T----G--C---G----A-AA---G---G----G-------AC----A--T---------C-------------A-----G--G---------T---A----------C----------------...A--A------- 6418
01_AE.TH.01.01TH_R2184       ---A-T--A-------A-------T---GG------G-G---AT---C---G-------------A---------T---------C-------------A-----G-------------T-------------------------T---...A--A----C-- 6257
01_AE.TH.02.OUR769I          --CA-T--A-------A---C--T----G------G-----A--------G----G-T----AG--------T---------C-------------A-----G-------------T---A-------------------------...A--A----C-- 6244
01_AE.TH.90.CM240            ---A-T--A-------A-------T----G------G-----AT--------G----G-------A---------T---------C-------------A-----G-------------T---------------------------...A--A----C-- 6622
01_AE.US.00.00US_MSC1164     ---A-T--A-------A-------T----G------G-----AT--------G----G-------AC--------T---------C-------------A-----G-------------T--T---------T---T---------T---...A--A----C-- 6253
02_AG.CM.02.02CM_4082STN     --------A-----------A-------G----GAG------GT-------------GT----C-A---------------------------C-A------------------------------------AC-------------...A---C------ 6229
02_AG.EC.x.ECU41             --------A-----------A---T----G-----------------------A---G-----C-AG---------------G--C----------C-A------A------C-------------------T---------G-...-----G----- 6119
02_AG.GH.03.GHNJ196          --------T-----------A-----G---------G------AT-------------G-T---C-GG--------T---------C----------C-A---------------------------------------AG----...-----G----- 7098
02_AG.NG.01.PL0710           --------A-----------A--------------G-C---GA-----G---------G-----C-AG--------T---------C----------C-A-----G-------------T-------------------------A-T...----C------ 6220
02_AG.NG.x.IBNG              --------A--------------------G----GG---AGG---------------G-----C-A--------------------C----------C-A--------------------------------------G---A...-----G--C-- 6566
02_AG.SE.94.SE7812           --------A-----C-----A-------G------G--G---A--------------G-----C-AG--------------------C----------C-A-------------------T-A----------C--------------...A---G----- 6408
02_AG.SN.98.MP1211           --------A-----------A-------G------G-G---GAT---G--------GT----C-AG--------------------C----------C-A--------G-------------------------------G---T...A---G----- 6235
03_AB.RU.97.KAL153_2         --------A-------------------G-----CG----A-A------C-------G-----------------T---------------------------A-----------------------------------T------------------...----------- 6240
04_cpx.CY.94.CY032           --------A-----------A---T----G------G-G--A-AT----C----TT---G-----C-----------T-G----G---C--T---------A-----G-------------T-------------A--CT-----ACG-------...----------- 6430
05_DF.BE.x.VI1310            -----T--A-------A------------G------G-G--A-A-----G---G-----------C-AG--------------------C-------------A-----G-----------GT----------------C------A------GT...A--A----C-- 6445
06_cpx.AU.96.BFP90           -----T--A---------------T----G----GGG-----GAT--------------------A---------T-----------------G----A-----G-------------T-A---C-G---------T---------G-A-C...A--AC-----A 7114
08_BC.CN.97.97CNGX_6F        --CA-T--A-------A------------G------G------TA-------G-------A---CCAT-----T-----G----------------------A-----G--------------A-----------T--C---------AG----...A--A------- 6230
09_cpx.GH.96.96GH2911        --------A--------------------G----GGG-----GA----------------AT-----A---------TT------------G-----G--------------G-------------T-A---C-G---------T-------------C...A--A------- 6241
10_CD.TZ.96.96TZ_BF061       --------A-----A-----A---A-----------G---G--A------------G--------C-A---CA----T-------G----------G---AA--------G------------T-A--G---------------------------...A--A------- 6415
11_cpx.GR.x.GR17             -----T--A--------------------G--C-GGG-----GA---------------------C-A------------------------------C-A------------------A---------------------------G---...---AGG----- 6315
12_BF.AR.99.ARMA159          ---A-T--A-------A---C--------G------G----A-A---------G--G--G-----CCAG--------------------------G----AA----G-------------T--T-------------C-------------T...A--A------- 7055
13_cpx.CM.96.1849            -----------------------------G------G--G--GA-------------G-----C-GG---------T---------C----------C-A-----------------------------------C----------G---A...A--AAG----- 6464
14_BG.DE.01.9196_01          ------------------------------------G----A-A-------------T----------------------------------------------------------------------------------------------A-----...---A------- 6565
15_01B.TH.99.99TH_MU2079     ---A-------------------------G-----C--C--A-A---T-----------------------------T-------------------------------------------------------------T--------------...---A----C-- 6414
16_A2D.KR.97.97KR004         --------A-----------A--------G-----GG-CCC-GAA-------------G-----C-AG----T--T---------------------------G-C---T-----T-----C-G-------------------G-A-TAAAAC-A------- 6414
17_BF.PE.02.PE02_PCR0155     --C--T--A---------------------------G----A-A---------G--G--------C-AGG--------------------------AA----G------CTCATTG-TG-T...--------CT-------------T...A--A----C-- 6253
18_cpx.CU.99.CU76            -----T--A--------------------G--C--C---C-AC----T----------G-----C-G-----T-------------------------A-----G-------------T-------------A---------AC----A--...---A-G--C-- 6350
19_cpx.CU.99.CU7             --------A--------------------G------G-----GA-------------GAA---CCGG-------------------C----------C-A--------G----------------------------------G-A--...--GAC------ 6186
20_BG.CU.99.Cu103            -----T--A-----G-----A-------------GGG-----GA---------------------A---------T--------------------------A-----G-------------T-A---C-G-------CT--------A-G--A...A--A-G----- 6512
21_A2D.KE.91.KNH1254         --------A-------------------------GAG-----CA-------------C-----C-A-----------------------------------A-----G--G-----------------G---------T-------G------...A-CA------- 6261
22_01A1.CM.01.01CM_0001BBY   --------A--------------------G-----AG-G--AGA-------------G-----C-AC--------T---------C-------------A-----G----------------T----G--------------------A------...A--A-G----- 6233
23_BG.CU.03.CB118            -----T--------------A-------------GGG-----GA---------------------G------T--T-------------------------A-----G-------------T-A---C-G-------CT--------A-G--A...A--A-G----- 6512
24_BG.CU.03.CB378            -----T--A-----G-A---A-------------GGG--G--GA----------A----------G--------T-------------------------A-----G-------------T-A---C-G-------CT-------CA-G-GA...A--A-G----- 6500
24_BG.CU.03.CB471            -----T--A-----G-----A-------------GGG--GC-GA---------------------A---------T-------------------------A-----G-------------T-A---C-G-------CT--------A-G--A...---A-G----- 6512
25_cpx.CM.01.101BA           -----T--A--------------------G-----GG------A-------------G-------A---------T----G-----C----------C-A--------G-----A------T--C---------CT------------------...A---------- 6304
27_cpx.CD.97.97CDKTB49       --------A---------------T----G-------------T-------------G----C------------T---------C-------------AA----G-------------T-A-------------------T------------...---A--G---- 6258
28_BF.BR.99.BREPM12609       --------------------C-------G--------------A--T--------------G-----------------------------------------------------G------------C-----------------------------...A--A------- 6416
29_BF.BR.02.BREPM119         --------------C----------G--G------G-----AA----------GT-----------------------------------------------------------------------G-----------------------------------...---A------- 6206
31_BC.BR.02.110PA            -----T--A--G----A------------G-----A--------AT------A-------------C-AT-----------------G---------------A------A---------------A------------T---------------T---...A--A------- 6532
32_06A1.EE.01.EE0369         -----T--A---------------T----G----GAG-----AT---------------------A---------T----------------------C-A-----G-------------T-A---C-----------T------G----AGT...A--A------- 6713
33_01B.MY.05.05MYKL007_1     ---A-T--A-------A-------T----G------G-----AT--------G----G-----C-A---------T---------C-------------A-----G-------------T-------------T---------------------...A--A-G--C-- 6442
34_01B.TH.99.OUR2478P        ---A-T--A-------A---------C--G------G-----AT--------G----G-------A---------T---------C-------------A-----G-------------T-------------------------------------...A--A----C-- 6235
35_AD.AF.05.05AF095          --------A-----------A--C-----------G-G--T-A--------------GT----C-A---------T---------C----------C-A-----G-------------------A-----G-----T-----------C-G-----...A--A--C---- 6220
36_cpx.CM.00.00CMNYU830      --------A-----------A-------G-----AG---ACA--------------G-T---C-A---------T---------C-------G---A-----G-------------T-A---------------------A--A-T--A...AC-A----C-- 6254
37_cpx.CM.00.00CMNYU926      --------A-------------------------G-----GAA-A-----------G-----C-GG--------T---------C----------C-------G-------------T----------------------------G-A--...---A-G----- 6236
38_BF1.UY.05.UY05_4752       -----T--A-----A-A---A------------------T--A-------G--G-------C------------------------------------AA--A-G--G---------T--T-------------C---T------G---A...A--A-G--C-- 6492
39_BF.BR.03.03BRRJ103        -----T--A--------------------G------G----A-A-------G--G--G-----C-A----------------------------------AA----G-------------T--T--------------------------AGT...A--A----C-- 6550
40_BF.BR.04.04BRRJ115        -----T--A-------A------------G--------------A--A-------G----G-----C----------------------------------AA----G-------------T--T-------------C---------A---T...A------C-- 6611
42_BF.LU.03.luBF_05_03       -----------------------------G------G----A-A---------------------A-------------------------------------------G-------------------------------CY-------------T...A--A----C-- 6582
43_02G.SA.03.J11223          -----T--A-----G-----R--------G-----GG-A----AT--------G-------A----A--------TT---------------------------AA----G-------------T-A--AC-------T---T-------------A...A---------- 6570
U.CD.83.83CD003              --------T--------------------G------G------A---------------------C-A--------------------C-------------------G--G---------T-A--A------------T--T--------------...---A------- 6524
U.CD.90.90CD121E12           --------A-------------------------G-----GA---T---------------------C-G--------------------C-------------------G--G---------T-A--A------------T--T----A--------...---A------- 6210
U.GR.99.GR303                ---A---A-------------------GG------G-GG--A--------------G-----C-A---------TT---------C----------C-A---------------------------C---------------A---G-------...---A------- 6335
U.NL.95.U_NL_95_H10986_D1    --------A-------------------------GG----A-G-------------G-----C-A-----------------------------------A-----G--G-----C--GT-A---C----------------------------A...A--A-G----- 6939
CPZ.CD.90.ANT                -----T--A--A--A-A---T--AT-G-G----G-AG--G--GAT--T-CA---GT---GATG----A-------TCAGT------T--C--T--------A--C---A-G--GG----ATGGT-A--A--------A-C-TACCA-AC-A-CACTTCA--------G-A 6498
CPZ.CM.01.SIVcpzCAM13        -----A--A--A--A-----A-----C---------G--GTA-AC--TTCA--G-A---AA----G------A--------C--T--------------C-AA---------A----C---T--A-CA-T-----G---------TAC----A-C...A--AC-G-A-- 6653
CPZ.CM.05.SIVcpzEK505        -----A--AC-A--A-----C---A-----------G--C-AGAT--T--------T--GGA----AC-----A--T--T-----------C--------A-A-----GA----C--A--GT--A-CC------------CT------AC-TC-A-T...A-C------A-- 6591
CPZ.CM.05.SIVcpzLB7          -----T--A-----------TC------GG-----C-----A--A-AT-----T-GT--T-------AC-----T--T--------C-----T----------AA----TA-T--T------T--A-CC-G-----T------------------A-A...A-CA-G----- 6589
CPZ.CM.05.SIVcpzMB66         -----T--A-----A-----CC-CT-G--G------G----A-AT--T-C---G-T-----AG----A------T--T--T--G--T----------C--C--A--T-----C--------GT-T--T--------A-CCT---A----A-G-T-ACA--CC--GA--- 6690
CPZ.CM.05.SIVcpzMT145        -----A--A--A--A-----AC-C------------G---TAGAT-A---A--G-AT-A-A-T----A------A--T--------T----------C----AA-----G-C-A-C--C------A-T--G-----G---AC------CA-TC--...ACT--GGCC-- 6572
CPZ.GA.88.GAB1               -----A--A--A--G-----A-----T------C---G----AGAC--TTCA----A---TAA------T-----A--T-----G--T----------------AA-----G--TA-C-----GT-A--TA-T-----A---T----T----G-A-C...A--ACTG-A-- 7103
CPZ.TZ.00.TAN1               -----T--A--A--C-----CC-GT---------GAG--C---AT----CA--G-A-----A---CT-C------TCAGT---T--C-----------G---TAT--TA-GA--G-C--AGC-T-A-AC--------GTCC--G-A----------A...AC-AA-GC-TA 6686
CPZ.US.85.CPZUS              -----A--AC-A--C-----C--C------------G-A--AGAT-----A----AG---GAG----A------T--T-----G--G-----T-----C--A-AA-----G--GA-T--A--G--CA-AA-----------T-----ACTA--A-T...--TACTG-A-- 7086
CPZ.CM.x.SIVcpzDP943         -----A--CC-A--------A--CA-G-----C---G-GCCAGAT--T-------------AA-----G----A-A--T--------------T-----C--C-A------G---A-T--A--RT-AA-TC-T-----A---------T--GA-CAGC...ACGAYTGY--- 6611
CPZ.CM.x.SIVcpzMB897         -----A--A-----A-----CC--T-----------G-CC---AT-----G----A---GA------G------T--------G--T-----------T----A---T-----C--T--C---T-C--C--G-----A--C--G-------G-A-C...AAGACTG-AGT 6592
N.CM.02.DJO0131              -----A--AC-A--C-----C---A-G---------G-AGCA-AT--T----------GGA-----A------G--T--T--------C-----------A-A-----GA----C-----GT-AA-CC-------G--CT---ATA--...--T...A-T--T----- 6523
N.CM.04.04CM_1015_04         -----A--AC-A--C-----C---A-G-R-------G-AGGA-AC--T-G-----A--GGA-----A------G--T--T--------C-----------A-A---ACGA----T-----GT-AA-CM-------G--C----ATAC---T--T...A-T--TM---- 6519
N.CM.04.04CM_1131_03         -----A--AC-A--C-----C---A-G-G-------G-AGGA-AC--T-G-----A--GGA-----A------G--T--T--------C-----------A-A-----GA----C-----GT-AA-CA-G-----G--C----AT--GAC---T...A-T--T----- 6554
N.CM.95.YBF30                -----A--AC-A--C-----C---A-G---------G-AGGA-AC--T-G-----AT---AGC-----------G--T--T--------C-----------A-A-----GA----C-----GT-AA-CC-------A--CT--AATAC---T-GA...A-T--T----- 6635
N.CM.97.YBF106               -----A--AC-A--C-----C---A-G---------G-AGGA-AT--T-G-----AT---AGC-----------G--T--T--------C-----------A-A-----GA----C-----GT-AA-CC-------A--CT---ATAC---T--T...A-T--T----- 6608
O.BE.87.ANT70                -----T--A--A--A-A---T--CT-T--G-----C-GC-CAGAA--T-----------CA----C-G---CA-A-CGGT---TACT-----------C---------AC----AGT-----G--AA-A--------G-CA--CT-TA---G-A-A...A--AG---G-T 7077
O.CM.91.MVP5180              -----T--AA-A--A-A---C--CT-T--G------G-C-CAGAC--T-----------C-T---CCAC---A-TTCAGTG--TACT-----------C--C-A----AC----AGT--------AA-AC-G-----G-CA--CT-TAG----A--...A--AG-----T 7076
O.CM.96.96CMABB637           -----T--A--A--A-A---C--CT-T--G------G-C-CAG-A--T-----------CTT---C-AG--CA-TTCAGT---TACT-----C-----T--C-A----AC----AGT--------AA-AC-------A-CAA-CT-T----G-A-A...A--AG---GCT 6564
O.CM.98.98CMA104             -----T--A--A--A-A---T--CT-T--G-----C----TAGAA--T-----------CA----C-AC--CA-TTCAGT---TACT-----------C--C-A----AC-A--AGT--------AA-AC-------G-CA--CT-CA---GGA-A...A--AG-----T 6533
O.CM.99.99CMU4122            -----T--A-GA--A-A---C--CT-T--G------G-C-CAGAA--T-----------T-A---C-G---CA-T-CAGT---TACT--C--------C----A----AC----AGT--------AG-A--------G-CA--CT-T----G-A-A...A--AG-----T 6514
O.FR.92.VAU                  -----T--A--G--A-A---C--CT-T--G------G-A-CAGGA--T-----------T-TC----A---C--T-CAGT---TACT-----------C--C-AA--AC----AGT--C-----AA-AC------GGCA--CT-TA-G-GGA-T...A--AC---C-T 6616
O.SN.99.SEMP1299             -----T--A--G--A-A---C--CT-T--G-----C-GC-CAGA---T-----------C-----C-AC--CA-T-CAG----TACT-----------C--C-AA--AC----AGT--------AA-A--------G-CA--CT-T----G-A-C...A--AG-----T 7115
O.US.99.99USTWLA             -----T--A--A--A-A---T--CT-T-----C---G-A-CAGG---T-----------C-------A---CA-TTCAGT---TACC-----------C--C-A----AC----AGC-----C--AA-A--T-----G-CAA-CT-TA---G-A-A...A--AG-----T 6536
```

```
                                                                 V3 loop start
B.FR.83.HXB2      ATCTGTCAATTTCACGGACAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTAGAAATTAATTGTACAAGACCCAACAACAATACAAGAAAAAGAATCCGTATCCAGAGAGGACCAGGGAGAGCATTT.........GTTACAATAGGA...AAAATAGGAAATA 7200
Env               __S__V__N__F__T__D__N__A__K__T__I__I__V__Q__L__N__T__S__V__E__I__N__C__T__R__P__N__N__N__T__R__K__R__I__R__I__Q__R__G__P__G__R__A__F__._____.__V__T__I__G__.__K__I__G__N__

A1.GE.99.99GEMZ011
A1.KE.00.KER2008
A1.KE.00.KNH1144
A1.KE.00.KSM4024
A1.KE.00.MSA4069
A1.KE.00.NKU3005
A1.RU.00.RU00051
A1.RW.93.93RW_024
A1.SE.95.SE889I
A1.SE.95.UGSE8131
A1.TZ.01.A173
A1.UA.01.01UADN139
A1.UG.92.92UG037
A1.UG.99.99UGA07072
A2.CD.97.97CDKTB48
A2.CY.94.94CY017_41
A.SN.01.DDI579
A.SN.01.DDI369
A.SN.96.DDI360
A.CD.02.02CD_KTB035
A.CD.97.97CD_KCC2
A.CD.97.97CD_KTB13

B.AR.04.04AR151516
B.AU.87.MBC925
B.BO.99.BOL0122
B.BR.03.BREPM2012
B.CA.97.CANB3FULL
B.CN.05.05CNHB_hp3
B.CO.01.PCM001
B.GB.83.CAM1
B.GB.86.GB8
B.GE.03.03GEMZ010
B.IT.05.SG1
B.JP.05.DR6538
B.KR.05.05CSR3
B.NL.00.671_00T36
B.RU.04.04RU128005
B.TH.00.00TH_C3198
B.UA.01.01UAKV167
B.US.04.ES10_53
B.US.99.PRB959_03

C.AR.01.ARG4006
C.BR.04.04BR013
C.BW.00.00BW07621
C.CN.98.YNRL9840
C.ET.02.02ET_288
C.GE.03.03GEMZ033
C.IL.99.99ET7
C.IN.99.01IN565_10
C.KE.00.KER2010
C.MM.99.mIDU101_3
C.MW.93.93MW_965
C.SN.90.90SE_364
C.SO.89.89SM_145
C.TZ.02.CO178
C.UY.01.TRA3011
C.YE.02.02YE511
C.ZA.04.04ZASK164B1
C.ZA.05.05ZAPSK240B1
C.ZM.02.02ZM108

D.CD.83.ELI
D.CM.01.01CM_0009BBY
D.KE.01.NKU3006
D.KR.04.04KBH8
D.TD.99.MN011
D.TZ.01.A280
D.UG.99.99UGK09259
D.YE.01.01YE386
D.YE.02.02YE516
D.ZA.90.R1

F1.AR.02.ARE933
F1.BE.93.VI850
F1.BR.01.01BR125
F1.ES.x.P1146
F1.FI.93.FIN9363
F2.CM.02.02CM_0016BBY
F2.CM.95.MP257
F2.CM.97.CM53657

G.BE.96.DRCBL
G.CM.01.01CM_4049HAN
G.CU.99.Cu74
G.ES.99.X138
G.GH.03.03GH175G
G.KE.93.HH8793_12_1
G.NG.01.01NGPL0669
G.PT.x.PT2695
G.SE.93.SE6165

H.BE.93.VI991
H.BE.93.VI997
H.CF.90.056

J.CD.97.J_97DC_KTB147
J.SE.93.SE7887

K.CD.97.EQTB11C
K.CM.96.MP535
```

```
                                 V3 loop start
B.FR.83.HXB2                 ATCTGTCAATTTCACGGACAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTAGAAATTAATTGTACAAGACCCAACAACAATACAAGAAAAAGAATCCGTATCCAGAGAGGACCAGGGAGAGCATTT.........GTTACAATAGGA...AAAATAGGAAATA 7200
Env                          __S__V__N__F__T__D__N__A__K__T__I__I__V__Q__L__N__T__S__V__E__I__N__C__T__R__P__N__N__N__T__R__K__R__I__R__I__Q__R__G__P__G__R__A__F__.__.__.__V__T__I__G__.__K__I__G__N__
01_AE.CF.90.90CF11697        -----AA---C----AA-------C-------------G--C--T--T-A-----------C--------C------TT---A---TG----C---TGCA......-GT-T---------ACA--T---C.........TA--A--C----AGC-T--C----G-C- 7167
01_AE.CN.05.FJ051            -----AG---C----A--------C-------------G--C--T----G-------AC---C--------C------TC---------------C---T--A.....ACT-T----------CA-CTG--C.........TA--G--C----GAA-T-------G-C- 7203
01_AE.CN.06.FJ054            -----AA---C----AA-------C-------------G--C--T--T-A---A--------C--------C------TC---------------C---T--A......-T-T---------ACA--T---C.........TA--G--C----GAA-T--C----G-C- 7242
01_AE.HK.x.HK001             -----AA---------AA-T-----C-------------G--C--T--TGA-------C----C-C------C------TC-------------G----TG-G......-GT-T---------ACA--T---C.........TA--G--C----GAA-T--C----G--- 6573
01_AE.TH.01.01TH_R2184       -----AA---C----A------A-C-------------G--C--T--T-A---------G--C---------C------TC---------------C---T--G.....ACT-T---------ACA--T---C.........TA--G--C----GAC-T-G---------- 6412
01_AE.TH.02.OUR769I          -----AA---A----AA--------------------G--C--T--T-A------------C-----C--C------TC---------------C---T--A......--T-T---------ACA--T---C.........TA--G--C----GAC-T-G-----G--- 6399
01_AE.TH.90.CM240            -----AAG--C----AA-------C-------------G--C--T--T-A------------C--------C------TC---------------C---T--A.....ACT-T---------AC---T---C.........TA--G--C----GAT-T----------- 6777
01_AE.US.00.00US_MSC1164     -----AA---C----AA-------C-------------G--C--T--T-A---G--------C--------C------TC---------------C-G-T--A.....ACT-T---------AGC--T---C.........TA--G--C----GAT-T--C----G--- 6408
02_AG.CM.02.02CM_4082STN     -----AA-G-A----AA-------C-----------------T-AGTT-AGC-----A-----G------TC-----T----------------G--T--G......-GT-T----------CA-A----C.........TA-G---C----GAC-T----------- 6384
02_AG.EC.x.ECU41             -----AG---A----A-------------------------T--G-T---C-----AG---A--------C-----T-------------G-----TG-A......-GT-T----------CA-A----C.........TA-G---C---TGAC-T------GG--- 6274
02_AG.GH.03.GHNJ196          -----AA------T-A--T-----C-----------------T-AGCT-AGC-----C-------------C-----T----------------C-G-T--A.......-T-TG---T----C--A----C.........TA-G---C---TGAC-T------GG--- 7253
02_AG.NG.01.PL0710           -----AA---A----AA-------C------T-----------T--GTT-ATC-----AG------C-----C-----T------------T----G-TG-A......-GT-T----------CA-A----C.........TA-G---C---TAAA-T------G---- 6375
02_AG.NG.x.IBNG              -----AA---A----AA-------C-----------------T--GCT-ATC-----A-------------C-----T-----------------G-TG-A.......-T-T----------CA------C.........TA-G---C---TGAC-T------GG--- 6721
02_AG.SE.94.SE7812           -----AA-----T--AA-------C----T------------T--C-TGA------A-------------C-----TGG-----------------TG-A......-GT-T----------CA-A----C.........TA-G---C---TGAC-T------GG--- 6563
02_AG.SN.98.MP1211           -----AA---A--T-AA-------C-----------------T-----T-AGC-----TGG-----------C-----T-G-----------------TG-A......-GT-T----------CA-A----C.........TA-----C---TGCAGT------GG-C- 6390
03_AB.RU.97.KAL153_2         --------------------A--------------------------AGA-C----G-----------------------------------------G-T--T.......-T-T---------------------.........TA-G---C----GAC-T--C----G--- 6395
04_cpx.CY.94.CY032           ----AAA---A----A------A-C----AT--------------TGCA-AGG-----A-------------C-----TGG------------------TG-A.......-T-T---G-----ACT-A---GG.........TA-G---C---TGAA-T-------G--- 6585
05_DF.BE.x.VI1310            ----CAA---A--T-A--T-----A----A-----------C--T--TGA-------C-T--------------G---------------------T--A.......-TTT---------ACA------C.........TA-G---C---TGAC-T-------G-C- 6600
06_cpx.AU.96.BFP90           -A---AA--CA----A------A-C----AT--------G-----T--TCAGC------------GG-----C------GGT------------------T--A.....TCCTTT--------ACA---G--C.........A--G---C---TGAC-T-------G--- 7269
08_BC.CN.97.97CNGX_6F        -----AA---C-G--AA------TC-----A-----------T--T--TCA-------------GTA-----------------T----------------T--A.....AG--T---------ACA-A----C.........TA-G---C----GAC-TC------G-C- 6385
09_cpx.GH.96.96GH2911        -----AG---A----AA------TC---GT------------T--GCTGATC-----ACG---G-----GTC---A-TGG--------------C---TG-A......-GT-T---------ACA-A-G--C.........TA-G---C---TGAC-T-------G--- 6396
10_CD.TZ.96.96TZ_BF061       -----AA---C----A--T----------AT--------G-----T--TGAG------AC---------CGTG-----------T----------G---T--A.....AG--T---------ACA-A----C.........TA-G---C-...GAC-T---------C- 6567
11_cpx.GR.x.GR17             -----AA--CC----AA------------------------------T----GTA-----AG-----------C-----T-----------------GG-T--A.......-T-T---------ACA---T---.........TA-G---C---TGAT-T------GG--- 6470
12_BF.AR.99.ARMA159          ----CAA--CA--T-A--T---A-A-------------T--T--T--TGA-------C-G-----C-----------------T-----------------T--A........--TT---------AC-G------.........TA-G---C----GAC-T------GG-C- 7210
13_cpx.CM.96.1849            -G--AAA--C---T-A-------T-GG--A---------------TGCT--GC-----AG-----------C-----T--------------------T--A......-G--T---------ACA------C.........TA-G---C---GGAT-T-------G--- 6619
14_BG.DE.01.9196_01          G----AA----------------------------------------A--T-A-------T---------YT--------------------------T--A......-CT-T-----------------------.........TA-G---C----CAA-T-------G-C- 6720
15_01B.TH.99.99TH_MU2079     ----AG---C------A-----------T--------------A--TGA-------------------------------------------------T--A.......--T-T---T---------T---GG.........TA-----C----GAA-T-------G--- 6569
16_A2D.KR.97.97KR004         -----CA---A-T--A-----A-----A---------------T-T-CT-AGC-----C-------------C------G--GT-GG-CA-----GG--TG-A...CG-ATTG--CC-GG-C-AGCCTTC-A-ACAAGGCAAACATAT-C-A-GCAAGC--A---GG--- 6581
17_BF.PE.02.PE02_PCR0155     ----CAA---A--TTC--T---A-A--G-------------T--T--TGA-------C-G----C-----------------T-----------------T--A......-G--T----A---ACA--------.........TA-----C----GRA-T-------G-C- 6408
18_cpx.CU.99.CU76            -----AA---A----A------A-C--G-A--------------T-CG--GC-----A-C----CC-----C-------G---T--C--G----G---T--A.......--T-T---------ACA------C.........TA-G---C---TGAC-T-------G--- 6505
19_cpx.CU.99.CU7             -----AA---A----A-------C------------------T-T-CTRAGC-----A--------------C----T-----------------C---TG-A......-GT-T---------ACA--T----.........TA-G---C----............-----GG-C- 6335
20_BG.CU.99.Cu103            -----AA--CC----A--------C---GT---------G-----T--T-A-AG-A------------------C---------------------G--T--A.....GGATTT--------ACA---G--C.........TA-G---C---TGAC-T--C----G--- 6667
21_A2D.KE.91.KNH1254         -A---AA---A---GCAGG----AA----G---------------T--TGAG------AC---------------G-----T------------G-G-T--A.......-T-T----------CA----A-A.........TA-G---CC--TAAG-T---------C- 6416
22_01A1.CM.01.01CM_0001BBY   -----AA--------AA----C--C----T------------TY-GT--AGC-----ACT----------GC-----TGG---------------C--T--A......-GT-T---------ACA--TC--C.........TA-G---GG--TGAT-T------GG--- 6388
23_BG.CU.03.CB118            -----AA--CC----AA-------C---GTT--------G-----T--T-GCAG-A--AC-----------C------------------------G--T--A.....ACATTT--------ACA------C.........TA-G---C---TGAC-T----------- 6667
24_BG.CU.03.CB378            -----AA--CC----AA-------C---GT---------G---T-T--T-A-AG-A----T-----------C-----TGG--------------G--T--A.....AGATTT--------ACA---G--C.........TA-G---C---TGCC-T------CG--- 6655
24_BG.CU.03.CB471            -----AA--CC----AA-------C---GT---------G---T-T--T-A-AG-A--A-------------T---------------------GG--TG-A.....GGA-TT--------ACA--TG--C.........TA-G---C---TGAC-T-------G--- 6667
25_cpx.CM.01.101BA           -----AA---A-TT-AA-------C----G---------G--TT-T--TGAGA-----AC-----------TC-----T-------------------TG-A......-GT-T---G-----ACA-A----C.........TA-G---C---TGAT-T-------G--- 6459
27_cpx.CD.97.97CDKTB49       --T--AA---C----AA-------C------------------C--T--T-A-------AT-----------C------TC--------------C---TG-A.....AGT-T---------ACA--T---A......TTCTA--G-GC----GAC-T-------G--- 6416
28_BF.BR.99.BREPM12609       -----AA---C----A------------T---------G---------TGA-A--------------------G------------G--AC---C-T--A.......--TCT-----GG------AT-A-G.........TA-GT--C--A-CGA-T-CA----GG-- 6571
29_BF.BR.02.BREPM119         -----AG---A-----------T----GT--------------------TCA-----------------------G-------------------------T--A.......--T-TG---TGG-----------------.........TAC----C----GAA-T-----ATG--- 6361
31_BC.BR.02.110PA            ----AAA---A-A--A-------TCC----A--------------T--GGA-C-----------------G----------T--------G----T--A.....AGA-T---T-----ACA-A----C.........TA-G---C----GAC-T-------G--- 6687
32_06A1.EE.01.EE0369         -----AA--CA----A-----A-C-------------G-----T--TCA---------C----TA-----C------T-T--T-----------G---T--A.......--C-TT----------CA---G---.........T--G---C---TGAC-T-------G--- 6868
33_01B.MY.05.05MYKL007_1     -----AA---C----A--T---A-C----A---------G--C--T--T-A-----------C---------------TC-----G--T-----C---T--G.....ACT-T---------ACAC-T---C.........TA--G--C----GAC-T-G-----G--- 6597
34_01B.TH.99.OUR2478P        -----AA---A----AA-------A-------------G--C--T--T-A-----------C--------C------TC---------------C---TG-A.......-T-T---------ACA------C.........TA--A-------GAC-T-------G--- 6390
35_AD.AF.05.05AF095          -----AA---A--T-A---------------------------T--T--TGAG------AC-----------C-----T-G-----------------T--A......-GT-T---------ACA-A----C.........TA-G---C--AT...-T------GG--- 6372
36_cpx.CM.00.00CMNYU830      -----AA---A----AA-------C--G---------A-----T--GT--AGC-----A-T----C----GTC--G--T-----------------G-T-CA......-CT-T---------ACA---T--C.........TA-G---G-AATGATGT------GG-C- 6409
37_cpx.CM.00.00CMNYU926      -----AA---A----AA----C--C-----------------AT-T--TGAGA------------------C--GTTT---------C-----------TG-A......-GC-T---------ACA-A----C.........T--G---C---TGAC-T------GG-C- 6391
38_BF1.UY.05.UY05_4752       ----CAA---A--T-A--T-----A-----------------T--T--TGA-------C-T----------------------T--------G-----T--A.......--T-T---------A--------C.........TA-----C----GAT-T-------G-C- 6647
39_BF.BR.03.03BRRJ103        ----CAA---A-AT-A--T---A-A-----------------C--T--TGA-C-----C-G-----------------A----G-------------G-C--A.......--TTT---------AC--------.........TA-----C----GAA-T-------G-C- 6705
40_BF.BR.04.04BRRJ115        G---CAA---C--T-A--T-----A----T------------C--T--TGAG---A--C-G-------------------------------------T--A.....GGT-T---------AC--------.........TA-G---C----GAC-T-------G-C- 6766
42_BF.LU.03.luBF_05_03       ----CAA---A-YT-A--------A-----------------C--T--TGA-------C-G-------C--G-----------T-------------T--A.......--T-T--C------AC--------.........TA-G---C----GAC-T-------G-C- 6737
43_02G.SA.03.J11223          ----AAA--C---T-A--------CC--GT-C-------G-----T--TRGC---A------------C--C------------------R-----T--A.....GGATTC--T--T--ACA-A-G--C.........TACG---C---TGAT-T-------G--- 6725
U.CD.83.83CD003              -----AA---A----AA-------C------------------T--TGAGA-----A------------C------GGATC-G-C-AG-AG-T----CAAA----A-GT-T---------A-A--T---C.........TA-G---A---TGGA-T--C-----.... 6681
U.CD.90.90CD121E12           -----AA---A----AA-------C-----------------T-T--TGAGA-----A-----------GC--------T--T--------GC--G-T--A.......--T-T---------AC---T---C.........TA-G---C-A-TAAC-T-------G--- 6365
U.GR.99.GR303                -----AA--------A--------------------------T--TGAG------TC------------C-----T-G-----------------T--A.......--T-T---------ACA---G--C.........T--G---C--AC...-T-------G--- 6487
U.NL.95.U_NL_95_H10986_D1    -A---AA--CA----A------A-C--G-T--------------T--TGA-A--A--AGCC-G--------C--------------------------T--A......-CC-TG--------AC---G--C.........TA-G---C---TGAC-T-------G--- 7094
CPZ.CD.90.ANT                TGG-CG---AAATGAATCTGTGCT-GT--GAT-TGG-AA-G-AT-CG-A-AC-TA...AC-----CA----T------AGGA--T-GG---GT--G--ATC-A........-A-T---------A-TGA-T--C.........TA--ACG---A-ATAGC--C----G-C- 6650
CPZ.CM.01.SIVcpzCAM13        -GTAAA----GC-T-TA-A---A-CC-TGATTGG--------A--A-G----G-A--GA-TC-G-CG----AG--GGTAGGG--T--C------GGC-A-G-A...........T---------A-TGA-C--C.........TA--AT--G-ACCAT-T-T-T--TG--- 6808
CPZ.CM.05.SIVcpzEK505        -AA-......AATT-AA-A......G-C--AC-CT-G--G--AT-A--TGA-AG-A-CCC---A-----C--------AGGA--T--G------GG-CA-G-A........-A-T------T--T-TGA----C.........TA--AC----A-AAT-T-------G-C- 6734
CPZ.CM.05.SIVcpzLB7          G-AC......AAGG-T.........G-T-AGG-TC----------C--T---AG-A--C-C--C--------T--GGTAGGA--T--G----T----G------......-CCCT---------ACAGCTG--C.........TA-GG--C--A-ACTGT-G-----G-C- 6729
CPZ.CM.05.SIVcpzMB66         GAA-.......G--T-AA-TG--ATGG-C--T--------GA-AT-A---GA-A-----CGTC-C-----C--T---A-AGGA--T--C----T----G-T--A......-CT-T-------A-TCAGAT----.........TA-GG-----A-ACTGT-------G-C- 6839
CPZ.CM.05.SIVcpzMT145        -ATA......GA-C-TTCA...-AG--TTTAGC----A-------C--AGATC-A---A-G----CA----GG-----AGGA--T--C------GGTCA---A.........-A-T---------CA-TGA-T---.........TA--AC----A-AATGT-G----GG-C- 6718
CPZ.GA.88.GAB1               -GTG-AG---AAGT-AA-A---A--G-TGT-TGG------------AGTAGA-G-A--TAGTC-A------CAT-----AGGA-----------GGG-GAGG-G..........---T---G-----A-TGA-C---.........TA--AT----A-AATGT-G-----G--- 7258
CPZ.TZ.00.TAN1               C-T-...GT-AAT--CTCAGT-AA--C-C--T--T-----A-AT-T--TGT---AA--A-TT-A-CG---GA----A-AGGA-----------GGTCA-G-A..........---T---T-----T-TGA-C---.........TA--AT----A-AATGT-G----GG-C- 6838
CPZ.US.85.CPZUS              -AG-AAA--C--TG-A----T-..........--TC-------AT-CTCAGAGGGA--CA-T--G-C-----TT-----AGGA---------GTGGG--ATG-A........-ACT---------A-TGA-T---.........TA--AC---CC-AAG-T-G-----G--G 7232
CPZ.CM.x.SIVcpzDP943         GA---AAS-CA-G-MAR-A-TG-A-R-G-----C---------T-TC-T-A-AGG--CC-G--G-CA----T------AGGA--T--C-GC---GG-CAG--A........-A-T---------CA-TGT-----.........TA--ACC---A-AAT-TT------G--- 6766
CPZ.CM.x.SIVcpzMB897         TAGG--A-GAAGT-AATCT---A-AG-G--T--C--T--G-----AGCT---G-AA--T-T--C--------C----TAGGT--T--A--C-T-G-GG-C--A......-CT-T---G-----ACA-ATT---.........TA--GG-CTAA-ACGGTGG----CG-C- 6747
N.CM.02.DJO0131              -AA-AAT-G-GGTGGA............-ATC-GT-G--G--ATG---TGAGATA--GAC---G-----------G--AGGG--T------G--GG-CAGG-A........-A-T------T-CT-TGA-----.........TA--AC----A-AAA-T-G-----G-C- 6666
N.CM.04.04CM_1015_04         -YR-SAT--WGAT...............-AT--GT-G--G--ATG---TMAGA-A--GTC---A-----------G--AGGG--T----Y-G--GG-CAGG-G..........---T------T-CT-TGA-----.........TA--AC----AGAAA-T------GG-C- 6659
N.CM.04.04CM_1131_03         -AA-...G--AGTGGTACT-TATTGGGT-AT---T-G--G--ATG---T-AGA-A---TT---A-----------G--AGGG--T------G--GG-CAGG-G..........---T------T-CT-TGA-----.........TA--AC----AGAAA-T------GG-C- 6706
N.CM.95.YBF30                -AA-...G--..........-G-CAC-GT-ATC-GT-G--G--ATG---TGAGA-A--GCC---A-----------G--AGGA--T------G--GG-CAGG-G..........---T------T-CT-TGA-----.........TA--AC----A-AAA-T-G-----G-C- 6778
N.CM.97.YBF106               -CA-................C---GGGGT-ATC-GT-G--G--ATG---TGAGA-A--GTC---A-----------G--AGGA--T------G--GG-CAGG-G..........---T------T-CT-TGA-----.........TA--AT----A-AAA-T-G-----G-CG 6748
O.BE.87.ANT70                GG-AAAAG--A-TTT---AGG--GA----AT--C-----GACC--A---T-TA-CC--A-C--G-CC---GA------AC-A-TAG-C...-T-C--GAG--G.....AGA-T---T---...-TG--C-GG.........TAC-GC--G---ATAGGGGG-AC-GCAG 7226
O.CM.91.MVP5180              GGGAAAA---A-T--A--ATCA--A--G-AT--C------ACC--A-----TC--A--A-C--G-CC--C-T----GAAGGA-TTGCAGAGGT-C--GAT--A.....T-T-C---T---...-TGAG--GG.........CGC-GT--GAC-CTT----G-A-T--C- 7228
O.CM.96.96CMABB637           GGGAAAA---C-TT-AA-T-C-...G-T-AT--CC----GACC--A--TT-CA--C--A-C--A-CC--CGAG-----AT--G-GC-GTT-GT-C--GA---G.....A---T---T--T...-TG--T-GG.........TAC-GC--G-A-GTAG--CG-AAC--G- 6713
O.CM.98.98CMA104             GGGAAAAG--A-T--A-CTGG-T--CGC-AT--CC----GACT--A--TT-TA--A--A-T--A-CC---GAG-----ATT---TCTG--GGT-G--CAG--G.....A---T---C---...-TG--C-GG.........TAC-GC----ATATAGGG-G------C- 6685
O.CM.99.99CMU4122            GGGAAAA---A-T--A----G--GA----AT--------GACC--AG-TT-TG--A--A-C--A-CC---GAG-----ATGG--TC-GT--GT-C--GAGC-G.....A-A-T---T---...-TGT-C-GG.........TAC-GC--G--CATA-GCTC-CC-CTCT 6666
O.FR.92.VAU                  GGGGAAGG--A-TT-A----G--GGG-G-A---CC-GA--ACC-----T--TGA-A-GAC---AGCA---GAG-----AGGA-GTC-G----T-CG--AG--A.....AT-GC---T---...-TG--T-GG.........TAC-GC--G-CCCTT-GT-ATAC---GG 6768
O.SN.99.SEMP1299             GGGAAAA---A-TT-A------ATG----AT--C-----GACC--A--TT-TA--A--A-C--G-CC---GTG----AAGGA--TC-GT--GT-C--GAG--A.......--A-T---T---...-TG--C-GG.........TAC-GC--GA-CCTGGCCCA--AGGGA- 7267
O.US.99.99USTWLA             GG-AAAA--CA-TT--AG---CT-AG-T-AT--CT----GACC--A--TT-TA--A--A-C--G-CC--C-TG--G--AGGA--T---T--GT-C--GAG--G.....AGA-T---T---...-TG--T-GG.........TAC-GT--G---CTTGGG-AC--GT--- 6688
```

```
                                                V3 loop end┤
B.FR.83.HXB2            TG...............AGACAAGCACATTGTAACATT...AGTAGAGCAAAATGGAATAACACTTTAAAACAGATAGCTAGCAAATTAAGAGAACAATTTGGAATAATAAA...................ACAATAATCTTTAAGCAATCCTCAGGAGGGGACCCAGAA 7334
Env                     M__.__.__.__.__.__R__Q__A__H__C__N__I__.__S__R__A__K__W__N__N__T__L__K__Q__I__A__S__K__L__R__E__Q__F__G__N__N__K__.__.__.__.__.__.__T__I__I__F__K__Q__S__S__G__G__D__P__E_
A1.GE.99.99GEMZ011      -A...............---A-----T-------TG-C...-AC---A--G-----------------C--A----AG--TAC--------A-CT-C---AAT--C--A...................------------G-A-C-------------TTT---- 6515
A1.KE.00.KER2008        -A...............---------T-------TG-C...----A-T--------G--G-T-----G---A-AG---T-TA-C--------A--T-C---AAC...--A...................-------------CTCC-------------TGT---- 6555
A1.KE.00.KNH1144        -A...............-----------------TG-C...-----T--C---------G------C-----G----GAAC---------------C---AA---C--A...................-------A-----CAGT------------T-T---- 6572
A1.KE.00.KSM4024        -A...............-----G-----------TG-C...-A-GC-T--G-T----G-G-AG-----CG-A--G---T-CTA-CTC-TCAG.......................................----A--T-T---- 6472
A1.KE.00.MSA4069        -A...............---------T-------TG-C...---------GC---------A------C--G--G-----AAC--------C--G-----AAT--C--A...................------GA---GCTA-CCA---------A--T-T---- 6554
A1.KE.00.NKU3005        -A...............-----------------TG-C...----A-T--G---------A------C---------T--TAC--------A-G--C-GGAAC...--A...................------A----CTG--C-------------TTT---- 6554
A1.RU.00.RU00051        CA...............---A-------------TG-C...------A--G----GGA---G--AC-G--G-----CG--AAC-----CA-ACTT-C---CC---C--A...................------G------A-C-----------A--TTT---- 6629
A1.RW.93.93RW_024       -A...............---A-------------TG--...------T--G--------G-A------C-----G---T--CTC---------GGT-C-----G--C--G...................-----G------CTA-GC-------------T-T---- 6539
A1.SE.95.SE8891         -A...............---------T-----G-TG-C...-A---A--G--------G-AG-----C--A--G---TC-A-C------A-AC---C---AAG--C--A...................-------------TAGC-------------T-T---- 6511
A1.SE.95.UGSE8131       -A...............---A-------------TG-C...------T-----------G-A----------A--G-G----TAC--------A--T-C-GGAAC...-CA...................------------CTA-C-------------TTT---- 6715
A1.TZ.01.A173           -A...............---------T------CTG-C...-A---A--G---------CA-----GC--A--G---GC-AAC----------TTT--ACTTT------C---...................-------T---CTAGC-------------T-T---- 6546
A1.UA.01.01UADN139      CA...............---A-------------TG-C...--C----T-GC-----------------C--A-----AG--TAC--------A--T-C---AAT--C--A...................------C------A-C-------------TTT---- 6539
A1.UG.92.92UG037        -A...............-----------------TG-C...---G-GT--C-----------A------C-C---G---T-GAAC---------A--T-T-GGAAC--C---...................-------------TAGC-------------TTT---- 7362
A1.UG.99.99UGA07072     -A...............---------T-----G-TG-C...-A---A--G----------A-----GC--G--G----CTA-C-----G--A----C-----G--C--A...................-----------GCTA-C--A---------T-T---- 6552
A2.CD.97.97CDKTB48      -A...............-----------------T--C...-C-T-A--G-------GC----C-----A--G---T-GAAC---------------C--CCCT-----A...................-------------TAGC---------------T---- 6656
A2.CY.94.94CY017_41     -A...............-----------------T--C...-AC-A-A--TT-------G---------C--A--G-----GAAC-----------GA----CCCT--G--A...................--C----------CTA-C---------------------- 6709
A.SN.01.DDI579          -A...............--G--------------TG-C...-------A---T----C-G-G-------C-----G-----AAC------A-AG---C---AAC...--A...................------------CTGGG----------A--TTT---- 6541
A.SN.01.DDJ369          -A...............-----------------TG-C...-------A-G-T---------A------T-TA--G-G-T-GAAC---------A-G--C---AAC--C-GA...................------A----CTGC-C------------TTT---- 6553
A.SN.96.DDJ360          -A...............----------------CTG-C...--A------C---------C-------C-----G-----GA-C---------A----C---AAC...--A...................------A----CTAGGC----------A--TTT---- 6526
A.CD.02.02CD_KTB035     -A...............---A----CTTC----TAG-C...-A-GC-A--G-----G---CA---A---GCAG-G-GAGGGAAC-----C---C---C--CCC------A...................-----G-T---GCTA-C---A------------T---- 7357
A.CD.97.97CD_KCC2       -A...............-----------C--C---C--...------A--GC------CG-T------T-CA-TG--T---AAGCC-----------C--CCC------A...................-------------TA----A-------------T---- 7330
A.CD.97.97CD_KTB13      -A...............--------CT-C--CG-AG-C...-A--A-A--C-------C-TG------C--A--G-----GA-C---------A--TTC--CCCG--C--A...................------A------T---A-------------TT---- 6616
B.AR.04.04AR151516      -A...............---------------------...-------A---------C----------G----C-------A--------G---------AAT-CG-CA...................---G----------T--C-------------T------ 6546
B.AU.87.MBC925          -A...............--G-----T------C-C---...-A-------GG------G--------------------GAA------G-----------AAG-----A...................-----G-A-----T------------------------ 7356
B.BO.99.BOL0122         -A...............--G------------------...---GC-A--G-------------------G---------TC-CA------G------------GGTA-------...................-------------T------------------------ 6526
B.BR.03.BREPM2012       -A...............-----------------T---...---G--A----T----C-----------GG---G-----CAA------GA----TTTAA-...-----A...................-----G------GT-C---T----------------- 6788
B.CA.97.CANB3FULL       -A...............---A-----------------...--------------------G-----G---------TG-AA------G-----------AAG-----A...................-----G-------T------------------------ 6616
B.CN.05.05CNHB_hp3      -A...............---A-------------C---...------C--C-----G--G-----C---G----G----T-G-A------------------------G--C--A...................-------------T-----TG----------T-----G 7353
B.CO.01.PCM001          -A...............-----------------CAC-...-A----A--------G-G--A------GG----G------A----C-------------A-AA------A...................-------------T-------------------T---- 6514
B.GB.83.CAM1            -A...............------------------C--...----TA----------C-------------------T--CA-------A-------------G-----G...................-------------T------------------------ 7331
B.GB.86.GB8             -A...............---------------------...---A--A------------T------C-C-------T--TAG----G--GA---------A-G-----A...................-----G-------T------------------------ 7325
B.GE.03.03GEMZ010       -A...............------------------C--...----T-ATC------G----------------C-------CA-------------------AAG-----A...................-----GC------C------------------------ 6520
B.IT.05.SG1             -A...............------------------C--...-----A--GCC------G-----C------T-G--T-GAA-------A----------CC---G--A...................-----G-------T-G----G-------------G--- 7076
B.JP.05.DR6538          CA...............-A-A-----T------CAC--...-A--A-A--C------------G---------G----G----GAA-----------------------G-----A...................-----G-------T------------------------ 7373
B.KR.05.05CSR3          -A...............-----------C--------C...------GAA--GC----C-G------------A--------GA---------T-----------G-----A...................-----G-------A-CC---------------------- 7388
B.NL.00.671_00T36       -A...............------------------C--...---------------------G-----------C-------T-G-A---------------TT--A-AAG-----A...................-----G------TTC---------------T------ 6901
B.RU.04.04RU128005      -A...............-----R------------C--...---G-----G-C----C---A------G---------T--AA------MA---------A-AAT...--G...................-----G-------------A-------A--------- 6812
B.TH.00.00TH_C3198      -A...............---A--------------C--...----CA----T---------------------------T--AA--------------------------A-------A...................-----G------------------------------G 6529
B.UA.01.01UAKV167       -A...............------------------C-C...----G---C------------A------G----A----T--TA------G--A---------AAC-G--CA-C-...................-T---CT-TGA-C-A---C--C---------------- 6595
B.US.04.ES10_53         -A...............---A-----------------...---T--T--GT------G----C---------G---T-GAA-----CA------C--CCC------A...................-----GC------AA-T----------A--------- 7330
B.US.99.PRB959_03       -A...............-----------------T---...-CGA-AA-GC----------A----------------------T--CA----------------A--AG-----A...................-----G-------T-----A----------------- 6541
C.AR.01.ARG4006         -A...............---G-------------T---...----C-A--GC----------A------C-TG-AT--AG--AA-------A-A-G-TC--CCCT-----A...................------CAA----CAA-CCA-------------T---- 6525
C.BR.04.04BR013         -A...............---------------------...----A----GC----------A------C--G-AG--AG--CA------GC---G--C--CCCT-----A...................-------AA----CAA-GAG------------AT---- 6831
C.BW.00.00BW07621       -A...............---------------------...----A-A-T-----------G--C--G-GCAG---T-GGGAA-------AG-----C--CCCT-----A...................--C----GA-----CTC---TG-----------TT---- 6681
C.CN.98.YNRL9840        -A...............---------------------...---GA---T---------G-A------C--AG-G---G-GAA------GC------C--CCCT-----A...................-------A-----CATC-----------------T---- 6541
C.ET.02.02ET_288        -A...............---------------------...---G---AC-G---------G---C--G--A-TG--AAGGAA-G----CA-A-G-----CCCT-----A...................------C-A---GCA-C-----------------T---G 6575
C.GE.03.03GEMZ033       -A...............-----G---------------...-A-GA-AAT----------CA------C--A--G-----GAA--G----AGA--T-C--CCCT-----A...................------GCA-----CTCG----------------T---- 6547
C.IL.99.99ET7           -A...............---------T------C----...-A-G--ATG---------G-A------C--GG-G---G-GAA------CA-A----C--CCCT----GA...................-------GG---G-A-C-----------------T---G 6518
C.IN.99.01IN565_10      -A...............---------T-----------....--TAGT--C------G-A------T-TA-TG--AG--AA------GC---C--C--CCCT-----A...................-------A-----CATC-----------------T---- 6720
C.KE.00.KER2010         -A...............---A-----T-----------...----A--AG-----------T------C-----G--AGG-AA------GA------C--CCCT--C--C...................------GAA-------C-CA-C-----------T------ 6538
C.MM.99.mIDU101_3       -A...............---------------------...----A-CA----------G-A------C--AG-G---G---A------GC-----TC--CCCT----GA...................------G-A-----CTC----C-------------T---- 6717
C.MW.93.93MW_965        -A...............---------------------...----C-ATT-----------A------C--GG-G---AG-AA------A------C--CCCT-----A...................------GAA-----A-C---------------T-T---- 6536
C.SN.90.90SE_364        -A...............---------------------....--T-GT--C------GCA------G--A--G--AAAG-A------CA------C--CCCTGG---A...................-AT----G----G-A-C-----------------T---- 6515
C.SO.89.89SM_145        -A...............---------------------...---G---AT--------C-GA------C-----G---GA-AA------GC------C--CCCT-----A...................-------AG---G-A-C---------------T-T---- 6560
C.TZ.02.CO178           -A...............---------------------...---GA-AGC-C----------G------C--TG-G---GA-TA------C-G---T-TGAACATTT--------...................--------A-----TTC---------------TT---- 6547
C.UY.01.TRA3011         -A...............---------------------...----A-AA-GC-------GG-A------C--A--G--AGA-AT------GC---G--C--CCCT-----A...................-------A-----CAA-GCA-------------T---- 6511
C.YE.02.02YE511         -A...............---------------------...---GA-AAG---------G----C---C--G------GA-AA-------A------C--CCCT-----A..................T------GG---G---C-CA-------------T---- 6526
C.ZA.04.04ZASK164B1     -A...............----------C----------...----A-AA-GC-------G-A-TCC--C--A--G--AAGGAA-------A------T--CTCTG----A..................GA-----A------C--T--------G-------T---- 6790
C.ZA.05.05ZAPSK240B1    -A...............---------------------...-------A--G-------C--A-----G--AG-G--AAA-AG------G-------C--CCCT-----A...................-------AG-----A-C-CA-------------T---- 6719
C.ZM.02.02ZM108         -A...............---GC----------------...---GA--GGCT-------C-A------C-CA-TG--AG-GAA-----C-----GT-C--CCCT-----A-C-...................-TG-ACT--ACAC-A-CCG-TG-----------TT---- 7365
D.CD.83.ELI             -A...............G---------------T----...-----------C------G---A------C----AG-------A------G--ACC-TTC--AAC--A-CA...................-T-----AG-----A-C---------------------- 6873
D.CM.01.01CM_0009BBY    -A...............---A---T--------T----...---G--AA-G-T-------------C--C--G--G-------A------G--A-T-TC-A-AAC...--A...................-------GT----GA-C---------------------- 6535
D.KE.01.NKU3006         -A...............----T----TT----G-G---...-A-G--A--G-G--------A------C-----G------AA------G--A-C-TT--CCCC--C--A...................-----C--T-----A-C------G---------T---- 6577
D.KR.04.04KBH8          -A...............---AG------------T--C...------CA----------G-A------C-----T----A-AA------G----C-TC---AAC--G-CA..................GA-----AT-------C---T-------------T---G 7259
D.TD.99.MN011           -A...............-----------------T---...---G--------------------C--C--G--G-------A------G---C-GCTC--AAC-CC-CA..................CA------T-----A-C---------------------- 6536
D.TZ.01.A280            -A...............---A-----T-----------...---G-----G-----G----A------C----AG------CA---C--G----C-TTC--AAC--G-CA...................-T-----AT----GT-C---------------------- 6531
D.UG.99.99UGK09259      -A...............---------------------...--C------------------A---C--G-----G------AA---------CAGC-TT---AACC-G-CA...................--------A----GA-C-----G---------------- 6535
D.YE.01.01YE386         -A...............---A------------CTG--...-------ACG----------A------C--A--G--A---AT-G----GA-A---TTA--A-C...--A...................---G---AG---C-A-C---------------------- 6532
D.YE.02.02YE516         -A...............-----------------T---...---G--A--C---------------C--C--A---------AA------G--A---TT-A-AAC--A-CA...................------G-T-----A-C-----------A--------- 6511
D.ZA.90.R1              -A...............---------------------...---GA--A---------G-GA------C-----G------TA------G--A-C-TTC--AAC--A-CA...................-------AT----GA-C-C-------------------- 6694
F1.AR.02.ARE933         -C...............---A-G-----------TG--...---G--ATGC----------CA--GC--G---G-G--AAGGCA-------AGCCT-TT--CCCT---GCA...................-----T-AA-----CTC-----------------T---- 6625
F1.BE.93.VI850          -A...............---A-G---------------...---G--A--C---------------C-GG-GT-TG--AAGGCAG------AGTCG--T--CCCT-------C-..................G------AA-----C------------------T---- 6653
F1.BR.01.01BR125        -C...............---A-G-----------G---...---G-----C----------A--G---GG-AG-G--AGGGCA--G----AGTCT--T--CCCTG--GCG..................G------AA-----CTC-----------------T---- 6791
F1.ES.x.P1146           -A...............---A-G-----------G---...---G-----C----------A---G--G---G----AG-GCA--G----AGTCTT-T--CCC------A...................------CAA-----CTC---------------T-T---- 6674
F1.FI.93.FIN9363        -A...............---A-G---------------...---G---A-C----------A------G-C-GAG--AAGGCAG-G----AGCTT--T---AAT...--A...................------CAA-----CTC-----------------T---- 6666
F2.CM.02.02CM_0016BBY   -A...............---A-G----------C----...-A-G--A--CTG------GCA--------TAG-G-----GCAG-GG-C-A-A-C-TCAC-AAT...-T-...................-------AA---G---C---A-------------T---- 6517
F2.CM.95.MP257          CA...............---A-G---------------...---GA-AA-C-G---T--G----------T-A---------CAG-G--C-A---C-----A-AAT...--A...................---G--GGG---C-A-C---AG-----------T-T---- 6539
F2.CM.97.CM53657        -A...............---A-G--T------C-----...-A-G--A--CTG--------A------G--GGAG------AAG-GG-CCA-AGC--CC--AAT...--A..................T------CA---GC--C---A--------------T---- 6520
G.BE.96.DRCBL           -A...............---A-G-----------TG--...---T-GA-----------G-G--GC---GGG-TG-CCAGGCA---C--CAG---T-C---ATT-----G...................-GC---GAA-----CTC---T--------------T---- 7294
G.CM.01.01CM_4049HAN    -A...............-----------------TG--...----AG-----------G-CA-TG------A-TG-CA-A-CACTTC--CATA-GA---A-AAT...--G...................--C------------TC----G------------TG--- 6538
G.CU.99.Cu74            -A...............---------------G-GG--...-A-G--A-----------G-G--AC--C-GA--G--AAAGAA---C---AG---AGC-ACAAC--A--G--CATAACATTTGGG......-AC----CA---G-A------A-------------T---- 6977
G.ES.99.X138            -A...............-----G---T-------T---...-A----A--G-----------T-TGA----TA-TG-GT---AA---C---A----ATC---AAC--C--G...................-GC----C------C-C-C---------------TT---- 6784
G.GH.03.03GH175G        -A...............----------------TTG--...-A-G---A----------C-GGTGA--C--A--G-CAAA-CAC-TC--GAG---ATC-A-AAT...--G...................--C----------GCTC---G-------------T---- 7407
G.KE.93.HH8793_12_1     -A...............-----G---------G-TG--...---G--AG---T----G-----TGA--G-GA--G--AAAGCAC--C------A--ATC---AAC...--G...................--C----C----G-CTC---TG------------T---- 6737
G.NG.01.01NGPL0669      -A...............-----------------TG-A...---G-G-A-------GGGG---TG-C-C-GA-AG--AAA-CA---C--CA---GGTC---AAC...--G...................-AT----C------CTC-----------------T---G 6569
G.PT.x.PT2695           -A...............-----------------T---...-------AT--G----G-A--T-TGAA-C--A-TG--A--GCA---C--GA-A--AT-----...--A--G...................-AC----C------C-C-C-TG-----------TT---- 7363
G.SE.93.SE6165          -A...............-----------------TG--...-C--A-AG----------AG-GG-C---C-GA-TG-C---GCAG--C--G-GA--ATC---AAT--G-GC-GCGAG................-AC----C------CTC---TG------------T---- 6764
H.BE.93.VI991           -A...............----G----T-------T---...---G--AA-C---------G-G--C---C-CA--G--ATC-C-------G--AGCT-C----AC-----A...................--------T--AC-A-C-C--G----------TAT---G 6730
H.BE.93.VI997           -A...............-----------------TG--...---GA--AG-----------G------C-------------CAC-------TA--T-C----TC----GA...................---T----T-----A-C-CA---------------TG--- 6660
H.CF.90.056             -A...............-----------------T---...-------A--G-C--------G------C-C---G---T--CAC-----G--AT---C--GAAC----GA...................-------G--------C-AA-------------ATG--- 6659
J.CD.97.J_97DC_KTB147   -A...............---------------------...---G---T------------T------C-TA-AG-----GAAG--------------GAACACTT---G--TAAAACA.............-T-G-CT-TGCACCAGC-AAT-------------AT---- 6549
J.SE.93.SE7887          -A...............---G--A--------------...---GA-AG-G-T----G----------CGTAGAG------CA---C-----------C---AAT...--A...................------A-----CATC-C-------------TAT---- 6638
K.CD.97.EQTB11C         -A...............--G--G--G------------...--CG---G-C-G--------A---G----C--AG--AAG-AAG-----G--A----C---AAC...--A...................--------A---C-A-C-------------------C-- 6547
K.CM.96.MP535           -A...............-----G--------C------...---G---A---------C-TG-------GCAGAG--AAGGAA--GC---A------T---AAG---GGA-C-...................-T--C-T-TAAACCA-C-AA-C-------A--------- 6539
```

```
                                V3 loop end┘
B.FR.83.HXB2               TG...............AGACAAGCACATTGTAACATT...AGTAGAGCAAAATGGAATAACACTTTAAAACAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAA.................ACAATAATCTTTAAGCAATCCTCAGGAGGGGACCCAGAA 7334
Env                        M__.__.__.__.__.__R__Q__A__H__C__N__I__.__S__R__A__K__W__N__N__T__L__K__Q__I__A__S__K__L__R__E__Q__F__G__N__N__K__.__.__.__.__.__.__T__I__I__F__K__Q__S__S__G__G__D__P__E_
01_AE.CF.90.90CF11697      -A.................---A-----T-----G-G---...-A-G--A-----------G-A-----------G--A-A-AA----------G--C---AAG-----G.................------------C-A-C-----------A--T------ 7298
01_AE.CN.05.FJ051          -A................--GA-----T-----G-G---...-A--A-A---C----C--G-A-----------G---T-G-A----------G--C---AAT...--A.................------------C-A-C-C---------A--T-T---- 7331
01_AE.CN.06.FJ054          -A.................---AG----T----C--G---...ACGA-A------------A-----G-----G--AGGGAA----------G--CA--AAT...--A.................-----G----C-A-C-C---------A--T-T---G 7370
01_AE.HK.x.HK001           -A................--GA-----T-----C-G---...-A-GC-A--------G-A--AGT----------G------AA----------G------AAT--G-CA.................-AC----GA---C-A-C-CA---------A--T-T---- 6704
01_AE.TH.01.01TH_R2184     -A.................---A-----T-----G-G---...-A-G--A---T---------AGT----------G-----GA--------A--G--C---AAT...--G.................----C-GC----C-A-C-C----------A--T-TG--- 6540
01_AE.TH.02.OUR769I        -A.................-----G---T-------G---...-A-G--A-----------G-A-----------G-----G-A----------G--C---AA-GGG-CAG-G.................-A----------C-A-C-C-A-T------A--T-T---- 6533
01_AE.TH.90.CM240          -A.................---A-----T-----G-G---...-A-G--A-------------AGT----------G--A--GAA-------A--G--C---AAT...--G.................------------C-A-C-C---------A--T-T---- 6905
01_AE.US.00.00US_MSC1164   -A.................---A-----T-C---G-G---...-A-GA-A-----------G-A--G----------G-------A-------A--G--C---AAT...--G.................-------A----C-A-C-C-T--------A--T-T---- 6536
02_AG.CM.02.02CM_4082STN   -A.................---------T------TG-C...-A--A-A--GC-----C---T------CGC---G-G---CAC---------A-GT-C---AAG---G-A-C-.................-----G------CTA-C--------------TTT---- 6518
02_AG.EC.x.ECU41           -A.................-----G-----C-----TG-C...-A---A--G-------C-G--G----C--A-----T--CA--------GA-GT-C---AAT-----A.................------------GCTA-C--------------TTT---- 6405
02_AG.GH.03.GHNJ196        -A.................-----------------TG--...---GC-AA-GCT-----G-T------C-----G-G--C-CAC-----G-GA-G--C-AC--TGG---C-C-.................-----C--A----CTA-CCA------------TGT---- 7387
02_AG.NG.01.PL0710         -A.................---------T-C----CTG--...---GA-A--G-------AG-AG-----C-G---G------CAC-------GA--A-C---AAG-----G.................------------G-TA--C----------A--TGT---- 6506
02_AG.NG.x.IBNG            -A.................-----------------TG-C...----A-A--G----------A------C-T---G----T--CAC------AGACGT-C---AAG--C-CC.................------------GCTA-CC---T---------TGT---- 6852
02_AG.SE.94.SE7812         -A.................-----------------TG-C...--CT-GCA-C-------C--A------C-CG-TG-G----CA--------G--GT-T---AAT----CC.................------------G-TG--C--------------TTT---- 6694
02_AG.SN.98.MP1211         -A.................-----------------TG-C...-------A-----------G-------C--A--G-G-T--CA------G-GA-G--T---AAT-G-TCC.................-A-----------CTA-C--------------TTT---- 6521
03_AB.RU.97.KAL153_2       -A.................-----------------------...----T-A----------------A-------------T--T----------A----------G-----A.................------G-------T----------------------- 6526
04_cpx.CY.94.CY032         -A.................---------------------...---G--AATG-T------G----C------GT----AG-GAAG----G-A-AG--TC--CCCT-----A.................-------AA---GCT-C-C--GT-----------T---- 6716
05_DF.BE.x.VI1310          -A.................---A-G------------G--...------G-A-C----------A--C----TC---G----G-AAG-G---CAGTCT--C--CCCT-----A.................-------AA-----CTC-----------------T---- 6731
06_cpx.AU.96.BFP90         -A.................--------T--------TG--...------------T----CAG-T-TAC--GG-G-AG-GAAGGTA---C--GA----GTT---AAT--C-CA.................CAC----C-------TC---TG------------T---- 7400
08_BC.CN.97.97CNGX_6F      -A.................---------------------...-----A--ATGC----T--G-A------C-GAG-G--AG--AA------GC------C--CCCT-----A.................-------AA--CGCATC-----------------T---- 6516
09_cpx.GH.96.96GH2911      -A................--GA---T-------C-G---...-A-G--A-----------G-A-TC---GG-A--G-----GAA-------A--G---C---AAT-----G.................------------CGA-C--------------TGT---- 6527
10_CD.TZ.96.96TZ_BF061     -A.................---------T---A--------...---G--A--G-----------A------C----AG------AA------G----C-TT---AACCGACCA.................--------T-----A-C------------------------ 6698
11_cpx.GR.x.GR17           -A................-A----------------G-C...----------G-C---------------C-----G-----GAAC-----CATA-TA-C---AAC...--A.................------G-A-----TG-GCAT-----------TTT---- 6598
12_BF.AR.99.ARMA159        -T.................---A-G-----------TG--...---G-----C------G---A--GG--G-----G--AAGGCA--G------TCT-----CAAT...--A.................-------AA-----CTC---------------------- 7338
13_cpx.CM.96.1849          -A.................-----G--------------C...------CA---------------C---C-----G-----GCAC-G------A--T-T--GAA-G------C-.................------------GCTA-TC-------------TTT---- 6753
14_BG.DE.01.9196_01        -A.................-------------------C--...-C-A----C-----G----A--C----G---AG-----G--G--------A---------AGG-AG-----.................-------G------CA----------------------- 6854
15_01B.TH.99.99TH_MU2079   -A.................-------------------C--...-----CA--------GG------------C------A--GAA--------------------G-----A.................------G-------C---C-------------------G 6700
16_A2D.KR.97.97KR004       -A.................-----------A-----T--C...--C--CAG-C--------G---------C--A--G-----GAGC-G------A--T-C--CTCT-----A.................-------------CCA-C-----G---------AT---- 6712
17_BF.PE.02.PE02_PCR0155   -C.................---A-G-----------TG--...---G-----C-----------A--A---G-----G--AAGGCA--G----AGTCTT-T-GGAAT...-CA.................------CAA---G-CTC-----------------T---- 6536
18_cpx.CU.99.CU76          -A.................------------------T---...-CGA-AA-G-T------G--------C-CG--G------CAC-----G---G---CC-GAAC--G-GC.................------TA------A-C---T--------A---AT---- 6636
19_cpx.CU.99.CU7           -A.................-----G-----------TG-C...----A-AA-G-G-----AG-AG-C---C--G--G------AAC------C--TC-TC---AAC...--A.................-------G----GCTA-C-----------A--TTT---- 6463
20_BG.CU.99.Cu103          -A.................---------TT----GTGG--...-A--A-A-----------G-G-TGC--C-GA--G--AAAGCA---C--GCT---AGC-ACAAT--G--C-T-ACA...............-AC----AA----GC-C-C--A------------T---- 6804
21_A2D.KE.91.KNH1254       -A.................-----------------T---...---GA----G-T---------A---------GG-G-G----TA------G----C-TT---AAC-CA-CA.................-A-----AT-----A-C------------------------ 6547
22_01A1.CM.01.01CM_0001BBY -A.................---A-----T-----G-GG-C...-A--A-A--G--------G-A------C-CGG-G---TA-AAC-G------A----T-GGAAC...--A.................-----------C-CTA-C--------------TTT---- 6516
23_BG.CU.03.CB118          -A.................---G-----TT----GTGG--...-A----A--GCC------G-G--GA-GC-GA--G--AAAGCA---C--GCT---AGC-ACAAC-CG-GC-C-GAA...............-AC----CA---GGA-C-C---------------T---- 6804
24_BG.CU.03.CB378          -A.................---------T-----G-GG--...-A----A--------G-GG-A-TGA--C-G---G--AAAGCAG--C--GCT---AGC-ACAAC--G-GC-T-GCA...............-AC----CA----GC-C-C--A------------T---- 6792
24_BG.CU.03.CB471          -A.................---------T-----GTGG--...-A----A--G--------G---TGA--C--A--G--AAAGTA---C--GCT---AGC-ACAAC--G-CC-C-GAA...............-AC----CA----GCAC-C--A------------T---- 6804
25_cpx.CM.01.101BA         -A.................-----------------TG--...------CATC-------C--G--G---C--G-AG------AAC-------CA-G--C--GAAC-TACCA.................GA----------GCTA-TCA---------A--TGT---- 6590
27_cpx.CD.97.97CDKTB49     -A.................-----------------T---...-A-GA-A--C-G---G-GG--------T-TA--G-----GCAC--C------A----CC-GAAC--G-GC.................-GCTC---A--C--A-C----------------AT---- 6547
28_BF.BR.99.BREPM12609     -AAGA.........CAAGC---TTGCA-CAT--GT-GAGAA-AAT-G-A---CACTTTA-GACAGG--GTTACA-A-TTAGAAG--CA-TTTAC-A----A---G-AC-ATTTAAAGATTTTAAAAATAGA------GC------T---------------------- 6732
29_BF.BR.02.BREPM119       -A.................---A-------------------...----A----------G-C---------------------A--------A----------AAT...--A.................------G-T-----T------------------T---- 6489
31_BC.BR.02.110PA          -A.................---------T--------------...---------GC----G-A--A------CT-G--G--AG--AA------TC---G--C--CCCT-----A.................-------A-----CAA-GCA--------------T---- 6818
32_06A1.EE.01.EE0369       -A.................--------T--------------...---G--------------G-G-TG---C-TA--G-GATAGCA-------AG---TCC---CCTGG---G.................-AC----C------TTC---GG------------T---- 6999
33_01B.MY.05.05MYKL007_1   -A.................-----G---T-----G-GG--...-A-G--A-G------------AGT----------G-GT--GAA-------A---G--C---AA------G.................------G-----CG--C-C--C-------A--T-T---- 6728
34_01B.TH.99.OUR2478P      -A.................---------T-----G-G---...-A-G--A-----------G-A-------------G-----G-A---------A-G--C---...------G.................-------G------A-C-C----------A--T-T---- 6518
35_AD.AF.05.05AF095        -A.................-----------------TG-C...----------GCC---------A-------G---G-----C-C--------C----C-GG...--C--A.................------G------GTA-C--------------TGT---- 6500
36_cpx.CM.00.00CMNYU830    -A.................---A-----------G-GG-C...-A-GC-A--G--------G-------GC-G---G---TACAA---------A--TTT-GG...--C-CA.................-----------C-CTG--C----------A--TTT---- 6537
37_cpx.CM.00.00CMNYU926    -A.................---A--------A---GCTG-C...-A--A-A--G-G--------AG-----C--A--G---T-GCAC--------A-C--T-----G--C-CA.................--------G---GCTA-C-----------A--TGT---- 6522
38_BF1.UY.05.UY05_4752     -C.................---A-G---------------...---G---A-C-----------A--A---G--AG-G--AAAGAA--G----AGTTTT-T--CCCT-----A.................------CAA-----CTC-----------------T---- 6778
39_BF.BR.03.03BRRJ103      -C.................---A-G-----------TG--...---G-----C-----------A--GA--G-----G--AAG--A--G----A-TCT--T--CCCT---GCC.................-------AG-----CTC-----------------T---- 6836
40_BF.BR.04.04BRRJ115      -C.................---A-------------TG--...------A--C-----CG---A-AA----G----G--AAAG-AG-G----AGTCT--G--CAATGGAGCA.................-AT----CA-----TTC-----------------TG--- 6897
42_BF.LU.03.luBF_05_03     -C.................-----G---T--------G--...-----G-A-C-----------A--A---G---GAG--AGGG-A--G-----GTCT-TT--C...-----A.................-------AA-----KTC-------------A---------- 6865
43_02G.SA.03.J11223        -A.................---G-------------TG-A...------A---G-------G-G--G-----GA-TG--AAGGCAC--C--CACA-GATC---AAC...--G.................--T---C-------CTC----AT-----------TR--- 6853
U.CD.83.83CD003            ...................---------------------...-CAGAT-GGG------GG--T------C-----G------T-GC-------AG-------AAT-----A.................T-------A-----CTC----------------AT---G 6807
U.CD.90.90CD121E12         -A.................-------------------C--...-CA-A-TGGC-------G-GAG-----C-----G-----GATGC-------AG---C---AAT...--A.................------G-A-----CTC----------------TT---- 6493
U.GR.99.GR303              -A.................-----------------TG--...TCAG------C------C------C-----AG-G---T-GCA---C----------G-----G--C--A.................---------------TTC---TA-----------TTT---- 6618
U.NL.95.U_NL_95_H10986_D1  -A.................---A-----T-----G-GG--...-A----A--------------TG----G-A--G-----GTA---C----------G---AATCT------CAAA...............---------------TTC---TG-------------T---- 7231
CPZ.CD.90.ANT              CT................--GA----GTTC----CAG-C...-A--AGA-GCT----G-AC-AG-ACGT--CA-A-C--AGCA-GTTC-TGCG--G--T-GGAA---AGTAG-CAACAAAACCAAT......G-G-A--CAA-ATG-AC--T-CA--AT--A--T--T--- 6796
CPZ.CM.01.SIVcpzCAM13      CA................--GA----CTT---CG-AC-C...-A-G-GA---C-------G-A--AC-GC--A-AG-TAGAGAA-GC----T-A-GG--A-AAA-GCA---GC-AATGGAACCTAC......-AT----CG---G---C-AG-AGT-----T--------- 6954
CPZ.CM.05.SIVcpzEK505      CT................--G------T-----G-GG-A...-AC--GA--...---G-AC-A-TA-GG--TACA-C-AAACAA-T-A---T-.........AAC--C-GA--G.................-AT--T-CA----TA-CCAATC----C--A----T---- 6856
CPZ.CM.05.SIVcpzLB7        CA.................---G-G--T------G-A---...-ACCAGA--C-T---T----A-TAC----C---G-GAAA--AG-GC---CCACGGTC--CAATG-ATC---CAAA...............--TG--CA----GCAA-CAGT-----T--A--T------ 6866
CPZ.CM.05.SIVcpzMB66       CA................--G--G--CTT---CC-GT-G...-AC-A-A--GT-----C---T--A--T---A-AG--AGGCA-GCC----AT---AC--A-AAGGGCT--CTCGGAAATGAA.........--T--C-CA---GGA-C-AG-A-G-----T---TT---G 6982
CPZ.CM.05.SIVcpzMT145      CT.................---A-----T-----G-G--A...-A-G-GA--C-G---GCA--AG-CC-C--TG-A-C-AAAGAAGTTC----GA-CAT-C-GA-G...--G.................-AC----G-----T-GTCC-T--T--G--A--T------ 6846
CPZ.GA.88.GAB1             CC.................---TCT--CT-C-----G--A...-A-G-GA-C-C-------CGA--GG-GG--G-AG--AAA-AGGCT---GC-ACCTC--CAAATCGA-CAGC-GCC...............-AC----CAC-C--CAG-G----T-----A--T-----G 7395
CPZ.TZ.00.TAN1             CC................--GA----TT-----TCAG-C...-A-GC-A---C----T-C-GG-AC---G-TTG-GCTATGGCTGCCA--ACAC-ACCA-GA-GGCC-GA--TGAAACGGTA.........CA-CA--CG--CC-ATGGCAGAGG-AT--A-----T--G 6981
CPZ.US.85.CPZUS            -A.................---G-----------------C...TCA-A-CTG-C----G-G............A-ACAGAGA-AAT-CACTTT---GAT-A-AAA---GG-AGC-AACCTGACA.........-AGG--GAG--A-TT-C-AATG--------A---------- 7363
CPZ.CM.x.SIVcpzDP943       CC.................---------T-------TR-C...--CGCC---G-----G-AC-A-GG--M--WG-C-CTRTG--GGCTA---A-AGT-TTAAGCC-GGC-G-...............-AC--C-C---C-CAG-TCAACA------A-----G--- 6897
CPZ.CM.x.SIVcpzMB897       CC.................---GG----G-G--C-GA---...-A-G-GA--GCC------G-G---C-G-G----G--AAGGAGGC-----ACA-CAC--A-C-----CTC--TCTCAGTCTCACA......GA-----A----G-AGGTG-A--C--C--A--T-T---- 6893
N.CM.02.DJO0131            -T.................-----------C-----TG-C......TCTAACG------GATCA-TG-GG--TA-A-C-AAAGAG---A---AGAGC-TCC-G-----C--C-T-.................---T-C-AGGC-C--G--AAGAAT-----A---------- 6797
N.CM.04.04CM_1015_04       -T.................-----------C-----TG-C......TCT-A-G-------ATCA-TG-GGG-TAGA-C-AAAGAG---A---AG-GC-TCC-G-----C--C-C-.................--CT-T-AG-C-CG-GTGAA-ATT-----A---------- 6790
N.CM.04.04CM_1131_03       -T.................-----------C-----TG-C...TC-GA--A-G-------ATCA-TG-GG--TAGA-C-AAAGAA---A---AG--CATTC-G-----C--A-C-.................--CT-TGAGGC-CG-GGGAA-GAT---A-T---------- 6840
N.CM.95.YBF30              -T.................---------T-C-----TG-C...TC--A--A-CT----G-ACCA-TG-GG--TAGA-C-AGAGAGG--A---AGA--ATCC-G--G--A--C--CATA...............--CT-C-GGGC-CGAG-GAGGAAT-A---A----T---- 6915
N.CM.97.YBF106             -T.................---------T-C-----TG-C......TCT-A-G----GGATCA-TG-GG--TA-A-C-AAA-AG--GA---A-AG--TCC-G-----C--C-C-.................--TT-C-AAGC-C--G-TAAAAAT-----A----T---- 6879
O.BE.87.ANT70              GAAAC......AGCTCA--GGC---TT----C--GTA-...-A-GCCA-TG-T---GGA--A-TA--------A-C----GAA-GG-ATTT-----T-G-AAAC---CAGGTAGTATT.............-AC--G-CA--C--T--CAG-AGC--T--A--T-T---G 7375
O.CM.91.MVP5180            ATACATCACCAAGATCA--GGT---TT------CATA-...-A--AGA-TGT----G-A--TG-CC--C----A-C-----TA-GG-ATTT-A-T-TTG-AAACC-A-CAG-GAATGTT.............--C-----A--C-GCAG-A-TAGT--T--A--TG----- 7383
O.CM.96.96CMABB637         CAGCCGCC...ACTCCC--GAC---TT----C-CATA-...-A-GTCT-TG-----G-A-GTG-A--------A------GAA-G--ATTT-----TTA-AAATC-G-CAG-CAATGTT.............G-TG-G-----CG-TAG-AGTA-T--T--A--T-----G 6865
O.CM.98.98CMA104          CAAGC......GACCCA---AG---TTT---C--AT--...GAAGA-A-C---------AG-AG-C----R---A-C----GAA-GG-ATTT-----TTG-GAAC----CA-T-GATAATGTT..........-GCG-G-----C--TAG-AG-AAT-AT--A--T-----G 6837
O.CM.99.99CMU4122          CAAAT......GCCTCA--GTT---TT----CCGGTA-...-A-GCCA-TG-G-----A--AG-C---------A------GAA-GG-ACTT-----TTG-AAAC----CA--CAATACAGRTGACAATGTT-GC--G--A--C-GTA--AG-A-T--T--A--TT----G 6827
O.FR.92.VAU                G-GAT.........ACA--GGC---TT--------TTA-...---GCCA-TG-C-----C--AG-C------A-C---A--GAA-G--ATTT-----TTG-A-A-T-----C--ACTGATGTT..........--C----AA--CGGTA-TCA-AGT--T--A--TG----- 6917
O.SN.99.SEMP1299           AACCAAAT...AACTCA--GAT---CT----C--GTA-...-AC-TCT-TG-T---G-A--AG-C-----------C----GAA-GG-ATTT---C-TTAGGAAC----CA--TACAGTT.............-AC----CA--CG---GCAG-ATT--T--A--TT----G 7419
O.US.99.99USTWLA           CAAAT......AGATCA---AT---TTT-----C-TA-...-A-GCCA-GG-------AG-A--C--GC--GGA------GAA-GG-ATTT-----TTG-AAATT---CAG--AAAATT.............-AC----CA--C---A-CAG-A-T-AT--A--TATC--- 6837
```

```
                                                       ⌐V4 loop start
B.FR.83.HXB2              ATTGTAACGCACAGTTTTAATTGTGGA...GGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGTACTTGG..................................TTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAAT 7465
Env                       _I__V__T__H__S__F__N__C__G__.__G__E__F__F__Y__C__N__S__T__Q__L__F__N__S__T__W__.__.__.__.__.__.__.__.__.__.__.__.__F__N__S__T__W__S__T__E__G__S__N__N__T__E__G__S__D__T__I
A1.GE.99.99GEMZ011        G--AC---A--T-----C---------...--A-----------T--C---A-----GGC-----C-----C------.........................GACG-A-A-GACACT-TCC------G---A--G-CT-A---A--T-- 6637
A1.KE.00.KER2008          ---AC---A--T-TG----C-------...--A--G--------T------A--T--GGC-----C-----C------............................CCTGTC-A-C-G-A--------G---AA-GTC--A--G--TT-- 6674
A1.KE.00.KNH1144          --CAC---A--T-----C---------...--A-----------T------A--T--GGT-----C-----C--C---.........................-A-ACC--C-TGTC-------G-----G--GAC--A-----T-- 6694
A1.KE.00.KSM4024          G-CAC---A--T---------------...--A-----------T------A--T--GGC-----------C------...................TC-TGGAATGAC----C-AATGGC-C-GCA-TG--GCC--A-AG---T-- 6597
A1.KE.00.MSA4069          --CAC---A--T---------------...--A-----------T------A--T-GGGC-----------C--C-AC.........................-A-G-C-C-GACATC-CTGA-----G---G--GTC--A------T-- 6676
A1.KE.00.NKU3005          ---AC---A--T---------------...--A--------T--T--C---A--T--G-T------------------...................CC-GTCAATGCC-GCATG-A--------G---G--GTC--A-----AT-- 6679
A1.RU.00.RU00051          ---AC---A--T---------------...--A-----------T--C---A-----GGC-----C----A-G-CA-CAACAAT.................................GGT--AA-CC-C-CA---C-T-- 6733
A1.RW.93.93RW_024         --CA----A--T-----------A---...--A-----------T------A--T--AGC-----------C------...................----AC---GCC--C-TGC-GAA--------G---G-GGTC--A------T-- 6667
A1.SE.95.SE8891           --AAC---A--T---------------...--A-----------T------A--T--GGC------------------............................AATG-C-----CA-CATGC-G--GTTAA-TA-C-CG-G--AT-- 6630
A1.SE.95.UGSE8131         ---AC---A--T---------------...--A-----------T------A--T--GGT------------G-----...................-A-GAAAATGAC--CA---TAAATT-C--------GTC--A------T-- 6840
A1.TZ.01.A173             ---AC---A--T-------------G.. .--A-----------T-----CA--T--GGC---------G-CCTG-TTAATGGCACCTGG.....................-A-GGCA-TCACA----C-G---G--GTT--A-AG---T-- 6674
A1.UA.01.01UADN139        G--AC---A--T-----C---------...--A-----------T--C---A-----G-------C-----C------.........................GGCGAA-A-GGCACTATCA------G---A--GACC-----A-AT-- 6661
A1.UG.92.92UG037          ---AC---A--T-----------C-..  .--A-----------T------A--T--GGC------------------............................GTAA-T--CA-T-CC-G--GCATGTC--A--G---T-- 7475
A1.UG.99.99UGA07072       --AAC---A--T---------------...--A-----------T-----CA-----A-C-----------AC-T----............................A-T-A-G-CAGCAT-CAG--GGCA-A-AGCAT--AC-----T-- 6671
A2.CD.97.97CDKTB48        ---AC---A--T---------------...--A-----------T--C---A-----GGC-----------C--A---.........................GAAAATG-C---A-CAAACAG---T----G--GTC--A------T-- 6778
A2.CY.94.94CY017_41       ---AC---A-TT-----------C-...--A-----------T--C---A-----GGC---------G----G---..........-GG--C-A-GG-AC-T-G-AC-GGCCC-AC-CACCT-A-A-CACC-A--GA-GT-- 6840
A.SN.01.DDI579            ---AC---A--T--C------------...--A-----------T--C---A--T--GGT------------------.........................---GAC--GGCC-TC-A---CACTATC--CGTGCAGA-TTC--C-T---AT-- 6669
A.SN.01.DDJ369            ---AC---A--T--C------------...--A-----------T--C---A--T--GGC-----------------TAT.................AA--G--C-TGGAAT-AC-----CA-CA--C-G--AT-AA-T--C-CG--G-A--- 6687
A.SN.96.DDJ360            ---AC---A--T--C------------...--A-----------T--C---A--T--GGT------------------.........................-A--GCACT-CC-TCA-GCA-AATTTA-CT-ACA-T-AC-CG-GGGT--- 6651
A.CD.02.02CD_KTB035       ---AC---A--T-C-------------...--A--------T--T--C---A--T--GGT---AA-----ACT-----..............................-G-ACTAA-----T--T--CCAAT--CAC--T--- 7467
A.CD.97.97CD_KCC2         G-GAC-CAA--T-TG----T-------...--A-----------T--C---A--T--GGC--A--------C-G----..........GCGCCC-ACG-C---TGGATCAACAAC-T-GG--G-AGG-A--A------T-- 7458
A.CD.97.97CD_KTB13        G--AC---A--T-T---------A---...--A-----------T------A--T--AGC-----------CT-----...................CCC-T-AAC--C---ACCA-CAAT--C-G-----CCA-T-A---G--C-- 6741
B.AR.04.04AR151516        -------------C-------------...--------------T------A------GT------------------............................C-TGAGGA-A-TA-TAGTG-AG-GTTCA-T-C--A---A-AT-- 6665
B.AU.87.MBC925            -------T-------------------...--------C--T------------------------G-----------..................---GA----GAA--GT-ATGGTCA-AT--C-CTG-A---AAT-A--------- 7484
B.BO.99.BOL0122           -------T-------------------...--------C------------A----G-------------------AAT...........TTTAC--G--CCTGGAAT-TG---AGT---------T--T---C--A---AA-T-TC-- 6663
B.BR.03.BREPM2012         --A-C--T-------------------...---------------------A------CG----------A--G-AA-TGGGTGAATGGT......................----GG----GT---ATGAAG-A--GA--T-- 6907
B.CA.97.CANB3FULL         -------T-------------------...---------------------A------G------------A------.........................GT-AAT------TGGAATAGC-C-G-A-GGTC-AAT--CACTGAT-- 6738
B.CN.05.05CNHB_hp3        G------T-----T--C----------...---------------------A-CT---G-------------------TATAATAATGGTACT...................TGGAATAATGC--GT----GGAATGA-AG---T-- 7478
B.CO.01.PCM001            -C-----T---T---------------...---------------------A--------------------------...................AA-G-G-A--G-ACTT-G-A--GGACTATT-C-G-A-A-TC-AAT-ACACT-AT-- 6645
B.GB.83.CAM1              -----G-T-------------------...---------------------A------G---------C---------...................C-GTT--A-GG-ACTT-G-A---TACTGA-GGGTTA-A-A-CACTGAAAGA-AT-- 7462
B.GB.86.GB8               -------T-------------------...--------------------GA--G-----------------------.........................-A-A-T-C-GGAA-T--TA-T-TA--GT-AA-TACC-C---A-TT-- 7447
B.GE.03.03GEMZ010         ---AC--T-TT--A-------------...--------------T------A------G----C---A----------............................CT--A--A---TACTGGG----G----A-GAAT-A--CT----- 6639
B.IT.05.SG1               -------T-------------------...-----G----------C---T------T--------------------.............................................................-- 7153
B.JP.05.DR6538            -------T---T------C--------...------------------------G-----------------------...................CA-GGAAAC-TAT-AA-T-T-----------T--G-AA--A-AG--AC-- 7498
B.KR.05.05CSR3            --A----T----------C--------...-------C-----------A-------------G---A----------.........................-G-AATG----CTGGAATGATCTA----ACTCTA-TTC-ATA-AT-- 7510
B.NL.00.671_00T36         -------T---------------A---...---------------------A------G-------------------.........................GA-GT--A-G--ACT---AATGG--CT--A---CC--A-AGT--T-- 7023
B.RU.04.04RU128005        ---AY--T-------------------...---------------------A----A---------------------...................GAA-A-AAT------A-TA-TA-TRGG--TGAGACT----A--RG----- 6937
B.TH.00.00TH_C3198        ---AC--T--T----------------...---------------------A--T--A--------------------............................AAT-C--A--GTACT-GG---G-T---AC----GA------T-- 6648
B.UA.01.01UAKV167         ---T---T-----C-------------...--------------T------A--T--A-G-----C------------..............AATAA-GG--C-TGGAC-----A-ATT-ATGAGGG---T---AC---GCCA---CAC-- 6729
B.US.04.ES10_53           -------T-----G-------C-----...---------------------A----G-CT------C-T-G-A-G-TACTTGGAAT..........................GGT-C-GCAGGG-C--AT-ACA-T----- 7443
B.US.99.PRB959_03         -------T-----C-------------...---------------------A--------------------------...............................-A-GG-ACC-AAGAGTTA--T-G-AC-AATGACA-A-TT-- 6657
C.AR.01.ARG4006           ---AC---A--T--C---------A--...--A-----------T--C---A------G-------------A-AT...................TRC---AATG----A---ATTA--GG-TGG--AA-C-ATTCA--MC-C-- 6650
C.BR.04.04BR013           ---AC---A--T--C---------ATG...--A-----------T--C---A--T--GGCT--------G----A-AC.........................-A-G-T-CAGAAAGCAATAATG-A-CAGAAA-T-C--AC-CA-AC-- 6953
C.BW.00.00BW07621         ---AC---A--T--C-----C---A--...--A--G--------T--C---A-----AG----------G----A-AT..............................-A--G-AC---AAATG---CA-A-TC-AC--A-TCA--C-- 6797
C.CN.98.YNRL9840          ---AC---A--T--C---------A--...--A-----------T------A--T--GGC---------G----A-AC.................................-TGCCTAATGGT-CA----G-A-CTC---CTCA-AC-- 6654
C.ET.02.02ET_288          --CAC---A--T--C---------A--...--A-----------T--C---A--T--A-----------CA--AG-A..........................................CTGTCT-AT-A--CA--TT- 6670
C.GE.03.03GEMZ033         ---AC---A--T--C------------...--A-----------T--C---A--T--A--------G-G-A--GAAT-.........................TT-AATGA---CT-CA--G-T--CT-------G-C--A-TCA-GC-- 6669
C.IL.99.99ET7             G--AC---A--T--C---------A--...--A--------------C---A--T--A-C-------------A-AC.................................-GGTCTAATAGT-C---T-G-ACTTC-GA-ACAT-C-- 6631
C.IN.99.01IN565_10        ---AC---A--T--C---------A--...--A-----------T--C---A--T--A---------------A-AC.........................-AA-AC-CG-A--GTACA-AC-G-T-T-A--GTCC---CCCA-AT-- 6842
C.KE.00.KER2010           G---C---A--T--C------------...--------------T------A--T--GGC---------GAA---CTA..............................-C-CGGTTGAATGAT-CAT---G-TCTAATG-GACA-AC-- 6654
C.MM.99.mIDU101_3         ----C---A--T--C---------A--...--A-----------T------A--T--GGC---------G-G--A-AC...................AATGA---AT-CA----T---G-T--A--TA-T-A-TCA--C-- 6836
C.MW.93.93MW_965          ---AC---A--T------------A--...--A-----------T------A--T--A-C-----.......-A-CA.................................-ACCTGTTTA-T--C--TCTAAC-AAT-C-ACA-AC-- 6643
C.SN.90.90SE_364          ---AC---A----C----------A--...--A-----------T--CG--A--T--A-R---------G----AACA..........................................CA-ACT-C--A-TCA-GC-- 6610
C.SO.89.89SM_145          ---AC---A--T--C---------A--...--A-----------T--C---A--T--GGC-------------A-AC.................................-A-CCTAATAAT-CAGGA-G---GTC--ACA--RAC-- 6673
C.TZ.02.CO178             ---AC---A----C-------------...--A-----------T--C---A--T--A-----------G----A-TC..........................-A-GG-ACATTC-A-AGTA-C-AT-CAGGA-AA--TTC--ACTCA...-- 6669
C.UY.01.TRA3011           ---AC---A--T--C---------A--...--------------T--C---A--T--AGC--------------A-ACAAT.............ACA-CAG-G--AGAA--C-G-A-T---A--G-A-G--G-A-T-A--CA-GA-GC-- 6645
C.YE.02.02YE511           ---AC--T---T--C---------A--...--A-----------T--C--CA--T--A--T----------G--A-AC.............A-GTT--A--G-ACATAC-TGCCTAATAGT-C---T-G-ACTTC--A-TCA--C-- 6657
C.ZA.04.04ZASK164B1       G--AC--T-----CA------------...--A-----------T--C---A--T---GT---------CAT-A-AT.........................GT-TT--A-A-TACTG-T-----T--A--G-C--A-C-A-AC-- 6909
C.ZA.05.05ZAPSK240B1      G--AC---A--T----------CA---...--A-----------T--C---A-T---A-C----------..................G-A-A--TAACA---GTAA-TATAA---G---A-A---TTC--CA----GC-- 6838
C.ZM.02.02ZM108           ---AC---A--T--C---------A--...--A-----------T--C---A--T--A---------C--GG-A-TT.........................-A-GG-ACATACTA-TGGAACAGT-CA-CAGGGA-TTC--ACTCA-AC-- 7490
D.CD.83.ELI               ---AC---A------------------...-------C-----------A--T--GG--------------A---..................AA--T----G-A----A--A-ATTACAGAGTCA--T-A-AGCAC--ACACA-AC-- 7004
D.CM.01.01CM_0009BBY      ---AC---A------------------...---------------------A--T--A-------------A--A---AGT...........ATTGG--CAT-GGAGGAT-----A-GGAATGAT-----T----C-TT--A---T----- 6672
D.KE.01.NKU3006           ---AC---A-----C---------A--...--A--------------C---A--T------------------G---.................GA-----A--AG------GA-AC-ACAAAT-G--CAGGG-G-AATGCGAGT-A--- 6708
D.KR.04.04KBH8            ---AC---A---------GT-------...-----------------C---A--T--GG--------------C---..........................-TGGTT-A--G-AC--T-AATC-G-CAGGGTT-AAT-A-AGT----- 7381
D.TD.99.MN011             --AAC---A------------------...---------T-----C---A--T--A--------G----AGAA-TT.........................-ACATA-CAGGGA-TAATAGT-C-----A-AGTAC-GA---T----- 6658
D.TZ.01.A280              ---AC---A--T--C------------...-----------------C---A--T---TC----------C---A---.............................ATT-AGGG-AC-CAAAAT----CAGAGAC-AAT-A-AGT----- 6650
D.UG.99.99UGK09259        ---AC---A-----C---------A--...-----------------C---A--G--AG-T--------TA-GAA---....................-G--A--A-ACATTAG--AGTACAGGG---G---G-A----G-A-----GT-- 6663
D.YE.01.01YE386           ---AC---A-----C------------...--------------T--C---A--T--G-----------CAA-G-GCATGG.................................GGAA-TAC------G...-- 6627
D.YE.02.02YE516           ---AC---A-----------C------...-----------------C---A--T--AC--------C---A--AATT...................AAC-CA-CA-A-GAT-A-----T-AATAGT-C-GGA-A-A-TAC--A---T----- 6642
D.ZA.90.R1                ---AC-T-A------------------...---------------------A--T--GGG-------------A---AATAGGACAGGGAAT.........................AATAGT-G--GT--A-GGTC--A---G----- 6816
F1.AR.02.ARE933           ---AC--T---T-----------A---...--A--------------C---A--T--GG----------..........................................G----A--T-AC-A--G---T-- 6714
F1.BE.93.VI850            ---AC--T---T-----------A---...--A--------------CG--A--T--GG---------GAC--AG-C.........................................TCCAAT-A--G---T-- 6745
F1.BR.01.01BR125          ---AC--T---T-----------A---...--A--------------CG--A--T--AC---------...........................GAC---G----A-T-TCC-A------T-- 6886
F1.ES.x.P1146             ---AC--T---T-----------A---...--A--------------C---A--T--A--T------CA-A--ACATA.........................................AGCTCC-A--GG--T-- 6766
F1.FI.93.FIN9363          ---AC--T---T-----------A---...--A--------------C---A--T---T-----------A---AGTA.........................................CCTAAT-A--G---T-- 6758
F2.CM.02.02CM_0016BBY     G--AC---A--T-------G-------...--A--G-----------CG-CA-----GC---------C.................-CA--ATT-CTTAAC-CA-CA-TG--TAAT-A--GT--T-- 6624
F2.CM.95.MP257            ---AC---A--T-----------A---...--A--------------C--CA---TATT---------CAC--AA-A..............GTAAATGAC-T-CTGTCTAAT---C----GAG--A--A------T-- 6661
F2.CM.97.CM53657          G--AC---A--T-----------A---...--A--G-----------C--CA--GT-GCT---------GCA---AAC......................ATGACTAATGCA-TG-ACAGGTCC-A--G--TT-- 6630
G.BE.96.DRCBL             ---AC---A--T-----C---------...--------------T------A--T--GG---A-------A--G-ATA......................TT-AA-AGT----TAT-A---AAT-A---T--T-- 7404
G.CM.01.01CM_4049HAN      ---ACC--A--T--C--C-----A---...--A-----------T------A--T--G--------.......................-A--T-AAT-TAT-AA-TA-TAGT-CA--T-G-ACT-AG-A---A--T-- 6651
G.CU.99.Cu74              ---AC--T---T-------TC---A--...--A-----------T------A--T--A--------------G-WAT...........................ATGKG---T-G-ACT--G-A--G-C--T- 7081
G.ES.99.X138              ---AC---A--T-----------A---...--A-----C-----T------A--T---G-------------G-CT-................TC--A-AGT--TAACCTG-CT-A-AGTA-T-A------T-- 6900
G.GH.03.03GH175G          ---AC---A--T-----C-----A---...--A-----C-----T------A--T--AGG----------.....................AATGA--C-GGG-A-AGCAC--ACACA-AT-- 7505
G.KE.93.HH8793_12_1       ---AC---A--T-----------A-G...--A-----------T------A--T--GG-----------A-GAGACC.........................................AT-A-C-A--GG--C-- 6829
G.NG.01.01NGPL0669        --AAC---A--T-T---------A---...--------------T------A--T--G-CT---------A---A-CA.........................................--GTAT-AG-G--AT-- 6661
G.PT.x.PT2695             ---AC---A--T-------T---A---...--A-G------T--T------A--T--GC--------------G-A-TCTGTCT.......................----GT-A-AGTA-T-A------C-- 7470
G.SE.93.SE6165            ---AC---A--T-----C-T---A---...--A--G--------T------A--T--GGC----------TCA-GCCT-.......................TT-AG-TCA--T-G-AGT-AG-A--GG--T-- 6871
H.BE.93.VI991             ---A----A--T---------------...--A--G--------T--C---A-----A--T----------C--C---ACCAAT.............AGCAGCT---CC-A-GAC-CCTA-A-CA-CAAC-GC-CAGAA--TATC-CA-GA-AT-- 6870
H.BE.93.VI997             G--AC---A--T--------C---A--...--A-----------T--C---A--T--GGG-------------G----..........................ACTG-GGACA-TATTAAT-TGCCA-A---CAC-G-AA-A-AT-- 6779
H.CF.90.056               G--AG---A--T-----------A---...--A-----------T--C---A--T--GGG-------------G----...................GAA-TGCAT-C--A-T-CACA------G----AA-G----AC--A-AC-- 6784
J.CD.97.J_97DC_KTB147     ---A-G-T---T-C-------------...--A--G---------------A--T--AT------C-------G----GATGAGAAT.........................-GC-CTGTGACTAT--A---A---G- 6659
J.SE.93.SE7887            -----G--A--T---------------...--A--------A--T------A--T--A-------C-------G----..............GA-AA--A--GCATT-AAG-T-CA--TGACAC-A-TNAC-CA--T-- 6760
K.CD.97.EQTB11C           G-CAC--G-----T------C---A--...--A-------C-------G--A-T---G-CAC-G--G--GAC--AGA-..............................................GAG----AGGA-AC------ 6642
K.CM.96.MP535             ---C---------TG----------C-...--A--------------C---A-----A-----------GAG--AG--...............................................-AG-A--GT--T-- 6628
```

```
                                                                                     V4 loop start
B.FR.83.HXB2                 ATTGTAACGCACAGTTTTAATTGTGGA...GGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGTACTTGG..................................TTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAAT 7465
Env                          _I__V__T__H__S__F__N__C__G__.__G__E__F__F__Y__C__N__S__T__Q__L__F__N__S__T__W__.__.__.__.__.__.__.__.__.__.__.__F__N__S__T__W__S__T__E__G__S__N__N__T__E__G__S__D__T__I
01_AE.CF.90.90CF11697        ---AC--T---TCA---------A--...-------------T--C---A-----A------------------..............................-CA--AAATGAA-T-ATG-A-GAGTT---GGGCAC-AACTCAAG---T-- 7423
01_AE.CN.05.FJ051            ---AC--T---TCA---------A--...-------------T--C---A-----A-T----------------......................................-A-GGAACCAT-GAGGG---TGGCACT-AT-A--G---T-- 7447
01_AE.CN.06.FJ054            -C-AC--T--TTCA---------A--...-------------T--C---A--T--A----------......................................................G-A--CAG---T-A--G---T-- 7459
01_AE.HK.x.HK001             ---AC--T---TCA---------AA-...-------------T--C---A-------------G------G----.........................................-A-AC-ACAA---CC--T-GG--GAACC--AGT-ATT- 6817
01_AE.TH.01.01TH_R2184       ---AC--T---TCA---------A--...-------------T--C---A-----A-----------A----........TGCATAAATGACACCGGGAACAATGAAACCAAAGGGGAGTG-.............................GAAA---AC-- 6665
01_AE.TH.02.OUR769I          ---AC--T---TCA-----C---A--...-------------T------A-----A--T---------C------..........................ACA-A-GGAACCAT-GA-C---T--TAC-TAAT-ACACTGAC-- 6652
01_AE.TH.90.CM240            ---AC--T---TCA---------A--...-------------T--C---A-----A-----------A----........TGC............CTAGGAAATGAAACCA-GGCGG-GTG-................................-A------T-- 7018
01_AE.US.00.00US_MSC1164     ---AC--T---TCA---------A--...-------------T--C---A-----A-G------------------......ACAGGAAATGAAAGCAGGGATGGGTGTAATGGC-C--CGTG-....................................-AG-- 6658
02_AG.CM.02.02CM_4082STN     ---AC---A--T----------------...--A-----------T--C---A--T--GG-----------C-G----...........................................GA--GCAAT--C-----G-AAA-C-----TT-A---AAT-A---A--T-- 6643
02_AG.EC.x.ECU41             G--AC---A-------------------...--A-----------T--C---A--T--GGC-----------C------...........................................GA-GT-AAC-A--GCAC-AA------CT----AATGCC--A------T-- 6530
02_AG.GH.03.GHNJ196          ---A-G--A--T-C--------------...--A-----------T--C---A--T--AG-----------C-A----.................................................AAA-ACGG-ACT-CCAGCTCA--TGGCACT-C--A-----TT-- 7506
02_AG.NG.01.PL0710           ---AC---A--T----------------...--A-----------T--C---A--T--A------------C---G--......................................................--TACCAAT-CC-CTG-CA-CTC--CA-GA--T-- 6616
02_AG.NG.x.IBNG              ---AC---A--T----------------...--A-----------T--C---A--T--A------------C------..............................................GA-AAT--C-A-AGCACTG-C--CC----A-GGTC--A------T-- 6974
02_AG.SE.94.SE7812           ---AC---A--T----------------...--A-----------T--C---A--T--A-T-----------A------AATCATAATGGCACTTGGAATGCA.........CCAGGACCGTT--A-GAC-C-GAG--TAAAA---TA--TGGCACT-AGGA-A-A----- 6852
02_AG.SN.98.MP1211           ---AC---A--T--C-------------...--A-----------T--C---A--G--G------------C------GCTAGCAAT............ACT.........AATGGC-T-T-GG--A-C-A--TAA-T-CTAGC-------AG--C-C--A------T-- 6667
03_AB.RU.97.KAL153_2         -------T--------------------...-------------------A-----A------------C------...........................................-A-GG-ACT-AAGAGTTA--T-ACACT-A-G-A--T-T-G- 6642
04_cpx.CY.94.CY032           ---AC---A--T--C---------AA-...--A-----------T--C---A------C--------------ACAC.............................................-TGCAAAATG----AA-CATTA---G--CAGA-TCTAC--A-TCA--C-- 6841
05_DF.BE.x.VI1310            ---AC--T---T-----------A--...--A--------------CG--A--T--A--T--------GCC--AGT-...............................------GA---AGT-TT--A--CCACAATGTT---T-A---CA-TGACA-G-AT-- 6862
06_cpx.AU.96.BFP90           ---AC---A--T----------------...--A-----C-----T------A--T--A-C----------CAT-AGAC..........................................CTGTT-A-CACC----GGGGT-A---CACC-ACAC---T-- 7516
08_BC.CN.97.97CNGX_6F        ---AC---A--T--C---------A--...--A-----------T------A--T--GGC---------G----A-AC.............................................-A-GG-ACATAC-TGCTTAATGGT-CA-GAGA-A-TTC---CTCA-TC-- 6641
09_cpx.GH.96.96GH2911        ---AC---A--T----------------...--A-----------T--CG--A--T--GGC--A-------AC------...................................-A-GG-ACTT--GA--------A-A-TA-C-TA--A-ATTA 6643
10_CD.TZ.96.96TZ_BF061       ---AC---A----CC-------------...------G----------C---A--T--G---------------A---............................................-CAT-AA-TAATA---G--CAGGG-T-AATG-GAGT----- 6814
11_cpx.GR.x.GR17             G--AC---A--T----------------...-----------------C---A--T--GGC-----------C------.............................................T--GCA-A-GAC-A---CAGCACACA----G---TGC-GTC--A------T-- 6726
12_BF.AR.99.ARMA159          ---AC--T---T-----C------A--...--A--------------C---A--T--GG----------GA---AATA..............................................................TCC-A---T--T-- 7427
13_cpx.CM.96.1849            G--AC---A--T----------------...--A-----------T--C---A--T--GCC-----------C------..................................GAA---GAC---GTC--C--GC----------G----A-GGTC--A------T-- 6884
14_BG.DE.01.9196_01          -------T-----C--------------...--A--------T---------A--------T------C---------.......................................AA-TC--A-AGTACTGAG---G-T--A--GAAC-A-AG-TT--- 6973
15_01B.TH.99.99TH_MU2079     ---A--GT---------C----------...-------------------A--T-------------------------...............................................TT-CAGAATG---A-AGTACT-GG---G-T----C-----A--G---T-- 6825
16_A2D.KR.97.97KR004         ---AC---A--T---G-C----------...--A-----------T--CG--A--T--GGC-----------C--G---.......................................CCTGCC-A--C-TCAAGGG-G--TGAG--GAA-GA-CG--ATG- 6831
17_BF.PE.02.PE02_PCR0155     ---AC------T-----------A--...--A--------------C---A--T--GG----------GAC---G-C...................................................---ATCTCC-A------T-- 6631
18_cpx.CU.99.CU76            ---AC---A--T-C--------------...--A-----------T------A-----GCCT-------------C---AAT............................AAAACARR-G----C----TA-AACT--ATAAC-C--CA-A-RCGCA--A--GT--T-- 6773
19_cpx.CU.99.CU7             --AAC---A--T----------------...--A-----------T------A-----G-C-------------GAA-A..................................................................G--T-- 6544
20_BG.CU.99.Cu103            ---AC---A--T-----------A--...--A-----------T------A--T--GG---A----------G-AAT.........................................................A-TAAT----GA-AT-- 6896
21_A2D.KE.91.KNH1254         ---AC--AA-------------------...--------C--T------G--A-----G------------C--A---.........................................TA--GGAATG----ACT-ATTAATGG--CA-AGTC-ATT-A--G------ 6672
22_01A1.CM.01.01CM_0001BBY   ---AC---A--T-----------A--...--A-----------T--C---A--T--GGC-----------C------...................................CA-GACA-CACAA--GGGTTAGA-A-CACG-A--G---C-- 6632
23_BG.CU.03.CB118            ---AC---A--T-------T---A--...--A-----------T------A--T--GG-----------C-G-AAT.........................................................A-TAAT---AGT-AT-- 6896
24_BG.CU.03.CB378            --CAC---A--T-----------A--...--A-----------T------A--T--A----------GAGT-ACT-.....................................CT-A-TA-TAGC-GG--TGA-AGTAAG-A---T--T-- 6905
24_BG.CU.03.CB471            --CAC---A--T-----------A--...--A-----------T--C---A--T--GG-------------G-AAT.................................-----T-G-AGT-AG-AG-GG-AT-- 6905
25_cpx.CM.01.101BA           ---AC---A-------------------...--A-----------T--C---A--T--GCCT--------G---G----.............................-CA-A-GAC-C-G--A-TAATGGG-CAG-A---AGCAAT-A---T--T-- 6715
27_cpx.CD.97.97CDKTB49       G--AC---A--T-C----------ATG...--A-----------T--C---A-----AC-T-------------C---.......................................-A-GA-ACCACTAAGGGC-G--A---GACC-A-AGT--T-- 6663
28_BF.BR.99.BREPM12609       -----G-T-----C--------------...------------------CA------TG--------C---------..............................-AGTGGAAT-AC---TGGAATAAC-C-GGG-AG-GT-CTGA-A-G-AT-- 6857
29_BF.BR.02.BREPM119         -------T------G-------------...-------------------A--T--A-G-----C-----------...............................-A-GCT-C--ACATG-AAGG----TC-----------A-TCT-AT-- 6611
31_BC.BR.02.110PA            ---AC---A--T--C---------AA-...--A-----------T--C---A--T--AG------------AG-A-ACTTT.................AAT.........AGTAC-KC-CACCTKAAT-----ARGTAACAGT-C-GGA------ACWTCAAC--TC-- 6958
32_06A1.EE.01.EE0369         ---AC---A--T----------------...--A-----------T------A--T--A-------------CT-A-CA...................................................................AG-A-T-A---G--T-- 7091
33_01B.MY.05.05MYKL007_1     ---AC--G---TCA---------A--...-------------T--C---A--GA-A-----------A------C.........................................-TA-G-AATGA--CC-TGGGA-GGT-T-A-A----T-- 6841
34_01B.TH.99.OUR2478P        ---AC--T---TCA---------A--...-----G--------T------A-----A--T-----------------CCAGAAAAT.................................GG--CC-TGGAG-GGA-T-A--G---T-- 6631
35_AD.AF.05.05AF095          ---AC---A--T----------------...--A--G--------T---G--A--T--GGC--------C-----C---AATATCACTGACATAGGA.................--GCCAAAT-GC-CA-A-AGCAC--A------T-- 6628
36_cpx.CM.00.00CMNYU830      G--A----A--T----------------...--A-----------T--C---A--T--GGC-----------C------GGGTATAACAATAGC.....................A-C-T-A---GGTTA-A-A-CACG-A------T-- 6662
37_cpx.CM.00.00CMNYU926      ---AC---A--T------------C-...--A-----C--------C---A--T--GGC-----C--------A---AACACTACTACCAGC.....................ACGCA-ATTCCA--T-GGACTAC-GA--GG-AT-- 6647
38_BF1.UY.05.UY05_4752       G--AC------T-----------A--...--A--------------C---A--T--A--------.........................................G------TA---AAC-A--GA--T-- 6867
39_BF.BR.03.03BRRJ103        G--AC--T---T-----------A--...-----------------C---A--T--GG----------AC--AGTA.....................................................TCC-A--G---T-- 6925
40_BF.BR.04.04BRRJ115        ---AC--T---T-----------A--...-----------------C---A--T--GG---------GAC--AGAA...............................-T-GAA-CTAATCAC-CA--TG--A-T-AT-A------T-- 7013
42_BF.LU.03.luBF_05_03       G--AC--T---T-A---------A--...--A--------------C---A-----G----------GAC--AGAA.............................................TCCAAT-ACACT-TC-- 6957
43_02G.SA.03.J11223          ---AC---A--T----------YA--...--A-----------T-----CA--T--AG-------..................................--ACTGTT---TGMATC---GTC---T-ATC- 6948
U.CD.83.83CD003              ---AC---A--T-C------C-------...--A--G--------T--C--CA--T--G-G------C-G---T----AATGGT...........................ACT-GG...GA--AGAATT-C---AGCACTGAG......-G-A-TT-C-C--GA-AT-- 6938
U.CD.90.90CD121E12           ---AC---A--T--C-----C-------...--A--G--------T-----CA--T--GGG--A-----C---------AATGGT...........................ACT-GGGC-GAG-A-AATT-C---A-CA-TGAG-G---T-GCAGTT-C-A---A--C-- 6633
U.GR.99.GR303                ---AC---A--T-----A----SY--GGATC-------------------A--T--GGC--------C--C-----.................................-GGCCA-ACGAC-TGCAGTGGTCAAATG-C-CGGGCA-TAT--CGC-TTAT-A 6749
U.NL.95.U_NL_95_H10986_D1    ---AC---A--T-----------ATG...--A-----------T------A------GC----------CA--ACACCTG.................................-A----------G---A-GACAC-CAAGAT-------G-ACCAAT--CA-A-AT-- 7365
CPZ.CD.90.ANT                G-AAA-GT---TT-G--------CCA-...--A-----C--T--T---G--AT-----CTTG---C---GCC--A-AC........................................A-GGGA---CTCATCAC--AC-GAG-CC- 6900
CPZ.CM.01.SIVcpzCAM13        --A-C--AT----TG--C--C-------...--A-----C--T-----CG--A-T-G-A--A--------.....................................G-A-GTGAACCTTTTCAC--G-AC-- 7046
CPZ.CM.05.SIVcpzEK505        G-AAC--ACTTA-TGA-C----------...--A-----C--T---------A-C-GT---T------CA-A-CAAAAT.......................................GGCA-CAC--C--G--AT-- 6951
CPZ.CM.05.SIVcpzLB7          G-G-C--AC-TTCA--------------...--A--------T--T-----CA-CT-GGCCT----C---GAC--C-T-...........................-TAC-C-A-A--ACC-TGCTC--TGAAACCTTGCACA-T-ACT- 6985
CPZ.CM.05.SIVcpzMB66         G-AAC--AC-TTCACC-C-TC-------...--A--G--C--T--T--C---A-CT-CAT---C-------CCT-CATC...................................ATCTT---TGAGACTAAGGA---T-AT-- 7086
CPZ.CM.05.SIVcpzMT145        G-GACT-AC--TCAC-----C--C--G...--A-----C-------------A---GTG-GA-CA-C----TC---AA-.......................................-TA-ACA--ACTGAGA-T-TG-C 6944
CPZ.GA.88.GAB1               G-CAC-CAT--T-TG-------------...--A-----------T------A-T-GC---A-A----C-GAC-ACATT..........................................AC--A--G--TT-- 7484
CPZ.TZ.00.TAN1               G-CACT-GCTT-T-G--C------CA-...--A-----C--T---------CTC---A-TTG-AC----....................-CCTGGAC--CTAAT-GA-C--A-A-TACTCA--GT--TC- 7088
CPZ.US.85.CPZUS              G-G----ATATG-TGC------------...--A--------T--T------A---TT-CC--A-----C-TG--C-AC........................................----A-ACC-AC-ACAC---T-- 7461
CPZ.CM.x.SIVcpzDP943         --A----ACATG-TG--C--------G...--A--G-----T--T-----CA-C--C-C-T----C----AC-G----.........................--AA-TACAA-T-GCTCAGA-AT-AAT--CTCA-AT-- 7010
CPZ.CM.x.SIVcpzMB897         G-AACT--A--TTA---C----------...--A-----C--T---------A--T--A-CT-A--C---..................................CAATCCACT-T-A---AT-- 6979
N.CM.02.DJO0131              G-AAC-CACTTA-TG--C----------...--A--G---A---T------A-TT-CAG-T-A--C---GAG-GCAT-.................................----CA-A--GGACT-A--GG--T-- 6898
N.CM.04.04CM_1015_04         G-AAG-CACTT--TG--C----------...--A--G------CTT-----CA-TT-CAG-T-A---..........................................GATG-G---GGGACT-T--A--GG--T-- 6882
N.CM.04.04CM_1131_03         G-AAC-CACTTA-TG--C----------...--A--G------CTT-----CA-TT-CAG-T-A------..............................-A-AGC--G---GAGACTAC--A--GG--T-- 6938
N.CM.95.YBF30                G-GAC-CACTTA-TG--C------A--...--A--G--------T-----CA-TT-CA--T-A------GAGGAA-TA.........................................CTT-AC--GAC-G----GC-T-- 7013
N.CM.97.YBF106               G-AAC-CACTTA-TG--C------RR-...--A--G--------T--C--CA-TT-CAG-T-A------GAG-GCGA-.....................................-ACA-GACT-A-A-G--T-- 6974
O.BE.87.ANT70                G-AACCCATTTACAC-----C---CAT...--A-----C--T--T-----CA--G-TA-GA--------TA---C-TT...TCATGTAACGGA.................ACCACCTG----...............-TTAGT...--TGT-AGTCA-G--A---ATGG 7503
O.CM.91.MVP5180              G-AAGCCATTTACA------C---CAT...--A-----C--T--T-----CA--T-TGGGA-------CTA-----TTATCAACTGTACAAAG.................TCCGGATGCCAG............--GATCAA-GGG-G--A---GACC-A-A-A-ATGG 7520
O.CM.96.96CMABB637           G-AAGTTATTTGCA------C---CAT...--A-----C--T--T-----CA---GTG-GA--------TAC--C-TT...AACTGTACCCCG.................ACCAAATG--A-...GAA----G-T-CAATAGT-CAGGT-G-T-CAAT---ACACTT-G 7008
O.CM.98.98CMA104             --A-CCTTTTTGCA------C--CCAT...--A--G--C--T--T-----CA--T-GGGGA--------TA---C-TT...ACATGTAACAAA.................ACCAACTGC-C-...............AACAAT-GG-CT...--TAAT--CACT-ATGA 6965
O.CM.99.99CMU4122            --A-CCCACTTGCA-----GC---CAT...--A--G--------T-----CA----C--GA--------TA---C-TC...TCATGTAACGGA.................ACCAGCTG--C-......-A--ACC--A-TAAC-G-G-K-ACA-GAATG-CACATTG-C 6967
O.FR.92.VAU                  G-AAC--ATTT-TT------C---CAT...--A-----C--T--T-----CA---AT--G---------CA---C-TT...TCCTGCAAGAAGAATATGACCTATAACAAGATCAA-TG--C-...................AAT-T--G--A-A-TA-C-A--G---TCA 7063
O.SN.99.SEMP1299             G-AACCCATTTGCA---C--C---CAT...--A--G--C-----T-----CA--T-TA-GA--------TA---C-TT...TCATGTATAGGA.................ACCAACTG--C-...............A-TAACC-A--T-GCAGCAAC---A-TGAC-C 7550
O.US.99.99USTWLA             G-AACCCATTTGCA------C---CAT...--A-----C--T--T-----CA---AT---A--------TA---A-TC...GAATGTAATAAC.................TCCAACTG---G.........-T-C-GAATGAC-----T---T-T-A--ACAGT----G 6974
```

```
                               V4 loop end┐
B.FR.83.HXB2             CACC...CTCCCATGCAGAATAAAACAAATTATA...AACATGTGGCAG...AAAGTAGGAAAAGCAATGTATGCCCCT..............CCCATCAGTGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGT.......... 7601
Env                      __T__.__L__P__C__R__I__K__Q__I__I__.__N__M__W__Q__.__K__V__G__K__A__M__Y__A__P__.__.__.__.__.__P__I__S__G__Q__I__R__C__S__S__N__I__T__G__L__L__L__T__R__D__G__G__.__.__.__
A1.GE.99.99GEMZ011       A--T...--A--------------G---------...--T---------...-G-------C-----------------..............--T----AA--A-T-A-G--GA------C--------A--A---C---T-----------AGGT....... 6776
A1.KE.00.KER2008         A--T...----G------------G-----C---...--T---------...-G-AC----C-----------------.............-----TCAA--AT--A-AG--GT------C--------A--A-G-----------------G......... 6810
A1.KE.00.KNH1144         A--T...----A------------G---------...--T---------...-G-AC----C-------A---------.............------CAG--GTG-A--G--GA------C--C-----A--A--G-----------------GGAG....... 6833
A1.KE.00.KSM4024         A--T...-----------------G---------...--T---------...-G-------C-----------------.............------CCA--GT---A--T--GT------C--------A--AA-----GT---------A--G......... 6733
A1.KE.00.MSA4069         AGTT...-----------------G---------...-GG---------...CG-------C-----------------.............------CCA--AT---A-----GAC-----C--------AT-A--------------------G......... 6812
A1.KE.00.NKU3005         A--A...----A------------G---------...--T---------...---------C-----------------.............----ACAA--GT---A--G--GA------C--------A--A--------------------A......... 6815
A1.RU.00.RU00051         A-TT...-----------------G--G------...--T--------A...-G------GC-G-----------T---.............-T---AA--ATC-A--G--GA------C--------A--A--------------------GAATAAA.... 6875
A1.RW.93.93RW_024        A--T...----A------------G---------...-GA--------A...-G-AC----C-G---------------.............------AA--G-T-A--G--GA------C--------A--AA-------------------G......... 6803
A1.SE.95.SE8891          A--T...---------------------------...--T---------...-G--C----C-------A---------.............------CAA--GT---A--G--GA------C--------A--AA-------------------G......... 6766
A1.SE.95.UGSE8131        A--T...----A------------G---------...--T---------...-G-AC----C------C----------.............------CCA-GGT---ACAG--AG------C--------A--A----------G--------GGTA....... 6979
A1.TZ.01.A173            A-TT...----A-----A------G---------...--T---------...-G--C----C-------A---------.............------CAA--GT---A-AG--GT---------------A--AA-----------------AG......... 6810
A1.UA.01.01UADN139       A--T...--A--------------G---------...--T---------...-G-------C-----------------.............-T----A--A-T-A--T--GT------C--------A--A--C-----------------A......... 6797
A1.UG.92.92UG037         A--T...-----------------G---------...--T---------...-G-------C-----------------.............------CAA--GT---A-AG--GA------C--------A--AA-------------------G......... 7611
A1.UG.99.99UGA07072      A--T...----A------------G---------...--T---------...-G-------C-------A---------.............------CAA--GTG-A-AG--GA------C--------A--AA-------------------G......... 6807
A2.CD.97.97CDKTB48       A--G...----A----------------------...------------...-G-----CCG-----------------.............-----TGCA--GT---A-AG--A-------C--------AA--A---G--------C-----G......... 6914
A2.CY.94.94CY017_41      A-T-...---------------------------...------------...-G--------G----------------.............-----TGCA--AT---A-AG--A-------C--------AA-AA---G---------------......... 6976
A.SN.01.DDI579           A--T...----------------G-------G--...C-T---------...-G-------C-----------------.............-----TCAA-GGT---A-AC--A-------C----------A--G------G--------G......... 6805
A.SN.01.DDJ369           A--T...---------------------G-----...--A--------A...-G-------C-----------------.............-----TCAA--GT---A-A---GT------C---------A-A--GC------------G--G......... 6823
A.SN.96.DDJ360           A-AT...----------------G-------G--...C-T---------...-G-------C-----------------.............-----T-AA-GGT---A-AC--A-------C----------A--GC------G--------G......... 6787
A.CD.02.02CD_KTB035      TCTT...T-AT-C---------C--------......--T---------...-G-----T-C--C--------AA----.............-----AGAA--TT---A-T----A----C--------CA--AT---G----C-----------......... 7600
A.CD.97.97CD_KCC2        G--T...--G------------------------...--T---------...-G-----A-C-----------------.............--T----AA--A-C--A-CT--GT------C--------A--A--G-----------------AACT....... 7597
A.CD.97.97CD_KTB13       A-TT...--G-----T-------T----------...--T---------...-G-C-----C-G---------------.............------AA--TT---A-AT--GT------C--------A--A--G--G-----G--------A......... 6877

B.AR.04.04AR151516       --YA...--------------------------C...--T-G-------...G----G---------------------.............--------A-----------------------------T-------------------AGTGGG.... 6807
B.AU.87.MBC925           ---G...--------T-A----------------...-----G------...G----------------A---------.............--------A---------A-----------C-------------C----------------C......... 7620
B.BO.99.BOL0122          ---A...---T-----------------------...------------...G--------------------------.............-T----A---------A---------------------------------------G......... 6799
B.BR.03.BREPM2012        ----...------------------------G--...-----G------...G--------------------------.............--------A---------AC--A-----C-A-------------C-------------------......... 7043
B.CA.97.CANB3FULL        --TA...--------------------------C...------------...GG---------C---------------.............--------C-------------------------------------------------------......... 6874
B.CN.05.05CNHB_hp3       T-TA...-----T------------------G--...-----G------...G-------G--------A---------.............-----TGCA---C----AT----------A--------T-A------------------G......... 7614
B.CO.01.PCM001           ---A...--------T-A------G---------...--T---------...-G-------------------------.............--------A--GT---A-----GT------C--C--------A--------------------GAAC....... 6784
B.GB.83.CAM1             T--A...--T------------------------...-----G------...G--------------------------.............-------CA--AC-----C------------------------------------------------......... 7598
B.GB.86.GB8              ---A...--------T---------------G--...------------...G--------------------------.............-------CA---------------G-----C----------------A--------------A---AGGGAG.... 7589
B.GE.03.03GEMZ010        ---A...-----G---------------------...----GA------...G--------------------------.............--------A---T-----------C---------------------A-----GT---------------......... 6775
B.IT.05.SG1              ---A...G----------------G------G--...------------...G--------------------------.............-----A--A--TGG---AC---------G-----------T------------------A-A-......... 7289
B.JP.05.DR6538           ---A...-----G--T-------G-------G--...------------...-C-------------------------.............-------C--T-C---A---A------A------------------A-------G-------------AAAG....... 7637
B.KR.05.05CSR3           ---A...---T--------------------G--...------------...G--------------------------.............------AA---------C-------------------------A-------------------AAGGACAATA 7656
B.NL.00.671_00T36        A--A...---------------------------...-----GA-----...G--------------------------.............-----TGCA---------C----------C--C---------------------G---------GGT....... 7162
B.RU.04.04RU128005       ---A...---------------------------...-----GA-----...G--------------------------.............------AA---------A--C---C---------------------A---------------------......... 7073
B.TH.00.00TH_C3198       ---G...-----T-----------G------G--...------------...G--------------------------.............------GCA--A-C--C--C--------------------T-AA--------T-G---------AATAAGAACG 6794
B.UA.01.01UAKV167        ---A...---------------------------...-----GA-----...G----------------A---------.............--------C---T----------------------A-A-----GT----------------......... 6865
B.US.04.ES10_53          ---A...----A-----------G----------...------------...G--------------------------.............------CCA--T-CC--AC------------------A--------------G-----------......... 7579
B.US.99.PRB959_03        ---A...---------------------------...------------...G--------------------------.............-T----A-------------------------------------------------------AATAAT.... 6799

C.AR.01.ARG4006          ---A...A--T-----------------------...-----------A...G-G-----CG-----------------.............------T--A---ATG--A-C----AG---------C-----AA-A---C-G---C-T-----A--CACA....... 6789
C.BR.04.04BR013          ---A...----------A------------G---...CG----------...G--------C-----------------.............----G-T-AG--ATG--A-C----AG---------C-----AT-A-----G---C-G-----A--CTCT....... 7092
C.BW.00.00BW07621        ---A...----A----------------------...------------...GGG-----------------------G.............-----TGCA--A-C--A-T---AA------C--------A--A-----G---C-T-----A---......... 6933
C.CN.98.YNRL9840         ---A...-----------------G---------...------------...G-G-----CG-----------------.............-----TGCA--A---A-C---AA---------C-----A--A-----GGT-C-T-----A--AACAGAGAATA 6800
C.ET.02.02ET_288         A---...A----------------G---------...------------...G-G-----CG-----------------.............-----T-AA--AT---A-C---AG---------C-----AT-AT--C-G---C-T-----G--GATG....... 6809
C.GE.03.03GEMZ033        ---A...---------------------------...------------...G-G-----CG-----------------.............-----TGCA--T-C--A-C---AAT---------C-----A--A--C--G---C-T-----A--CAAT....... 6808
C.IL.99.99ET7            --TA...---------------------------...------------...GGG-----CG-----------------.............-----TGAA--A-C--A-C---AA-----------------A--A-----G---C-T-----A--A......... 6767
C.IN.99.01IN565_10       ---A...A--------------------------...------------...G-G-----CG-----------------.............-----TGCA--A-C--A-C---AA---------C-----AA-A-----GGT-C-T-----G--AAAA....... 6981
C.KE.00.KER2010          ---A...--------T------------------...------------...--G-----CG-----------------.............-----TGCA--A-C--A-C---A--------A--C-----A--AT----GCA-C-T-----A--G......... 6790
C.MM.99.mIDU101_3        ---A...-----------------G---------...------------...G-G-----CG-----------------.............--T--TGCA--A-C--A-C---AA-------C--C-----A--A-----GGT-C-T-----A--A......... 6972
C.MW.93.93MW_965         ---A...---------------------------...------------...GGG-----CG-----------------.............-----TGCA--A-C--A-C---AA--G----C-----A--A---C-G---C-T-----A--A......... 6779
C.SN.90.90SE_364         ---A...A---A---------------G------...------------...GG------CG-------A---------.............-----TGCA--A-C--A-C---AA---------C-----A--A---C-G---C-T-----A--A......... 6746
C.SO.89.89SM_145         ---A...----A----------------------...------------...G-G-----C------------------.............-----T-AA--A-C--A-C---A----------C-----A--A-----GCTGC-T-----A--AAATGATCCAA 6819
C.TZ.02.CO178            --T-...........--------G----T-----...----CA------...--G-----CG-----------------.............-----AGCA--A-C--A-C---AAC---GC--------A--A-----G---T-G-----A--CAGTAGTAAA. 6805
C.UY.01.TRA3011          T--A...A--------------------G-----...-----------A...GGG-----CG-----------------.............-----TGAA--A-C--A-C---AGG---------C-----AT-A---C-G---C-G-----A--A......... 6781
C.YE.02.02YE511          ---A...-----T---------------------...------------...GGG-----CG-----------------.............--T--TGCA--A-C--A-C---A-------C--C-----A--A-----G---C-T-----A--A......... 6793
C.ZA.04.04ZASK164B1      ---A...--G------------------------...------------...G-G--G--CG-------A---------.............-----TGCA--A----A-C---AA---------C------T-A-----G---C-T-----A--AAATTCA.... 7051
C.ZA.05.05ZAPSK240B1     ---A...---T-A-----G---------------...------------...G-G-----CG-----------------.............-----TGCA--A-C--A-C---AAT---C---C-----A--A--------C-T-----A--AAAC....... 6977
C.ZM.02.02ZM108          ---A...A---------A----------------...------------...G-G-----CG-----------------.............-----TGCA-GA---A-C---G----------C-----A--A-----G---C-T-----A--AGTATCAAATA 7636

D.CD.83.ELI              ---A...----A----------------------AAG-TGG--GCAGGCAGG---.........-----A---------.............--T---GAAA--A-C---CT------------------A----G-----------------......... 7137
D.CM.01.01CM_0009BBY     A---...--------T------------------...------------...G--------------------------.............------GCA--------A---A------A---------------------AT---G--------------AATAATAATT 6818
D.KE.01.NKU3006          --TA...---------------------------...------------...G--------------------------.............-----T--A--GCC--C-AC---T-----G-------A--AT-G--G-----------A--G......... 6844
D.KR.04.04KBH8           ---A...----A----------------------...-----------A...G---------G----------------.............-----TGAA--GT---C-A-----------C-----------AA---G-------------CA......... 7517
D.TD.99.MN011            A---...A--------------------------...------------...GG-------------------------.............------GCA---------A-------------------------A----G-----------------......... 6794
D.TZ.01.A280             ---A...A--------------------------...------------...GG-------------------------.............------GCA---T---C-----A-----A---------------A--AT--C-----------------AAT....... 6789
D.UG.99.99UGK09259       ---A...----A----------------------...------------...GG-------------------------.............-----TGAG---T---C-CC--------------------------A--AT-G--G-----------------GCA....... 6802
D.YE.01.01YE386          ---A...----A--T---------------C---...------------...GG-------------------------.............-----TGAA---T---C--G--------G----------C-----------AT----G-----------G--C......... 6763
D.YE.02.02YE516          A---...A-------T------------------...--T---------...GG--------G----------------.............------GCA--AT-----A------------------------A-----G-----------------AGT....... 6781
D.ZA.90.R1               ---A...--T-------------G----------...------------...GG-------------------------.............-----AGAA---T--------------------------------AA---G-----G---------GAT....... 6955

F1.AR.02.ARE933          ---T...--------TC--------------G--...------------...G-------CG------ACA--CA----.............---CATGCA--A-C----CC--GGT---------------T--A-----G----------------ACTGGGAATT 6860
F1.BE.93.VI850           ---T...--------TC-G------------G--...------------...GG------CG-----------A--T--.............-----TGCA--A-C----CC--AAC---------------TT-A-----G-----------------......... 6881
F1.BR.01.01BR125         ---T...--------T-A-------------G--...------------...G----G-GCG--------------G--.............-----TGCA--G-----CC--AG---------------T--A-----G--------------CCGAAAT.... 7028
F1.ES.x.P1146            ---T...--------T-----------G------...------------...GG-------C-----------A--G-C.............-----TGCA--A-C----C---AAC---------------CT-A--CA-----------------CAGGAG.... 6908
F1.FI.93.FIN9363         ---T...--------TC-----------T--G--...------------...G-------CG--------------G--.............-----TGCA--A-C----CC--AAC---------------T-------G-----------------CAG....... 6897
F2.CM.02.02CM_0016BBY    --TT...A-T-----T---------------G--...---G-------A...-G-----GCG-----------------.............-----TGCA--A----CAG-CAAC---------C-----T--A-----G---------------A......... 6760
F2.CM.95.MP257           ---T...---T----T---------------G--...-----------A...-G------C-G----------------.............-----TGCC--A----CAG--AAC---------C-----T--A-----------T---------G......... 6797
F2.CM.97.CM53657         ---T...---T----T-------G-------G--...-----------A...-G------CG--------------G--.............-----TGC---------CAG--AAC---GC--C-----T--AA---G---------------G......... 6766

G.BE.96.DRCBL            ---A...--TAAT--T-AG------------G-G...-GA---------...-G---G---C-----------------.............-----TGCA--A-C----C----AG-----C--------A--AA-------------------G......... 7540
G.CM.01.01CM_4049HAN     ---A...--------T-A-------------G--...-GA---------...-G---G---C-----------------.............--T---GCC--A-C----C----AA-----C--------A--AA-------------------AAAT....... 6790
G.CU.99.Cu74             --AA...--------Y-AG--------G---G-G...-GG---------...-G-------C-----------------.............------GCA--A-C----C---GT------C--------A--AT-------------------AACT....... 7220
G.ES.99.X138             ---A...---T----T---------------G-G...-GA---------...-G---G---C-----------------.............------GCA--A-C----C---AGG-----C--------A--AT----G-T-------------R......... 7036
G.GH.03.03GH175G         ---A...--------T-A-------------G--...-GA---------...-G---G---C-----------------.............------GAA--G----------GAC----CA--------A--AA-------------------GAACAGT.... 7647
G.KE.93.HH8793_12_1      ---A...--------TG-G-K----------G-G...-GAC--------...-G---G--GC-----------T-----.............------GCAA-A-C----C----AA-----C--------A--A--------------------AAATGCT.... 6971
G.NG.01.01NGPL0669       ---A...--------T--G------------G-G...-GA---------...-G-------C-----------------.............-----TGCA--G-C----C----AG-------------------A-----GTG----------C......... 6797
G.PT.x.PT2695            ---A...---T----T--G------------G-G...-GA---------...-G---G---C-----------------.............------GAA--A-C----C----ATG-------------A--A--------------------AGAAAAT.... 7612
G.SE.93.SE6165           ---A...--------T-AG------------G-G...-GA---------...-GG--G---C-----------------.............------GCA--A-C--GA---AAC----GC--------A--AA-------------------GAATAATAACA 7017

H.BE.93.VI991            ---A...--A-A-----A-------------G-G...------------...-G-------C-----------------.............--------A--A-C--C-C----AT---------------A--AA---G---TTT--CC-G............ 7003
H.BE.93.VI997            T--A...--G---------------------G--...--T---------...-G-------C-------------T---.............-----T-AA--AGC----C----GT---------------AT-AA-CC-G--TTAT--C-AA-A-......... 6915
H.CF.90.056              T--A...--G---------------------G--...------------...-G-------CG----------------.............------CAA--A-C----TG--GT---------------A--AA---G----TT--C-AG---......... 6920

J.CD.97.J_97DC_KTB147    ---A...--------T-A-----G-------G-G...-GA---------...-G-------C-------A---------.............------GCA--A-C--C-C---CA-----------------T-A-----G---C-T-----G--AAATACC.... 6801
J.SE.93.SE7887           ---A...A-A-----T-A-------------G-G...-GA--------A...-GGAC----C-------C---------.............-----TGCA--A-C---CG--CA-----------------AT-A-----G---C-T-----A--AAACAGGGGGA 6906

K.CD.97.EQTB11C          ---A...A-------T--G--------G--A---...------------...-----G--GC-------A---------.............----C-GCA--A-C----CC--AG---------------AA--A----G--------------AATGAC.... 6784
K.CM.96.MP535            ---A...--------T--------G--G------...------------...-----G-----------A---------.............-----TGCA--AGT----AC--AGC---------------AA--A----G---------------......... 6764
```

```
                                    V4 loop end
B.FR.83.HXB2                CACC...CTCCCATGCAGAATAAAACAAATTATA...AACATGTGGCAG...AAAGTAGGAAAAGCAATGTATGCCCCT..............CCCATCAGTGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGT.......... 7601
Env                         __T__.__L__P__C__R__I__K__Q__I__I__.__N__M__W__Q__.__K__V__G__K__A__M__Y__A__P__.__.__.__.__.__P__I__S__G__Q__I__R__C__S__S__N__I__T__G__L__L__L__T__R__D__G__G__.__.__.__
01_AE.CF.90.90CF11697       ---A...-T-------G----G--G--G--...-------------...G-------------------T--...............----------A-GC-G---AT--GT--------------AA-AA------------------.......... 7559
01_AE.CN.05.FJ051           --TA...-T-------G----G-----A--...-------------...G--------C-------------T--...............-------------AGC---AT--GT--------------AA-A----G------------------.......... 7583
01_AE.CN.06.FJ054           --TA...-T-------G----G---------...-------------...GG---G---C-------------T---..AATGTATGCTCCT------.-G--AT----AT--GT------C--------AA-A----G---------------ACT....... 7610
01_AE.HK.x.HK001            --TA...-T-------G--------------...-------------...G--AC------------------T--...............--------G--AGC--CAG--GT---------------AA-A----G---------------AATAAT.... 6959
01_AE.TH.01.01TH_R2184      --TA...-T------AG--------------...-------------...GG-------C-------------T--...............--------G--AGC--C-AT--GT---------------AA-A----G---------------ACT....... 6804
01_AE.TH.02.OUR769I         --T-...A-T-------G----G-----G--...-------------...GG-------C-------------T--...............-------------AG---GAT--GT---------------AA-A-------------------------.......... 6788
01_AE.TH.90.CM240           ---A...-T------AG---G----------...-------------...GG--C----C-------------T--...............-------------AG----AT--GT---------------AA-A----G-----------------.......... 7154
01_AE.US.00.00US_MSC1164    --TA...-T------A-T---G---------...-------------...G---C----C-------------T--...............-------------AGT---AT--GT---------------AA-A----G---------------GCT....... 6797
02_AG.CM.02.02CM_4082STN    A--T...----A-----------G-----G--...-------------...-G------C--G-----------..............------CAG--AT---A--G---GA-----C--------A--A-----------------A--G.......... 6779
02_AG.EC.x.ECU41            A--T...----A---T-------G---C------...--T----------...-G------C-G------------..............------C-A--AT---A--G---AAC-----C-----T--AT-A-----------------A--G.......... 6666
02_AG.GH.03.GHNJ196         A--T...----A----------GG----------...--T----------...----------------------..............------CCA--GT---A--G---GAG-----C--------A--A-----------------A--GAAA....... 7645
02_AG.NG.01.PL0710          AGAG...----A-----------G-----G--...--T----------...---------C------A-------..............------CAA--AT---A--G---GAT-----C--------A--A-----------------A--G.......... 6752
02_AG.NG.x.IBNG             A--T...----A-----------G-----G--...--T----------...---------C----------------..............------CAA--AT---A-----GAT-----C--------A--A-----------------A--G.......... 7110
02_AG.SE.94.SE7812          A--T...----A-----------G-----G-G...CGT----------...---------C----------------..............------CCA--G---A--G---GA-----C--------A--A-----------------A--G.......... 6988
02_AG.SN.98.MP1211          A--T...---AA----A-----------------...--T----------...---------C------A--------..............------GAA--GT---A--G---GAT-----C--------AA-A-----------------A--G.......... 6803
03_AB.RU.97.KAL153_2        ---A...-----------------------...-------------...G-------------G-----------..............------GCA--------------------A------------------------.......... 6778
04_cpx.CY.94.CY032          ---A...----A------C-------T--G--...-GG----------...G----G--GC---------C---T-C..............-----TGCA--GAGC----AC--CAGC---G----------AA-AA----G-----------------.......... 6977
05_DF.BE.x.VI1310           --TT...--------T-A-----------AG--...-GA----------...GG------C-------------G--..............-----TGCA---A-C---GCC---AAC----C----------A-A-----GG-------------AAATGGTAAT. 7007
06_cpx.AU.96.BFP90          ---A...--------T-AG-----------G--...-GA----------...-G---G---C----------------..............-----TGCA---A-C----C----GT------C--------AA-AA----------------G............ 7649
08_BC.CN.97.97CNGX_6F       ---A...A-A--------------G----------...-------------...G-G-----CG----------------..............-----TGAA---A-C--A-C----AA---------C-----A--A-----GGT-C-T-----A--AAGA....... 6780
09_cpx.GH.96.96GH2911       ---T...A---AC----------G------G-G...-GA----------...-G------C-----------------..............------CCA---G----A--G---AG------C--------A--AA-----------------G.......... 6779
10_CD.TZ.96.96TZ_BF061      ---A...----------------G------------...--T----------...GG--------------A---------..............-----TGCA----TG--C-AT-------------------A--AG-G--G--------------GCG....... 6953
11_cpx.GR.x.GR17            A--T...-----------------G------------...--T----------...-G------C-----G----------..............------CAA---G--C-A--G---GAC-----C--------A--A---------------------GGAGGGG.... 6868
12_BF.AR.99.ARMA159         T-TT...--------T-A-----------G--...---C----------...G---G---CG------------G--..............-----TGAA---A-C----CC---AAC--------------T--A-----G---------------CTT....... 7566
13_cpx.CM.96.1849           A--T...-----------------G---T--G--...-------------...-G------C-----------------..............------CAA---AT---A--G---GA------C--------A--A------T--------------GTG....... 7023
14_BG.DE.01.9196_01         ---A...-----------A------------G--...-------------...G------------------------..............--------C-------------C---------------------------T---------------.......... 7109
15_01B.TH.99.99TH_MU2079    ---G...-----T---------G--------G--...-------------...G--------G----------------..............-------AA--------------------------T----------------------AAAAATGAG. 6970
16_A2D.KR.97.97KR004        G--T...------------------------G--...-------------...-G------CG----------------..............-----T-AC---AC---A-AG---A---C--C--------AA--A----------------GAACAGT....... 6970
17_BF.PE.02.PE02_PCR0155    ---T...----A---TC-------------G--...-------------...G---G---CG------------G--..............-----TGCA---A-C----CC---AGC--------------T--T-----G---------------GGT....... 6770
18_cpx.CU.99.CU76           T--A...--G-A-----------------G--...-GA----------...-G-AC----C-----------------..............----G---A---A-T----C---------------------------AA----G-----T--CCT-AT-.......... 6909
19_cpx.CU.99.CU7            A--T...----AG-------------G--G--...--T----------...---------C-----------------..............------CCA---GT---ACAG---AA---T------------A--A------------------G.......... 6679
20_BG.CU.99.Cu103           ---A...------------G--C-----G---G-G...-GA----------...-G------C-----------------..............------GCA--GAGC----C----AT------C--------A--AT--A---------------GAAT....... 7035
21_A2D.KE.91.KNH1254        ---A...----A--------------------...--T--------A...G--------G----------------..............------GCA--GAT----AC---A--------------------T-AA----G-------------AATACCAGTA 6818
22_01A1.CM.01.01CM_0001BBY  A-TT...--T--G-----------G------------...--T----------...-G--C----C-----------------..............------CCA---AC---A-CG---RAG-----C--------A--A-----G-----G--------G.......... 6768
23_BG.CU.03.CB118           ---A...------------G--------G---G--...-GA----------...-G------C-----------------..............------GCA---AGC----C----GT------C--------A--AA--A---------------G.......... 7032
24_BG.CU.03.CB378           --TA...------------G--------G---G-G...-GA----------...-G------C-----------------..............------GAA---AGC---------GA------C-----------AA--A---------------GACT....... 7044
24_BG.CU.03.CB471           ---A...------------G--------G---G-G...-GA----------...-G------C-----------------..............------GCA---AG-----AT---GTC-----C--------A--AA--A---------------GAATAGG.... 7047
25_cpx.CM.01.101BA          A--T...-----G-----------------G--...-------------...-G------C-----------------..............------CAA---GT---A-AC---A-------C--------AT--A--C-G-------------TA.......... 6851
27_cpx.CD.97.97CDKTB49      T--A...--G--------------G------G--...-------------...-G-AC----CG-----------T---..............-------------GT-----AT---G----------------A--A-----G---C-----------AATAAT.... 6805
28_BF.BR.99.BREPM12609      ---A...----T--------------------...----G---------...G-------------C-------------..............------GAA--------------------------------------------GTCTAAT.... 6999
29_BF.BR.02.BREPM119        ---A...-------------------G--------------...-------------...G------------------------..............-------AA-------------C---------------------A------------------GAAC....... 6750
31_BC.BR.02.110PA           GGTA...A--------G---------------------...--T----------...G-G------CG--C----------------..............---G-T-AA---AT---A-C---AG---------C--------AT--C-G-----G---A-A--CGTG....... 7097
32_06A1.EE.01.EE0369        ---A...--------T-A-----G------G--...-GA--------A...-------C-CG---T------------..............-----TGCA---A-C----C----A-------C--------A--A-----------------GTC-.......... 7227
33_01B.MY.05.05MYKL007_1    --TA...-----------------G---------...-------------...GG--C----C-------------T---..............-------AG---AT-----AT---GT---------------AA-A---------------------GCTAAT.... 6983
34_01B.TH.99.OUR2478P       --TA...--T-------G-----G------------...-------------...GG-------C-------------T---..............-------------AC-----AT---AT---------------AA-A--G--G-----------------.......... 6767
35_AD.AF.05.05AF095         A--T...-----------------G-----------...--T----------...-G-AC----C------A----------..............------C-A---G----C--G---GT------C--------A--AT-G--G---------------ACT....... 6767
36_cpx.CM.00.00CMNYU830     A-TT...-----G--T-A------G-----------...--G----------...-G------C-----------------..............------GCA---AGG--A--G---GAG---------------A--A-----G----------------G.......... 6798
37_cpx.CM.00.00CMNYU926     A--T...-----T-----------G------G-C...--T----------...-G------C-----------------..............------CCA---GT---A--G---GAG--C--C--------A--AA-G-----------------G.......... 6783
38_BF1.UY.05.UY05_4752      --TT...--------T-A-----------AG--...-----------A...G---G---CG------------G--..............-----TGCA---A-C---GTC---AAC----C---------T--A--C-----------------CATGGTAAT. 7012
39_BF.BR.03.03BRRJ103       T--T...--------TC-----------------G--...-------------...G--------CG-----------AAA-..............---G-TGAA---AGC----CC---AGC--------------T--A-----G---------------AGGGAAAAAGA 7071
40_BF.BR.04.04BRRJ115       --TT...--------T-A---------G--G---...-------------...G--------CG-----------TG-C..............---G-GGCA---A-C----TC---AAC--------------A--A----------------------AATAAGACT. 7158
42_BF.LU.03.luBF_05_03      T--T...--------TC------------------...-------------...G---G---CG-----------TG--..............-----TGCA---A-C----CC-------------------T--A-----G---------------CTGGCTAAT. 7102
43_02G.SA.03.J11223         ---A...A-------Y--G-----------G-G...-GA----------...-G---G---C-----------------..............-----AGAA---AT----C----AG------C-----------AT------------------G.......... 7084
U.CD.83.83CD003             T--A...-----------G---------G-GG--...-GA-CA-------...GG------C-----------------..............--T---GAA--GAC-----G--C-----------------T--A-----G---------------AAT....... 7077
U.CD.90.90CD121E12          T--A...A-T-------------------G--...-GA-CA-------...GG------C-----------------..............--T----AA--GAC----TG--CAAC----G---------A--A----------------------GAT....... 6772
U.GR.99.GR303               -T-T...TC-ATG...CAG-AT--G-----------...--T----------...-G------C-----------------..............------CAA---G----A-AG---AAT----CC-----T--A-----------------------.......... 6882
U.NL.95.U_NL_95_H10986_D1   ---A...--------TTAC-----------G-G...-GA--------A...-G----A-C-----------------..............------GCA---AG----CA----AT---------------AT-A--G-----C--------G---AAGAAT.... 7507
CPZ.CD.90.ANT               --TA...GCA-AT--------T--G--G--AG-T...--TCAT---GGC...-T---TTC-----GC--T--CTTAG-C..............--T-GG--A--GA-TG--TCC---A-T--C-GC--A--T--AA-TA-G--GGA-G-TC-A................ 7030
CPZ.CM.01.SIVcpzCAM13       G--A...A-A-----------T-G-----AG-G...--TTCA---AT-...-G--------GG-G---C-----A---..............--A--ACCA-----T--A-CT---AAC--TCTA-----T-----CA----G---C----CCA--T-.......... 7182
CPZ.CM.05.SIVcpzEK505       ------...T-A-A-----------G-------G--...--TT-A---ACA...-G-----------G---T-----T--G..............--A--T-AA----CT----AC---CTC--T--C--C--T---A-AA--C-TGACTATACA................ 7081
CPZ.CM.05.SIVcpzLB7         G--A...--A--T--T--G--T-G-------G--...---C-A----TA...-G--------G-G---A-T---T--A..............--T--T--A--TA-C--CC-C---AAT--C-C------T---T--A-CC--GA--A-................... 7112
CPZ.CM.05.SIVcpzMB66        ---T...A-T--T-----------GG------G--...-GGC-A-----A...-G--------G--G---A-T-CTT--C..............-----T--A---AC---A-AC---ATC--G------C------A-AT-G--TG-GCAGC-AAAAACA.......... 7222
CPZ.CM.05.SIVcpzMT145       A-TA...A----T--T-------G------AG-G...--TTCC---AT-...CGT-----G---G---T-T---A--A..............--A--T--A--TA-C----CC--CA-T--T-----C-----TA-----C--GAG-T-C--AAGAACCGAGAAGATC 7090
CPZ.GA.88.GAB1              A-TA...--G-----------T-G--G--AG-T...-GTTCA---AT-...-GG-------G--G---A-----A--C..............--A-----A---A-C--C-CT--CAAT--T--C-----T--A--T--C-----TTC---CAC-CC-GTG....... 7623
CPZ.TZ.00.TAN1              TGTT...GCA---------C-G-GG--G--AG--...--TCAT---GGT...-T---GTC-----GGG-T--CCTT--C..............--A-GG--G---AC-G-A-A----CAC-----C--C-----A--TA-CA-G---GC---A................ 7218
CPZ.US.85.CPZUS             ---A...--TAAG--T--------G--------G--...--TCA----AT-...-G-----------G---C-T------A..............--A----AA--TGTGC-A--T---AAC---------A--G--AA--A-TC-TGAC-T-AGCATAA-CGCAGTC.... 7603
CPZ.CM.x.SIVcpzDP943        A-TA...A-AAAC--T--------GG------G--...---TCC---AT-...-G---G--C--G-G---T--C-----A..............--A--T--A---ACT--A--T--CA--------------T-----C--GC-CGAGTT---CA-A--ACTCCCGAGAA 7156
CPZ.CM.x.SIVcpzMB897        T--ACACA-T--G----------G--------G--...--TCA------A...GG-----------G---C-T---A---..............--T--T--A---AC----CA----AAT--C-CA--C-----C--A--C------------------.......... 7115
N.CM.02.DJO0131             T--T...--G--C--T--------G--------G--...--TT-----ACA...-G-----------G---T-----A--G..............--A--TC-G---A-TC---CC---AAC--C-C------T--A---A-TC--GA-CATAG----.............. 7031
N.CM.04.04CM_1015_04        T-TT...--G--C--T-----------------G--...--TT-----ACA...-G-----------G---T-----A--A..............--A--TC-G---A-T----CC---AAC--C-G-----T--A--AA-TC--GA-...................... 7006
N.CM.04.04CM_1131_03        T-TT...--G--C--T-----------------G--...--TT-A---ACA...-G-----------G---T-----A--A..............--A--TC-G---A-T----CC---AAC--C--------T--A---A-TC--GA-TAT.................. 7065
N.CM.95.YBF30               T--T...--G--T--T--------G---G---G--...--TT-----ACA...-GG----------G---T-----A--A..............--A--TC-G---GTTC---AC---A-C--C--------T--A---G-TC--GA-TATAG------G.......... 7149
N.CM.97.YBF106              T-TT...--G--C--T-----------------G--...R-TT-----ACA...-G-----K------G---T-----A--A..............--A--TC-G---A-TC----C---ARC--C-G---C--T--A---A-TC--GA-CATAG----.............. 7107
O.BE.87.ANT70               ---T...--A--T----A-C-G-G---GG-GG--...-GGTCA---ATA...-GG-G-CAGTCG-G-C-C-----A---..............------AA--TA-TC-A-C----ATG-----C--A--T--AA--A-CC--CA--TG---AACACATGG....... 7642
O.CM.91.MVP5180             T--T...A-A--T----AGT---G---GC-AG--...-GATCA---AT-...--G-G--AGTCGAG---C-----A---..............-----CCC--CA-CT-A-C----CAT--C--C--A--T--AA--A-TC--CAGTT----CAACCA.......... 7656
O.CM.96.96CMABB637          GCGA...A-A--T----AG--------GG-AG--...-GGTCA---ATA...--G-G--AGTCG-G-C-C-----A--A..............-----T-AA--TT-CC-A-CC--CAT------C--A--T--AA--A-TC--CA---G--CA--CCA.......... 7144
O.CM.98.98CMA104            ---A...A-A--T----AGT-G-G--GG-AG-G...-GGTCA---ATA...C-G-GG--GTC--G-C-C---------..............-------AA--TACCC-A-C----ATG-----C--A--T---A-A-TC--GA--TG--AC-GTCATGGAACGGGA 7111
O.CM.99.99CMU4122           AGAG...A-A--T----AG--------GG-AG--...-GGTCA---ATA...-GG-GR-AGTCG-G-C-C-----A---..............-------AA--T-CTC-A-A----GAG-----C--A--T--AA--A-TC--CA--A-----AACCATGGAAC.... 7109
O.FR.92.VAU                 GG-A...A-A--T-----GT-G-G----G-AG--...-GGGAC---AT-...-GG-G----TCG-G-C-T-----A---..............------CCA---A-CC-AGT---CAGG-----C--A--T--AA--A-TC--CA-TTG--CACGCCATGGAAT.... 7205
O.SN.99.SEMP1299            TCGG...A-ATAT-----G---------GG-GG--...-GGTCA---ATA...C-G-G---GTC--G-C-C-----A---..............----GG-AA--TA-CC-A-C---CAGC---CTC--A--T--AA--A-TC--CA-TTG---ATGCCATGGAAC.... 7692
O.US.99.99USTWLA            A--A...A-ATAT---C-GT-G-G---GG-AG--...-GGTCA---AT-...-GG-G---GTC--GCC-C-----A---..............--T----AA--TAGTC-A-CC--CAGT-----C--A--T--AT-AA-CC------------CACTAATAAAMAGA. 7119
```

```
                              V5 loop start              V5 loop end                                                  Rev Responsive Element (RRE) region
B.FR.83.HXB2 ....................AATAGCAACAATGAGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTG......... 7742
Env .__.__.__.__.__.__.__N__S__N__N__E__S__E__I__F__R__P__G__G__G__D__M__R__D__N__W__R__S__E__L__Y__K__Y__K__V__V__K__I__E__P__L__G__V__A__P__T__K__A__K__R__R__V__V__.__.__._
A1.GE.99.99GEMZ011 ....................GGA-CT--T-GCAGCAAT----C----------AT-------A--------------------------C-------G------------C-----------A-------------T------------------------......... 6917
A1.KE.00.KER2008 .......................-AT-GT-C--TCAAT--A-CT-----------------------------------------------------G------------------------C------------------------G-------------......... 6948
A1.KE.00.KNH1144 ....................G-G-AG----G-ACAAAT--A-----------------------------------C--------------------G------------------------C------------------G-----G-------------......... 6974
A1.KE.00.KSM4024 .......................-AT--TGT-ACAAAT--A-C------------------------------------------------------G-----G------------------C--------------------------------------......... 6871
A1.KE.00.MSA4069 .......................GATG-T-----AA-T--A-C---T--------------------------T-----------------------G------------------------C------------------G------A------------......... 6950
A1.KE.00.NKU3005 .......................TCT----G-ACAGAA--A-C------------------------------------------------------G-----------G--------G--------G-------------G-------------------......... 6953
A1.RU.00.RU00051 ....................---G-A-G---C-G-A------C----------AT------------------------------------------G------------------------A-------------T----G-------------------......... 7016
A1.RW.93.93RW_024 ....................---GAT-G--G-CCAGAT--A--------------------------------------------------------G------------C----------TC------------------------GA------------......... 6944
A1.SE.95.SE8891 ....................---GAT--------AAGT--A--------G--------------C--------------------------------G-----G------------------C-------------------------A------------......... 6907
A1.SE.95.UGSE8131 ....................-C--A---T-CAA-TAAT--A-C---------------------------------------------------C--G---------------C--------C-----------------------C--------------......... 7120
A1.TZ.01.A173 ....................-GG-A---T-G-CTAAAT--A-C------------------------------------------------------G------------------------C-------------T----G-------------------......... 6951
A1.UA.01.01UADN139 .......................-CT--T-GCAGCA-T----C----------AC----------------------------------C-------G------------G-----------A-------------T------------------------......... 6935
A1.UG.92.92UG037 ....................GT--AT-G--G---CAGT--A-C------------------------------T-----------------------G-----G------------------C------------------------G-------------......... 7752
A1.UG.99.99UGA07072 ....................GG--AT------ACAAAT--A-C------------------------------------------------------G------------------------C-------------T----G-------------------......... 6948
A2.CD.97.97CDKTB48 .......................-AG----G-ATTAAT----C------------------------------------------------------G------------------------C-----A----------AG------G-------------......... 7052
A2.CY.94.94CY017_41 .......................-A---TGGGACTAAT----C----------------------------------------------------------------------C--------C-------------T----G------A------------......... 7114
A.SN.01.DDI579 ....................G---AT-GT-G-ACAAAT----CA------------------A------A----------------------------------------------------C----------------------------G---------......... 6946
A.SN.01.DDJ369 ....................--CCTT--T-GCTCA-TT----CG------------------A------A----------------------------------------------------C--------------------------------------......... 6964
A.SN.96.DDJ360 ....................G---AT--T-G-ACAGAT----AG------------------------------------------------------------------------------C------------------G-------------------......... 6928
A.CD.02.02CD_KTB035 ....................GG---G--T--CAGTAA-----C------------G----------------------------------------GG------------C-------------------------------------------G------......... 7741
A.CD.97.97CD_KCC2 ....................---GAA-CT-GG---AAT----C----------AT---------G-----A--------------------------G------------------------C--------------------------------------GAG...... 7741
A.CD.97.97CD_KTB13 ....................G----G----G-AGT-AT----C----------AT-------A------A-------------------C-G-----G------------------------C-------------T----------G---G---------......... 7018
B.AR.04.04AR151516 ....................--CGAGTC---AC-AAAT--A-C---------------------------------------------------C---------------------------------------------G--------------------......... 6948
B.AU.87.MBC925 ....................CC-GA---T--A--CA-T------------------------------------------------------------------------------------------A----------T---------------------......... 7761
B.BO.99.BOL0122 ....................G---AT----GAAGCA------C---T---------------A-------A----------------------------------------------------------------------G-----G-------------......... 6940
B.BR.03.BREPM2012 ....................-T--A-C-G-CCA-CAAT--A-C---------A-C------------------------------------------------------G--------G----------------------------G-------------......... 7184
B.CA.97.CANB3FULL ....................-T--A-----G-ACCAAT----C-----A------------------------T--------------------C---------------C--G-----------------------------------------------......... 7015
B.CN.05.05CNHB_hp3 ....................-T--AT--G-C---CA------C---------------------------A--------------------------------------G-------------------------------G-------------------......... 7755
B.CO.01.PCM001 ....................--CGAT-CT--CA-CAAT----C-------------------A-C--------T---------------------------------------------------------------------------------------......... 6925
B.GB.83.CAM1 ....................-GAG--G-A------A------C----------------------------------------------------------------------------------------------------------------------......... 7739
B.GB.86.GB8 ....................--C-AT-C---C---A------C----------------------------------------------------------------------------------------------------------------------......... 7730
B.GE.03.03GEMZ010 ....................-----T-GTGC---TCAG-----------------G------A----A-A-------------------C-----------------------C----G------------------------------------------......... 6916
B.IT.05.SG1 .......................-AT--T-----CA-T------------A-C--G--------------------------------------C---------------------------------A----------T---------------------......... 7427
B.JP.05.DR6538 ....................GGA-AT----CG--CA-G---G----------------------------------------------------C---------------------------------A------------------G-------------......... 7778
B.KR.05.05CSR3 CTACCGGC............-GC-ATG-G-CAA-AGGG----------------C--------------------------------------------------------G-----------------------------------G-------------......... 7805
B.NL.00.671_00T36 ....................GG-GAG-----AGCA-----------T-----------------------A-----------------------------------------------G-----G------------------------------------......... 7303
B.RU.04.04RU128005 ....................-G--A--G--CCA-CAA-----C-----------C------------A-----------------------------G-------------------------------------------T-------------------......... 7214
B.TH.00.00TH_C3198 GGAAC...............--G-C--C------TA-A----C---T------------------------------------------C-------------------G---------------------------------------------------......... 6940
B.UA.01.01UAKV167 ....................----AT-GT-G-A-CAA-----CT-----------------GA------A------------------------------------------------G------------------------------------------......... 7006
B.US.04.ES10_53 ....................C---A--CG---ACTA-T------------------------------------------------------------------------------------------A------------------G-------------......... 7720
B.US.99.PRB959_03 ....................--C-A--G-G--ACTC-T----C------------------------------T---------------------------------------G--------------A--------------------------AC----......... 6940
C.AR.01.ARG4006 ....................---G---CA---A-CA-A-----A-----G--------------------------------------------------Y-----G---G----CA-C---C-----A----------TR-T-----A--G---------......... 6930
C.BR.04.04BR013 ....................--C-C--CG---A-CA-A--A-CA-----G---------------------------------------C----------------G---G-----A-----------A----------TG-------A--G---------......... 7233
C.BW.00.00BW07621 ....................G-AG-TG-A---ACAA-A----CA-----G----C---------------------------------------------------G---G-----A-G---------A----------T-GC-----A--G---------......... 7074
C.CN.98.YNRL9840 AT..................-CAGAG-CA---A-TA-A----CA--------------------------------------------------------------G---G-----A-G-----G--------------TCCA-----A--G---------......... 6943
C.ET.02.02ET_288 ....................--C-A--CA--CA-CA-A----CA-----------------------------------------------------G--------G---G-----A-G--------------------TGG------A--GGC-------......... 6950
C.GE.03.03GEMZ033 ....................GG---T-GT---TGCA-A----CA--------G-----------C-----------------------------------C-----G---G-----A-G-----------------------------A--G---------......... 6949
C.IL.99.99ET7 ....................-CGGA---TG-AACAATA----CA-----------------------------T--------------------C---------A-----G-----A-G--------------T--A--TGC------A--G---------......... 6908
C.IN.99.01IN565_10 ....................-CA-AGG-T-----CA-A-----A--------------------------------------------------------------G---G-----A-G-----G--------------TG-A-----A--G---------......... 7122
C.KE.00.KER2010 ....................--C--T-CA---A-AA-T----CA-----G------------A------A------C-----------------C-----------G--GG-----A-G-----G---A----------TG-------A--G---------......... 6931
C.MM.99.mIDU101_3 .............................----GTA-AC---CA------------------A------A---T--------------------------------G---G-----C-G---------A----------TGGA-----A--G---------......... 7104
C.MW.93.93MW_965 .......................GAA--T---AG-A-T----CA-------------------------A---------------------------------G--G---G-----A-G--------------------TGGT-----A--G---------......... 6917
C.SN.90.90SE_364 ....................-CACTT--T-----TA-A----AA------------------------------------------------------------------G-----A-G---------A----------T--------A--G---------......... 6887
C.SO.89.89SM_145 AAAGAACT............G-AGAGC-A---A-CA-A----GA--------------------------A--------------------------------G--G---G-----A-----------A----------TGG------A--G---------......... 6968
C.TZ.02.CO178 ....................G-A-CA-G---CACCA-A----CA---------TT---------C--------T---------C----------------------G---G-----A-G-----G---A----------TGG------A--G---------......... 6946
C.UY.01.TRA3011 ....................G-CG-G--TGGCACAA-A----CA--------------------------------------------------C-----------G---G-----A-----------A----------T-CT-----A--G---------......... 6922
C.YE.02.02YE511 ....................GGA-AAG----C-----A----CA-----------------------------T--------------------------------G---G-----A-----------A----------TGCA--------G---------......... 6934
C.ZA.04.04ZASK164B1 ....................--CGAT-CT---A-CA-A-----A----------C----------------------------------C----------------G---G-----A-G--GC----------------AG-------A--GG--------......... 7192
C.ZA.05.05ZAPSK240B1 ....................-TC-CG--T--CACAGAG-----A--T---------------A------A---T--------------------------C--G--G---G-----A-------------------------------A--G---------......... 7118
C.ZM.02.02ZM108 AG..................-CAGATGCA---A--A-A----CA----------AT-------A----------------------------------------G--G---G-----C-G---C----C--G--------T--------A--G---------......... 7779
D.CD.83.ELI ....................-TA-AT--T-G-ACTAA-----C---T--------------------------------------------------------G------C-----------C------------------G-------------------......... 7278
D.CM.01.01CM_0009BBY CTGGTAAT............---TCTGGA---TCTA-A-----------G------------A------A----------------------------------A------G----------C------------------G-------------------......... 6967
D.KE.01.NKU3006 ....................----T---T-G-ACTAAT----C------G-----------G--------A-----------------G-----C-----------------G--C------C----TC-----------G--------------------......... 6985
D.KR.04.04KBH8 .......................-AT--T---AC-AAT----CT------------------A------AA------------------C--------------------C--G--------A-----C-------T------------------------......... 7655
D.TD.99.MN011 .......................-AA--TGC-TCTAAT----C--------------------------------------------------------------------G----------C------------------------G-------------......... 6932
D.TZ.01.A280 ....................GTA-AT--T-G-AGAGAG--------------------------------A-----------------------C----------------G---A------A------------------G-------------------......... 6930
D.UG.99.99UGK09259 ....................----AT--T-G-C--AAT----C------------------------A--------------------------C------------------C--------C----TC------------G-------------------......... 6943
D.YE.01.01YE386 ..........................--TG-AA-TAAT----C------G--------------C-----------------------------C----------------G----------C----TC------------G-------------------......... 6898
D.YE.02.02YE516 ....................--C-AT-CT--CAGTA-T----C------------------------------T-----------------G-------------------G----------C-------------------------A------------......... 6922
D.ZA.90.R1 ....................-----T-GTTC-C-CAAT----C---------------------------------------------------------C----------G----------A-----T------------G-------------------......... 7096
F1.AR.02.ARE933 GCACTGATGGT.........GG--CTGGG---TGCA-T----C------------G------A------AA--T--------------------------C---------G-----------C-------------------A----GA--CA-------A......... 7012
F1.BE.93.VI850 .......................-ATG-G-G-A-CATT----C-----------AG------A------AA------------------C-------G------------G-------------------------------------A--CA--------......... 7019
F1.BR.01.01BR125 ....................-T---T-CTG--C--C-T-----------------G------A------AA-----C---------------------------------G-------G---C-----------------G-------A--CA--------......... 7169
F1.ES.x.P1146 ....................GGGC-T--TG--ACTGAA--A--A-----------G------A------AA------------------C--------------------G-----A-----C-----A-----------G------G---CA--------......... 7049
F1.FI.93.FIN9363 ....................-G--AT--T-G---T--T----C------------G-------------AA------------------C--------------------G-----------C------------------G-C-C--A--CC--------......... 7038
F2.CM.02.02CM_0016BBY ....................-G--AGGCA---A-CA-A--C---C--------AT---G-----G-----A--------------------------------G------C-----C-----C-----A--------T---G------A--CA--------......... 6901
F2.CM.95.MP257 .......................GAGGGA-------------C-C-------------G-----------A------------------C-------------G------------------C-------------------------A--CA--------......... 6935
F2.CM.97.CM53657 ....................--A-AT--T-CCA-TAAT--C-C-C-------------G-----------A------------------C-G-----------G------------------C-------------------------A--CA--------AAG...... 6910
G.BE.96.DRCBL ....................G---A---T-GCACTAGT-------------------------------AAA-------------------------G------AC---------CA--T--C-----A------------G-----G-------------......... 7681
G.CM.01.01CM_4049HAN ....................-C---GG-T---ACC--A--------------------------A-----A-------------A------------T------AC---------CA--T---------------------G-----G-------------......... 6931
G.CU.99.Cu74 ....................----AT-CT---AGAA-T---G------------C---------C-------------------------------GG------------------A-----C------------------R-----G-------------......... 7361
G.ES.99.X138 ....................-C-GAT-CA----G-A-T-----T----------C---------C----A---------------------------G------A-----------A-----C------------------G-----G-------------......... 7177
G.GH.03.03GH175G ....................--C-AT-CT---A-TA-A----C------------------------------------------------------G--C---AC---------CA--T--C-----A------------G-----G----C--------......... 7788
G.KE.93.HH8793_12_1 ....................--C-ATGCA-G----A-T----C-----------C-------A------A--------------A------------G------------------A-----C------------------------G-------------......... 7112
G.NG.01.01NGPL0669 .......................-TA--T-GCACAAAT--A-C--------------------------A---------------------------G------AC---------CA--------------------------------------------......... 6935
G.PT.x.PT2695 ....................---G-T-CA------A-T-----T----------T---------C--------------------------------G------------------A-----C--------A---------G-----G-------------......... 7753
G.SE.93.SE6165 AT..................-CA-AT-CA-G----AGT-----------------------------------------------------------G------AC---------CA--T--C------------------G-----G-------------......... 7160
H.BE.93.VI991 .......................-AT--T-CAA-TAAT-TT-C-------------------------------------------------------------------------------C-G-----------A---G------G-------------......... 7141
H.BE.93.VI997 ....................--GG-A--T-----TAAT-TT-C-------------------------------------------------------------------------------C-G---------------G------G-------------......... 7056
H.CF.90.056 ....................--CGCGTCTGCA--AAATT-T-C------------------------------------------------G------------------------------C-G---A--------------A---G-------------......... 7061
J.CD.97.J_97DC_KTB147 ....................--GGAG-GT---AGTAGT----C----------AC--------------A--------------A---G---------------A----G--GG----G----------------------G------A------------......... 6942
J.SE.93.SE7887 AT..................GGC--TG-G----GCA-T----C----------AC-------A------A----------------------------------------G-G-----G---C-------------------------A------------......... 7049
K.CD.97.EQTB11C ....................----AT-CT-GGACTGAG----C---------------------C-----A------------------C--------------------C-G---------C-----A------------G-----G-------------......... 6925
K.CM.96.MP535 .......................-AT--T-C-C-TAAT----C---------------------C------------------------C-G------------------C-G---------C-----A------------G---C-G-------------......... 6902
```

```
                              V5 loop start        V5 loop end                                                               Rev Responsive Element (RRE) region
B.FR.83.HXB2                 ....................AATAGCAACAATGAGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTG......... 7742
Env                          .__.__.__.__.__.__.__N__S__N__N__E__S__E__I__F__R__P__G__G__G__D__M__R__D__N__W__R__S__E__L__Y__K__Y__K__V__V__K__I__E__P__L__G__V__A__P__T__K__A__K__R__R__V__V__.__.__._
01_AE.CF.90.90CF11697        ..........................--TGCGACTAAT----C-------------------A------A----------------------------------------C-----------C-----A------------G----CG-------------......... 7694
01_AE.CN.05.FJ051            ....................GG--AT-CT-CA-GTAA-----C-------------------A--G-A-AA-------------A-------------------------C----A------C-----A--------------------------------......... 7724
01_AE.CN.06.FJ054            ....................----CT-GTGGGAC-AAT----C-------------------A------A----------------------------------------C--C--------C-----C--------------------------------......... 7751
01_AE.HK.x.HK001             ....................---G-GTCT--C---A------C---T--------------------A--A---------------------------------------C-----------C------------------G-------------------......... 7100
01_AE.TH.01.01TH_R2184       ....................----TT-GT--GAGTAA-----C-------------------A----A-A----------------------------------------C-----------C-----A------------G-----G-------------......... 6945
01_AE.TH.02.OUR769I          .......................CTT--T---ACTAA-----C-------------------A----A-A----------------------------------------C-----------C-----A------------G-------------------......... 6926
01_AE.TH.90.CM240            ....................GT--AT--T-CG--TAA-----C-------------------A-C--A-A----------------------------------------C-----------C-----A------------G-------------------......... 7295
01_AE.US.00.00US_MSC1164     ....................GC---T--G-----TA-A------------------------A----A-AA---------------------------------------C-----------C------------------G-------------------......... 6938
02_AG.CM.02.02CM_4082STN     ....................G-A-AT--T-G-ACAAATC---C------G-----------------------------------------------G------------------------A----T------------C-T-----A------------......... 6920
02_AG.EC.x.ECU41             .......................-AT--T-G-ACAAAT----C------G--------------------------------G---------------------------------------C----T------------C-T-----A------------......... 6804
02_AG.GH.03.GHNJ196          ....................----CT-GTGG--TAAAT----CT----------A-------A------AA--------------------------G---------A--------------C----T------------CGT----GA------------......... 7786
02_AG.NG.01.PL0710           ....................-T-GAG--TGCA--TAAT----C------G--G--------------------------------C-----------G------------------------C----T-----------A--A-----A------------......... 6893
02_AG.NG.x.IBNG              .......................-AT--T-G-ACAAAT--A-C------G-----------------------------------------------G------------------------C----T------------CGT-----A------------......... 7248
02_AG.SE.94.SE7812           ....................---GAT--T---A-TA-A----C------G------------------------------------------------------------------------C----T---------T--C-T-----A------------......... 7129
02_AG.SN.98.MP1211           ....................G---AT-CT----GTGAT----CT-----G--------------------------C------------C-------G------------------------C----T---------T--C-T-----A------------......... 6944
03_AB.RU.97.KAL153_2         ....................--CCAG-G-----TTA-------T--T---A----------------------------------------------------------T-------------------------------G-------------------......... 6919
04_cpx.CY.94.CY032           ..........................-CT---A-TA-T----C---------------------C------------------------C----------------------G---------A---------------AT-------G----C--------......... 7112
05_DF.BE.x.VI1310            ....................G-----TC---C--CA-A-----A--------------------------A------------------C--------------------G-G---C-----C-----A------------G------A---CA-------......... 7148
06_cpx.AU.96.BFP90           ....................----ATG-G----TCAGT--A-C----------------------------------------------C-G-----G------------------A-----C-----A-----------TG-----G---------A---......... 7790
08_BC.CN.97.97CNGX_6F        ....................-CAGAGCCA---A-TA-A-----A-------------------------AA-------------A---------------------G---G-----A-G-----G--------------TGCA-----A--G---------......... 6921
09_cpx.GH.96.96GH2911        ..........................-GTG----CAAT----C----------AC-------A------A---------------------------G--------G---------------C-------G----------G-----G-------------......... 6914
10_CD.TZ.96.96TZ_BF061       ....................----AT--T-G-C--AAT----C---------------------------------------------------C---------------------------A----TC------------C-------------------......... 7094
11_cpx.GR.x.GR17             ....................---GAT-C--T--G-AAA----C----------AC--------------AA--------------------G-----G------------G-G---A-----C-------------T-----------A------------......... 7009
12_BF.AR.99.ARMA159          ....................---GAG-CT---C--A-T--A-C------------G-------A------AA-----------------------------------G---G-----------C------------------G------A---CA-------......... 7707
13_cpx.CM.96.1849            ....................GG--AT--T-GGACAAAT----CA------------------A------AA--------------------------G----------------------------------------G--GA--G-G-------------......... 7164
14_BG.DE.01.9196_01          .......................-AT------A--AG----G----------G---------A------A-----------------------------------------------------------------------G-------------------......... 7247
15_01B.TH.99.99TH_MU2079     ....................----C--CG---ACCA------------------C-------------------------------------------------A--------------------------------------------------------......... 7111
16_A2D.KR.97.97KR004         ....................GGGG----TGCAACCAAT----C------------------------------------------------------G---------------C-A------C-------------T----G-----GA------------......... 7111
17_BF.PE.02.PE02_PCR0155     ....................G--G-T-GTG-CC--A-A----C------------G-------A------AA-----------------------------------G---G------------------------------T------A------------......... 6911
18_cpx.CU.99.CU76            ....................Y--CATKNT-CA---AAT-TT-C------------G--------------A-----------------------C--------------G------------C----------------T-------G---------A---......... 7050
19_cpx.CU.99.CU7             ................................A-TAA-----C---T-------CT-------G---------------------------------G--------------------G---C--------------------------------------......... 6808
20_BG.CU.99.Cu103            ....................--C--T-CT---AG-A-T----------------C---G-----C--------------------------------G------------------A-----C------------------G-------------------......... 7176
21_A2D.KE.91.KNH1254         TTAAT...............GG-GAA--TG--A-AAAT--A-C--------------------------------------------------------------------G----------C-G---A------------------G-------------......... 6964
22_01A1.CM.01.01CM_0001BBY   ....................G---AA--T-G--CAAAT----CT--------------------------A---------------------------------------------------C-----A-------A----G------------G------......... 6909
23_BG.CU.03.CB118            .......................CCT----G-ASTAAT----C--------T--C---------C-----A--------------------------G------------------A-----C------------------------G-------------......... 7170
24_BG.CU.03.CB378            ....................----AT--A---ACTA-T----------------C---------C--------------------------------G------A------G---CA-----C-----A-----------G------G-------------......... 7185
24_BG.CU.03.CB471            ....................----CT-CT---A-TA-T---G------------C---------C--------------------------------------------------CA-----C------------------------G-------------......... 7188
25_cpx.CM.01.101BA           ....................GG-G-T-G------AAAT--A-CT----------------------------------------A------------G-------------G----------C------------------------G------------A......... 6992
27_cpx.CD.97.97CDKTB49       ....................---GAT-CT----GTA-T----C-C-----------------A-------------------------GC--------------------C-----------------A-------------------A------------......... 6946
28_BF.BR.99.BREPM12609       ....................--CGAG-C---CAT-A------C-----------C---------------------------------------------------------------G----------------------T-----G-------------......... 7140
29_BF.BR.02.BREPM119         ....................--CGCG-G--GC---A------C----------------------------------------------------------------A--------------------A--------------------------------......... 6891
31_BC.BR.02.110PA            ....................---GT--CT--CTCAA-A----CA-----G-------------------A------------------------------------G---G--G--A-----------A----------T--------A--G---------......... 7238
32_06A1.EE.01.EE0369         ....................--GGAA-GT-----AA-T-----T-C----------------A------A---------------------------G------------------A-----C----CA--------T---------G---------A---......... 7368
33_01B.MY.05.05MYKL007_1     ....................---GCA-GT------A------C---------------G---A------A---------------------G------------------C-----------C-----A------------G-------------------......... 7124
34_01B.TH.99.OUR2478P        ....................-----T-CT-C---TA-A----C-----------A-------A----A-A------------------G--------------------GC-----------C-----A------------G------A------------......... 6908
35_AD.AF.05.05AF095          ....................----A--GT-G-CT-AAT----C---------------------------A-----C--------C--------C------------------C--------C------------------G-------------------......... 6908
36_cpx.CM.00.00CMNYU830      ....................--A-AT----G-ACAAAT----CT--------A-----------------------------G--------------G-----G------C-----------C----T--G---------CGA-----------G------......... 6939
37_cpx.CM.00.00CMNYU926      ....................-T--A---A-CA--TAAT----C------------G-----------------T-----------------------G------------------------A-----A------------------G-------------......... 6924
38_BF1.UY.05.UY05_4752       ....................GGC-CA-T----A--A-A-----------------G-------A------AA---------------------------------------G-----------------A------------G-----G-------------......... 7153
39_BF.BR.03.03BRRJ103        ATACTACT............-C-GCT--T--GACTGAG-----------------G-------A------AA---------------------------------------G-----------C-------------------------A---CA-------......... 7220
40_BF.BR.04.04BRRJ115        ....................C---ATG-T-C-AG-AAT--------T---------------A------A-----------------------------------------G------G----------------------G-----G-------------......... 7299
42_BF.LU.03.luBF_05_03       ....................C-G-CTGGA---C--AAT--C-C------------G--------------A---------------------------------------G-------G---------A------------CA------------------......... 7243
43_02G.SA.03.J11223          ....................----TT--T-G-ACTAAT--C-C---------------------C----A---------------------------R------------------A-----Y-----R------------------G-------------......... 7225
U.CD.83.83CD003              ....................GGC-ATGCA-C-C-AAAT----C---T-----------------C----AA--T-----------------G-----G------------------------C------------------G------A------------......... 7218
U.CD.90.90CD121E12           ....................---GAG--TG--ACAAAA----C---T-----------------C-----A--T-----------------------G--C-----G---------------C------------------G------A------------......... 6913
U.GR.99.GR303                ....................--C--T-GT-----C-------C-------------------------------------S-----------------------------------------C-----------------GG------A------------......... 7023
U.NL.95.U_NL_95_H10986_D1    ....................---GAA-CT----GAAAT----C-----------C-------A-C----A------C------------C----C-----------------G---------C-----A----------T-------G-------------......... 7648
CPZ.CD.90.ANT                .......................-TAT-T-----AA-T-TT-AAG-GTC-----CT-C-...AGAG-AGCA---C-A------GCG--G--G-CC-GG--CC-G--G---G-G----RT--C--GTC------C--A--A-CA-GN--A--GCC--AAA-AAAA...... 7168
CPZ.CM.01.SIVcpzCAM13        ..........................--T---ACTAATA-T-CA--T-----AAT---------C----AAA---T------G------C-C--------C-----G--GCGC--A--------TCT--G------A-----A-----A--GCATACA--AGGA...... 7320
CPZ.CM.05.SIVcpzEK505        ....................--GTCTGGA-C----AAGT-C-C-A-ATAT--CAC---G-----C----CTA--CTC-----GCAG--G--G-----G-----G--G--C-GC--A-----CA-------------AGGA--A--T--A---CACACA---ACT...... 7225
CPZ.CM.05.SIVcpzLB7          ....................C-C-CA----CA--CAATATC-C----C-T--AAT------G-------CA---TC---------------------C------A---------------------------G---T----G---G-GA--G--GACT---GAT...... 7256
CPZ.CM.05.SIVcpzMB66         ....................G-C--G-TG--CA-A--A-CC--G----C-----TT--G-----A-----AA-T--------------------C----------------G---------C--------------------------A-----AC-----CAT...... 7366
CPZ.CM.05.SIVcpzMT145        AGGGAGAA............G--CAGG-TC-GA-TAA-ACATA-G-TT-T-TAAC---G--TA-C----AA--T-T-------------C----C-----C---A----TG-G--AC----C--------------T--T--AAGT-GA---TAT-CA--AGAAAAA... 7245
CPZ.GA.88.GAB1               ....................-C--A-----G--GTAA-CT--CA--T------AC-------A-C----AA--T-T------------GC-C-----G-----------TCG---A--G--T--TC---G------A--A--A----GA--GCATACA---GCAAGA... 7770
CPZ.TZ.00.TAN1               .......................-AAG----CA-TAATAGTTATAC-CCC-AATTTTCT-CT-TAG-AGAA---T------A-GTA------GCA-G---------G--GG-----C-G--C--GTC--G--T---A-GGCCA-G---A--GCCT-AAA-TAAGGCC... 7362
CPZ.US.85.CPZUS              ....................----A-G-T-G-AG-AATATA-CAG-G-TG---AC--------------C--CTTTA----A--A-------C----G-----G--G--C-GC--A-----TA------G------AGGT-----C--A--GCATACA---......... 7744
CPZ.CM.x.SIVcpzDP943         ATGGG...............-C--AT-C----A-CAAGTTA-CAC-ATAC---AC---------G----A--TTTG------GCTG------C-------C--------C-GC--A-----CA------G------A-GT--A--T--A--GCACACA---......... 7302
CPZ.CM.x.SIVcpzMB897         ..........................--A------A-T--A-CT--T------AC---------C-----AA-T--------------------C-----C--G-----------------CC-----T-------A-----------A--G--ACA----......... 7250
N.CM.02.DJO0131              .............................GGCAGTAAT-GC-CAG-ATAT--CAC-------A-C---GTCA-TCTC------CAA--G--G-----G--C---AC---C-GT--A-----CA-------------AGGT--A--T--A---C-CACA---......... 7163
N.CM.04.04CM_1015_04         ..........................GTT-G--GAAATAGT-CAG-ATAT---TC-------A-C---GTTA-TCTC------CAA--G--G--C--G--C--------T-GC--A-----CA-------------AGGT--A--T--A---C-CACA---......... 7141
N.CM.04.04CM_1131_03         ....................-G--AT-G--CAA-TGGTACAG-AG-ATAT---TC-------A-C---GTTA-TCTC------CAA--G--G--------C--------G-GT--A-----CA-------------AGGT--A--T--A---C-CACA---......... 7206
N.CM.95.YBF30                .......................CCTG---CCA--GAAACA--AG-ATAT--CTC-------A-C---GTTA-TCTC------CAA--G--G-----G--C--------T-GC--A-----CA-------------AGGT--A--T--A---C-CACA---......... 7287
N.CM.97.YBF106               ..........................G-A--C-GTAA-A---CAG-ATAT--CTC-------A-C---GTTA-TCTC------CAA---C-G-----G--C--------T-GT--A-----CA-------------AGGT--A--C--A---C-CACA---......... 7242
O.BE.87.ANT70                ....................--C----G---CA-CAAT-TA-CA--T-----AAT---G-----C----AA--T-TA-------C------G-TC--C--C----------GGG-AA----T--TA-T--G--------ACGTATTGCA--GCC---CA-AAGCACT... 7789
O.CM.91.MVP5180              ....................TGG-ATTC--CA-GTGAAA-T-CAC-T-----A-T---G----------AA--T-TA-------C-A----G--C--C--C--------C-G---AA----T--TA-T--------T--A--AATGTCA---CC-A-AA-AAACATT... 7803
O.CM.96.96CMABB637           ................................ACCAATT-C-C---T--G--A-T---G-----Y----AA--T-TA-------C-------GCC--T--C--------------AA----T--TA-T--------T--A--AATT-CA--GCC---CA-AAGCACT... 7279
O.CM.98.98CMA104             GTGCTACAAAGGACATTAATGC--CATGG--CATTAAT-CT-CA--------AAT---G----------AA--T-TA-------C-C------TC--C-----G--G----GGG-GA----T--TA-T--------T--A--AATTGCA--GCC---CA-AGGCATGGGC 7281
O.CM.99.99CMU4122            ....................--A-CA-----CACACA--CC-CA--T-----A-----G----------AA----TA-------C------G--C-----C--G-------GGG-AA----T--TA-T--G-----T--A--AATTGCA--GCC---CA-AGGCACT... 7256
O.FR.92.VAU                  ....................--A-CAC-TCC-A-CAG-ACC-C-C-T-----A-----G----------AA--T-TA-------C-C----G-TC----------------G-G-AA----T--TA-T--------A--A--AATTGCA--GCC-ACTA-AGGAACT... 7352
O.SN.99.SEMP1299             ....................-GC-C------CAGCAA--CC-CA--T-----AAC---G----------AA--T-TA-------C------G-TC-----C-----------GG-AA----T--TA-T--G-----T--A--AATTGCA--GCC---CA-AGGCACT... 7839
O.US.99.99USTWLA             ....................-RCTC--G---CCCCAA-ATC-CA--T-----AAC--------------AR--T-TA-------CCC----G--C--T--C----------G-G-AA--T-T--TA-T--------T--T---ATT-GT---CC---TA-AGGCACT... 7266
```

```
                           Env gp120 end Env gp41 start
B.FR.83.HXB2             .........CAGAGAGAAAAAAGAGCAGTGGGA...ATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCA...GGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGA.........CAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTT 7888
Env                      _.__.__.__Q__R__E__K__R__A__V__G__.__I__G__A__L__F__L__G__F__L__G__A__A__.__G__S__T__M__G__A__A__S__M__T__L__T__V__Q__A__R__.__.__.__Q__L__L__S__G__I__V__Q__Q__Q__N__N__L
A1.GE.99.99GEMZ011       .........G-----------------A-T---...C-G------GCC--------------A---------...-----T-----------G--G---C-----------------------.........--------------C--------A-----A-G------ 7063
A1.KE.00.KER2008         .........GGA-----------------T---...---------G-T--------------A--------T...--C--------------G--G-----A---------------------.........------C-------C--------A-----A-------- 7094
A1.KE.00.KNH1144         .........GGA-----------------T---...---------G-T--------------A---------...-----------------G--G-----A--------------------G.........--------------C--------A-------G------ 7120
A1.KE.00.KSM4024         .........G-------------------T---...---------G-C--------------A---------...-----------------G--G-----A---------------------.........-----------------------A-----A-G------ 7017
A1.KE.00.MSA4069         .........G-------------------T---...-----G-----TG-------------A---------...--------------------A-----A---------------------.........--------------C--------A-----A-G------ 7096
A1.KE.00.NKU3005         .........G----------------G--TAC-...C--------G-A---A----------A---------...--------------------G---G-A---------------------.........------C-------C--------A-----A-G------ 7099
A1.RU.00.RU00051         .........G----------G------A-T---...C-G------GCC--------------A---------...-----T-----------G--G-----A---------------------.........------C-------C-----A--A-----A-G------ 7162
A1.RW.93.93RW_024        .........--------G---------A-T---...---------G-C--------------A---------...-----------------G--G-----A---------------------.........--------------C--------A-----A-G------ 7090
A1.SE.95.SE8891          .........G-----------------A-T---...---------G-C---A--------------------...--------------------G-----A--------------------G.........--------------C--------A-----A-G------ 7053
A1.SE.95.UGSE8131        .........A-------------------T---...C-G-C----G-A---T----------A---------...-----------------G--G-----A---------------------.........-----------------------A-----A-------- 7266
A1.TZ.01.A173            .........G-------------------C---...C-G------G-C---A----------A---------...-----------------G--G---G-A---------------------.........--------------C--------A-----A-G------ 7097
A1.UA.01.01UADN139       .........G-----------------A-AATTGGAC-G------GCC--------------AA-T------...--------------------G---C-----------------------.........--------------C--------A-----A-G------ 7084
A1.UG.92.92UG037         .........G-------------------TAC-...C-G------G-A---A----------A---A----T...--------A--------G--G-----A---------------------.........A-------------C--------A-----A-G------ 7898
A1.UG.99.99UGA07072      .........G-------------------T---...--G------G-CC--A----------A---------...-----------------G--G-----A---------------------.........--------------C--------A-----A-G------ 7094
A2.CD.97.97CDKTB48       .........--------------------T---...T-G------G-T------------------------...-----------------G--G-----A---------------------.........---------A----C--------A-----A-G------ 7198
A2.CY.94.94CY017_41      .........G-------------------T---...C-G------G-C------------------------...-----------------G--G---C-A---------------------.........-----------------------A-----A-G------ 7260
A.SN.01.DDI579           .........G-------------------T---...C--------G-C--------------A---------...-----T-----------G--G---------------------------.........--------------C-----A--A-------G----C- 7092
A.SN.01.DDJ369           .........G----------------G--C---...C-G------G-C--------------A---------...-----T-----------G--G-----A----------C----------.........--------------------A--A-------G----C- 7110
A.SN.96.DDJ360           .........--------------------T---...C--------G-C---T----------A---------...-----T-----------G--G-----A----------C----------.........--------------C-----A--A-------G----C- 7074
A.CD.02.02CD_KTB035      .........G-----------------A-T---...--G--------C--------------A---------...--------------------G-----A---------------------.........-----------------------A-----A-G------ 7887
A.CD.97.97CD_KCC2        .........GG------------------T---...--G--------C--------------A---------...-----------------G--G-----A---------------------.........---------G-------------A-----A-G------ 7887
A.CD.97.97CD_KTB13       .........G-------------------T---...-----------C---------A----A---------...-----------------G--G-----A--------------A-----G.........---G-------------------A-------G------ 7164
B.AR.04.04AR151516       .........------------------------ACA---------A--------------------------...--------------------G---C-----------------------.........-----------------------A-------------- 7097
B.AU.87.MBC925           .........------------------------...C-----------------------------------...--------------------G-----A---------------------.........------------------------------------C- 7907
B.BO.99.BOL0122          .........--A---------------------ACG---------A--------------------------...--------------------G---G----------------A------.........------------------------------------C- 7089
B.BR.03.BREPM2012        .........--A---------------------ACA---------A--------------------------...--------------------GG--------------------------.........-T---------------------A-----------C-- 7333
B.CA.97.CANB3FULL        .........------------------------...---------G--------------------------...--------------------G---G-----------------------.........-----------------------A-------------- 7161
B.CN.05.05CNHB_hp3       .........--A---------------------GCA--------GA--------------------------...--------------------G-----A---------------------.........-----G---A-------------A-----A-GA----- 7904
B.CO.01.PCM001           .........--------------------AAC-...C--------A--------------------------...--------------------G-----A---------------------.........-----------------------A------------C- 7071
B.GB.83.CAM1             .........------------------------GCG---------C--------------------------...-------------------TGG--C-----------------A-----.........-----------------------A------------C- 7888
B.GB.86.GB8              .........------------------------ATG---------A--------------------------...--------------------G---C-----------------------.........-T---------------------A-----A--T----- 7879
B.GE.03.03GEMZ010        .........------------------------GTGT-T-----CA--------------------------...--------------------G-----A---------------------.........-----------------------A-------G------ 7065
B.IT.05.SG1              .........----------G-------------...------------------------------------...--------------------G---C-----------------------.........-T---------------------A-----A-------- 7573
B.JP.05.DR6538           .........--------------------A......---------A--------------------------...--------------------G---G-----------------------G.........-----------------------A------------C- 7921
B.KR.05.05CSR3           .........---------------------AC-...C-G-----GA----------T---------------...--------------------G---------------------------.........--------A--------------A-----A-----C-- 7951
B.NL.00.671_00T36        .........------------------A-AACG...C--------A--------------------------...--------------------G------G--------------------.........-----------------------A-----A-------- 7449
B.RU.04.04RU128005       .........------------------------...C--------G-------G------------------...--------------------G---------------------------.........-----------------------A-----A--Y----- 7360
B.TH.00.00TH_C3198       .........------------------------...------------------------------------...--------------------G---------------------------.........-----G-----------------A-----A-GA----- 7086
B.UA.01.01UAKV167        .........---------------AGTCCAACG...C-T------C--------------------------...--------------------G-----A---------------------.........-----------------------A-----------C-- 7152
B.US.04.ES10_53          .........------------------------...---------G--------------------------...--------------------G-----A--------------------G.........-T---------------------A-----A-----C-- 7866
B.US.99.PRB959_03        .........------------------------...---------G--------------------------...--------------------G-----A---------------------.........-T---------------------A------------C- 7086
C.AR.01.ARG4006          .........G-----------------A-A-TGGGAC--------G--------------------------...--------------------A-----A---------------------.........-----------------------------A-G------ 7079
C.BR.04.04BR013          .........------------------------...---------G--------------------------...--------------------A-----A---------------------.........--------------A--------A-----A-G------ 7379
C.BW.00.00BW07621        .........G-------------------A-TGGGA---------G---------A----------------...--------------------A-----A-------T-------------.........---C-G-----------------A-----A-------- 7223
C.CN.98.YNRL9840         .........G------------------C----...---------G--------------------------...--------------------G-----A----------C----------.........-----G-----------------A-----A-G------ 7089
C.ET.02.02ET_288         .........GG-----------------CA......C-------------------------------G---...--------------------G-----A---------------------.........-----------G-----------------A-G------ 7093
C.GE.03.03GEMZ033        .........G---A-------------------...---------G-AC-----------------------...--------------------G-----A---------------------.........---G-G--------A--------A-----A--T----- 7095
C.IL.99.99ET7            .........G------------------C----...---------------------C--------------...--------------------G-----A--------------------G.........-----G-----------------A-----A-G------ 7054
C.IN.99.01IN565_10       .........G------------------C-......C--------A--A-----------------------...--------------------G-----A---------------------.........-----G-----------------A-----A-G------ 7265
C.KE.00.KER2010          .........G-----------------------...---------G-AC-----------------------...-----------------G--G-----A---------------------.........---C-G-----------------A-----A-G------ 7077
C.MM.99.mIDU101_3        .........G------------------CA......---------A--------------------------...-----------------G--G--G---G--------------------.........-----G-----------------A-----A-G------ 7247
C.MW.93.93MW_965         .........G-------G---------------...---------G-----------------A--------...-----------------G--G-----A---------------------.........-----------------------A-----A-G------ 7063
C.SN.90.90SE_364         .........G-----------------------...---------G--------------------------...-----------------G--A-----A---------------------.........-----G-----------------A-----A-G------ 7033
C.SO.89.89SM_145         .........G-----------------------...---------G--------------------------...-----------------G--G-----A---------------------.........---C-G-----------------A-----A-GT----- 7114
C.TZ.02.CO178            .........G-----------------------...C--------A--------------------------...--------------------G--G--A----------C----------.........---G-GC----------------A-----A-G------ 7092
C.UY.01.TRA3011          .........G------------------C-......C--------A--------------------------...--------------------G-----A---------------------.........--------------G--------A-----A-G------ 7065
C.YE.02.02YE511          .........G------------------CA---...--G------------------------TC-------...--------------------A-----A--------------------G.........-----G-----------------A-----A-G------ 7080
C.ZA.04.04ZASK164B1      .........G-------------------A---...---------G--------------------------...--------------------A-----A---------------------.........-----G--------A--------A-----A-G------ 7338
C.ZA.05.05ZAPSK240B1     .........G-------------------A---...C--------C--------------------------...--------------------A-----A---------------------.........---G-G-----------------A-----A-G------ 7264
C.ZM.02.02ZM108          .........G-----------------------...---------G--A-----------------------...--------------------A---C-A---------------------.........-----------------------A-----A-G------ 7925
D.CD.83.ELI              .........G-A---------------A-A---...T--------A--------------------------...--------G---------CGG---G-----------------------.........------A----------------A-----A-------- 7424
D.CM.01.01CM_0009BBY     .........G-A---------------A-A---...T--------A--------------------------...--------G-----------G---------------------------.........-T---------------------A-----A--T----- 7113
D.KE.01.NKU3006          .........G-A---------------A-A---...C-G-----------------A---------------...--------G-----------G--G------------------------.........---G-------------------A-----A-------- 7131
D.KR.04.04KBH8           .........G-A--G------------A-A--G...C--------A--------------------------...--------A-----------G---G-----------------------.........--G--------------------A-----A-------- 7801
D.TD.99.MN011            .........G-A---------------A-A---...T--------A--------------------------...--------G-----------G---G-----------------------.........-----------------------A-----A-------- 7078
D.TZ.01.A280             .........G-A---------------A-A---...C--------A--------------------------...--------G-----------G---T-----------------------.........--G--------------------A-----A-------- 7076
D.UG.99.99UGK09259       .........G-A---------------A-A---...C--------A--------------------------...--------G-----------G--CC-----------------------.........--G--------------------A-----A--T----- 7089
D.YE.01.01YE386          .........G-A---------------A-A---...C-----------------------------------...--------G-----------G---C-----------------------.........--G--------------------------A--T----- 7044
D.YE.02.02YE516          .........G-A---------------A-A---...C--------A--------------------------...--------G-----------G---G-----------------------.........-----------------------A-----A--T----- 7068
D.ZA.90.R1               .........GCA--G------------A-A---...C---CGAGC-G-------------------------...--------A-----------G--TG-----------------------.........--G--------------------A-----A-------- 7242
F1.AR.02.ARE933          .........A-----------------------...--G------------A------------C-------...---G-------------G--G-----A---------------------.........--G---C-------A--------A-----A-------- 7158
F1.BE.93.VI850           .........-------------------C----CTAGG--CTTTG--CC-T..G--T--------GA-----...--G--.-----------G--G--G--A---------------------------------------------A--------A-----A-------- 7165
F1.BR.01.01BR125         .........A-----A------------C----...---AC-------------------------------...-------------------------------------------A-CAGACGAATTATTGGACT-G-C-ATA-TGCAGT-..-GCGG-------C- 7322
F1.ES.x.P1146            .........A------------------CA---...C----CT-----------------------------...--------------------G-----A---------------------.........--T-----------A--------A-----A-------- 7195
F1.FI.93.FIN9363         .........AG--------G-----------C-...---------G-A---------------A-C------...-----------------G--G---C-A---------------------.........--------------A--------A------------C- 7184
F2.CM.02.02CM_0016BBY    .........A---------G---------A---...---------G--------------------------...--------------------G-----A---------------------.........-----------------------A-----A-------- 7047
F2.CM.95.MP257           .........------------------------...--G------A-------------------------C...-----------------G--G-----A---------------------.........A-C--------------------------A-G------ 7081
F2.CM.97.CM53657         ......AGAG-A---------------------...---------G--C-----------------------...-----------------G--G---------------------------.........-----------------------A-----A-------- 7059
G.BE.96.DRCBL            .........G-------------------T---...G--------A-C--------------A---A-----...--------------------G-----A----------------T----.........--------------C--------A-----A-G------ 7827
G.CM.01.01CM_4049HAN     .........G-------------------T---...---------G-C--------------A---------...-----------------G--G-----A----------------T----.........--------------C--------A-----A-------- 7077
G.CU.99.Cu74             .........G-------------------A---...--GA-----G-C--------------A---------...-----------------G--G-----A----------------T---G.........--------------C--------A-------G------ 7507
G.ES.99.X138             .........G-------------------T---...C-G----GG-CC-------------A---A-----...-----------------G--G-----A----------------T----.........--------------C--------A-----A-G------ 7323
G.GH.03.03GH175G         .........G-------------------T---...C-G------G-C--------------A---------...-----------------G--G-----A----------------T---G.........--------------C--------A-----A-G------ 7934
G.KE.93.HH8793_12_1      .........GG------G-----------T---...G-G------G-C--------------A---------...-----------------G--G-----A----------------T----.........-----G--------C--------A-----A-G------ 7258
G.NG.01.01NGPL0669       .........--------------G-----T---...--G------G-CC-----------------------...-----------------G--G-----A----------------T----.........--------------C--------A-----A-G------ 7081
G.PT.x.PT2695            .........GG------------------T---...C-G------G-CC-------------A---A-----...-----------------G--G-----A----------------T----.........--------------C--------A-----A-G---C-- 7899
G.SE.93.SE6165           .........G-------------------T---...C-G------G-C--------------A---------...-----------------G--G-----A----------------T----.........--------------C--------A-----AGG---C-- 7306
H.BE.93.VI991            .........G-------------------A---...--G--------T------------------------...--------------------G--T--A---T-----------------.........-----------------------A-----A-G------ 7287
H.BE.93.VI997            .........G----------------T------...--G--------T-----------------------T...--------------------G-----A--------------------G.........-----------------------------A-G------ 7202
H.CF.90.056              .........G-----------------------...--G-------CT------------------------...-----------------G--G-----A--------------------G.........-----------------------------A-G------ 7207
J.CD.97.J_97DC_KTB147    .........G-----------------A-A-TGGGA--G------G--------------------A-----...--------------------G-----AG--------------------.........--G--G--------C--------A-----A-G------ 7091
J.SE.93.SE7887           .........G-----------------------...---------G-------------------A------...-----T--------------G-----A----------------T----.........-----G--------C--------A-----A-G------ 7195
K.CD.97.EQTB11C          .........-----------G--------A---...--------G---------------------------...-----------------G--G-----A---------------------.........-----------------------A-----A-------- 7071
K.CM.96.MP535            .........--------------------A---...T--------G-----T----A---------------...-----------------G--G-----A---------------------.........-----------------------A-----A-G------ 7048
```

Env gp120 end Env gp41 start

```
B.FR.83.HXB2                 .........CAGAGAGAAAAAAGAGCAGTGGGA...ATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCA...GGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGA.........CAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTT 7888
Env                          _.__.__.__Q__R__E__K__R__A__V__G__.__I__G__A__L__F__L__G__F__L__G__A__A__.__G__S__T__M__G__A__A__S__M__T__L__T__V__Q__A__R__.__.__.__Q__L__L__S__G__I__V__Q__Q__Q__N__N__L
01_AE.CF.90.90CF11697        .........G-----------------------...---------A--A--T----------A---------...-----------------G--G---C-A---------------------.........-----------------------A-----A-G------ 7840
01_AE.CN.05.FJ051            .........G-A--G------------------...---------A--A--T----------A---------...--------------------A-----A---------------------.........-----------------------A-----A-G------ 7870
01_AE.CN.06.FJ054            .........G-----------------------...---------A--A--T----------A---------...-----------------G--G-----A---------------------.........----------------------GA-----A-G------ 7897
01_AE.HK.x.HK001             .........G-------------------A-TGGGA---------A--A--T----------A---------...-----------------G--G---T-A---------------------.........------C----------------A-----A-G------ 7249
01_AE.TH.01.01TH_R2184       .........G-----------------------...---------A--A--T----------A---------...-----------------G--G-----A---------------------.........-----------------------A--A--A-G------ 7091
01_AE.TH.02.OUR769I          .........G-----------------------...---------A--A--T----------A---------...--------------------G-----A---------------------.........-----------------------------A-G------ 7072
01_AE.TH.90.CM240            .........G-------------G---------...---------A--A--T----------A---------...-----------------G--G-----A---------------------.........-----------------------A-----A-G------ 7441
01_AE.US.00.00US_MSC1164     .........G-----------------------...---------A--A--T----------A---------...-----------------G--G-----A---------------------.........-----------------------A-----A-G------ 7084
02_AG.CM.02.02CM_4082STN     .........G------G------------T---...C-G------G-C--------------A---------...-----------------G--G-----A---------------------.........-----------------------A-------G----C- 7066
02_AG.EC.x.ECU41             .........G------------------CT---...C-G------G-C--------------A---------...-----------------G--G-----A---------------------.........--G--------C-----------A-----A-G------ 6950
02_AG.GH.03.GHNJ196          .........GGA---------------A-A--TGGAC-G------GCCC-----------C-A---------...-----------------G--G-----A---------------------.........-----------------------A-------G----C- 7935
02_AG.NG.01.PL0710           .........G-------------------C......G-T------C-C--------------A---------...-----------------G--G-----A---------------------.........-----------C-----------A-------G------ 7036
02_AG.NG.x.IBNG              .........G----------------G--T---...C-G------G-C--------------A---------...-----------------GCGG-----A---------------------.........-----------C--C--------A-------G------ 7394
02_AG.SE.94.SE7812           .........G-----------------C-T-TTGGAC-G--------C---T----------A---------...-----------------G--G-----A---------------------.........-----------------------A-------G----C- 7278
02_AG.SN.98.MP1211           .........G-------------------T---...C-G------G-C--------------A---------...--------------------G-----A---------------------.........-----------------------A------------C- 7090
03_AB.RU.97.KAL153_2         .........--------G---------------...---------G--------------------------...-----------------G--G-----A---------------------.........-----------------------A--------T---C- 7065
04_cpx.CY.94.CY032           .........------------T-----------...-----G--CA--------------------------...--------------------G---------------------------.........-----------C--C----------------G------ 7258
05_DF.BE.x.VI1310            .........A-----------------A-----...---------A------------------C-------...-----------------G--G-----AG--------------------.........--------------A--------A-----A-------- 7294
06_cpx.AU.96.BFP90           .........GG------------------T---...C-G------A----------------A---A-----...--------------------G-----A----------------T----.........--------------C--------A-----A-G------ 7936
08_BC.CN.97.97CNGX_6F        .........G-----------------------...C--------G--------------------------...-----------------G--G-----A---------------------.........-----G-----------------A-----A-G------ 7067
09_cpx.GH.96.96GH2911        .........G-------------------T---...--G------C-T--------------A---------...--------------------G-----A---------------------.........--------------C--------A--A----G----C- 7060
10_CD.TZ.96.96TZ_BF061       .........G-A---------------A-A---...C-G------G--------------------------...--------A--------G--G--GC-----------------------.........---G----------------------A--A-------- 7240
11_cpx.GR.x.GR17             .........G-----------------------...---------G--C-----------------------...-----------------G--G-----A---------------------.........--------------C--------A-----A-------- 7155
12_BF.AR.99.ARMA159          .........AC----------------------...------------------------------------...-----------------G--G---G-----------------------.........--------------A--------A-----A-----C-- 7853
13_cpx.CM.96.1849            .........GCA-----------------A-TTGGAC----G---G-C--------------A---------...----------------TG--G---C-A---------------------.........--------------C--------A--A---------C- 7313
14_BG.DE.01.9196_01          .........------------------------...------------C--T--------------------...--------------------G---------T-----------------.........---C--C----------------A------------C- 7393
15_01B.TH.99.99TH_MU2079     .........------------------------ACG---------A--------------------A-----...--------------------G-----A--------------------G.........-----G-----------------A-------GA---C- 7260
16_A2D.KR.97.97KR004         .........G----G-------------CA......G-T---CTG--C------------------------...-----------------G--G---G-----------------------.........--------------C--------A-----A-G------ 7254
17_BF.PE.02.PE02_PCR0155     .........--A---------------------...--G-C----A--------------------------...--------------------G-----A---------------------.........--------------A--------A-------G------ 7057
18_cpx.CU.99.CU76            .........GG---------------C--T---...C-G------G-C--------------A---------...--------------------A-----A---------------------.........--------------C--------A-----A-G------ 7196
19_cpx.CU.99.CU7             .........G-----------------A-----...C-G------G-C--------A-----A---------...-----------------G--G--C--A---------------------.........-----------C--C--------A-------G------ 6954
20_BG.CU.99.Cu103            .........--------------------T---...C-G------G-C--------------A---------...-----------------G--G-----A----------------T---G.........--------------C--------A-------G----C- 7322
21_A2D.KE.91.KNH1254         .........ACA---------------A-A---...C--------A--------------------------...--------G-----------G-----A---------------------.........--G--------------------A-----A-GT----- 7110
22_01A1.CM.01.01CM_0001BBY   .........G-----------------C-----...T-T------G-C--------------A---------...-----------------G--G-----A---------------------.........-----G--------C--------A-----A-G------ 7055
23_BG.CU.03.CB118            .........G------G------------T---...C-G------G-C--------------A---------...-----------------G--G-----A----------------T---G.........--------------C--------A-------G----C- 7316
24_BG.CU.03.CB378            .........G-------------------C---...C-G------G----------------A---A-----...--------------------G---C-A----------------T---G.........--G---C-------C--------A-------G----C- 7331
24_BG.CU.03.CB471            .........GG------------------T---...C-G------G----------------A--------G...--------------------G-----A----------------T---G.........------C-------C--------A------GG----C- 7334
25_cpx.CM.01.101BA           .........G-------------------T---...--G------G-C--------------A---------...--------A-----------R-----A----------------T----.........--------------C--------A-----A-G------ 7138
27_cpx.CD.97.97CDKTB49       .........G-C---------------------...--------CA--A--T----------A---------...--------------------G-----A---------------------.........-----G-----------------A-----A-G------ 7092
28_BF.BR.99.BREPM12609       .........------------------------GCAC--------C--------------------------...--------------------G--------------------A------.........------A----------------A-----A-------- 7289
29_BF.BR.02.BREPM119         .........------------------------ACAC--------G--------------------------...--------------------G-----A---------------------.........-----------------------A-------G----C- 7040
31_BC.BR.02.110PA            .........G------------------CA......C--------A--------------------------...--------------------G-----A---------------------.........--------------G--------A-----A-G------ 7381
32_06A1.EE.01.EE0369         .........GGA-----G-----------T---...C-G------G-T--------T-----A---A-----...--------G-----------G-----A----------------T----.........--------------C--------A-----A-G------ 7514
33_01B.MY.05.05MYKL007_1     .........G-----C-----------------...-------T-A--A--T----------A---------...-----------------G--G--C--A---------------------.........-----------------------A-----A-G------ 7270
34_01B.TH.99.OUR2478P        .........G-----------------------...---------A--A--T----------A---------...-----------------G--G-----A---------------------.........-----------------------A-----A-G------ 7054
35_AD.AF.05.05AF095          .........G-------------------C---...C-G------A-----A----------A---------...-----------------G--G---------------------------.........--------------C--------A-----A-G------ 7054
36_cpx.CM.00.00CMNYU830      .........G-------------------T---...C-G------G-C--------------A---------...-----------------G--G-----A--------------A------.........------C----C-----------A-----A-G------ 7085
37_cpx.CM.00.00CMNYU926      .........G-------------------T---...C-G------G-C--------------A---------...-----------------G--G--T--A-----------------T---.........--------------C--------A-------G---CC- 7070
38_BF1.UY.05.UY05_4752       .........--------T---------------...--G-C----A--------------------------...--------------------G-----A---------------------.........--------------A--------A-----A-G------ 7299
39_BF.BR.03.03BRRJ103        .........A-----------------------TTAT-G------A--------------------------...-----------------G--G-----A---------------------.........--G-----------A--------A------------C- 7369
40_BF.BR.04.04BRRJ115        .........--------------------AAC-...--G------A--------------------------...--------------------G---C-A---------------------.........-----------------------A------------C- 7445
42_BF.LU.03.luBF_05_03       .........------------------...-TG..........................................--------------------G---G-A---------------------.........-----------------------A-------------- 7350
43_02G.SA.03.J11223          .........--------------G-----T---...--G------G-C--------------A---------...-----------------G--G-----A----------------T----.........--------------C--------A-----A-G------ 7371
U.CD.83.83CD003              .........G----------------------G...--G-----------T--C-----------------C...-----------------G--G------G---------C----------.........-----------------------------A-------- 7364
U.CD.90.90CD121E12           .........G-----------------------...--G-----------T--C-----------------C...--------------------G-----AG---------C----------.........-----------------------A-----A-------- 7059
U.GR.99.GR303                .........G-------------------T---...--G-CG---A-CA-----------------------GCA-----------------G--G---G-----------------------.........--------------C-------GA-------G----C- 7172
U.NL.95.U_NL_95_H10986_D1    .........--------------------A---...C-G------------T--------------------...-----------------G--G-----A---------------------.........-----------------------A-----A-------- 7794
CPZ.CD.90.ANT                ...CAACACTCC---C---------GCA-T---...-----GCTG--C---T-G--TC-TC-CA-T-----T...--C--T--A-----------G-----AG---------C----A----G.........A-T--GY-CCA------T--A--A------GC----C- 7320
CPZ.CM.01.SIVcpzCAM13        ......GAAAG----C-G--G-------CCTTTGGCC-G-----AC-------------TC----T------...--------A-----------GG--G-C------------------C--.........--G---C-A--G--A--------A--A--------CC- 7472
CPZ.CM.05.SIVcpzEK505        ............---C------------CCTTTGGAC----T--GC----T--------TC-T---------...--G-----------------G-----A------------------C-G.........A-------A-----G-----------------T---C- 7371
CPZ.CM.05.SIVcpzLB7          ............----------------CA--G...C-G----TG--T------------------------...--G-----------------A---T-A---------------------.........---------A----C--------A-------GT----- 7399
CPZ.CM.05.SIVcpzMB66         ............--G-------------CC-T-GGAT-G---------------------------------...--------------------G---C-----------------------.........---C-GC------------------------G----C- 7512
CPZ.CM.05.SIVcpzMT145        ...CAACAC--T--------G-----CC-T---...C-------AC----T---------C-----------...--------------------A---G-CGT-------------A---C-G.........-----G--AA-G--C-----------------T---C- 7397
CPZ.GA.88.GAB1               ...CAGAAAG-C---C-G----------CCTTCGGTC-G-----AC--------------C----T------...--------------------GG--G-A------------------C--.........--G--GC----G--------------A-----T---C- 7925
CPZ.TZ.00.TAN1               ...AATCATACT--GTC--G-----AT-----C...------CTG-----T-----A--TC-TA-T------...-----T--A-----------G-----AG---------C----------.........GG----C-C--------T--A--------AC-A--CC- 7514
CPZ.US.85.CPZUS              .........A-A----------------CCTTCGGAC----T--AC-------G-----TC-T---------...--------------------A---G-AGT-----------------C--.........-----------A--G--------A--------T---C- 7893
CPZ.CM.x.SIVcpzDP943         .........ACA----------------CCTTTGGAC----T--GC----T--------TC-A---------...--------------------A---C-A------------------C-G.........-----------A--G--------A-----------CC- 7451
CPZ.CM.x.SIVcpzMB897         .........--A--------G--G-----A---...C-G----TGA----------A---A-----------...--------------------GG--C-----------------------.........A-TC--C-C--------------A-----A-G------ 7396
N.CM.02.DJO0131              .........AGT----------------CCTTTGGAC----T--GC----T--G-----TC-T---------...--G-----------------G-----A------------------C-G.........AC---G--A-----G--------A--------T---C- 7312
N.CM.04.04CM_1015_04         .........A-T--------G-------CCTTTGGAC----T--GC----T--------TC-T--T------...--G-----------------G-----A------------------C-G.........A-C-----A-----G--------A--------T---C- 7290
N.CM.04.04CM_1131_03         .........AGT--------G-------CCTTTGGAC----G--GC----T--------TC-T--T------...--G-----------------G-----A------------------C-G.........AC------A-----G--------A--------T---C- 7355
N.CM.95.YBF30                .........AGT----------------CCTTTGGAC----T--GC----T--------TC-T---------...--G-----------------G-----A------------------C-G.........AC------A-----G--------A--------T-T-C- 7436
N.CM.97.YBF106               .........AGT--------G-------CCTTTGGAC----T--GC----T--------TC-T---------...--G-----------------G-----A--------------A---C-G.........AC------A-----G--------A--------T---C- 7391
O.BE.87.ANT70                ...AGAACT--T-----------------A---...T-G---ATGC-A---T-G---G-TC-AA-T------...--T-----------------GG---CA------G--------A--CAC.........ACT--GC--AAG-----------A------G----CC- 7941
O.CM.91.MVP5180              CACACCCCT--C--G--------------A---...T-G---ATGC-A---T-G---G-GC-AA-T------...--T-----------------GG---CAG------------G-A--CAC.........AGTG--C--AAG-----------A------G----CC- 7958
O.CM.96.96CMABB637           ...GGCAAT--T---A-G-----------A---...T-G---ATGC-A---T-A---G-TC-AA-T------...--T-----------------GG--G-----------------A--CA-.........ACT--G---AAG-----------A------G----CC- 7431
O.CM.98.98CMA104             ACTGGTACT--T--G-----G--------A---...T-G---ATGC-A---T-A---G-TC-AA-T------...--T-----------------AG---CAG--------------A--CAC.........ACT--GA-AAAG-----------A------G----CC- 7436
O.CM.99.99CMU4122            ...GGTACT--T--G--G-----------A---...T-G---ATGC-----T-G---G-TC-AA-T------...--T---------------A-GG---CA---T--G--------A--CAC.........ACT--GA--AAG-----------A------G----CC- 7408
O.FR.92.VAU                  ...AGCACT--T-----G----------CA--T...T-G-C-ATGC-A---T-G---A-TC-AA-T------...--------------------GG---CAG------------G-A--CAG.........--TC-GA-AAAG-----------A------G-T--CC- 7504
O.SN.99.SEMP1299             ...GGCACT--A-----------------A---...T-G---ATGC-A---T-A---G-TC-AA-T------...--------------------GG---CAG-----G--------A--CA-.........TCT--GA--AAG-----------A------G----CC- 7991
O.US.99.99USTWLA             ...AACCAT--A-----------G-----A---...T-G---ATGC-A---T-G---G-TC-AA-T------...--T-----------------GGG---A------T------GAA--CAC.........TC----A-AAGG-----------A------G----CC- 7418
```

```
B.FR.83.HXB2 GCTGAGG...GCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGA 8055
Env __L__R__.__A__I__E__A__Q__Q__H__L__L__Q__L__T__V__W__G__I__K__Q__L__Q__A__R__I__L__A__V__E__R__Y__L__K__D__Q__Q__L__L__G__I__W__G__C__S__G__K__L__I__C__T__T__A__V__P__W__
A1.GE.99.99GEMZ011 -------...-----A-----T--G-----GT--C--A-------G-----------T--A---------------G-------------------------------G--------A--A--------------------------C---------AA------C---- 7230
A1.KE.00.KER2008 -------...-----A-----T-----A--A------A-------G-----------T--A---------------G------------------------G---------------A--A--------C-----------------C---------AA----------- 7261
A1.KE.00.KNH1144 -------...-----A-----T-----A---A-----A-------G-----------T--A---------------G----T-------------------G---------------A--A--------C-----------------C---------AA------C---- 7287
A1.KE.00.KSM4024 -------...-----A-----T-----A---------A-------------------T--A---------------G------------------------G---------------A--A--------C-----------------C---------AA------C---- 7184
A1.KE.00.MSA4069 -------...-----A-----T-----A--AA-A---AG------G-----------T--A------------C--G--------------GC--------G---------------A--A--------C-----------------C---------AA------C---- 7263
A1.KE.00.NKU3005 -------...-----A-----T---------A-----A-------G-----------T--A---------------G-------------G----------G------------------A--------C-----------G-----C---------AG------C---- 7266
A1.RU.00.RU00051 -------...-----A-----T--G---------C--A-G-----G-----------T--A---------------G----------------------------------------A--A--------------------------A---------AA---A--C---- 7329
A1.RW.93.93RW_024 -------...-----A-----T-----A---------A-------G-----------T--A---------------G----------------------------------------A--A-----------------G--------C---------AA------C---- 7257
A1.SE.95.SE8891 -----T-...-----A-----T---------------A-------G-----------T--A------------CA-G------------------------G---------------A--A--------C-----------------C---------A------------ 7220
A1.SE.95.UGSE8131 -------...-----A-----T---------A-----AG------G-----------T--A---------------G--A----------------------------------T--A--A--------C-----------GA----C---------------------- 7433
A1.TZ.01.A173 -------...-----A-----T---------------A-------G-----------T--A---------------G----------------------------------------A--A--------C-----------------C---------AA------C---- 7264
A1.UA.01.01UADN139 -------...-----A-----T--G------T--C--AG------G-----------T--A-----A---------G-------------------------------G--------A--A--------------------------C---------AA------C---- 7251
A1.UG.92.92UG037 -------...-----A-----T---------------A-------T-----------T--A---------------G------------------------G---------------A--A--------C-----------------C---C-----AA------C---- 8065
A1.UG.99.99UGA07072 -------...-----A-----T---------------A-------G-----------T--A---------------G----------------------------------------A--A--------C-----------------C---------AA------C---- 7261
A2.CD.97.97CDKTB48 -----A-...-----A-----T-----A--GA-----AGG-----------------T--A---------------G--------C--------------C----------------A--A--------C-----------------C---G------A----AGA---- 7365
A2.CY.94.94CY017_41 ----CA-...-----A--A--T---------------A-------------------T--A-----------G--GG-------------------------------------------A--------C-----------------C---G-----A-------C---- 7427
A.SN.01.DDI579 -------...-----A-----T--------CA-----A-------------------T--A---------------G--------A-A-----------------------------A--AC----------------G--------C---------AA------C---- 7259
A.SN.01.DDJ369 -------...-----A-----T--------AA-----A-------G-----------T--A---------------G----------------G-----------------------A--A--------------------------C---------AA----------- 7277
A.SN.96.DDJ360 -------...-----A-----T--------CT-----A-------------------T--A---------------G------------------------G---------------A--AC----------------G--------C-----------------C---- 7241
A.CD.02.02CD_KTB035 -----A-...-----A-----T-----A---------A-------------A-----T--A--------------------------A-----------------------------AA-A----AA--A-----------------C---------AA------C-A-- 8054
A.CD.97.97CD_KCC2 -----A-...-----A-----T-----A---------AG------------------T--A---------------G----------------------------------------A--A--------C-----------------C---------A------------ 8054
A.CD.97.97CD_KTB13 -----A-...-----A-----T-----A---T-----A----------------------A---------------G----------------------------------------A--A--------C--------G--------C---------AA----------- 7331
B.AR.04.04AR151516 -------...-----------R---------A--------------------------------------------G------------------------G---------------A-----------------------------C---------A------------ 7264
B.AU.87.MBC925 -------...------------------------------------------------------------------C-------------------------------------------A-----------------G------------------A------------ 8074
B.BO.99.BOL0122 -------...---------------------A--------------------------------------------C------------------------G---------------A-----------A---------------------------AA------C---- 7256
B.BR.03.BREPM2012 ------A...---------------------A-------G------------------------------------G------------------T-----G---------------A---------------------------------------------------- 7500
B.CA.97.CANB3FULL -------...------------------------------------------------------------------C----------A-----------------------------A---------------A------------------------------------ 7328
B.CN.05.05CNHB_hp3 ------A...--------------------------------------------------A---------------G--T-------------------------------------A---------------------------------------------------- 8071
B.CO.01.PCM001 -------...-----------------A-----------------------------T------------------G----------A--------------------------------A--------------------G---------------A------------ 7238
B.GB.83.CAM1 -------...------------------------------------------------------------------G----------------------------------------A-----------------------------C---------------------- 8055
B.GB.86.GB8 -------...----------------------------G-------------------------------------G----------------------------------------A---------------------------------------A------------ 8046
B.GE.03.03GEMZ010 ------A...-----------------------------------------------T------------------G----------A------------C---------G------A---C-----------------------------------A------------ 7232
B.IT.05.SG1 ------A...---------------------A-A---------T--------------------------------G--------A-A-------------G-------------------------------------------------------T------------ 7740
B.JP.05.DR6538 -----AA...------------------------------------------------------------------G--------A----G----------G---------------A---------------------------------------------------- 8088
B.KR.05.05CSR3 -------...------------------------------------------------------------------G----A--------G----------G---------------A-----------------------------C---------A----A------- 8118
B.NL.00.671_00T36 -------...------------------------C----------------------T------------------G--------A-A------------C----------------A--A--------------------G---------------A------------ 7616
B.RU.04.04RU128005 -------...-----------------A------------------------------------------------G---------------------------------------------------------------------C-----------------A---- 7527
B.TH.00.00TH_C3198 ------T...-------------------G----------------------------------------------G----------------------------------------A---------------------------------------------------- 7253
B.UA.01.01UAKV167 ------A...-----------------------------------------------T------------------G--------------------T--GG------G--------A---------------G-------------C---------A------------ 7319
B.US.04.ES10_53 ------A...--------------------------------------------------A---------------G-------------G----------G------------------A--------------------------C---------AA----------- 8033
B.US.99.PRB959_03 -------...--------------------CA--------------------------------------------G----A--------G----------G---------------A---------------------------------------T------------ 7253
C.AR.01.ARG4006 -------...-----A---------------A-------------G-----------T--------T---A-----G--------C-A--G----------G---------------A--A--------C-----------------C---------AA----------- 7246
C.BR.04.04BR013 -------...-----A---------------A-------------G-----------T--------T---A-----G--------C-A-------------G---------------A-----------C-----------------C---------------------- 7546
C.BW.00.00BW07621 ------A...-----A---------------A-------------G-----------T------------A-----G-T------A-A-----------------------------A-----------C-----------------C-----------C---------- 7390
C.CN.98.YNRL9840 -----T-...-----A-----------------------------G-----------T------------A-----G--------A-A-----------------------------A-----------C-----------------C-----T--------A------- 7256
C.ET.02.02ET_288 ------A...-----A---------------A-------------G-----------T----------G-A-----G-------GA-A------------C----------------A-----------C-----------------C---C-----A------------ 7260
C.GE.03.03GEMZ033 -------...-----A---------------A-------------------------T--A---------------G--T-----A-A------------C----------------A--AC-C-----------------------C------T-CTT----------- 7262
C.IL.99.99ET7 -------...-----A---------------A-------------------------T--------G---A-----G--------A-A-----------------------------A-----------C-----------------C---------AA----------- 7221
C.IN.99.01IN565_10 -----A-...-----AA-T------------A-------------G-----------T------------A-----G--------A-A-----------------------------A-----------C-------------G---C--------------A-----A- 7432
C.KE.00.KER2010 -------...-----A--A------------A-------G-----G-----------T------------------G--------A-A-----------------------------A-----------C-----------G-----C---------AA----------- 7244
C.MM.99.mIDU101_3 -------...-----A---------------A-------------------------T------------------G--------A-A-----------------------------A-----------C-----------------C-----T--------A------- 7414
C.MW.93.93MW_965 -------...-----A---------------A-------------G-----------T------------------G--------A-A-------------G---------------A-----G-----C-----------------C---------------------- 7230
C.SN.90.90SE_364 ------A...-----A---------------A-------------------------T------------A-----G--------A-A-----------------------------A-----------C------------A----C---------------------- 7200
C.SO.89.89SM_145 ------A...-----A---------------A-------------G-----------T----------G-------G--T-----A-A------------G----------------A-----------C-----------------C--------------A------- 7281
C.TZ.02.CO178 -------...--C--A-----------------------------G-----------T------------------G--T-----A-A-----------------------------A-----------C-----------------C---------------------- 7259
C.UY.01.TRA3011 -----A-...-----A---------------A-------------G-----------T--------T---A-----G--------A-A-----------------------------A--------------------G--------C---------------------- 7232
C.YE.02.02YE511 -------...-----A-----------------------------G-----------T------------A-----G--------A-A-----------------------------A-----------C-----------------C--------------A------- 7247
C.ZA.04.04ZASK164B1 -------...-----A------------------C----------------------T------------A----------------A-----------------------------A---T-G-----C-----------------C--------CT------------ 7505
C.ZA.05.05ZAPSK240B1 -------...-----A---------------A-------------G-----------T------------------G--------A-A------------C----------------A-----G-----C-----------------C---------------------- 7431
C.ZM.02.02ZM108 ------A...-----A---------------A-------------G-----------T------------------G--------A-A--------------A--------------A-----------C-----------------C---------------------- 8092
D.CD.83.ELI -------...-----A-----------------------------G-----------T--A-------------------------------------------------------A--A----------------------A-------------AA------C---- 7591
D.CM.01.01CM_0009BBY ------A...-----A-----------A-----------------------------T--A--------------------------------------G-----------------A--A----------------------A---C---------AA------C---- 7280
D.KE.01.NKU3006 -----A-...-----A---------------A-------------------------T--A-------------------------------------------------------A--A----------------------A---C---------AA----------- 7298
D.KR.04.04KBH8 -------...-----A-----------------------------G-----------T--A---------------G--------C-A-----------------------------A--A------------------------------------A------------ 7968
D.TD.99.MN011 -------...-----A-----------------------------G-----------T--A---------------C----------------------------------------A--A----------------------A-------------A------------ 7245
D.TZ.01.A280 -------...-----A---------------T-------------------------T--A---------------G----------A-----C-----------------------A--A----------------------A---------------------C---- 7243
D.UG.99.99UGK09259 ---A-T-...-----A-----------------------------------------T--A---------------G--------C-------C-----------------------A--A----------------------A-------------AA------C---- 7256
D.YE.01.01YE386 -------...-----A--------------C------------------------G-T--A---------------G----------------C-----------------------A--A----------------------A-------------AA------C---- 7211
D.YE.02.02YE516 -------...-----A-----------A-----------------G-----------T--A------------------------------CA------------------------A--A----------------------A-------------AG---A--C---- 7235
D.ZA.90.R1 ------A...-----A-----------------------------------------T--A---------------G------------------------G---------------A--A----------------------A-------------A-------C---G 7409
F1.AR.02.ARE933 ------A...--------A------------A-------G-----------------T--A---------------C------------------------G---------------A---C-------C-----------------C---------A-------C---- 7325
F1.BE.93.VI850 -------...--------A--------------------G-----------------T--A---------------G----------------------------------------A-----C-----C-----------------C---------AA------C---- 7332
F1.BR.01.01BR125 ------AGG.A-------A--------------------G-----------------T--A---------------C----------------------------------------A---C-------C-----------------C---------AA------C---- 7491
F1.ES.x.P1146 -----A-...--------A--A---------A-------G-----------------T--A------------C--G--------C-------------------------------A-----------C-A---------------C---------AA---C--C---- 7362
F1.FI.93.FIN9363 ----CA-...--------A------------A-------G-----------------T--A---------------G----------------------------------------A---C-------A-----------------C---------AA------C---- 7351
F2.CM.02.02CM_0016BBY -----A-...-----A--A--------------------G-----------------T--A-----------G--------------------------------C-----------A-----------C-----------G-----C---------AA------C---- 7214
F2.CM.95.MP257 -----AA...--------A--------------------G-----------------T--A-----------G--------------------------------C-----------A-----------C-----------------C---C-A---A------------ 7248
F2.CM.97.CM53657 -----A-...-----A--A--------------------G-----------------T--A-----------G--------------------------------C-----------A-----------A-----------------C---------AA------C---- 7226
G.BE.96.DRCBL ------A...-----A--------G--------------------------------T--A-------G-------G--------C-------------------------------A-----------C-----------------C---------AA------C---- 7994
G.CM.01.01CM_4049HAN -------...-----A--------G--------------------------------T--A---------------G------------------T---------------------A-----------C-----------------C---------AA------C---- 7244
G.CU.99.Cu74 -------...-----A--------G--------------------------------T--A---------------G----------A-----------------C-----------A---C-------C-----------------C---------AA------C---- 7674
G.ES.99.X138 -------...-----A--------G--------------------------------T--A---------------G-----------------------G----------------A-----------A----------GG-----C---------AA------C---- 7490
G.GH.03.03GH175G -------...-----A---------------A-------------------------T--A---------------G----A-------------------G---------------A-----------C-----------------C---------AA---A--C---- 8101
G.KE.93.HH8793_12_1 ------A...-----A--A-----G--------------G-----------------T--A---------------G--------C-A-----------G-G---------------A-----------C-----------------C---------AA------C---- 7425
G.NG.01.01NGPL0669 -------...-----A-----------A-----------------------------T------------------G-----------------------C----------------A--A--------C-----------------C---------A-------C---- 7248
G.PT.x.PT2695 -------...-----A--------G--------------------------------T--A---------------G----------------------------------------A-----------C----------GG-----C--T------AA------C---- 8066
G.SE.93.SE6165 -------...-----A--------G--------------------------------T--A---------------C----------A-----------------------------A-----------C-----------------C---------AA------C---- 7473
H.BE.93.VI991 ------A...-----AC--------------A-------G-----G--------T--T--A--A------------G----A-----------------------------------------------C-----------------C---------AA----------- 7454
H.BE.93.VI997 ---A--A...-----AC-------------CA-------G-----G--------TG-T--A---------------G----A-----------------------------------------------------------------C---------AA----------- 7369
H.CF.90.056 ------A...-----AC------G-------A-------G-----G-----------T--A---------------G----C-------------------G---------------------------C-----------------C---------AA----------- 7374
J.CD.97.J_97DC_KTB147 -----A-...-----A---------------------AG------T-----------T---------------------T-------------------------------------A-----------C-----------------C---------AA------C---- 7258
J.SE.93.SE7887 -----A-...-----A---------------------A-------T-----------T------------------G----------------------------------A-----A-----------C-----------------C---------AA------C---- 7362
K.CD.97.EQTB11C ------A...-----A--A--------A--AA----A--G--T--G-----------T--A-------G-------G----------A--G----------G---------------A--A--------C--------------T--C---------AA------C---- 7238
K.CM.96.MP535 ------A...-----A--A--------A--------A--G-----G-----------T--A-------G------------------A-----------------------------A-----------C-----------------C---------AA------C---- 7215
```

```
B.FR.83.HXB2                 GCTGAGG...GCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGA 8055
Env                          __L__R__.__A__I__E__A__Q__Q__H__L__L__Q__L__T__V__W__G__I__K__Q__L__Q__A__R__I__L__A__V__E__R__Y__L__K__D__Q__Q__L__L__G__I__W__G__C__S__G__K__L__I__C__T__T__A__V__P__W__
01_AE.CF.90.90CF11697        -------...-----A--------G--------------------------------T--A---------------G-------------------------A------A--T----A--AC-------C------------A----C---------AA------C---- 8007
01_AE.CN.05.FJ051            ------A...-----A--------G--------------------------------T--A---------------G-------------------------A------A--T----A--AC--C----C------------A----C-----------------A---- 8037
01_AE.CN.06.FJ054            ----CA-...-----A-----------------------------------------T--A---------------G-------------------------A------A--T----A--AC-------C------------A----C---------AA------C---- 8064
01_AE.HK.x.HK001             -------...-----A---------------A-------------------------T--A---------------G--------------------------------AG-T----A--AC-------C--------G--GA----C-----------------C---- 7416
01_AE.TH.01.01TH_R2184       -------...-----A--------G--------------------------------T--A---------------G----------------------G---------A--T----A--AC-------C------------A----C-----------------C---- 7258
01_AE.TH.02.OUR769I          -------...-----A--------G--------------------------------T--A---------------G--------------------------------A--T----A--AC-------C------------A----C-----------------C---- 7239
01_AE.TH.90.CM240            -------...-----A--------G--------------------------------T--A---------------G--------------------------------A--T----A--AC-------C------------A----C-----------------C---- 7608
01_AE.US.00.00US_MSC1164     -------...-----A--------G--------------------------------T--A---------------G--------------------------------A--T----A--AC-------A------------A----C---------AA------C---- 7251
02_AG.CM.02.02CM_4082STN     -----AA...-----A-----T-----A---------A-------G-----------T--A---------------G--------C-----G--------C----------------A--A--------C-----------------C--------------A------- 7233
02_AG.EC.x.ECU41             -----A-...-----A-----T-----A---A-----AG------G-----------T--A---------------G--------C----------------------------------A--------C-----------------A--T------AA---A--C---- 7117
02_AG.GH.03.GHNJ196          -------...-----AA----T-----AG-A------AG------G-----------T--A---------------G--------C----GG---------G------G--------A--A--------A----------G------C---------AA---A--C---- 8102
02_AG.NG.01.PL0710           ------A...-----A-----T-----A---------A-------G-----------T--A---------------G--------C---------------G---------------A--A--------C-----------------C---------A----A------- 7203
02_AG.NG.x.IBNG              -----AA...-----A-----T-----A---------A-------G-----------T--A---------------G--------C---------------GA--------------A--A--------C-----------------C---------A----A--C---- 7561
02_AG.SE.94.SE7812           -----A-...-----A-----T-----A---------AG------G-----------T--A---------------G--------C-A---GC------------------------A--A--------C-----------------C---------A----A--C---- 7445
02_AG.SN.98.MP1211           -----A-...-----A-----T-----A---T-----A-------G-----------T--A--------A------G--------C-------------------------------A--A--------C-----------------C---------A----A------- 7257
03_AB.RU.97.KAL153_2         -------...-----------------A--------------------------------------------G---G-------------G------------------------------------------------------------------------------- 7232
04_cpx.CY.94.CY032           -------...-----A-----T-----A---------AG------G-----------T--A---------------G--------C-------C-----------------------A--A--------C-----------------C---------AA---A------- 7425
05_DF.BE.x.VI1310            ----CA-...--------A--------------------G--------------T--T--A---------------G----------------------------------------A---C-------CA------------A---C---T-----AG------C---- 7461
06_cpx.AU.96.BFP90           ------A...-----A--------G--------------------------------T--A--------------------------------------------------------A--A--------C-----------------C--TC-----AA-----TC---- 8103
08_BC.CN.97.97CNGX_6F        ------A...-----A---------------A-------------G-----------T--------A---A-----G--------A-A-----------------------------A-----------C--------------T--C-----T--------A------- 7234
09_cpx.GH.96.96GH2911        ----C--...-----A-----T--------------AA-------G-----------T--A---------------G----------------G-----------------------A--A--------C-----------G-----C---------AA------C---- 7227
10_CD.TZ.96.96TZ_BF061       -------...-----A-----------------------------------------T--A---------------G-------------G--C-----------------------A--A--------A----------G--A-------------AA------C---- 7407
11_cpx.GR.x.GR17             -----A-...-----C----TA-----A---------A-------T-----------T--A---------------G----------A-----------------------------A-----------C-----------------C---------AA------C---- 7322
12_BF.AR.99.ARMA159          ----CA-...--------A--------------------G-----------------T--A---------------G----------------------------------------A---C-------C-----------------C---G-----AA------C---- 8020
13_cpx.CM.96.1849            -----T-...-----A-----T---------------A-------G-----------T--A--------A---------------A-A-----------------------------A--A--------A-----------------C---------AA----------- 7480
14_BG.DE.01.9196_01          -------...---------------------A-------------------------A------------------G--------------------------------------------------------------------------------A------------ 7560
15_01B.TH.99.99TH_MU2079     -------...-----------A------------------------------------------------------G----------------------------------------------------------------------C---------------------- 7427
16_A2D.KR.97.97KR004         -----A-...-----A-----T---------A-----AG------G-----------T--A---------------G-----------------------C-------------------A--------C-----------------C--------CTT------C---- 7421
17_BF.PE.02.PE02_PCR0155     -------...--------A---------------C----G-----------------T--A-----------C---G--------------------------------------------------------------------------------------------- 7224
18_cpx.CU.99.CU76            ---A---...-----A-----T---------A-----A-------G-----------T--A---------------G----A-----------------------------------A--A--------C-----------------C---C-------G---------- 7363
19_cpx.CU.99.CU7             -----A-...-----A-----T-----A--G------AG------C-----------T--A---------------G--------C----------T--------------------A--A--------C-----------------C---------A-------C---- 7121
20_BG.CU.99.Cu103            -------...-----A--------G--------------------------------T--A---------------G----------A-----------------C-----------A---C-------C-----------------C---------AA------C---- 7489
21_A2D.KE.91.KNH1254         -------...-----A-----------A-----------------G-----------T--A---------------G-TT-------------------------------------A--A----------------------A------TC-----AA----------- 7277
22_01A1.CM.01.01CM_0001BBY   -------...-----A-----T--------CA-----A-------------------T--A---------------G--------------------------------------T-A--A--------C-----------------C--------CA-------C---- 7222
23_BG.CU.03.CB118            -------...-----A--------G--------------------------------T--A---------------G--------C-A-----------------C-----------A---C-------C-----------------C---------AA----------- 7483
24_BG.CU.03.CB378            -------...-----A--------G--------------------------------T--A---------------G----------A-------T---------C-----------A---C-------C-----------------C---------AA------C---- 7498
24_BG.CU.03.CB471            -------...-----A-----------------------------------------T--A---------------G----------A-----------------C-----------A---C-------C-----------------C---------AA------C---- 7501
25_cpx.CM.01.101BA           -------...-----A--------G--------------------------------T--A---------------G----------------------------------------A-----------C-----------------C---------AA----------- 7305
27_cpx.CD.97.97CDKTB49       -------...-----A---------------A-------G--T--------------T--A---------------G-------------------------A-----G---T----A--A--------C------------A----C---------------------- 7259
28_BF.BR.99.BREPM12609       ------A...--------------G-----------------------------------------------G--------------A-----------------------------A-----------------------------C---------AA----------- 7456
29_BF.BR.02.BREPM119         -----A-...-----------------A---A--------------------------------------------G--------A-A--------------------------------A--------------------------C---------A------------ 7207
31_BC.BR.02.110PA            ------A...-----A---------------A-------------G-----------T--------T---A-----G--------A-A-----------------------------A--A--------C------------A----C---------------------- 7548
32_06A1.EE.01.EE0369         ------A...-----A---------------T-------------------------T--A---------------G-------GT----R------G-------------------A--A--------A---A-------------C---C-----AA------C---- 7681
33_01B.MY.05.05MYKL007_1     -------...-----A--------G--------------------------------T--A---------------G--------C--------------C-A------A--T----A--AC-------C------------AC---C---------A-------C---- 7437
34_01B.TH.99.OUR2478P        -------...-----A--------G--------------------------------T--A---------------------------------------------------T----A--AC-------C------------A----C-----------------C---- 7221
35_AD.AF.05.05AF095          -------...-----A-----T---------A-----A-------G-----------T--A---------------G----------------------------------------A--A--------C-----------------C---------A-------C---- 7221
36_cpx.CM.00.00CMNYU830      -------...-----A-----T-----A---------A-------G-----------T--A---------------G--------C---------------G---------------A--A--------C-----------------C---------AA------C---- 7252
37_cpx.CM.00.00CMNYU926      ---A---...-----A-----T-----A---------A-------G-----------T--A---------------G----A---C-------------------------------A--A--------C-----------------C---CA----A-C-----A---- 7237
38_BF1.UY.05.UY05_4752       -------...--------A------------A-------G-----------------T--A---------------G-------------------------------------------A--------------------------C---------AA---C-G----- 7466
39_BF.BR.03.03BRRJ103        -------...--------A------------T-------G-----------------T--A---------------G----------------------------------------A-----------C-----------G-----C---------AA------A---- 7536
40_BF.BR.04.04BRRJ115        -----A-...--------A--------A-----------G-----------------T--A---------------G----A-----A------------C----------------A---C-------C-----------------C---------AA----------- 7612
42_BF.LU.03.luBF_05_03       -----A-...--------------------C---------------------------------------------G-----------------------C----------------A---C-------C-----------G-----C---------AA----------- 7517
43_02G.SA.03.J11223          -------...-----A--------G--------------------------------T--A---------------G----------A-----------------------------A-----------C-----------------C---------AA----------- 7538
U.CD.83.83CD003              -------...-----A-----------------------------------------T--A---------------G-----------------------G--AG------------A---C-------C-----------------C---------A-------C---- 7531
U.CD.90.90CD121E12           -------...-----A-----------------------------------------T--A---------------G-----------------------G---G------------A---C-------C-----------------C---------A-------C---- 7226
U.GR.99.GR303                ----CA-...--C--A-----T---------------AG------G-----------T--A---------------G-------------------------A--C-----AA----A-AA-------C------------------C---------A-------C---- 7339
U.NL.95.U_NL_95_H10986_D1    -------...-----A--A--------------------G--------------------A------------------------A-------------------------------A-----------C-----------------C---------AA------C---- 7961
CPZ.CD.90.ANT                ----CAA...--C--A---A-A--G--A------C-A--G---T-G--------AG-A--A--A--------------G--T--A--C--G-AG-------GA--------A--AT--A-CC-C--------TG---AC--GG-G-CC--TCA----A-G---------- 7487
CPZ.CM.01.SIVcpzCAM13        -T----A...--A--A-----C-----------A--------A-----------AG-A--A--A-----A--C---T-A--T-----A------------C-------G---A-T--A--CC-------C-----------GTCT--C--TTAT--CA------------ 7639
CPZ.CM.05.SIVcpzEK505        ------A...--A--A-----------A--CT--C--------T----T--------T--A--------A------G----T---A-A-----------G-G---C--G--AA----A--CC-A-----C-----A------TCTG-G--TTA---CAA----------- 7538
CPZ.CM.05.SIVcpzLB7          ------A...-----A--------G--A--CT--C--------T----------------A--------A------G----T---A-C-----------------------A-----A--A--A-----------A-----------C---------T-------A---- 7566
CPZ.CM.05.SIVcpzMB66         ------A...--C--A---------------------------T-------------T--A-----T------C-GG-T--T--------G------------------------T-A--AC-A-----C---------------------------T------------ 7679
CPZ.CM.05.SIVcpzMT145        ---A--A...--A--A--A--T-----A-----------G--TT-T--------A-----A-----------G---G-T--T-----------G-------G------G--A-----A--TC-A-----C--TA-A------ACT-----TC----------A-GC---- 7564
CPZ.GA.88.GAB1               -T---AA...--A--A-----T--------C--A--------AT--A-------AG-A--A--A--A--A--C---T-G--T-----A-----G-----GC-------G---A-T-----CC-G-----C-----------GGCTG-C--TTAT--CA-G---------- 8092
CPZ.TZ.00.TAN1               ---TCA-...--C--A--A--------A--CT-------G---T-T--A--------T--------------C-----G--T--A--A--G-A----A---GA--C-----------AA-CC-C-----A--TG--AAC---T-GG-G--TCA--G-AG------A---- 7681
CPZ.US.85.CPZUS              ---A--A...--A--A-----C---------T-A--------AT----T--------T--A-----T--A--C---G-A--T-----------------T--------G---A----A--TC-G-----C-----A------ACA-----TTAT--CA------------ 8060
CPZ.CM.x.SIVcpzDP943         ------A...--A--A---------------T-----------T----A--------T--A--------A------G----T-----------------T--A-----G--AA----A--TC-G-----C-----A------TC---C--TTAT---A------------ 7618
CPZ.CM.x.SIVcpzMB897         -----AA...-----A---------------------------T-------------T--A-----T-----GC--G-T--T-----------G--T--------------A-----A--AC-A-----C-----------------------T---AA----------- 7563
N.CM.02.DJO0131              -G----A...--A--A-----------A---T-----------T--A----------T--A-------GT----A-G----T---A-A-----------T-G------G--AA----AA-TC-A-----C-----A------ACT--A---TAT--CA------------ 7479
N.CM.04.04CM_1015_04         -T----A...--A--A-----------A--AT-----------T--A----------T--A--------A----A-G----T---A-A-----------T-G------G--AA----AA-TC-A-----C-----A------ACT--A---TAT--CA------------ 7457
N.CM.04.04CM_1131_03         -T----A...--AW-A-----R-----A--MT-----------T--A----------T--A-------RA----A-G----T---A-A--R--------T-G------G--AA----AR-TC-A-----C-----A------ACT--A---TAT--CA------------ 7522
N.CM.95.YBF30                -T----A...--A--A-----------A---T-----------T--A----------T--A-------------A-G----T---A-A-----------T-G------G--AA----AA-TC-A-----C-----A------ACA--A---TAT--CA------------ 7603
N.CM.97.YBF106               -T----A...--A--A-----------A---T-----------T--A----------T--A-------GA----A-G----T--GA-A-----------T-GA-----G--AA----AA-TC-A-----C-----A------ACT--A---TAT--CA------------ 7558
O.BE.87.ANT70                ---A--A...--A--AC----C--G-----AT--C--AGG--AT-TR-A-----T----GA--A----GA--TC-CC-G--A--CT-A----CC-TA---C--A----G--A-----AA-CC-A-----C--TAAA-----G--AG-C---TA---AT-A--AAAA---- 8108
O.CM.91.MVP5180              ------A...--G--AC----C--G--A--CT--C--AGGT-AT-T--A-----T--T-GA--A----GA--TC-CC-G-AA--CT-A----CCCTTA--C--A----G--A-G---AAACC-A-----C--TAAA--------A--C--TTA---AT-A--AAAA---- 8125
O.CM.96.96CMABB637           ---A--A...--A--AC----C--G--A--GT--C--AGGT-AT-T--A-----T--T-GA--A----GA--TC-CC-G--A--CT-A----CC-TTA--C--A----G--A-----AAACC-C-----C--TAAR----GG--A------TAT--AT-A--AAAA---- 7598
O.CM.98.98CMA104             ------A...--A--AC----T--G--T--A---C--A-G--AT-T--A-----T----GA--A----GA--TC-CC-G--A--CT-A----CC-TTA--C--A----G--A-----AAACC-A-----C--TAAA-----T--A--C---TA---AT-A--AAAA---- 7603
O.CM.99.99CMU4122            ---A--A...--A--AC----C--G-----AT--C--AGG--AT-T--A-----T----GA--A----GA--TC-CC-G-AA--CT-A----CC-TAA--C--A----G--A-----AAACC-A--------TAAG----GG--A--C--TTA---AT-A--AAAA---- 7575
O.FR.92.VAU                  ---A--A...--A--AC----C--G--A--CT--C--AGG--AT-T--A-----T--T-GA--A----GA--TC-CC-G--A--CT-A----CC-TTA--C--A----G--A----ATAACC-G-----C---AAGAAT-G---A--C---TA---AT-A--AAAG---- 7671
O.SN.99.SEMP1299             ---A--A...--A--AC----C--G-----AT--C--AGG--AT-T--A-----T----GA--A----GA--TC-CC-G--A--CT-A----CC-TAA--C--A-C--G--A-----AAACC-A-----C--TAAG-----G--A--C---TA---AT-A--AAAA---- 8158
O.US.99.99USTWLA             ------A...--A--AC----C--G--A------C--AGGT-AT-T--A-----T--T-GA--A----GA--TC-CC-G-AA--CT-A----CCCTTA-GC--A----G--A-----AAACC-A-----C--TAAA------TCA--C---TA---AT-A--AAAA---- 7585
```

```
                          RRE end┤
B.FR.83.HXB2              ATGCTAGTTGGAGT.............................AATAAATCTCTGGAACAGATTTGGAAT......CACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAG 8189
Env                       N__A__S__W__S__.__.__.__.__.__.__.__.__.__N__K__S__L__E__Q__I__W__N__.__.__H__T__T__W__M__E__W__D__R__E__I__N__N__Y__T__S__L__I__H__S__L__I__E__E__S__Q__N__Q__Q__E__K__
A1.GE.99.99GEMZ011        --T-----------.............................---------TACAGTG---A--G--......A--T---------C-A------AG--G--T----------A-A-----T-TGAGC----------------------G------ 7364
A1.KE.00.KER2008          -CT-C---------.............................-----G----A-AGTG---A--G-C......A--T-------C--C-A-----T-----------------T---GA-A-----T-TAGAC--C-------T-------------G------ 7395
A1.KE.00.KNH1144          -CT-----------.............................----------A--TG-A--A----C......A--T-------C--C-A-----T-A--------G-----------A-C-----T-TAGTC----------------------G------ 7421
A1.KE.00.KSM4024          -CT-----------.............................-----------AA-TG---A--G--......A--T-------C--C-A-----T-AG-------G----------CAAA-----T-TAATC--------A-------G--------G------ 7318
A1.KE.00.MSA4069          -CT-----------.............................---------TATA-TG-A--A--G--......A--T-------C---A-----G---------G----------GATG-----T-TAA-C--C------C------G--------G------ 7397
A1.KE.00.NKU3005          -CT-----------.............................---------TACA-TG---A--G--......A--T-------CA-C-A-----A---------G-A---------GA--AC---T-TAA-C-G-------------------------G------ 7400
A1.RU.00.RU00051          -CT-----------.............................----------A-AGTG-T--A--G--......A-TCT----------C-A----------------C-------T---GA-A-----T-TGATC--------A-----------------G------ 7463
A1.RW.93.93RW_024         --A-----------..........................AAA------------AGTG-A--A--G--......A--T-------C--C-A-----T-A--------G-----------A-C-----T-TAA-C----------------G--------G------ 7394
A1.SE.95.SE8891           -CT-----------.............................G--GT---C-A-AGTG---A--G-G......A--T-------C--C-A-----T-A--------G----------CAGA-----T-TAGTC-------------A--G--------G------ 7354
A1.SE.95.UGSE8131         -CT---CC------.............................---------TACACT--A--A--G--......A--T-------C--C-A-----T-A--------G----------GA-A-----T-TCAAC-----------------T-------G------ 7567
A1.TZ.01.A173             --T-----------.............................--------C-AT--TG-C--A-------......A--T-------C--C-A-----T-A--G---G----------CAGC-----T-TAATC-------------A--G--------G------ 7398
A1.UA.01.01UADN139        -CT---A-------.............................-G--------AT--TG---A-------......A--T----------C-A------AG--G--T-----------A-A-----T-TGGTC--------A-----------------G------ 7385
A1.UG.92.92UG037          -CT-----------.............................-----------A--TG-A--A--G-A......A--T-------C--C-A-----T-A--------G-----------T-AA----T-TGAGC----------------G-T------G---G- 8199
A1.UG.99.99UGA07072       -CT-----------.............................--------C-A-AGTG---A--G-C......A--T-------C--C-A-----T-A--------G----------CAGA-----T-TAATC----------------G-T------GA----- 7395
A2.CD.97.97CDKTB48        -CT-----------.............................-----GA--A--G-----------A......A--T-------T--C-A-----T-AG-------G--C-------GA-A-----T-TATGC--C----G--------G--------G------ 7499
A2.CY.94.94CY017_41       -CA-----------.............................-----G---A--TG-------G-C......A--T-------T--C-A-----T-A--------G-----------A-A-----T-TAGG---C--------------------------G------ 7561
A.SN.01.DDI579            -CT---A-------.............................-GA--G---A--GG-A--A--GG-......A-T-T-------C--C-A-----T-A--------G----------GA-A-----T-TAGTC-------------A--G--------G------ 7393
A.SN.01.DDJ369            -CT-----------.............................---------TATAGTG-A--A--G--......A-T-T---T---C--C-A-----T-A--------G----------GA-------T-TAATC----------------G--------G------ 7411
A.SN.96.DDJ360            -CT-----------.............................----------A-A-TG-A--A--G-G......A-T-T-------C-AC-A-----T-A--------G----------GA-A-----T-TAATC----------------G--------G------ 7375
A.CD.02.02CD_KTB035       -CTA-----A----.............................----G-----A---G-T-----A-GA......A--T-----A----C-A-A---T-A---G---G--G--------A-AC----T--ATGC----------------GC-A--A--G----G- 8188
A.CD.97.97CD_KCC2         -CT-----------.............................----------A---TG---------C......A--T----------C-A-----T-A---------T---------GAAA-----T--GGGC-C--------------G--A-----G--G--A 8188
A.CD.97.97CD_KTB13        -CT----C------.............................----------A--GTG-A-----G--......A--T----------C-A-----T----------G----------GATA-----T--AAG-------G-------GG-A-----G--G--- 7465
B.AR.04.04AR151516        --------------.............................-----------A-GTG-C------G-G......A--T-----------------T----------G----------GA-------T--AAY--GC------------------A--------- 7398
B.AU.87.MBC925            --------------.............................--------------TA-----------......A--T--------G---------A-A------G----------GA---G---T--A-----------C------G----------------- 8208
B.BO.99.BOL0122           --AG----------.............................-----------AGTA-------G--......A--T-----------------A--------G----------G--------T-TA-T-------------A--G-----A--------- 7390
B.BR.03.BREPM2012         --------------.............................---------T--AGTG-------G--......A--T-----------------A-A------G----------G--------T--------C------C-----------A--------- 7634
B.CA.97.CANB3FULL         --A-----------.............................-----------A-TG-A------G--......A--T---------C------A--------G--G-------ATGAC---T--AGA---------------------------A--------- 7462
B.CN.05.05CNHB_hp3        --------------.............................----------AC--TG-----------......A--T---T-------------A---C-----G-------T---GCAGA---T--A----------C-------G-----A--------- 8205
B.CO.01.PCM001            -------C-----C.............................-----------A--TAT----------......A--T-----------C-----A-A------G-----------G--------T-TA------------------G----A--G------ 7372
B.GB.83.CAM1              --------------.............................--------------TA-----------......A-T-T----------------G--------G------------A-------T--A------------------G-----A--------- 8189
B.GB.86.GB8               ---T----------.............................--------------TA-----------......A--T-----------------A--------G------------A-AC----T--A------------------G-----A--------- 8180
B.GE.03.03GEMZ010         --------------.............................-----------AA-TG-A------G--......A--T---------C------G--------G-T----------G--A-----T--AAT---------AT---A---C---A--------- 7366
B.IT.05.SG1               --A-----------.............................------A--A-A-T---------G--......A--T---------C------A-A------G----------------G---T--A----C---G---------C------------- 7874
B.JP.05.DR6538            --A-----------.............................-------------G-G-C--C---G--......A--T--A------C------A--------G------------A-------T--A----C------------------A--------- 8222
B.KR.05.05CSR3            --T-G-------A-.............................--C-G-AAC---AGT---------G--......A--T---------C------A-AG-----G-----GT------TG-----T--AGG---------A-------G--------------- 8252
B.NL.00.671_00T36         --------------.............................--------------T---------G-G......A--T---------C------G--------G-------------T-----T--A-TC----------C-----G--A--A--------- 7750
B.RU.04.04RU128005        --------------.............................-----------A-TG--------G--......A--T---T-----C------R--------G-----------R-AC----T--AA----C--------------------------------- 7661
B.TH.00.00TH_C3198        --AA----------..........................AAT---------TAT--G------------......A-T-T---T-----C------A-C------G-----------AAGA----T--A-----------C-------G-----A--------- 7390
B.UA.01.01UAKV167         ------C-------.............................------A----A-G---------G--......A--T-----------------G---------------------GA-A-----T--AAT---------C----A--------A--------- 7453
B.US.04.ES10_53           --AT----------.............................------AA---ATCT---------G--......A--T-----------------A--------G-G--------CAAAC----T--AA-------------------G--------------- 8167
B.US.99.PRB959_03         --AG----------.............................----G-AA----ACTGC-------G--......A--T---------C-A----A--------G-----------A------TT--AG----C------------------A--------- 7387
C.AR.01.ARG4006           -CT-----------.............................-----------A--TA-T------G--......A--T---A------C-------T-----------T--------GA-AC----T--AGG--GC----------------------G--GC-A 7380
C.BR.04.04BR013           -CT-----------.............................---------AAA---GCT------G--......A--T---------C-------T---------G-T--------G--AC----T--ATG--GC----------------------G------ 7680
C.BW.00.00BW07621         -C------------.............................---------GAA-G-G-T----------......A--T---------C-------T----------GT--C-------A-AC----T--AGG--GC----A-C-------T------G---C-A 7524
C.CN.98.YNRL9840          -CT-C---------.............................G-C-------AA--TG-T-----GG-......A--T---------C-------T----------GT---------GA-AC----T--AGG--GC----------------------G---G-A 7390
C.ET.02.02ET_288          -CT-----------.............................-----G----AA--GG--------G-A......A--T---------C-------T--------C-G-----T-------AC----T--AGG--GC------TC--A------------G--G--A 7394
C.GE.03.03GEMZ033         -CT-C-C-------.............................----------A-AC-G-T------G-A......A--T---------C-------T----------GT---------G-AAC----T--CAA--GC----G--TG------G------G--G--A 7396
C.IL.99.99ET7             -CT-----------.............................----------AA--GG-T-------G-......A--T---T-----C-------T----------GT--G-------A-AC----T--GGG--GC------TC----------------G---G-A 7355
C.IN.99.01IN565_10        -CT-C---------.............................--C--------AAC-G-T--C---G-C......A--T---------C-A-----T----------GT---------GA-A-----T--AGG--GC-------C----------------G--G--A 7566
C.KE.00.KER2010           -CT-C---------.............................----------AAAC-G-T--C-------......A--CT---------C-------T-A--------T---------GA-AC----T--AGG---C-------------G-G------G----GA 7378
C.MM.99.mIDU101_3         -CT-C---------.............................--C-------AA---G---------C......A--T-----------------T-A--------T---------GA-AC----T--AGG--GC-------C---------------G----GA 7548
C.MW.93.93MW_965          -CT-----------.............................------A--TACA-TG-T-----G--......A--T---------C-------T----------GT---------GA-AC----T--AGG--GC-------C---------------G----GA 7364
C.SN.90.90SE_364          -CA-----------.............................-----------A--GG-T-----G--......A--T---T-----C-------T----------T-------------A-----T--AG---GC----------------------G--G--A 7334
C.SO.89.89SM_145          -CT-C---------.............................----R-A--AA---G-A------G-A......A--CT---------C-A-----T-A--------GT---------GA-AC----T--AGG--GC------TC----------------G-----A 7415
C.TZ.02.CO178             -CT-C---------.............................----------AAAC-G-A-----GG-......A--T---------C-------T----------GT-----------TGCC---T--AGA--GC-------C-------T------G--GG-A 7393
C.UY.01.TRA3011           -CT-C---------.............................---------GAA--GG-T----------......A--CT---------C-------T-A--------GT-------------AC---T--AGGC-GC-------C-----G---------G--G--A 7366
C.YE.02.02YE511           -C------------.............................----------AA-GGG-T-----G--......A--T---T-----C-------T----------GT---------G-AA-----T--AAG--GC----------------------G---G-A 7381
C.ZA.04.04ZASK164B1       -CA-C---------.............................---------TATAC-G-T-----G--......A--T---------C-------T---------G-TG--------GAAAC----T--AGG--GC----------------------G---CGA 7639
C.ZA.05.05ZAPSK240B1      --T-----------.............................------AA--AAACTG-------G--......A--T---------C-------T-A--G---GT--------TATACT---T--AGG--GC-------C-------T------G-----A 7565
C.ZM.02.02ZM108           -CT----C------.............................----G-----AA--GG-------G-C......A--T---------C-A-----T-A---G--GT-----T--GA-AC----T--AAG--GC------CC----------------G--GC-A 8226
D.CD.83.ELI               -CT-----------.............................----G-----AA-TG--------C-G......A--T-----------------A---------G-----------G--------T-TAG---------G-------G-C------------- 7725
D.CM.01.01CM_0009BBY      --T-----------.............................-----------A--TG-T-----G-G......A--T---------C-A-----A-A-----G-G---------G-TG-G---T--AAT-----------G-------T-----------A 7414
D.KE.01.NKU3006           -C-----C------.............................-----------A-GGT-T-----G-G......A--T-----------------A----G--G-----------G-G------T--AA----C--------------T------------- 7432
D.KR.04.04KBH8            -CT----------C.............................----G---G-A--TACA-----G--......A--T---------C-------A---------G-----------G--A-----T-TGA----C----------A------------------ 8102
D.TD.99.MN011             -CT-----------.............................----G-----AA-TG-T------C-A......A--T---------C-A-----A-A------AG-A---------G-TG-----T--AA-------------A----T--------G--A 7379
D.TZ.01.A280              -CT-----------.............................-----G--CT-A--TG-C----------......A--T-----------------A-A------G-----------G-AG-----T--AG-----------------GT------------- 7377
D.UG.99.99UGK09259        -CT-----------.............................------A----AA-TA-------G--......A-T-T-----------------A-A------G-----------GAG------T--AA------------------C------------- 7390
D.YE.01.01YE386           -CT----C------.............................------A----A--TG-------GG-......A-A-T-------G-C------A-A------G-----------G-AG-----T--AG------------------C------------A 7345
D.YE.02.02YE516           --T-----------.............................----G------A--TGCT------C-G......A--T--A-------A-----A-A------G-T---------G-T------T--G-----C---------A---T---------C--A 7369
D.ZA.90.R1                -CC--GAA------.............................------A--TAC--TG-------GGG......A--T-----------------A--------G-----------G-A-AC---T--A---------------A--GGT------------- 7543
F1.AR.02.ARE933           -CT-----------.............................------A--A-A-GG-------G-A......A--T-----------------A-AG--G---G------T---ATGA----T--AGG-----------GG---G---------G------ 7459
F1.BE.93.VI850            -CT-----------.............................----------A---G------------......A--T-----------------A-AG--G---G------T---A-A-----T--AAA--G--------G--A-----------G------ 7466
F1.BR.01.01BR125          -CAA----------.............................------AACTACAGTG-------G--......A-A-T-----------------A-A--G---G------T---ATGA----T-TAGG-----------G----G-T------G--G--- 7625
F1.ES.x.P1146             -CT-----------.............................------A-CTAT--C--------G-C......A--T-----------A-----A-A--G--GGT------T---A-AC---CT--CGG---C-----C-----------------G----TA 7496
F1.FI.93.FIN9363          -CT-----------.............................----------A---TG-----------......A--T---------C------A-A--G---G------T---AAAC----T--ATG---------A-------G-G------G---CG- 7485
F2.CM.02.02CM_0016BBY     --T-----------.............................----------A---G-A-----GGA......A--T--G------C------A-A--G-CG-T---------GA-AC----T--AGA----A--G--GG--------------G------ 7348
F2.CM.95.MP257            --CT----------.............................----------AA--TG-A-----GGA......A--T--A-------A-----A-AG--G--GG----------GA-AC----T--AGA---------AGTG-A-----T-----G------ 7382
F2.CM.97.CM53657          --T-----------.............................----------A-A-TG-A-----G-G......A--T---------C------A-A--G-C-GT---------G--AC----T--AAA----A---A-TG-A-----------G------ 7360
G.BE.96.DRCBL             -CA-----------.............................---------TATA-TG-------G-G......A--T--T-----A--A-----A--G-----G-------------TA-CAC---T--AG-C-G------C--------T-----G------ 8128
G.CM.01.01CM_4049HAN      -CA-----------.............................--C------T-T--TG-A----------......A--T--T-----A--A-----A--G--------T---------CAACA----T--AA-C--------AT-----G-G-----G------ 7378
G.CU.99.Cu74              -C------------.............................---------G-T--TG-A-----GGA......A--T-----------A--A-----G--G-------------------GCACA----T--AG-----------------G-------G------ 7808
G.ES.99.X138              --------------.............................---------TATA-TG-A-----G--......A--CT---A---G-AC-A-----A--G----C-G-----T---CAACA----T--AA-C--C-------C-----G-------G------ 7624
G.GH.03.03GH175G          -CA-----------.............................------A--TATA-TG-------G--......A--T--T-----AC-A-----A-AG------------------AACA----T--AG-C----------------G-------G------ 8235
G.KE.93.HH8793_12_1       -C------------.............................------A--TATA-TG-C------G--......A--T---------AC-A--------G-------G---------CAACA----T--AG-C----------------G-------------A 7559
G.NG.01.01NGPL0669        -CA-----------.............................---------TAT--GG-T-----GG-......A--T---------AC-A-----A--G--G---G----------G--CA----T--A-C--------C------------------G------ 7382
G.PT.x.PT2695             --------------.............................---------TATA-T--A-----G--......A--CT---A---G-AC-A-----A--G-------G------T---CAACA----T--A--C--C----G--------G-------G------ 8200
G.SE.93.SE6165            -C-T----------.............................---------TATA-TG-A-----G--......A--T---T-----A--A-----A--G----C-------------TA-CA----T--AG-C--C------------A--G-------G------ 7607
H.BE.93.VI991             -CT-----------.............................--------------TG-A-----G-C......A--T---T---------------A-C-----------------GATGA----T--AGA--GC------TC----------------G------ 7588
H.BE.93.VI997             -CT---C-------.............................-------------CTG-A-----G-C......A--T---T--------------T---C-----G----------GAGG-----T--AGA--GC-----CTC-------C-----G---C-- 7503
H.CF.90.056               -CT-----------.............................---------A-A-AGTG-A--C---G-C......A--T---T-------A-----T-A-C------G----------GAGGA----T--AGG--GC------TC-------C-----G------ 7508
J.CD.97.J_97DC_KTB147     -CT-----------.............................---------C-AT--TG------------......A--T-------G-A--A-----A---------G---------------AA----T--AATC-------G-T---A-----------G------ 7392
J.SE.93.SE7887            --------------.............................---------TAT--GG-C------G-G......A--T---------AC-A-----A--G------------------G-AA-----T--AGTC-----------G-A-----------G-----T 7496
K.CD.97.EQTB11C           -CT----C------.............................----------A-AGTG-------G-G......A--T---------C-------A-AG--G---GT---C-T------AC----T-TAGG-----------G-------T---A--G------ 7372
K.CM.96.MP535             -CT----C------.............................---------TG----G---------C......A--T-----------------A-A---G---GG-------T--GA-AC----T-TAAG--------------A----C------G------ 7349
```

```
                              RRE end┤
B.FR.83.HXB2                  ATGCTAGTTGGAGT..............................AATAAATCTCTGGAACAGATTTGGAAT......CACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAG 8189
Env                           N__A__S__W__S__.__.__.__.__.__.__.__.__.__.__N__K__S__L__E__Q__I__W__N__.__.__H__T__T__W__M__E__W__D__R__E__I__N__N__Y__T__S__L__I__H__S__L__I__E__E__S__Q__N__Q__Q__E__K_
01_AE.CF.90.90CF11697         -CA-C-------------------------------------------TAT-C-G-A------G-C......A---T---A-----A--A-----T-A--------------------A-CA----T-TGAGC--C--AC-A----A--G--------G------ 8141
01_AE.CN.05.FJ051             -CCGC-C-----------------------------------------TAT---G--------G-C......A---T---A-----A--A-----G----------G-------------CA-G--T-TGA-A--C--AC---------G--------G--C-G- 8171
01_AE.CN.06.FJ054             -CT-C-C------..............................------A--TAT--C---------G-C......----T---A----CA--A-----G----------G----------GA-C-----T-TGA-A--CG-AC---------G--------G--T-GA 8198
01_AE.HK.x.HK001              -CAAC-C-------..............................---------TAT---G-A--------C......A---T---A----CAC-A-----G----------G----------GA-CA----T-TG-TA--C--AC------------------G--C--A 7550
01_AE.TH.01.01TH_R2184        -CT-C-C-------..............................---------TAT---G----------C......A---T---A-----A--A-----G----------G-----------A-CA----T-TGAGA--C--AC-A------GG--------G--C-G- 7392
01_AE.TH.02.OUR769I           -CTAC-C-------..............................---------TAT--TG----------C......A---T---A-----A--A-----G----------G-------------CA----T-TGAGA--C--AC---------G--------G--C-G- 7373
01_AE.TH.90.CM240             -CT-C-C-------..............................----G----T-T---G----------C......A---T---A-----A--A-----G----------G-----------A-CA----T-TGAGA--C--AC---------G-----A--G--C-G- 7742
01_AE.US.00.00US_MSC1164      -CT-C-C-------..............................---------TAT---G--------G-C......A---T---A-----A--A-----G----------G-----------A-CA----T-TGAGA--C--AC-A-------G--------G--C-G- 7385
02_AG.CM.02.02CM_4082STN      --T-----------..............................------AA-TATAC-G-C--A---G--......A---T-------C-AC-A-----A-A-------GG----------G--A-----T-TAAAC----------------G--------G------ 7367
02_AG.EC.x.ECU41              -C------------..............................------AA-TATACGG-C--A---G--......A---T-------C--C-A-----T-A--------G----------GAAA-----T-T--TC----------G-----G--------G------ 7251
02_AG.GH.03.GHNJ196           -CT---C-------..............................------A--TATA-TG-C--A---GGG......A---T-------C----A-----T----------G----------GA-A-----T-TAATC---------T---------------G------ 8236
02_AG.NG.01.PL0710            -CT-----------..............................-----GA--TAT--GG-T--A---G-A......A---T-------T--C-A-----T-A--------GT-----T---CAGA----TT-TCAGC----------------G--------G------ 7337
02_AG.NG.x.IBNG               -CT----C------..............................------A--T-TA-TG-C--A---G--......A---T---------AC-A-----G-A--------G----------GA-A-----T-TAATC----------------G----G---G------ 7695
02_AG.SE.94.SE7812            -CT-----------..............................------A--TAT--T--C--A---GG-......A---T-------C--C-A-----T-A--------GT--C------CA-A-----T-TGATC----------------G--------G------ 7579
02_AG.SN.98.MP1211            -CT-----------...........................AAT------A--T-TA-GG-C--A---G--......A---T-------C--C-A-----T-A--------GT---------GAAAA----T-TA-T------------G-------------G------ 7394
03_AB.RU.97.KAL153_2          --A-----------..............................------C----A--TG-------A---......A---T------------------A-----------T---------G-T------T--AAT----A------------G-----A--------- 7366
04_cpx.CY.94.CY032            -CT-----------..............................---------TATA-TG-T--A---G-C......A-T-T-------T--C-A-----T-A-------------------CAAA-----T-TGGG---C----------A--G--------G------ 7559
05_DF.BE.x.VI1310             -CT-----------..............................---------GA--GGG--------G-C......A---T------------------A-A---G----G-------T-----AC----T--AGG---------C-G-----G-T------G------ 7595
06_cpx.AU.96.BFP90            -CA-----------..............................------A--TAT--TG-A------GG-......A---T---------A--A-----T--G------------------CAACA----T-----C--------CT--------CT-----G------ 8237
08_BC.CN.97.97CNGX_6F         -CT-C---------..............................--C-------AAC--G--------G--......A---T----------C-------T-AG-------GT-----------A-AC----T--AGG--GC-------C--------------G----GA 7368
09_cpx.GH.96.96GH2911         -CT-----------..............................----------AAC-TG-A--A---G-G......A-T-T-------C----A-----T----------GT---------CAAC-G---T-TAGGC---A------------G--T-----G------ 7361
10_CD.TZ.96.96TZ_BF061        -CT----C------..............................-----------A-GGG-A------G--......A---T------------------A--G------G-----------G-TC-C---T--AA--------------------T------------- 7541
11_cpx.GR.x.GR17              ---T----------..............................------A--TATA-TG--------G-GAACGAGA---T---T-----AA-A-----A--G------G-----------CAAAC----T--A--C-G-----------A--G--------G------ 7462
12_BF.AR.99.ARMA159           -CT-----------..............................----------A---GG-A---------......A---T----------C-------A-A---G----G-------T--CATGC----T--AGG-----------T-----G--------G------ 8154
13_cpx.CM.96.1849             -CT----C------..............................----------A-AGTG-A--A---G--......A---T-------C--C-A-----T-A----G---G----------GAAA-G---T-TAGGC--C-------------G--------G------ 7614
14_BG.DE.01.9196_01           --------------..............................--------------TG-C---------......A---T------------R-----A-A-------G-----------G--------T--AAT---------C----A--G--------G----G- 7694
15_01B.TH.99.99TH_MU2079      --------------..............................------------A-TG--------G--......A---T---T--------A-----A----------G-----------AAGA----T--A-----------C-------G-----A--G------ 7561
16_A2D.KR.97.97KR004          -CT-----------..............................----GGA-C-A---TG-C--------C......A---T-------T--C-A-----G---------G-----------G--AC---TT--AGGC--C----------A--G--------G------ 7555
17_BF.PE.02.PE02_PCR0155      --------------...........................AAT------A-----------------G--......A---T------------------G-----G----G-------T----AGA----T--AGG--------------A--G--------------- 7361
18_cpx.CU.99.CU76             -CT-----------..............................------A--TACA-TG-C---------......A---T-------T--C-A-----T-A--------G----------CAA------T-TA-AC----------------G--------G------ 7497
19_cpx.CU.99.CU7              -CT----------C..............................----C---CTAT--KG-T--A-----C......A-T-T---A---C-A--A-----T-A----------------C--CAGA-----T-TAATC----A-----------G--T------------ 7255
20_BG.CU.99.Cu103             -CA-----------..............................---------TATA-TG-A------G-C......A---T---------A--A-----G--G------------------CAACA----T--A--C----------------G--------G------ 7623
21_A2D.KE.91.KNH1254          -CT-----------..............................------A---AA-TA-C-------C-G......A---T----------C-------A-A-------G-G---A------G-TAC----T-TAG------------------G-T------------C 7411
22_01A1.CM.01.01CM_0001BBY    -CT-----------..............................--C-------A----G----A------......A---T-------C-AC-A-----T-A--------GT---------GA-A-----T-TAGTC----------------G--------G------ 7356
23_BG.CU.03.CB118             -C------------..............................------A----T---G-A------G-C......A---T---T-----AC-A-----G-G------------------CA-CA----T--AA-C----------------GC-A-----G------ 7617
24_BG.CU.03.CB378             -CT-----------..............................---------GWTA-TG-A--------C......A---T---------A--A-----G--G------G-----------CAACA----T-TA--C-GC-------------G-C------------- 7632
24_BG.CU.03.CB471             -CT-----------..............................---------T-TA-TG-A------G-C......A---T---------A--A-----A--G--------T-----------CACAC---T--A--C-------------A--G--------G------ 7635
25_cpx.CM.01.101BA            --A-----------..............................---------T-T--TG-T---------......A-A-T---------AC-A-----A--G------G-------T---CTACA----T-TGA-C--------C-------G-----A--G------ 7439
27_cpx.CD.97.97CDKTB49        -CT-----------..............................---------TAT--GG-A--------C......A--CT---------A--A-----T-A--------G-------------CA----T-TGAGC--C---T-A------GG--------G------ 7393
28_BF.BR.99.BREPM12609        --T-----------..............................--------------G---------G--......A---T----------C-------A-A-------G-----------GA-A-G---T--A-----------A--G----------A--------- 7590
29_BF.BR.02.BREPM119          --------------..............................--------C--A--T-----------C......A---T--------G---------A---------G----C------G--------T--A-----C-----A-C-----G-----A--------A 7341
31_BC.BR.02.110PA             --T-----------..............................----------AA--GG-T------G-G......A---T----------C-------T-----------T----------GA-AC----T--AGG--GC------TC--------------G-----A 7682
32_06A1.EE.01.EE0369          ---A----------..............................------A--T-T---G-A---------......A---T---------A--A-----A---------------------TCACAT---T--A--C-------------------T-----G------ 7815
33_01B.MY.05.05MYKL007_1      --T-C-C-------..............................----G----TAT---G----------C......A---T---A-----A--A-----G------------------------CA----TTTGAT---C--AT------------------G----G- 7571
34_01B.TH.99.OUR2478P         -CT-C-C------C..............................----G----TAT---G-A--------C......A---T---------AA-A-----G----------G----C-----G--CA----T-TCAGA--C--AC------A--G--------G--C--- 7355
35_AD.AF.05.05AF095           -CA----------C..............................------AAC-A-AGTG----A---G-C......A-A-T-------C--C-A-----T-AG-------G----------GA-AC----T--A-TC--C--------G-A-----------G------ 7355
36_cpx.CM.00.00CMNYU830       -CT-----------..............................------A--TAT--TAGC--A---G-A......A---T-------C--C-A-----T----------G----------CAGA-----T-TGAGC-------------A--G--------G----G- 7386
37_cpx.CM.00.00CMNYU926       -CT----------C..............................-----------CA-TG-T--A---G-A......A---T---T---C--C-A-----T-A-------G-A------T---CAGA-----T-TAATC------------------T------------- 7371
38_BF1.UY.05.UY05_4752        --A---C-------..............................------A-CTAT--GG--------GG-......A---T------------------A-----GG---------------ATGA----T--AGA-------G---------G-T------G------ 7600
39_BF.BR.03.03BRRJ103         --T----C------..............................----------A-C-GG----C---G-G......A---T---A--------------A-----G----G---C---T---ATGA----T--AGG---------AT--------T------G------ 7670
40_BF.BR.04.04BRRJ115         --AAG-A-------..................AGTACTTGGAGG-C------------TG-C------G-C......A---T------------------G--G------G-------------AAC----T-TAA------------------G-----A--G------ 7758
42_BF.LU.03.luBF_05_03        -CT-----------..............................------A--TATC-GG-A------GG-......A---T------------------A-AG--G--C---------T--CAAGA----T-TAGG-----------C-----G--------G------ 7651
43_02G.SA.03.J11223           -CA-----------..............................---------TAT--TG-R------GG-......A---T---------AC-A-----A--G--G-CGG-----------CAGCAC---T-TGATY-------------R-----------G-----A 7672
U.CD.83.83CD003               -CT----C------..............................--------CT---TA-C-------G-C......A-TCT------------------T----------G----------CAAG-----T-TGGG--GC-------C--A-----A-----G------ 7665
U.CD.90.90CD121E12            -CT----C------..............................----G---CT-AA-TG-C------G-C......A-T-T------------------T-A--------G----------CAAA-----T-TGGG-----------C--------A-----G------ 7360
U.GR.99.GR303                 -CT-G---------..............................----------AA...G-AC-AACAT--GGGAAAA---T-------C-A--A-----T-A--------GT--------GCAACA----T-TGAAC----------------G-T------------- 7476
U.NL.95.U_NL_95_H10986_D1     -CT----C------..............................----------A---GG--------G-C......A---T---------T--------A-----G----G---A------G--AC----T-TAGG-----------G--A-----------G------ 8095
CPZ.CD.90.ANT                 --AA-TCC---GTA...AACTTCACGCAAACATGTGCAAAGAAC-GC-GTGA-A-AC--TGT------G-A......A-T-T---A---CAA--A---------TT-G-AC-G----CA---G-ACAG---T-TAATA-CT-AC--AT-G-A--TG-G--A-----G-GA 7648
CPZ.CM.01.SIVcpzCAM13         -CAAA-CA---TCA..............................GGC--G---A--AGTG-C--------C......A-TTTA--A---CAAC-A-------A-TTG--C-CT----------G-GAC---TTTTGGT---T-A-----GG-A---TCA--A-----G--A 7773
CPZ.CM.05.SIVcpzEK505         --A-C-C------C...........................AAC-----C--CTAT--CACA--A---GGA......A-T-T---A---CAAA-T-----TGAGC--G-A-GA-----TT--G-TG-C--TTTTGGGC-GC-A--GC--G-A---G-G--A---AGC-TA 7675
CPZ.CM.05.SIVcpzLB7           --AGA-C-------..............................--C---A-ATATA-TG-A------G-C......A---T---A--------A--------G---G-A-GA---------GAAA-T--TT-TGGG-----A--GC--G-A---G-------G-----T 7700
CPZ.CM.05.SIVcpzMB66          -CA---CC----C-..............................---------TAT--TG-T------T--......A---T---A------C-A-------A----G---G-----------GATG-T--CT-TAATC--T-A---A-GG-A----C------G-----T 7813
CPZ.CM.05.SIVcpzMT145         -CAAG-CA---G-A...........................AAT-T--GTGACTATC-GGTT---------......A-TTAC--A---CAAC-A-----T------G-A--T---------G-AC-C--CT--A-A--GC-A--G--GG-TA-T-CA-----G--G--A 7701
CPZ.GA.88.GAB1                -CAACTC----CC-...........................GGG-GC--T--CACA--TG-C------GGG......A-TCTA--A---CA-C-A-----T-A-TT-G-GTCT--C------G-GAA---TTTTGGTC-GT-A--G---G-A---TCA--A-----G--A 8229
CPZ.TZ.00.TAN1                -CCTC-CC---GC-GAAGATTCTACAAAGTGCAATCACAGTGATGCA--G-ACTAT--CTGT--A-----C......A-TTT---T---CA---A-----TC--TT-G-AG-A--C-CT---G-AACC---T-----C-GT-A--GA--G-A----CA--A--G--G--A 7845
CPZ.US.85.CPZUS               ---A--CC------...........................AAC--CCTC--CTAT--TGCT------GGC......A-TCTA--T---CAA--A---------A--G-A-GA--C--TT--G-TACT--TTTTAGTC-T--A---C--G-A---G-A-----GA-T-CA 8197
CPZ.CM.x.SIVcpzDP943          --AG--C----G--...........................GCT----CC--CTAT--TG-A--C------......A--CTA--A---CAA--C-------AGAG-G-A--A-----TT--G-AG-C--CTTTGA-C-T--A--GC--G-A---G-G--A--RA-C-CA 7755
CPZ.CM.x.SIVcpzMB897          --AA--C-------........................AAAAAATTAGG-A-AAACCT-TCCT-----G--......A---T---A------C-A-----G-A---G---G-T--C------GAAACT--CT-TGAAC-TC-AAC-CG---T--G--T--A--G---GTC 7703
N.CM.02.DJO0131               ---AG-C------C...........................AAT----CC---TAT--TACW--C---GG-......A-TTTA------CA-C-A-----T--GA--G-A-GA--C--TT--G-TG-C--TTTTGAGC-T--A--RA-GG-A---G-A--A--GA-C-CA 7616
N.CM.04.04CM_1015_04          ---AG-C------C...........................AATC---CC--CTAT--TTCA--C---GGA......A-TTTA------CAAC-A-----TGAGA--G-A-GA--C--TT--G-TG-C--TTTTGA-C-T--A--GC-GG-A---G-G-----GA-C-CA 7594
N.CM.04.04CM_1131_03          ---AG-C------C...........................ARTM---CC--YTAT--KRCA--C---GGW......A-TTTA------CARM-A-----TRAGA--G-A-GA--C--TT--G-TRYT--TTTTGAMC-TM-A--GC-GG-A---G-A--A--RA-C-CA 7659
N.CM.95.YBF30                 ---AG-C------C...........................AAC----CC---TAT--TACA--C------......A-TTTA------CAAC-A-----TGAGA--G-A-GA--C--TT--G-TG-C--TTTTGGAC-T--A---C-GG-A---G-A--A--GA-C-CA 7740
N.CM.97.YBF106                ---A--CY-----C...........................AGT----CC--CTAT--TACA--C---RG-......A-TCTC------CAAC-A-----TC-GA--G-A-GG--C--TT--G-TG-C--TTTTGATC-T--A--GC--G-A---G-A--A---A-C-CA 7695
O.BE.87.ANT70                 --AGA-CA----TA..............................---GG-...AAC---AGC------G-C......AC-TTA--A---CA---A-----TC-GC-G--A-G---CATA-GCTC-ACC---T-TGAGGA---AC--A-GG-A---GTA-----G---C-A 8236
O.CM.91.MVP5180               -CA-ATCA---TCA...........................GGA-GAT-TAA-GAT--CAGT------G-C......A--CTT--A---CA-C-A------CA-C-C--A------GTA-GCTC-A-T---T-TGATGA---AC---C-G-A---G----A--G------ 8262
O.CM.96.96CMABB637            -CAGC-CA---G-A..............................---GG-GGAAAT--GTCA------G--......AG-CTA--A---CA-C-------TCA-C-G--AG-----GTA-GCTC-A-C---T-TGAAGA---AC--A-GG-A---G-G--A--G---C-A 7729
O.CM.98.98CMA104              -CAA--CA---G-C...........................GGAG---GGGACT-A-G-A-C------G-C......A---T---A---CA---A-----TCAGC-G--A-G---CATA-GCTC-ACC---T-TGAAGA---AC--A-GG-A---GTA-----G---C-C 7740
O.CM.99.99CMU4122             --AAC-CA---GCA...........................AACGGC-CTAA---AA-TACA------G-C......A--CTA--A---CA-C-A-----TCARC-G--AG---CATA-GCGA-G-T---T-TGGTGA---AC--A--G-A--GG-------G---C-A 7712
O.FR.92.VAU                   --AAA-CA---G-A..............................---GG-GA-AAT---TCA------G--......G-GTTA--A---CA-C-------TCA-C-G--A-----CGTA-GCTC---C---T-TGAAAA---AC----GG-A---G-A--A--G--G--A 7802
O.SN.99.SEMP1299              -CA-G-CA----CA..................AATTGCACAAAT-CC--TAAGT-A--TG-T------G-C......A-GCTA--A---CA-C-A-----TCAGC-G--A-G---CGTA-GTTC-A-C---T-TGAGGA---ACG-A-TG-A---GTA-----G---C-A 8304
O.US.99.99USTWLA              ---AA-CA---G-A..............................------GGAAATCTCTCA------G-C......AG-TTA--A---CA-C-A-----TCA-C-GG-ATC----GTA-GCTCT--C---T-TGA-AA---AC-----G-A---G-A--A--G--GG-A 7713
```

```
B.FR.83.HXB2                AATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATA 8359
Env                         _N__E__Q__E__L__L__E__L__D__K__W__A__S__L__W__N__W__F__N__I__T__N__W__L__W__Y__I__K__L__F__I__M__I__V__G__G__L__V__G__L__R__I__V__F__A__V__L__S__I__V__N__R__V__R__Q__G__Y
A1.GE.99.99GEMZ011          -----------T-------C---G-----G------G--C---------------G----T--------T-A------------A----T--------------------AA----A--------------------G---------A-A--------C----------- 7534
A1.KE.00.KER2008            -----------C-------C---G--C--G--------CC------C-------G--C--T----C-----------C------A----T-----------------T--AA----A--------------------G---G--G--A-A-------------------- 7565
A1.KE.00.KNH1144            -----------C-------C---G-GC--G-------A-C------C-------G-----T---------------------G-A----T--------------------AA----A--------------------G---G--G--A-A--G----------------- 7591
A1.KE.00.KSM4024            --------G----------C---G--C--G------G-CC--------------G----------G------------------A----T--------------------AA----A--------------------G---G-----A-A-------------------- 7488
A1.KE.00.MSA4069            -----------C-------C---G--C--G---------C------G--------G----------------------------A----T--------------------AA----A--------------------G---------A-A--A----------------- 7567
A1.KE.00.NKU3005            -----------C-----A-C---G--C--G---------C------G--------G----T----A------------------A----T--------------------AA----A-----G--------------G------G--A-A-------------------- 7570
A1.RU.00.RU00051            -----G-----T-------C---G-----G---------C--------------G-----T----------A------------C----T--------------------AA----A--------------------G------G--A-A--------C----------- 7633
A1.RW.93.93RW_024           -----------C-----A-C---G-----G-------A-C-T----C-------G-----T-----------------------A-------------------------AA----A--------------------G---A--G--A-A-------------------- 7564
A1.SE.95.SE8891             -----------C-------C---G--C--G-------A-C--------------G-----T----A------------------A----T--------------------AA----A------------A-------GA--------A-A-------------------- 7524
A1.SE.95.UGSE8131           ------A-G--------------G--C--G-------A-C--------------G-----T-C--C------------------A----T--------------------AA----C--------------------G---A--G--A-A--A----------------- 7737
A1.TZ.01.A173               -----------G-------C---G--C--G---------C--------------G-T---T---------------------GGA----T--------------------AA----A---------A-------A--G---------A-A-------------------- 7568
A1.UA.01.01UADN139          -----------T-------C---G---------------C------G-------G-----T--------T-A------------A----T--------------------AA----A--------------------G---------A-A--------C----------- 7555
A1.UG.92.92UG037            ------A----C---C----G--G--C--G---------C--------------G-----T----A------------------A----T------------------C-AA----A--------------------G------G--A-A-------------------- 8369
A1.UG.99.99UGA07072         -----------C-------C---G--C--G-------ACC--------------G-----T----C--R-------------G-A----T--------------------AA----A---------------A----G-----C---A-A-------------------- 7565
A2.CD.97.97CDKTB48          -----------C-------C------C----------A----------------G---------GA------------------A----T---------------------A----A------------A-----A--A-----G----A--A----------------- 7669
A2.CY.94.94CY017_41         -----G-----C-------C------C---------GA--------G------C-----TT--C-C----------------G-A----T---------------------A----A------------------A--A--A--G----A-------------------- 7731
A.SN.01.DDI579              -----------C-------C---G--C------------C------G-------G-----T----G---T-A------------A----T--------------G-----AA----A--------------------G---------A-A-----------------T-- 7563
A.SN.01.DDJ369              -----------C-------C---G--C--G---------C-A------------------T----------A-----C----G-A----T--------------G-----AA----A---------------A--------------A-A------------A------- 7581
A.SN.96.DDJ360              -----------C-------C---G--C----------A-C--------------G-----T----A---T-A------------A----T--------------G-----AA----A--------------------G---------A-A-----------------T-- 7545
A.CD.02.02CD_KTB035         --------G--T-------C---G--CG-G-A-------C---A-----A----G-----T----A-----A-A--------G-A----T--------------------AA----G-----------------A--G------G--A-A------------A------- 8358
A.CD.97.97CD_KCC2           ------TTT--G-----------G--C--G-------A-C--------------G-A---T----------A------------A----T--------------------AA----A-----------------A--G-----------A-------------------- 8358
A.CD.97.97CD_KTB13          -------TTA-G-------C---G--CC-G---------C--------------G-----T----------A----------G-A----T--------------------AA----A-----------------A--G------G----A-------------------- 7635
B.AR.04.04AR151516          -----------C-------C---G-----G-------A----------------G----------C------------------A-------------------------AA--------------------A----------------R-------------------- 7568
B.AU.87.MBC925              -----------------A------------------------------------G-------C---------------------A--------------------------A----C---------------AT----------G------------------------- 8378
B.BO.99.BOL0122             -----------------A------C-----------------------------G----------G-----A------------A----------------------T--------------------C--------G-----------A-------------------- 7560
B.BR.03.BREPM2012           -----------C-------C---G-----G------------------------------T----C------------------A----------------------T---A--------------------A----G------G----A-------------------- 7804
B.CA.97.CANB3FULL           -----------------------------G------------------------------T----C----------------G----------------------------A--------------------------------G------------------------- 7632
B.CN.05.05CNHB_hp3          -----GTT----C----------G-----G--------C--------------------G---C-A--------------T---A--------------------------A--------------A----------------------A-------------------- 8375
B.CO.01.PCM001              ------A----------------G--C--G---------C------CA------G----------C----------------G-C----------A---------A----A---------------A----------------------------------------C-- 7542
B.GB.83.CAM1                ------A----C-----------G----CG------------------------G-----------------------------A--------------A-------T--AA--------------------A--A--------T------------------------- 8359
B.GB.86.GB8                 -----------------------G-----G-------A----------------G----------C------------------A--------------A----------AA--------------A-----------A------------------------G------ 8350
B.GE.03.03GEMZ010           --------T-----------------C--G-----------A------------G-----T----C------------------A----------------------T--AA--------------------A---------------CA-------------------- 7536
B.IT.05.SG1                 -----G---------GA------G-------------A-G--------G-----G--------C-G----------------G-A--------------------T-----A-C---------G---------T---G-C---------------G-----------G-- 8044
B.JP.05.DR6538              ---------------C-A-----------G------------------------G--------CGA------------------A----T---------------------A----------G----------TGA---------------------------------- 8392
B.KR.05.05CSR3              -----------------------G-------------A-C--------------G-----------------------------A--------------------------A----------------C---A------------------------------------- 8422
B.NL.00.671_00T36           -------------------C---G----C--------------------------G----------------------------A----------------------T---A----G---------------AT----------------------------A------- 7920
B.RU.04.04RU128005          -----------------------G------------------------------G----------C------------------A----T-----A--------------------------------------------G----------------------------- 7831
B.TH.00.00TH_C3198          -----------------------G-----C----------------G-------G----------G----------------GGA----------A---------------------------------------------G--G----A-----------------G-- 7560
B.UA.01.01UAKV167           -----------------------G-------------------------------G---------C-----------------GA----------------------T--AA--------------A-----A-----T------------------------------- 7623
B.US.04.ES10_53             ------A----------A------------------------------------G----------C------------------A--------------------------A--------------------A----G-------------------------------- 8337
B.US.99.PRB959_03           --------------G--------G-------------------------------G--------GC---------------------------------------------A-------------------------G-----------A--A----------------- 7557
C.AR.01.ARG4006             ------A----T-----A-C---G--C--G---CA--A-C--------------------T----C------------------A--------------A-----------A--------------A----------G--A----------------------------- 7550
C.BR.04.04BR013             --------G--T-----A-C---G--C--G---CA--A-C--------------G---------G----T--------------A----T---------------------A----C---------A----------G--C----------------------------- 7850
C.BW.00.00BW07621           ------A-G--T---C-A-C---G--C-GT---AA--A-C------G-------G-----T-----------------------A--------------------------A----C---------A----------G--A--------A--------C----------- 7694
C.CN.98.YNRL9840            ------A----T-----A-C---G--C-GT---AA--A-C-A----G-------------------------------------A--------------------------A--------------A----------G--C---G------------------------- 7560
C.ET.02.02ET_288            --------G--T---C---C---G--C------AA-GA-C------G--------G-------C-G------------------A--------------------------A--------------A----------G--C---C-G--A-------------------- 7564
C.GE.03.03GEMZ033           ------A-G--T-----A-C---G--C-GT---AAT---C-------------C-G----T-----------------------A--------------------------A----C---------A----------G--A----------------------------- 7566
C.IL.99.99ET7               ------A--------C-A-C---G--C-GG---AA--A-C--------------G----------------------------GA--------------------------A--------------A----------G--C-------CA-------------------- 7525
C.IN.99.01IN565_10          ------A----T-----A-----G--C-GT---AA--A-C---------------G-------C--------------------A--------------------------A--------------A----------G---------A-A-------------------- 7736
C.KE.00.KER2010             ------A----T-----A-C---G--C-GT---AAC-A---A------------G---------GG-----------------------------A-------T-A-----A--------------A------T---GT-C----------------------------- 7548
C.MM.99.mIDU101_3           ------A----T-----A-C---G--C-GT---AA--A-C-A------------------------------------------A--------------------------A--------------A-------------C---G------------------------- 7718
C.MW.93.93MW_965            ------A----T---C-A-C---G--C-GT---AAC-A-C-------------------------C------------------A--------------------------A--------------A----------G--C--------A-------------------- 7534
C.SN.90.90SE_364            ------A----T---C-A-C---G--C-GT---AAC-A-C-------------------------G-----A------------A--------------------------A--------------A----------G--C--------A-------------------- 7504
C.SO.89.89SM_145            ------A----T-----A-C---G--C-GT----G----C--------------G----------A------------------A----------A---------------A--------------A------T---G--C----------------------------- 7585
C.TZ.02.CO178               ------A----T-----A-C---G--C-GT---AAC-CCC--------------G-----T--G------------------G-A----------A---------------A--------------A----------G--C----------------------------- 7563
C.UY.01.TRA3011             ------A-G--T-----A-C---G-----G---CA--A-C------C-------G-T--------G------------------A--------------------------A--------------A----------G--C--------A-------------------- 7536
C.YE.02.02YE511             ------A----TC--C-A-C--CG--C-GT---AAC-A-C--------------------------------------------A--------------------------A-----G--------A-----A----G---------A-A-------------------- 7551
C.ZA.04.04ZASK164B1         ------A----T-----A-C---G--C-GT---AAC-A-C----------------T---T----C---T--------------A----T---------------------A-------G------A----------G--AG-------A--C----------------- 7809
C.ZA.05.05ZAPSK240B1        ------A----T---C-A-C---G--C-GT---AAC-A-C-A-------------G---------A---T--------------A--------------------------A--------------A----------G--C---C------------------------- 7735
C.ZM.02.02ZM108             ------A----TC--C-A-C---G--C-GT---------C-----TC-------------------------------------A-------------------------AA--A----------------------G-----------A-------------------- 8396
D.CD.83.ELI                 ------A-------G--------G--C--G-------------------------G-------C-A------------------A--------------A-----------A-------------------------G------T----A-------------------- 7895
D.CM.01.01CM_0009BBY        ------A----------A-----G-----G-------------------------G------T--C----------------G-A-------C-----------------AA----C--------------------G--G-----G--A-----------------G-- 7584
D.KE.01.NKU3006             ------------C-T---C----G--C----------------------------G-------C-A---------------------------------------------A-------------------------GT-----T----A-------------------- 7602
D.KR.04.04KBH8              ------A----C--G--------G--C----------------------------G--------GA---T--------------A----T-----A---------------A--------------A----------G------T------------------------- 8272
D.TD.99.MN011               ------A----------A-----G--C--G-------------------------------GTC-A------------------A-------C-----------------AA-------------------------G--------G--A-------------------- 7549
D.TZ.01.A280                ------A-----C----------G--C----------------------------G---------A---------------------------------------------A---------A----------A----GT-----T----A-------------------- 7547
D.UG.99.99UGK09259          ------------C-----C----G--C-----------C----------------G---------A------------------A-------------------------AA-------G--G--------------G------T----A-------------------- 7560
D.YE.01.01YE386             ------------C-----C----G------------------------------G-----T---------------------G------T---------------------A----------G--------------G------T----A--A-A--------------- 7515
D.YE.02.02YE516             ------A----------A-----G--C--G-------------------------G-----GTC-A------------------A----T--C------------A----AA-----C-------------------G--------G--A-------------------- 7539
D.ZA.90.R1                  ------A-------G--------G--C--G---------C-------------------------C-----A------------A--------------------------A-------------------------G------G----A-------------------- 7713
F1.AR.02.ARE933             -----------------A-C---G--C------------C---------------G---------C-----------C-----------------A--------------CA----C--------------------G-----------A------------A------- 7629
F1.BE.93.VI850              -------------------C------C------------C--------------G-----T----C------------------A----------------G---------A----C--------------------------------A------------A------- 7636
F1.BR.01.01BR125            ------------C----A-C---G--C------A-----C------G-------------T---------------------G-A----T---------------------A----C--------------------G-----------A------------A------- 7795
F1.ES.x.P1146               -----------------ATCC--G--C--G---------C------G-------G----------------------C----G-A--------------------------A----C-----G---------------T--------T-A------------A------- 7666
F1.FI.93.FIN9363            --------------G--------G--C--G----AT---C------G-------G-T--------C---T-A------------A--------------------------A----C--------------------G-----------A------------A------- 7655
F2.CM.02.02CM_0016BBY       --------G--C-------C---G--C--G----AC---C------G--------CT--------C---T-------C----G-A-------C---G--------------A----A--------------------G---------A-A-----G-----------T-- 7518
F2.CM.95.MP257              --------G--C-------C---G--C--G----AC-A-C--------------TCT-------GA---T-------C----G-A--------------A-----A-----A----A---------------A----G---------A-A-------------------- 7552
F2.CM.97.CM53657            --------G--C-------C---G--C-------AC-A-C------G--------CT------------T-------C---------------------------------A----A--------------------G---G--G--A-A-----G-------------- 7530
G.BE.96.DRCBL               -----------C-------C---G--CC-G---------C------G--------G---CT----C-----A----------G-A----TG-------------------AA-------------------------G--A--------A-----------------C-- 8298
G.CM.01.01CM_4049HAN        -------------------C---G--C--G----AC---C------G-------G----------A-----A------------A----T-----------------G--AA-------------------------G-----------------GA------------- 7548
G.CU.99.Cu74                -----------C-------C---G--C--G----------------C-------G-----T----C-----A------------A----T--------------------AA--------------------A----G------T--A-A------------A------- 7978
G.ES.99.X138                -----------C-------C---G--C--G------------------------G-T---T--C-C-----A------------A----T-----------------T--AA-------------------------G---------A-A------------A------- 7794
G.GH.03.03GH175G            -------------------C---G--C-------------------G-------G----------A-----A------------A----------------------T--AA--------------------A----G-----------A-----------------G-- 8405
G.KE.93.HH8793_12_1         -----G-----C-----A-C---G--C--T----------------C-------G----------A-----A------------A----T-----------------TC-AA--A------A----A----------G-----------A------------A----G-- 7729
G.NG.01.01NGPL0669          ------------------TC---G--CC-G-------AC---------------G-----T----C-----A------------A----------------------T--AA----G--------------------G-----------A-------------------- 7552
G.PT.x.PT2695               -----------C-----A-C---G--C--G------GA----------------G-----T---GA-----A------------A----T-----------------T--AA-------------------------G---------A-A------------A------- 8370
G.SE.93.SE6165              -----------C-------CC--G--CC-G------------------------GG--------GA-----A------------A----T-----------------T--AA-------------------------G-----------A-------------------- 7777
H.BE.93.VI991               -----------C-----A-C---G--C----------A-C-A-------------GT-------------------------G-A----------------------TA-AA----C--------------------G-----------A-------------------- 7758
H.BE.93.VI997               -----------T-----A-C---G--C-------AC-------------------G---------C------------------A----------A---------CT----A----G---------A----------G-----------AGG------------------ 7673
H.CF.90.056                 --------G--C-----A-C---G--C------------C------C-------G-----T-C---------------------A----------------------T--AA--------------A----------G-----------A-------------------- 7678
J.CD.97.J_97DC_KTB147       -----------C---C---C-C-G--C--G---A-----C------G---------G---T---------------------G-A-------------------------AA--------------A-C--------G---G-------A-------------------- 7562
J.SE.93.SE7887              ------A----C-------C---G--C--G---A---A----------------------T----C-----------C--C---A--------------A-------T--AA--------------A----------G---G-------A-------------------- 7666
K.CD.97.EQTB11C             -----------C-------C---G--C--G---------C--------------G-----T----C-----A------------A--------------------------A--------------------A----G------G----A--------C----------- 7542
K.CM.96.MP535               -----------T-------C---G--C------------C--------------G----------A-----A------------A--------------A-----------A---------------C---------G------G----A--------C----------- 7519
```

```
B.FR.83.HXB2 AATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATA 8359
Env _N__E__Q__E__L__L__E__L__D__K__W__A__S__L__W__N__W__F__N__I__T__N__W__L__W__Y__I__K__L__F__I__M__I__V__G__G__L__V__G__L__R__I__V__F__A__V__L__S__I__V__N__R__V__R__Q__G__Y
01_AE.CF.90.90CF11697 --------G--T--G--A-C---G--C--G-------A-C---------------G----------------------------A-----------G----------T--AA----C--------------------G-----------A-------------------- 8311
01_AE.CN.05.FJ051 ------A-G--T-----A-----GA---C---------CC--------------G----------G-----------C------A----T-----A-----------T--AA--------------A----------G-------------------------------- 8341
01_AE.CN.06.FJ054 ------A-G--T--GC-A-----G---------------C--------------G---------GA------------------A----T--------------G--T--AA--------------A----------G-----------A-------------------- 8368
01_AE.HK.x.HK001 -----GA-G--T--G--------G--CC-----------C---------------G---------G------------------AC---T-----------------T--AA----A---------A-C--------G-------------------------------- 7720
01_AE.TH.01.01TH_R2184 ------A-G--T--G--A-----G--C------------C--------------G-----------------------------A----T--------------------AA--------------A----------G-----------A-------------------- 7562
01_AE.TH.02.OUR769I ------A-G--T--G--A--G--G---------A-----C--------------G-----T----G------------------A----T-----------G-----T--AA-------------------------G-------------------------------- 7543
01_AE.TH.90.CM240 ------A-G--T--G--A-----G---------------C--------------G-----------------------------A----T-----------------T--AA--------------A----------G-----------A-------------------- 7912
01_AE.US.00.00US_MSC1164 ------A-G--T--G--A-----G---------------C--------------------------------------------A----T-----------------T--AA--------------A----------G-----------A-------------------- 7555
02_AG.CM.02.02CM_4082STN -----------C---C---C---G--C--G-------A-C--------------G----------G---T-A------------A----T---------------------A--------C-----------------T--AA----A-A-------------------- 7537
02_AG.EC.x.ECU41 -----------C-------C---G--C--G---------C--------------G----------A------------------A----T-----------G-----T---A-------------------------GT--A-----A-A-------------------- 7421
02_AG.GH.03.GHNJ196 -----------C-------C---G--C--G---------C------G--------G----------------------------A----T--------------------AA-----------------------------A-----A-A-------------------- 8406
02_AG.NG.01.PL0710 ------A----C-------CC--G--C--G---------C--------------G----------------A------------A----T-----------------T--TA--------------------A----GT--AA----A-A-------------------- 7507
02_AG.NG.x.IBNG -----------C--G----C---G--C--G---------C-A------------G----------------A----------G-A----T-----------------T--AA-------------------------G---A-----A-A-------------------- 7865
02_AG.SE.94.SE7812 -----------C-------C---G--C--G---------C-----G---------G----T---G------A----------G-A----T-----A-----------T--AA----C-----------C------------G-----A-A-------------------- 7749
02_AG.SN.98.MP1211 -----------C-------C---G--C--G---------C--------------G----------------A---------TT-A----TC-----G----------T---A-----------------------------G-----A-A-------------------- 7564
03_AB.RU.97.KAL153_2 ------------A------C----------------------------------G-----T----A------------------A-----------------------------------------A------------------------------G------------ 7536
04_cpx.CY.94.CY032 -----G-----C-------CC--G--C--G--------CC---------------G---------A-----A------------A----T---------------------A----C---------A----------G-----------A-------------------- 7729
05_DF.BE.x.VI1310 ------------------TC------CC-G--------CC--------------G----------A-----A------------A--------------------------A----C---------------A----G-----C------------------A------- 7765
06_cpx.AU.96.BFP90 -----------C-------C---G--C--G-------A-C-----TC-------G----------C---T-A-----C------A----T--------------------AA-------------------------GA----------C-------------------- 8407
08_BC.CN.97.97CNGX_6F ------A----T-----A-C---G--C-GT---AA--A-C-A----G-------G-----------------------------A--------------------------A--------------A----------G-------------------------------- 7538
09_cpx.GH.96.96GH2911 -----------C-------C---G--C-GC---------C------G-------G---------------------------G----------------------------A--------------A----------G------G----A--------C----------- 7531
10_CD.TZ.96.96TZ_BF061 ------------C-----C----G-----------------A----G--------G---------C------------------A-------------------------AA------------------G------G------T----A-------------------- 7711
11_cpx.GR.x.GR17 -----------C-----A-C---G--C------------C--------------G-----T----------A----------G-A----T--------------------AA--------------A----------G-----------A------TG---------C-- 7632
12_BF.AR.99.ARMA159 -------------------C---G--C------------C--------------G-----T----A----------------G-A--------------------------A----C--------------------G-----------A------------A------- 8324
13_cpx.CM.96.1849 -----G-----C-------C---G--CC-G---------C---------------GT---T----A------------------A----T-----------------T--AA-------------------------G-----------A--A--T-----------G-- 7784
14_BG.DE.01.9196_01 -----------------------G-----------------------G--------G---------------R----------G-A----T---------A-------TK-AA-------------------------G-----K---A-A------------A------- 7864
15_01B.TH.99.99TH_MU2079 -----------------A-C---G-----G------------------------G----------C------------------A--------------------------A----C--------------------G-------------------------------- 7731
16_A2D.KR.97.97KR004 ------A----C-------C------------------C---------------G-----T----C------------------A--------------------------A----A------------A-G---A--A-----G----------------------G-- 7725
17_BF.PE.02.PE02_PCR0155 -------------------C---G--C------------C---------------G----T----C-----------C----G-A----------A---------------A----C--------------------G-----------A------------A------- 7531
18_cpx.CU.99.CU76 -----------C-------C---G--C------------C-A----C-------G-------R-------------------G-A-------------------------AA-------------------------G-----------A--G----------------- 7667
19_cpx.CU.99.CU7 -------------------C---G--C-------GC-AG---------------G-T---------------------------A----T-----------------T--AA----------------C--------------------A-------------------- 7425
20_BG.CU.99.Cu103 -----------C-------C---G--C--G---------C--------------------T----C-----A---------------------------------------A--------------------A----G---------A-A-------------------- 7793
21_A2D.KE.91.KNH1254 ------A-G--------------G--C--G-----------------G-------G----------A----------------G-A-----------G--------------A-------------------------G------T------------------------- 7581
22_01A1.CM.01.01CM_0001BBY -----------C-------C---G--C--G---T---A-C------C-------G-----T----------A--------T---A----T--------------------AA----A--------------------G---------A-A-------------------- 7526
23_BG.CU.03.CB118 -------G---C-------C---G--C------------C-------C------G-----T----C-----A------------A----T-----A--------------AA-----------------A-G-----G---------A-A-------------------- 7787
24_BG.CU.03.CB378 -----------C-------C---G--C--G---------C------C-------G----------A-----A------------A----T--------------------AA--------------------A----G---------A-A-------------------- 7802
24_BG.CU.03.CB471 -----------C-------C---G--C--G--------CC------C-------G----TT----------A----------G-A----T--------------------AA-------G-----------------G---------A-A-------------------- 7805
25_cpx.CM.01.101BA -----------C-------C---G--C--G---------C--------------G-----T----G-----A----------G-A----T-----------------T--AA-----------------------T-G-----------A-------------------- 7609
27_cpx.CD.97.97CDKTB49 ------A----T--G--A-C---G--C--G--------CC--------------G-----T----------A------------A----T-----------------G---A------------------G------GTG-------A-A-------------------- 7563
28_BF.BR.99.BREPM12609 -----------C------C----G----CG--------C---------------G----------C------------------A-------------------------A-------------------G--------------------------------------- 7760
29_BF.BR.02.BREPM119 --------G--------------G-----G-------A---A------------G--------C-C------------------A-------------------------AA--------------------A--------------A-A------------A----G-- 7511
31_BC.BR.02.110PA --------G--T-----A-C---G--CC-G---CA--A-C--------------GG---------G------------------A--------------------------A--------------A----------G--C----------------------------- 7852
32_06A1.EE.01.EE0369 ---------------C-A-C---G--C--G----------------G-------G-----T----C------------------A----T--------------------AA-------------------------G-----------C-------------------- 7985
33_01B.MY.05.05MYKL007_1 ------A-G--T--G--A-----G-----G---A-----C------G--------G---------------A----------G-A----T-----A-----------G--AA--------------A----------G-----------A-------------------- 7741
34_01B.TH.99.OUR2478P -------GG--T--G--A-----G---------A-----C--------------G-----T-----------------------A----T-----------------T--AA--------------A----------G-----------A-------------------- 7525
35_AD.AF.05.05AF095 -----------C-------C------C--G-------A-C--------------G-----T----A-----------------------T--------------G-----AA----A--------------------G---G-----A-A-------------------- 7525
36_cpx.CM.00.00CMNYU830 -----------C-------C---G--CC-G---------C--------------G----------A-----C------------A----T-----------------T---A----------------C--------G---AA-G--A-A-------------------- 7556
37_cpx.CM.00.00CMNYU926 -------------------CC--G--C------A---A-C------G--------G---------C----------------G-A----T--------------------AA----------------C--------G-----------A-------------------- 7541
38_BF1.UY.05.UY05_4752 -----------C-------C---T--C-----------CC------G-------G-T---T---GA------------------A----------A---------------A----C---------A----------G-----------A-------------------- 7770
39_BF.BR.03.03BRRJ103 -------------------C---T--C------------C------G-------G-----T---GC----------------G---------------------------AA----C---------------T----G-----------A------------A------- 7840
40_BF.BR.04.04BRRJ115 -----------C-------C---GA--------A-------------G-------G--------C--------------------A-------------------C-----A--------------------------G-------------------------------- 7928
42_BF.LU.03.luBF_05_03 -----------------A-C---G--C------------C--------------G-T---T----C---T-------C------A-----------------------------------------------T----------------A-------------------- 7821
43_02G.SA.03.J11223 -----------R-------C-C-G--C--G------R----------C------G-----T----------A----------G-A----T-----------------T--AA-------------------------G-------------RA----------------- 7842
U.CD.83.83CD003 -G----A----T---C-------G-----G---------C-------C------G--------------T--------------A--------------------------A----C---------------A----GT--------A-A-------------------- 7835
U.CD.90.90CD121E12 ---------------C-------G-----G---------C--------------G-----T--------T--------------A----T-----A---------------A--------------------A----G---------A-A-------------------- 7530
U.GR.99.GR303 -----------CC------C---G--C-CG--K------C-A--------------A---T----------A----------G-A----T--------------------AA----A--------------------G-----GG----A--A-----C----------- 7646
U.NL.95.U_NL_95_H10986_D1 -----------C-------C---G--C-CG---------C------G------------TT-----------------------A-------------------------AA--------------------A----G---------A-A-------------------- 8265
CPZ.CD.90.ANT ---A--A-G-------AT---C----C------AGCTCA--A------------G--------C-A-----A------------A----T--T-------------CTA-T-----AC-------TT-GC---TAT--G--AG-TGCT-A-GA-AG-------------- 7818
CPZ.CM.01.SIVcpzCAM13 ------A-G--C--------G--G---C--------TC---------C------G-------C--C-----------C------A-T---C-T---GC---------GA-TA-------------TA-CA-GT-A--TG-C--AG-GA-C-GG-----C--------T-- 7943
CPZ.CM.05.SIVcpzEK505 ------A-GTC-C-C--------G---C-----T-------------C------G----T-----A-----------C------A-T---------G------C----A-T-----CA--------A-AAGCATAA--A-G-----G--AGCA----------------- 7845
CPZ.CM.05.SIVcpzLB7 -----GA--A--C-T--A---C-G---C-T---A----CC--------------G------GTC-C-----------C------A-------T------A----------AA---TG-GT--G---A----------G--CG-------A--------C----------- 7870
CPZ.CM.05.SIVcpzMB66 ------A-----C-CC-A--------C--G-----G--CC-C------------G--------TCA------------------A----T--C--A-----------A--AA-----------------------T-G-----A-----A-------------------- 7983
CPZ.CM.05.SIVcpzMT145 ------A-G--G-----------G---TC------T-A---A----G-------G--------------T-------C-----GA-G---T----AG-----------A-TA----C---------TGC-----AA--GGA--AC--A-C-----G------A------- 7871
CPZ.GA.88.GAB1 ------AGG--CC----------G---C-------C--C---------------G----------A------------------A-T---C-T---GC---------AA-CA----G--------TA--A-GA-A--TT-C--AG----C-GG--------------C-- 8399
CPZ.TZ.00.TAN1 --CA-------G--G--A--------C------AGC---C-T---G--------G-T------C-A------------------A--GCT-----C-------C---A--A-----AC-T-----TC-CA-GTT-A--G--AA-G----T--GCA------------T-- 8015
CPZ.US.85.CPZUS ------A--TC-C-C--------G---C-----T-----C-A-----C------G-T--T--C--C------------------A----TT----AG------C-A----A-----AA-C-----T--AGG--TGA--T----AC----AGCA-A------------T-- 8367
CPZ.CM.x.SIVcpzDP943 -----GAG-AGC-----------G---C-----T--------------------G-T--TT-T--------------C------A----TC-----GC-----C------AA----CA----------AGGA-T-A-CA-G--CT--A-AGCA-AG-------------- 7925
CPZ.CM.x.SIVcpzMB897 -----G--------G--A-CTC-G---------T----CC-------------CG-T------C-G-----------------------TG-C-----------------AA----CA-----------------A-G-----------C-------------------- 7873
N.CM.02.DJO0131 -----GA--TC-C-C--------G---C-G---------C-------C-------GT--C-----C-----------C--C--GA--GCT------G------C----A-TA----CA-------CA-AAG--TAA--A-AA-----A-AGCA----------------- 7786
N.CM.04.04CM_1015_04 -----GA--TC-C-C--------G---C-----A-----C-------C------G-T--TT----C-----------C------A--GCT------G------C----A-TA----CA-------CA-AAG--TAA--A-AA-----A-AGCA----------------- 7764
N.CM.04.04CM_1131_03 -----GA--TC-C-C--------G---C---------R-C-------C------G-T--T----GC-----------C------A--GCT------G------C----A-TA----CA-------CA-AAG--TAA--A-AA-----A-AGCA----------------- 7829
N.CM.95.YBF30 -----GA--TC-C-C--------G---C------AC---C------GC------GGT--T-----A------------------A--GCT-------------C----A-T-----CA-------CA-AAGATA----A-AA-----A-AGCA----------------- 7910
N.CM.97.YBF106 -----GA---C-C-C--------G---C-----------C-------C------G-T--T-----A-----------C------A--GCT------G------C----A-TA----CA-------TA-AAG---AA--A-AA-----A-AGCA----------------- 7865
O.BE.87.ANT70 -----GA--A-G--GC----G------G-------CTC-A-T---------C--G-------T--A---T--------------A--GCA-----C----------CAC-A-----GG-G---G-TA-CA-GATA------AA---------A-ACA------------- 8406
O.CM.91.MVP5180 ----T-A---C---G-----GC-----G-------CTC-C-T------------G-------T--A---T--------------A--GCT-----C-----G----CAC-AA-----A-----G-TA--A-GATAA-----AA-C-------G-ACA------------- 8432
O.CM.96.96CMABB637 -----GA--AGG---C----G------G-------CTC-C-T----G----C--G-------T--A---T--------------A--GCT-----C----------CAC-AA-----G-G-----T--CA-G-TAA-----AA-C-------G-ACA------------- 7899
O.CM.98.98CMA104 -----GA--A-G---C-A--GC--AG-G-------TTC-A-T------------G-------T--A---T-------C------A--GCT-----C----------CAC-A------G-G-----T--CA-GATA------AACC------GA-ACA------------- 7910
O.CM.99.99CMU4122 -----GA-GA-G--GC----G---A--G-------TTCAC-T---G-----C--G-------T--A---T--------------A--GCA-----C----------CAC-AA-----A-------T--CA-GATA------AA-C------GA-ACA------------- 7882
O.FR.92.VAU -----GA-------GC----G------G-------CTC-A-T---------C--G-------T--A---T-------C------A--GCT-----C----------CAC-AA-----G-----G-----A-GTTA------AA-C----A--G-ACA------------- 7972
O.SN.99.SEMP1299 -----GA--A-G--GC----G------G-------TTC-A-T-----C---C-GG-------T--A---T--------------A--GCA-----C--------G-CAC-AA----GG-G-----T---A-GATA------AA-C-------A-ACA------------- 8474
O.US.99.99USTWLA ------AGG-CC--GC----G------G--------TC-A-T---------C--G-T-----T--A---T--------------A--GCT-----C----------CAC-AA-----G-------T--CA-GATA------AACC----A--G-ACA------------- 7883
```

```
                         Tat Rev intron end  Tat Rev exon 2 start              Tat premature stop in HXB2                    Tat end
B.FR.83.HXB2             TTCACCATTATCGTTTCAGACCCACCTCCCAACC......CCGAGGGGA...CCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCCTTGGCACTTATCTGGGACGATCTGCGGAGCCTGTGCCTCT 8520
Rev exon 2                                  N  P  P  P  N  .  .  P  E  G  .  T  R  Q  A  R  R  N  R  R  R  R  W  R  E  R  Q  R  Q  I  H  S  I  S  E  R  I  L  G  T  Y  L  G  R  S  A  E  P  V  P  L
Tat exon 2                                   P  T  S  Q  P  .  .  R  G  D  .  P  T  G  P  K  E  *  K  K  K  V  E  R  E  T  E  T  D  P  F  D  *
Env                        S  P  L  S  F  Q  T  H  L  P  T  .  .  P  R  G  .  P  D  R  P  E  G  I  E  E  E  G  G  E  R  D  R  D  R  S  I  R  L  V  N  G  S  L  A  L  I  W  D  D  L  R  S  L  C  L
A1.GE.99.99GEMZ011       C-----C--G--A--G-------TTGC----CA-......--AGA----...--A----------G-A----CA-----------------CA--G------------G-------------G-----T---A------GC----------------------------- 7695
A1.KE.00.KER2008         C-----T--G-----C-------TTAC-----A-......--A----A-...-T-----------G-A----C------------------CA--------G------G-----C-------G-----T---A------GC---------C------------------- 7726
A1.KE.00.KNH1144         C-----TG-G--A-------T---TGC-----A-......--AG----T...-T-----------G-A----C------------------CC--G-----G------G-------------G-----T---A------GC-------------A---A----------- 7752
A1.KE.00.KSM4024         C-----T--G--------------TTC-----G-......--A-----T...-T-------------A----C----------A-------CA--G-----G------G-------------------T---A--GA--GC------------A---------------- 7649
A1.KE.00.MSA4069         C-----T--G--A--C--------TAC-----A-......--A----A-...-T-------------A----C-------G----------CA--G------------G-------------------T---A------GC---------C------------------- 7728
A1.KE.00.NKU3005         C-----TC----A------------AC-----G-......---G-----...---------------A----C--C-------------G-CA--G------------G-----------C-G-----T---A------GC----------------------------- 7731
A1.RU.00.RU00051         C-----C--G--A--K-------TTAC-----A-......--AGA---W...--A----------G-A----C-----C-G----------YA--G--A---------G------------RG-----T---A------GCY------------Y--------------- 7794
A1.RW.93.93RW_024        C-----T--G--A-------T---TAC-----A-......--AG---AT...-------------G-A----C------------------CA-------A-------G-----------A-G-----T---A--G---GC---------------------------T- 7725
A1.SE.95.SE8891          C-----T--G-----C----T---TAC-----G-......--AG-----...-T-----------G-A----C------------------CA-------A-------G-----C-------------T---A------GC---------C--CA--------------- 7685
A1.SE.95.UGSE8131        C-----TC-G-----C----T---TAC---C-G-......--A----AT...-------------G-A----C------------------CA--G------------G-----C-----C-G-----T---A--G---GC-------------A------------G-- 7898
A1.TZ.01.A173            C-----T--G--A-----------TAC-----A-......--A----AT...-T-----------G-A----C------------------CA--G------------G-------------G-----T---A------GC------------C---------------- 7729
A1.UA.01.01UADN139       C-----C--G--A--G-------TTAC----CA-......--AGA----...--A----------G-A----C------------------CA--G------------G-C-----------G-----T---A------GC------------------------T--T- 7716
A1.UG.92.92UG037         C-----CC-----------------TAC---G-A-......--A------...-T-----------G-A----C------------------CA--------G------G--A--C-------G---G-T---A------GC---------C-------A----------- 8530
A1.UG.99.99UGA07072      C-----T--G-----C--------TAC-----A-......--A-----T...-T---------A-G-A----C------------------CA---------------G-----------A-G-----T---A------GC---------------------------T- 7726
A2.CD.97.97CDKTB48       C--------G--A-------T--CTAC-----A-......--AGA----...-T-----------G-A----C------------------CA--G------------G-------------G-----T---A------GC---------C--A---------------- 7830
A2.CY.94.94CY017_41      C-----TG-G--A-------T--CTAC-----G-......--AGA---T...------------AG-----CC-------G----------CA--G------------G-----------------G-T---C------GC---------C--A---------------- 7892
A.SN.01.DDI579           C-----TC-C-------------TTAC-----A-......--C------...---------------A----C------------------CA---------------G-----C-------G-----T---A--------------------A---------------- 7724
A.SN.01.DDJ369           C-----TC-G--A-----------TTC-----G-......--AG-----...-T----------AG------C------------------CA---------------A-------------G-----T---A------GC----A----C--A--------------T- 7742
A.SN.96.DDJ360           C-----TC-C--------------TAC-----A-......--AG-----...-T-----------G-A----C------------------CA---------------G-------------G---C-T---A------G----------C--A---------------- 7706
A.CD.02.02CD_KTB035      C--------G-----C----T--CTAC-----A-......--AGT----...GT-----------G-A----C-------G--A-------CA---------------G-------------G-----T---A------GC------------A--------------T- 8519
A.CD.97.97CD_KCC2        C-----T----------------CTA------G-......--AGA----...-------------G-A----C-------G----------CA--G----AC------G-------------G-----T---A-A----GC------------A--------------T- 8519
A.CD.97.97CD_KTB13       C-----T----------------TTAC-----G-......--AGA----...-T-----------G-AA---C-------G----------CA-------AC------G-----C----C--G-----T---A------GC------A--C--A--------------T- 7796
B.AR.04.04AR151516       C-----------A--C---------------G-T......---------...--A----------G--R---C-------------------------------------GA--C-----C-CT----T---A-----C-----------C-------------T----- 7729
B.AU.87.MBC925           C----------------------G-T-----G--......---------...----G---------------C------------------------------------GAGA--------G------T---A------T------TG--C--A---------------- 8539
B.BO.99.BOL0122          C-----------A----------G----------......---------...-----------A--------C-------------------------------C---GG-CG--------C-T----TA--A-----C--------------A---------------- 7721
B.BR.03.BREPM2012        C--------------C----T--G----------......-A-------...-------------G------C------------------C-----------------GAAA---------CT----T---A-----C-----------C-------------T----- 7965
B.CA.97.CANB3FULL        C------A---------------T-------G-T......---------...--------------------C------------------------------------GGA--C-------------T---A--------------------A----------T----- 7793
B.CN.05.05CNHB_hp3       C--------G--A--G---------T--------......-A-------...----GA--------------C------------------C---------G-------GAGA---------------T----A-------------G-----A---------------- 8536
B.CO.01.PCM001           C-----------A------------------G--......G--------...--------------------C----G------------------------------GGG----------C-T----T---A-G-G-C------------------------------- 7703
B.GB.83.CAM1             C--------G--A----------G-T-----GTT......--A------...--------------------C----------------G--------------C----GGG--------A-CT----T---A-----C-----------C------------------- 8520
B.GB.86.GB8              C-----------A--G------------------......--A------...--------------------C------------------CA------------------CA----G-----T----T---A--G---T----------C------------------- 8511
B.GE.03.03GEMZ010        C--------G--A--C-------G------G---......-A-------...-------------------C-------------------CA-----------------GGA-C----C--CT----T---A-A---C-----AT-------AA--------------- 7697
B.IT.05.SG1              C--------------C--------TT--G--G--......--AC-----...--------------------C---------------------A-------TG------G---C------G-T----T---A---A-C-------T---C--AT------------G-- 8205
B.JP.05.DR6538           C-----------A--G-------TTA-----GT-......T-------CCA---------------C-----C----------------------G-------------GGAGT--------------T---T------------A----C------------------- 8556
B.KR.05.05CSR3           C-----T------------------------G--......--C------...--------------------C-------------------------------------G--A--C---C-T---G-T---A-A-A-C-------T----------------------- 8583
B.NL.00.671_00T36        C--------G--A--C---------------G-T......---------...--------------------C------------------------------G----GGG-C----------T----T---A---A-C-------T---C------------------- 8081
B.RU.04.04RU128005       -------------------------------G--......AG-------...-------------------C--------------------------------A----GGA--C------G-T----T---A--GA---------T---C---M--------------- 7992
B.TH.00.00TH_C3198       C--------G--A--A---------T-----G--......-A-------...----G---------------C----------------G-C---------G-------GAGA--------------------A-------------G-----A---------------- 7721
B.UA.01.01UAKV167        C--------G--A--C-------G----------......-A-------...--------------------C----G-------------CA-----------------GGA-C-------CT----T---A---A-C-------T---C------------------- 7784
B.US.04.ES10_53          C--------------C----T--G--C----G--......AG-------...--------------------C----G--------------AC---------------GGA-CC------G-T----T---A---A---------T----------------------- 8498
B.US.99.PRB959_03        C--------------A-------T-------G-T......---------...--------------------C----G-------------CA----------------------------G-T----TG--A-----CT----------C------------------- 7718
C.AR.01.ARG4006          C-----C--G--A----------TTA------A-......---------...--A-G-G---T--G------C------------------CA-----A-----------------------G-----T---AT-----GC---------C--A---------------- 7711
C.BR.04.04BR013          C-----C--G-----A-------TTAC-------......---------...--A-------T--G------C------------------CA-----------------------------G-----T---AC-----GC------------A--------------T- 8011
C.BW.00.00BW07621        C-----T-----A----------TTAC-----A-......-A-------...----------T--G------C------------------CA-----------------------------G-----T---A-----GT----------C-----------------T- 7855
C.CN.98.YNRL9840         C-----T--G-------------TTAC---G-A-......--AG-----...----------T--G-A----C------------------CA-----A-A--G------------------------T---A------GC---------C-------A-------A-T- 7721
C.ET.02.02ET_288         C------C-C-------------TTAC-----A-......-------A-...----------T-AG------C------------------CA---------------------C-------------T---A--GA--GC---------CA-C---------------- 7725
C.GE.03.03GEMZ033        C-----TC---------------TTAC---G-A-......--A----A-...-T--------T-AG------C------------------CA-----------------------------G-----T---A------GCT---------------------------- 7727
C.IL.99.99ET7            C-----T----------------TTA------A-......-A-------...-T--------T--G-A----C-------G----------CA-------A----------A----------------T---A------GC----------------------------- 7686
C.IN.99.01IN565_10       C-----T--G-T-----------TTAC---G-A-......--C------...----------T-AG------C------------------CA-----A---------TG------------------T---A------GC---------C-----------------T- 7897
C.KE.00.KER2010          C-----T--G-------------TTAC-----A-......---------...----------T--G-A----C------------------CA-----A----G------------------------T---A--G---GC---------C-----------------T- 7709
C.MM.99.mIDU101_3        C-----T--G-------------TTAC---G-A-......--AGA----...----------T-AG-A----C------------------CA-----A-AG------------------------G-T---A------GC---------C-------A----------- 7879
C.MW.93.93MW_965         C-----T----------------TTAC-----A-......---------...-T--------T--G------C------------------CA------------------------G----G-----T----C-----GC---------C------------------- 7695
C.SN.90.90SE_364         C-----T--G-----C-------TTA------A-......---------...----------T--G-A----C------------------CA-----------------------------------T------GA--GC---------C------------------- 7665
C.SO.89.89SM_145         C-----T--G-------------TTAC-----A-......---------...----------T--G-A----C------------------CA-----A-----------------------------T---A------GC---------C--A---------------- 7746
C.TZ.02.CO178            C-----T--G--A--C----T--TTAC---G-A-......-A-------...----------T--G------C------------------CA-----------------------------------T---A--G---GC----A----C------------------- 7724
C.UY.01.TRA3011          C-----T--G--A----------TTAC-----G-......---------...--A-------T--G--A---C------------------CA-----------------------------G-----T---AC-----GC------------C----T----------- 7697
C.YE.02.02YE511          C-----T--G-T-----------TTAC-----A-......T--------...--A-------T--G------C------------------CA-----A--------------A---G-C--------T---A--TG--G-----A----C------------------- 7712
C.ZA.04.04ZASK164B1      C-----TC-G-----G----T--CTTC--AG-A-......--A------...-T--------T--G-A----C-------G----------CA-----A---G-----A-------------G-----T---A------GC---------C-----------------T- 7970
C.ZA.05.05ZAPSK240B1     C-----T--G--A--A-------TTA--------CCA...A--------...----------T--G------C------------------CA-----A------------CA---------G-----T----T----------------C--A----A----------- 7899
C.ZM.02.02ZM108          C-----T--G--A----------TTA------A-......---------...----------T----A----C------------------CA--------G---------CA---------G-----T---A------GC------A----------A---------T- 8557
D.CD.83.ELI              C-----TC-G-------------T-------G--......---------...-------------------C-------------------C---G-------------G-GA---GC----------T--C------------------C------------------- 8056
D.CM.01.01CM_0009BBY     C-----TC----A----------T-------G--......---------...---------------------------------------CA--------G--CC-----A-----G----------T--CA-------------------------A----------- 7745
D.KE.01.NKU3006          ------TC-G-------------T-------G--......---------...---------------------------------------CA--G-----G---------------G----------TA-CA---------------------A---A---------T- 7763
D.KR.04.04KBH8           C-----TC-G-------------C-T-----G--......---------...-------------------C--------G----------CA--------G--A---A--------G----G-----TA-CA-------------------------A----------- 8433
D.TD.99.MN011            C-----CC-G--A--A-------T-------G-A......---------...---------------------------------------CA--------G---------------G----------T--CA-------------------------A------T---- 7710
D.TZ.01.A280             ------TC-G-----G-------T-------G--......T--------...--A------------------------------------CA--G-----G------G--------G----------T--CA-------------------------A----------- 7708
D.UG.99.99UGK09259       ------TC-G--A--C-------T-------G-T......---------...--A------------------------------------CC--G-----G--AC-----A-----G----------T--CA-----------------CT--A---A----------- 7721
D.YE.01.01YE386          ------TC-G--A----------T-------G--......---------...-------------------------------------------------G---------------G----------T--CA---------------------A---A----------- 7676
D.YE.02.02YE516          C-----TC-------A-------T-------G--......---------...---------------------------------------CC--G--A--G--CC-----------G----------T--CA-----CT-------G----------A----------- 7700
D.ZA.90.R1               C-----TC-G-------------T-------G--......-A-------...---------------------------------------C---G--------A---A--------G--------G-T--ATT------------------------A----------- 7874
F1.AR.02.ARE933          C-----T--G--A--A-------TTA------G-......-------A-...--------------------C-------G----------CA---G-A----------G-G----------C-----T---A--T---GC---------C---A---A----------- 7790
F1.BE.93.VI850           C-----T--G--A--A-------TTA------G-......---------...--------------------C-------G----------CA---G-A----------G-G--C-------C-----T---A--T---GC-------------A---A----------- 7797
F1.BR.01.01BR125         C-----T--G--A--G-------GTA------G-......--AGA--A-...--------------------C----G--G-------AC-CA---G-A-----C----G-GA---------------T---A--T---G----------C---A---A----------- 7956
F1.ES.x.P1146            C-----TC----A--G-------TTA------G-......---------...--------------------C-------G----------CA-----A------------AA---------------T---C--T---G-------G--C---T---A----------- 7827
F1.FI.93.FIN9363         C-----T--G--A--A-------TTA-----G--......----C--A-...--------------------C-------G----------CA---G-A----------G-G----------------T---A--T---G----------CT--A---A----------- 7816
F2.CM.02.02CM_0016BBY    ------T--G--A--A-------TTA------G-......---------...-------------G------C------------------CA-----A----------G-GA---------G-----T---A------GC----------------------------- 7679
F2.CM.95.MP257           C-----T--G--A--G-------TTA------A-......T--------...-----A-------G------C------------------CA-----A------------CA---------G-----T---A------GC---------CT-CA-----------G--- 7713
F2.CM.97.CM53657         C-----T--G-----A-------TTAC-----G-......-G----A--...---AC--------G------C------------------CA-----A-AC-------G-CA---------G-----T---A------GC---------C--A---------------- 7691
G.BE.96.DRCBL            C-----T-----C----------TT----ACCA-......-A-----A-...-------------C------C-------G----------CA-----------------------------G-----T---A------GC---------C--A---------------- 8459
G.CM.01.01CM_4049HAN     C-----T--G--A--C-------TTAC--ACCA-......-A-----A-...--------A-T--G-A----C------------------CA-----A-A----------A----------G-----T---A------GC---------C--A--------------T- 7709
G.CU.99.Cu74             C-----C--G--A--C-------CTAC--ACCA-......-A-----A-...---------------A----C-------G------C---CA-------A---------------------G---G-T---A------GC------------R---------------- 8139
G.ES.99.X138             C--------G--A--C-------TTAC--ACCA-......-A-----A-...-------------G-A----C----------A-------CA-----A-----A-----------------G-----T---A--G---GCG--------C--A----------T----- 7955
G.GH.03.03GH175G         C-----T--G--A--C-------TTAC--ATCAG......-A-----A-...--A-------TA---A---CC-------G------C---CA------------G---G-G--C-------G---T-T---A--G---GC---------C--A---------------- 8566
G.KE.93.HH8793_12_1      C-----T--G--T--C-------TTAC--ACCA-......-A-----A-...---------------A----C------------------CA-----A-----------------------G-----T---A--T---GC----------------------------- 7890
G.NG.01.01NGPL0669       C--------G--A----------TTAC--ACCA-......-A-----A-...--------------------C------AG----------CA--G------------------------A-------T---A--GA--GC---------CG----------------T- 7713
G.PT.x.PT2695            C--------G--A--C-------TTAC--ACCA-......-A-----A-...--------A----G-A----C-------G-A--------CA-----------A-----------------G---G-T---A------GC------------A---------------- 8531
G.SE.93.SE6165           C-----T--G--A----------TTAC--ACCA-......-A-----A-...--------------------C-------G------C---CA--G-----G------AG-G----------G-----T---AC-G--------------C------------------- 7938
H.BE.93.VI991            C-----TC-G--T--A-------TTA----G-A-......-A-C-----...------------AG--A------------------C---CA------------------------G--------G-T----C-----G-------G--C-------A----------- 7919
H.BE.93.VI997            C-----T--G--T----------TTA----G-A-......---C-----...-----------------------------------C---CA--------G-------G-G----------------T---AC-GA--G----------C--C---------------- 7834
H.CF.90.056              C-----T--G--T----------TTG----G-A-......--AC-----...-------------------C--------G------C---CA----------------G-GA---------------T---AC--G--G----------C--C----------CA---- 7839
J.CD.97.J_97DC_KTB147    C-----C--G--A--A-------TTA------A-......--A-C----...GT---------A-G-A----C-------G----------CA--G----AC------G-------------G-----T---A--G---GC---------C-----------------T- 7723
J.SE.93.SE7887           C-----T--G--A----------TTA------A-......--A-C--A-...G------------G------C-------G----------CA--G----AC------G-------------------T-C-A--T---GC---------C-------A----------- 7827
K.CD.97.EQTB11C          C-----T--G--A--C-------TTAC-----G-......---------...--------------------C-------G----------CA-----A----------G-GA-------C-G---G-T---A-T----GC-------------A---A----------- 7703
K.CM.96.MP535            C-----T--G--A----------TTA--------......T--------...G-A-----A-----------C------------------CA-----A-A-------AG------------G---C-T---A--G---GC-----------------A----------- 7680
```

```
                               Tat Rev intron end  Tat Rev exon 2 start          Tat premature stop in HXB2                                  Tat end
B.FR.83.HXB2                   TTCACCATTATCGTTTCAGACCCACCTCCCAACC......CCGAGGGGA...CCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCCTTGGCACTTATCTGGGACGATCTGCGGAGCCTGTGCCTCT 8520
Rev exon 2                                         N  P  P  P  N  .  .  P  E  G  .  T  R  Q  A  R  R  N  R  R  R  R  W  R  E  R  Q  R  Q  I  H  S  I  S  E  R  I  L  G  T  Y  L  G  R  S  A  E  P  V  P  L
Tat exon 2                                           P  T  S  Q  P  .  .  R  G  D  .  P  T  G  P  K  E  *  K  K  K  V  E  R  E  T  E  T  D  P  F  D  *
Env                              S  P  L  S  F  Q  T  H  L  P  T  .  .  P  R  G  .  P  D  R  P  E  G  I  E  E  E  G  G  E  R  D  R  D  R  S  I  R  L  V  N  G  S  L  A  L  I  W  D  D  L  R  S  L  C  L
01_AE.CF.90.90CF11697          C-----T--G--A--C-------TTAC--ACCA-......-A-----A-...---------------A----C-------G------C--AGA--G--A-------------C-------G-----T----T-G---GCG----------A--------------- 8472
01_AE.CN.05.FJ051              C-----T--G--A--C-------CTA---ATCAT......-A-----A-...--------------------C-------G------C---CA--G------------G-GA---------G-----T---A--T---GCG---------G-A--------------- 8502
01_AE.CN.06.FJ054              C--------G--T--C----T--CTTC--ATCAT......-A-----A-...--A------------AA---C--------------C---CA----------------G-G----------G-----T---A------GCA------------------------- 8529
01_AE.HK.x.HK001               C-----C--G--T--C----T--CT----ACCAG......-A-----A-...--A-----A---G-A----C--------------C-G-CA----------------G-AA---------G-----T---A--T--GT-G-----------A----A----------- 7881
01_AE.TH.01.01TH_R2184         C-----T--G--T--C-------CTAC--ATCAT......-A-----A-...---------A-----A----C-------G------C---CA--------------GG-GA---------G-----T---A------GC------------A--------------- 7723
01_AE.TH.02.OUR769I            C-----T--G--T--C----T--CTTC--ATCAT......-A-----A-...---------------A----C-------G------C---CA----------------G-G----------G-----T---A--G---GC------------A--------------- 7704
01_AE.TH.90.CM240              C-----T--G--T--C-------CTTC--ATCAT......-A--A--A-...---------A----------C-------G------C---CA--G------------AG-G----------G-----T---A------GC------------A--------------- 8073
01_AE.US.00.00US_MSC1164       C-----T--G--T--C-------TTTC--ATCAT......-A-----A-...--------------------C-------G----------CA--G--------------G-G----------G-----T---A--G---GC------------A--------------- 7716
02_AG.CM.02.02CM_4082STN       C-----T--G--A--C-----------A--ACCA-......-A-G---A-...----G--------G-A---C--C-----------C---CA-----------G---G-GA---------G-----T---A--T---GC---------C------------------ 7698
02_AG.EC.x.ECU41               C-----T-----T----------TTAC--ACCA-......-A-----A-...---------------A----C-------G------C---CA----------A---G-G----------G-----T---A------GC---------C------------------ 7582
02_AG.GH.03.GHNJ196            C-----T--G--A--C-------TTAC--ACCA-......-A-----AT...--A-G---A--A---A----C--------------C---CA--C-----C------G-G----------G---C-T---A--T---GC---------C--AA--------------- 8567
02_AG.NG.01.PL0710             C-----T--G-------------TTAC--ACCA-......-A-G---A-...-----------G----A----C-------G------C---CA----------------GG-G--C-------G-----T---A------GC---------C-------A------T--T- 7668
02_AG.NG.x.IBNG                C-----T--G-----C-------TTAC--ACCA-......-A-----A-...---------A-----A----C-------G----------CA----A-----------G-G----------G-----T---A------GC------------------------------ 8026
02_AG.SE.94.SE7812             C-----T--G--T--C-------TTAC--ACCA-......-A-----A-...----G----------A----C-------G------C-G-CA----------------G-G----------G-----T---A------GC------------------------------ 7910
02_AG.SN.98.MP1211             C-----T--G--A--C-------TTAC--ACCA-......-A---------...G-A--------------C-------G------C---CA-----------G---G-G----------G-----T---A------GC---------C------------------ 7725
03_AB.RU.97.KAL153_2           ------------A----------G-------G--......-A---------...--------------------------A---------------------C-------------------A---T---A----------------C---A-----------T--- 7697
04_cpx.CY.94.CY032             C-----T--G--T--G-------TTA-------AACC...-AAC-------...-T----------G-----C---------------C---CA-------AG------------C-------------T----C----------------C-------A----------- 7893
05_DF.BE.x.VI1310              C-----TC-G--A----------C-T-----G--......----------...-------------------A--C-------G---------CA-------G-----A-------G-------G--------C-T--CA--------------------A---A----------- 7926
06_cpx.AU.96.BFP90             C-----T--G-----G-------TTA------A-......--A-C------...G-A---------G--A--C-------G---------CA--G----AC------G---------------------G-T---A------GC------------------------------ 8568
08_BC.CN.97.97CNGX_6F          C-----T--G----------T--TTAC---G-A-......--AG-------...--A-G-----T--G-A---C------------------CA----A-AC---C-----------------------------T---A--G---GC---------C--A----A----------- 7699
09_cpx.GH.96.96GH2911          C-----T--G--A--C-------TTAC-----A-......-------A-...--------------------C-------G---------CA----A--T-----A----------G-----T---C--G---GC---------C---A---A----------- 7692
10_CD.TZ.96.96TZ_BF061         ------TC-G-----G--C----T-------G--......----------...------------------------------------G-----G---------------G---G-----T----A-G----------------C---A---A----------- 7872
11_cpx.GR.x.GR17               ------T--G--C------G---TTAC-----G-......-A-CA--A-...-------------G-----CCA------G---------CA--G---AC---C--------------C-G-----T---A------GC---------------A---A----------- 7793
12_BF.AR.99.ARMA159            C-----T--G--A--A--------TA------G-......-------A-...--------------------C-------G---------CA--G--A--T-------G-G-------------------TA---A----------------C---A------------T- 8485
13_cpx.CM.96.1849              C-----T--G--A----------TTAC--ATC--......AA-----A-...--A-------------A----C-------G---------CA---------------------G--C-G-----T---A------GC------------------------------ 7945
14_BG.DE.01.9196_01            C--------G--A--C-------TTAC--ACCA-......-A-----A-...--------------G-A---C------------------CA----A----A------------------G-----T---A--G---GCG---------Y--A--------------- 8025
15_01B.TH.99.99TH_MU2079       C-----T--G--TC-C-------CTA---ATCAG......-A-----A-...G-------A------A---C-------G------C---CA--G----A--G----G-GA---------G-----T---A--C---GC------------------------------ 7892
16_A2D.KR.97.97KR004           C-----TG----A-------T--CT-C---G---......--AGA--AT...-----------A--G-A---C-----C-G---------CA--G-------------G-G----------G-----T---A--G---G-------A--C--A--------------- 7886
17_BF.PE.02.PE02_PCR0155       C-----T--G--A-----------T-C-----G-......-------A-...--------------------C-------G---------CA--G-------------G-GA----------------T---A-----GT---------C------------------ 7692
18_cpx.CU.99.CU76              C-----T--G--T--C-------TTA----G-A-......-A-CT----...------------------C-----G---------C---CA--G-----G------AG-G----------G-----T---AA-----GC------G--C------------------ 7828
19_cpx.CU.99.CU7               C-----T--G--A----------TTAC--G-CA-......AGAG--......----------A----A----C-------G---------CA----A-------------------------------T---A------GC---------------A------------T- 7583
20_BG.CU.99.Cu103              C-----C--G-----C--------TAC--ACCA-......-A-----A-...--------------------C----G--G---------CA----AC---------G-G--C-------G-----T---AC-----GC---------C--C------------------ 7954
21_A2D.KE.91.KNH1254           C-----T--G--A--C-------T-------G--......----------...----------------------------------C-----------------------------------T--CA----------------C---A---A----------- 7742
22_01A1.CM.01.01CM_0001BBY     C-----C--G--A----------TTAC-----A-......-AAGA----...-T----------G-A---C-------G---------CA------A---A---G-----C-------G-----T---A--G---GC------------------------------ 7687
23_BG.CU.03.CB118              C-----C--G--A--C-------TTAC--ACCA-......-A-----A-...--------------------C-------G---------CA-------G-----------------CC-G-----T----C-GA--G---------C--C------------------ 7948
24_BG.CU.03.CB378              C-----C--G--A--C-------TTAC--ACCG-......-A-----A-...----------------A--C-------G---------CA----A--G---------------------G---G-T---AC-----GC---------C------------------ 7963
24_BG.CU.03.CB471              C-----C--G--A--C--------TAC--ACCA-......-A-----A-...--------------------C--------------C---CA-------C-------G------------G-----T---AC-----GC---------C--A--------------T- 7966
25_cpx.CM.01.101BA             C-----T--G-----C--------TAC--ACCA-......-A-----A-...--A----------------C--------------C---CA-----C--------G-CA---------G-----T---AC-T---G----------CT--A---A----------- 7770
27_cpx.CD.97.97CDKTB49         C-----TG-G--A-------T---TAC-----A-......--AGA---T...-T-A--------G-A---C-G-----------------CA--G-----------G-C------G-C--------T---A------G----------C--A---A---------T- 7724
28_BF.BR.99.BREPM12609         C--------------C-------G-------GT-......-A-----A-...------------G------C-----------------------AC--------------C----CAG-T----T---A---A-C--------A--C---A-------------T- 7921
29_BF.BR.02.BREPM119           C-----------A--C-------T-------G--......G---------...--------------------C-----------------------A--G-----A-----------AC-T----T---A-----CGC---------C---A-----------A--- 7672
31_BC.BR.02.110PA              C-----C--G--A----------TTAC-----R-......----------...--A------T--G-----C------------------GA----A-------------------------G-----T---A------GC------------A-------------T- 8013
32_06A1.EE.01.EE0369           C-----TG-G--A--G-------GTA------A-......--A-C------...G-----------G--A----------G---------CA--G--ACAC--A---G-----------CC-C-----T---A------G----------C------------------ 8146
33_01B.MY.05.05MYKL007_1       C-----T--G--T--C-------CTA---ATCAG......-A-----A-...--------A-T---A----C--------------C---CA--G---------G--TG-G--C-------C-----C-----T---A--G---GC------------A--------------- 7902
34_01B.TH.99.OUR2478P          C-----T--G--T--C-------CTAC--ATCAT......-A-----A-...--A------------A----C-------G------C---CA--G----A--------G-GA---------C-----T---AT-T---GC------------A--------------- 7686
35_AD.AF.05.05AF095            C-----T--G-----C---------AC-----G-......--AG---AT...-T----------G-A----C------------------CA--G----AC-----G-----------C-G------C-G-----T-----G--GC---------C------------------ 7686
36_cpx.CM.00.00CMNYU830        C-----T--GC-A--C-------TTAC--GCCAT......-A-----A-GAA-------------A----CA-------G------C---CA-----------A----G-G-----------------T---A------GC----------T--A--------------- 7720
37_cpx.CM.00.00CMNYU926        C-----T--G-----C-------TTAC--ACCA-......-A-----A-...---------------A----C-------G------C---CA--G-------------G-G--C-------G-----T---A------GC------------------------------ 7702
38_BF1.UY.05.UY05_4752         C-----C--------A----T----GC-----G-......T------A-...--------------------C-------G---------CA--G--A-----------G-GA----------------T----A-T---G----------C-------A----------- 7931
39_BF.BR.03.03BRRJ103          C-----T--G--A--A-------TTA------G-......--A---A-...--------------------C-------G---------CA--G--A----------TG-CA----------------T---C--T---G----------CT--A---A----------- 8001
40_BF.BR.04.04BRRJ115          C--------G--A--C-------G-------G-A.....----------...------------------C-----G---------------------------GGAA--------CT-----T---A-----C-----------C--A--------------T- 8089
42_BF.LU.03.luBF_05_03         Y--------G-------------G--A----G--......-A---------...-------------------YC----------------------------------GA-CW------G-T-----T---AA------------T---C---A--------------- 7982
43_02G.SA.03.J11223            C-----C--G--A----------TTAC--ATCA-......-A-C-A-A-...--------------------C-------G---------CA-------------------------------------T---A--G---GC------A--CG------A----------- 8003
U.CD.83.83CD003                C-----T--G--T--C-------T----------......----------...-----------A-G-A---CC----------------AGA----ACA-------G-------------------T---A------GC------A--C-----------------A--- 7996
U.CD.90.90CD121E12             C-----T--G--T--C-------TTA--------......----------...-------------G-----CC-----------------A-A----ATAC------G-------------------T---A--G---GC--------C--A--------------- 7691
U.GR.99.GR303                  C-----T--G--A--C----T--CTA------G-......----------...-T-----------------C------------------CA----A----------------------G---------CT-----GC------------A---A----------- 7807
U.NL.95.U_NL_95_H10986_D1      C-----T--G--A--C-------TTT------G-......T---------...-------------------CC------AG---------CA-------------------------------------T---A------GC---------------A----------- 8426
CPZ.CD.90.ANT                  -CAT--TC-G--A-------T--CTAC--A--A-......-A-CA--AT...--A--GCA---A---A---AG-----------AGA-A-----GAT---G-GG-GGGCC--GCA-C----G-T---C-----CT-G----TG--C---AC----A-AAT--AG- 7979
CPZ.CM.01.SIVcpzCAM13          C-----CC-C-----G-------TTA-----G--......-A-----A-...--A-GA-C--T--G--A--C--GC-G-------------CC--G---C-G-------G-GA-------A-CA--G-G--G-C-G--C-----------C--C--C-A----GT-A--- 8104
CPZ.CM.05.SIVcpzEK505          ---T--CC-C-----G-------TTA-------G.....A-TC-------...--A--------A---A---C--------C-C-------CT----AC-G-------G-G--C-------G-----T---A--T---GCG-----G--CT-CA---A---TCTA---- 8006
CPZ.CM.05.SIVcpzLB7            C-----TC-G--A----------TTTC---C-T-......----------...-T--CTCA----G-----C-------------------CA--G--C---------AG-GA---GC--G-------GC--G-----T---A------GCT---A----C---A--GA----------- 8031
CPZ.CM.05.SIVcpzMB66           ------TC-G--A----------TTA-----G--......-G-----AT...-GA--------A----A---C-------G----------CC----AC-T-----A----C-------G-C-------G-----T---A------GC----A----C--A---GA------T---- 8144
CPZ.CM.05.SIVcpzMT145          C-----T--G--A--G-------TTA-----G--......AATCA-----...--A---G--T--G--A--C----A---G-----------AC-T----G-----A--CA----G----G----G-----G--G-T----C-T---G-------A-----T----A---CCT-AG-- 8032
CPZ.GA.88.GAB1                 C-----TC-C-----G-------TTA-----GT-......-A-----A-...-AA-GA----T--G--A---C--C----G------C---CA-------AG------AG-GA--------G-A--G-A--G-G----C-----------------C--C--A-A----G-GA-T- 8560
CPZ.TZ.00.TAN1                 -A----CC---TT-CA----T--CTAC--A-G-G.....GA-CA--AT...--A--ACA---A-G------C-C--G--G---------GC-------ACAT---G-GG-CG-CC-CGCCAGCA----T---CAGTA-CG-------G--C--CA---A---CCT-A--- 8176
CPZ.US.85.CPZUS                ---T--CC-C-----A-------TTT-------A.....A-TC---A-...--A--------A-------C--------GC-C---A-A-C----AC-T------A-CGA-------C-G-----T---A-----G-------A----T---A---CCTGA--- 8528
CPZ.CM.x.SIVcpzDP943           ---T--CC-C-----G-------TTA----G--A.....A-A-------...--A--------A-G----C------------------CC--G----G--C-----CG--C-------G-----C-------G-----T-----T--GC------------C---A---CCTA---- 8086
CPZ.CM.x.SIVcpzMB897           C-----T--G-----C-------CTT-----GT-......-G-----AC...-A--------A--------C-------G---------A-AC---AC-T-----AG-G--C-------G-----C-------G-----T---A------GC---------CT--A---A----------- 8034
N.CM.02.DJO0131                ---T--CC-C-----G-------TTA--------AACAACAA-C-------...--A--------A---A--C-------GC-T-----GACA--G-------------G-GA-------C-G-------C-G-----T----A-----G-------G--CT-CA---A-----TGA--- 7953
N.CM.04.04CM_1015_04           ---T--CC-C-----G-------TTA-------A.....A-C-------...--A--------A------C-------GC-T------CA--------G------TG-G-----------------T---A------G-------G--CTGCA---A----GTGAC-C 7925
N.CM.04.04CM_1131_03           ---T--CC-C-----G-------TTA-------A.....A-C-------...--A--------A----R--C-------GC-T-----R-CA--------G-------G-G---------R-----------R-----T---A------G-------G--CTKCA---A----GTGA--C 7990
N.CM.95.YBF30                  ---T--CC-T-----G-------TTA-------A.....G-A-------...--A--------A---A--C-----G--GC-T------CA--------G-------G-G---------G-----------G-----T--CA--T---G-------G--C--C----A-----TGA--- 8071
N.CM.97.YBF106                 ---T--CC-C-----G-------TTA----G--AGCA...G-C-------...--A--------A---A--C-------GC-T-----G-CA--------G-------G-G---------G-----------G-----T-----G-----------G--C--CA---A-----TGA--- 8029
O.BE.87.ANT70                  -CA---CC-C-----A----T--C-AA--ATCA-......-AAGA--A-...G-A-GA-C---A-G-A---C--G--G--G---------AGA--G---GCC---G-GG--A-CC-CGCC-C-A--G-T----C----GT-G-ACACG--C--CA-A-CAA-AATAT-G- 8567
O.CM.91.MVP5180                -CA---CC-C-----G----T--CTG-----CA-......-G-CA--A-...G-A--A-C---A-G-A---C--G-----------------AG-------GCC--AG-GG-CAGCC--GCCACCA----T----CA--AGT-G-ACACG-----CA---CAA-AAT-T-G- 8593
O.CM.96.96CMABB637             -CA---CC-C-----A-------T-AA--AGCA-......-AACC--A-...G-A--A-C---A-G-----C--G--G----------GCGA------GC----G-GG-GA-CC--GCA-C-A--G-T----CA---GT-G-ACACA-----CA---CGA-AAT-T-G- 8060
O.CM.98.98CMA104               -CA---CC-C----------T--C-A---ATCA-......-AATC--A-...GTA-GAGT---A-G-----C--G-----G---------AG-------GCC---G-TG--A-C---GCA-C-A----T----CA---GT-G-ACACG--C--CA-A-CAA-AAT-T-G- 8071
O.CM.99.99CMU4122              -CA---CC-C----------T--C-AA--ATCA-......-AATC--A-...G-A-GAGC---A-G-----C--G------------------AG-------GCC-----TG--A-CC--GCC-CCA--T-T----CAT--GT-G-ACACG--C--CA---CAA-AAT-T-G- 8043
O.FR.92.VAU                    -CA---CC-C-----A----T--C-A---A-CAA......-AAGC--A-...GTA-GA-C---A-G-----C--G-----G---------CGA-----GCG---G-GG-C--C---GCC-C-A--G-T----CAT--GT-G-ACACG--C--CA---CAA-AAT-T-G- 8133
O.SN.99.SEMP1299               -CA---CC-C--A---A----T---A-A-ATCA-......-AACC--A-...G-A--AGC---A-G-----C--G-----G---------AC---G--TGCC--CG-TG--A-CC-GGCC-C-A----T----C----GT-G-ACACG--C--CA---CAA-AAT-T-G- 8635
O.US.99.99USTWLA               -CA---CC-C-----A----T--C-AA--ATCA-......-A-CA--A-...GTA-GA-C---A---A---C--G-----G---------AGA--G---GCG---G-GG--AGCC-CGCCACC-----T----CA---AT-G-ACACG--C--C----CAA-AAT---G- 8044
```

```
                                                                                                                                                          Rev end (TAA) in some lineages
B.FR.83.HXB2            TCAGCTACCACCGCTTGAGAGACTTACTCTTGATTGTAACG.....................AGGATTGTGGAACTTCTGGGACGC.....................AGGGGGTGGGAAGCCCTCAAATATTGGTGG...AATCTCCTACAGTATTGGAGTCAGGAACTA 8645
Rev exon 2               Q  L  P  P  L  E  R  L  T  L  D  C  N  .  .  .  .  .  .  .  E  D  C  G  T  S  G  T  .  .  .  .  .  .  .  Q  G  V  G  S  P  Q  I  L  V  .  E  S  P  T  V  L  E  S  G  T
Env                     F  S  Y  H  R  L  R  D  L  L  L  I  V  T  .  .  .  .  .  .  .  R  I  V  E  L  L  G  R  .  .  .  .  .  .  .  R  G  W  E  A  L  K  Y  W  W  .  N  L  L  Q  Y  W  S  Q  E  L
A1.GE.99.99GEMZ011      -------------A-----------CA---C-----C-G--.....................----C-------------------AGCAGTCTCAAGGGACTGAGACT---------G-G----------CT-G-A...-----T--GGGA--C---G----------- 7841
A1.KE.00.KER2008        ------------A------------CA---------C-G--.....................----------G-------------.....................-------------G------G---CT----...--------GATA------G----------- 7851
A1.KE.00.KNH1144        ----------T--------G-----CGC---------TG--.....................----C-----------------A-AGCAGTCTCAAGGGGTTGAGGCT---------G-G------G---CT----...--------GGTA--C---------------- 7898
A1.KE.00.KSM4024        -------------------------CA---------C----.....................----C-------------------AGCAGTCTCAAGGGGTTGAGACT-----------G------G---CT----...--------G-TA------G---G------- 7795
A1.KE.00.MSA4069        ---------G---A-----------CA---C------C--A.....................----C-----------------A-AGCAGTCTCAAGGGGCTGAGACT---------G-G------G---CT-G--...--------GATA------GT--G------- 7874
A1.KE.00.NKU3005        -------------------------CA---------CTG--.....................----C-----------------A-AGCAGTCTCAAGGGGTTGAGACT-----------GA-----G---CT-G--...--------GGTA----------G------- 7877
A1.RU.00.RU00051        --CT---------------------CA---------C----.....................----C-----------------A-AGCAGTCTCAAGGGACTGAGACT-----------G----------CT----...--------GGTA------R---G------- 7940
A1.RW.93.93RW_024       --T----------------G-----CAC--------C-G--.....................----C-----------------A-AGCAGTCTCAAGGGGTTGAGACT-----------G------G---CTTG--...--------GTT-------G----------G 7871
A1.SE.95.SE8891         --T------G---A-----------CA-----------G--.....................----C-----------------A-AGCAGTCTCAGGGGGCTGAGACT---------G-G------G---CT-AA-...-----T---TCA------G---G------- 7831
A1.SE.95.UGSE8131       -------------------------CA-A-C-----CC---.....................----------------------AG.................................-GA-----G---CT-G--...-----T--GTT--------T--G------- 8011
A1.TZ.01.A173           -------------------------CAC---------TG--.....................----C-----------------A-AACAGTCTCAAGGGGTTGAGACT---------G-G------G---CT-G--...--------GGT-------G-A--------- 7875
A1.UA.01.01UADN139      -------------A-----------CA-A-CA----C-G--.....................----C-------------------AGCAGTCTCAAGGGACTGAGACT---------GAG----------CT-G--...-----T--GGGA------G----------- 7862
A1.UG.92.92UG037        -------------A-----------CA---------C-G--.....................----C-----------------A-AGCAGTCTCAAGGGGTTGAGACT-----------GA-----G---CT-G--...--------GTT-------G---G------- 8676
A1.UG.99.99UGA07072     -------------A-----------CA-----------G--.....................----C-----------------A-AGCAGTCTCAAGGGACTGAGACT-----------G------G---CT-G--...-----T---ATA------G----------- 7872
A2.CD.97.97CDKTB48      ----------------------T-GCA---------C-G--.....................---------------G------A-AGCAGTCTCAAGGGGCTGAGACT-----------G-------C--CT----...-----T--GGTA------G--------T-G 7976
A2.CY.94.94CY017_41     ----T-----------------T-GCA-A-------C-G--.....................----C-----------------A-TGCAGTCTCAAGGGACTGAGACT-----------GT-----GA--CT----...-----T--GTTA--C---G---G------G 8038
A.SN.01.DDI579          ----T--------------------CAC--------C-G--.....................----C-----------------AG.................................-GT--T------CT----...---G-T--G-TA------G-GAG------- 7837
A.SN.01.DDJ369          -------------------G-----CACA-------C-G--.....................----C-----------------A-AGCAGTCTCAAGGGACTGAGACT---------G-G---TC-----CT-G--...-----T--GTTA------G--AG------- 7888
A.SN.96.DDJ360          ----T--------------------CAC--------C-G--.....................----C-----------------AG.................................-GT--T------CT----...-----T--G-TA------G-GAG------- 7819
A.CD.02.02CD_KTB035     -------------A-----------CA----A----C-G--.....................----C------G----------A-AGCAGTCTCAAGGGACTGAGACT---------GAG----------CT----...--------GGT-------G-----C-G--G 8665
A.CD.97.97CD_KCC2       -TGT-----------C---------CA---------C-G--.....................----C-----------------A-AGCAGTCTCAAGGGACTGAGACT---------G-G----------CT----...-----T--GGT------------------G 8665
A.CD.97.97CD_KTB13      -------------------------CA-A-------C-G--.....................----C-----------------A-AGCAGTCTCAAGGGACTGAGACT-----------G---------CCT----...-----T--GACA------T----------- 7942

B.AR.04.04AR151516      ---------G--A------------------------G---.....................----------------------A-.....................------Y-------T--------CCT----...-----K------------------------ 7854
B.AU.87.MBC925          ----------------------------------------A.....................------------------------.....................--------------------G---------...------------------------------ 8664
B.BO.99.BOL0122         -----------------------------------------.....................------------------------.....................-----------GAT---------C------...------------------G----------- 7846
B.BR.03.BREPM2012       ------------------------------------C-G--.....................------------------------.....................--------------A---------CT----...-----T-------------T---------- 8090
B.CA.97.CANB3FULL       -----------------------------------------.....................------------------------.....................------------------------------...------------------------------ 7918
B.CN.05.05CNHB_hp3      --------------------------------------G--.....................------------------------.....................--A------------------------G--...--------G----------T---------- 8661
B.CO.01.PCM001          ------------T-----------C-A-----------G--.....................--------------G---------.....................-------------------GG---------...-G------G--------------------G 7828
B.GB.83.CAM1            --------------------------------------G-A.....................------------------------.....................------------------------------...--------G--------------------- 8645
B.GB.86.GB8             -------------A-----------------------G---.....................----------G-------------.....................--------------------G---------...--------G---------TC---------- 8636
B.GE.03.03GEMZ010       ----------T-------------------------C-G--.....................--------C---------------.....................------------------------------...-----------A------G----------- 7822
B.IT.05.SG1             ----------------------------A-----AC-GG--.....................----C---------------T---.....................------------CT------G---------...-----------------------------G 8330
B.JP.05.DR6538          ---------G--A------------T----------C-G--.....................------------------------.....................------------------------------...--------G--A------------------ 8681
B.KR.05.05CSR3          --CT----------------------------------G--.....................------------------------.....................-----------GAT------G---------...-G---T-----A------------------ 8708
B.NL.00.671_00T36       --------------------------------------GT-.....................------------------------.....................-----------------------C------...--------G--A-------T---------- 8206
B.RU.04.04RU128005      -----------------------------------------.....................------------------------.....................------------AT---------C------...--C-----C--------------------- 8117
B.TH.00.00TH_C3198      -------------------C----------------C----.....................------------------------.....................--------------------------T---...------------------TC------G--- 7846
B.UA.01.01UAKV167       --------------------------------------G--.....................------------------------.....................-----------G------------------...--------G-----------C--------- 7909
B.US.04.ES10_53         -------------A---------------------------.....................------------------------.....................--------------------G---------...--------G----------T---------- 8623
B.US.99.PRB959_03       --------------------------------------G--.....................----------G-------------.....................------A-----AT----------------...--C--T--------------------GA-- 7843

C.AR.01.ARG4006         ---------G--AA------------A-A--A----C-G--.....................--AGCG-C-----G----------.....................--C---------AT---T-------TAG-A...-GC--TG-G---------G---TA--G--- 7836
C.BR.04.04BR013         -------------A------------A-A--A----C-G--.....................---GCA------------------AGCAGCCTCAGGGGACTCCAG-----------GAT---T--G----TAG-A...-GC--TG-G--A------G---TA--G--- 8157
C.BW.00.00BW07621       -------------A-----------CA-A---G---C-G--.....................--AGCG------------------AGCAGCCTCAAGGGACTCCAG-----A------AT---T------CT-G-A...-T---TG-G---------G---TA--G--- 8001
C.CN.98.YNRL9840        -------------A-----------CA-A---G-GAC-G--.....................--AG-G-----G------------AACAGTCTCAGGGGACTACAG-A-------C-------T------CT-G-A...-G---TG-G---------G---T---G--- 7867
C.ET.02.02ET_288        --------------------------A---------C-G--.....................--AGCG------------------AGCAGCCTCAAGGGACTACAG------------AT---T--G---CT-G-C...-G-A-TG-G---------G---TA--G--- 7871
C.GE.03.03GEMZ033       ----T--------A-----------CA-A-------C-G-C.....................--AGCA-C---G------------AGCAGTCTCAAGGGACTACAG-----------------T-------T-G-A...-G---GG-G---------G---T------- 7873
C.IL.99.99ET7           ------------A------------CA-A---------G--.....................--AGCA-C----------------AGCAGCCTCAGGGGAATACAAC----------------T--G---CT-G-A...-G---TG-G---------G---TA--G--- 7832
C.IN.99.01IN565_10      -------------A-----------CA-A--AG-GAC-G--.....................--AGGA------------------AGCCTTCTCAGGGGACTACAG------------AT---T------CT-G-A...-G---TG-----------G--TT---G--- 8043
C.KE.00.KER2010         -------------A----------GCA-A---------G--.....................---G-G------------------AGCAGTCTCAGGGGACTACAG-A---------------T--G---CT-G-A...---TGTG-G---------G-------G--- 7855
C.MM.99.mIDU101_3       --T----------A----C------CA-A-CAG-GAC-G--.....................--AG-G-----G------------AGCAGTCTCAGGGGACTACAGG----------------T------CT-G-A...GG---TG-G---------G---T---G--- 8025
C.MW.93.93MW_965        ------------AA-----------CA-A---------G--.....................--AGCA------------------AGCAGTCTCAGGGGACTGCAG-----------------T--G---CT-G-A...-G---TG-G---------G-G-TA--G--- 7841
C.SN.90.90SE_364        ---------G---A-----------CA-A---------G--.....................--AGCA--------C---AT--AG.....................------------A----T--G---CT-G-A...-GC--TGYG---------G--TT---G--- 7790
C.SO.89.89SM_145        -------------A-----------CA-A---G-G-C-G--.....................--AGCG------------------AGCAGTCTCAGGGGACTACAG--A--------------T--G---CT-G-A...-----TGGG---------G---T---G--- 7892
C.TZ.02.CO178           -------------------------CA-A-------C-G--.....................--AGCG------------------AGCAGTCTCAGGGGACTACAG------------A----T--G--CCT-G-A...GG-TG-GGG---------G--T-------G 7870
C.UY.01.TRA3011         ----T--------A------------A-A--A----C-GT-.....................--AGCG------------------AGCAGCCTCAGGGGACTACAG------------AT---T--G----TAG--...-GC--TG-G--A------G-C-TA--G--- 7843
C.YE.02.02YE511         -------------A-----------CACA---------G--.....................--AGCA--------A---------AGCAGTCTCAAGGGAATACAG--------------T--T--G---CTAG-A...-G---TG-G---------G-G-TA--G--- 7858
C.ZA.04.04ZASK164B1     ----T-------AA-----------CA-A---G-GA--G--.....................--A-CCT-----------------.....................------------A----T--G---CT-G-A...-GC--TA-----------G---TA--G--- 8095
C.ZA.05.05ZAPSK240B1    -------------A-----------CA-A-------C-GT-.....................--AGCA----------------A-AGCAGTCTCAAGGGACTACAG------------AT---T------CT-G-A...-G----G----A------G---TA--G--- 8045
C.ZM.02.02ZM108         --G----------A-----------CA-A--------G---.....................--AGCA----------------A-AGCAGTCTCAAGGGACTACAG------------AT---T--G---CT-G-A...-G---TG-G---------G---TA--G--- 8703

D.CD.83.ELI             --------------------------A----A----C-GT-.....................--------A---------------.....................-----------CAT----------CT----...-----------------------------G 8181
D.CM.01.01CM_0009BBY    --------------------------A----A------G--.....................----C-------------------.....................------------------------CT----...------------------G----------G 7870
D.KE.01.NKU3006         --------------------------A----A----C----.....................------------------------.....................---------------A----G--C-T----...--------G--A-------T---------- 7888
D.KR.04.04KBH8          -------------A--------------A--A----C----.....................------------------------.....................--------------AG----G---CT-G--...-----T--G--A------TA---------- 8558
D.TD.99.MN011           ------------AA------------A----A----C-G--.....................----C----A-C------------.....................--------------T---------CT----...----------------------G------G 7835
D.TZ.01.A280            ---------G--A-------------A----A----CC--A.....................C---------G-----------A-.....................---------------A----G---CT----...--C-----G--A------T----------G 7833
D.UG.99.99UGK09259      ------------------------------------C-G--.....................------------------------.....................---------------A----G----T----...--------G---------T----------- 7846
D.YE.01.01YE386         --------------------------A----A----C----.....................-----------G------------.....................---------------A----G--CCT----...--C--T--G---------T----------- 7801
D.YE.02.02YE516         A-----------AA------------A----A------G--.....................----C------C------------.....................------------------------CT----...--C--T--------------C-G------G 7825
D.ZA.90.R1              ---------G--------------------------C-G--.....................------------------------.....................------------------------CT----...--------G--------------------- 7999

F1.AR.02.ARE933         ---------G--A------------CA-A--A----C-G--.....................----C------C.....................CGGGGACTGAGG-----T-----G----------TGCT-G--...------ACG-T----------------T-- 7915
F1.BE.93.VI850          ---------G--A------------CA-A--A----C-G--.....................-----------C.....................AGGGGCCTGAGG--------------T---------CT-G--...------ACG-G------------------- 7922
F1.BR.01.01BR125        ---------G--A------------CA-A--A----C-G--.....................---------A-C.....................AAGGGACTGAGG------------AT-------CT-CT-G--...--C---GCG-T---------------GA-- 8081
F1.ES.x.P1146           G-----G--G--T------------CA-A--A----C-G--.....................----C------C.....................AGGGGACTGAAA------------------------CT----...------ACG---------G---G------C 7952
F1.FI.93.FIN9363        ---------G--A------------CA-A--A----C-G--.....................-----------C.....................AGGGGACTGAGG------------------------CT-G--...--CA--A----------------------- 7941
F2.CM.02.02CM_0016BBY   ---------G--A------------CA-A-------C-G--.....................----CA-----C.....................AGGGGACTGAAAG-----------T-----------CT----...------GCT---------G---G------G 7804
F2.CM.95.MP257          -----------T--------A----CA-A--A----C-G--.....................----C------C.....................AAGGGACTGAAA------------T-----------CT----...------GCG---------G----------- 7838
F2.CM.97.CM53657        ---------G--A------------CA-A--A----C-G--.....................----C----A-C.....................AAGGGACTAATA------------AT----------CT-G--...------GCG---------G-C-G----A-- 7816

G.BE.96.DRCBL           -------------------------CA---------C-G--.....................----CG------------------AACAGTCTCAAGGGACTGAGATT-----------------------T----...-----T--GTT-------GC--G------- 8605
G.CM.01.01CM_4049HAN    -------------A----------GTG---------C-G--.....................----CA-C--------------A-AGCAGCCTCAAGGGACTGAGACT-----------G------G-----TG--...--------G-T-C-----G----------- 7855
G.CU.99.Cu74            --------------------------G---------C-G-A.....................----CA----------------A-AGCAGTCTCAAGGGACTGAGACT-----------G----------CT----...--------GGT-------G----------- 8285
G.ES.99.X138            ------------------------------------C----.....................----CA--R-----YT------A-AGGAGKCTGAAGGGACTGAAACT---------T-G-----------K-G--...---G-T---TT-A-----G----------- 8101
G.GH.03.03GH175G        ------------------------TGC---------C-G--.....................----CA------------------AGCAGTCTCAAGGGACTGAGACT-----------G----------CT----...--------GTT-------GT---------- 8712
G.KE.93.HH8793_12_1     -------------------------CA-------A-C-G--.....................----CG----------------A-AACAGTCTCAAGGGACTGAGACT---------G-G-----------T----...-----T--GTT-------G---G---G--- 8036
G.NG.01.01NGPL0669      -------------------------CA---G-------G--.....................----CA--A---T-A---------AGCAGTCTCCAGGGACTGAGACT---------G-G-------C--CT----...--------G-T-------G----------- 7859
G.PT.x.PT2695           ------------A------------CA---------C-G--.....................----CA--------CT--------AGCAGCCTCAAGGGACTGAGCCT---T-----G-G----------CT----...-----T--GTT-------G---G------- 8677
G.SE.93.SE6165          ------------------------CCA-----------G--.....................----CA------------------AGCAGCCTCAAGGGACTGAGACT-----------GT--T-------T----...-----T--GTT-------G-CAG------- 8084

H.BE.93.VI991           ---------G--------------------C-------G--.....................----C------------......................GGGAGA---------------------CT-CT-G--...--------G-T-------G-A-----G--- 8044
H.BE.93.VI997           ---------G--T-----------C--------------TA.....................----C------------......................GGGAGA-----AA----G-----------CCT----...-----T------------G-A--------- 7959
H.CF.90.056             ---------G--T-------------------A------GT-....................----C------------......................GGGAGA------A----------------CTC----...-----T-----A--C---G-A--------- 7964

J.CD.97.J_97DC_KTB147   -------------A-----------CG-----------G--.....................----C------GAC---......................GGACGC-----------GAT----------CT-G--...-----GG-GTGT------G-G--------- 7848
J.SE.93.SE7887          ----T--------A-----------CG---------C-G--.....................----C-----G-AC---......................GGACTC-----------GAT----------CT-GT-...--C--TG--TG-------G-G--------- 7952

K.CD.97.EQTB11C         ---------G--A-------------G----A----CG---.....................------C----C.....................AGGGGACTGAAGG--A-T-----------------CCT----...--C--GA---T-------G-------GA-- 7828
K.CM.96.MP535           ---------G--AA------A-----A----A-----G---.....................-----CC-----.....................AGGGGACTGAGGG----T------------------CT----...--C--TG----------------------- 7805
```

```
                                                                                                         Rev end (TAA) in some lineages
B.FR.83.HXB2                 TCAGCTACCACCGCTTGAGAGACTTACTCTTGATTGTAACG.....................AGGATTGTGGAACTTCTGGGACGC....................AGGGGGTGGGAAGCCCTCAAATATTGGTGG...AATCTCCTACAGTATTGGAGTCAGGAACTA 8645
Rev exon 2                    _Q__L__P__P__L__E__R__L__T__L__D__C__N__.__.__.__.__.__.__.__E__D__C__G__T__S__G__T__.__.__.__.__.__.__.__Q__G__V__G__S__P__Q__I__L__V__.__E__S__P__T__V__L__E__S__G__T__
Env                          F__S__Y__H__R__L__R__D__L__L__L__I__V__T__.__.__.__.__.__.__.__R__I__V__E__L__L__G__R__.__.__.__.__.__.__.__R__G__W__E__A__L__K__Y__W__W__.__N__L__L__Q__Y__W__S__Q__E__L__
01_AE.CF.90.90CF11697        ------------------------CA-A-------C-G--......................----C----------------A-AGCAGTCTCAAGGGACTGAGAC----------G-G----------CT-G--...-----T--GTCA------G---------- 8618
01_AE.CN.05.FJ051            ------------T-----------CA---C-----C-G--......................----C----------------A-AGCAGTCTCAAGGGACTGAGAC------------G-----------T-G-A...-----T--GTT-------G---------- 8648
01_AE.CN.06.FJ054            ---------G--------------CA-A-C--G-------......................---GC-------------G---AGCAGTCTCAAGGGTCTGAGAC------------G----------CT-G--...--C--T--GATA-------C--------- 8675
01_AE.HK.x.HK001             ----------T-------------CA--------------......................----C----------------A-AGCAGTCTCAAGGGACTGAGACT-----------G----------CT-G--...-----T--GTTA------G-G--------- 8027
01_AE.TH.01.01TH_R2184       ------------------------CA---------C-G--......................----CA---------------A-AGCAGTCTCAAGGGACTGAGAC------------G----------CT-G--...-----T--GTTA------G-CA-------- 7869
01_AE.TH.02.OUR769I          ------------------------CA---------C-G--......................----C----------------A-AGCAGTCTCAAGGGACTGAGAC-----------AG----------CT-G--...-----T--GGTA------G-C--------- 7850
01_AE.TH.90.CM240            ------------G-----------AC---------C-G--......................----CG---------------A-AGCAGTCTCAAGGGACTGAGAC------------G----------CT-G--...-----T--GTTA------G-C--------- 8219
01_AE.US.00.00US_MSC1164     ------------------------CA---C-----C-G--......................----C----------------A-AGCAGTCTCAAGGGACTGAGAC------------G----------CT-G--...-----T--GTTA------G-C--------G 7862
02_AG.CM.02.02CM_4082STN     -------------A----------TG---------C-G--......................----CC-C-------------A-AACAGTCTCAAGGGACTGAGACT----------------------CT-G--...-----T---TCA--C---G---------- 7844
02_AG.EC.x.ECU41             --T---------AA----------TG-----------GT-......................--A-C--------C-------A-AGCAGTCTCAAGGGACTGAGACT----------------------CT-G--...---G-T---TCA--C---G---------- 7728
02_AG.GH.03.GHNJ196          -------------A----------------------....................................-------A-AGCAGCCTCAAGAGCCTGCAACT----------------------CT----...-----T---ACA--C---G---------- 8692
02_AG.NG.01.PL0710           -------------A---A------G----------C-G--......................----C------------------AGCAGTCTCAAGGGACTAAGACT-------------G---CT----...--CA---GTCA--C---G---------- 7814
02_AG.NG.x.IBNG              -------------A----------A----------C-G--......................----C----------------A-AACTGTCTCAAGGGACTGAGACT--------G-------------CT----...-----TA--TCA--C---GT--------- 8172
02_AG.SE.94.SE7812           -------------A----------CG--C--------G--......................----C----------------A-.......................-------------------CT----...-----T---TC---C---G---------- 8035
02_AG.SN.98.MP1211           --CT---------A----------TG---------C-G--......................----CA---------------A-AGCAGTCTCAAGGGACTGAGACT------------T------C--CT----...-----T--GTCA--C---G---------- 7871
03_AB.RU.97.KAL153_2         ---T--------A----------------------C-G--......................------------------------.......................---------------------...------------------T---------- 7822
04_cpx.CY.94.CY032           ---------G--A-------A----------A-----G--......................----C-------------G....................ATA-------------G---CT----...--CT----G-T-------G-A-----G--- 8018
05_DF.BE.x.VI1310            ------------A-------------AC---A-----GT-......................------------------------.......................-----------------CT----...-GC----CG--A---------G------- 8051
06_cpx.AU.96.BFP90           ----T-------A-----------CGG--------C-G--......................----C------A-------------.......................-----AT----------CT-G--...--C--GA--TGT------G-A--------- 8693
08_BC.CN.97.97CNGX_6F        ----T-------------------CA-A--AC-GAC-G--......................--AGGG------------------AACAGCCTCAGGGGACTACAG-----------------T------CT-G-A...-G---TG-G---------G---T---G--- 7845
09_cpx.GH.96.96GH2911        ------------------------CA----A----C-G--......................----C------A--------A-.......................----T----C-GAT----------CT-G--...-----TGCG--A------G---------- 7817
10_CD.TZ.96.96TZ_BF061       -------------------------A----A----C----......................------------------------.......................-T---------G---A--------CT----...------------A-------T--------- 7997
11_cpx.GR.x.GR17             --CT--------AA----------CA-----------G--......................-----------AC------------.......................-----------GAGT--------CCT-G--...-----AGC------C---G-C--A-----G 7918
12_BF.AR.99.ARMA159          --------------------------G-------------......................------------------------.......................-------------------------...-----------------------G------- 8610
13_cpx.CM.96.1849            ------------------------CA---------C----......................----C----------------A-AGCAGTCTCAAGGGACTGAGACT----------G-G---------CCT----...--------GGTA------G---G------- 8091
14_BG.DE.01.9196_01          ------------------------CA---------C-G--......................----CA--------CT--------GGCAGTCTCAAGGGACTGAGACT----------G-G----------CT----...-----T--GTT-------G---------- 8171
15_01B.TH.99.99TH_MU2079     ------------------A-------G---C----C----......................----C-------------------AGCAGTCTCAAGGGACTGAGACT------------G----------CT-G--...-----T--GTT-------G-C--------- 8038
16_A2D.KR.97.97KR004         ----------------------T-GCA---C----C-G--......................----C----------------A-AGCAGCCTCAAGGGACTGAGACT------------G----------CT----...--C--T--GTT-------G---G------G 8032
17_BF.PE.02.PE02_PCR0155     --T---------AA---------------------------......................------------------------.......................------------A-------------------...------------A------------------ 7817
18_cpx.CU.99.CU76            ---------------------------G----C-G--......................----C----------T--------AGCAGTCTCAAGGGACTGAGACT----------G--T-----G--CCT----...-----T--GGT-------G-A-----GA-- 7974
19_cpx.CU.99.CU7             ------------------------CA---CA----C-G--......................----C----------------A-.......................--C--------G-G----------CT----...-----T--GTT-------G-------G--- 7708
20_BG.CU.99.Cu103            -----------------------------------C-G--......................----CA-------------------AGCAGTCTCAAGGGACTGAGACT------------G------G---CT----...---------GGG-------G---------- 8100
21_A2D.KE.91.KNH1254         -------------------------A----A--A---GT-......................------------------------.......................----ACA---G-G---------CCT----...--------G-----------G------G 7867
22_01A1.CM.01.01CM_0001BBY   ------------------------CA---------C-G--......................----C----------------A-AGCAGTCTCAAGGGACTGAGACT------------G----------CT----...---------GATA------G-G--------- 7833
23_BG.CU.03.CB118            -----------------------------------C-G--......................----CAT------------------AGCAGTCTCAAGGGACTGAGACTA---------G-G----------CT----...---------G-T-------G---------- 8094
24_BG.CU.03.CB378            --T--------------------------------C-G--......................--AGCA---------------A-AGCAGTCTCAAGGGACTGAGACT----------G-GT---------CT----...-----T--GTT-------G---G------- 8109
24_BG.CU.03.CB471            -------------------------------------G--......................---GCA---------------A-AGCATTCTCAAGGGACTGAGACT----------GAG----------CT----...-----T--GTT-------G---G------- 8112
25_cpx.CM.01.101BA           ------------------------CA-A-------C-G--......................----CA---------------A-AGCAGTCTCAAGGGACTGAGACT------------G----------CT-G--...-----T--GTT---C---G---------- 7916
27_cpx.CD.97.97CDKTB49       -------------A----------TG---------.........................................-------A-AGCAGTCTCAAGGGACTGAAACT----------G-G-------C--CT----...--CA----G-TA--C---G-A-G---G--- 7849
28_BF.BR.99.BREPM12609       -------------A--------------------------......................------------------------.......................-------------G-------------...---------G---------G---------- 8046
29_BF.BR.02.BREPM119         ------------T-------------------C-------......................----C-----G---------C---.......................-------------------------...--C------------------------- 7797
31_BC.BR.02.110PA            -------------A-----------A-A--A----C-G--......................--AGCG------------------.......................-----AT---T--G----T-AAA...-GC--AG-G---------G---TA--G--- 8138
32_06A1.EE.01.EE0369         -------------A----------CGG----------G--......................----C-------AC-------A-.......................-----GAT---------CCT-AAA...--C--GG-GTGT------G-A--------- 8271
33_01B.MY.05.05MYKL007_1     ---T---------A-----------A---------C-G--......................----C---------------CAAG............GGACTGAGACA-----------G----------CT-G--...--C--T--GTCA------G-A-T------- 8036
34_01B.TH.99.OUR2478P        -------------------------------------G-A......................---------------------A-.......................------------GG-----------...---------G---------T---------- 7811
35_AD.AF.05.05AF095          ------------A-----------CA---------CG--......................-A--C----------------A-AACAGTCTCAAGGGGTTGAGACT------------G------G---CT----...--C-----GTT-------G---G------- 7832
36_cpx.CM.00.00CMNYU830      ------------------------CA---------C-G--......................----C----------------A-AACAGCCTCAAGGGACTGAGGCT----------G-G------G--C-T----...---------GGTA------G-A--------- 7866
37_cpx.CM.00.00CMNYU926      ------------------------CA----A----C-G--......................----C---A--GA---------A-AGCAGCCTCAAGGGACTGAGACT---T-----G-G----------CT----...--C-----GGT-------G---------- 7848
38_BF1.UY.05.UY05_4752       --CT--------T-----------CA-A--A----C-G--......................---------A-C....................AGAGGACTGAAC------------AT------------G--...---------G--A----------G------- 8056
39_BF.BR.03.03BRRJ103        ---------G--A-----------CA-A--A----C-G--......................--A-------GC....................AGGGGGCTGACG--A---------------CT-CT-G--...--CA----G-TA------G---------- 8126
40_BF.BR.04.04BRRJ115        -------------------------------------G--......................------C-----------------.......................-----------CT------G----T----...-GC-----G-----C---------------G 8214
42_BF.LU.03.luBF_05_03       ---TT------------------------------C-G--......................------------------------.......................-------------G-------------...---------G-----------A---------- 8107
43_02G.SA.03.J11223          ------------------------CA---------C-G--......................----CA--A-----C---------AGCAGTCTCCAGGGACTGAGACT----------G-G---------CCT----...---------GTT-----------G------- 8149
U.CD.83.83CD003              -------------------------A-----------GTA..............................................AAGGGACTGCGAC-------------A--------CCT-G--...-----TG-G-T-------G---------- 8109
U.CD.90.90CD121E12           ------------------------CA---C-------G--..............................................AAGGGACTGAGAC-------------A---------CT-G--...-----TG-G-------------------- 7804
U.GR.99.GR303                ----------T--A----------CA-----------G--......................----C-------A------GACA..................CAG--------C-GAY----------CT-G--GA.-CCT-T--GTTA---G--T.---------G 7935
U.NL.95.U_NL_95_H10986_D1    ---------G--A------------------------GT-......................------C-A--G....................AGGGGACTGAGGG--A-T-------------------CCT----...-GC--TG-G------------------ 8551
CPZ.CD.90.ANT                GG-T------GAT--GC---AC--GT-------ACC-TTG-....................GCAG-CC-CC---AC--CT-CA-AATTACTTTCAGACTGTGCAACCATCT-GA-A-CAAT----GCACCCTC---...-CAA-AA-C.................... 8104
CPZ.CM.01.SIVcpzCAM13        GG-------GGA------C-AG----GC--GC-----TTG-......................--ACAGT-----AAGT------AACTGTTAATCAACATCCTCAGACT-ATACA---GAAA--G-CTCTACT-A--...GGGA-AA-C--A------G-AAGA---T-- 8250
CPZ.CM.05.SIVcpzEK505        --CT-------------C-----GT----CA--CC-G-G-......................----CCT-----------A---AGAACATCCACAAGGGACTGCAACTACTTAAT---CT-AGA-T-----T-----...GGAA-AA--GCT--C---G-AAG---G--- 8152
CPZ.CM.05.SIVcpzLB7          --CTT--------A----C-------G----A--A-CG-G-......................----CAT----GA--G------AATACACTCTGAAGGGACTGAGACT--T-----------ACT----CT-CA-...GGAA-T------------G-AA----G--- 8177
CPZ.CM.05.SIVcpzMB66         --CTT--------A----------------------A--GCT-......................----CAT-------GG------AGACCCTCCTCAAGGGACTGAGAC----AC-----------A-T-C-CCT-A--...GGCA-T-----A------G-A-----G--- 8290
CPZ.CM.05.SIVcpzMT145        --CT---------AA---------GTGCA-CTC-GA-TTG-......................--AC-AC----GA-A-------AATACAGCCTCAGGGGAGTACAACAAATAG-TACTCT---AC-CC-ACT-A-A...GGGAC---G---------CAACA---A-C 8178
CPZ.GA.88.GAB1               GG--------GA------C-AG----GC--GC-A---GTG-......................--ACAAC-CA-GACGT------ATCTAATACTTCACAGCCTCAGACT-CTAC----GAGA--GTGCCTGCT-G--...GGAA-TA-T--A------G-AA----G--- 8706
CPZ.TZ.00.TAN1               GG-TA-----GAC---TCA-A----CA---G---CC-CTG-ATCAGCCTGCAAGCACTGAAACA-GGGA-AATCAGCT---C--A-AGCCTAGTAATAGTGCATAGA-CTATCATA-T--GAG-T-G-C-GATCATT...G-GTGGAGCAGTA--ACTTA-GCTAGCT-- 8343
CPZ.US.85.CPZUS              --CT---------A---GAG---------A--A---C-G-G-......................----CA--AC--A-C-------AGAACATCAACAAGGGACTGCAACT-TT-AAT---CT-AGAGC-CGC--C-----...GGGG-AA-CGCT--C---GCAAG---G--- 8674
CPZ.CM.x.SIVcpzDP943         --CTT--------A----T------------CA--CC-G-G-......................----CAT-----A---------AGAACATCAACAAAGGACTGCAGT--CT-A----CAT-AG--G-C-CCT-----...GGAG-GA--GCA------G-AAG---G--- 8232
CPZ.CM.x.SIVcpzMB897         -----------------------------G---A--G-G-......................----CGC----GA--G------AGAAGTTCCTYGCAGGACTGAGACT--T-------A----A-C---C-T-AA-...-GCA-T--C---------G-A--A--G--G 8180
N.CM.02.DJO0131              --CT-------------C-----C----------CC-GCA-......................----C-C----G--C--------GAGCCTCATCAGGGGACTGCAACTACT-AAT---CT-AGA-C-CGC-T-----...GGAA-AA--GCA------G-AA-A--GT-- 8099
N.CM.04.04CM_1015_04         --CT------------------------------AAC-CA-......................----C-C----GA-C-------AGAGCCTCAACAGGGGACTGCAACTATT-AAT---CT-AGA---CGC-T-----...GGAA-A---GCA------G-AA-A--GT-- 8071
N.CM.04.04CM_1131_03         --CT--G--R--------S---------------AM--CA-......................----C-C----G--C-------AGAGTCTCAGCAGGGGACTGCAACTACT-AAT---CT-AGR-W-CGCYT-----...GGAR-A---GCA------G-AA-A--GT-- 8136
N.CM.95.YBF30                --CT-------------C-----C----------AC-G-G-......................----C-C-------C-------AGAGTCTCAGCAGGGGACTGCAACTACT-AAT---CT-AGA-C-C-C-T-----...GGAA-A--TGCA------G-AA-A--GT-- 8217
N.CM.97.YBF106               --CT------------GC-----C----------CA---G-......................----C-C----GA-C-------AGAGTCTCAGCAGGGGACTGCAACTACT-AAT---CT-AGA-T-CGC-T-----...GGAA-AA--GCA------G-AA-A--GT-- 8175
O.BE.87.ANT70                GG-CT-------T------CA-----GCA-CAGGGA-CCA-.....................-A-G-GA-CAGCTA----A-G-TTGGACTGTGGATCCTAGGGCAG-A-ATAATTA-T-TTTG--G-AT---TGCA...GC-G-AAC---A--C---CTA--A---T-G 8713
O.CM.91.MVP5180              GG-CT-------T------CA-----A-A-CAGGGA-CCG-.....................---C-GA-C--CTAC-------TGGGACTGTGGATCCTGGGACAA-A-ACAATT-----TTGT-G-CT---TG-A...GC-G-AA-G--A------CTA--A---T-G 8739
O.CM.96.96CMABB637           GG--T----G--T------CAG----GCA-CAGGGA-CCG-.....................-ACG-GACCAGTTAC-------TTGGACTGGAGATCCTAGGGCAG---ACAATT-----TTG--G-C-AACAG-A...GC-G-GR----A--C---CTA--A---T-G 8206
O.CM.98.98CMA104             GG-CT-------T------CA-----GCA-CAGGGACCCA-.....................--AG-GA-CAGC-A-------ATTGGACTGTGGATCCTAGGGCAG-A-ATAATT--T-GTTG----CT---TAAA...GC-G-AA----A--C---CTA--A---T-G 8217
O.CM.99.99CMU4122            GG--T-------T------CA-----GCA-CAGGGA-CCA-.....................-C-G-GA-CAGC-A--------TGGGACTGTGGATCCTAGGACAG-A-ATAATTA-T--TTG--G-AT---TGCA...GC-G-AG----A--C---CTA--A---T-G 8189
O.FR.92.VAU                  GG-TT-------T------CA-----GC--CAGAGA-CCA-.....................-A-T-GA-CAG--AC-------TTGGACTGTGGATCATAGGGCAG---ACAATT-----TTG--G-CTC-TTAAA...GC-A-AA----A--C---CTA--A---T-G 8279
O.SN.99.SEMP1299             GG----------T------CA-----GCA-CAGGGA-CCA-.....................-C-G-GA-CAGC-A--------TTGGACTGTGGACTCTGGGACAG-A-ATAATTAGT--TTG--G-CT---CATA...GC-G-AA----A--C---CTA--A--GT-G 8781
O.US.99.99USTWLA             GG--T-----GAT------CAG----ACA-CAGGGA-CCA-.....................-A-T-GA-CAGC--G-------TTGGACTGTGCACTCTATGGCAA---ATAACT------TG--G-CT---TGCA...GC-G-AAC---A--C---CTA--A---T-G 8190
```

```
                          Rev end
B.FR.83.HXB2              AAGAATAGTGCTGTTAGCTTGCTCAATGCCACAGCCATAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAGTAGTACAAGGAGCTTGTAGAGCTATTCGCCACATACCTAGAAGAATAAGACAGGGC..TTGGAAAGGATTT......................... 8787
Rev exon 2                K  E  *
Env                       _K__N__S__A__V__S__L__L__N__A__T__A__I__A__V__A__E__G__T__D__R__V__I__E__V__V__Q__G__A__C__R__A__I__R__H__I__P__R__R__I__R__Q__G_##_L__E__R__I__.__.__.__.__.__.__.__.__._
A1.GE.99.99GEMZ011        ----G-------A---ATC--A-AG--A---T---A------------G-T-------------------A---G--G----TT--------------TA---------G-----C---------..GCA--G-A-GC--.......................... 7983
A1.KE.00.KER2008          --A-C----------------TG--A-------A----T--------GCT-------------C-----A---G----A--AT-G----G---T--AT-A----------------T--------T..--T------GC--.......................... 7993
A1.KE.00.KNH1144          --A-C-------A----T---G-TG--A-T-T---A------------GCT-----------A-C-------A---G-------AT-G--------T---T-----------------C---------..--A-----GC--.......................... 8040
A1.KE.00.KSM4024          --A-T-------A-C-AT---G-T---A---T---A----T--------GCT---------------------A---C--------T-G-------------T-----------------C---------..--A-----GC--.......................... 7937
A1.KE.00.MSA4069          --A-T-------A---AT---G-TG--A-T-T---A----T--------GCT-----------A---------A---G----A--AT-G-------------T-----------------C---------..--T----AGC--.......................... 8016
A1.KE.00.NKU3005          --A-T-------A-C-AT-----TG--A---T---A------------GCT----------------------A---G----A--AT-G-------C----T-A-------------G-C--G-------..--C-----GC--.......................... 8019
A1.RU.00.RU00051          ----G-------A---AT---G-AG--A---T---A------------G-T----------------------A---G----A---T-----------YYTT---------K-----C---------..--C----A-GC--.......................... 8082
A1.RW.93.93RW_024         --A---------A---AT---T-TG--A---T---A------------GCT-------------------------G----A---T-G-------T-G-T----G---------G--C---------..--T-----GC--.......................... 8013
A1.SE.95.SE8891           --ACT-------A---AT-----TG--A---T---A----T-A------G-T-------------------A---G-------TT--------------TT--..-------------GC-A--------..--AA----GC--.......................... 7971
A1.SE.95.UGSE8131         ----T-------A----T---T-TG-CA---T---A------------GCT-------------------A---G----A--AT-G-------------T---T---------------T---------..--T-----GC--.......................... 8153
A1.TZ.01.A173             --A-G-------A---AT-----TG--A---T---A------------GCT--------------------A---G----A--AT--------------TA-A-----------------C---------..--A-----GC--.......................... 8017
A1.UA.01.01UADN139        ---GG-------A---ATC----AG--A---T---A------------G-T-----------------------G----A--TT--------------T-TA----------G-----C---------..GCA----A-GC--.......................... 8004
A1.UG.92.92UG037          --A-T-------A---AT-----TG--A---T---A------------GCT--------------------AC-------A-GCT-G-------------T-A----------------C--G------..--C-----GC--.......................... 8818
A1.UG.99.99UGA07072       --A-T--A----A---AT-----TG--A---T---A------AC-------T--------------------A---G----A--CT--------------T-----------------C-------C-..--A-----GC--.......................... 8014
A2.CD.97.97CDKTB48        ----C-------A----A-----TG--A---T---AG-------------T--------------------A---G----A------C--------C---A-----------------C---------..C-C-----GCG-.......................... 8118
A2.CY.94.94CY017_41       ------------A-------AT-TG--A-T-T---AG-------------T--------------------A---GG---A-----TC----------T-A-----------------C---------..--A-----GC--.......................... 8180
A.SN.01.DDI579            --A-GC------A-CTCT-----T-----T-----A----T---------T--------------------A---G----A-----G-------C---T---T-----------G-C---------..--A-----GC--.......................... 7979
A.SN.01.DDJ369            --A---------A--TCT-----TG-C--T-T---A---A------G-G-T-----------------------G----A-----G-------C---T-----------------T---------..---------GC--.......................... 8030
A.SN.96.DDJ360            --A-GC------A--TCT-----T-----T-----A----T---------T-----------------------G----A-----G-------C---T-----------------C---------..--A-----GC--.......................... 7961
A.CD.02.02CD_KTB035       --A-G-------A----T--AT-AG--A--CT---A------------C-T------------G------A---G----A--AT-G-------T---TTA--G-------------C---------..--C----AGC--.......................... 8807
A.CD.97.97CD_KCC2         --A-C---------C--TGCAT-T-----TTT---A---A----T-----T------------C-------------G----A--AT-G-------T---T-A-----------------C---------..--C----AGC--.......................... 8807
A.CD.97.97CD_KTB13        --A-C-------A----T--AT-TG--A-TCT---A--------------T-----------A-----------G--CA--AT-G-------T---T-A-----------------C---------..--C----AGC--.......................... 8084
B.AR.04.04AR151516        ----------------T-----T--C-------T------------------------C-----------T-----A-CAT-G----G-------T-----------------------------..--A-----TC--.......................... 7996
B.AU.87.MBC925            -------------------------T--------------------------------------------T-----A--AT--A----------AT----------C------------------..--T-----GC--.......................... 8806
B.BO.99.BOL0122           -------------------------C-------T---A-----------A--A---------A-C-----A---------A------A----------T----------C------------------..--A-----GC--.......................... 7988
B.BR.03.BREPM2012         ------------A-----T-----T---TA------T---------------------------A-C---------------A--A--G----------T-----------------------------..--A-----GC--.......................... 8232
B.CA.97.CANB3FULL         ---------------------A-T-------T-----------------------------A-----------------A-----T------G---T---G---------------------------..C-------C--.......................... 8060
B.CN.05.05CNHB_hp3        ------------A---------T--C--------------------G---A------------A------------C---A-------G-----------T-A---------C------C--------..C-C----AGC--.......................... 8803
B.CO.01.PCM001            ----------------------T----T------T-----------------------------A---------G------A-----------------------A--G--------------------..--A-----GC--.......................... 7970
B.GB.83.CAM1              -G--------------T---T-TG--A---T---T-----T------------------A----------------------G--A------------------T-------------------A---..--A------C--.......................... 8787
B.GB.86.GB8               ------------A----------T---A-------T-----T-----------------------------------A------A-----------T----------C------------------..--A-----GC--.......................... 8778
B.GE.03.03GEMZ010         ----------------AT-----T-----TG----------------------------AA--G-------------A--AT-G-----------AT-----------------------------..--T-----TC--.......................... 7964
B.IT.05.SG1               --------------------------A---T---T-------------------------------------AC---AG-A-----------------T---------GT-----------------..---------GC-A.......................... 8472
B.JP.05.DR6538            ----G------AA---A------T---A-------T----------T---------A---------------------G--A-G---G-----------T----------------------------..--A--------.......................... 8823
B.KR.05.05CSR3            --------------------T--G-C--------A------A-------------------------------C------A-----G--------C-C-TT---------------------------..--A-----GC--.......................... 8850
B.NL.00.671_00T36         -G-GG------------T---T-T-------T---T-----T-----------C----------------AC-A------A-----T-------C----T------------------------------..--A-----A---.......................... 8348
B.RU.04.04RU128005        ------------A------------C---------------------------------------A-------T--T-----A-----G-----------AT---T--------C-------------..---------GC--.......................... 8259
B.TH.00.00TH_C3198        --A-------------T------------------TG------------AA-----------A-----------G--T-A-----A-----------T----------C------------------..--A------C---.......................... 7988
B.UA.01.01UAKV167         ----G-------A----T---T-T-----T-T---T-------T-----------------A-A--------A---C----AC-AT-G-C----------T-----------------------------..C-C------T---.......................... 8051
B.US.04.ES10_53           ------------A-C--T---T-T-------T---T-----T------------------------------A-------A---T---C-----------------------------------A---..----------G--.......................... 8765
B.US.99.PRB959_03         -------------------------C-----------------------------------A-----------T-----A--AT-G-----------C-T---------------------------..--T-----GC--.......................... 7985
C.AR.01.ARG4006           --A--G------A----TC----TG--A---T---A--------------A--A---------A-----------A-----C--AT---GGA-AT---CT-TA-------------------------..--T---GCAGC--.......................... 7978
C.BR.04.04BR013           --A--G------A----TC----TG--A---T---A--------------A--A---------A---------T--A-------AT---G--------C-T-AG---------------------T---..--T---GCAGC--.......................... 8299
C.BW.00.00BW07621         --A--G------A----T---T-TG--A---T---A------------------A---------A---------CGA-----A--AT-----------CT--A-----------------------A--T..--T---GCAGC--.......................... 8143
C.CN.98.YNRL9840          --A--G------A----TC----TG--A---T--------T---------A--A---------A-----A--T-----T-CA-CAT------------C-----------------------------..--T---TTAGC--.......................... 8009
C.ET.02.02ET_288          --A--G------A----TC----T---A-------A----T---------A--A--------AA--------T--------A--TT---G--------CT--A--GC----------------------..--T---GCAGC--.......................... 8013
C.GE.03.03GEMZ033         --A--G------A---ATC----TG--A---T---A--------------A--A---------A--------T--A-----A--AT--------G---CTA-A--G---------------------..--T---GCAGC--.......................... 8015
C.IL.99.99ET7             --A--G------A----TC----TG--A---T---A---A----------A--A--------AA----------------A---G---G-----C-CT--A-------------------------..--T---GCAGCA-.......................... 7974
C.IN.99.01IN565_10        --A--G------A----TC--T-TG--A---T---A--------------A--A---------A--C-----T--------A--AT-------------CTA-A-T------------------------..--T---GCAGC--.......................... 8185
C.KE.00.KER2010           --A-GG------A----TC----TG--A---T---A--------------A--A---------A--------T--A-----A--AT-------------C-TTA---------C---------------..C-T---GCAGC--.......................... 7997
C.MM.99.mIDU101_3         --A--G------A----TC----TG--AG--T-------A----------A--A---------A------C--------CA--AT-------------C---A---------------------------..--T---GCAGC--.......................... 8167
C.MW.93.93MW_965          --A-GG------A----TC----TG--A---T---A---A----------A--A--------AC------------------A--AT-C-G-------CT--A---------------------------..--T--G-CAGC--.......................... 7983
C.SN.90.90SE_364          --A--G------A----TC----TG--A---T---A---A----------A--A--------AA--------T--------A--ATC-----G-----CA--A----------------------A..--T---GCAGCG-.......................... 7932
C.SO.89.89SM_145          --A--G------A----TC----TG--A---T---A---A----------A--A---------A--------T---C----A--AT--------G-G-C---A---------C-----------------..--T---GCAGC--.......................... 8034
C.TZ.02.CO178             --A--G-----CA----TC--T-TG--A---T---A------------------A---------A--------T--------A--AT-G-------.......................--C-----------..--T----CAGC--T.......................... 7992
C.UY.01.TRA3011           --A--G------A----TC--T-TG----------A--------------A--A---------A-----------C-----A--AT---G--------CT--A---------------------------..--T---GCAGC--.......................... 7985
C.YE.02.02YE511           --A--G-----------TC--G-TG--A-T-T---A--------------A--A--------AA--------A--------A--AG---------C-CT--A---------------------------..--T---GCAGCA-.......................... 8000
C.ZA.04.04ZASK164B1       --A-GG------A----TC----TG--A---T---A---A--AC----------A---------A--------C----G-T-A--AT------------C---A--------GC---------------..--T---GCAGC--.......................... 8237
C.ZA.05.05ZAPSK240B1      --A--G---------A-C----TG--A-------A-------------A--A---------C-------C-CA-----A-CTT--------G---C---A-----------------------------..-G----GCAGC--.......................... 8187
C.ZM.02.02ZM108           --A--G-----------TC----TG--A---T---A------------------A--------AA--T-----T-GC-----A--AT-------G-----C---A---C---------------------..--T---GCAGCG-.......................... 8845
D.CD.83.ELI               -G---C------AG-------T-TG------T---A---T--A------------------------A--------A--A-----A-------C-----G--TTA------C------------------..--A-----TC--.......................... 8323
D.CM.01.01CM_0009BBY      ------------A--------TAT--C--------AG------AG-----------------------------A------C---GG--A------T------GC--TT--T-----C-C----------A--T..--A-----GC--.......................... 8012
D.KE.01.NKU3006           ------------A----------TG----------A---A-----A--A--------G-------------C-------A-------A--T-GT-----------TTA------C-C-C----------------..---------GC--.......................... 8030
D.KR.04.04KBH8            ------------A----A---T-TG----------A-----A---------------------------------------G-A------T------G---TT-----------------------------..--C-----GC--.......................... 8700
D.TD.99.MN011             -----------CA----------TG-C--------A-----A--------------------------A--------A-------A----GC------------TTA------C-------------A..--A----AGC--.......................... 7977
D.TZ.01.A280              ------------A----T---------T------A------A-------------------------A--------A--A-------A-----T------G---TT-------C------G-------------..---------GC--.......................... 7975
D.UG.99.99UGK09259        ------------A--------T-TG----T-----AG-----AG-----------------A---------A-C------A----------------T-------.....TA--G---C---------------A--..GC-------GC--.......................... 7983
D.YE.01.01YE386           ------------A----T---T-T---A-------A----T-----A---T-------------A-C------A---C---A--T-G-----------ATT--------C----------------------..---------T---.......................... 7943
D.YE.02.02YE516           ----T-------A--------T-T--C--------A-----A------------------A--------AA-------A----------A-G--G-------------TTA-------C-----------A--T..--A-----GC--.......................... 7967
D.ZA.90.R1                ------------A----------TG--A-T-T---A-------AC-----------------------------------A-------C-----G---TTA--G---CGT------C-----------..--A-----C---.......................... 8141
F1.AR.02.ARE933           -GA-T-------A--------T-T---A-------A---T-----------A--A---------A----------CT--G---A-----GT------G--T-TA------------------------------..--------AGC--.......................... 8057
F1.BE.93.VI850            ------------A-C---C--T-T---A-------A---T-----------A--------------AA---------TT-G---A-----G-------G---T-A--------------------------..GCA-----GC--.......................... 8064
F1.BR.01.01BR125          ------------A---A------T---A---G--A---T--------------------------A-----------CTT-G---A----GG-------G---TTA-------------------------..ACA-----GC--.......................... 8223
F1.ES.x.P1146             ------------A--------T-T---A---T---A---T--------------C--A---------------TT-G---A--AT-G---G-----AT-A--G------------C------------..--T-----ATC--.......................... 8094
F1.FI.93.FIN9363          ------------A--------T-T---A-------A---T------------------------A---------CTT-G---A-----GT------G---T-A---------------------A-A..G-A-----GC--.......................... 8083
F2.CM.02.02CM_0016BBY     --A-T-------A----TC----T--CA-------A---TG---------A-----------A--C------A-TC-G---A-----G-------------T---------------------------T..--C-----GC--.......................... 7946
F2.CM.95.MP257            ------------A----TC--T-GG--AGA-----A-------------A--A--------AA-C------A-TT-G---A-----G-------G---T-A------------------------T..C-T----AGC--.......................... 7980
F2.CM.97.CM53657          ------------A--GATC----T---A-------A---TG--------A-----------AA-C---------TC-G---A-----G-------------T-------------------------T..GCA-----GC--.......................... 7958
G.BE.96.DRCBL             ------------A---AT-----TG--A-A-T---A------------A-CT---------------------C---A-----G-------G---T-A------------------------------..--A----AGC--.......................... 8747
G.CM.01.01CM_4049HAN      -----------CA---AT-----TG--A-ACT---A------------A-TT---------------------C---------G-------------T-------------------------A---..--A----AGC--.......................... 7997
G.CU.99.Cu74              ------------A---AT---G-TG--A-AGT---AG-----------A-CT-----------------A--------A--A------G---T---T-A-------------G--------A---..--T-----GCGC.......................... 8427
G.ES.99.X138              -----------CA---AT---T-TG--R-ATT---A------------R-CT---------A-C-------T---GG---A---T-G---R-------T-A--------GT------------R--A..---------A----.......................... 8243
G.GH.03.03GH175G          ------------A----T-----TG--A-AGT---A------------A-CT----------------------A----A-----G-C-----------TTA-T------C---G--------A---..--T----AGC--.......................... 8854
G.KE.93.HH8793_12_1       ------------A---AT-----TG--A-A-T---A------------A-CT-----------------A--------C-----T---G---T---TTA------C-C---G--------A---..--A----AGC--.......................... 8178
G.NG.01.01NGPL0669        --A-T-------A----T-----TG--A-AGT---A------------A-CT-----------------T---C---A-----G-------------TTA-------------G--------A---..GCA----AGC--.......................... 8001
G.PT.x.PT2695             -----------CA---AT---A-TG--A-AGT---A------------A-CT--------A-C-----------G---A---T-G---G-------T-A--------GT------------A---..---------AC---.......................... 8819
G.SE.93.SE6165            ------------A----T-----TG--A-AGT---A------------A-TT---------------------C---A-----------------C-T-A--------C---G--------A---..---------AGC--.......................... 8226
H.BE.93.VI991             ------------A----T--A--T---A-T-----A-----T-----------A---------A--------C----G---A-------G-------C-T-------------------------------..--T------GC--T.......................... 8187
H.BE.93.VI997             ------------A---AT-----T--CA-------A----T------------A--------AA--------A----G---A-------G-----G---T--------------------------------..C-T----------T.......................... 8102
H.CF.90.056               ------------A--GAT-----T--CA-------A--------------A-----G--A-----T-A----G---A-------G-----------T-------------------------------..--T-----A-GC-T.......................... 8107
J.CD.97.J_97DC_KTB147     ------------A----T-----T-----------A-------------CA--A--------AA-C------A----------GATA-T---------C-TA---------------------A.................................... 7976
J.SE.93.SE7887            ------------A----T-----T---A-------A-------------A--A--------AA-C------A---C---A-----T-----------TT-------------------------T..--A----AGC--.......................... 8094
K.CD.97.EQTB11C           -----------CA---A------T---A-T-----A-------------A--A--------AA-C-----GA-----T-CA------T-------T-A-TT---------------------------..--T-----AC---.......................... 7970
K.CM.96.MP535             -----------CA----------T---A-------A-----------GA--A--------AA---------A---G---A-------T-------T-A-TT-------------------------T..C-C-----AGC--.......................... 7947
```

```
                           Rev end
B.FR.83.HXB2               AAGAATAGTGCTGTTAGCTTGCTCAATGCCACAGCCATAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAGTAGTACAAGGAGCTTGTAGAGCTATTCGCCACATACCTAGAAGAATAAGACAGGGC..TTGGAAAGGATTT......................... 8787
Rev exon 2                 K  E  *
Env                        _K__N__S__A__V__S__L__L__N__A__T__A__I__A__V__A__E__G__T__D__R__V__I__E__V__V__Q__G__A__C__R__A__I__R__H__I__P__R__R__I__R__Q__G_##_L__E__R__I__.__.__.__.__.__.__.__.__._
01_AE.CF.90.90CF11697      --A-C-------A---CT---T-TG----T-T---AG---------G-G-T--------------------------GA-------G------C--AT-----------------C---------.--A------GC--......................... 8760
01_AE.CN.05.FJ051          --A-T-------A--TCG-----TG--A-------A---------A-G-T---------------------C----A--AT---G-----CT---T-----------------C--------T..C-T-----AGC--TGCAATAAGAGCT............. 8803
01_AE.CN.06.FJ054          --A-T------CA--TCT---T-TG----TT----A------A---A-G-T-----------------A---C-----------G-----C----T-----------------C---------.--A------GC--......................... 8817
01_AE.HK.x.HK001           --A-T-------A--TCT-----T---A-T-----A-----------A---T---------------------C----A-------G--G-----C-TT----------------CC--------.--A-------C--......................... 8169
01_AE.TH.01.01TH_R2184     --A-T-------A--TCT-----TG--A-T-----A-----------G-G-T----------------C--C----A-------G---------T-----------G-----T---------.--A------GC--......................... 8011
01_AE.TH.02.OUR769I        -GA-T-------A--TCT---T-TG----T-T---A-----------G-G-T----------------C--C----A-------G---------T-----------------C-----A---.--A------GC--T....................... 7993
01_AE.TH.90.CM240          --A-T-------A--TCT-----TG----T-----A-------C---G-G-T-------C---------------C-----------G-----C----T-----------------C---------.--A-------C--......................... 8361
01_AE.US.00.00US_MSC1164   --A-T-------A-CTCT-----TG----T-T---A------A---G-G-T---------------------C----A-------G-----C---------------C-----T---------.--A------GC--......................... 8004
02_AG.CM.02.02CM_4082STN   ------------A---A-C--T-TG--A-A-T---A------------A-CT-------C--A---------A---G----A-----G-------------TA-T-------C------C---------.C-A------GC--......................... 7986
02_AG.EC.x.ECU41           ------------A---AT-----T---A-A-----A------------A-C--------C--A---------C---------A--AT-A--G----------A-------------G--C---------.C-C------GC--......................... 7870
02_AG.GH.03.GHNJ196        -G----------A----T-----TG--A---T---A------------A-CT-------C--A--C------A---G-----A--AT-GC----------T--A-----------------C--------T..C-T------GC--......................... 8834
02_AG.NG.01.PL0710         --------C---A---AT-----TG--A-A-----A------------A-CT-------C--A--A------A---G---GA-----G--G----------A-----------------C---------.--A-------C--......................... 7956
02_AG.NG.x.IBNG            ------------A---AT-----T---A-A-T---A----T-------A-CT----------A-C-------A---G----A---TGG--------------A-----------------C---------.--C------GC--......................... 8314
02_AG.SE.94.SE7812         ------------A----T-----TG--A---T---A----T-------A-TT-------C--A---------C--------A-----G-----------T-A-----------G-----C---------.--T------GC--......................... 8177
02_AG.SN.98.MP1211         ------------A---AT-----TG--A-A-----AG-----------A-CT----------A---------A--------A--A--G-------A---T--A-----------------C---------.--A------GC--......................... 8013
03_AB.RU.97.KAL153_2       ----G-------A---ATC--A-AGG-A---T---A-------------G-T--------------------A---G----A--TT------------G--TA-----------G-----C----------.GCA----A-GC--......................... 7964
04_cpx.CY.94.CY032         ------------A---AT---T-T---A-------A-----------------A---------A---------C-----GA-----------------T--A-----------------C---------.C-T-----AGC--......................... 8160
05_DF.BE.x.VI1310          ------------A----------T---A-------AG---T--------------------A----------CTT-G---A----G-------------T-A-------------------------.--A------GC--......................... 8193
06_cpx.AU.96.BFP90         C-----------A----T--AT-TG----AG---A------------A-CT--------------GC-------------A--AT--T-------T---TTA--G-----------------A--A..--T-----AGC--......................... 8835
08_BC.CN.97.97CNGX_6F      --A--G---A--A----TC--G-TG-------T---A---------------A--A----------A-----A-CAG----------AT------------C-A-A--G---------------------.--T---GCAGC--......................... 7987
09_cpx.GH.96.96GH2911      ------------A----T-----T---A-------A----T-----C-----------------------------TT-G---A-G---G----------C-TT--T-----------------C---------.--T------GC--......................... 7959
10_CD.TZ.96.96TZ_BF061     ------------A----------TG--A---T---A----A--------GCT--------------------A--------A-----GTC----------TTA---------C-C--------------.-----G---GC--......................... 8139
11_cpx.GR.x.GR17           ----G-------A----T-----T-----------A---------------A--A----------A-------G----C---TA------TA--------A-T-A-T-------------------------.--T------GC--......................... 8060
12_BF.AR.99.ARMA159        ------------A----------T---A-------A----T--------G------------AA--T------CTT-G---A--AT-G--------G---T-A--GC----------G----------A..---------GC--......................... 8752
13_cpx.CM.96.1849          --A-T-------A---C------TG--AGT-----A--T---------------T-----------A-------GAC--G----A--AT-G-G----GC----T-----------------G--C---------.--C------GC--......................... 8233
14_BG.DE.01.9196_01        ------------CA---AT---T-TG--ATA-T---A------------A-CT---------R-A-C-------A----G---A--AT-G----G---T---T-A---------C------------A---.------R-AKC--......................... 8313
15_01B.TH.99.99TH_MU2079   --A-T-------A--TCT-----TG--A-T-----A-----------G-G-T---------------------C----A-------G-----C----T--G-----------------C--------T..--C------GC--......................... 8180
16_A2D.KR.97.97KR004       ------------A----T--AT-T-----------A-----------------T----------A---------A---------A-------C----------AT-A-----------------C---------.--A--G--AGC--......................... 8174
17_BF.PE.02.PE02_PCR0155   ------------A----------T---A-------A----T-------------------------A----------CTT-G---A---T-G----------CTT-A---------------------G-------.C-C------------......................... 7959
18_cpx.CU.99.CU76          ----G-------A---AT-----TG--A-AGT---A------AC----A-CT--------------------A--------A--A--G------------TTA-----------------C----------.GCC-----AGC--......................... 8116
19_cpx.CU.99.CU7           --A-C-------A----T-----T---A-T-----A----T------G-G-T-----------C--------A--A-----A-------G-----G---T-----------------C---------.--A------GC--......................... 7850
20_BG.CU.99.Cu103          ------------A---AT-----TG--A-AGT---A------------A-CT----------AA---------A--------A---GA-T-------T---TG--------------G--------A---.--T------GCGC......................... 8242
21_A2D.KE.91.KNH1254       ------------A----T-----T-----------A-----------C--A---------------C------A--T-----A-G---G-C------T-C-T-A-----------------C---------.--C------GC--......................... 8009
22_01A1.CM.01.01CM_0001BBY --A---------A---C---A--TG----T-----A---A-------G-GAT---------------------A--AC----A-------G------T---TG--T-------C-------T---------.----------GC--......................... 7975
23_BG.CU.03.CB118          ------------A---AT-----TG--A-AGT---A------------A-C------------A---------A--------A---GA-T-------T---TG-----------------------A---.--T------GCGC......................... 8236
24_BG.CU.03.CB378          ------------A---AT-----TG--A-AGT---A------------A-C------------A---------A--------A---TA-T-------T---TG--------------G--------A---.C-T------T-AC......................... 8251
24_BG.CU.03.CB471          ------------A---AT---T-TG--A-AGT---A------------A-C------------A---------A--------A---GA-T-------T---TG--------------G--------A---.--T------GCGC......................... 8254
25_cpx.CM.01.101BA         ------------A---AT---G-TG--A-A-T---A------------A-CT--------------------G-A-----A-----G----------TAT--------------------------------.--C-----AGC--......................... 8058
27_cpx.CD.97.97CDKTB49     ------------A---AT-----TG--A-A-T---A------------A-CT----------------------------A-----G-----------T-A-------------G--------A---.--T----TAGC--......................... 7991
28_BF.BR.99.BREPM12609     -----------------------T-----------T-----------------------------A---------A--T-----A--AT-----------------A---------C-------------------.--A------GC--......................... 8188
29_BF.BR.02.BREPM119       ----------------T------T---A---T---T-----------------------------------G-------C--G---A---T-G-C------G---TT----------C-------------------.--A------GC--......................... 7939
31_BC.BR.02.110PA          --A--G------A----TC----TG--A---T---A--------------A--A---------A-------------------A--AT.--C---------CTATA---------------TA-GACA..----..--T---GCAGCA-T....................... 8278
32_06A1.EE.01.EE0369       ------------A-A-AT---A-TG--A-A-T---A--------------CT----------A--------------G---A---T--T-------T---TTA-T-------------------------A--A..--A----TAGC--......................... 8413
33_01B.MY.05.05MYKL007_1   --A-T-------A--TCT-----T-----------A-----------G-G-T---------------------C----A-------G-----C----T-----------------C--------T..--A------TC--......................... 8178
34_01B.TH.99.OUR2478P      ----------------------------------T-------------------------------A---------------A-----A-----------T-----------------C---------.--A------GC--......................... 7953
35_AD.AF.05.05AF095        --A-C-------AC---T---G-TG--A--TT---A--------------GCT----------T---G-----A---G---A--TT-G-------------T-A-------C--------T---------.C-T-----AGCA-......................... 7974
36_cpx.CM.00.00CMNYU830    --A-C-------A-C-CT-----TG----T-----AG--A-------A-G-T----------A---------------C---A--A--G-------------TTA--G------C------C--------A..C-A------GC--......................... 8008
37_cpx.CM.00.00CMNYU926    --A-T-------A-C-AT-----TG--A-T-----A----T------A-G-T---------------------A---C---A----------------------T-----------------C--------T..--T-G--A-GGC-T....................... 7991
38_BF1.UY.05.UY05_4752     -G----------A----------T--CA-------A----T---------------------A--------G---TT-G---A--AT-G--------G---TTA-----------------------------.--C------GC--......................... 8198
39_BF.BR.03.03BRRJ103      ------------A----------T---A---T---A----T---------------------A---------CTT-G---A--TT-G-----G---C-T------------------------------.--T------GC--......................... 8268
40_BF.BR.04.04BRRJ115      ----------------------G---A---T---T-----------------A---------A---------C-A-----A--AT-G---G-GC--C-TA--T--C---------------A---.--T-------G--......................... 8356
42_BF.LU.03.luBF_05_03     ------------A----TC----T-----------T-------------------------------------Y-T-----A--AY-G---G--R-C-T---T-------C-------------------.--A------GC--......................... 8249
43_02G.SA.03.J11223        ------------A---AT---G-TG--A-A-T---A--------------GCT---------------------C----A-----------------AT-A--------------G--------A---.--T-----AGC--......................... 8291
U.CD.83.83CD003            ------------A----T-----T-----------A----T-----------A---------AA---------G-G---A--AT-----G------T-A-T--------------------------T..--C------GC--......................... 8251
U.CD.90.90CD121E12         --A---------A----T-----T-----T-----A----T---------------------A--------T-------A---T-G---G-----C-T---T---------G-------------T..--C------GC--......................... 7946
U.GR.99.GR303              C----G------AC---T-----TG--AT------A----T--------G-T-------------C-----AC--G--GA--AT-G-------------T-----------------C-----A---..--C------GC--......................... 8077
U.NL.95.U_NL_95_H10986_D1  ------------A---A------T---A-------A-------------------A--A-------AA-------A---G---A-----T------G---TG-------------------------T..--T------GC--......................... 8693
CPZ.CD.90.ANT              -GA-C-GAGATCA---AGAACA-TG-CAGACTT--T--TTGG----GG--AAAA--------CA-AC-TCT--CTC-C---ACTATAGTC---ATC--AA-GG-AG-------GC-C--C-----A--G.--------TTGCA-T......................... 8247
CPZ.CM.01.SIVcpzCAM13      -GA-C-------ACA-----AT-AG----A-----T--T-------GA--A--C---------AA-C---AG--C---G---AT--TA-TC---ATC--AG--A-------CC-GC-C---------A---..C-A--G--A-C--TACTTTAAATCTTGCTTTAAAATAGT 8418
CPZ.CM.05.SIVcpzEK505      --A-T-------A-A-A------TG--A-T-----TG-T-------A--A-----------AA-C-----GC----G---A--ATAG-AC-G-GA--A-TA-----------G-----C--------A--T..C-A-----AGCAC......................... 8294
CPZ.CM.05.SIVcpzLB7        --A-C------CAC---------TG--A-A-----A------------------A-------AA-C-----A--------A--AT-G-----GA--ATTA--T-----------G--T-------A--T..--C------GC--......................... 8319
CPZ.CM.05.SIVcpzMB66       --A-C------CA-A--------TG-CA-A-----T----------A--A--A---------AA--------A---C--GA--TT-G-----GC--A-TGA---C-----------C-----A---..C-T-----AGCA-......................... 8432
CPZ.CM.05.SIVcpzMT145      ------------CA-A-----AT-AG--A-T--C--T--T-------GA--A--A--T-----AA-CT------C-A--AC-C-CTTGG-C--G-GC---TTA--T-------------T-----A---..--A-----AGC--......................... 8320
CPZ.GA.88.GAB1             --A-T-------A-A---C----TG----A-----T--T-------A--A--T---------AA-A-------CTT-T----TTA-ACT----AT----A-AA----C-----GC-C--------A--G..C-------AGC--TACTTTAAGCAC.............. 8860
CPZ.TZ.00.TAN1             -GAGT-TTGCTAA-ACAAGCCA-AG-CAGACTT--T-ACTTTAC--GGTG-T--------TTAA-C-------G---GGTTTACATAGCC--G-GA--CA-AA-T--T-----------T--------T..C-----CTAGCC-......................... 8485
CPZ.US.85.CPZUS            --AGT-------ACA-----AT-AG--A-A-----A----------------A--A--T-----AA--------T--AC-AG-A-GCTC-T-CT--G----ATA--------A--------T--G--A---..C-A--G--A-GC-......................... 8816
CPZ.CM.x.SIVcpzDP943       C-A-T-------ACA-----AT-AG--A-A-----A--T-------------A--A-------AA----------T--G---A--ATAG------G-----TA--T-----A---C-C--T-----A---..--A--G--A-GC-......................... 8374
CPZ.CM.x.SIVcpzMB897       --A-C-------A-A--------TG-C--------A---------------A-----A-------AA-Y------A---C----A---T-G-C----GC---TTG--T--------------T-----A--T..--A-----AGC--......................... 8322
N.CM.02.DJO0131            ---G--------A-C--------T--CA-A-T---T--T-T------A-----A-----C---C--------T---C----A--ATAG-A--G-GA--ATTA-----------------C---------.C-A-----AGCAC......................... 8241
N.CM.04.04CM_1015_04       --AG--------A-C--------T---A-A-----T--T--------A-G---R----------C--------T---C----A--ATAG-A--G-GA--ATTA-----------------C---------.C-A-----A-CAC......................... 8213
N.CM.04.04CM_1131_03       --AG--------A-C--------TR--A-A-Y---T--T--------A-----A---------M-Y------T---C----A--ATAG-A--G-GA--ATTA--G-C-----------T---------.C-A-------CAC......................... 8278
N.CM.95.YBF30              -G-G--------A-C--------T---A-A-----T--T-T------A--A--A----------A--------T---C----A--ATAG-A--G-GA--ATTA-----------------C-----A---.C-A-----AGCAC......................... 8359
N.CM.97.YBF106             --AG--------A-C--------T---A-A-----T--T-T------A-----A----------T--------T---C----A--ATAG-A--G-GA--ATTA-----------------C---------.C-A-----AGCAC......................... 8317
O.BE.87.ANT70              C-----------ACA--------AG-CA-ACTT--AG-G--------CA-TT----T--CG-CA-A--C-C--GGA-----A--ATAG-A-C--GA-----TA----C--A--G-----T---------.--A-----A-G--......................... 8855
O.CM.91.MVP5180            --A---------ACA-A-C----TG--A-T-TT--AG-GT----T--CA-TT----T--CG-CA-C--CTT--GTC-----A--ATAG-ACA--GAT-C-TT-----C--A-------T-----A--T..GCA-----A--C-......................... 8881
O.CM.96.96CMABB637         C----A------ACA------G-AG--A---TT--A--------T-GCA-TT----T-----CA-C--CTT--G-A-----A--ATAG-A--G-GA---TTA-T--Y--AC-------T---------.--C---C-CTGC-......................... 8348
O.CM.98.98CMA104           C-----------ACA--TC-T--AG--A---TT--AG-G-------GCA--T----T--CGACA-A--CTT--GGA-----A-GATAG-A----GA--A-TTA----C--A-------T--------M..--A-----A-GC-......................... 8359
O.CM.99.99CMU4122          C-----------ACA-A-C-AT-RG--A--CTT--AG-------T--CA-TA-A--T--C--CA-AT-CTT--GGA-----A-GATAG-A----GA----TTA----C--A-------T---------.--A---CA-GC--......................... 8331
O.FR.92.VAU                C-A-C-------ACA-ATC-A--AG--A-TGTT--AG-G-----T---A-TT----T--C--CA-C--CCT--GTA-----A-CATAG-G----GA----TTA-------A-----G--T---------.C-T--GC-AC-C-......................... 8421
O.SN.99.SEMP1299           C-----------ACA---C-A--AG--A---TT--AG-G-----T--CA-TT----T-TC-CAA-A--TTT--GGA-----A-GATAG-A----GA----TTA----C--A--------T---------.--A-----A-G--......................... 8923
O.US.99.99USTWLA           --A-T-------ACA-----A--AG--A-TGTT--AG-G-----T-GCAGTT----T--C-ACA----CTC--GT--G---A-GATAG-A----GA----TTA----C--A--G-----T--------T..GCA-----A--C-......................... 8332
```

HIV-1/SIVcpz Complete Genomes

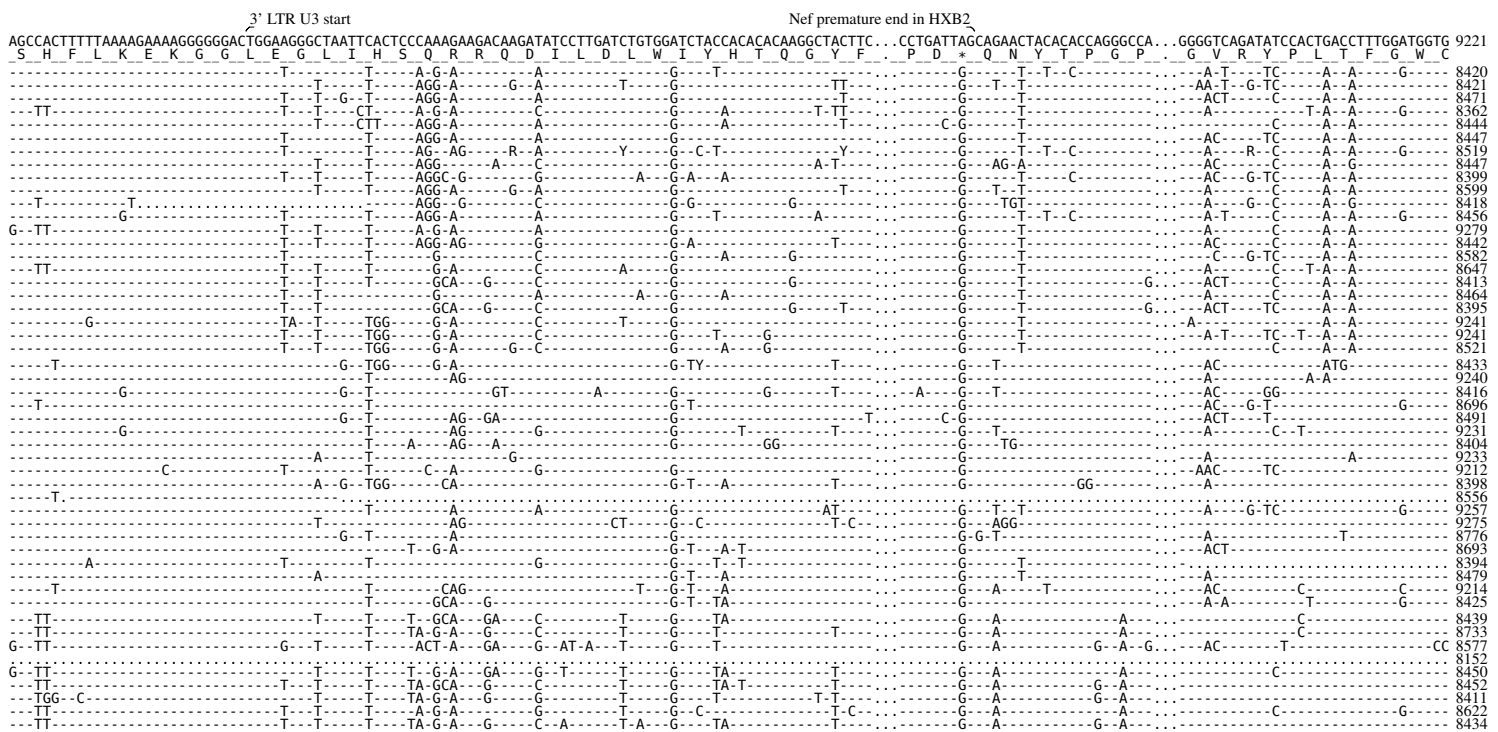
```
                       Env gp41, gp160 end  Nef start
B.FR.83.HXB2           ....TGCTATAAGATGGGTGGCAAGTGGTCAAAAGTAGTGTGATTGGATGGCCTACTGTAAGGGAAAGAATGAGACGA...........................................................GCTGAGCCAGCAGCAGATAGGGTGGGAGCAG 8893
Nef                                 _M__G__G__K__W__S__K__S__S__V__I__G__W__P__T__V__R__E__R__M__R__R__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__A__E__P__A__A__D__R__V__G__A__
Env                    _.__L__L__*_
A1.GE.99.99GEMZ011     ....---A----C-----G-----A-------------CA-AG-G---------CAG--T---------T-A------..........................................GCTC---CT---------CCAG-A--A---C--- 8092
A1.KE.00.KER2008       ....----G---C-------------------------CA-AG-G---------CAGA-T-----------A-----T............................................A--CCA---------CCAG-A--A------- 8099
A1.KE.00.KNH1144       ....----G---C-------------------------CA-AG-G---------GAGA-T-----------------T.........................................ACTC---CT--TC------TAG-A--A------- 8149
A1.KE.00.KSM4024       ....--------C-----GA----------------A--CACA--G-A-------GAG--C------C---------T-T...............................................-CTT---------CCAG-A--A------- 8040
A1.KE.00.MSA4069       ....--------C-----G-------------------CA-AG-G---------GAG--T-----G-----------T...........................................-T-CCT---------TCAG-A--A------- 8122
A1.KE.00.NKU3005       ....--------C------A------------------CA-AG-G---------GAGA-T---------C-A------............................................A--CCTG--------A--G-A--A--G---- 8125
A1.RU.00.RU00051       ....---A----C-----G-----A-----------C--CA-AG-G---------GAG--T-----------A------.........................................GCTC---CT---------AGAG-A--A---C--- 8191
A1.RW.93.93RW_024      ....--------C----------------------------...AAA--G--C------GAGA-T-----------------C..................................GCTCCT---CCTT--------ACAG-A--A------- 8122
A1.SE.95.SE8891        ....---A----C-----G-------------------CA-AG-G--G-------AG--T--------G---GC---............................................A--CCT---------A-AG-A--A------- 8077
A1.SE.95.UGSE8131      ....-A------C------AAT------------G--..............------GAG--T--A--------A----A-GCTCGA....................GCTCCAGCTCATACTCCA---CCTA--------ACAG-A--A------- 8271
A1.TZ.01.A173          ....--------T-----G-------------------CA-AG-G---------GAGA-T-----------A------............................................A-CCCT-------T-AGAG-A--A------- 8123
A1.UA.01.01UADN139     ....---A----C-----G-----A-------------CA-AG-G---------CAGA-T-----------A------GCT...........................CCTGCTCCAGCAGCAC---CT---------ACAG-A--A---C--- 8128
A1.UG.92.92UG037       ....-A------C------AA--------------T-CA-AG-G---------GAG--T-----------A----A-ACTCCTACA.............GCAGCAAGGGAAAGAACAAGACAA---CCTA--------A-AG-A--A------- 8957
A1.UG.99.99UGA07072    ....--------C-----G-------------------CA-AG-G---------GAGA-T-----G-----A------............................................A--CCAG--------A-AG-A--A------- 8120
A2.CD.97.97CDKTB48     ....-A------A-----G-----------------A-CCA-AG-G---------GAGA-T------------------ACTCCTCCA................GCAGCAGAAGGAGTAAGACCAA--CCT-----T-----AG-A--A------- 8254
A2.CY.94.94CY017_41    ....-A------A-----G-----------------G--CA-ACCA---------G--A-T-----G---------A--ACTCCTCCA................ACAGCACAAAGAACAGAAGCA-TGTCT---------CCAG-A--A------- 8316
A.SN.01.DDI579         ....--------C----------------C--------ACAAGAG-------A--AG--T-----T----------A-.................................ACTCCCC-AACA-----------AG-A--A------- 8091
A.SN.01.DDJ369         ....--------C----------------C-------...........-------G--GAAA-T---------------AA-.......................ACTCCTCCAGCAC-AACA---------A-AG-A--A------- 8142
A.SN.96.DDJ360         ....--------C----------------C--------CC-AGCG----------AG--T-----T--G-------A-.................................ACTCCTC-AACA---------A-AG-A--A------- 8073
A.CD.02.02CD_KTB035    ....--------A-----G-------------------AC-AGGG---------G----CT-----------CAG---............................................A--CC------------AG-A--A------- 8913
A.CD.97.97CD_KCC2      ....--------A-------------------------CC-AG-G---------G----T-----C-------------............................................A--CCTG----------AG-A--A-----T- 8913
A.CD.97.97CD_KTB13     ....--------A--------------------G--CC-AGGG------T--G--A-T-----------------...............................................A--CCT-----------AG-A--A------- 8190
B.AR.04.04AR151516     ....-----------------T-----------GTA--TCAGAGG----------G----------------ATT-A--............................................A----------------G----------- 8102
B.AU.87.MBC925         ....--------A--------------------------A-AT---------------------------------...............................................-----------------G----------- 8912
B.BO.99.BOL0122        ....-------------------------------AG...........------G--G-------------------AA-...........................................-----------------G----A------- 8088
B.BR.03.BREPM2012      ....---A----A------A--------------CA--T--G-GAG---------G-------A--C------------GCTAGGCCA................GCAGCAGAAAGAAGGAGACAA-----------------G----------- 8368
B.CA.97.CANB3FULL      ....--------------------T-------C-C...A---GG-A----G--------------------CT----..............................................-----A-----------G----------- 8163
B.CN.05.05CNHB_hp3     ....---------------A--------------G--T--A--GC-----------G---G-------------GA----...........................................-----------------G----------- 8909
B.CO.01.PCM001         ....-A------A---------------------C-A-AG-CAGC--AG---G--GA---------------GA----.............................................-----A-----------AG----------- 8076
B.GB.83.CAM1           ....--------A---------------------C-----C--GG------T--G------------------CA----.......................GCTGAGCCACGA-----------------GG----------- 8905
B.GB.86.GB8            ....---A-------------------T-------C-----A--T----C---T---GG--------C------CA--A-...........................................-----------------AG----------- 8884
B.GE.03.03GEMZ010      ....-----------------------------GC-GCT-AC-GG----------G-----------------GA--A-............................................---CT-----------AG----------- 8067
B.IT.05.SG1            ....--------A-----------------C--.......................................A--................................................................................. 8505
B.JP.05.DR6538         ....-----------------T--A-------------AGT-AGG-------------A-----------------G---...........................................A-----------------G----------- 8929
B.KR.05.05CSR3         ....-----------------------------------A........----------G----G-------------GA-A--........................................-----------------G------G---- 8947
B.NL.00.671_00T36      ....-A--------------A-------------CA--AG-G-GG----------G-G--G-------------CA----............................................-----------------CG--------C--- 8454
B.RU.04.04RU128005     ....--------------W-----A-----T---C-C...AARG---A------CG-----------------GA-AAG..............................................-T----------------G----------- 8362
B.TH.00.00TH_C3198     ....--G-------------------------------T--------G---------AG-GG-----------------A---.........................................-----------------G----------- 8094
B.UA.01.01UAKV167      ....-----------------------------------C-----AC-GG----------G-------A----------GA----........................................-----------------AG----------- 8157
B.US.04.ES10_53        ....-------------------------------......A--TC--AG---T---T-------------------A--GCT.................GAGCCAAGAATGAGACGA-------------...-CAG----------- 8883
B.US.99.PRB959_03      ....-------------------------------T----CA--GG-------A--G---------------------AA-.........................ACTGAGCCA--A-------------C-G----------- 8100
C.AR.01.ARG4006        ....---A----A-----GAA---------------......T--...---------G------------------------GCTCCTCCT.............GGTGTAGGAGACAGMAGGAGACAA-A--CT-----T-----GG-A--A---C--- 8108
C.BR.04.04BR013        ....---A----A-----GAA---A----------T----CCCAG-G------T--G--A----A--------A-------..........................................-----A-----------GG-A--A---C--- 8405
C.BW.00.00BW07621      ....---A----A-----G--------------------C---A-AG-----------G-------A--------A---A--.........................................A----T-----------GG-A--A------- 8249
C.CN.98.YNRL9840       ....---A----A-----G-------------------A---CAAAG---A-------G--A----A--G-------------.........................................A-----------------GG-A--A------- 8115
C.ET.02.02ET_288       ....--------T-----GA--GC-----------T----CCCA-GG------T----CA----A-------------A-...........................................A--CCA-----------GG-A--A------- 8119
C.GE.03.03GEMZ033      ....---A----A-----G-----A------------AC---A-AG-----------G-------A--G-----A--GA--..........................................A-------T---------AG-A--A------- 8121
C.IL.99.99ET7          ....---A----T-----G---------------TCC---C-AG-C--G------G-G-----A--C-----A--CA--............................................A----T------------AG-A--A------- 8080
C.IN.99.01IN565_10     ....--------A-----G---------------GC-C---.....---------G--A----A---------------.......................GCTGAGCCAGCA--A--A-----------A--A--A------- 8297
C.KE.00.KER2010        ....--------C-----A--------------------C--CA-AG-----------G--------------------------......................................---------------G--GG-A--A------- 8103
C.MM.99.mIDU101_3      ....---AG---A-----G---------------GC----CA-AG-----------G-CA----A--G------------...........................................A-----------------AG-A--A------- 8273
C.MW.93.93MW_965       ....--------------G--------------------C---A-A-----------GA------A--------A----A-...........................................A--AGT-----------GG-A--A------- 8089
C.SN.90.90SE_364       ....---A----A-----G-----------------T-C---A-AG-G--------AGAG-----A--------A---A--..........................................A--A-T----------ACGG-A--A------- 8038
C.SO.89.89SM_145       ....--------A-----G---------------G--C--AA-AG--------G---A-A----A--------A---A--...........................................A-----------------GG-A--A------- 8140
C.TZ.02.CO178          ....---A---GC-----G-----A---------C-----A-AG-----------GAAA----A--------A-----TACT.......................AGGCCAGCAGCAGAG-GA-T-------------GG-A--A------- 8116
C.UY.01.TRA3011        ....---A----A-----GAA---A----------T----CCAG-G---------GA-A----A--------A---A--............................................A----T-----------GG-A--A---C--- 8091
C.YE.02.02YE511        ....-A-A----A-----G---------------TCC---A-AG----G------GTG---------C-----A--GA--ACTAATCCAGCA.............GCAGCAGGAGTAGGAGCAGCGTCTCAA--CT----------CAG----A------- 8142
C.ZA.04.04ZASK164B1    ....--------A-----A--------------------C---A-AG-----------G--A--------------A---A--ACT....................AGGCCAGCAGCAGAG-GA-CAGG----------AG-A--A---A--- 8361
C.ZA.05.05ZAPSK240B1   ....--------A-----A--------------------C-AG......A--------GAAA----A--------------A-........................ACTAGGC-A--AG------------G-A--A------- 8293
C.ZM.02.02ZM108        ....--------A-----GAA-------------------------............------G-------A--------A---A--ACT...............AGAACTGGGGCAGCAA-----------------GG-A--A------- 8957
D.CD.83.ELI            ....-A--T---A-----------A---------------A-AG-G---------G--A-------------A---A--............................................A--A-T-------------G----A------- 8429
D.CM.01.01CM_0009BBY   ....--------A-----------A---------------A-AG-----------G--A---------A---A---A--.............................ACTAATC--ATT-------------G----A------- 8124
D.KE.01.NKU3006        ....----G---------------A---------------A-A-----------G--A-------------A---A--.............................................AT---T-----------AG------------- 8136
D.KR.04.04KBH8         ....--A-----C-----------A------------.........................------G--A-----------------A--..............................A-----------------G----A------- 8791
D.TD.99.MN011          ....-A------A-----------A---------------A-AG-----------G--A-------------A---A--............................................A----T------------GG----A------- 8083
D.TZ.01.A280           ....-A------A-----------A---------------A-AG-----------G--A-------------A---AA-............................................-----T-----------CG------------- 8081
D.UG.99.99UGK09259     ....--------------------A-----------C-AGA--GG-------T--G-------------------A--.............................................-----------------AG------------- 8089
D.YE.01.01YE386        ....-A------A-----------A---------------A-AG-----------G--A-------------A---A--............................................A----A-----------AG------------- 8049
D.YE.02.02YE516        ....--------A-----------A---------------A-AG--------------A---------A---A---AA-............................................A----T-------------G----A------- 8073
D.ZA.90.R1             ....--------A-----------A---------------A-AG-----------G--A-------------A---A--............................................A----T-------------G------------- 8247
F1.AR.02.ARE933        ....-C------A---------------------------A-AG-----------GAAA------------------G.............................................A-CCCT-----------AG------------- 8163
F1.BE.93.VI850         ....----G---A-----G--------------------C---A-AG-----------G-----G----------------A-........................................A-CCCTA----------AG------------- 8170
F1.BR.01.01BR125       ....--------A------AA----------------C-----A-AG-----------G--A-----------G--A----A-.......................................A-CCCTA----------AG------------- 8329
F1.ES.x.P1146          ....--------A--------------------GG---...C-AGC---G------G---------------A-----GACCCCTCCAACA..............GCAGAAAGGGTGGAGAGAAGACAGAACCCT------------AG------------- 8233
F1.FI.93.FIN9363       ....--A-----A---------------------------A-AG-----------G--A---------------------..........................................CCCC--CCAG-----------AG------------- 8192
F2.CM.02.02CM_0016BBY  ....--------A---------------------------A-AG------------AG-------------------A-............................................A-CCCTGT----------AG---------AAG- 8052
F2.CM.95.MP257         ....--------A---------------------------A-AG-----------G--A-------------A---.............................................................---------AG----A------- 8074
F2.CM.97.CM53657       ....--------A---------------------------AAAG-------------A-------------A----A-.............................................A-CCCTG--------CCAG------------- 8064
G.BE.96.DRCBL          ....--------A-----AAA-------------G-AA--AGC---------CGAG------------C-A----A-..............................................CACCCTG-----------AG-A--A------- 8853
G.CM.01.01CM_4049HAN   ....--------A-----A---------------------CA-AG-------------AGA---A---------A----A-..........................ACCCCTATAGCAGCA-AA--AGA---------CAG-A--A---.... 8114
G.CU.99.Cu74           ....---AG---A-----A-----A---------------A-AG-----------G-G--------------A--G-A-............................................A-CCCT----------CAG-A--A------- 8533
G.ES.99.X138           ....-A------C-----G---------------T--TCCA-AG----G------GAKA-------------A-A----............................................R-CCCT----------TAG-A--R------- 8349
G.GH.03.03GH175G       ....--------A--A--A----------------A--CA-A-AG-A--------AG--G-A----CAG--A----A-.............................................--CCCTG-----------AG-A--A------- 8960
G.KE.93.HH8793_12_1    ....--------A-----A-----A----------A.........---------G-G-------------A----A-...............................ACCCCAATAAGACAAA-CCCA-----------AG----A------- 8290
G.NG.01.01NGPL0669     ....--------A-----AA----A--T---------......T--...---------CG-G-------------A--G-A-ACC.....................CCTCCAACAGAAGAAAGA--A-GAG-----------AG-A--A------- 8119
G.PT.x.PT2695          ....--T-----C-----G---------------T--TCCA-AG----G------GAG--------------A-------...........................................--CCCTT---------TAG-A--A------- 8925
G.SE.93.SE6165         ....--------A-----A--------------------CA-AG----------CGAG--------------A---AAC............................................A-CCCTA-----------AG-A--A------- 8332
H.BE.93.VI991          ....GCTATA-CA-----A-----A---------G--T--A-A-G----------G----------------A--G-A-............................................A-----------------AG----A------- 8293
H.BE.93.VI997          ....GCTATA-CA-----A-----A--------------A-AG---G------G----------------A--G---..............................................---C--------------G----A------- 8208
H.CF.90.056            ....GCTATA-TA-----A-----A---------------GA--GG---G---T-----A-----------------G---..........................................-----A----T------AG----A------- 8213
J.CD.97.J_97DC_KTB147  .....................................ACAGC-------T--GAGA-----------------G---...............................GCTACTCCT-A--CT-----------AG-A--A------- 8055
J.SE.93.SE7887         ....--------A-----GAA---A----------------.........................CAG-----------------G---....................................GCTC-C-CT-------------G-A------------- 8191
K.CD.97.EQTB11C        ....-A------A---------------------------A-AG--------T-------------------------AA-..........................................A-ACCT-------------G------------- 8076
K.CM.96.MP535          ....-A------A---------------------------A-AG-----------G--A----------------------GCACGA..................CCAGCAGCAGACAGGGTGGGAA-AC-A------------CG------------- 8080
```

```
                                  Env gp41, gp160 end  Nef start
B.FR.83.HXB2                  ....TGCTATAAGATGGGTGGCAAGTGGTCAAAAAGTAGTGTGATTGGATGGCCTACTGTAAGGGAAAGAATGAGACGA.........................................................GCTGAGCCAGCAGCAGATAGGGTGGGAGCAG 8893
Nef                                        _M__G__G__K__W__S__K__S__S__V__I__G__W__P__T__V__R__E__R__M__R__R__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__A__E__P__A__A__D__R__V__G__A__
Env                           _.__L__L__*_
01_AE.CF.90.90CF11697         ....--------C----------------------A---AA-AG-G---------CAG--C-----------A--G---.........................................................A--CCTG----------AG-A--A------- 8866
01_AE.CN.05.FJ051             ....-TGCTAT-A-----G--------T------G----CAAAG-G------T--CAGA-C-----------A--G-A-.........................................................A--CCT------A----AG-A--A------- 8909
01_AE.CN.06.FJ054             ....----G---C-----AAA----------------.............-----CGG--C-----------A--G-A-.........................................................A--CC-A-----A----AG-A--A------- 8911
01_AE.HK.x.HK001              ....--G-----C-----A-----A--------------CA-AG-G---------CAG--C--A--------A--G-A-.........................................................A--CCC------A----AG-A--A------- 8275
01_AE.TH.01.01TH_R2184        ....--------C-----AA-------------------CA-AG-G---------CAG--C-------A---A-AG-A-.........................................................A--CCT------A----AG-A--A------- 8117
01_AE.TH.02.OUR769I           ....GCTATA--C-----A---G----------------CA-AG-G---------CAAA-------------A--G-A-.........................................................A--CCT-----------AG-A--A------- 8099
01_AE.TH.90.CM240             ....--------C-----AA-T-----------------CA-AG-G---------CAG--C-------A---A-AG-A-.........................................................A--CCT------A----AG-A--A------- 8467
01_AE.US.00.00US_MSC1164      ....--G-----C-----A------------------.............-----CAG--C-------A---A--G---.........................................................A--CCC-----------AG-A--A------- 8098
02_AG.CM.02.02CM_4082STN      ....--------C-----G-----------------C--CT-AG-G---------CAG--T-----------A----A-ACCCCA.................................GCAAGAGGGGAAAGAATAAGAC-AA-CC-----A-AG-A--A------- 8116
02_AG.EC.x.ECU41              ....--------C-----------------------C--CA-AG-G---------GAG--C-----------A------ACC.......................................CCTCCGGAAAGACAAA-CCCT-----------AG-A--A------- 7994
02_AG.GH.03.GHNJ196           ....--A-----C-----CA-----CTT--------CC-CA-AG-G------G--GG---T---------C-A------ACCCCTCCAACAGCAGAAAGAGTAAGACGACCCCCTCCAGCAGCAGAAGGGGCAGGAGCAA-ATCT-A--------TAG-A--A------- 9000
02_AG.NG.01.PL0710            ....--------C-----------------------C--CA-AG-G----------AGA-T-----T-----A----A-.............................................ACCCCCCCAGAAC-A--------G---A-AG-C--A------- 8074
02_AG.NG.x.IBNG               ....--------C-----------------------C--CA-AG-G----------AG--T-T-A------------A-.........................................................A-CCCTA--------ACAG-A--A------- 8420
02_AG.SE.94.SE7812            ....--------C-----------------------C--CA-AG-G---------CAGA-T-----C-----A----A-.........................................................A-CCCT---------AGAG-A--A------- 8283
02_AG.SN.98.MP1211            ....---A----C-----------A-----------C--CT-AG-G---------CAG--T-----------A---A-G................................................ACACAACCCA-CCCTT--------ATAG-A--A------- 8128
03_AB.RU.97.KAL153_2          ....---A----C-----G-----A--------------CA-AG-G---------CAG--T-----------A------......................................................GCTC---CT-----G---AGAG-A--A---C--- 8073
04_cpx.CY.94.CY032            ....----T---A-----A-----A--------------CA-AG-----------GAGA----------------G---GCTCGAGCTGAG....................CCAGAAAGAATGAGGCGAGCTCAA-----------------CAG-A--A------- 8302
05_DF.BE.x.VI1310             ....--------A------------------------------AG-----------G--A---------A---------C.........................................................A-CCCT-----------AG------------ 8299
06_cpx.AU.96.BFP90            ....--------A-----AAA---A---------.............-----T--CAG-----------G---------................................................ACCCCACCAA-A--AAG---------AG-A--A------- 8938
08_BC.CN.97.97CNGX_6F         ....---A----A-----G--------------------CA-AG-G---------G--A----A--------A------.........................................................A-----------------G------------ 8093
09_cpx.GH.96.96GH2911         ....--------C-------------------------RCA-AG-G---------GAGA-T-----------A---A--...................................................ACCTCCC-MTCC-------..--AG------------ 8068
10_CD.TZ.96.96TZ_BF061        ....-A------A-----------A---------------A-AG-----------CAG------------------A--.........................................................A--AGT-----------GG-A--A------- 8245
11_cpx.GR.x.GR17              ....--------C--------A--T--------------CA-AG-G---------GAGA-T---------T-A-----C...................................................ACTCCTC-AACAG----------AG-A--A------- 8172
12_BF.AR.99.ARMA159           ....--------A-----------------------C---A-AG-----------GAGA-----------T--------.........................................................--CCCT-----------AG------------ 8858
13_cpx.CM.96.1849             ....-A-A----C---------------------C----CACAG-G--------AG----------------A........................................................................------ACAG-A--A------- 8324
14_BG.DE.01.9196_01           ....-K------C-----G---------------M--TCCC-A-C---G------GAG--------------A------.........................................................A-CCCT----------TAG-A--A------- 8419
15_01B.TH.99.99TH_MU2079      ....--------C-----AA-----------------.............-----CAG--C-------A---A-AG-A-.........................................................A--CCT-----------GG-A--A------- 8274
16_A2D.KR.97.97KR004          ....-A------A-----G-----------------A--CT-ACCA---------G--A-T--------G---------ACTCCT..............................CCAGCAGCAGAACGAACCCCTC-A-CAG----------AG-A--A------- 8307
17_BF.PE.02.PE02_PCR0155      ....---A----A------A-------------G------A-AG-----------GAGA--------------------GCTGGA.................................CCAGGAACGAGACGAGCT-AGCCAG----------AG------------ 8089
18_cpx.CU.99.CU76             ....---A----T-----G-----A--------------CC-AG-A---------GAGA-T-----------A---A--ACTCAAGAACTA....................CCTAGAACTCCTCCAGCAGAGTCCCAA---GAGCG-.................... 8239
19_cpx.CU.99.CU7              ....--------A.----G-----A---------------A-A------------G-----------------CA-A--.........................................................------------------G----A------- 7955
20_BG.CU.99.Cu103             ....--------A-----GAA---A------------.............-----G-G------------C-A--G---.........................................................A-CCCT----------CAG----A------- 8336
21_A2D.KE.91.KNH1254          ....-A------A-----G-----------------A--CA-AG----------------T--A--------A--GAA-.........................................................A---T-------------G----A------- 8115
22_01A1.CM.01.01CM_0001BBY    ....--------C-----------------------C--CA-AG-G---------CAG--C-----------A--G-A-.........................................................A--CCTG-G--------AG-A--A------- 8081
23_BG.CU.03.CB118             ....--------A-----GAA---A------------............---G--G-G-----------------G---..........................................GCCCCTCCAGCAGRAAACCCT----------CAG----A------- 8345
24_BG.CU.03.CB378             ....--------A-----GA----A------------...--ACC-------T--GTGA----A-----------G---ACCCCTCCAACAAGAACA.........AGGGGAGGAGCCTCTCAAAATTTAGCTAGCCAT-AAC-TT---------TAG-A--A------- 8405
24_BG.CU.03.CB471             ....-------GA-----GAT---A--C------CC-.........--G------GAGA----A-----------G---..........................................ACCCCAGCTAGACAG-AACCC----------TAG-A--A------- 8366
25_cpx.CM.01.101BA            ....-------GA-----G--------------G-----CA-AG-----------GAGA--------------------......................................................ACCC--ACAG----------AG-A--A------- 8167
27_cpx.CD.97.97CDKTB49        ....--------C-----G-----------------A--CA-AG----G------G---------------------A-.........................................................A-CCCTGT----------G-A--A------- 8097
28_BF.BR.99.BREPM12609        ....------------------------------C--G--AC-GG-------GAGG-----------------------.............................................ACTAGGCCAGCA------------------G------------ 8306
29_BF.BR.02.BREPM119          ....---A----A----------------------A----A-A------------G----------------A------.........................................................A-CCCT-----------AG----A------- 8045
31_BC.BR.02.110PA             ....GCAATA--C-----GAA---A---------T-----CCA..G-TGG-ATGGC---CT-TAAG-GA--GA-TGA--......................................................CGA-----------------GG-A--A---C--- 8385
32_06A1.EE.01.EE0369          ....--------A-----A--------------------CA-AG-A--T------GAGA-----------T-A------.........................................................A-CCC------------AG-A--A------- 8519
33_01B.MY.05.05MYKL007_1      ....--------C-----G-----A--------------CA-AG-G----------AG--C-----------A--G-A-.........................................................A--CCT----------GAG-A--A------- 8284
34_01B.TH.99.OUR2478P         ....-AA-----C-----A--------------------CA-AG-G---------CAG--C--A--------A--G--G.........................................................A--CCT-----------AG-A--A------- 8059
35_AD.AF.05.05AF095           ....--------C--------------------------CAAAG-G---------GAG--T-------A--------A-.........................................................--CCCTG--------C-AG-A--A------- 8080
36_cpx.CM.00.00CMNYU830       ....--------C-----------------------C--CA-A--A---------CAGA-T-----C--------G-A-.........................................................A--CCT---------ATAG-A--A------- 8114
37_cpx.CM.00.00CMNYU926       ....---A----C-----------A--------------CC-AG-G---------CAG--G-----------A-A--A-.........................................................A-CCCT--------G--AG-A--A------- 8097
38_BF1.UY.05.UY05_4752        ....--------A------A--------------CC-...AGAG-G------T--GAGA--------------------............................................GCTGAGCCAGCAAA----------------AG------------ 8313
39_BF.BR.03.03BRRJ103         ....--------------------------------G-A-AAAG---A-------G------------A---AGA--A-AC..........................................CCCAATAGAACAAACCC-.....................----- 8368
40_BF.BR.04.04BRRJ115         ....------------------------------G--C-----CCC---------GAGA-----------------A--............................................GCTGAGCCAGCA----------------AGGG------------ 8474
42_BF.LU.03.luBF_05_03        ....--------------------R---------------T--G-----------G--------------------AA-.........................................................A----A--G-------T-G--------C--- 8355
43_02G.SA.03.J11223           ....--------A-----A-----A-----------G--CA--GG-------T-CGAGA----A--G-----A------ACACCC.................................CCTACAGCAGCAGAAGGA--A-GAA----------AG-A--A------- 8421
U.CD.83.83CD003               ....-A--G---A------AA---A---------CAA.............-----G--A-------G------------.............................................GCTAGGCCAGCA------------------G-A--A------- 8357
U.CD.90.90CD121E12            ....-A--G---A------AA---A---------CAA.............-----G--A------------------A-.............................................GCTAGGCCAGCA------------------G-A--A------- 8052
U.GR.99.GR303                 ....--------T----------------------------A-AG-G---------GAAA-T-----------A------ACT.....................................CCTCCAAAGGATAGACAAC--CTC----------AG-A--A------- 8204
U.NL.95.U_NL_95_H10986_D1     ....--------A----------------------------A-AG-----------G-------A-----G--A--G---ACT....................................CCAGCAGCGGAGGGAAGGCGA-CT------------G-A--A--G---- 8820
CPZ.CD.90.ANT                 ....AAA-TA--C------TCTGCA-----T----TC-AGTG-G-----GCAAGACAA-C--TTAGG-A---A...-AT............................................................--AA--AATC-----GATA-A--GC-CT 8347
CPZ.CM.01.SIVcpzCAM13         CT..----T---------A-CTC------T---------CC-AG-A---------GAAA-C--AA-------A-A-GAG.........................................................---CCAAT---------AG-A--A----AG- 8526
CPZ.CM.05.SIVcpzEK505         ....-TT-----A-----AAAAGTT--------G-----CC-TG-G--G-----ACAA--C-TA--T---------A--.........................................................CAGC-AGATC-------AG----A------- 8400
CPZ.CM.05.SIVcpzLB7           ....-AT-----A-----AA----A--------------CC--G-A---------GAA--T--AAG------T----A-.........................................................-AAACTGT---------AG-A--A---C--- 8425
CPZ.CM.05.SIVcpzMB66          ....-AT-----A-----A-----------------C---C--G-G--C-----AGAG-----A--C---C-T--GA-G.........................................................A-AC-AA-T--------AG-A--A---C--- 8538
CPZ.CM.05.SIVcpzMT145         ....----T---------A-C------------------CC-AG-A--------AGAAA----A--G-----A-A-A--.........................................................---CCAA-----------G-A--A--G-A-- 8426
CPZ.GA.88.GAB1                ....---AT---------AAC--A-----T----------C--G-A---------GAG--C--AAG------A--GGA-.........................................................---CCAA----------GG-A--A----A-- 8966
CPZ.TZ.00.TAN1                ....-AAAT---A-----AAA--TA-TT.....G-----ATG-CC---GGCC-GG-AA-CC-TC---GATC-T...-AT............................................................A-CA-CT--AGT--GCCT--A---CAG- 8579
CPZ.US.85.CPZUS               ....-AT----.------AA---------T--G------CA-AG-A--------AGAA--C--AA-C---T------A-...................................................ACTCAGA-AACAG----------AG-A--A---C-T- 8926
CPZ.CM.x.SIVcpzDP943          ....--------C-----CAA---A-----------C--CC--G-A--------AGAA-----A-CT-----A----A-.........................................................---C-A-----T---CCAG----A---T--- 8480
CPZ.CM.x.SIVcpzMB897          ....--------------AAA---A--------------CC--G-A---------G-A--------C---C-T---A-G.........................................................A-ACCT-----------AG-A--A--GC--- 8428
N.CM.02.DJO0131               ....-T------------AAAA-TT--------G--C--CA-AG-A--------AGAAA-C--A------------A--CAA.....................................CGACCTCACGAGCCAGCA-TA-----...----TAG-A--A------- 8365
N.CM.04.04CM_1015_04          ....-TT-----------GAAG-GT--------G--C-----AG-A--------AGAAA-C--A------------A--CAA.....................................AGACAAACTCAAGCAGCA--A-CA.....----TAG-A--A------- 8334
N.CM.04.04CM_1131_03          ....-TG-----------GAAA-TT-T------G-AC--A--AGCA------T-AGAAA-C--A------------A-GCAA.......................................ACTCAAGAGCCAGCTCAA-------------TAG-A--A------- 8402
N.CM.95.YBF30                 ....--A-----------AAAG-TT--------G--C--CC-AG-A--------AGAAA-C--A------------A--CAA........................................ACGCAAGAACCAGCA-TA-----...----TAG-A-CA------- 8480
N.CM.97.YBF106                ....-TT-----A-----AAAA-TT--------G--C--CC-AG-A--------AGAAA-C--A--------A---A--....................................................CAAACTC-A-----...----TAG-A--A------- 8426
O.BE.87.ANT70                 ....-AT-G---C-----AAA-GCA-T-AG----G---AAT-TGAG-------G-AG-A----A------------A--ACT.......................................AGAACTTTCCCTGAGT----A---TGC---CC-G-A--A---CAGA 8979
O.CM.91.MVP5180               ....-AG-G---C-----GAATGCA---AGC-----C-AAT-TGCA------T-AGAA-----A--T------------..........................................TCCTCCTCTGATCCTCAAC-A---TGT---CC-G-A--A-----T- 9002
O.CM.96.96CMABB637            ....-AT-G---C-----GAATGCA---AGT----AC-AAT-TGCA------T-AG-A-----A--T---------A--ACCAGA..............................ACCTCCAGAACCTCCCCTAATT----A---TGT---CC-G----A----A-A 8481
O.CM.98.98CMA104              ....-AT-G---C-----AAA-GTT-T-GGG-----C-AGT-TGCA------T-A--A-----A--G---------A--..........................................GCTCCTGMRCCTGAAC-A--A---TGT---CC-G-A--A---MA-C 8480
O.CM.99.99CMU4122             ....-AG-G---C-----AAA-GCA-T-AGR-----C--AT-TCCA------T-AG-CA----A--R---------AA-.............................................ACTTCCCCTGAGC----A---TGT---YC-G-A--A---C--- 8449
O.FR.92.VAU                   ....--T-----C-----AAATGTA-TAGG--------AATG-GCA------T-AGAA-----AAG-----------A-ACTTCC.................................CCTGACCCCCAGTCCTCT-ACACC-A-C-T---CC-G----AA------ 8551
O.SN.99.SEMP1299              ....-AT-----C-----AAA-GTA-T-GGT---GA--TAT-T-AG------T-AG-A--C--A------------G--.............................................ACTTCCCCTGATC----A---TGT---CC-G-A--A---CA-A 9041
O.US.99.99USTWLA              ....-AK-----C-----AAA--TA-TAA----G--C-AAT-TGAA--------AG-A-----AAG-----------A-.............................................ACTTCCCCTGATC--C-A---TGT---CC-G-A--A----AG- 8450
```

```
B.FR.83.HXB2           CATCTCGAGACCTGGAAAAACATGGAGCAATCACAAGTAGCAATACAGCAGCTACCAATGCTGCTTGTGCCTGGCTAGAAGCACAAGAGGAGGAGGAG......GTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTT 9057
Nef                    A__S__R__D__L__E__K__H__G__A__I__T__S__S__N__T__A__A__T__N__A__A__C__A__W__L__E__A__Q__E__E__E__E__.__.__V__G__F__P__V__T__P__Q__V__P__L__R__P__M__T__Y__K__A__A__V__D__L_
A1.GE.99.99GEMZ011     T-----A---TT----T--GT--------G-------C--T-------------T-------AC--C--------G--G--G--------A-----........--A--C----------AG--G-----------G-----------A--T---G----T-C-----C 8256
A1.KE.00.KER2008       T-----A---TT-A--T--------------T--------T---......-TA-ATC-CC--AG---CA-----G-G--------------A------........--A--C---------CGG--A--A-----AC---G-------------A-----C-G-----C 8257
A1.KE.00.KNH1144       T-----A---TT-A-CT--------------T-------------......-TA-ATC-CC--AG---C--------G-----GT-T-----T------........--A--C----------GG--A---------C-------------------G----A-------C 8307
A1.KE.00.KSM4024       T-----A---TT-A--T-----C------G-T-------------......-TG-ATC-CC--AG-----T------G-------------A-A-----........--A--C----------GG--A--A-----CC-------------------A------T-T-----C 8198
A1.KE.00.MSA4069       T-----A---TT-A--T---T--------G-T--------T---......-TA-A-C-CC--AG---C-T------G-----G-G-----A-------........--A--C-----G----GG--A--A-----C---G-----------------G----A-------C 8280
A1.KE.00.NKU3005       T-----A---TT-A---G-------------------C--T--......A-A-ATC-CC--AG---C-T------G-----G-----A--C--------........--A--C----------GG--A--A-------------------------G----C-G-----Y 8283
A1.RU.00.RU00051       T-----A---TT----T---Y------Y-G-------C-AT-------------T-----------C--------G--G--G--------A------........--A--C----------AG--G--A-------------------A--T---G----K-------C 8355
A1.RW.93.93RW_024      T-----A---TT-A-CT-G----------G-------C----------.....A-C-ATC-CC--AG---C-T------G-----G-----A--A------GAG...--A--C----------GG--A--A--G---C-------------------------G-----C 8283
A1.SE.95.SE8891        T-----A---TT-A--T------------G-------C----------.....ATA-ATC-CC--AG---C-------G-G-----G--------A-------........--A--C----------GG--A--------C-------------------G------G-----C 8235
A1.SE.95.UGSE8131      T-----A---TT-A------------------------C--T---......ATC-ATC-CC--AG---CA-------G----------CAC-A--A--TGAGGAG--A--C----------GG--A--A-------------------------G---C-G-----C 8435
A1.TZ.01.A173          T-----A---T--A--T-------------------AC-------......ATA-ATC-CC--AG-----T------G--------------T-----T..........----C----------GG--A--A------C-------------------------T-T-----C 8281
A1.UA.01.01UADN139     T-----A---TT----T--G---------G-------C--T-------------AT-------A--C--------G--G--G--------A-----........--A--C----------AG--G-----------G-----G-----A--T---G----T-C-----C 8292
A1.UG.92.92UG037       T-----A---TT-A--T------------G-------C----------......-TA-ATC-CC--AG---C-T------G-----G--G-----A--A---........--A--C----------G---A---------C---G---------------------T-C-----C 9115
A1.UG.99.99UGA07072    T-----A--GTT-A--TC--T----------------C----------......ATA-ATC-CC--AG---C--------G-----G--------C------........--A--C--------T-GG--A---A-------------------------G---T-T-----C 8278
A2.CD.97.97CDKTB48     TG----A---TT-A-CC-G----------G-------C--T-------------A---C---A--C--------G-----G-----A-------A........--A--C----------GG--A-------------GG-G-----C-------G-----------C 8418
A2.CY.94.94CY017_41    TG----A---TT-A-CT-CT---------G-------C--T-------------T---C---A--C-----G-G-----G-----A------AGTGAA...--A--C--C-------GG--A-------------G-------C-T-----G---GT-T-----C 8483
A.SN.01.DDI579         T-----A---TT-AAGT------------G-A-------CT---......CAAG-A-G---AG--CA-------G-----G--------A-----A.......--A--C----------G---A--A-----------G--C--------T--A-G---T-T-----C 8249
A.SN.01.DDJ369         T-----A---T-CA--T--G---A-----G---------GAA---......--A-A--G---AG--CA--A-----------AG--------A-----A.......--A--C---------CGC--G--A---------------C--------T--A-G---T-T-----C 8300
A.SN.96.DDJ360         T-----A---TT-A--T---T--------G----------CT---......CAAG-A-G---AG--CCAT------G-----G--T-----C-----A.......--A--C----------G---G--A-----------G--C-------TT--A------T-C-----C 8231
A.CD.02.02CD_KTB035    T---------T--A--T-G-------------T-----C--T-C-----T-----T--CT-A-A------T-----G--G--G--------A-G---........--A--C----------AG--G--A-G---------G------G-C--T------C-AT-T-----G 9077
A.CD.97.97CD_KCC2      TC----A---TT-A--T---T--------G-------C-AT-------------T--C---A--C--------G---------------A------........--A--C--C-------GG--A--------AAC---------------T-----------G--C--- 9077
A.CD.97.97CD_KTB13     T-----A---TT----T---T--------G-------C-AT-------------T--CT---AC--C------G-G--------------A-----GAA...--A--C----------GG--A---------------------------T--A-G---T-T--C--C 8357
B.AR.04.04AR151516     -------G---T-----G-------------------------C----------T------A--C------------------------G---GAG...--A-------------G---------G-----G-----------C-----------G--------- 8269
B.AU.87.MBC925         ------------------------A-------------------------A--A-------A-------------------------A--T-----.........---------------A------------------------------------------------- 9076
B.BO.99.BOL0122        -------------------------------------------------AA--T-----A-A-GTA---------------------------------.......----C----------G---------------------------------A------A-------C 8252
B.BR.03.BREPM2012      -----------T---C-G-G-G---------------------------------T-------A----------------------T--A--T-----A.......--A----------GG--------G--------G--------C-----------GT-T------ 8532
B.CA.97.CANB3FULL      T--------------------G------------------T--------------T-------A--------------------------T--A--A.......--A----C--------A------------------------------------------------- 8327
B.CN.05.05CNHB_hp3     T--------------G--G---------------------------......A-AG-A----A---C----TA-------------------------A.........----C----------G-----------------------------------G---GT-------C 9067
B.CO.01.PCM001         --------------A---------------------------C-----------T------A------W------------T-----T--A-----.......--A--C----------G---------C------------------CG----------A-----A-G 8240
B.GB.83.CAM1           T----A-----T--------------------------------------------T------A---------------------------------.......----C----------G------------G---------------------------T-----A-- 9069
B.GB.86.GB8            T------------------------------------------------------T------A---------------------------------.......----C----------G---------------------------------------G--------- 9048
B.GE.03.03GEMZ010      ---------------------T--------C-----------------------T------AG-------------------------A---GAG...-----C----------GG--------G--C----------------------G---CC-T------ 8234
B.IT.05.SG1            ........................................................................................................................................................---A-----T-- 8518
B.JP.05.DR6538         T--------------------T-------------------------G---AT------A---------------------------A--........-A------------G------------------G----------TT----A----T-T------ 9093
B.KR.05.05CSR3         ------------------G-T---------C-----------------------A----C---A--TG------------------------------A.....----------------G---------G--------------------------------G------ 9111
B.NL.00.671_00T36      T----------T-------------------------------......A--AT-------------T----------------------A.....----------------G-----------------------------------G----T-G------ 8612
B.RU.04.04RU128005     -T----G-------------------------------------C-------------A-----------C-------------------GAA...------------------G------A------------------------------------------ 8529
B.TH.00.00TH_C3198     T-----------------------------------------------------T---A--C--------T-------------------------GAG...------------------G----------------------------------------------A-- 8261
B.UA.01.01UAKV167      T----------T-----GG---------------------------......-T--AT-G---------A-------------------------A---.........----C--------CG------A--G-----------------------------CC-T------ 8315
B.US.04.ES10_53        ---------------G---------G-----------------------------A--T-------A----------G-------C-----A--A-G---GAG...--A-------------G-----------------------------------G------------- 9050
B.US.99.PRB959_03      T-------------------------------------A---......-T--AT-------------A---------C-------C-A----T---GAG...-------------T-G------A--------------------------A-G----T-------- 8261
C.AR.01.ARG4006        -T----A----T-A--T-G-T----G---T-G-----C--T--C---C-T--C-AT-------AC-----------G-----G--C-----------AGAG...--A--C----------G-----------------------------TT----G---AT-C-----C 8275
C.BR.04.04BR013        -----A-----T-A--T-G------G----CG-----C-----C---C----C-AT------A-----------G-------G----C--A--T.......--A--C----------G---------------------T-------TT--------AT-T-----C 8569
C.BW.00.00BW07621      ------A----T-A--T-G-----G----T-T-----C-----C--GC-CA-C--T--C------------------G-----G-----------T--A.......--A--C----------G------------------G-----------T----G---AT-------C 8413
C.CN.98.YNRL9840       -G----A----T-A--T---T--------C-T-----................................................................................................................................. 8152
C.ET.02.02ET_288       -G--C------T-A---C-------G---C-A----CC-----C---C---AG-AT-------------------GC----G----------A--AGAG...--A--C----------G-----------------------------T--------AT-C------ 8286
C.GE.03.03GEMZ033      -C---------T-A--T------------C-T----CC--T--C-----C-AG-A-------A-----------GA---G---------------GAA...--A--C----------G------------G--------T--------T----G---AT-T-----C 8288
C.IL.99.99ET7          -G----A----T-A--T--------G---C-T-----C-----C----TT--C-----AAA-A------------G-----G-----C---A--AGAG...--A--C----------G-----------------------------------G---A--------C 8247
C.IN.99.01IN565_10     -G----A----T-A--T-----..........------C-----C----ACA-CG-T--------------------C-AG--AG-----C---------AGAAGGA--A--C--------A-G---------G--------------------T----G---AT-T-----C 8458
C.KE.00.KER2010        ------A----T-A--T---T--------C-T--C--C-----C-----T-GC-A--------AC-----------GC----------------A--AGAA...--A-------------G---------G------------------CGT----G---AA-C-----C 8270
C.MM.99.mIDU101_3      -G----A----T-A--T---T--------C-T-----C-----C----ACA-CT-T-----A-AC-----T-----GAG------------A--A--AGAA...--A--C-----A---G---------G-----------------G-TT----G---AT-T-----C 8440
C.MW.93.93MW_965       -G----A----T-A--T--------G---C-T----C-----------TCAGC-A----------------------G-----G-----------T-GAGAC...--A--C--------T-G-----------------------------T--------AT-C-----C 8256
C.SN.90.90SE_364       -G----A----T-A--T---T----G---C-T-----C-----C-----C--C--T---AG---------CA---G-----G------C---A--AGAG...--A--C----------G-----------------------------T----G---AT-T-----C 8205
C.SO.89.89SM_145       -----AA----T-A--T---T--------C-T----AC-----C---C-CA-C-AT-------A------T----GC----G------------A---GAA...--A--C----------A---------G--------G-----------T--------AT-T-----C 8307
C.TZ.02.CO178          -G----A----T-A--T--------------------C--T--C-----GC--C--T--------A-A------------G------------------AGC---GAA...--A--C----------G---------G--G-----------------T--AT----A--------C 8283
C.UY.01.TRA3011        -G----A----T-A--T---T----G---C-G-----C-----C---C-T--C-AT-------A--------------------G---C------CA--AGAG...--A-------------G-----------------------------TT---G---AT-C-----C 8258
C.YE.02.02YE511        -G----A----T-A-CT--------G---C-T-----C-----C---C-CA-C-AT----A------------------G-----G-----C---C--AGAG...--A--C----------G-----------------------------------G---A--C-----C 8309
C.ZA.04.04ZASK164B1    ------A----T-A--C------------C-T----CC-----C-----CA-C-AT--------C-----------G-----G-----------T---GAT...--A--C----------G-----A----------------------------G---AT-T-----A 8528
C.ZA.05.05ZAPSK240B1   ----A-AG---T-A--T--G-----G---T-T----CC-----C-----CAGC-A--------A-C-----------GC--------------A--AGAC...--A--C----------G------------G-------------------T---G---AT-C-----C 8460
C.ZM.02.02ZM108        ----C-A----T-A--C------------T-T-----C-----C-----CA-C--T------A------------G-----G-------------A--.......--A--C--------T-GG-----------------------------T--------AT-T-----C 9121
D.CD.83.ELI            T------------------------G------------------------AG---T------AC------------------------A---AGC--CGAG...-----C----------G---C-----------------------------A-A----C-------C 8596
D.CM.01.01CM_0009BBY   T----AA---T--------------G-------------------------A---T-------C------------------------A----A----........----C----------G------A----------------------------G----T-------- 8288
D.KE.01.NKU3006        T------G----------G------G---G--------------------CAG----C---AC-----------------G--G-------C-G-GAA...----------------G------A-----A----------------------G----C-------G 8303
D.KR.04.04KBH8         -------G----------G------G--------------A--------------T-------A------------------------A--A------GAA...----------------GG--------G--AC--------------------------C-------- 8958
D.TD.99.MN011          T---------T-------G------G----------------------------A---T-------C------------------------A-GT------......----------------G-----------------------------T------------------- 8247
D.TZ.01.A280           T-----------------G------G----------------T--------CAC--T------AC-----------------G----A---CA--CGAA...----------------G--------CG--A-----------------T---------T-G---T-G 8248
D.UG.99.99UGK09259     T------G------------T----G----------------CA---------CAG--T---C--AC-----------------------A----A--GAA...----------------G-----------A---------------------G----A--------A 8256
D.YE.01.01YE386        -------G-----------------G--------------------------CAG--T--AG--AC-----------------------A-----TA-TGCA...--A--------------G-----------A--------------------A-G----T-------A 8216
D.YE.02.02YE516        T---------T--------------G----------------------TAATAA-C-TGCC......-----------------------A--C-----A.......----------------G------------------------------------------------- 8231
D.ZA.90.R1             T------------------------G--------------------C-----CAG---T-------AC------------------------A---AGT---GAA...------------T-A-----------------------------T---A-A------------- 8414
F1.AR.02.ARE933        TG----A----T-A--C-G--GG--------T-----C-------TAA-AA--AT---C---AC-TG--------G-----------A-----A.......--A--C----------G----------------------------TT----G----C-.-----C 8326
F1.BE.93.VI850         TG---------T-A--C-G--GG--------T-----C-------TAG-A---T---C---AC-TG--------G-------------------A.......--A--C----------G------------G-------------C---T-----------------C 8334
F1.BR.01.01BR125       TG----A----T-A----G-GA--G------T-----C-------TAG--A--AT---C---AC-TG--------G------------A-----A.......--A-------------G-----------------------------T----G---C--------C 8493
F1.ES.x.P1146          -----------T-A-CT-GTAGA--G-----T-----C-------G-G------T---C---A-GTG--------G-------------T-----T.......----C----------G------------G-----------------T--------C--------- 8397
F1.FI.93.FIN9363       TG----A----T-A---G--GG--G------T-----C-------T-G-----T---C---AC-TG--------G-------------C-----A.......--A--C----------G---C-------------------------T----G----T-T-----C 8356
F2.CM.02.02CM_0016BBY  TG----A----T-A--T--------G-----T-----C-------T-GG-----T---AA--ACATA--------G-----G-----A--T-----A.......--A-------------G-----------------------------T-----G------C-----C 8216
F2.CM.95.MP257         TG----A----T-A--T----G---G-----T----AC-------T-GG-----T-------AC-TG---------------G-----------A...........--A-------------G-----------------------------T---------CC-G-----C 8235
F2.CM.97.CM53657       TG----A----T-A--T--------G-----T-----C---C-G--TAGG-A---T---AG--AC-TG--A-----------G---------T------........--A----C--------G-----------------------------T----G----T-C-----C 8228
G.BE.96.DRCBL          T-----A---TT-A--T-GG---------C----------A-----------GC-----C--A---------------G---------------CTCAGAG...--A--C----------G--A------------------T--------T-----T--GT-T-----C 9020
G.CM.01.01CM_4049HAN   .....................................................-----C-AT---C---A-A------------G----------------TTCA...GAG--A--C----------G--A---------C-------------T-----T---T-T-----C 8234
G.CU.99.Cu74           T-----A---TT-A-CT-GG-----------T-----C--T-----------C-AT--CC---A-----------GC-------------A--TTCAGAA...--A--C----------G--A----------G-----------C--T-----T---T-T-----C 8700
G.ES.99.X138           T-----A---TT-A-CT-GG-----------T-----C--------------C-AT---C---A-----------G-----------A--A--TTCAGAG...--A--C----------G--A--------C-R---------------T---GT---T-T-----C 8516
G.GH.03.03GH175G       ----C-A---TT-A--T--G--------GC------C-------------A-AT---C---AC-----T--------------------A--TTCAGAG...--A--C----------A--A--A--------G----T--------T-----T---T-T-----C 9127
G.KE.93.HH8793_12_1    T-----A---TT-A-CT-GG-----------------C-------------A-C-AT---C-C-A--------------G--G-----A-----TTCAGAG...--A--C--------T-G--A--A-----G-G----------C--T-----T---T-T-----C 8457
G.NG.01.01NGPL0669     ------A---TT-A-CT-GG-----------T-----C--------------C--T---C---A----------G----------C------CTCAGAG...--A--C--------T-G--A------------G--------------T--A-G---T-T-----C 8286
G.PT.x.PT2695          T-----A---TT-A-CT-GG-----------T-----C-------------C-AT---C---A-----------G--------A-T--A--TTCAGAG...--A--C----------G--A---------C------------------T---AGT---T-G-----C 9092
G.SE.93.SE6165         T-----A---TT-A--T-GG-----------------C-------------C-AT---C---A-----------G-------------------CTCAGAA...--A--C----------G--A--A--------G--------------TT---GT---T-T-----C 8499
H.BE.93.VI991          TG----A---TT-A--T-G--G-------G------TC-AT--C-T----T---AT-------AC-CC--------G-------------A--A---GAA...--A--C----------GG--G--------------------C--T----G----T-T-----C 8460
H.BE.93.VI997          T----------T---T--G--GA--G---G------TC-AT----------------T---C---A-GT--------G-----------------C---GAA...--A--C----------GG--C----------------------C--T---------C-G-----C 8375
H.CF.90.056            TG--------TT---T--G--GC--G---G------T--AT--------T----T---CG--A-GCC--------G---------------C-G---GAA...--A--C----------GG----------------------C--T----G----T-T-----C 8380
J.CD.97.J_97DC_KTB147  -G----A---TT-A--T-GG----G------------C----------T--AA--A---AG---C------------G--G-----ACA--A------GAG...--A--C----------G-----A----A---------------------G---CT-T-----C 8222
J.SE.93.SE7887         TG----A---TT---CT--G----G------------C-----------------A--------AC-----------G--------ACA--------------.........--A--A----------A------A----A--------------A-------G------T------- 8355
K.CD.97.EQTB11C        T-----A---------T--------G---G-------C-------T---TTC-AT---C---A------------G--------G--A--C-----C.......----------------G-----------A-----------------TT----G----T-T----G 8240
K.CM.96.MP535          T-----A--------CT-G------G---G-------C-------TT--CAC-AT---C---A------------G--------G-----A-----A.......--A-------------G-----------A-----------------T---------T-T-----C 8244
```

```
B.FR.83.HXB2               CATCTCGAGACCTGGAAAAACATGGAGCAATCACAAGTAGCAATACAGCAGCTACCAATGCTGCTTGTGCCTGGCTAGAAGCACAAGAGGAGGAGGAG......GTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTT 9057
Nef                        A__S__R__D__L__E__K__H__G__A__I__T__S__S__N__T__A__A__T__N__A__A__C__A__W__L__E__A__Q__E__E__E__E__.__.__V__G__F__P__V__T__P__Q__V__P__L__R__P__M__T__Y__K__A__A__V__D__L_
01_AE.CF.90.90CF11697      ------A---T--A--T-------------A--------T---......ATG-AT-------A-----------G----G-------A-----......--A--C---------GG--A----G---C--------------T----G----T-T------ 9024
01_AE.CN.05.FJ051          T-----A---T-CA--T------------G-A--------T---......ATG-AT-----A-A-A---T-------AG--------------A--AGAAGAG--A--C---------GG--G---------C--------------TT----G----T-T------ 9073
01_AE.CN.06.FJ054          T-----A---T--A--T------------G-A--------T---......ATG-GT-------A-----T------G--------------A-----......--A--C---------GG--A---------C-----------G---TT----G----T-T------ 9069
01_AE.HK.x.HK001           T-----A---T--A--T------------G-A------C-AT---......ATG-AT----A--A-A---T------GAG------------A--T---GAAGGG--A--C---------GG--A---------C----------------T----AC---T-T------ 8439
01_AE.TH.01.01TH_R2184     T-----A---T--A--T------------G-A--------T---......ATGGAT-------A-----T------GAG------------A-----......--A--C---------G---G---------C----------------T----G----T-T-----C 8275
01_AE.TH.02.OUR769I        T-----A---T--A--T------------G-A---GT---T---......ATG-AT---A---A-A---T------GAG-----------C--A---......--A--C---------GG--G--------------------------T----G----T-T------ 8257
01_AE.TH.90.CM240          T-----A---T--A--T--------------A--------T---......ATAGAT-------A-----T------GAG-----------C-----......--A--C---------TG--G---------C----------------T----G----T-T------ 8625
01_AE.US.00.00US_MSC1164   T-----A---T--A--T------------G-A--------T---......ATG-AT-------A-A----------G-G------------A-C---GAGGGG--A--C---------GG--G---------C----------------T----G----T-T------ 8262
02_AG.CM.02.02CM_4082STN   T-----A---TT-A-GT-G----------G-------C--A----------C--T-------A-AT---------G-------------A-----......-----C---------GG--G-----------G-------------T----G------C-----C 8280
02_AG.EC.x.ECU41           ------A---TT-A--T-----------C--------C--T--------T---T-------A-----A-----G-----------A--C----A.......-A--A---------GG--G-----------G--G-----------T--------A--C-----C 8158
02_AG.GH.03.GHNJ196        ------A---TT-A-CG-G------------------C--T------T--AG---T-------A------------G-------------A----AGAGGAG--A--C---------GG--A-----------G-------------T-----------C-----C 9170
02_AG.NG.01.PL0710         ------A---TT-A--T-G----------C------CC--T--C---T--CAA-AT---C---A------------G-----------A--A-GA-G-GAG...--A--C---------G---A-----G---C-G-------------T----GG---C-C-----C 8241
02_AG.NG.x.IBNG            ------A---TT-A--T-G------------------C--T---------CAA--T--CC---A------------G-------------T---A-T.......-A--C---------GG--A-----G-----G-------------T----G------C-----C 8584
02_AG.SE.94.SE7812         ------A---TT-A--T-G------------------C--T--------------T--------------------G-------------A-----......--A--C---------GG--G---------C-G-------------T-----------C-----C 8447
02_AG.SN.98.MP1211         T-----A---TT-A--T-G------------------C--A---------CA---T----A-A-----------G-------------A-----......--A--C---------G---A-----------G-------------T--------G--------C 8292
03_AB.RU.97.KAL153_2       T-----A---TT----T--GT--------G-------C--T-------------AT-------A---C--------G--G--G--------A-----......--A--C---------GG--G-----------G-----G-----A--T----G----T-C-----C 8237
04_cpx.CY.94.CY032         TG----A----T----C--------G----------T--A--------------T---C---ACAAAA-------G-----------A-----A--AGAG...--A------------GG--A-------------G---------C-TT--A-G----T-------C 8469
05_DF.BE.x.VI1310          TG-T--A----T-A--C--T-GG--G---G-T----TC--A-------T---C--T---C---A-ATG------G----------------------GAA...--A--C----------A-----------------------------T----------------C 8466
06_cpx.AU.96.BFP90         TG----A---TT----T--G-----G----------C-----------A---A-------C--------G-G--------ACA----A---.......--A--C----------G------------G-G-------------T--A--T---T-T-----C 9102
08_BC.CN.97.97CNGX_6F      T-----------------------------------------C-------A---T-------A---------------A-------------------.......----------G--------------------------TT----G---AT-C-----C 8257
09_cpx.GH.96.96GH2911      -G----A---TT-A-GT-GR-----G---C-T-----C-----C--T--CT----T---CA--AC-----------G--G-----G--A--T-----C.......-A--G--C--------G---C--------------------------T----G------C-----C 8232
10_CD.TZ.96.96TZ_BF061     -G----A----T-A-CT---TT---G--CC-G-----C-----C---C-G-GC-AT-------A-----------G-----G-----A-C---A--AGAG...--A--C----------G---------T--C---------------TT----G---AT-C-----C 8412
11_cpx.GR.x.GR17           -----AA---TT-A----G------------------C--T------AGCCAG-AT-----------C-----------------------T--A-G-......--A--C----------G---G--------------------------T----G----T-T------ 8336
12_BF.AR.99.ARMA159        TG----A----T-A-----C-GG--G-----T-----C--------TAG---C-AT---C----C-TG--------G-----G--------A-----A.......-A--C---------GG-------------------------------TT----G----T-C------ 9022
13_cpx.CM.96.1849          T-----A---TT-A--T-GG--------CG-------C--T---......CTAG-A---AAG-A--AC-T--------------------A-G---A.......-----C----------A-----------A-----------------T--A-G----T-T------ 8482
14_BG.DE.01.9196_01        T-----A----T-A-CT--G-----------T-----C---------R---SC-MT---C---M-----------G---------ACA--A--CTCAGAG...--A--C----------G---A---------C-G-------------T---RGT---T-T-----C 8586
15_01B.TH.99.99TH_MU2079   T---G-A---T--A--C--C----------G-A--------T---......ATG-AT-------A-A---T------GAG---G-G-----C-----......--A--C---------GG--A---------C----------------TT----A----T-T------ 8432
16_A2D.KR.97.97KR004       T-----A---TT-A-CT-C--G---G---G------T--AT--------------------A-A----------G-----G--------A------GAA...--A--C--C-------GG--A----------------------C--------G---GT-T-----C 8474
17_BF.PE.02.PE02_PCR0155   TG----A----T-A-G-----GG--G-----T-----C--------TAGC----AT---C---AC-TG--------G-------------C--A--A.......-A--C----------G------------------G---------------G----A--------- 8253
18_cpx.CU.99.CU76          ..................--G-----GT--GGA----AC----G-----------T---C--A--CAC-G-C-A--G----AGG-G-----A-----......--A--C--C-------GG--C---------C-------T-----C--T----G----T-T--C--C 8386
19_cpx.CU.99.CU7           T-----------------------G--------------------CAG------T-------A-----------------------A----CA--CGAA...--A-------------GR-------------------------------------G----T-------C 8122
20_BG.CU.99.Cu103          T-----A---TT-A-CT--G-----------T-----C--------------C-A----C---A-----------G-----------A--A--CTC-GAG...--A--C----------G---A---------C-G-----------C-TT-----T---T-T-----C 8503
21_A2D.KE.91.KNH1254       -----A----------C---G---G-------T------------------AT--CC-A-A---C--------G-----G--G------A-T---GAA...--A--C--C-------GG--A-----G-----------------C-T-----G----T-T-----C 8282
22_01A1.CM.01.01CM_0001BBY T-----A---T--A--T------------G-A--------T----TG......-ATC----AG-----T------G-------------CA-T--A.......-A--C---------GG--G---------C----------------T----G----T-T------ 8239
23_BG.CU.03.CB118          T-----A---TT-A-CT-GG-----------T-----C--------------C--T-------A-C--------G-G-----------A--A--TTCAGAG...--A--C----------G---A--------C--G-----------C--T-----T---T-T-----C 8512
24_BG.CU.03.CB378          T-----A---TT-A-CT-GG-----------T-----C----------T---C-AT---C---A-----------G-----------A--A--TTCAGAG...-----C----------G---A---------C-G-----------C-TT-----T---T-T-----C 8572
24_BG.CU.03.CB471          T-----A---TT-A-CT-GG-----------T-----C--------------C-AT---C---A-----------G-----------A--A--TTCAGAG...-----C----------G---A-----------G-----------C--T---AT---T-T-----C 8533
25_cpx.CM.01.101BA         T-----A---TT-A-CT-G------------------C-AT---------AG---T---T---------------G-----------A--ATCA---GAA...--A--C----------G---A-----------G-------------T----GT---T-T-----C 8334
27_cpx.CD.97.97CDKTB49     T-----A---TT-A-CT-GG---------G-------C--A---------A--AT---C---A-----------G---------C-------AAGTGAG...-----C----------G---C-------------------------T--A---------C-----C 8264
28_BF.BR.99.BREPM12609     T----A-----T------G-----------T------------------CA--AT-------AC---------------------A--------A.......-----C--C-------G----------------------------------A-G----T-------C 8470
29_BF.BR.02.BREPM119       T---------------G--------------C--------------T---A-------A-----------------------------------......-----C---------G----------------------------C------T----GT--------- 8209
31_BC.BR.02.110PA          ------A----T-A--T-G-T----G---C-G-----C-----C---C-T--C--T-------------------G------..CA-CAG-AG-A--ATG...AG-A-G--C-NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAG.----C-----C 8549
32_06A1.EE.01.EE0369       TG----A---TT-A--C--G-----------------C---C-----C--A---A-------------------G-------------T-----......--A--C---------G---------G----------------C--T--A-GT---T-T-----C 8683
33_01B.MY.05.05MYKL007_1   T-----A---T--A--T------------G-A--------T---......ATG-AT-------A--C--T------GAG------------A---A-CGAAGAG--A--C---------GG--A---------C----------------TT--A------T-T------ 8448
34_01B.TH.99.OUR2478P      T-----A---T--A--T------------G-A--------T---......ATG-AT-------A-----T------GAG------------AA-T--AGAGGAG--A--C---------GG--G---------C----------------TT---------T-T------ 8223
35_AD.AF.05.05AF095        T-----A---T--A-------------------G--C--T--C......ATG-ATC--C--AGC--C-T------G-----G-----A--A-----......--A--C---------GG--A--A---T-AC-G--G--G-------------G---CT-C-----C 8238
36_cpx.CM.00.00CMNYU830    T---------TT-A--T-G------------------C---------A--AG---T-------A--A--------G-------------A-G---......--A--C----------G---G---------C-------------C--T--A-GG---T-T------ 8278
37_cpx.CM.00.00CMNYU926    T-----A---TT-A-CT--G-------------------A------T--A-A--T---C---A-----------G---------G------CCCAGAT...--A--C----------G---------------C----------------T---------T-T-----C 8264
38_BF1.UY.05.UY05_4752     TG---------T-A--G-G--GG--------T-----C--------TAAT----AT---C---ACGTG--------G-------------A--A--A.......-A--C----------G---C-----------------------------G----T-------- 8477
39_BF.BR.03.03BRRJ103      TG----A----T-A----G--GG--G-----T-----C--------TAG-----AT---C--AA--TG--------G----------------T--TGAA...--A--C----------G------------G-------------T----G----T-------C 8535
40_BF.BR.04.04BRRJ115      -----------T----G---T---------------------------C--T---C---A----------------------------T----A.......-A--C----------A-----------------------------T----G-----------C 8638
42_BF.LU.03.luBF_05_03     -----------T----------------C------C-------------A---AG---------GT---------G-----------A---------......-----G--Y-------G-----------------------------T----G---AT-------- 8519
43_02G.SA.03.J11223        ------A---TT-A-CC-G------------------C--T------C--CAA----------A-----------G-------------T-----......--A--C----------G---A-G-----------------------T-----T--CT-T------ 8585
U.CD.83.83CD003            -----------T-A--G---T----G-----------C-------------AA--T----AG--C------CA---------GCGC-------C-G-GAA...--A--C----------G-----------------------------C--T---------T-T-----C 8524
U.CD.90.90CD121E12         T----------T----T-G------G-----------C-----------------T---AG---C------CA-----------G--A--A--A-GAGAT...--A--C----------A-----------------------------C--T--A--------T-----C 8219
U.GR.99.GR303              T-----A----T-A--T-G-----------G-------------------AGG----AT------AA-CTC--------G-----G-------------------......-A--C----------G---G---T-----C----------------T----A---AT-------C 8368
U.NL.95.U_NL_95_H10986_D1  TG-------------T-GG-----G-----------C------------------G-----A-AC-------------------ACA-----A--A.......-A--G----------G-------------A-----------------T---------T-T-----A 8984
CPZ.CD.90.ANT              GTGGAAAC--AT---CC-GTAGA--G---C-A--CTCCTCT-C--T--GTA-AGAA--A-A--TAATCA---ATTC-----AT--TACA-----A-G-......ACA--C-----T--G-G------A--C--CA-G-----C-----AG-A--ATT---------C--G 8511
CPZ.CM.01.SIVcpzCAM13      TT---AAG--T--A------T-------T-----T-----AC-----C--AG---T---CAAA---TA--A---T----T-A-ATGA-AC-TC-T--AGAAGAA-----C--------T-G---A-----G--A-------------------A-G---CT-T--C--- 8696
CPZ.CM.05.SIVcpzEK505      TT----A---T--A-CT--T-GA--G--T--T--C-TA--A-----TAAG-AG-AT---CAA--CCTA--T----------AG---A-A-----A-CA.......-A--C--------ATGT-----------A---------------A------A-A----T-T------ 8564
CPZ.CM.05.SIVcpzLB7        -C--A-A------A-CT-GG-----G-----T--------T-----CC--CAA--T----AGA-CAT-------T----G-AG---A-------A--A......A-AT-G--------A-A-------------A-----G-----------TT--A-AG---T-T------ 8589
CPZ.CM.05.SIVcpzMB66       T---A--G-----A-C--G------------A--C-----A--C---T-CCAG--A----AGA-CCT-------T------A-GT-C-AA-T--A--A.......-A--C--------A-GG-----A-----C-------------C-TT----G----T-T-----C 8702
CPZ.CM.05.SIVcpzMT145      T---A-A----A-A-C--GC-G-------G------T--A-----CC--CAA-------AGA---TA--T----------A-ATGC----T--A--A.......-----A----------A----A--A--G--C--------T-----C-TT----GT--CT-T-----C 8590
CPZ.GA.88.GAB1             TT--GAAG-----A----G---C------T------T-----G--C--CC---AG--T---CAAA--CTA--T---------G-A-ATG--CA-T--A--A.......-A--A--C------A-G---A------T--AACG--------------T--A-------T-T------ 9130
CPZ.TZ.00.TAN1             -C--A-A------CC-G--TA-A----GTC-----T-C--A--CCCT--GTA-CT-AGCA-A--TG-TA-AA-ACTCT-C--AC--TACT--A--A--A.......-A------------G-----AGCA-----CA-G-----C-----AG-G---CT---AA-------G 8743
CPZ.US.85.CPZUS            TC----A----T-A-C-G-------G--T-----T-C---A------C--CAG-A---CCAAA--CTG--A---T----T-AGATGACTA-TC-T--AAGTGAA--A--C--C--T--TCGC-----A-----A-----G--C------------ACA---AT-C---T-A 9096
CPZ.CM.x.SIVcpzDP943       TC--A-----TT-A--GGC------------A--C-C---A-----CC--CAA-A---CAAA--CTA--A-----------A-ATGC---GAC-T--AGAGGAG--A--C--------T-G---A--A-----CC-T-----T-----A-TT--A-G----T-T---T-A 8650
CPZ.CM.x.SIVcpzMB897       TC--G--------A--G-G------G--T--A--------G-----TC--CAG--A----AGA--CT-------T----G-AG---AG---C--A--A.......-A--A--------A-G---C--A--G---------G--C-----C--T--A--G--CT-T------ 8592
N.CM.02.DJO0131            -T----A---T--A-CC--T-GA--G--CC-----C-CA-------TAG-A-C-AT---C-AA--GTA--T---G-------------A--A--A-G-GAA...--A--C--------ACGC-----------A--G--G--------C--T--A--G-----T------ 8532
N.CM.04.04CM_1015_04       -T----A------A-CT--T-GA--G--C-----C-CA-------TAG--A--AT---AAA--GTA--T---G-------------A--A--A---GAA...--A--C--------ACGC-----A-----A--G--G--------C-TT--A-------T-T------ 8501
N.CM.04.04CM_1131_03       -T----A------A-CT--T-GA--G--C-----C-CA-------TAG--A--AT---AAA--GTAA-T--------------G--A-----A---GAA...--A--C--------ACGC-----A-----AA---G---------C--T--A-------T-T------ 8569
N.CM.95.YBF30              -T----A---T--A-CT--T-GA--G--C-----C-TA--A-----TAG--AC-AT---AAAG-ATA--T------------------A--A--A---GAA...--A--C--------ACGC-----------A-----G-----A--C--T--ACAG---T-T------ 8647
N.CM.97.YBF106             TT----A------A-CT--T-GA--G--C-----C-CA-------TAA--A--AT---CAAA--GTA--T------------------A--AS-----......--A--C--------ACGC-----A-----GC-G--G---------C--T--A--G---T-T------ 8590
O.BE.87.ANT70              TC--CA-G--AT-A-C-GCTAGA----GG--AC-----TC-C----TC-TCAA-A------A--CCT---A-TC------AGT--CC-A-----A--A.......-A-----------AG-------A--G---C----G---------C--T--A-G---AT-T--C--C 9143
O.CM.91.MVP5180            TC--CA-G--GT-A-C--CTAGA--G-G---AT-----TC-C-C--TC-TCAA-A------A--CCT---A-TC-----CAGC--CA-A--T-----T.......-A--C--C-----A-G-----A--G---C----G---------C-TT--A-----CT-T--C--C 9166
O.CM.96.96CMABB637         TC--CA--C-GT-A-C-G-TAGA--G-G---AC---T-TC-T----TC-TCAA-A------A--CCT---ACTC--C---AGC--CC-A---AGT--AGAA...--A--G--C-----A-G-----A--G---C----G---------C--T--A-G---AT-T--C--C 8648
O.CM.98.98CMA104           TC--CA-G--AT-A-C-GCTAGA----GG--AC-----TC-T-C--TC-TCAA-A------A--CCT---A-TTT-----AGT--CC-A--T--A--TGAA...--A-----C-----A-GG-----A------G----G---------A-TT--A-G---AT-C--C--C 8647
O.CM.99.99CMU4122          TC--CA-G--GT-A-C-GCTAGA----GG--AC-----TC-T----TC-TCAA-A------A--CCT---A-TC------AGT--TTCA--T--A--A.......-A-----C-----A-G------A--T---G--------------C-TT--A-----AT-T--C--C 8613
O.FR.92.VAU                T---CA-G-CAT-A-C-G-CAGAA---GG--AC-----TCTT-C--TC-TCAA-A----CAA--CCT---A-TC----G-AGC--CC---GTA-A-GA.....A-A--G--C-----A-G---C--A--G---C----G--G-----C--T--A-G---AT-T--C--C 8715
O.SN.99.SEMP1299           TC--CA-G--AT-A-C-GCTAGA----GG--AC-----TC-T----TC-TCAA-A------A--CCT---A-TC------AGT--CC-A--T--A--A.......-A--C--C-----A-G------A--G---C----G---------C--T--A-G---AT-T--C--C 9205
O.US.99.99USTWLA           T---CA-G-CAT---C-GGTAGA----GG--AC---A-TC-T----TC-CCAG-A------A--CCT---A-TC--------CT-CC----T-----C.......-A-----C-----A-G------A--G---C--CAG---------C-TT--A-----ACAT-----C 8614
```


HIV-1/SIVcpz
Complete Genomes


3' LTR U3 start
Nef premature end in HXB2
B.FR.83.HXB2 AGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTC...CCTGATTAGCAGAACTACACACCAGGGCCA...GGGGTCAGATATCCACTGACCTTTGGATGGTG 9221
Nef _S_H_F_L_K_E_K_G_G_L_E_G_L_I_H_S_Q_R_R_Q_D_I_L_D_L_W_I_Y_H_T_Q_G_Y_F_._P_D_*_Q_N_Y_T_P_G_P_._G_V_R_Y_P_L_T_F_G_W_C
A1.GE.99.99GEMZ011 ---------- 8420
A1.KE.00.KER2008 ---------- 8421
A1.KE.00.KNH1144 ---------- 8471
A1.KE.00.KSM4024 ---------- 8362
A1.KE.00.MSA4069 ---------- 8444
A1.KE.00.NKU3005 ---------- 8447
A1.RU.00.RU00051 ---------- 8519
A1.RW.93.93RW_024 ---------- 8447
A1.SE.95.SE889I ---------- 8399
A1.SE.95.UGSE8131 ---------- 8599
A1.TZ.01.A173 ---------- 8418
A1.UA.01.01UADN139 ---------- 8456
A1.UG.92.92UG037 ---------- 9279
A1.UG.99.99UGA07072 ---------- 8442
A2.CD.97.97CDKTB48 ---------- 8582
A2.CY.94.94CY017_41 ---------- 8647
A.SN.01.DDI579 ---------- 8413
A.SN.01.DDI369 ---------- 8464
A.SN.96.DDI360 ---------- 8395
A.CD.02.02CD_KTB035 ---------- 9241
A.CD.97.97CD_KCC2 ---------- 9241
A.CD.97.97CD_KTB13 ---------- 8521
B.AR.04.04AR151516 ---------- 8433
B.AU.87.MBC925 ---------- 9240
B.BO.99.BOL0122 ---------- 8416
B.BR.03.BREPM2012 ---------- 8696
B.CA.97.CANB3FULL ---------- 8491
B.CN.05.05CNHB_hp3 ---------- 9231
B.CO.01.PCM001 ---------- 8404
B.GB.83.CAM1 ---------- 9233
B.GB.86.GB8 ---------- 9212
B.GE.03.03GEMZ010 ---------- 8398
B.IT.05.SG1 ---------- 8556
B.JP.05.DR6538 ---------- 9257
B.KR.05.05CSR3 ---------- 9275
B.NL.00.671_00T36 ---------- 8776
B.RU.04.04RU128005 ---------- 8693
B.TH.00.00TH_C3198 ---------- 8394
B.UA.01.01UAKV167 ---------- 8479
B.US.04.ES10_53 ---------- 9214
B.US.99.PRB959_03 ---------- 8425
C.AR.01.ARG4006 ---------- 8439
C.BR.04.04BR013 ---------- 8733
C.BW.00.00BW07621 ---------- 8577
C.CN.98.YNRL9840 ---------- 8152
C.ET.02.02ET_288 ---------- 8450
C.GE.03.03GEMZ033 ---------- 8452
C.IL.99.99ET7 ---------- 8411
C.IN.99.01IN565_10 ---------- 8622
C.KE.00.KER2010 ---------- 8434
C.MM.99.mIDU101_3 ---------- 8604
C.MW.93.93MW_965 ---------- 8420
C.SN.90.90SE_364 ---------- 8369
C.SO.89.89SM_145 ---------- 8471
C.TZ.02.CO178 ---------- 8447
C.UY.01.TRA3011 ---------- 8422
C.YE.02.02YE511 ---------- 8473
C.ZA.04.04ZASK164B1 ---------- 8692
C.ZA.05.05ZAPSK240B1 ---------- 8624
C.ZM.02.02ZM108 ---------- 9285
D.CD.83.ELI ---------- 8760
D.CM.01.01CM_0009BBY ---------- 8421
D.KE.01.NKU3006 ---------- 8467
D.KR.04.04KBH8 ---------- 9122
D.TD.99.MN011 ---------- 8411
D.TZ.01.A280 ---------- 8412
D.UG.99.99UGK09259 ---------- 8420
D.YE.01.01YE386 ---------- 8380
D.YE.02.02YE516 ---------- 8395
D.ZA.90.R1 ---------- 8578
F1.AR.02.ARE933 ---------- 8338
F1.BE.93.VI850 ---------- 8498
F1.BR.01.01BR125 ---------- 8657
F1.ES.x.P1146 ---------- 8561
F1.FI.93.FIN9363 ---------- 8522
F2.CM.02.02CM_0016BBY ---------- 8349
F2.CM.95.MP257 ---------- 8399
F2.CM.97.CM53657 ---------- 8392
G.BE.96.DRCBL ---------- 9184
G.CM.01.01CM_4049HAN ---------- 8367
G.CU.99.Cu74 ---------- 8864
G.ES.99.X138 ---------- 8680
G.GH.03.03GH175G ---------- 9291
G.KE.93.HH8793_12_1 ---------- 8621
G.NG.01.01NGPL0669 ---------- 8450
G.PT.x.PT2695 ---------- 9256
G.SE.93.SE6165 ---------- 8663
H.BE.93.VI991 ---------- 8624
H.BE.93.VI997 ---------- 8539
H.CF.90.056 ---------- 8544
J.CD.97.J_97DC_KTB147 ---------- 8386
J.SE.93.SE7887 ---------- 8519
K.CD.97.EQTB11C ---------- 8404
K.CM.96.MP535 ---------- 8408

```
                               3' LTR U3 start                                                            Nef premature end in HXB2
B.FR.83.HXB2                   AGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTC...CCTGATTAGCAGAACTACACACCAGGGCCA...GGGGTCAGATATCCACTGACCTTTGGATGGTG 9221
Nef                            _S__H__F__L__K__E__K__G__G__L__E__G__L__I__H__S__Q__R__R__Q__D__I__L__D__L__W__I__Y__H__T__Q__G__Y__F__.__P__D__*__Q__N__Y__T__P__G__P__.__G__V__R__Y__P__L__T__F__G__W__C
01_AE.CF.90.90CF11697          ---TT---------------------------T---------T-----A-G-A---------G--------CT-A---G----T--T---------------...-------G---A------------------...---A-------C------TGT----------- 9188
01_AE.CN.05.FJ051              ---TT--------G------------------T---T-----T-----A-G-A---------A---------T-A---G----TA-T----------T----...-------G---A------------------...---A-------C------TGT----------- 9237
01_AE.CN.06.FJ054              ---TT-C------G------------------T---------T-----A-G-A---------A---------T-A---G-----A-T----------T----...-------G---C------------------...---A-------C---T-ATGT----------- 9233
01_AE.HK.x.HK001               ---TT---------------------------T---------T-----A-G-A---------G--------CT-A---G-T--TA------------T----...-------G----------------------...---A-T------------TGT----------- 8603
01_AE.TH.01.01TH_R2184         ---TT--------G------------------T---------T-----A-G-A---------G--------CT-A---G----T-------------T----...-------G-G-C------------------...-............................... 8408
01_AE.TH.02.OUR769I            ---TT--------G------------------G---------T-----AGG-A---------G---------T-A---G----TA-T----------T----...-------G---C------------------...-............................... 8390
01_AE.TH.90.CM240              ---TT---------------------------T---------T-----A-G-A---------G--------CT-A---G----TA-T----------T----...-------G---A------------------...---A------TC------TGT----------- 8789
01_AE.US.00.00US_MSC1164       ---TT---------------------------T---------T-----A-G-A---------G---------T-A---G----TA-T----------T----...-------G---C------------------...-............................... 8395
02_AG.CM.02.02CM_4082STN       ------------------------------------T-----T-----A-G-AG----T---A-----------A---G----TA----------A------...--A----G----------------------...-............................... 8413
02_AG.EC.x.ECU41               ------------------------------------T-----T-----A-G-A---------A---------------G----T-----------A-T----...--A--C-G----------------------...-............................... 8291
02_AG.GH.03.GHNJ196            ------------------------------------T-----T-----A-G-A------G--G---------------G----T-----------A-T-C--...--A----G----------------------...---A-------C-------------------- 9334
02_AG.NG.01.PL0710             ------------------------------------T-----T-------G-A---------G---------------G----T-----------A------...--A----G------------T---------...--AA-T-----C---------C---------- 8405
02_AG.NG.x.IBNG                --------------------------------T---T-----T-----A-G-A---------A---------------G----TA----------A-T----...--A----G----------------------...---ACT----TC--------------G----- 8748
02_AG.SE.94.SE7812             ------------------------------------T-----T-----A-G-A---------A---------------G-T--T-----------A-T----...--A----G----------------------...---ACT----TC--------------G----- 8611
02_AG.SN.98.MP1211             ------------------C-----------------T-----T-----A-G-A---------G---------------G----T-----------A-T----...--A----G------T---------------...---A-T-----C--------------G----- 8456
03_AB.RU.97.KAL153_2           --------------------------------T---T-----T-----A-G-A---------A---------------G----T------------------...-------G---.........------.......A-T----TC-----A--A-----G----- 8386
04_cpx.CY.94.CY032             ------------------------------------------T-----A-G-A---------G---------------G----T-----------T-T----...-------G-G-T--------------A---...----AG----TC------TG------------ 8633
05_DF.BE.x.VI1310              --T---------------------------------T-----T-----A-G-A---------------------T---G-----A----------T----C-...-------G----------------------...---A-T-----------------G--T----- 8630
06_cpx.AU.96.BFP90             ---TT---------------------------T---------T-----A-G-A------G--G---------------G----TA----------T-T----...-------G---A------------------...---AC-----TC--------T----------- 9266
08_BC.CN.97.97CNGX_6F          ---TT-------------------------------T-----T----TA-G-A---G-----G---------T-----G----T------------------...-----C-G---C--------------A---...----------TC--------T-----G----- 8421
09_cpx.GH.96.96GH2911          --TTT---------------------------T---T-----T-----A-G-A---------A---------------G-----------------------...--A----G---A--------------A---...----------T----T---------------- 8396
10_CD.TZ.96.96TZ_BF061         G--TT---------------------------T---T-----T----T--G-A---GA----C---------T-----G----TA------------T----...-------G---A--------------A---...---AC---------G----------------- 8576
11_cpx.GR.x.GR17               ---TT---------------------------T---------T-----A-G-A---------G--------C------G-TC-TA-T--------A-T----...-------G------T---------------...---A-------C------TGT----------- 8500
12_BF.AR.99.ARMA159            ------------------------------------T-----T-T---A-G-A---------G--------------------------------T------...-------G----------------------...----------TG--------------G----- 9186
13_cpx.CM.96.1849              G--TT---------------------------T---------------A-G-A---------C--------C------G----T--T---------------...-------G---A-----------------G...-----------C------TGT----A------ 8646
14_BG.DE.01.9196_01            ---TT---------------------------T---------T-Y-----GM-------------------C------G----TA-T--------A-T----...--A----G----------------------...---AYT----TC--------------G----- 8750
15_01B.TH.99.99TH_MU2079       ---TT---------------------------T---------T-----A-G-A---------G---------T-A---G-T--TA-T----------T----...-------G---A------------------...---A-------C------TG------------ 8596
16_A2D.KR.97.97KR004           -------------G------------------T--------------A--G-A---G-----C---------------G------------------T----...-------G----GGT---------------...--AA-T-----C-----A--A----------- 8638
17_BF.PE.02.PE02_PCR0155       ------------------------------------------T-----A----------G--A---------------G-----A-----------------...-------G----------------------...---A----------C--------G--G----- 8417
18_cpx.CU.99.CU76              ---TT---------------------------T---T-----TT--T---G-A---G-C---G---------T-----G-T-T-A------------TTC--...-------G------T---------------...--A-------TC--C--A-------------- 8550
19_cpx.CU.99.CU7               ------------------------------------A--G--TGG--T--G-A---------C-----------T---G----T-----------T------...-------G---T--T---------------...---A----G----------------------- 8286
20_BG.CU.99.Cu103              ---TT---------------------------T---------T-----A-GCA---------C--------C------G----TA-T---------------...--A----G----------------------...---ACA----TC--------------G----- 8667
21_A2D.KE.91.KNH1254           ---TT---------------------------T---T-----T------CG-A---G-----A---------------G-----------------------...-----C-G------T---------------...---A-------C-----A--A----------- 8446
22_01A1.CM.01.01CM_0001BBY     ---TT---------------------------T---------T-----A-G-A---------G---------T-A---G----TA-----------------...-------G---A------------------...-............................... 8372
23_BG.CU.03.CB118              ---TT---------------------------T---------TGG---A-GCA---------C--------C------G----TA-T----------T----...--A----G----------------------...---A-A--------------------G----- 8676
24_BG.CU.03.CB378              G--TT---------------------------T---------T-----A-GCA------G--C--------C------G----TA-T--------A-T----...--A----G----------------------...---ACA-----C--------------G----- 8736
24_BG.CU.03.CB471              ---TT-------------------------------------T-----A-G--------G--C--------C------G----TA-T--------A-T----...--A----G----------------------...---A-A----T---------------G----- 8697
25_cpx.CM.01.101BA             G--TT---------------------------T---------------A-G-----------C---------------G-T-----------G---------...-------G----------------------...---A-A-----C-----A--A-----G----- 8498
27_cpx.CD.97.97CDKTB49         -----T--------------------------T-------C-T-----A-GCA-------------------------G----TA----------T------...-----C-G---A--------------A---...-----G--------C----------------- 8428
28_BF.BR.99.BREPM12609         -------------G-----------------------------------C--A---------G---------------G-T---------------------...-------G----------------------...---AC-----TC-------------------- 8634
29_BF.BR.02.BREPM119           -----T---------------------------------G--TGG-----G-A-------------------T-----G-----------------------...-------G---A------------------...---A----------------------G----- 8373
31_BC.BR.02.110PA              ---TT-------------------------------T--.--T-----A-G-A---G-----C---------------G-T--TA-----------------...-------G---A------------------...---AC---------C----------------- 8712
32_06A1.EE.01.EE0369           G--TT---------------------------T---------T-----A-G-A---------G---------------G----T-----------T------...-------G---A--------C---------...---A-------C---T---------------- 8847
33_01B.MY.05.05MYKL007_1       ---TT---------------------------T---------T-----A-G-A---------G--------CT-A---G----TA-T----------T----...-----C-G---A------------------...---AC-----TC------TGT----------- 8612
34_01B.TH.99.OUR2478P          ---TT---------------------------G---------T-----A-G-----------G---------T-A---G----TA-T----------T----...-------G---A------------------...---AC-----TC------TG------------ 8387
35_AD.AF.05.05AF095            ---TT---------------------------T---T-----TGG---AGG-A---G--G--A---------------G-----A-------G---------...-------G------T---------------...--A-------T----T-A--A----------- 8402
36_cpx.CM.00.00CMNYU830        ---TT------------------A--------T--------C--------G-A------------------CT-A---G----T--T----------T----...-------G---A--------------A---................................... 8410
37_cpx.CM.00.00CMNYU926        ---TT---------------------------T---T-G--------T----A---------C---------------------A-----------------...-------G------T---------------................................... 8396
38_BF1.UY.05.UY05_4752         -----T------------------------------------T-----A-G-A-------------------T-----G----------------G------...-------G------T---------------...---A------T---------------G----- 8641
39_BF.BR.03.03BRRJ103          ------------------------------C--------G--TGG-----G-A-------------------------G-T---------------------...-------G------T---------------...---ACT----TC--------------G----- 8699
40_BF.BR.04.04BRRJ115          --T---------------------------------------T-----A-G-A---------G--A-----C------G-T--TA----------T-T----...-------G----------------------...-----T--------------------G----- 8802
42_BF.LU.03.luBF_05_03         ---------------------------------------G--T-------------------C--------------------T------------------...-------G------------C---------...---A----------C----------------- 8683
43_02G.SA.03.J11223            G--TT----------------------------------G--T-----A-G-A------G--A---------T-A---G----T-----------A-T----...--A----G---T-----T--T---------...---A-T-----C--------------G----- 8749
U.CD.83.83CD003                G--TT---------------------------T---------T-----A-G-A---------G---------------G----TA----------T------...--A----G----------------------...---A--------------------------CC 8688
U.CD.90.90CD121E12             ---TT-------------------------------------T-----A-G-A---------G---------------G----TA-----------------...--C----G----------------------...---A---------------------A------ 8383
U.GR.99.GR303                  ---------------------------------T--T-------------GCA---------C---------------T-----A-----------------...--A----G------T---------------...--AA-------C-----A--A-----G----- 8532
U.NL.95.U_NL_95_H10986_D1      ---TT---------------------------T-----------------G-A------G------------------G----T-----------T------...-------G---A------------------...---A-------C-----A--A----------- 9148
CPZ.CD.90.ANT                  TCATGN--------------------------TN--------T-NN-A-TT-A----GC--C------A---A-------G-TTA----C--G--AAT----...-------G------T------GA--A----...--A--T-----------CTGTAGA-------- 8675
CPZ.CM.01.SIVcpzCAM13          TC-----------G--G-------------------T--G--T-------G-A---G-----A---T-A-----C---G-T--T--T----------T----...-----C-G------------C-----A---...--A-AG-----C----CCTG------------ 8860
CPZ.CM.05.SIVcpzEK505          TC-TT--------G-------------------------G--TGG--AACC-AG--------A--T--A--C--C---G-T--T--T--T-----AAT----...-----C-G------------T-----A---...--C--------C--C----------------- 8728
CPZ.CM.05.SIVcpzLB7            G--T--------------------------------T--G--T----TA---AG--------A--------C--CG--G-T--T-----------AAT----...-----C-G------T-----C---------...---A-T---------A-A--------CG---- 8753
CPZ.CM.05.SIVcpzMB66           ---------------------------------------G--T-----A---AG--------G-----------T---G-T--T-----------A-T----...-------G---A------------------...---A----------------------C----- 8866
CPZ.CM.05.SIVcpzMT145          TCT---------------------------------T----CTGG--AAGG-A------G--G---------T-A---G-T--T-----------T------...-----C-G---A--------C--G------...--A--GC-C--C-----A-----C--G----- 8754
CPZ.GA.88.GAB1                 TCA---------------------------------T--G--T-----AGG-----------G--------C--C---G----T-------------T----...-----C-G--------------A---A---...--AACA----TC-----CTGT----------- 9294
CPZ.TZ.00.TAN1                 TCATGG--C-----------------------T-----T---TT---T-C--A----GC--CC-----A--CACC-----G--TA-T-----G--TGT---T...--A--C-G------------C--T--A---...--AA-------C------TGTAGG-------T 8907
CPZ.US.85.CPZUS                G--TT-------------------------------T--G--T-T---AG------------G-----A-----C---G-T-----T--------AAT----...-----C-G------T-----C--------T...--A--A--------C--A--A-A---G----- 9260
CPZ.CM.x.SIVcpzDP943           G--T--------------------------------------T-----AG------------G---T-G-----C---G----T--T--------A-T----...-------G------T---------------...--A--------------A--A-----G----- 8814
CPZ.CM.x.SIVcpzMB897           -----------------------------------------CT-------------G-----G-----------C---G----T--T----------T----...-------G---A--------------A---...-----------C--GG-------C--C----- 8756
N.CM.02.DJO0131                TC-TT-------------------------------------TGG---AG--A---G-----CG-TT-A-----C---G-T--T-------------T----...-----C-G----------------------...--AACT----TC--C----------------- 8696
N.CM.04.04CM_1015_04           TC-TT----------------------------------G--TGG---AG--A---G-----G--T--A--C--C-----T--TA-----------------...-----C-G----------------------...--A-----G-----C-----A----------- 8665
N.CM.04.04CM_1131_03           TC-TT----------------------------------G--TGG---AG--AG--G-----G--T--A--C--C---G-T--T-------------T----...-----C-G----------------------...--A--T-----C--C----------------- 8733
N.CM.95.YBF30                  TC-TT------------T---------------------G--TGG---AG--A---G--------T--A--C--C-----G--T------------AT-C--...-----C-G---T------------------...--AA-T-----C--CG-A-------------- 8811
N.CM.97.YBF106                 TC-TT----------------------------------G--TGG---AG--A---G-----A--T--A--C--C---G-T--T-----------T-T----...-----C-G----------------------...--A--T-----C--CY--TG------------ 8754
O.BE.87.ANT70                  ---TT------------------A------------T-----T-------T-A----GC---A-----G-----T---G-G--TA----T--G--A-T----...-------G------------------A---...--AAC---G-TC--------A----------T 9307
O.CM.91.MVP5180                ---TT------------------A--------T---T-----T-------T-AG---GC---A-----G-----C-----A-T------T--G--A-T----...-------G----TGT--------G--A---...--ACCT----TC--------A----------T 9330
O.CM.96.96CMABB637             G--TT------------------A--------T---T-----T------CTGA----GC---A-----G-----T---G-G--T-----T--G--A-T----...-------G------------------A---...--AACT-----C--------A---------CC 8812
O.CM.98.98CMA104               ---TT------------------A--------C---T-----T----AG-TGC----GC---A-----G-----T---G-GC-TA----T--G--A-T-M--...-------G------------------A---...--AAC-----TC--C-----A---------AT 8811
O.CM.99.99CMU4122              --TGT---------------T---------------T-----T----T--TG-G---GC---G-----G-----T-----A--TA----T--G--A-T----...-------G----------------------...--AACT----TC--------A-----G----T 8777
O.FR.92.VAU                    ---TT------------------A--------T---T-----T-------T-A----GC---G---T-A--C--T---G-GC-T-----T--G--A-T----...-------G---A--------------A---...--AAC-----TC---T-A--A---------CT 8879
O.SN.99.SEMP1299               ---TT------------------A--------T---T-----T-------T-A----GC---G-----G-----T---G-A--T-----T--G--A-T----...-------G----GG------------A---...---CCA----TC--------A-----G----T 9369
O.US.99.99USTWLA               --T--T--------------C--A--------T---------T-------T-A----GC---A-----G-----C---G-A--T-----T--G--A-T----...-------G------T-----------A---...--ACCT-----C--------A-----C---CT 8778
```

```
B.FR.83.HXB2          CTACAAGCTAGTACCAGTTGAGCCAGATAAGATAGAAGAGGCC...AATAAAGGAGAGAACACCAGC...TTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCACGTGGCCCGAG 9385
Nef                   __Y__K__L__V__P__V__E__P__D__K__I__E__E__A__.__N__K__G__E__N__T__S__.__L__L__H__P__V__S__L__H__G__M__D__D__P__E__R__E__V__L__E__W__R__F__D__S__R__L__A__F__H__H__V__A__R__
A1.GE.99.99GEMZ011    --------------------C---TC-G-AG------------...-CAG-----------A----...C-A---------A-AT--AA--C--A--------TGA---A-----------AT-----A---------T-----G--TC-AACA---AGA------- 8584
A1.KE.00.KER2008      --T------G--G--------T------G-AG-----AGA--T...-C-G-G----------A-----...C---------GA-AT-T-AA----A--------TGAA-----------A---A-----CA---------------G------ACA---CAA---A--- 8585
A1.KE.00.KNH1144      --T-----------------T------G-AG-----A----T...-C-G-G----------A-T--...C---------GA-AT-T-AA----A--------TGAA--G-------C---AG-----A-----------A---G---C-AG-A---A-C-----TA 8635
A1.KE.00.KSM4024      --TT----------------T------G-AG----GA----T...-C-G-G--------T-A-----...C-A-------GA-AT-T-AA----A--------TGA-----A------A---C-----A-----------A---G---AGAA-A---AGA--T-TT- 8526
A1.KE.00.MSA4069      --T-G------------G--T--T---G-AG----GA----T...-C-G-G----------A-----...C---------GA-AT-T--A----A-----G--TGA-----A------A--GA-----A---------------G---C-AAGA---AGA-----T- 8608
A1.KE.00.NKU3005      --T--------------A-CT-----KG-AG----GA----T...---G-G---------AT----...C-C--------A--T---A-----A--------TGA---T-----ACA---AT-----A---------------G---C-AA-A---AGA----A-- 8611
A1.RU.00.RU00051      --------------------C---RT-G-AG------------...-CGG-G---------A-T--...C-A---------A-AT--AA--C--A--------TGAT--AG-------A---AT-----A---------T-----G--TC-AA-A---AGA----A-- 8683
A1.RW.93.93RW_024     --T-----------------T------G-AG----GA---AG...-C-G------------A-----...C-A---------A-AT--AA----A--------TGA-----A-----------A-----GT---------------G---C-GA-A---AGA----A-- 8611
A1.SE.95.SE8891       --T-----------------T-----CG-AG----GA----T...-C-G-G----------A-----...C-A-----T---A-AT--AA----A--------TGA----------ACA---AT-----A---------------G---C-CACA---AGA---A--- 8563
A1.SE.95.UGSE8131     --T-----------------T------G-AG----GA----T...-C-G-G----------A---T...C-A-----T--AA--T-T-AA----A--------TGAA-------GACA---AT-----A------CC--A---G-------A-A---AGA---TTC- 8763
A1.TZ.01.A173         --T---------T-------T-----AG-AG----GA----T...-C-G-------------C--...C-A---------A-A----A-----A--------TGAT---------ACA---AG-----CA---------AAG--G-----AA-A---A-A---A--- 8582
A1.UA.01.01UADN139    --------------------C----C-G-AG-G----------...-CAG-G---------A-----...C-A---------A-AT--AA--C--AG-------TGA-----AG--G-------AT-----A---------T-----G--GC-AACA---A-A------- 8620
A1.UG.92.92UG037      --T--------------G--TGA----G-AG----G--A--T...-C-GG-----------AT----...C-A---------A-AT--AA----A--------TGA-----A---ACA---AG-----A-----------A-----G---CGAGTA---AAA--AA--- 9443
A1.UG.99.99UGA07072   --T-----------------T------G-AG-G--GA----T...-C-G-G----------A-----...C---------A-AT--AA----A--------TGA----------ACA---AT-----A---------------G-----AA-A---AGA----T-- 8606
A2.CD.97.97CDKTB48    --T-----------------T---TC-G--G----------T...-CGG-G--G-------A---T...--A--------CA-AT--AA----AGCA--G-----T--A----------------A-----A------T--T-----G--GC-GAGA---C-A---A--- 8746
A2.CY.94.94CY017_41   --T--------------A------TC-G--G-------A--T...-C-C-G----------A-----...--A---------A-AT--AA----AG-A---------T--A-------------AGA---GA------T--AA-----G---CGGAGA---AGA------- 8811
A.SN.01.DDI579        --T-----------------T-A--C-TCAG------------...-C-G-G------............C-----------AT--AA----A--------TAA-----A------A---AT----C----------T-----G---C-AA-A---ACA--T---- 8568
A.SN.01.DDJ369        --T-----------------T-A--C-G-AG--------A-T...-C-GGG----------A-----...C-----------AT--AA----AG-------TA-----G----G-------AT-----A---------T-----G---C-AA-A---A-T--TGC-- 8628
A.SN.96.DDJ360        --T--------------G--T-A--C-GCAG------------...-C-GGG----------A-----...C----------A-AT--AA----A--------TGA-----A------A---AT----C----------T-----G---C-AA-A---A-A--T-A-- 8559
A.CD.02.02CD_KTB035   --T-----------------TT-----GCAG----G-----T...-C-G--A---------A-----...C-CC------GA-A-A--AA--A-A--------T--T----AG---------AT----CA---------T-----G---C-GA-A---C-A----A-- 9405
A.CD.97.97CD_KCC2     --T---T-----------------AGCAG------------T...-C-GGG---------AT----...C----------A-AT--AA----A--------T--T------------------AT-----A---------T-----G-----AA-A---AGA---AA-- 9405
A.CD.97.97CD_KTB13    T------T-----------------AGCAG-----AG---T...-C-G-G--------T-A-T--...C-A---------A-AT--AA----AGC---G--T--T-----------------AT-----A---------T-----G---C-GA-A---C-A----A-- 8685
B.AR.04.04AR151516    --T-----------------TAA---GC--G-----A-----...---G-----------A-T-T...------------G---------------G---T------A-----------ATA----A---------------G------------A-------- 8597
B.AU.87.MBC925        --T----------------------G---------------...---G-----------A----...-----------A-----------------G----------------------------------------C--C------A-------- 9404
B.BO.99.BOL0122       --T-------------C----T-----CC-AG-----AGA---...---G-----------A-----...C--C-G------A------AA-------------------------A---------G-T------A-----------------C---------TA-------- 8580
B.BR.03.BREPM2012     --T-------------T-A--------GG--G-----A-----...---G-------A---A-----...---C--------A-A-----------------G---G-----G---------G-T------------------------G-GC---C-A-------- 8860
B.CA.97.CANB3FULL     --T----------------------G---------------...---G-----------ATGT-...-----------A-----------------G---------G--------------T-----A-----------------------------A-------- 8655
B.CN.05.05CNHB_hp3    --T-----------------T-----GC--G-------A----...-C----A---------A-T--...C--C----T---A------AA----A-CA-------A-------------C-GAT-----A---------------------C--C------A-------- 9395
B.CO.01.PCM001        --T---------------C--------GG--G-G---A-----...---G-----------A-----...-----------A---------------------A---A-------C-GC------A----------------C-----C-----------A---A--- 8568
B.GB.83.CAM1          --T----------------------GC--G-------------...------A-------G-----...-----------A-----A-------------------A---A-----AC--AT-----A--------------------------------A-------- 9397
B.GB.86.GB8           --TT--A-------------------GG--G-----A-----...---G-----------A-T-T...-----------A-----A-------A--G---------A-----------T-----A---A----------------------------A-------C- 9376
B.GE.03.03GEMZ010     --T-----------------T-A---GC--G-----A-----...---G-----------A-T-T...----------AA-A---A-------G---------A---A-----AC--AT-----A---------AAG-------------T--A---AA-- 8562
B.IT.05.SG1           ..................................................................................................................................................................... 8556
B.JP.05.DR6538        --T-------------A-G--------GC--G----G-----...------------------A-T--...C----------A-----------------A-----A----------------T-----A---------A-------C-------AG-------- 9421
B.KR.05.05CSR3        --T----T--------------AA---G---G---------T-...-C-G--A---------------...C--C--------A-A---A-----------A-G-----------------------AT-----A---------------G--------------A------- 9439
B.NL.00.671_00T36     --T-----------------T-----CC--G-----A-----...---G-----------A-T-T...-----------A-----AA--------A-----------------A---A-----T-----A--------------------------------A-------- 8940
B.RU.04.04RU128005    --T-----------------T-A--G--AG------------...---G-----------A-----...-----------A------A---------------------A---A-----------AT-----A---------------------T-----C---A-A------ 8857
B.TH.00.00TH_C3198    ..................................................................................................................................................................... 8394
B.UA.01.01UAKV167     --T-----------------T------C--G-----A-----...---G-----------A-T--...C----------A-----A--------------------A---A-----------AT-----A--------------------------------A-------- 8643
B.US.04.ES10_53       --T--------------G---A----G---G------------...--CG-----------A-----...-----------A-A-----------------------A---A-----------C------A-----------------------C--C-GC------------ 9378
B.US.99.PRB959_03     --T----T----G-----A-----------G------------...---G-G-----------A--T-...--------T--CA-A-----------------G---------A-----------C------A-----------A------C----G---A-------- 8589
C.AR.01.ARG4006       --T-----------------C---AGGG-AG-R----------...-Y-C-----------A-----...C--C--------A-----------A----G--TT-A--CG------A---C------A---------T-T-------CGCAGA---A-------C- 8603
C.BR.04.04BR013       --T---------G--------C---A-GG-AG-G----------...---G------T----A-----...C--C--------A-----------A----G--TGAAC-CG---------GC------A---------T-T-------CGCAGA---A-------C- 8897
C.BW.00.00BW07621     --T-----------------C---A-GG-AG------------...---G---G--A---A-T-T...---C--------A-----A---A-A--A----TGATC-C-------A---AT-----A---------TA--------CGCAGA-----------C- 8741
C.CN.98.YNRL9840      ..................................................................................................................................................................... 8152
C.ET.02.02ET_288      --T---------G--------C---AG-G-AG------------...---G-----------A-T--...C--C-G------A-A-----------A-------TGAA--C-A-----A---AT-----A--------T-TC-------CACAGA---C-------C- 8614
C.GE.03.03GEMZ033     ------------G-----G--C----GGG-AG--------A-...--C-------A---A-T--...---C--------A-----AA----AG-A--------A--T-A-----A---AT-----A----------A---------CGCAGA---C-------C- 8616
C.IL.99.99ET7         --T---A----------G--T---A-GG-AG-------A---...--CG-------A---A-T-T...---C--------A--------------A--A--G--TGAA--C-------A---AG-----A---------T-AT------CGCAGA---A-------C- 8575
C.IN.99.01IN565_10    --T-----------------C---AGAG-AG------------...--CG-------A---A-C-T...C--C--------A--T---A-----A--A--G--TGGA--C-------C---AT-----A-----------A------CGCA-A---A-------C- 8786
C.KE.00.KER2010       --T---------G--------C---AGGG-AG------------...--C----A---A---A-T-T...---C--------A-----A-----A--A--G--TGAA-----------A---AG----CA---------T-T-------CACAGA---C-------C- 8598
C.MM.99.mIDU101_3     --T-----------------C---AGGG-TG------------...-CC--------AG--A-T-T...---C-------A--TT---A-----A-----G--TGAAC-CG--------C-GC------A---------T-AA------C-CAGA---AG------C- 8768
C.MW.93.93MW_965      ------------G-----C--C---AGGG-AG------------...---G-------A---A-T-T...---C--------A-----A-----A--A--G--TGAA--C------A-A---AGA----A---------T-AA------CGCAGA---C-------C- 8584
C.SN.90.90SE_364      --T-----------------C---AGGG-AG------------...---G-------A---A-T-T...---C--------A-----------------A-----G--TGAA--M--------A---AT-----A--------TT-AA------CACAGA---A-------C- 8533
C.SO.89.89SM_145      --T-----------------C---AGGG-AG---------A-...--C-------A---A--T...---C--------A-----A-----A-----G--TGAA--C--------C---AT-----A---------T-A-------CGCAGA--T--------C- 8635
C.TZ.02.CO178         --T---------G--------C---AG-G-AG------------...---G-----------A-T--...---C--------A-----------A-----G--TGAA--C--------AC--A------A---------T-T-------CGCAGA---A-------C- 8611
C.UY.01.TRA3011       --T---------G--------C---AGGG-AG-G----------...---G-------A---A-----...---C--------A----TA--------A-----G--T--AC-C-A------A---C------A---------T-T---G---CA-AGA---A-------C- 8586
C.YE.02.02YE511       --T-----------------C---AGAG-AG-------A---...--CG---------T-A-----...---C--------A-A-----------A-----G--TGGT--CG------A---AT-----A-----------------GCACAGA-----------C- 8637
C.ZA.04.04ZASK164B1   --T---A-------------C---AGGG-AG------------...-GC-GT-------C--A-T--...C--C------GA-A-A--A-----A--------TGAA--C--------------T-----A-------------------CACATA---A-------C- 8856
C.ZA.05.05ZAPSK240B1  A-T-----------------C---AGGG--G----G-------...---G-------AG---A-TTT...--AC--------A-----------A-----G--TGAA--AG------A---A-----A-------------A------CGCAGA---C-------C- 8788
C.ZM.02.02ZM108       --T-----------------C----GGGGAG-------A---...-G-G-G------G---A-T-T...---C--------C--A--A-----A--------TGAAC-C--------A---A-----A---------T-T-T-----CGCAGA---A-------C- 9449
D.CD.83.ELI           ----G---------------T---C-GG--G-------A-A-...-C-G---------C--A----...-----------A-AT---A-----A-----G-----------C---------A-A-----A---A-----A-A---------G-G---AA-------- 8924
D.CM.01.01CM_0009BBY  ..................................................................................................................................................................... 8421
D.KE.01.NKU3006       --T-G----------------AT-A-GG--G---------AAT...-C-G-----------A-T--...-----------A--T---A-----A-----G-------------------------AT-----A---A----TA-A-----C---G-A---AA-------A 8631
D.KR.04.04KBH8        --T-G----------------AT----GG--G------------...---G-----------A-T--...C-----------A-----A-----A--A--G-----------C-----C--A------A---A-----AAA---------G-A---AA-------A 9286
D.TD.99.MN011         --T-G---------------C--CA-GGT-G---------AG...-C-G-----------A-T-T...---C--------A--A--A-----AGG---C---------A-------C---T------A---A-----A-A---------G-A---AAA---AA-- 8575
D.TZ.01.A280          --T-G---------------C---A-GG--G--------A---...-C-G-----------G-T--...-----------A-A----A-----A-----G---A-----A-------------TA-----A---A-----A-A---------G-A---AA-------A 8576
D.UG.99.99UGK09259    --T----------------------G---G------------...-C-G-----------A-T--...---C--------A-A-A--A-----A-----G-----T---A-------------ATA-----A-----------A-A---------G-G---AAA----A-C 8584
D.YE.01.01YE386       --T-G---------------T---A-GG--G---------A-...-C-G-----------G-T--...-----------A--T---A-----A-----G---AA----A-----A-----A---AT-----A---A-----AAG---------G-A---AA----AA-A 8544
D.YE.02.02YE516       --T-G---------------T--CA-G-T-G-------A-AG...--CG--------G---AGT-T...---C--------A--CA--A-----A----GG----A---A-----------------T-----A---A-----AAA---------G-A---AAA---A--A 8559
D.ZA.90.R1            --T--------------G--T---C-GG--G-----C------...-C-G-G-------C--A-T--...-----------A---A--A-----A-----G---------------------------A-----AA--------A-A---------G-G---AA-------- 8742
F1.AR.02.ARE933       ..................................................................................................................................................................... 8338
F1.BE.93.VI850        --T-----------------T----AG--G-G---A-----...---G-----------A-T--...-----------CA------AA----A-----G--TGAA--C--------AC-GAG-----A------------A----G---C-GAGA---A-A---A--- 8662
F1.BR.01.01BR125      --T-----------------C-----GG--G-----A-----...-G-G-G-----------A-T--...---C-------CA-A-A--A-----A-----G--TGAA--C--------AC--C------A-----------------------------A-------- 8821
F1.ES.x.P1146         --T---------------CCT------G--G-----A----T...---G---------CT-A-T--...---C-------CA------AA----A-----A--TG-A--CG--------AC-GAT-----A-----------ATG--G------AGA-----A---AA-- 8725
F1.FI.93.FIN9363      --T----------------------GG--G-----A-----...---G-----------A-T--...---C-------CA------AA----A-----G--TGAA--C--------AC-GA-A----A-----------------G---C-GA-A---A-A---A--- 8686
F2.CM.02.02CM_0016BBY ..................................................................................................................................................................... 8349
F2.CM.95.MP257        --T-----------------T----GG--G-----A-----...---G-----------A-T-T...-----------A------A----A-----G--TGAC--C-AG-----------A-----CA---------------------GC-GAGA--TA-A---A--- 8563
F2.CM.97.CM53657      --T----------------------GG--G-----A-----...---G-----------A-T-T...--A--G------A------AA----A-----G---GAA--C--G-----------A---------A-----------TC-------CGGAGA---AGA---A--- 8556
G.BE.96.DRCBL         --T---A--G------A-G------TCAG--G----G--A---...-----G-----------A-----...C----------CA-CT---A-----A-----G---GAA--TG--------C-G-TA-----A--------TA----G---CGGAGA---T-A------- 9348
G.CM.01.01CM_4049HAN  ..................................................................................................................................................................... 8367
G.CU.99.Cu74          --T---A---------C-G--T----CAG-------G---------...-C--G-----------A---T...--A-----T--CA-CT---A-----A-----G---GAT--C--------TC-GATA-----A--------TA--------CGGAGA---T-A------- 9028
G.ES.99.X138          --T---A----A----A-G--T----CAG--G----G--A--T...-C-G------------A---T...--A--------CA-CT---A-----A-----G--TG-A--C-AT--G---C-GATA----AA--------TRA-------CGGAGA---A--------- 8844
G.GH.03.03GH175G      --T---A---------A-A--T----CAG--G----G----------...--------------C--A---T...C-C-------CA-CT---A-----A--A--G--TGAA--C-A-------C-G-TA----A--------TA--------CGGAGA--TC-T------- 9455
G.KE.93.HH8793_12_1   -CT---A------------G--T----CAGT-G----G--A---...-C--C--A---------A---T...C-A-------CA-TT-G-A-----A-----G---GA---C---------C-GATA-----A--------TA--------CGGAGA---C-A------- 8785
G.NG.01.01NGPL0669    --T---A---------A-GA-T----CAG-------G--A---...------------------A-----...--A-------CC-CT---A--C--A-----A--G-A--C---------C-G-TA-----A--------TA--------AGAAGA-----A------- 8614
G.PT.x.PT2695         --T---A---------A-G--T----CAG--G----G--A---...------------------A---T...--A-------CA-CT---A-----A-----G--TG-A--C-AG--G---C-G-TA-----A--------TA--------CGGAGA--TA-A------- 9420
G.SE.93.SE6165        --T---A---------A-G--T----CAG--G----G--A---...------------------A---T...C-A-------CA-CT---A-----A-----G---GAA--C---------C-G-T-----A--------TA--------CGGAGA---A-A------- 8827
H.BE.93.VI991         --T---A----G-----A--T---C-GG-TG-G---A------...---G-----------A-----...-----------A--T---AA----A--A--G--T--A-----------AC--AT-----A---------T--A------C-GAGA---AGA---AA-- 8788
H.BE.93.VI997         T-T-------A-------A--T---C-GG--G-----AG-----...---G-----------A-T--...C--C--T-----A-AT---A-----A-----G---GAA---G--------C-GAT-----A---------T--------C--CACA---ACA------- 8703
H.CF.90.056           --T---------------AA-T---C-GG--G-----C-----...---G-----------A-----...---C-------CA-------------A-----G--TGAC-G---------C-GAT-----AA--------T--A-------GACA---T--------- 8708
J.CD.97.J_97DC_KTB147 --------------------T---GGG-AG-----AG-----...---G-------AG---A-C--...C---------CA-CT---A-----A-----G--TGAA--C-------------AT-----A--------CTTC---G---CGGAGA--TA--------- 8550
J.SE.93.SE7887        --------------------T---AGCG-AG----G--A--T...---G-------A---A-T--...C---------CA-AT---A-----A--A--G--TGAA--A-------------C------A-----------TC-------CGGAGA---A-A------- 8683
K.CD.97.EQTB11C       --------------------T---AGAG-AG------------...-C-G-----------A-T--...C-------------A-A--A-----A-----G--TGAAC-C--------A---A------A---------TTC---G---CGAA-A--T--A------- 8568
K.CM.96.MP535         --------------------T----CAG-AG--------A--...-CAG--------G---A-T--...C---------CA-A-A--A-----A-----G--TGAAC-C------A-A---AT-----A---------TTC-------CGAAGA-----A------- 8572
```

B.FR.83.HXB2	CTACAAGCTAGTACCAGTTGAGCCAGATAAGATAGAAGAGGCC...AATAAAGGAGAGAACACCAGC...TTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCACGTGGCCCGAG	9385
Nef	__Y__K__L__V__P__V__E__P__D__K__I__E__E__A__.__N__K__G__E__N__T__S__.__L__L__H__P__V__S__L__H__G__M__D__D__P__E__R__E__V__L__E__W__R__F__D__S__R__L__A__F__H__H__V__A__R__	
01_AE.CF.90.90CF11697	--------------------T---A-AG-AG----G--A-A-...--C----AT--A-G--A-T--...C----------CA------AA-----AG-A-----TGAA--A---------C-GAT-----A---------TTCT------CGAAGA---A-A-------	9352
01_AE.CN.05.FJ051	--T-------A---------C---A-AG-AGC--------A-...-CC--------A----G-T--...C----------CA------AA-----A----GG--TGAAC-CG--------C-GAT-----A---------TTC-------CGAA-A---A-A-------	9401
01_AE.CN.06.FJ054	--T---------------C--C---A-GG-CG----G---AA-...-CC--------A------T--...C----------CA-A----A------AC-A--G---GAA--AG--------C-GAT-----A---------TGC-------CGAAGA-----A-------	9397
01_AE.HK.x.HK001	--T---A-------------C---AGAG-AG----G----A-...--C--------A----A-T--...C----G-----C-CA----A------A--A------GAA--A---------C-GAT-----A---------TGC-------CGAA-A---C-A-------	8767
01_AE.TH.01.01TH_R2184	....	8408
01_AE.TH.02.OUR769I	....	8390
01_AE.TH.90.CM240	--T-----------------C-A-AGAG-AG----G----A-...--C--------A----A-T--...C----------CA------A------A--A--G---GAA--A---------C-GAT-----A---------TGC-------CGAA-A-----A-------	8953
01_AE.US.00.00US_MSC1164	....	8395
02_AG.CM.02.02CM_4082STN	....	8413
02_AG.EC.x.ECU41	....	8291
02_AG.GH.03.GHNJ196	--T---A---------A-A--T----CAG--G----G--A---...---G------------A-GT-...--A--------CA-CT---A------A---------TGAA--C---------AC-G-TC----A---------------G-----CACA---ACA--AA---	9498
02_AG.NG.01.PL0710	--T---A---------A-G--T----CAG--G----G--A--T...-C-G------------A---T...--A--------CA-CT---A------A-----G--GATC-C---C-----C-G-TC----A---------T-AA--G-----CAGA--T--A--T-AC-	8569
02_AG.NG.x.IBNG	--T---A---------A-G--T----CAG--G----G--A---...---G------------A---T...--A--------CA-CT-T-AA-----A-----G--GAC--T---------C-GATC-----A--------TA-A--G-----CAGA---ACA--TA---	8912
02_AG.SE.94.SE7812	--T---A---------A-G--T----CAG-C-----GA-A-A-...-C-G------------A---T...--A--------CA-CT-T-A------A-----G--GAA--C---------C-G-TC-----A--------TA-A--G-----CACA---AAA--T----	8775
02_AG.SN.98.MP1211	------A---------A-G--T----CAG--G----G--A---...---C------------A---T...--A-----T--CA-CT-T-A-----A-----G--GAA--C-A-------C-G-TC-----A----------A-G--G-----CAGA---ACA--T----	8620
03_AB.RU.97.KAL153_2	------------------C--C----C-G-AG-----------...-CAG-G----------A-----...C-A----------A-AT---AA--C--A---------TGA---A-A-----------AT-----A---------T-----G--TC-AACA---AGA-------	8550
04_cpx.CY.94.CY032	--T---------------A--T---C-GG--G-G---------...-C-G---------------T-T...---C-G------A-A----A------A-----G--TGAA-----------A---A------A---------T-----G----ACA-G-----A-------	8797
05_DF.BE.x.VI1310	--T----T-----------A-T-----GG--G----GA-----...---G--------G---A-T--...---C-------CA-------------A-----G--TGAT--C---------AC-GC------A-----------A----G-----AAGA---A-A---A---	8794
06_cpx.AU.96.BFP90	--------------------T-----GG-AG-------ACTT...-C--------------A-T--...C----------CA-AT---A------AGCA--G--TGAA--A-----------------A------A----------TC-------AGGAGA---A-A---A---	9430
08_BC.CN.97.97CNGX_6F	--T-----------------C---AGGG-AG------------...--CG-------AG---A-T-T...---C----------CT---A------A-----G--TGAAC-C-------A---A------A---------T-AA------CACAGA---AG------C-	8585
09_cpx.GH.96.96GH2911	--T----T----G-----A--T---A-GG-AG----G------...-C-GG------------A----...--------------A--T---AA-----AG-----TCA...-GAG--------C-GAT-----------------AC-------C-GA-A--YA-A---A---	8557
10_CD.TZ.96.96TZ_BF061	T-T----T------------C---AGAG-AG-G----------...---------------A-T--...---C-------CA-------------A-----G--TAAAC-TG-------A----T-----A----------AC-------CACA-A---A-A-----C-	8740
11_cpx.GR.x.GR17	--------------------C---AGAG-AG------------...---G------------A-T--...C----------CA------A------A--A--G--TGAA--T------------AGA---A---------TTC-------CGAAGA---A-A-------	8664
12_BF.AR.99.ARMA159	--T-----------------C----TGG--G-----A------...---G------------A-T--...---C-------CA------AA-----A-----A--TGAA--C---------AC-GATA---A---------------G--GC-AAGA---T-A---A---	9350
13_cpx.CM.96.1849	--T----T--------A----T--CA-AG-AG----G------...---G--------A---A-T--...C--C-------AA------A------A--A--G--TG-A--A---------A---AT-----A---------TTC-------CGAA-A---A-A-------	8810
14_BG.DE.01.9196_01	--T---A-----------G--T----CAG-------G--A--T...-C--------------A---T...--A--------CA-CT---A------A-----G--TR-A--C-AT--G---C-GAGA----CA--------TA--------CGGAGA-----A-------	8914
15_01B.TH.99.99TH_MU2079	--T-----------------C---AG-G-TG----G---AA-...--C--------A----A-T--...C----------CA------A------A-----G---GAA--A---------C-GA-----CA---------TGC-------CGAA-A---A-A-------	8760
16_A2D.KR.97.97KR004	--T---------------CAGT----C-G--G----G----T...-C-G-G-------C--A----...--------------A-AT---AA-----A--A--G-----T--------------A-----GT------T--T-A---T---C-GGTA---AAA-------	8802
17_BF.PE.02.PE02_PCR0155	--T-----------------A-C-----GG--G-----A----...---G------------A-T--...---C-------CA------AA-----A-----G--TGAT--C---------AC-GAT-----A----------MT---G------AGA---C-A---A---	8581
18_cpx.CU.99.CU76	--T---A--------------TT--A--G-AG----------T...---G-G----------A---T...C------------A-AT---AA-----A---------TGA------------A---A------A---------TGC---G---CGGAGA---A-A--AA---	8714
19_cpx.CU.99.CU7	--T-G-------------G--T---A-GG--GC---GA-A-A-...-C-G------------A-T--...------------------C-T-T-AA-----A------------A----A------------T------A---A-----A-A---------G-R---AA----A--A	8450
20_BG.CU.99.Cu103	--T---A---------A-G--T----CAG--G----G--A---...-----T----------A----...--A-----T--CA-CT---A--------------G--TAAT--A---------C-GATA-----A--------TA--------CGGAGA---T-A-------	8831
21_A2D.KE.91.KNH1254	------A--------------T---TC-G--G---------AG...-C-C-G----------A----...--A-----------A-AT---AA-----AGCA--G-----C-----------A---AGA---GA------T--T-AA--G-G-C-GAGA-----A-------	8610
22_01A1.CM.01.01CM_0001BBY	....	8372
23_BG.CU.03.CB118	--T---A-----------A--T----CAG-TG----G--A---...-G---T----------A---T...--A--------CA-CT---A------A-----G---GAT--C-AG-------C-GAT------A--------TA--------CGGAGA---A-A-------	8840
24_BG.CU.03.CB378	--T---A---------A-G--T----CAG--G----G--A---...----GG----------A---T...--A-----T--CA-C----A------A-----A--TGAT--AG--------C-G-TA-----A--------TA----G---CGGAGA---A-A-------	8900
24_BG.CU.03.CB471	--T---A---------A-G--T----CAG--G----G--A---...----G-----------A---T...--A--G--T--CA-CT---A------A-----A--TG-T--AG--------C-G-TA-----A--------TA----G---CGGAGA---A-A-------	8861
25_cpx.CM.01.101BA	------A---------A-G--T-----AG-------G--A---...-------A---T----A----...--A--------CC-CT---AR-----A-----C--TG-TC-C-----------G-TA-----A--------TTC-------CGAAGA---A-A-------	8662
27_cpx.CD.97.97CDKTB49	------------------A--T---C-GG-AG-----------...---------------GT---...---C----T---A-AT---A------A-----G---GGA--CG-------A---A-----GT---------T---------C-GA-A---C-A-------	8592
28_BF.BR.99.BREPM12609	--TT----------------T-----AC--G------C-----...---G----------T-A----...C-----------A------A------A------------A----A-------C--AT-----A---------------------------A---------	8798
29_BF.BR.02.BREPM119	--T-----------------C-----G---G-G-----A----...-C-G------------A-T--...-----------------AT--------------G-----A----A-----------A-A----A-----------AAA--G------------T---------	8537
31_BC.BR.02.110PA	--T---------G--------C---AGGG-AG-G---------...---G------------A----...C--C----T---A----------C--A-----G--TGAAC-CG-------A---C------AA--------T-T-------CA-A-A---A-A-----C-	8876
32_06A1.EE.01.EE0369	---------G-----------T---AGGG--G-------A--T...-C--------------A-T--...C----------CA-AT----------AGCA--G--TGAA--AG--------------AT-----------------TC-T-G---CGGAGA---A-A---A---	9011
33_01B.MY.05.05MYKL007_1	--T-----------------C---A-AG-AG----G-------...--C----A---A--T-A-T--...C----------CA------A------A--A------GAA--C---------C-GAT-----A---------TGC-------CGAA-A-----A-------	8776
34_01B.TH.99.OUR2478P	--T-----------------C----GAG-AG--A-G---................--A----G-T--...C----------CA------A------A--A--G---GAA--A---------C-GAT-----A---------TGC-------CGAA-A---A-A-------	8539
35_AD.AF.05.05AF095	------A--------------T----T-G-AG----GA----T...--CG-G----------A----...C------------A--T---AA-----A---------TGA---AG-------A---AT-----A---------------G-----AA-A---AGA----A--	8566
36_cpx.CM.00.00CMNYU830	....	8410
37_cpx.CM.00.00CMNYU926	....	8396
38_BF1.UY.05.UY05_4752	--T-----------------C------G--G-----A------...---G-G--G-------A-T--...---C-------CA------AA-----A--A-----TG----C---------AC-GC------A---------------G------ACA-----A--TAA--	8805
39_BF.BR.03.03BRRJ103	--T------------T-----C------G--G-----AG----...--CG-------A---A-T--...---C-------CA------AA-----A---------TG-A--C-AG-------C-GA-----GT-----------A-----------G-----A--------..	8861
40_BF.BR.04.04BRRJ115	--T----------------A-C-----GG--G-----A-----...---G------------A-T--...---C-------A------AA-----A-----G--TGAA--C-A-------C-GC------A-----------AAA-------G-------------TCA	8966
42_BF.LU.03.luBF_05_03	--T--------------------MR----C--G-----A----...-C-G-----------------...-----------A------------A-----A---A-A---------------G-T-----A-----------AAG---------------------AC-	8847
43_02G.SA.03.J11223	--T---A---------A-G--T----CAG-TG----G--A---...-------A--------A---T...--A---------A-CT---AA-----A-----A--TGAA--C---------C-G-TA-----A--------TA--------CGGAGA--T--A-------	8913
U.CD.83.83CD003	--T-----------------T---A-GG-AG------------...---G------------A-T--...------------A-CT---A------A---------TGAA----A------------AT-----A---------TTC-------CGAAGA---C-A-------	8852
U.CD.90.90CD121E12	--T-----------------T----TGG--G------------...---G------------A-----...--A---------A-AT---A------A-----G--TGAA----A-------C--AT-----A------T--T-A------CGAAGC---AA---------	8547
U.GR.99.GR303	---T---T----------G--T---AGGG-AG..TT-----AGGCC--CG--------A---A-T-T..T------------A-T----AA-----A-----G---A-A-------------C--ATA---A---------------------C-AAGA---AGA---A---	8698
U.NL.95.U_NL_95_H10986_D1	--------------------T---AGAG-AG------------...---G------------A-T--...C----------CA------A------A---------TG-A--C--------------AGA---GA----------TC-------CGAGTA---CAA-------	9312
CPZ.CD.90.ANT	T-TT---T-----------A--AT--CCA...CCT--T--T..................----GG-A--TA...C-CC-C--------CTT-TACA-----AGAT-GA-----AC-T-A---GA-C---AGA---GAA-----TGCAA----GATGAGAAGA---A-T--AA---	8824
CPZ.CM.01.SIVcpzCAM13	--TT--A--------TT-GACAA-T--GC-AG-G----CA---...---G-G-----T-----A-T--...--AC-G------A-AT-T-A------A-----G--TGAAAGT-A-------C---ATA---C-C----------A-A--G--T--AAGA--TA-A--AA---	9024
CPZ.CM.05.SIVcpzEK505	--T-------------C-GTCA--T--GG--G-------A--T...---C-------C---ATGT-...C-------T--CA--T---AA-----A-----A--T--A--C-A-------C-G-T----C-C----------TC-------AGAGTA---AGA--AA---	8892
CPZ.CM.05.SIVcpzLB7	--T-------------C-AACC-----AG--G-G--G-----T...-----------------ATCTA...C-AC-G-----AA-AT---A------A-----A--TGAA--C-A-------C-GATA---AA-A--------AG--C--CC-AAGA---A-A--T----	8917
CPZ.CM.05.SIVcpzMB66	--T---A---------T-A--A-----AG-AG-G---AGA--A...---G--------C----A--TA...---------T--CA-CT---A---C--ACAA--G--TGAA-CA---------C----TC----CA-----------G--T---C-GA-G---AGA--T----	9030
CPZ.CM.05.SIVcpzMT145	--TT--A--G--C--TT-GCCA--T--GG--G-G--GA-A---...---G--------A---A-T--...C--C----T--CA--T---A------A-----G--TGA---A--------AC-G-TA----CT------T---AAG--G--TC-CAGA--TC-A--AA---	8918
CPZ.GA.88.GAB1	--T---A-----G--CC-GACAGAG--GC--G-----C-A---...---G-------T----A-T--...C----G--T--CA-AT-T-A-------------A--TGA---C-A---G--C--G-TC---C-C----------A-G--G--C--AAGA--TA-T---A---	9458
CPZ.TZ.00.TAN1	A-TT---T-G-----G--A--C---CCAG-AGAT--T---AAG...................-A--T-...---C----T--A-CCT-TAGC-----A-CTACC--T--A--TG----GACTC-GATC---C-C----------A--------AGAAGG---A-A---A---	9056
CPZ.US.85.CPZUS	--TT---T----C--TC-CACAGA---GG-AG----GC-----...--------------C--A--TA...C--C-G-----CA--T---A------A-----A--TGAAC-TG-C-----C-----C---ATC---CA--------CTGAA--G--TCGGAGA---AGA--TAA--	9424
CPZ.CM.x.SIVcpzDP943	--TT--A--G--T---C-GACA-----AG-AG-G---A-A---...---G-------T-C--A--TA...C--C-T-----CA-CT---A-----AA--A--A--TGAAC-TG-------C--G-TT-----A-A-------A-G-----T--GAGA---A-T--TA---	8978
CPZ.CM.x.SIVcpzMB897	--T---------G--T--A--T---ACAG-TG-G--G--A-A...---G--------C----ATGTG...C-A-----T---A--T---A---C--ACAA--A--TGAAC-C---------C-GAT-----CA-------------A-T-G--CC-GACA---AGA-------	8920
N.CM.02.DJO0131	T-T---A---------T-GTCAG-T--AGCAG---------T...---G--------C----ATGC-...C-C-----T--CC-AT---AA-----AG-A-----TGGAC-C-A-------C-GC-----CA-----------TC---G---AGAAGA---A-A--AA---	8860
N.CM.04.04CM_1015_04	--T---A---------T-GTCAG-T--AGCAG-------A--T...---G--------C----ATGC-...C-C-----T--CA-AT-T-AA-----AG-A-----TGATC-C-A-------C-G-T----C-----------TC---G---AGAAGA---A-A--AA---	8829
N.CM.04.04CM_1131_03	--T---A---------T-GTCAG-T--AGCAG-------A---...---G--------C----ATGC-...C-CC-G-----CA-AT-T-AA-----AG-A-----TGA-C-C-A-------C-G-T----C-----------TC---G---AGAAGA---T-A--AA---	8897
N.CM.95.YBF30	--T---A---------T-GTCAG-T--AG-AG---------T...---G--------C----ATGC-...C-C---------CA-AT-T-AA-----AGCA-----TGATC-T-A---------G-T----C-A---------TC-------AGAAGA--T--A--AA---	8975
N.CM.97.YBF106	--T---A---------T-GTCAGAG--AGCAG-------A--T...---G--------C--T-ATGC-...C-CC-G--T--CA-AT-T-AA-----AG-A-----TGATC-C-A-C-----C-G-T----C-----------TC---G---AGAAGA-----A--AAA--	8918
O.BE.87.ANT70	G-TT--A------------GTCAGA---AG--GC----AGACTAGGA----C-T-T----GGG-T-AT...C-CC-G--T--A-CAT-TGCC-----AT-T--A--TA-AC-T-A----A-AC-GAT-----A------T--ATCT----GCAACACC--T--T--TATGA	9474
O.CM.91.MVP5180	G-TT--A--G----------GTCAG----AG--GC---GAGACTGGGT----C-AAT--AG-TG-T--T...C-TC----T--A-CTT-TAAT-----AGCT--G--TG-AC-CG-G--GA-AC--A-A---CA------T--ATCAT---GC--AACA--TA-A----TGC	9497
O.CM.96.96CMABB637	A-T--G-------------GTCAGA--C-G--GC----AGACTGGGC----CGT-T--A-GGG-T-TG...C-CC----T--A-CTT-TAAT-----T-GT--C--TG-AC-TG-----A-ACCGA-A---CA------T-CTTCAT----TCGCACC--T--T--TTT-A	8979
O.CM.98.98CMA104	G-TT---------------GTCAGA--C-G--GC------ATTGGGC-----GCAT--A-GAG-T-AG...C-CC-G--T--A-CCT-TGCT-----AT-T-GA--TG-AC-CG-G---A-A----A----CA------T--ATCA----GCC-CACC--T--T--TAAGA	8978
O.CM.99.99CMU4122	A-TT---------------GTCAGA--C-G--GC------ACTAGGA-----GT-T-T--CAG-T-AA...C-CC-G--T--A-CTT--AAC-----TTCT--A--T--AC-CG-----A-AC-GA-----CA------T--ATCA----GCAACATA--T--T--TAAG-	8944
O.FR.92.VAU	A-TT---------------GTCAGA--C-G--GC-----CACTGGGC-----GT-T--A-GGG-T---...C-CC-G--T--A-CTT-TAAT-----TTAT--A--T-AAC-C-AG--A-AC-GA-A---CA------T--ATCA-----CAACACC--T--A--T-T-A	9046
O.SN.99.SEMP1299	A-TT---T-----------GTCAGA--C-G--GC---G--ACTAGGA--C--GT-T----GGG-T---...C-CC-G--T--A-CTT--AAC-----CT-T--A---AACC-CG-GC--A-AC-GA-A---CA------T--ATCA----GCAGCACC--T--T--TATG-	9536
O.US.99.99USTWLA	A-TT---T----------GACAGA-C-AG--GC----C-ACTGGGC-----G-AT--AGGGG-T-TG...C-CC----T--A-CAT-TAAT-----T-GT--G--T-ATC-CG-----A--C-A-A---CA------T---TCA----GCA-CACC--T--A---ATG-	8945

```
                                          TCF-1 alpha binding
                                                        Nef end                                                                                     NF-κ-B-II
B.FR.83.HXB2               AGCTGCATCCGGAGTACTTC...AAGAACTGCTGA....................................CATCGAGCTTGCTACA..................................................AGGGACTTTCCGCT...................... 9447
Nef                        E__L__H__P__E__Y__F__.__K__N__C__*_
A1.GE.99.99GEMZ011         ------------------A-...--AG--------...................................CACAAACTGC----GACAG................................................------------K...................... 8647
A1.KE.00.KER2008           --A-A-------------A-...--AG--------....................................--CA--AG-----GACTG.................................................----------A--...................... 8648
A1.KE.00.KNH1144           --AAA----------T--A-...--AG--------....................................--CA--AG-----GACAG.................................................---------T---...................... 8697
A1.KE.00.KSM4024           ---------------T--A-...------------....................................--CA--AGC----GACAG.................................................-------------...................... 8588
A1.KE.00.MSA4069           --A-------------T-A-...--AG--------....................................--CA--AG-----GACAG.................................................-------------...................... 8670
A1.KE.00.NKU3005           ---------------T--A-...--AG--------....................................--CA-GAG-----GACAC.................................................-------------...................... 8674
A1.RU.00.RU00051           ------C-----G--T--A-...--AG--------....................................-MCA-GAG-----GMCAG.................................................-------------...................... 8745
A1.RW.93.93RW_024          ---------------T--A-...--A-----A---....................................--CA--AGA----GACAG.................................................---------T---...................... 8673
A1.SE.95.SE8891            --T-A----------T--A-...--AG--------....................................--CA--AG-----GACAG.................................................-------------...................... 8626
A1.SE.95.UGSE8131          ----------------T-A-...--------A---....................................T-CA-CAG-----GACTG.................................................-------------...................... 8825
A1.TZ.01.A173              -------------C-T--A-...--AG--------....................................--CA--AG-----GACAG.................................................-------------...................... 8644
A1.UA.01.01UADN139         --AA-----------T--A-...--AG--------........................CACAGGCTGCTGACAT--AG-----GACAG.................................................------------G...................... 8694
A1.UG.92.92UG037           ---------------TT-A-...--AG--------....................................--CA-GAG-----GACTG.................................................-------------...................... 9505
A1.UG.99.99UGA07072        --T------------T--A-...--AG--------....................................--CA--AG-----GACAG.................................................-------------...................... 8668
A2.CD.97.97CDKTB48         ---A-----------T--A-...--AG--------.....................CACAGAAGATGCTGA--CA--AG-----GACAG.................................................----------A-....................... 8822
A2.CY.94.94CY017_41        ------------------A-...--AG--------....................................--CA--AG-----GACGG.................................................-------------...................... 8873
A.SN.01.DDI579             --A------------T--A-...--AG--------.....................CACAGAAGTTGCTGA--CA--AG-----GACAG.................................................-------------...................... 8645
A.SN.01.DDJ369             --AA-T-------C-GG-A-...--AG--C-----.................CACAGAAGTTGCTGACTGA--CA--AG-----GACAG.................................................------------G...................... 8709
A.SN.96.DDJ360             ---------------T--A-...--AG--------....................................--CA--AG-----GACAG.................................................-------------...................... 8621
A.CD.02.02CD_KTB035        ---A------A----T--A-...--AG--------....................................--CA--AG-----GACAG..................................................AA-----------...................... 9468
A.CD.97.97CD_KCC2          ---------------T--A-...--AG--------....................................--CA--AG-----GACAG.................................................-------------...................... 9467
A.CD.97.97CD_KTB13         --A------------T--A-...--A-----A---....................................--CA--AG-----GACAA...........................GGGACTTTCCGCTGCTGAC.A-------------G...................... 8769
B.AR.04.04AR151516         ------------------A-...--AG--------....................................--------.-T-----AG.................................................-------------...................... 8658
B.AU.87.MBC925             ------------------A-...---G--------....................................--G-------T-----...................................................-------------...................... 9466
B.BO.99.BOL0122            -------C--A----T--A-...--AG-T--A-TG.........CTGACAGTGTTCTACAGAGAGTGCTGA--GA---T--T-----AG.................................................----------A--...................... 8669
B.BR.03.BREPM2012          --AA-----------T--A-...--AG-T--A---....................................--C----T--T-----AG.................................................-------------...................... 8922
B.CA.97.CANB3FULL          --AA--------------A-...--AG--------.......................................................................................................................................... 8687
B.CN.05.05CNHB_hp3         --AAA-----T----T--A-...--AG--------....................................--C-------T-----AG.................................................-------------...................... 9457
B.CO.01.PCM001             ---------------T----...--AG--------....................................--C-A-----TG----AG.................................................-------------...................... 8630
B.GB.83.CAM1               --AA-----------T--A-...--AG--------....................................----------T-----AG.................................................-------------...................... 9459
B.GB.86.GB8                -A-------------TT-A-...--AG--------....................................----CG-AG-.-----AG.................................................-------------...................... 9436
B.GE.03.03GEMZ010          ------------------A-...--------A---....................................--C----T--T-----AG.................................................-------------...................... 8624
B.IT.05.SG1                ....----------C---A-...------------....................................---------CT-G---AG.................................................-------------...................... 8614
B.JP.05.DR6538             ---AA----------TT-A-...--AG--------....................................-------T--C-----AG.................................................-------------...................... 9483
B.KR.05.05CSR3             -C------------A---A-...------------...........CATGGAGCCCTACAGGAATTGCTGA----------T-----AG.................................................-------------...................... 9526
B.NL.00.671_00T36          --A-A-------------A-...------------..........CACCGGAGTACTACAAGAAATGCTGA--C-----.-T-----...................................................--------------..................... 9027
B.RU.04.04RU128005         --AAA-----A---------...-----T--A---.........CATCGAGCTTCCTACAAGAACCGCTGA---A------T-----AG.................................................-------------...................... 8946
B.TH.00.00TH_C3198         ............................................................................................................................................................................. 8394
B.UA.01.01UAKV167          --A-A-------------A-...--AG--------....................................--C-------T-----AG.................................................-------------...................... 8705
B.US.04.ES10_53            ------------------A-...------------....................................--CT-----AT-----...................................................-----------A-C...................... 9440
B.US.99.PRB959_03          ------------------A-...--AG--------..............CATTAACTACCAAGACTGCTGA----------T--G--AG.................................................-------------...................... 8673
C.AR.01.ARG4006            ----A-----------T-A-...--AG--------................................CACAG-AG-GA---T-CG-...................................................T----------A--...................... 8668
C.BR.04.04BR013            ----A-----------T-A-...--AG--------................................CACAG-AG-GA---T-CG-...................................................T----------A--...................... 8962
C.BW.00.00BW07621          ----A--------A--T-A-...--AG--------................................CACAG-AG-GA---T-CG-...................................................T----------A--...................... 8806
C.CN.98.YNRL9840           ............................................................................................................................................................................. 8152
C.ET.02.02ET_288           ----A-----------T-A-...--AG----A---................................CACAG-AG-GA---TTC....................TGCTGACACAGAAGGGACTTTCTGT........T----------A--...................... 8702
C.GE.03.03GEMZ033          ----A----------TT-A-...--AG--------................................CACAG-AG-GA---T-CG-...................................................T----------A--...................... 8681
C.IL.99.99ET7              ----A--------C-TT-A-...--AG--------................................CACAG-AG-GA---T-CG-...................................................T----------A--...................... 8640
C.IN.99.01IN565_10         ----A----------TT-A-...--AG--------...........GGTCTTACAAAAGACTGCTGACACAG-AG-GA---TTCG-...................................................T----------A--...................... 8872
C.KE.00.KER2010            ---AA-----------T-A-...--AG-T------................................CACAG-AG-GA---TT-G-...................................................C----------A--...................... 8663
C.MM.99.mIDU101_3          ----A----------TT-A-...--AG--------................................CACAG-.G-GA---T-CG-..................................................GG----------A--...................... 8833
C.MW.93.93MW_965           ---AA----------T--A-...--AG--------................................CACAG-AG-GA-----CG-...................................................T----------A--...................... 8649
C.SN.90.90SE_364           ----A----------TT-A-...--AG--------................................CACAG-AG-GA---T-CG-...................................................T----------A--...................... 8598
C.SO.89.89SM_145           ----A-----------T-A-...--A---------................................CACAG-AG-GA---T-CG-...................................................T----------A--...................... 8700
C.TZ.02.CO178              ------------------A-...--AG--..................................................................................................................................... 8637
C.UY.01.TRA3011            ----------------T-A-...--AG--------.........TACAGAAGTACAAAGACTGCTGACACAG-AG-GA---T-C--...................................................T----------A-C...................... 8674
C.YE.02.02YE511            ----A----------TT-A-...--AG--------................................CACAG-AG-GA---T-CG-...................................................T----------T--...................... 8702
C.ZA.04.04ZASK164B1        ----A----------GG-A-...--AG----A.....(........................)CTGCTGACACAG-AG-GA---T-.....................TGCTGACACAGAAGGGACTTTCCGC........T----------A--...................... 8947
C.ZA.05.05ZAPSK240B1       --A-A-----------T-A-...--AG--------................................CACAG-AG-GA---T-CG-C..................................................T----------A--...................... 8854
C.ZM.02.02ZM108            ----A-----------T-A-...--AG--------................................CACAG-AG-GA---T-CG-...................................................T----------A--...................... 9514
D.CD.83.ELI                --A------------T--A-...--A-----A.................................TGACACCG-G....---T-----A.................................................-------------...................... 8986
D.CM.01.01CM_0009BBY       ............................................................................................................................................................................. 8421
D.KE.01.NKU3006            C-A---------------A-...--AG--------.....................................CA-CGAG-...........................................TTCTACA.......--------------...................... 8692
D.KR.04.04KBH8             T--------------TT-A-...--AG--------.....................................CA-TGAG..........................................TTTTCTACA.......G-------------...................... 9348
D.TD.99.MN011              --AA-----------T---T...--AG--------.....................................CA-CGAG-AAA..................GACTGCTGACACCAAG....CCTTCTACA.......-----------A--...................... 8657
D.TZ.01.A280               GCT---------------A-...--AG--------.....................................CATCGA.......................GACTGCTGACACCGAG...TTTTGCTACA.......--------------...................... 8654
D.UG.99.99UGK09259         ----A-------------A-AAC--AG--C-----.....CACCACCGAGTTATACTGCTGA..........CA-TGAG--TTCTACAAG...............................................-------------...................... 8671
D.YE.01.01YE386            T--------------T--A-...--AG--------.....................................CA-AGAAG-TGCTG-CAG...............................................-------------...................... 8606
D.YE.02.02YE516            --AC-----------T---T...--AG--------.......CACTGAGTGAAGACTGCTGA..........CA-C-AG--TTCTACAAG...............................................-------------...................... 8641
D.ZA.90.R1                 ----A-------------A-...--AG--------.....................................CA-CGAG--TTGTACAAG...............................................-------------...................... 8804
F1.AR.02.ARE933            ............................................................................................................................................................................. 8338
F1.BE.93.VI850             --AGA----------T--A-...C-AG----A.....................GACTGCTGA..........CA-AGAGA-TGCTG-CACAGAAGAATCTAAAAG................................----------A--...................... 8745
F1.BR.01.01BR125           ------------------A-...--------A---.....................................CA-CGAG...CTTTAGACTGCTG..............ACAC...T.GAGCTTTCTACAAG......-------------...................... 8904
F1.ES.x.P1146              --ACA----------T--A-...C-AG---AA.....................GACTGCTGA..CACAGAGATC-AGAGA-TGCTG-CATAGAAGATTCTAAAG.................................-------------...................... 8815
F1.FI.93.FIN9363           --AGA----------T--A-...CGAG----A.....................GATTGCTGA..........CA-AGAGA-TGCTG-CACAGAAGAATCTAA...................................----------A--...................... 8767
F2.CM.02.02CM_0016BBY      ............................................................................................................................................................................. 8349
F2.CM.95.MP257             --AGA-------------A-...--AG--............................................................................................................................................... 8589
F2.CM.97.CM53657           --A---------------A-...--AG----A...........................................--AAC-.GCTG-TACAGAAGATTCTAAAAG................................-------------...................... 8626
G.BE.96.DRCBL              ------------------A-...--AG--------.....................................CA-AGAAG-TGCTG-CAA...GGGGACTTTAAGTTGCTGACA.......................-A-----------C...................... 9433
G.CM.01.01CM_4049HAN       ............................................................................................................................................................................. 8367
G.CU.99.Cu74               -------------T-T--AT...--AG--------.....................................CA-AGAAG-TGCTG-CAA...............................................------------C...................... 9091
G.ES.99.X138               -----------A-T-T--A-...--AG--------...............CATAGCTGCTGA..........CA-AGAAG-TGCTG-CAAAG.............................................------------C...................... 8920
G.GH.03.03GH175G           ------------------A-...------------...............CACAGAAGGTGTTGACTGCTGACA-AGAAG-TGCTG-CAAG..............................................------------C...................... 9540
G.KE.93.HH8793_12_1        ---------------T--AT...------------.....................................CA-AGAAG-TGCTG-CAAGG.............................................------------C...................... 8849
G.NG.01.01NGPL0669         ------------------A-...--A-----A---.....................................CA-AGAAG-TGCTG-CACAGGAGTTGCTGACAA................................G-----------TC...................... 8693
G.PT.x.PT2695              -----------A-T-T--A-...--AG--------...............CACAGCTGCTGA..........CA-AGAAG-TGCTG-CAA...............................................G------------C...................... 9496
G.SE.93.SE6165             ------------------A-...--AG--------.....................................CA-AGAAG-TGCTG-CAA...............................................G----------A-C...................... 8891
H.BE.93.VI991              ---------------T--A-...--AG--------.....................................CA-AGAAG-TGCTG-CACAGAAGATTCTATACA................................-------------...................... 8865
H.BE.93.VI997              --AAA----------T--A-...--AG--------.....................................TA-AGAAGATTCTG-.................................................C-------------....................... 8764
H.CF.90.056                TAAA----------...-A-...--AG--------.....................................CA-AGAAGAT.CTG-A.................................................-------------...................... 8765
J.CD.97.J_97DC_KTB147      ---------------TT-A-...--A---------...............CAAGACTGCTGA..........CAAAGAAG-TTCTA-.................................................AG------------....................... 8623
J.SE.93.SE7887             ---------------TT-A-...--AG--------.....................................CAAAGAAG-TTCTAG.................................................CG------------....................... 8744
K.CD.97.EQTB11C            --A---------------A-...--AG--------.......................................................................................................................................... 8600
K.CM.96.MP535              -------------C----A-...--AG--------.......................................................................................................................................... 8604
```

```
                                          TCF-1 alpha binding
                                                        Nef end                                                                                                      NF-κ-B-II
B.FR.83.HXB2               AGCTGCATCCGGAGTACTTC...AAGAACTGCTGA....................................CATCGAGCTTGCTACA.................................................AGGGACTTTCCGCT..................... 9447
Nef                        E__L__H__P__E__Y__F__.__K__N__C__*_
01_AE.CF.90.90CF11697      ------G---A-------A-...--AG-T-------------------------------------------CAGAGAAG-TTCTA-CC................................................A----.-------..................... 9413
01_AE.CN.05.FJ051          -AAA-T-C--AC---T--AT...--AG--------.....................................CAAAGAAG-TTCTA-CT................................................A-A--.-------..................... 9462
01_AE.CN.06.FJ054          --A-------A-------AT...--AG--------....................................ACAA-GAAG-TTCTA-CT................................................-----.---G-TG..................... 9459
01_AE.HK.x.HK001           -A--------A----T--AT...--AG--------.....................................CAAAGAAG-TTCTA-CT................................................A----.-------..................... 8828
01_AE.TH.01.01TH_R2184     ........................................................................................................................................................................... 8408
01_AE.TH.02.OUR769I        ........................................................................................................................................................................... 8390
01_AE.TH.90.CM240          -A-A------A-------AT...--AG--------.....................................CAA-GAAG-TTCT.-CT................................................A-A--.-------..................... 9013
01_AE.US.00.00US_MSC1164   ........................................................................................................................................................................... 8395
02_AG.CM.02.02CM_4082STN   ........................................................................................................................................................................... 8413
02_AG.EC.x.ECU41           ........................................................................................................................................................................... 8291
02_AG.GH.03.GHNJ196        ---------------TT-A-...--AG----T---.....................................CA-AGAAG-TGCTG-CAG............................GGACTTTCTG......CCT-------------..................... 9574
02_AG.NG.01.PL0710         ------------------A-...--AG-T------.....................................CA-AGAAG-TGCGG-CAAG...............................................-A----------..................... 8632
02_AG.NG.x.IBNG            ------------------A-...--AG--------.....................................CA-AGAAG-TGCTG-CAG................................................---------A--..................... 8974
02_AG.SE.94.SE7812         --A------------TT-A-...--AG--------................CACTTCTACAAAGACTGCTGACA-AGAAG-TGCTG-CAG............................GGACTTTCTG......AC-----------A--..................... 8872
02_AG.SN.98.MP1211         ------------------A-...--AG--------.....................................CA-AGAAG-TGCTG-CAA...............................................-------------..................... 8683
03_AB.RU.97.KAL153_2       ---------------T--A-...--AG--------.........................CACAGACTGCTG-CATGAAG-TGCTG-CAG................................................-----------G..................... 8624
04_cpx.CY.94.CY032         ---------------TT-A-...--AG--------.....................................CA-AGAAG-TGCTG-CAAAG..............................................-----------C..................... 8861
05_DF.BE.x.VI1310          --AGA-----A----T--A-...C-AG----A.--CTGCTGACACAGAAACTGCTGACACAGGAAGTGCTGACA-AGAAGCTT-CT-CAA...............................................-------------..................... 8893
06_cpx.AU.96.BFP90         --AAA----------TT-A-...--AG--------.........................CAAGGCTGCTGACAAAGAAG-TTCTA-TGG................................................------------..................... 9504
08_BC.CN.97.97CNGX_6F      -A--A----------TT-A-...--AG--------.....................................CA-AGAAGG-ACTTTCCGCG..............................................---------A-C..................... 8649
09_cpx.GH.96.96GH2911      --AAA----------T--A-...--AG--------.....................................CA-AGAAG.TTCTAGCAGG...............................................----------AA..................... 8619
10_CD.TZ.96.96TZ_BF061     ----A-----------T-A-...--AG-T------.....................................CA-AGAAGG-ACTTTCCGCTGG.............................................--------AA-..................... 8805
11_cpx.GR.x.GR17           -------------C-T--A-...--AG---AA...............................GACTGCTGACA-AGAAG-TGCTG-CATAAAAGTTGCTGACAG.................................------------..................... 8747
12_BF.AR.99.ARMA159        --AAA-----A----GG-AT...--AG----A...............................GACTGCTGACA-AGAGACTGCTG-CACAGAAGAATCTAAAG..................................------------..................... 9432
13_cpx.CM.96.1849          ---------------T--A-...--AG--------.....................................CA.AGAGA-TT-.........................................................ACA----T-..................... 8862
14_BG.DE.01.9196_01        -------------T-T--A-...--AG--------.........................CACAGCTGCTGACA-AGAAG-TGCTG-CAAGG..............................................-----------C..................... 8990
15_01B.TH.99.99TH_MU2079   -A--------T-------AT...--AG--------.....................................CAAAGAAG-TTCTA-CTG................................................---.--------..................... 8821
16_A2D.KR.97.97KR004       ---------------T--A-...--AG--------......................CACAGAAGCTGCTGACA-AGAAG-TGCTG-CAG................................................---------A-C..................... 8879
17_BF.PE.02.PE02_PCR0155   --AAA----------T--A-...C-AG----A...............................GACTGCTGACA-AGAGA-TGCTG-CACAGAAGAATCTAA...................................----------A--..................... 8662
18_cpx.CU.99.CU76          -A-AA----------T--A-...--AG--------.....................................CA-AGAAG-TGCTG-CAG................................................-----------C..................... 8776
19_cpx.CU.99.CU7           TAAA---C--A----T--A-...--AG--------.....................................CA-C-AG--TTCTACAAG................................................-GAC--T-CGCG..................... 8512
20_BG.CU.99.Cu103          ---------------T--A-...--AG--------.....................................CA-AGAAG-TGCTG-CAAGG.............................................-----------CT..................... 8896
21_A2D.KE.91.KNH1254       --A------------T--AT...--AG-T------..............CACGGAAGCC.AGAAGCTGCTGACA-TGAG.-TTCTACAAG................................................---------A--..................... 8693
22_01A1.CM.01.01CM_0001BBY ........................................................................................................................................................................... 8372
23_BG.CU.03.CB118          ---G---------C-T--A-...--AG--------.....................................CA-AGAAG-TGCTG-CAAGG.............................................-----------CT..................... 8905
24_BG.CU.03.CB378          ---------------T--A-...--AG--------.....................................CA-AGAAG-TGCTG-CAAAG.............................................-----------CT..................... 8965
24_BG.CU.03.CB471          ---------------T--A-...--AG--------.....................................CA-AGAAG-TGCTG-CAAAG.............................................-----------CT..................... 8926
25_cpx.CM.01.101BA         --AA-----------GG-A-...--AG--------.....................................CA-AGAAG-TGCTG-CAAGG.............................................---------A-CT..................... 8727
27_cpx.CD.97.97CDKTB49     --AC-----------T--A-...--AG--------.....................................CA-AGAAG-TGCTG-CAA...............................................------------C..................... 8655
28_BF.BR.99.BREPM12609     ----------------T-A-...--AG---AA...........................AGACTGCTGACATCGAGA-TTCTGCAAG................................................-GAC--T-CGCT..................... 8867
29_BF.BR.02.BREPM119       ------------------A-...------------.....................................CATCGAGACTTCTACAAG................................................------------..................... 8599
31_BC.BR.02.110PA          ----A-----------T-A-...--A----A----........................................................................................................................................ 8908
32_06A1.EE.01.EE0369       --AA-----------TT-A-...--AG--------.....................................CAAAGAAG-TTCTA-CG.................................................------------..................... 9072
33_01B.MY.05.05MYKL007_1   -A--------A--A----AT...--AG--------.....................................CAAAGAAG-TTCTA-CTA................................................-A--.-------..................... 8837
34_01B.TH.99.OUR2478P      -A-------------T--AT...--AG--------.....................................CAAAGAAG-TTCTA-CTA................................................----.-------..................... 8600
35_AD.AF.05.05AF095        ---------------T--A-...--AG--------.....................................CA-AGAAG-TGCTA-CAG................................................------------..................... 8628
36_cpx.CM.00.00CMNYU830    ........................................................................................................................................................................... 8410
37_cpx.CM.00.00CMNYU926    ........................................................................................................................................................................... 8396
38_BF1.UY.05.UY05_4752     --AAA-----T----T--A-...C-------A..............GACTGCTGACACAGAGATTGCTGACA-AGAAGAATCTG-AG.................................................------------..................... 8887
39_BF.BR.03.03BRRJ103      ........................................................................................................................................................................... 8861
40_BF.BR.04.04BRRJ115      TCA-----------------...--AG-TC-----.....................................CA-CGAGC-TTCTACAAG................................................------------..................... 9028
42_BF.LU.03.luBF_05_03     --A---------------A-...--AG--------.....................................CA-C-AG--TTCTACAAG................................................------------..................... 8909
43_02G.SA.03.J11223        -A-----C--A-------AT...---G--------............................CGCTGCTGACA-AGAAG-TGCTG-CACAGATGTTGTTGACAAAG...............................-----------C..................... 9001
U.CD.83.83CD003            ---------------TT-A-...--AG--------..........CATAGACTGCTG...........ACA--GAAGTTC-AAG-G...................................................---------AAAG..................... 8927
U.CD.90.90CD121E12         ----------------T---...--AG--------..........CACAGACTGCTG...........ACA--GAAGTTC-AAGGG...................................................---------A-AG..................... 8621
U.GR.99.GR303              GAGAA-------------A-...--AG--------..................................CA-TGAACCTC-AAA-G.....................................................---------C-..................... 8758
U.NL.95.U_NL_95_H10986_D1  T---A-------------A-...--AG--------.................................TAAAGAA-TTTC-AACTG....................................................------------..................... 9374
CPZ.CD.90.ANT              -AAGA--C--T--A--T---...-GAG---AA....GACTGCTAACAGGCGGATATAAAACCGCTGACAGAGC--AC-TG-GGTCT................................................----------A--AT................... 8918
CPZ.CM.01.SIVcpzCAM13      --AAA-----------T-A-...--AG----A.....................CTGGCGCATGCGCCTA-GAAGTGC--ACTGTG................................................C----------AAG..................... 9099
CPZ.CM.05.SIVcpzEK505      ---------------T--A-...C-----------.....................................CAA-G-AC-TTCTACTTGG...............................................---------AGC..................... 8955
CPZ.CM.05.SIVcpzLB7        --------T---------A-...------------.....................................CAA---AC-TT-CTGACCAG..............................................---------A-C..................... 8981
CPZ.CM.05.SIVcpzMB66       ----A-------------A-...--AG-T------.....................................CAAAG-AC-TG-G-T..................................................G---------A-A..................... 9091
CPZ.CM.05.SIVcpzMT145      --AAA--------A--T-A-...-GAG----A....CCACAGCGGCGCATGCGCACTAAGAACTGCTGAT..C-A-A-TTGCTGACAAG.............................................----------A....................... 9008
CPZ.GA.88.GAB1             -A-AA-------------A-...--AG----A............................................C-TTAGACTGGCGCATGCGCACAAGAACTGCTGACTCTGC..................----------A....................... 9538
CPZ.TZ.00.TAN1             -AAGAT--------------...-AT-AG-A..........CTTCCGGGTGCCATGACTCAGAACTGCTGACAGAG-AC-TT-GG-CT.................................................................................. 9132
CPZ.US.85.CPZUS            ----------A---------...CG------A..............TGGCTGTAACCGCGCAGGCGCAAT-AAACTGC--ACTG...................................................A-------TA........................ 9502
CPZ.CM.x.SIVcpzDP943       -A-AA-------------A-...C--G----A..............AGCCCAGAGCCCAGCGCCTGCGCAG-A--AAC----GACAA................................................---------TA....................... 9059
CPZ.CM.x.SIVcpzMB897       ------------------A-...--AG--------.....................................CAAAG-AC-TTGTA-.................................................-----------A-A..................... 8981
N.CM.02.DJO0131            -------------C-TT-A-...------------.....................................CAA-G-AC-TTACT-ATG...............................................------------C..................... 8923
N.CM.04.04CM_1015_04       -------------C-T--A-...------------.....................................CAA-G-AC-TTAC..............TGCTGACAA........GGGACTTTACACAT......G-----------TC..................... 8911
N.CM.04.04CM_1131_03       -------------C-TT-A-...------------.....................................CAA-G-AC-TTACACATG...............................................------------C..................... 8960
N.CM.95.YBF30              ---------------TT-A-...------------.....................................CAA-G-AC-TTAC..............TGCTGACAA........GGGACTTTATACTT......G------------C..................... 9057
N.CM.97.YBF106             -------------C-T--A-...------------.....................................CAA-G-AC-TTACACGG................................................------------C..................... 8979
O.BE.87.ANT70              TAACT--C--A---CT----...C----GGA--A-.............AAACTGCTGAC..........CTG-AGATTGC--AC--TGT................................................--A------AGCAA.................... 9552
O.CM.91.MVP5180            -AAA---C--A---CT----...CCC--G-AA............................................C--AC--TGC................................................----------........................ 9546
O.CM.96.96CMABB637         TAACT--C--A---CT----...--C--GGA--A-..............AACTGCTGACC..........TG-AGATTGC--AC--TGC................................................----------AGCGG................... 9056
O.CM.98.98CMA104           TAACA--C--A---CT----...CCC--GGA----..............ACCTGCTGATC...........T-AA..............................................................----------AGCAAAGACTGCTGACACTGCG.. 9057
O.CM.99.99CMU4122          TAACC--C--A---CT----...CTC--GGA--A-..............AACTGCTGACC..........TG-AGA.TGC--AC--TGC................................................----------AACAG................... 9020
O.FR.92.VAU                TAACT--C--A---CT----...TCC--AGA--A-..............AACTGTCAACA..........GG-AGACTGC--ACG-TAC................................................----------AGCAA................... 9123
O.SN.99.SEMP1299           TAACCA-C--A---CT---T...--C--GGA--A-........................AACTGCTGACCTG-AGATTGC--AC--TGC................................................----------AGCAG................... 9613
O.US.99.99USTWLA           TAACT--C--T---CT----...-TC--GAA--A-..............AACTGCTAACCAGA.......AG-AGATTGC--AC-ACGC................................................----------AGCT.................... 9024
```


HIV-1/SIVcpz
Complete Genomes

```
                                                    NF-κ-B-I          Sp1-III            Sp1-II       Sp1-I                      TATA box                 TAR element start
                                                                                                                                                  3' LTR U3 end   3' LTR repeat start
B.FR.83.HXB2             ..........................GGGGACTTTCCAGGGAGG..............CGTGGC...CTGGGCGGGACT...GGGGAG...TGGCGA...GCCCTCAGA...TCCTGCATATAA.GCAGCTGCTTTTTGCCTG...TACTGG...GTCTCTCTGGTT.. 9552
A1.GE.99.99GEMZ011       ..........................-------------GA-...........................-GT-...------.----T-...A--------...-G---------.------------C-----...---G--...---------T----.. 8742
A1.KE.00.KER2008         ..........................-------------GA-G............T----T...T------AGT-...------.----T-...A--------...AG---------.-----C------C--T--...------...---------T----.. 8754
A1.KE.00.KNH1144         ..........................-------------GA-G............T----TTG...------AGT-...------.----T-...---------...-G---------.------------C-----...------...---------T----.. 8802
A1.KE.00.KSM4024         ..........................-------------GA-G............T----T...T------AGT-...------.----T-...A--------...-G---------.------------------...------...---------T----.. 8694
A1.KE.00.MSA4069         ..........................-------------GA-G............T----T...T------AGT-...------.----C-...A--------...-G---------.------------C-----...------...---------T-G-.. 8776
A1.KE.00.NKU3005         ..........................-------------GA-G............T-C--T...T------AGT-...------.----T-...A--------...-G---------.-----------C---T--...------...---------T----.. 8780
A1.RU.00.RU00051         ...........................----------C---GA-G............T----T...T------AGA-...------.----T-...A--------...-G--.......................................................... 8806
A1.RW.93.93RW_024        ..........................-------------GA-G............T----T...T------AGT-...------.----T-...---------...-G---------.------------C-----...------...---------T-G-.. 8779
A1.SE.95.SE8891          ..........................-------------GA-G............T----T...T------AGT-...------.----T-...A--------...-G---------.-----------CC--T--...------...---------T----.. 8732
A1.SE.95.UGSE8131        ..........................-------------GA-G............T----T...T------AGT-...------.----T-...A--------...-G---------.-----C---C-C-----...------...---------T----.. 8931
A1.TZ.01.A173            ..........................-------------GA-G............T----T...T------AGT-...------.----T-...A--------...-G---------.------------C-----...------...---------T-G-.. 8750
A1.UA.01.01UADN139       ...........................------------GA-G............T----T...T------AGT-...------.----T-...A--------...-G---------.----------C-C--A--...------...----------------.. 8799
A1.UG.92.92UG037         ..........................-------------GA-G............T----T...T------AGT-...------.----T-...A--------...-G---------.----------C-C--T--...------...---------T-G-.. 9611
A1.UG.99.99UGA07072      ..........................-------------GA-G............T----T...T------AGT-...------.----C-...A--------...-G---------.------------C-----...------...---------T-G-.. 8774
A2.CD.97.97CDKTB48       ..........................----------G------............T----T...GA-----AGT-...------.----T-...A--------...-G---------.-----------GC-----...------...--------CTTG-.. 8926
A2.CY.94.94CY017_41      ..........................-----------------............T----T...G------AGT-...------.----T-...A--------...-G---------.----------C-C--A--...------...---------T----.. 8978
A.SN.01.DDI579           ..........................-----------------............-----T...T------AGT-...------.----C-...---------...-G---------.-----------C------...------...---------T----.. 8750
A.SN.01.DDJ369           ...........................----------G---..............G----T...T------AGT-...------.----T-...---------...-G---------.-----------C------...------...---------T----.. 8810
A.SN.96.DDJ360           ..........................-----------------............T----T...T------AGT-...------.----C-...---------...-G---------.-----------C------...------...---------T----.. 8726
A.CD.02.02CD_KTB035      ..........................AAA---------A-GAAG............TAA--T...T-A----AGT-...AA----.----T-...A--------...-G---------.----GCTGC---CTG--T...GTACTAA..--------TAG-.. 9575
A.CD.97.97CD_KCC2        ...........................----------G---GA-G............T----TT..TG--C---AGT-...------.----T-...A--------...-G---------.-----------C---T--...------...---------T-A-.. 9573
A.CD.97.97CD_KTB13       ...........................------------GA-G............T----T...T------AGT-...------.----T-...A--------...-G---------.-----C---C-C---T--...------...---------T-G-.. 8874
B.AR.04.04AR151516       ..........................-------------GA-G............T-----...------K---C...------.------...---------...-G---------.------------------...------...-------T-A-C-.. 8764
B.AU.87.MBC925           ..........................-------------GA-G............------...------------...------.------...---------...-G---------.------------------...------...------------.. 9572
B.BO.99.BOL0122          ..........................-------------GA-G............-----A...-----------A...------.------...---------...-G---------.------------------...------...------------.. 8775
B.BR.03.BREPM2012        ..........................-----------------............T-----...-----------C...------.------...---------...-G-----------------------........................... 9003
B.CA.97.CANB3FULL        ......................................................................................................................................................................... 8687
B.CN.05.05CNHB_hp3       ..........................-----------------............----A-...A----------C...------.------...---------...-G---------.-----C------------...------...---------A--C.. 9561
B.CO.01.PCM001           ..........................-------------GA-G............-----A...------------...A-----.------...---------...-G---------.-----------C------...------...------------.. 8736
B.GB.83.CAM1             ..........................-------------GA-G............T-----...------------...------.------...---------...-G---------.------------------...------...------------.. 9565
B.GB.86.GB8              ..........................-------------GA-G............------...-----------C...------.------...---------...-G---------.------------------...------...---------A----.. 9542
B.GE.03.03GEMZ010        ..........................-----------------............T-----...-----------C...------.-----AT...---------...-G---------.------TGC-CAC--T--...------...------------.. 8729
B.IT.05.SG1              .........................G----------G----A-G............--C---...------------...------.------...---------...-G---------.------------------...------...------------.. 8721
B.JP.05.DR6538           ..........................---------------A-G............------...------------...------.------...---------...-G---------.------------------...------...---------A----.. 9589
B.KR.05.05CSR3           ..........................----------GAAATCC.............G-AC-............................................................................................. 9549
B.NL.00.671_00T36        ..........................-----------------............T-----...---------A--A...------.---T--...A--------...-G---------.------------------...------...--------CA---.. 9132
B.RU.04.04RU128005       ..........................-------------GA-G............-----A...------------...------.------...---------...-G---------.------------------...------...------------.. 9051
B.TH.00.00TH_C3198       ......................................................................................................................................................................... 8394
B.UA.01.01UAKV167        ..........................-------------GA-G............T-----...-----------A...------.------...---------...CG---------.------------------...------...------------.. 8811
B.US.04.ES10_53          ..........................-------------GA-G............T-----...-----------A...------.------...---------...-G---------.------------------...------...------------.. 9546
B.US.99.PRB959_03        ..........................-------------GA-G............-----A...------------...------.------...---------...-G---------.------------------...------...------------.. 8779
C.AR.01.ARG4006          ..........................---G-G--------G-..............A----T...---------T-C...------.---TC-...A--------...-G---------.-----------GC-----...------...-------T-A-G-.. 8771
C.BR.04.04BR013          ..........................---G-G--------G...............A----T...------------...------.----C-...A--------...-G---------.------------------...---A.................. 9050
C.BW.00.00BW07621        ..........................---G-G-------AG-..............A----T...------------...A-----.----C-...---------...-G---------.------------C-----...------...---------A-G-.. 8909
C.CN.98.YNRL9840         ......................................................................................................................................................................... 8152
C.ET.02.02ET_288         ..........................---G-G--------G...............CA-AAC...------------...------.----C-...A--------...-G---------.------------C-----...------...---------A-G-.. 8803
C.GE.03.03GEMZ033        ..........................A---G-G--------G...............GAG--T...-----------C...------.----T-...---------...-G---------.------------C-----...------...--------T-A-G-.. 8784
C.IL.99.99ET7            ..........................---G-G-------AG-..............A----T...-----------G...A-----.----T-...---------...-G---------.------------C--T--...------...---------C-G-.. 8742
C.IN.99.01IN565_10       ..........................--G-G---G--AG-..............T----T...------------...------.---TC-...A--------...-G---------.-----------GC-----...------...---------A-G-.. 8973
C.KE.00.KER2010          ..........................---G-G-------AG-..............T----T...------------...A-----.---TC-...A--------...-G---------.------------C-----...------...---------A-G-.. 8766
C.MM.99.mIDU101_3        ..........................---G-G-------AG-..............T----T...------------...------.---TC-...A--------...-G---------.------------C-----...------...---------A-G-.. 8935
C.MW.93.93MW_965         ..........................---G-G-------AG-..............T----T...------------...------.----C-...---------...-G---------.------------C-----...------...---------A-G-.. 8752
C.SN.90.90SE_364         ..........................---G-G-------AG-..............T----T...------------...------.----C-...---------...-G---------.------------C--T--...------...---------A-G-.. 8701
C.SO.89.89SM_145         ..........................---G-G-------AG-..............T----T...------------...A-----.---TC-...A--------...-G---------.------------C--T--...------...---------A-G-.. 8803
C.TZ.02.CO178            ......................................................................................................................................................................... 8637
C.UY.01.TRA3011          ..........................---G-G-------AG-..............A-GA-T...-----------C...T-----.----C-...A--------...-G---------.-----------GC-----...------...-----------G-.. 8777
C.YE.02.02YE511          ..........................---G-G-------AG-..............T----T...---------A--C...------.----C-...---------...-G---------.------------C-----...------...---------A-GA.. 8805
C.ZA.04.04ZASK164B1      ..........................---G-G-------AG-...............AGT-GT...-----------G...------.----C-...A--------...-G---------.-----A------C-----...------...---------A-G-.. 9049
C.ZA.05.05ZAPSK240B1     ..........................---G-G-------AG-..............A----T...-----------A...T-----.---TC-...A--------...-G---------.------------C--T--...------...-----A--A-GC.. 8957
C.ZM.02.02ZM108          ..........................---G-G-------AG-..............A----T...-----------G...------.----C-...---------...-G---------.------------C--T--...------...---------A-G-.. 9616
D.CD.83.ELI              ..........................-----------------............-----A...------------...------.----T-...A--------...-G---------.------------------...------...------------.. 9091
D.CM.01.01CM_0009BBY     ......................................................................................................................................................................... 8421
D.KE.01.NKU3006          ..........................-----------------............CGTA-C..AG-------AC...T-----.----T-...A--------...-G---------.------------------...------...---------T----.. 8797
D.KR.04.04KBH8           ..........................-----------AAGA-GC..............---A-...------------...------.----T-...-------A-...-G---------.-----------C------...------...------------.. 9453
D.TD.99.MN011            ..........................-------------GA-GT..............------...------------...------.----C-...A--------...-G---------.---------C--------...------...------------.. 8763
D.TZ.01.A280             ..........................----------G.--GA-GC.............--AA-...---------G--...------.----C-...A--------...-G---------.-----------C------...------...---------A-C-.. 8758
D.UG.99.99UGK09259       ..........................-----------.--GA-G............A--AA-...AG----------...------.----C-...A--------...-G------A--.-----C------C--T--...------...---------T----.. 8775
D.YE.01.01YE386          ..........................-------------GA-G............T----T...T------AGT-...------.----T-...A--------...-G---------.------------------...------...---------T----.. 8711
D.YE.02.02YE516          ..........................A------G--G.--GA-G............-----A...----------TC...------.----C-...A--------...-G---------.---------C-----T--...------...------------.. 8746
D.ZA.90.R1               ..........................-----------.--GA.G............GTGT-GT...-----------...T-----.----T-...A--------...-G---------.------------C-----...------...-----------G-.. 8909
F1.AR.02.ARE933          ......................................................................................................................................................................... 8338
F1.BE.93.VI850           ..........................-----------.----G.............--GTC-A..GA---------.------.-----TC..A--------...-G---------.-----C-----C-----...------...---------A----.. 8850
F1.BR.01.01BR125         ..........................---A---------A---..............-CC---...----C---A--...-----A...-C-A-..C------A-...-G-----........................................................ 8968
F1.ES.x.P1146            ..........................-----------.----G.............--GTT-A..GA--------A...A-----.----T-...A--------...-G---......................................................... 8877
F1.FI.93.FIN9363         ..........................-----------.----G.............--GTC-A..GA---------.------.-----TC..A--------...-G---------.-----C-----C--T--...---A--...------------.. 8872
F2.CM.02.02CM_0016BBY    ......................................................................................................................................................................... 8349
F2.CM.95.MP257           ......................................................................................................................................................................... 8589
F2.CM.97.CM53657         ..........................-----------.----G..............---TC-A..GA-------C...------.----T-...A--------...-G----C---.-----------GC-----...------...-----------C-.. 8731
G.BE.96.DRCBL            ..........................C-------------GA-G............--C---...-----A---G--...------.----T-...A--------...AG---------.-----C----C-C-----...------...---------T----.. 9539
G.CM.01.01CM_4049HAN     ......................................................................................................................................................................... 8367
G.CU.99.Cu74             ..........................T---.................................................................................................................................... 9098
G.ES.99.X138             ..........................T---------G.--GA-G............--C---...-----A---G--...------.----C-...A--------...A......................................................... 8978
G.GH.03.03GH175G         ..........................T---------G.--GA-G............--C---...-----A---G--...------.----T-...A--------...AG---------.-----C---C-C--T--...------...-------TCT-C-.. 9645
G.KE.93.HH8793_12_1      ..........................T---------G.--GA-G............--C---...A---A------...A-----.----T-...A--------...AG---------.-----C---C-C-----...---G--...---------T----.. 8954
G.NG.01.01NGPL0669       ..........................T---------G.--GA-G............--C---...A-------GACT..----A.......................AG---------.-----C-----C-----...------...---------T-C-.. 8786
G.PT.x.PT2695            ..........................T---------G.--GA-G............--C--T...-----A--G--...------.----C-...A--------...AG---------.-----C---C-C-----...------...---------T-C-.. 9601
G.SE.93.SE6165           ..........................T---------G.--GA-G............--C---...-----A--AGT-...------.----T-...A--------...AG---------.-----C---C-C-----...------...---------T----.. 8996
H.BE.93.VI991            ..........................T--------TG.--GA-G............T----T...------------...A-----.----T-...A--------...-G---------.------------C--T--...------...------------.. 8970
H.BE.93.VI997            ..........................T---------G.--GA-G............--C--T...----------GA-...T-----.----TT..A--------...-G---------.----------C-C--T--...---G--...-------T-----.. 8869
H.CF.90.056              ..........................T---------G.--GA-G............----AT...----------GA-...------.----C-...A--------...-G---------.------------C--T--...------...-------T-----.. 8870
J.CD.97.J_97DC_KTB147    .........................T-----------.--GA-G............T---C-...T-------GT-...------.----T-...A---G-A-...-G------A--.-----C------C--T--...------...------------.. 8729
J.SE.93.SE7887           .........................T-------------GA-G............A-----...--------GT-...------.----T-...A--------...-G---------.------------C--T--...------...---------T----.. 8851
K.CD.97.EQTB11C          ......................................................................................................................................................................... 8600
K.CM.96.MP535            ......................................................................................................................................................................... 8604
```

```
                                                
```

```
                                                                                                    Extensive secondary structure in this region?
                                                     TAR element end          Poly-A signal   3' LTR R repeat end  3' LTR U5 start
B.FR.83.HXB2              .AGACCAGATCTGAGCCTGGGAGCTCTCTGG...CTAACTAGGGAACCCACTGCTTAA.GCCTCAATAAA.GCTTGCCTT.GAGTGC.TTCAAGTAGTGTGTGCCCGTCTGTTGTGTG.ACTCTGGTAACTAGAGATCCCT.CAGACCCTTTTAGTCAGTGTGGAAAATC 9712
A1.GE.99.99GEMZ011        .---------A.-------------------...---G-G---------------................................................................................................................... 8792
A1.KE.00.KER2008          .-------G--.-----C-------------...---G-G---------------................................................................................................................... 8804
A1.KE.00.KNH1144          .---------T.-------------------...---G-----------------................................................................................................................... 8852
A1.KE.00.KSM4024          .----------.-----C-------------...---GTG---------------................................................................................................................... 8744
A1.KE.00.MSA4069          .----------.-------------------...---G-.-A-------------................................................................................................................... 8825
A1.KE.00.NKU3005          .---------T-.............................................................................................................................................................. 8791
A1.RU.00.RU00051          .......................................................................................................................................................................... 8806
A1.RW.93.93RW_024         .---------TGAGC-TG-....................................................................................................................................................... 8797
A1.SE.95.SE8891           .----------.-------------------...---G-G------------------................................................................................................................ 8785
A1.SE.95.UGSE8131         .----------.-------------------...---G-G------------------.-----------.---------.---G--.GCATGC............................................................................ 9016
A1.TZ.01.A173             .---------TGAGC-T......................................................................................................................................................... 8766
A1.UA.01.01UADN139        .----------.-------------------...---G--------------------................................................................................................................ 8852
A1.UG.92.92UG037          .----------.-------------------...---G-G------------------.-----------.---------.------................................................................................... 9690
A1.UG.99.99UGA07072       .--T-------.-------------------...---G-A---------------................................................................................................................... 8824
A2.CD.97.97CDKTB48        .T-GA-CAGATA----------------A--...---G-AG-------------.................................................................................................................... 8972
A2.CY.94.94CY017_41       .G--------.A-----C-------------...---G-AG-----------------.-----------.---------.---A--.---............................................................................... 9060
A.SN.01.DDI579            .---------T--------------------...---G-A---------------................................................................................................................... 8801
A.SN.01.DDJ369            .---------T--------------------...---G-A---------------................................................................................................................... 8861
A.SN.96.DDJ360            .---------T--------------------...A--G-A---------------................................................................................................................... 8777
A.CD.02.02CD_KTB035       .---------A.------A-----------A...---C-G-AA---------------.-----------.---------.----C-.--T---------------T-----------.C--T---------AG-------.---C---C----AAAGAA---A---TCT 9734
A.CD.97.97CD_KCC2         .---------A.-------------------...---G-A------------------.-----------A--------.T------.-------G----------------------.G---------------------.------ACC----CAGAGTGAA---TCT 9733
A.CD.97.97CD_KTB13        .---------A.-------------------...T-GG-A------------------.-----------.---------.------.--T--------------------C------.----------------------.------AC--AGACAGAGTGAA---TCT 9033
B.AR.04.04AR151516        .----------------C-------------...--GG---A-------------................................................................................................................... 8815
B.AU.87.MBC925            .------------------------------...---G--------------------.-----------.---------.------.------------------------------.----------------------.---------------------------- 9732
B.BO.99.BOL0122           .------------------A-----------...--G---------------------................................................................................................................ 8829
B.BR.03.BREPM2012         .......................................................................................................................................................................... 9003
B.CA.97.CANB3FULL         .......................................................................................................................................................................... 8687
B.CN.05.05CNHB_hp3        .----------C-------------------...--G-T-------------------.-----------.---------.------................................................................................... 9641
B.CO.01.PCM001            .----------A-------------------...---G-----------------................................................................................................................... 8787
B.GB.83.CAM1              .-------------------------................................................................................................................................................ 9590
B.GB.86.GB8               .----------------C-------------...---G--G-----------------.-----------.---------.------.--T---------------------------.----------------------.------...................... 9679
B.GE.03.03GEMZ010         .------------------------------...--G------------------................................................................................................................... 8780
B.IT.05.SG1               .------------------------------...------------------------.-----------.---------.------.----.............................................................................. 8805
B.JP.05.DR6538            .------------------------------...--GG----A---------------.-----------.----.----.------................................................................................... 9667
B.KR.05.05CSR3            .......................................................................................................................................................................... 9549
B.NL.00.671_00T36         .------------------------------...---G--G-----------------.--------CT-.TAGG-T-AC.CTAAAT.ACACT-NNN-T--GTTGG-G--TCAAGC-A.GTA-CA--TGA-CC---C-AGG.T---ANAGGCC-A-G-AG-N-AG--CAA 9292
B.RU.04.04RU128005        .----------A-------------------...---G----................................................................................................................................ 9089
B.TH.00.00TH_C3198        .......................................................................................................................................................................... 8394
B.UA.01.01UAKV167         .------------------------------...---G--------------------................................................................................................................ 8865
B.US.04.ES10_53           .----------A-------------------...---G--------------------.-----------.---------.------.--T---------------------------.----------------------.---------------------------- 9705
B.US.99.PRB959_03         .-----TAC-GG-TCT--CT-GTTAGA-CA............................................................................................................................................ 8808
C.AR.01.ARG4006           .------------------------------...---T-----------G-----................................................................................................................... 8822
C.BR.04.04BR013           .......................................................................................................................................................................... 9050
C.BW.00.00BW07621         .------------------------------...---T----A---------------.-----------.---------.------.--T---------------------------.----------------------.------...................... 9047
C.CN.98.YNRL9840          .......................................................................................................................................................................... 8152
C.ET.02.02ET_288          .----------------C--...................................................................................................................................................... 8822
C.GE.03.03GEMZ033         .------------------------------...---T----A------------................................................................................................................... 8835
C.IL.99.99ET7             .---------TA-------------------...---T-----------------................................................................................................................... 8793
C.IN.99.01IN565_10        .------------------------------...---T--------------------.-----------.------T-G.AGTGCT.-................................................................................. 9054
C.KE.00.KER2010           .----------------C-------------...---T-----------------................................................................................................................... 8817
C.MM.99.mIDU101_3         .----------A-------------------...---T--------------------.-----------.---------.------.-CTG--C---------------A-------.----------------------.------...................... 9073
C.MW.93.93MW_965          .----T-----------C--...................................................................................................................................................... 8771
C.SN.90.90SE_364          .----------------C--...................................................................................................................................................... 8720
C.SO.89.89SM_145          .-------------------...................................................................................................................................................... 8822
C.TZ.02.CO178             .......................................................................................................................................................................... 8637
C.UY.01.TRA3011           .---------T--------------------...---T-----------------................................................................................................................... 8828
C.YE.02.02YE511           .G-----------------------------...-C-----------------..................................................................................................................... 8854
C.ZA.04.04ZASK164B1       .------------------------------...---T--G-----------------.---------...................................................................................................... 9112
C.ZA.05.05ZAPSK240B1      .----------------C-------------...A--C----T-----T---------.---------...................................................................................................... 9020
C.ZM.02.02ZM108           .G---------------C-------------...T-CT--------------------.-----------.---------.------................................................................................... 9696
D.CD.83.ELI               .---------T--------------------...---G--------------------.-----------.---------.------.-----............................................................................ 9176
D.CM.01.01CM_0009BBY      .......................................................................................................................................................................... 8421
D.KE.01.NKU3006           .---------T--------------------...---G-----------------................................................................................................................... 8848
D.KR.04.04KBH8            .----------GT---------------GT-..CA--G-A.................................................................................................................................. 9490
D.TD.99.MN011             .------------------A-----------...---G-----------------................................................................................................................... 8829
D.TZ.01.A280              .---------T-------........................................................................................................................................................ 8775
D.UG.99.99UGK09259        .------------------------------...---G--A--G-A--CACTGC-................................................................................................................... 8827
D.YE.01.01YE386           .---------T--------A-----------...---G--------------------................................................................................................................ 8765
D.YE.02.02YE516           .---------T--------A-----------...---G-----------------................................................................................................................... 8797
D.ZA.90.R1                T------................................................................................................................................................................... 8916
F1.AR.02.ARE933           .......................................................................................................................................................................... 8338
F1.BE.93.VI850            .---------T----------G---------...---G--G---------------.................................................................................................................. 8903
F1.BR.01.01BR125          .......................................................................................................................................................................... 8968
F1.ES.x.P1146             .......................................................................................................................................................................... 8877
F1.FI.93.FIN9363          .---------T--------------------...---G--.-A------G--------................................................................................................................ 8925
F2.CM.02.02CM_0016BBY     .......................................................................................................................................................................... 8349
F2.CM.95.MP257            .......................................................................................................................................................................... 8589
F2.CM.97.CM53657          .---------T--------------------...---G-----------------................................................................................................................... 8782
G.BE.96.DRCBL             .---------T--------------------...---G-AG-------.--------G.AG-CTC-AT--.AGC-TG-C-T------.------------------------------.----------------------.------AC-C---AT-----AA---TCT 9699
G.CM.01.01CM_4049HAN      .......................................................................................................................................................................... 8367
G.CU.99.Cu74              .......................................................................................................................................................................... 9098
G.ES.99.X138              .......................................................................................................................................................................... 8978
G.GH.03.03GH175G          .------------------------------...---GTAGA----------------.-----------.---------.------................................................................................... 9725
G.KE.93.HH8793_12_1       .G--------T--------------------C..---G-AG-A---------------.AG-CTC-AT--.AGC-TG-C-.TGAG-G.CG--T-C-AGCC...................................................................... 9047
G.NG.01.01NGPL0669        .---------T--------------------...---GTAG--------------................................................................................................................... 8837
G.PT.x.PT2695             .---------T-------------------A...---G-AG-----------------................................................................................................................ 9655
G.SE.93.SE6165            .---------T--------------------...---G-AG-----------------.AG-CTC-AT--.AGC-TG-CA.AGCC..................................................................................... 9074
H.BE.93.VI991             .----------A-------------------...---G--------------------.-----------.---------.---G--.GCATGC............................................................................ 9056
H.BE.93.VI997             .----------------------T-------...--GG-C-A----------------.-----------.---------.---G--.GCATGC............................................................................ 8955
H.CF.90.056               .------------------------------...---G---A----------------.-----------.---------.------.--T............................................................................... 8953
J.CD.97.J_97DC_KTB147     .----------............................................................................................................................................................... 8739
J.SE.93.SE7887            .------------------------------...------------------------.-----------.---------.---G--.GCATGCA--CCG...................................................................... 8943
K.CD.97.EQTB11C           .......................................................................................................................................................................... 8600
K.CM.96.MP535             .......................................................................................................................................................................... 8604
```


HIV-1/SIVcpz Complete Genomes

```
                                                                                                                  Extensive secondary structure in this region?
                                              TAR element end        Poly-A signal    3' LTR R repeat end  3' LTR U5 start
B.FR.83.HXB2                  .AGACCAGATCTGAGCCTGGGAGCTCTCTGG...CTAACTAGGGAACCCACTGCTTAA.GCCTCAATAAA.GCTTGCCTT.GAGTGC.TTCAAGTAGTGTGTGCCCGTCTGTTGTGTG.ACTCTGGTAACTAGAGATCCCT.CAGACCCTTTTAGTCAGTGTGGAAAATC 9712
01_AE.CF.90.90CF11697         .-------G--.-----C-------------...---G-A------------------.AG-CTC-AT--.AGC-TG-C-.TGAGTG.C-T--AGT-GTGTGTG--CGTCTG--ATA-.................................................... 9628
01_AE.CN.05.FJ051             .---------T.-------------------...---G-AG--------------G-A-----------.---------.------.C-A---CG---------A-----GT-AG-.----------------------.------AC-C---A-T-A--AA--TC-- 9728
01_AE.CN.06.FJ054             .-------G--GAGC--C-T----------G..--G--AC----G----------TA-----------.---------.------.--A----G---------------GT-AG-.----------------------.------AC-C----TT-A--CA---TCT 9726
01_AE.HK.x.HK001              .-------G--GAGC--G--AGCTCTCTG-C...--G-A-----------------G.AG-CTC-AT--.AGC-TG-C-.TGAGTG.C-TG-AGT-GTGTGTG--CGTCTG---TA-.GACTCT-GT-ACT.................................... 9056
01_AE.TH.01.01TH_R2184        .......................................................................................................................................................................... 8408
01_AE.TH.02.OUR769I           .......................................................................................................................................................................... 8390
01_AE.TH.90.CM240             .-------G--.-----C-------------...---G-A------------------.AG-CTC-AT--.AGC-TG-C-.TGAGTG.C-T--.................................................................................... 9203
01_AE.US.00.00US_MSC1164      .......................................................................................................................................................................... 8395
02_AG.CM.02.02CM_4082STN      .......................................................................................................................................................................... 8413
02_AG.EC.x.ECU41              .......................................................................................................................................................................... 8291
02_AG.GH.03.GHNJ196           .---------T------------------A...---G-AG.------------------.-----------.---------.------................................................................................ 9758
02_AG.NG.01.PL0710            .---------T.----------------G-A...---G-AG-A------------.................................................................................................................... 8787
02_AG.NG.x.IBNG               .------------------------------...---G-GGA-----------------.-----------.---------.------.-----------------------A-TGTGT.GACTCT-GT-AC..................................... 9201
02_AG.SE.94.SE7812            .---------T------------------A...---G-AGA-----------------.-----------.---------.---G--.GCATGC............................................................................ 9063
02_AG.SN.98.MP1211            .---------T--------------------...---G-AGA-----------------.-----------.--------......................................................................................... 8851
03_AB.RU.97.KAL153_2          .----------.-------------------...---G---A-----------------.-----------.---------.------................................................................................ 8808
04_cpx.CY.94.CY032            .------------------------------...---G---------------------.-----------.---------.------.---............................................................................ 9050
05_DF.BE.x.VI1310             .---------T--------------------...---G---------------------.--TCA-TA--G.CT-GC-T-G.AG-GCG.CATGC......................................................................... 9083
06_cpx.AU.96.BFP90            .---------T------------------A...---G-AG-------------------.-----------.---------.------.--AT--------------------------.----------------------.------AC-C---A------AA---TCT 9769
08_BC.CN.97.97CNGX_6F         .------------------------------...---G-----------------.................................................................................................................. 8802
09_cpx.GH.96.96GH2911         .----------.A------------------...---G-A-------------------.-----------.---------.------.C-A---------------------------.----------------------.------AC-C---A-GAG--AA---TCT 8884
10_CD.TZ.96.96TZ_BF061        .-----------------C------------...---T--G-------------C---.-----------.-------........................................................................................... 8976
11_cpx.GR.x.GR17              .------------------------------...---G---------------------.-----------.-------AA.-CCGAA.--............................................................................. 8935
12_BF.AR.99.ARMA159           .G--------TA-------------------...---G---------------------.-----------.---------.------.-----------------A-T----------.----------------------.------A--C----------------- 9697
13_cpx.CM.96.1849             .......................................................................................................................................................................... 8916
14_BG.DE.01.9196_01           .......................................................................................................................................................................... 9062
15_01B.TH.99.99TH_MU2079      .-------G--.-----C-------------...---G-A------------------.AGGCTC-AT--.AGC-TTGC................................................................................................. 8999
16_A2D.KR.97.97KR004          .G--------A.-------------------...---G-AG-----------------.AG-CTC-AT--.AGC-TG-C-.TGAGTG.C-GC............................................................................... 9064
17_BF.PE.02.PE02_PCR0155      .------------------------------...---G-----------------.................................................................................................................. 8818
18_cpx.CU.99.CU76             .------------------------------...---G-----R------YK-Y-W--.R-T---TA-..................................................................................................... 8945
19_cpx.CU.99.CU7              .......................................................................................................................................................................... 8512
20_BG.CU.99.Cu103             .......................................................................................................................................................................... 8935
21_A2D.KE.91.KNH1254          .---------T--------------------...---G---------------------.-----------.---.............................................................................................. 8865
22_01A1.CM.01.01CM_0001BBY    .......................................................................................................................................................................... 8372
23_BG.CU.03.CB118             .......................................................................................................................................................................... 8964
24_BG.CU.03.CB378             .......................................................................................................................................................................... 9025
24_BG.CU.03.CB471             .......................................................................................................................................................................... 8987
25_cpx.CM.01.101BA            .------------------------------...---G-----------------.................................................................................................................. 8882
27_cpx.CD.97.97CDKTB49        .------------------------------...---G---------------------.-----------.---------.------.--T---------------------------.----------------------.------AAA--------C-A------AT 8923
28_BF.BR.99.BREPM12609        .......................................................................................................................................................................... 8886
29_BF.BR.02.BREPM119          .......................................................................................................................................................................... 8680
31_BC.BR.02.110PA             .......................................................................................................................................................................... 8908
32_06A1.EE.01.EE0369          .---------T------------------A...---G-AG-A------------................................................................................................................... 9229
33_01B.MY.05.05MYKL007_1      .-------G--.-----C-------------...---G-A------------------.AG-CTC-AT--.AGC-TG-C-.TGAGTG.C-T--AGT-GTGTGTG--CGTCTG---TA-.GACTCT-GT-ACTAGAGAT--C.TCAGA--........................ 9081
34_01B.TH.99.OUR2478P         .......................................................................................................................................................................... 8649
35_AD.AF.05.05AF095           .......................................................................................................................................................................... 8703
36_cpx.CM.00.00CMNYU830       .......................................................................................................................................................................... 8410
37_cpx.CM.00.00CMNYU926       .......................................................................................................................................................................... 8396
38_BF1.UY.05.UY05_4752        ..........................................................T-T............................................................................................................. 8944
39_BF.BR.03.03BRRJ103         .......................................................................................................................................................................... 8861
40_BF.BR.04.04BRRJ115         .......................................................................................................................................................................... 9123
42_BF.LU.03.luBF_05_03        .------------------------------...---G---------------------.-------....................................................................................................... 9063
43_02G.SA.03.J11223           .......................................................................................................................................................................... 9085
U.CD.83.83CD003               .-----------------C------------........................................................................................................................................... 9060
U.CD.90.90CD121E12            .------------------------------...T--G---------------------.----......................................................................................................... 8784
U.GR.99.GR303                 .----------.-------------------...---G-A-----------...................................................................................................................... 8909
U.NL.95.U_NL_95_H10986_D1     .------------------------------...---G--------------------.AG-CTC-AT--.AGC-TG-C-T------.--A---------------------------.----------------------.------A---AGTCAGTGTG-A---TCT 9640
CPZ.CD.90.ANT                 GGAG-------A-------------------...TA-TT-TT--C-GGAGACC--GCT..T--T.......................................................................................................... 9068
CPZ.CM.01.SIVcpzCAM13         .G----------------C-----------C...-AGT..GA----------------.G-CTC-AT--.AGC-TG-C-.TGAGTG.C-TC-............................................................................... 9284
CPZ.CM.05.SIVcpzEK505         .---------TA-------------------...---G-.-A----------------.G-CTC-AT--.AGC-TG-C-.TGAGTG.C-T--............................................................................... 9132
CPZ.CM.05.SIVcpzLB7           .---------TCTGAG-CT------------...---G--A-----------------.AG-CTC-AT--.AGC-TG-C-.TGAGTG.C-GC-.............................................................................. 9170
CPZ.CM.05.SIVcpzMB66          .---------TCTGAG-CT------------...---G--A-----------------.G-CTC-AT--.AGC-TG-C-.TGAGTG.ACTT-A.............................................................................. 9281
CPZ.CM.05.SIVcpzMT145         .------------------------------...-----A------------------.G-CTC-AT--.AGC-TG-C-.TGAGTG.C-GC-.............................................................................. 9208
CPZ.GA.88.GAB1                .G--------T------------------A....-C-GTG-A----------------.-----------.---------.-----TA--TG--C---------TA---A--CA-AC.----------------------.----TTAAA---TAGTCAAG-A---TCT 9805
CPZ.TZ.00.TAN1                .CTG---------------------------....--GTGGCT-GCTAG-GAC-GCTG..-T-A-CGCTC.AA-AAAGCC.TGCCTG.AGAGT--TA......................................................................... 9326
CPZ.US.85.CPZUS               .---------AA-----C-------------...---GTGTA-.---------------.-----------.---------.------.--T-------------TA--CT-CAGTG..----------------------.----A-TC--A-AAGTAG---A---TCT 9775
CPZ.CM.x.SIVcpzDP943          .---------TA-----C-------------...TC-GT.--A---------------.G-CTC-AT--.AGC-TG-C-.TGAGTG.ACTTG............................................................................... 9247
CPZ.CM.x.SIVcpzMB897          .------------------------------...----TAG-A---------------.G-CTC-AT--.AGC-TG-C-.----AA.C-TG............................................................................... 9167
N.CM.02.DJO0131               .......................................................................................................................................................................... 8938
N.CM.04.04CM_1015_04          .......................................................................................................................................................................... 8926
N.CM.04.04CM_1131_03          .......................................................................................................................................................................... 8975
N.CM.95.YBF30                 .G--------TA---T--------ATAT---...G....................................................................................................................................... 9182
N.CM.97.YBF106                .......................................................................................................................................................................... 9045
O.BE.87.ANT70                 .G-----G--------C--------C----...-CTCTAGCT------G--CG----.CG---------.---------.-----A..G................................................................................. 9753
O.CM.91.MVP5180               .G-----G--------C--------C----...-CTCTAGCT------G---------.CG---------.---------.-----...AG---C---------T-A------CA....--C-----GT---------........................................ 9792
O.CM.96.96CMABB637            .G-----G--------C--------C----...-CTCTAGCT------G--....................................................................................................................... 9225
O.CM.98.98CMA104              .G-----G-.------C--------C---C...-CTCTAGCT------G---------.CG............................................................................................................... 9243
O.CM.99.99CMU4122             .GA----G--------C--------C----...-CTCTAGCT------.......................................................................................................................... 9186
O.FR.92.VAU                   .GA----G--------C--------C----...-CTCTAGCT------G---------.CG---------.---------.-----A.GAAGCA.......................................................................... 9330
O.SN.99.SEMP1299              .GA----G--------C--------C----...-CTCTAGCT------G---------.CG---------.---------.-----...AG---C---------T-A------CG....--C------T---------........................................ 9858
O.US.99.99USTWLA              .G-----G--------C--------C----...-CTCTAGCT------.......................................................................................................................... 9186
```

	3' LTR U5 end	
B.FR.83.HXB2	TCTAGCA.	9719
A1.GE.99.99GEMZ011	.....	8792
A1.KE.00.KER2008	.....	8804
A1.KE.00.KNH1144	.....	8852
A1.KE.00.KSM4024	.....	8744
A1.KE.00.MSA4069	.....	8825
A1.KE.00.NKU3005	.....	8791
A1.RU.00.RU00051	.....	8806
A1.RW.93.93RW_024	.....	8797
A1.SE.95.SE889	.....	8785
A1.SE.95.UGSE8131	.....	9016
A1.TZ.01.A173	.....	8766
A1.UA.01.01UADN139	.....	8852
A1.UG.92.92UG037	.....	9690
A1.UG.99.99UGA07072	.....	8824
A2.CD.97.97CDKTB48	.....	8972
A2.CY.94.94CY017_41	.....	9060
A.SN.01.DDI579	.....	8801
A.SN.01.DDI369	.....	8861
A.SN.96.DDI360	.....	8777
A.CD.02.02CD_KTB035	CTACCA..	9740
A.CD.97.97CD_KCC2	CTAGCA..	9739
A.CD.97.97CD_KTB13	CTAGCAG.	9040
B.AR.04.04AR151516	.....	8815
B.AU.87.MBC925	-------.	9739
B.BO.99.BOL0122	.....	8829
B.BR.03.BREPM2012	.....	9003
B.CA.97.CANB3FULL	.....	8687
B.CN.05.05CNHB_hp3	.....	9641
B.CO.01.PCM001	.....	8787
B.GB.83.CAM1	.....	9590
B.GB.86.GB8	.....	9679
B.GE.03.03GEMZ010	.....	8780
B.IT.05.SG1	.....	8805
B.JP.05.DR6538	.....	9667
B.KR.05.05CSR3	.....	9549
B.NL.00.671_00T36	CTGTTTG.	9299
B.RU.04.04RU128005	.....	9089
B.TH.00.00TH_C3198	.....	8394
B.UA.01.01UAKV167	.....	8865
B.US.04.ES10_53	-----...	9710
B.US.99.PRB959_03	.....	8808
C.AR.01.ARG4006	.....	8822
C.BR.04.04BR013	.....	9050
C.BW.00.00BW07621	.....	9047
C.CN.98.YNRL9840	.....	8152
C.ET.02.02ET_288	.....	8822
C.GE.03.03GEMZ033	.....	8835
C.IL.99.99ET7	.....	8793
C.IN.99.01IN565_10	.....	9054
C.KE.00.KER2010	.....	8817
C.MM.99.mIDU101_3	.....	9073
C.MW.93.93MW_965	.....	8771
C.SN.90.90SE_364	.....	8720
C.SO.89.89SM_145	.....	8822
C.TZ.02.CO178	.....	8637
C.UY.01.TRA3011	.....	8828
C.YE.02.02YE511	.....	8854
C.ZA.04.04ZASK164B1	.....	9112
C.ZA.05.05ZAPSK240B1	.....	9020
C.ZM.02.02ZM108	.....	9696
D.CD.83.ELI	.....	9176
D.CM.01.01CM_0009BBY	.....	8421
D.KE.01.NKU3006	.....	8848
D.KR.04.04KBH8	.....	9490
D.TD.99.MN011	.....	8829
D.TZ.01.A280	.....	8775
D.UG.99.99UGK09259	.....	8827
D.YE.01.01YE386	.....	8765
D.YE.02.02YE516	.....	8797
D.ZA.90.R1	.....	8916
F1.AR.02.ARE933	.....	8338
F1.BE.93.VI850	.....	8903
F1.BR.01.01BR125	.....	8968
F1.ES.x.P1146	.....	8877
F1.FI.93.FIN9363	.....	8925
F2.CM.02.02CM_0016BBY	.....	8349
F2.CM.95.MP257	.....	8589
F2.CM.97.CM53657	.....	8782
G.BE.96.DRCBL	CTAGCA..	9705
G.CM.01.01CM_4049HAN	.....	8367
G.CU.99.Cu74	.....	9098
G.ES.99.X138	.....	8978
G.GH.03.03GH175G	.....	9725
G.KE.93.HH8793_12_1	.....	9047
G.NG.01.01NGPL0669	.....	8837
G.PT.x.PT2695	.....	9655
G.SE.93.SE6165	.....	9074
H.BE.93.VI991	.....	9056
H.BE.93.VI997	.....	8955
H.CF.90.056	.....	8953
J.CD.97.J_97DC_KTB147	.....	8739
J.SE.93.SE7887	.....	8943
K.CD.97.EQTB11C	.....	8600
K.CM.96.MP535	.....	8604

3' LTR U5 end

```
B.FR.83.HXB2                    TCTAGCA.  9719
01_AE.CF.90.90CF11697           .......   9628
01_AE.CN.05.FJ051               -AGCA..   9733
01_AE.CN.06.FJ054               CTAGCA..  9732
01_AE.HK.x.HK001                .......   9056
01_AE.TH.01.01TH_R2184          .......   8408
01_AE.TH.02.OUR769I             .......   8390
01_AE.TH.90.CM240               .......   9203
01_AE.US.00.00US_MSC1164        .......   8395
02_AG.CM.02.02CM_4082STN        .......   8413
02_AG.EC.x.ECU41                .......   8291
02_AG.GH.03.GHNJ196             .......   9758
02_AG.NG.01.PL0710              .......   8787
02_AG.NG.x.IBNG                 .......   9201
02_AG.SE.94.SE7812              .......   9063
02_AG.SN.98.MP1211              .......   8851
03_AB.RU.97.KAL153_2            .......   8808
04_cpx.CY.94.CY032              .......   9050
05_DF.BE.x.VI1310               .......   9083
06_cpx.AU.96.BFP90              CTAGCA..  9775
08_BC.CN.97.97CNGX_6F           .......   8802
09_cpx.GH.96.96GH29I1           CTAGCAG.  8891
10_CD.TZ.96.96TZ_BF061          .......   8976
11_cpx.GR.x.GR17                .......   8935
12_BF.AR.99.ARMA159             -------.  9704
13_cpx.CM.96.1849               .......   8916
14_BG.DE.01.9196_01             .......   9062
15_01B.TH.99.99TH_MU2079        .......   8999
16_A2D.KR.97.97KR004            .......   9064
17_BF.PE.02.PE02_PCR0155        .......   8818
18_cpx.CU.99.CU76               .......   8945
19_cpx.CU.99.CU7                .......   8512
20_BG.CU.99.Cu103               .......   8935
21_A2D.KE.91.KNH1254            .......   8865
22_01A1.CM.01.01CM_0001BBY.     .......   8372
23_BG.CU.03.CB118               .......   8964
24_BG.CU.03.CB378               .......   9025
24_BG.CU.03.CB471               .......   8987
25_cpx.CM.01.101BA              .......   8882
27_cpx.CD.97.97CDKTB49          CTCTAGC.  8930
28_BF.BR.99.BREPM12609          .......   8886
29_BF.BR.02.BREPM119            .......   8680
31_BC.BR.02.110PA               .......   8908
32_06A1.EE.01.EE0369            .......   9229
33_01B.MY.05.05MYKL007_1        .......   9081
34_01B.TH.99.OUR2478P           .......A  8650
35_AD.AF.05.05AF095             .......   8703
36_cpx.CM.00.00CMNYU830         .......   8410
37_cpx.CM.00.00CMNYU926         .......   8396
38_BF1.UY.05.UY05_4752          .......   8944
39_BF.BR.03.03BRRJ103           .......   8861
40_BF.BR.04.04BRRJ115           .......   9123
42_BF.LU.03.luBF_05_03          .......   9063
43_02G.SA.03.J11223             .......   9085
U.CD.83.83CD003                 .......   9060
U.CD.90.90CD121E12              .......   8784
U.GR.99.GR303                   .......   8909
U.NL.95.U_NL_95_H10986_D1       CTAGCA..  9646
CPZ.CD.90.ANT                   .......   9068
CPZ.CM.01.SIVcpzCAM13           .......   9284
CPZ.CM.05.SIVcpzEK505           .......   9132
CPZ.CM.05.SIVcpzLB7             .......   9170
CPZ.CM.05.SIVcpzMB66            .......   9281
CPZ.CM.05.SIVcpzMT145           .......   9208
CPZ.GA.88.GAB1                  CTAGCA..  9811
CPZ.TZ.00.TAN1                  .......   9326
CPZ.US.85.CPZUS                 CTAGCAGT  9783
CPZ.CM.x.SIVcpzDP943            .......   9247
CPZ.CM.x.SIVcpzMB897            .......   9167
N.CM.02.DJO0131                 .......   8938
N.CM.04.04CM_1015_04            .......   8926
N.CM.04.04CM_1131_03            .......   8975
N.CM.95.YBF30                   .......   9182
N.CM.97.YBF106                  .......   9045
O.BE.87.ANT70                   .......   9753
O.CM.91.MVP5180                 .......   9792
O.CM.96.96CMABB637              .......   9225
O.CM.98.98CMA104                .......   9243
O.CM.99.99CMU4122               .......   9186
O.FR.92.VAU                     .......   9330
O.SN.99.SEMP1299                .......   9858
O.US.99.99USTWLA                .......   9186
```

III

HIV-2/SIV Complete Genomes

Contents

III-1	Introduction	**149**
III-2	Annotated features	**150**
III-3	Sequences	**152**
III-4	Alignments	**154**

III-1 Introduction

The HIV-2/SIVsmm alignment contains a non-redundant set of all HIV-2, SIVsmm, and related species (SIVmne, SIVstm and SIVmac). Compared to HIV-1, only a small number of HIV-2 complete genomes have been determined, so this alignment contains all but a few duplicte sequences available in the database. The web alignment has a few additional sequences.

The HIV-2/SIVsmm family is presented together in spite of their different hosts, because their genomic structure is the same—the complete SIV alignment (presented in the next chapter along with a few HIV-1 sequences) is much more stretched out because of the presence of a vpx and absence of vpr genes in some of these viruses.

As with HIV-1/SIVcpz, this alignment was generated by an iterative process between automated alignment using HMMER and manual editing using BioEdit and Se-Al. As in previous years, the alignment presented is not suggested to be an "optimal alignment" with the absolute minimum number of gaps and mismatches. It is a compromise between optimal alignment, readability, and an attempt to keep insertions and deletions from altering the protein reading frame presentation. Most gaps have been introduced in multiples of 3 bases to maintain open reading frames when translated directly from the alignment.

As before, the MAC239 sequence is the master sequence in this set.

III-2 Annotated features

Features of HIV-2 annotated in the alignment that follows.

Feature	Location	Page
5' LTR U3 start	1	154
TCF-1 alpha	309-324	155
NF-κ-B-II	415-424	156
NF-κ-B-I	441-449	157
TATA Box	488-493	157
TAR element start	518	157
5' LTR U3 end	520	157
+1 mRNA start site	521	157
5' LTR R repeat begin	521	157
TAR element end	642	158
Poly-A signal	671-676	158
5' LTR R repeat end	694	158
5' LTR U5 start	695	158
5' LTR U5 end	818	159
Lys tRNA primer binding site	821-839	159
Gag and Gag-Pol start	1053	161
Gag p17 Matrix end	1457	163
Gag p24 Capsid start	1458	163
Gag p24 Capsid end	2144	167
Gag p2 Spacer start	2145	167
Gag p2 Spacer end	2195	167
Gag p8 Nucleocapsid start	2196	167
Gag p8 Nucleocapsid end	2351	168
Gag p1 Spacer start	2352	168
Gag-Pol -1 ribosomal slip site	2354	168
Pol start	2354	168
Gag p1 spacer end	2393	169
Gag p6 start	2394	169
Pol Protease start	2555	170
Gag p6 end	2585	170
Gag end	2585	170
Pol Protease end	2851	172
Pol p66 and p51 RT start	2852	172
Pol RT end	4168	179
Pol p15 RNAse H start	4169	179
Pol RNAse H end	4528	182
Pol p31 Integrase start	4529	182
Vif start	5340	186
Pol, Gag-Pol, and p31 Integrase end	5410	187
Vpx start	5812	189
Vif end	5984	190
Vpx end	6150	191
Vpr start	6151	191
Tat exon 1 start	6302	192
Vpr end	6456	193
Rev exon 1 start	6528	193
Tat Rev exon 1 end	6597	194
Tat Rev intron start	6598	194

Feature	Location	Page
Env signal peptide start	6604	194
Env signal peptide end	6669	194
Env gp120 start	6670	194
V1 loop start	6733	195
V1 loop end	6774	195
V2 loop start	6904	196
V2 loop end	7074	197
V3 loop start	7183	198
V3 loop end	7392	199
V4 loop start	7534	200
V4 loop end	7635	200
V5 loop start	7786	201
V5 loop end	7980	202
Rev Responsive Element (RRE) start	8123	203
Env gp120 end	8178	204
Env gp41 start	8179	204
RRE end	8479	205
Tat Rev intron end	8805	207
Tat Rev exon 2 start	8806	207
Tat end	8902	208
Rev end	9059	209
Nef start	9077	209
Env gp41, gp160 end	9243	210
Premature stop in SMM239	9355	211
3' LTR U3 start	9462	211
TCF-1 alpha binding	9770-9785	213
Nef end	9868	214
NF-κ-B-II	9876-9884	214
NF-κ-B-I	9891-9900	214
TATA box	9949-9955	215
TAR element start	9980	215
3' LTR U3 end	9978	215
3' LTR repeat start	9979	215
TAR element end	10103	216
Poly-A signal	10132-10137	216
3' LTR R repeat end	10155	216
3' LTR U5 start	10156	216
3' LTR U5 end	10279	217

III-3 Sequences

Sequences included in the HIV-2/SIV complete genome alignment.

Name	Accession	Country	Author	Reference
MAC.US.x.239	M33262	United States	Kestler, H.	*Science* **248**(4959):1109-1112 (1990)
H2A.CI.88.UC2	U38293	Cote D'Ivoire	Barnett, S.W.	*Virology* **222**(1); 257-61 (1996)
H2A.DE.x.BEN	M30502	Germany	Kirchhoff, F.	*Virology* **177**(1):305-311 (1990)
H2A.DE.x.PEI2	U22047	Germany	Kraus, G.	*ARHR* **14**(1); 65-77 (1998)
H2A.GH.x.GH1	M30895	Ghana	Hasegawa, A.	*ARHR* **5**(6):593-604 (1989)
H2A.IN.NNVA.NEW	EU980602	India	Desai, A.	Unpublished
H2A.IN.95.CRIK_147	DQ307022	India	Santhosh, C.V.	*ARHR* **24**(10):1315-1317 (2008)
H2A.GM.x.MCR35	AY509260	Gambia	Schmitz, C.	*J Virol* **78**(4):2006-2016 (2004)
H2A.GM.87.D194	J04542	Gambia	Kuehnel, H.	*Proc Natl Acad Sci USA* **86**(7):2383-2387 (1989)
H2A.GM.x.ISY	J04498	Gambia	Franchini, G.	*Proc Natl Acad Sci USA* **86**(7):2433-2437 (1989)
H2A.GM.x.MCN13	AY509259	Gambia	Schmitz, C.	*J Virol* **78**(4):2006-2016 (2004)
H2A.GW.86.FG	J03654	Guinea-Bissau	Zagury, J.F.	*Proc Natl Acad Sci USA* **85**(16):5941-5945 (1988)
H2A.GW.87.CAM2CG	D00835	Guinea-Bissau	Tristem, M.	*J Gen Virol* **72**(PT 3):721-724 (1991)
H2A.GW.x.ALI	AF082339	Guinea-Bissau	Azevedo-Pereira, J.M.	Unpublished
H2A.GW.x.MDS	Z48731	Guinea-Bissau	Becker, M.	Unpublished
H2A.SN.85.ROD	M15390	Senegal	Clavel, F.	*Nature* **324**(6098):691-695 (1986)
H2AB.CI.90.7312A	L36874	Cote D'Ivoire	Gao, F.	Unpublished
H2B.CI.88.UC1	L07625	Cote D'Ivoire	Barnett, S.W.	*J Virol* **67**(2):1006-1014 (1993)
H2B.CI.x.EHO	U27200	Cote D'Ivoire	Rey-Cuille, M.A.	*Virology* **202**(1):471-476 (1994)
H2B.GH.86.D205	X61240	Ghana	Dietrich, U.	*Nature* **342**(6252):948-950 (1989)
H2B.JP.01.KR020	AB100245	Japan	Kusagawa, S.	*ARHR* **19**(11):1045-1049 (2003)
H2AB.CM.03CM_510_03.NEW	EU028345	Cameroon	Yamaguchi, J.	*ARHR* **24**(1):86-91 (2008)
H2G.CI.92.ABT96	AF208027	Cote D'Ivoire	Brennan, C.A.	*ARHR* **13**(5):401-404 (1997)
H2U.FR.96.12034	AY530889	France	Damond, F.	*ARHR* **20**(6):666-672 (2004)
SMM.SL.92.SL92B	AF334679	Sierra Leone	Chen, Z.	*J Virol* **70**(6):3617-3627 (1996)
MAC.US.x.MAC239_87082	AY600249	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.251_1A11	M76764	United States	Marthas, M.L.	*J Med Primatol* **18**(3-4):311-9 (1989)
MAC.US.x.251_32H_PJ5	D01065	United States	Rud, E.W.	*(in) Brown,F, Chanock,RM and Ginsberg,HS(Eds); VACCINES* **92**: MODERN APPROACHES TO NEW VACCINES INCLUDING PREVENTION OF AIDS: 229-235; Cold Spring Harbor Laboratory Press, New York (1992)
MAC.US.x.251_BK28	M19499	United States	Hirsch, V.	*Cell* **49**(3):307-319 (1987)

Name	Accession	Country	Author	Reference
MAC.US.x.17EC1	AY033233	United States	Anderson, M.G.	*Virology* **195**(2):616-626 (1993)
MAC.US.x.17EFR	AY033146	United States	Flaherty, M.T.	*J Virol* **71**(8):5790-5798 (1997)
MAC.US.x.1937	AY611495	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.2065	AY611493	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.MM142	Y00277	United States	Chakrabarti, L.	*Nature* **328**(6130):543-547 (1987)
MAC.US.x.SMM142B	BD131285	United States	Alizon, M.	Patent: JP 2002030099-A 2 29-JAN-2002; INSTITUT PASTEUR
MNE.US.82.MNE_8	M32741	United States	Kimata, J.T.	*J Virol* **72**(1):245-256 (1998)
MNE.US.x.MNE027	U79412	United States	Kimata, J.T.	*J Virol* **72**(1):245-256 (1998)
SMM.US.x.F236_H4	X14307	United States	Hirsch, V.M.	*Nature* **339**(6223); 389-92 (1989)
SMM.US.x.H9	M80194	United States	Courgnaud, V.	*J Virol* **66**(1):414-419 (1992)
SMM.US.x.PBJ14_15	L03295	United States	Dewhurst, S.	*Nature* **345**(6276):636-640 (1990)
SMM.US.x.PBJA	M31325	United States	Dewhurst, S.	*Nature* **345**(6276):636-640 (1990)
SMM.US.x.PBJ_143	M80193	United States	Courgnaud, V.	*J Virol* **66**(1):414-419 (1992)
SMM.US.x.PBJ_6P6	L09212	United States	Novembre, F.J.	*J Virol* **67**(5):2466-2474 (1993)
SMM.US.x.PGM53	AF077017	United States	Novembre, F.J.	*J Virol* **72**(11):8841-8851 (1998)
SMM.US.x.SME543	U72748	United States	Hirsch, V.	*J Virol* **71**(2):1608-1620 (1997)
STM.US.x.STM	M83293	United States	Novembre, F.J.	*Virology* **186**(2):783-787 (1992)
MAC.US.x.80035	AY611486	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.81035	AY599200	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.85013	AY611490	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.92050	AY603959	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.92077	AY599201	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.93057	AY611492	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.93062	AY607704	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.95058	AY611494	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.95086	AY607703	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.95112	AY588946	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96016	AY607701	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96020	AY611488	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96072	AY611491	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96081	AY597209	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96093	AY611489	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96114	AY588945	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96123	AY611487	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96135	AY607702	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.97009	AY599199	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.97074	AY599198	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.r80025	AY576480	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.r90131	AY576481	United States	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)

```
                             5' LTR U3 start
MAC.US.x.239                 TGGAAGGGATTTATTACAGTGCAAGAA.GACATAGAATCTTAGACATATACTTAGAAAAGGAAGAAGGCATCATACCAGATTGGCAGGATTACACCTCAGGACCAGGAATTAGATACCCAAAGACATTTGGCTGGCTATGGAAATTAGTCCCTGTAAATGTATCAGATGA  169
H2A.CI.88.UC2.U38293         ----------G-T-------AGGGAT-.-----------C-----T-G---C---------G-----GG-A--------------A----T--TCAT--G------G-G--G-------T-TGC--C--G-----G------C----A--A-----CA-G---C-A--  169
H2A.DE.x.BEN.M30502          --C-------G-T-------AGG--G-.-----------C-----------C------A--G-----G--A--------------A----T--TCAT--G------G-A--G-------T-TAC--C--G-----G-----GC----AT-A---G-AC-C---C-A--  169
H2A.DE.x.PEI2.U22047         ----T-----G----------AG----G-------------------A-G-------------G--A--T------------A-C--T--TCAT--G------G-A--G---------TTC-----G----------GC----A--A---G-C--CC--C-A-G  170
H2A.GH.x.GH1.M30895          ..........................................................................................................................................................................  0
H2A.IN.NNVA.NEW              ----------G----------AG----.---GAGA----A----TT-----------------G--G--A--T-----C------A-C--T---AGT--G--------A-----T----T-TTCCA---G----------GC----A--A--------CC--C-A--  169
H2A.IN.95.CRIK               ----------G-T--------A----T...............................................................................................................................................  27
H2A.GM.x.MCR35.AY509260      ..........................................................................................................................................................................  0
H2A.GM.87.D194.J04542        ..........................................................................................................................................................................  0
H2A.GM.x.ISY.J04498          ..........................................................................................................................................................................  0
H2A.GM.x.MCN13.AY509259      ..........................................................................................................................................................................  0
H2A.GW.86.FG.J03654          ..........................................................................................................................................................................  0
H2A.GW.87.CAM2CG.D00835      --------C-G-T--------A-----.-----------C-G--T-------------A-------G--A--TG----------A-C--T--TAGT--G------G-A----------T-TTC-----G----------GC----A--A---G--ACC---C-A--  169
H2A.GW.x.ALI.AF082339        ----------G-T--------AG----.--------------------------------------G--A--T-----------A-C--T--TCAT--G--------A--G-----G-T-TTC-----G-----G-----GC----A--A---G----CC--C-A--  169
H2A.GW.x.MDS.Z48731          ..........................................................................................................................................................................  0
H2A.SN.85.ROD.M15390         ..........................................................................................................................................................................  0
H2AB.CI.90.7312A.L36874      ------------T---T----A-----.-------G--A-----T-C---T--T--G--T--GAG---T--AG-TGGT-GA------A-C-----A-AT--G-----G--A----------ACAC----------G------C----G--A--GG-A--GG---CA-T  169
H2B.CI.88.UC1.L07625         -------------C--T----AG----.----------AC------C---TC----G--T--G--------TG--T-T-GA-----AA-C--T--A-AT--G-----G--A--G--T----G---C-----G-----G-----GC----G--A---G--A--C----G--  169
H2B.CI.x.EHO.U27200          ------------TC--T----AG----.-G----A---AC------C-----------T----------TG-GT-T-GA------A-C-----ACAT--G------G-A-----T--C---TTC-----------------GC----G--AA-----A-GAT--CA--  169
H2B.GH.86.D205.X61240        ----------G-----T----AG----.----C-----A-------C---T--T--G--T----------TG-GT-T-GC-----AA-C--T--TCAT--G-----G--A--GC-T--C--ATAC-----T-----G-----GC-G--A--A---G-G--GC---CA-C  169
H2B.JP.01.KR020.AB100245     ..........................................................................................................................................................................  0
H2AB.CM.03CM_510_03.NEW      ..........................................................................................................................................................................  0
H2G.CI.92.ABT96.AF208027     ..........................................................................................................................................................................  0
H2U.FR.96.12034.AY530889     ..........................................................................................................................................................................  0
SMM.SL.92.SL92B.AF334679     ..........................................................................................................................................................................  0
MAC.US.x.MAC239              ..........................................................................................................................................................................  0
MAC.US.x.251                 ----------------T----------.---------------------------------------------------------------------------------------------------------------------------------------------  169
MAC.US.x.251_32H             ---------------------------.---------------G------------G-------------------------------------------------------------------------------------C-----------------------  169
MAC.US.x.251                 ---------------------------.---------------G---------------------------------------------------------------------------------------------------------------------------  169
MAC.US.x.17EC1.AY033233      ---------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  169
MAC.US.x.17EFR.AY033146      ---------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  169
MAC.US.x.1937.AY611495       ..........................................................................................................................................................................  0
MAC.US.x.2065.AY611493       ..........................................................................................................................................................................  0
MAC.US.x.MM142.Y00277        ..........................................................................................................................................................................  0
MAC.US.x.SMM142B.BD131285    ..........................................................................................................................................................................  0
MNE.US.82.MNE                ..........................................................................................................................................................................  0
MNE.US.x.MNE027.U79412       ..........................................................................................................................................................................  0
SMM.US.x.F236                ------------------A--AG----.----------A-----T--G-----G------------A--A-------------AA------A-----G---------------T-T-CAC-A---G-----C-------------A---G----C--------  169
SMM.US.x.H9.M80194           ..........................................................................................................................................................................  0
SMM.US.x.PBJ14               ..........................................................................................................................................................................  0
SMM.US.x.PBJA.M31325         ..........................................................................................................................................................................  0
SMM.US.x.PBJ                 ..........................................................................................................................................................................  0
SMM.US.x.PBJ                 ...........................................................................-----------AA-------A-----G--------C--------T-T-TTC-----A-----C-------------G---G----C--------  95
SMM.US.x.PGM53.AF077017      ...........................................................................-----------AA-------A-----G--------C--------T-T-TTC-------------------G-----A----G---C--------  95
SMM.US.x.SME543.U72748       ------------------A--AG-A--.----------A-----T--G---A-G-------------A--A-------------AA-------A-T---G-------C---------T-T-TAC-----G-----C-------------A---G----C--------  169
STM.US.x.STM.M83293          ---------------------AG----...............................................................................................................................................  27
MAC.US.x.80035.AY611486      ..........................................................................................................................................................................  0
MAC.US.x.81035.AY599200      ..........................................................................................................................................................................  0
MAC.US.x.85013.AY611490      ..........................................................................................................................................................................  0
MAC.US.x.92050.AY603959      ..........................................................................................................................................................................  0
MAC.US.x.92077.AY599201      ..........................................................................................................................................................................  0
MAC.US.x.93057.AY611492      ..........................................................................................................................................................................  0
MAC.US.x.93062.AY607704      ..........................................................................................................................................................................  0
MAC.US.x.95058.AY611494      ..........................................................................................................................................................................  0
MAC.US.x.95086.AY607703      ..........................................................................................................................................................................  0
MAC.US.x.95112.AY588946      ..........................................................................................................................................................................  0
MAC.US.x.96016.AY607701      ..........................................................................................................................................................................  0
MAC.US.x.96020.AY611488      ..........................................................................................................................................................................  0
MAC.US.x.96072.AY611491      ..........................................................................................................................................................................  0
MAC.US.x.96081.AY597209      ..........................................................................................................................................................................  0
MAC.US.x.96093.AY611489      ..........................................................................................................................................................................  0
MAC.US.x.96114.AY588945      ..........................................................................................................................................................................  0
MAC.US.x.96123.AY611487      ..........................................................................................................................................................................  0
MAC.US.x.96135.AY607702      ..........................................................................................................................................................................  0
MAC.US.x.97009.AY599199      ..........................................................................................................................................................................  0
MAC.US.x.97074.AY599198      ..........................................................................................................................................................................  0
MAC.US.x.r80025.AY576480     ..........................................................................................................................................................................  0
MAC.US.x.r90131.AY576481     ..........................................................................................................................................................................  0
```

```
                                                                                                                                                                                      TCF-1 alpha
                                                                                                                                                                                     |-----------|
MAC.US.x.239            GGCACAG...GAGGATGAGGAG............CATTATTTAATGCATCCAGCTCAAACTTCCCAGTGGGATGACCCTTGGGGAGAGGTTCTAGCATGGAAGTTTGATCCAACTCTGGCCTACACTTATGAGGCATATGTTAGATACCCAGAAGAGTTTGGAAGCAAGT 324
H2A.CI.88.UC2.U38293    ----G--...--T------ACC............A-C-GC--------C-----A-----AAG-AGA-AT------ATACAT--G---ACA----TG---CG----A-CT-C-TG--------TGAG--CA----C-TCAC-CT------------A-----GCA----- 324
H2A.DE.x.BEN.M30502     ----G--...--A-------CC............A-C-GC---G-A--C-----A-----AAG-AGACAT-----TGAGCAT--G---ACAT---TG---C-------CT-C-TG--------T-AC--CA----C-TCAC-CTG-----------------GCA----- 324
H2A.DE.x.PEI2.U22047    T.............--A-----C..............--C-GC---C-A--C-----A-----AAG-GG--CT---------CAT--G--AACAT--ATG----G------C--T-GG--------TGAG---AC---T-T-AA-C-------------A-----GTAT---- 316
H2A.GH.x.GH1.M30895     .......................................................................................................................................................................... 0
H2A.IN.NNVA.NEW         -..........-GA--A--CACT...GAAACTCACTGC-TAA-GCAC-CAG--CAAAC--G-AAG...-TT------AGACAT--G---ACA----T-----GA--------C--AT-A-----TGA-----TA--C-T-AGGCT-C-------G--A-----GCA----- 324
H2A.IN.95.CRIK         ...........................................................................................................................--CAT---C-TCAGATT-----T---G--A------CA----- 70
H2A.GM.x.MCR35.AY509260 .......................................................................................................................................................................... 0
H2A.GM.87.D194.J04542   .......................................................................................................................................................................... 0
H2A.GM.x.ISY.J04498     .......................................................................................................................................................................... 0
H2A.GM.x.MCN13.AY509259 .......................................................................................................................................................................... 0
H2A.GW.86.FG.J03654     .......................................................................................................................................................................... 0
H2A.GW.87.CAM2CG.D00835 --G-G--...--CAC----ACTGACACTGAGACT--C-GC---T-A--C-----A-----AAG-AG-CAT------ATGCAT--G---ACA--G-TC-----------CT-C-TG------CTT-AG--------C-T-AC-C-------------A-----GCA----- 336
H2A.GW.x.ALI.AF082339   A-GGG--...--CAC----ACT..............--C-GCC-GC-A--C----TA-----AAG-AG-CAT------A-GCAT--G---ACAT---TT----GA-----C--T-AG-----TC-TGA---CA-A--C-T-A--CT-C-------G--A-----GTA----- 324
H2A.GW.x.MDS.Z48731     .......................................................................................................................................................................... 0
H2A.SN.85.ROD.M15390    .......................................................................................................................................................................... 0
H2AB.CI.90.7312A.L36874 AA-C-G-GAG--A--G---ACC...............----G-C--G----C-----A--G--CG--GCA--------T--CAT---GAGA.C---T-TC---C--------T-CCTC-----A--TT-AG-------C-TCAAC--G-T---------------GTATC--- 326
H2B.CI.88.UC1.L07625    --A-AG-...-GA-CG--AACC............AG--G-C-GG----C-----A--G-TC--TTCA---------ATCCAC--G---ACC--T--C---CG------C--CCTC-----AC-TGAC----TA--C-TCAAC--G----------------GTATC--- 324
H2B.CI.x.EHO.U27200     AC--G--...--C--G--AACC...............----G-C-GG----------A--G--C---GCA-----------CCAC-A----ACC--T-TC---C-------CT-CCTC--A--A--TGAC----T---C-TCAGC--G-T------G--------GTATC--- 324
H2B.GH.86.D205.X61240   -A-C-GAGAG-----G--AACC...............----GCC-------C--G--A--G-TC---TCA---------ATCCAT--G---AC---TATC---C--------T-CCTC-----A--TGA-----T---T-TCAA---G-TT-----------C--GTA-C--- 327
H2B.JP.01.KR020.AB100245 ......................................................................................................................................................................... 0
H2AB.CM.03CM_510_03.NEW .......................................................................................................................................................................... 0
H2G.CI.92.ABT96.AF208027 ......................................................................................................................................................................... 0
H2U.FR.96.12034.AY530889 ......................................................................................................................................................................... 0
SMM.SL.92.SL92B.AF334679 ......................................................................................................................................................................... 0
MAC.US.x.MAC239         .......................................................................................................................................................................... 0
MAC.US.x.251            -------..-------------...............---...-----------------------------------------------------C-----------C---------------------------------------------------------------- 324
MAC.US.x.251_32H        -------..---------AG-...............---...-------------G-------------A-----------------------------------------------A--------------------------------------------------- 324
MAC.US.x.251            -------..---------AG-...............---...-------------G-------------A--------------------------G-----------------A--------------------C----------------.---------- 323
MAC.US.x.17EC1.AY033233 -------..-------------...............---...-------------------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.17EFR.AY033146 -------..-------------...............---...-------------------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.1937.AY611495  .......................................................................................................................................................................... 0
MAC.US.x.2065.AY611493  .......................................................................................................................................................................... 0
MAC.US.x.MM142.Y00277   .......................................................................................................................................................................... 0
MAC.US.x.SMM142B.BD131285.......................................................................................................................................................................... 0
MNE.US.82.MNE           .......................................................................................................................................................................... 0
MNE.US.x.MNE027.U79412  .......................................................................................................................................................................... 0
SMM.US.x.F236           A--T---...--A--C---ACA...............----GCC-GG-------G--A--G----AT-------------C----------A--G-----------------GAAT-A--T--T-GC---A------T-A---AG----------------T--T---- 324
SMM.US.x.H9.M80194      .......................................................................................................................................................................... 0
SMM.US.x.PBJ14          .....................................................................................--G-------------------AGT-A--T--T-AC---A------T----GAGCG---------------T--TC--- 79
SMM.US.x.PBJA.M31325    .....................................................................................--G-------------------AGT-A--T--T-AC---A------T----GAGC----------------T--TC--- 79
SMM.US.x.PBJ            .......................................................................................................................................................................... 0
SMM.US.x.PBJ            A--T---...--A--C---ACA...............-----C--G-----------A-------T-------------C-----------A--G-------------------AGT-A--T--T-AC---A------T---GAGC----------------T--TC--- 250
SMM.US.x.PGM53.AF077017 A--T--A...--A--C---ACA...............-----C---G-------------G-TA--------------T--C---------------G---------------T--CAAT-A-------GA---------T-T-A-----C----------------T--T---- 250
SMM.US.x.SME543.U72748  A--C---...--A--C---ACA...............----GCC-GG----------A--G---CAT--------------C-----------A--G-------------------GAAT----T--T-GC---A------T-A---AG-----------------T--T---- 324
STM.US.x.STM.M83293     .......................................................................................................................................................................... 27
MAC.US.x.80035.AY611486 .......................................................................................................................................................................... 0
MAC.US.x.81035.AY599200 .......................................................................................................................................................................... 0
MAC.US.x.85013.AY611490 .......................................................................................................................................................................... 0
MAC.US.x.92050.AY603959 .......................................................................................................................................................................... 0
MAC.US.x.92077.AY599201 .......................................................................................................................................................................... 0
MAC.US.x.93057.AY611492 .......................................................................................................................................................................... 0
MAC.US.x.93062.AY607704 .......................................................................................................................................................................... 0
MAC.US.x.95058.AY611494 .......................................................................................................................................................................... 0
MAC.US.x.95086.AY607703 .......................................................................................................................................................................... 0
MAC.US.x.95112.AY588946 .......................................................................................................................................................................... 0
MAC.US.x.96016.AY607701 .......................................................................................................................................................................... 0
MAC.US.x.96020.AY611488 .......................................................................................................................................................................... 0
MAC.US.x.96072.AY611491 .......................................................................................................................................................................... 0
MAC.US.x.96081.AY597209 .......................................................................................................................................................................... 0
MAC.US.x.96093.AY611489 .......................................................................................................................................................................... 0
MAC.US.x.96114.AY588945 .......................................................................................................................................................................... 0
MAC.US.x.96123.AY611487 .......................................................................................................................................................................... 0
MAC.US.x.96135.AY607702 .......................................................................................................................................................................... 0
MAC.US.x.97009.AY599199 .......................................................................................................................................................................... 0
MAC.US.x.97074.AY599198 .......................................................................................................................................................................... 0
MAC.US.x.r80025.AY576480 ......................................................................................................................................................................... 0
MAC.US.x.r90131.AY576481 ......................................................................................................................................................................... 0
```

```
                                                                                                                                                 NF-κ-B-II
MAC.US.x.239                CAGGCCTGTCAGAGGAAGAGGTTAGAAGAAGGCTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGA....................................AGGAAACTCGCTGAAACAGCAGGGACTTTCCACAAG........................ 429
H2A.CI.88.UC2.U38293        ----AT--C-----A-----TGG-AGGC--AA--G-AA-------GA-A-CAT-T-GT-AAT...-ACAGGAACAACCATACTTGGTCAAAGCAGGAAGTAGCTACTAA...G--A-A-----G--T-------------G---........................ 464
H2A.DE.x.BEN.M30502         ----AT--C-----A----ATGG-AGGC--AA--G-AA-------GA-A-CAT-T-GT-AAT...-ACAGGAACAACCATACTTGGTCAAGGCAGGAAGTAGCTACTAA...G--ACA-----GG-T-------------G---........................ 464
H2A.DE.x.PEI2.U22047        --------C----A------TGG-AGGC--AA--G-AA-------GA-A-CATTT-GTTAAA...GACAGGAACAGCTATATTTGGTCAGAACAGGAAGTA..............GAT-A------T-------------G---........................ 445
H2A.GH.x.GH1.M30895         ......................................................................................................................................................................... 0
H2A.IN.NNVA.NEW             -------AC-----------TGG-AGGCG--AT---AG-----G--A--A-CATTT-ATTAAAGG-GACAGGAACAACTATACTTGGTCAGGGCAGGAAGTAGCTATT...----ACA-----G--T-------------G---........................ 466
H2A.IN.95.CRIK              -----T-AC-----------TGG-AGGCG--A--G-AA-----G--AA-A-CCTTT-GCTAA-..-GACAGGAACAACTATACTTGGTCAGAACAGGAAGTAGCTACT...----ACA-T---G--T-------------G-T-........................ 210
H2A.GM.x.MCR35.AY509260     ......................................................................................................................................................................... 0
H2A.GM.87.D194.J04542       ......................................................................................................................................................................... 0
H2A.GM.x.ISY.J04498         ......................................................................................................................................................................... 0
H2A.GM.x.MCN13.AY509259     ......................................................................................................................................................................... 0
H2A.GW.86.FG.J03654         ......................................................................................................................................................................... 0
H2A.GW.87.CAM2CG.D00835     -------AC----A--T---TGG-AGGCG-AA--G-AA-------GA-A-CATTT-GTTAA-...G-CAGGAACAACCATACATGGCCAGGGCAGGAAGTAGCTACT...----ACA-----G--T-------------G---........................ 475
H2A.GW.x.ALI.AF082339       --------C----A--T---TGG-AGGC---A--G-AA-------GA-A-CATTT-GTAAGA....ACAGGAACAGCTG.ATTTGGTCAGGGCAGGAAGTAACTACT...----ACA-----G--T-------------G---........................ 461
H2A.GW.x.MDS.Z48731         ......................................................................................................................................................................... 0
H2A.SN.85.ROD.M15390        ......................................................................................................................................................................... 0

H2AB.CI.90.7312A.L36874     ----AT-AC-----A-G--ATGG-AGGCT--A----AA---------A-A-C--CAGA-TAGACAGGA-AGCAGCAGCATAAAGAGGAACTAGCTGACGCTGCATAAGAA--------G-----C--T-------------G---........................ 472

H2B.CI.88.UC1.L07625        ----AT-AC-----A-G---TGA-AGGCT--A----AA-------GA-A-C--CAGACTAG-CA-GAG.ACAGCAGCATAAACAGGAACTAGCTGACACTGCACAAG.AG-------A-----C--C-------------A---........................ 468
H2B.CI.x.EHO.U27200         ----AA--C-----A-----TGG-AGGCT-AA--G-GA-------AA-A-C--CAGA-TAGACAGGA-.ATAACAGCACAAAGAGGAACTAGCTAACACTGCATAGAGA--------A-----T--T--.A----------G---........................ 468
H2B.GH.86.D205.X61240       ----G--AC-------G---TGG-AGGCT--A----AA-------GA-A-C--CAGATTAG-CA-GAG.ACAGCAGCATAAACAGGAACTAGCTGACACTGCACAAG.A--------A--A--C--T-------------A---........................ 471
H2B.JP.01.KR020.AB100245    ......................................................................................................................................................................... 0

H2AB.CM.03CM_510_03.NEW     ......................................................................................................................................................................... 0

H2G.CI.92.ABT96.AF208027    ......................................................................................................................................................................... 0

H2U.FR.96.12034.AY530889    ......................................................................................................................................................................... 0

SMM.SL.92.SL92B.AF334679    ......................................................................................................................................................................... 0

MAC.US.x.MAC239             ......................................................................................................................................................................... 0
MAC.US.x.251                -------------------------------------------------------------------.....................................--------A-A---G--------------------........................ 429
MAC.US.x.251_32H            --------C----------------------------------------------------------.....................................---------------G--------------------........................ 429
MAC.US.x.251                -------------------------------------------------------------------.....................................G--------------G-T------------------........................ 428
MAC.US.x.17EC1.AY033233     -------------------------------------------------------------------.....................................-------------------------------------........................ 429
MAC.US.x.17EFR.AY033146     -------------------------------------------------------------------.....................................-------------------------------------........................ 429
MAC.US.x.1937.AY611495      ......................................................................................................................................................................... 0
MAC.US.x.2065.AY611493      ......................................................................................................................................................................... 0
MAC.US.x.MM142.Y00277       ......................................................................................................................................................................... 0
MAC.US.x.SMM142B.BD131285.  ......................................................................................................................................................................... 0

MNE.US.82.MNE               ......................................................................................................................................................................... 0
MNE.US.x.MNE027.U79412      ......................................................................................................................................................................... 0

SMM.US.x.F236               -----T---------------A-AG---------------------T-A--.------------.....................................-------AA-----G--------------------........................ 428
SMM.US.x.H9.M80194          ......................................................................................................................................................................... 0
SMM.US.x.PBJ14              -----T-----A---------GCAG---------------------T-A--A------------.....................................--A----AA---A-G-------------------AGA..................... 186
SMM.US.x.PBJA.M31325        -----T-----A---------GCAG---------------------T-A--A------------.....................................--A----AA---A-G-------------------AGA..................... 186
SMM.US.x.PBJ                ......................................................................................................................................................................... 0
SMM.US.x.PBJ                -----T-----A---------GCAG---------------------T-A--A------------.....................................--A----AA---A-G-------------------AGACAGCAGGGACTTTCCACAAAGA 379
SMM.US.x.PGM53.AF077017     -----T------G--------A-AG---------------------T-A--A------------.....................................-------GA-----G-------------------A....................... 355
SMM.US.x.SME543.U72748      -----T---------------A-AG-------------------A--TA---A-----------.....................................G------AA-----G---------------------GGACTT.................. 435

STM.US.x.STM.M83293         ..............................-------------------A--G------------.....................................-------AA-----G--------------------........................ 102

MAC.US.x.80035.AY611486     ......................................................................................................................................................................... 0
MAC.US.x.81035.AY599200     ......................................................................................................................................................................... 0
MAC.US.x.85013.AY611490     ......................................................................................................................................................................... 0
MAC.US.x.92050.AY603959     ......................................................................................................................................................................... 0
MAC.US.x.92077.AY599201     ......................................................................................................................................................................... 0
MAC.US.x.93057.AY611492     ......................................................................................................................................................................... 0
MAC.US.x.93062.AY607704     ......................................................................................................................................................................... 0
MAC.US.x.95058.AY611494     ......................................................................................................................................................................... 0
MAC.US.x.95086.AY607703     ......................................................................................................................................................................... 0
MAC.US.x.95112.AY588946     ......................................................................................................................................................................... 0
MAC.US.x.96016.AY607701     ......................................................................................................................................................................... 0
MAC.US.x.96020.AY611488     ......................................................................................................................................................................... 0
MAC.US.x.96072.AY611491     ......................................................................................................................................................................... 0
MAC.US.x.96081.AY597209     ......................................................................................................................................................................... 0
MAC.US.x.96093.AY611489     ......................................................................................................................................................................... 0
MAC.US.x.96114.AY588945     ......................................................................................................................................................................... 0
MAC.US.x.96123.AY611487     ......................................................................................................................................................................... 0
MAC.US.x.96135.AY607702     ......................................................................................................................................................................... 0
MAC.US.x.97009.AY599199     ......................................................................................................................................................................... 0
MAC.US.x.97074.AY599198     ......................................................................................................................................................................... 0
MAC.US.x.r80025.AY576480    ......................................................................................................................................................................... 0
MAC.US.x.r90131.AY576481    ......................................................................................................................................................................... 0
```

```
                                                
```

```
                                                                          TAR element end┤                                           Poly-A signal5' LTR R repeat end┼5' LTR U5 start
MAC.US.x.239               CAGGTAGAGCCTGGGTGTTCCCTGCTAGACTCTCACC.AGCACTTGGCCGGTGCTGGGCAGAGTGACTCCACGCTTGCTTGCTTAAAGC..CCTCTTCAATAAAGCTGCCATTTTAGAAGTAAGCTA..GTGTGTGTTCCCATCTCTCCTAGCCGCCGCCTGGTCAACTC  740
H2A.CI.88.UC2.U38293       -C---------C-------------------------.--TG--C------CA---------.C-G---------------------A..------A-TA--G.------A--AGA-GCA-GTTA-..----------------------T-------------TTCG  775
H2A.DE.x.BEN.M30502        --------------------------G----------.--T----------A---------.C-G---------------------A.C-TCT-CA-TA--G.------G--AGA-GCA-GTTA-..----------------------T-------------TTCG  776
H2A.DE.x.PEI2.U22047       -------------------------------------.--T----------A---------.C-G---------------------A..------A-TA--G.------G--AGA-GCA-GTTA-..----------------------T-------------TTCG  755
H2A.GH.x.GH1.M30895        -------------------------------------.--TG---------CA---------.C-G---------------------A..------A-TA--G.------A--AGA-GCA-GTTA-..----------------------T-------------TTCG  220
H2A.IN.NNVA.NEW            ------------------...A---------------.--TG---------C----------C.-G---------------------A.C-TCT-CA-TA--G.------G--AGA-GCA-GT-A-..----------------------T-------------GTCG  770
H2A.IN.95.CRIK             --------------------------G----------.--TG---------CA---------.C----------------------A..------A-TA--G.------G--AGA-GCA-GTTA-..------------------T----T-------------TTCG  521
H2A.GM.x.MCR35.AY509260    -------------------------------------.G-TG---------CA---------.C-G--------------------AGA.------A-TA--G.------G--AGA-GCA-GTTA-..----------------------T-------------TTCG  221
H2A.GM.87.D194.J04542      -------------------------------------.--TG---------CA---------.C-G---------------------A..------A-TA--G.------G--AGA-GCA-GTTA-..----------------------T-------------TTCG  220
H2A.GM.x.ISY.J04498        --------------------------G----------.--TG---------C----------.C-G--------------------AGA.------A-TA--G.------G--AGA-GCA-GTTA-..----------------------T-------------TTCG  221
H2A.GM.x.MCN13.AY509259    -------------------------------------.G-TG---------CA---------.C-G--------------------AGA.------A-TA--G.------G--AGA-GCA-GTTA-..----------------------T-------------TTCG  221
H2A.GW.86.FG.J03654        -------------------------------------.--TG---------C----------.C-G---------------------A..-----CA-TA--G.------A--AGA-GCA-GTTA-..-------C--------------T-------------T-CG  219
H2A.GW.87.CAM2CG.D00835    -------------------------------------.--TG---------CA---------.C-G---------------------A..------A-TA--G.------G--AGA-GCA-GTTA-..----------------------T-------------TTCG  786
H2A.GW.x.ALI.AF082339      -------------------------------------.--TG---------CA---------.C-G---------------------A..------A-TA--G.------A--AGA-GCAGGTTA-..AG--------------------T-------------TTCG  771
H2A.GW.x.MDS.Z48731        -------------------------------------.--TG---------CA---------.C-G---------------------T..------A-TA--G.------G--AGA-GCA-GTTA-..----------------------T-------------TTCG  220
H2A.SN.85.ROD.M15390       -------------------------------------.------------------------.C-G-C------------------AA..------A-TA--G.------G--AGA-GCA-GTTA-..-------C--------------T-------------TTCG  220
H2AB.CI.90.7312A.L36874    -------------------------------------.------------------------.C-G--------------------A-..------A-TA--G.------A--AGA-GCA-GT-A-..----------------------T-------------CTCG  783
H2B.CI.88.UC1.L07625       -------------------------------------.------------------------.C-G--------------------A-..------A-TA--..------A--AGA-GCA-GTTG-..----------------------T-------------TTCA  777
H2B.CI.x.EHO.U27200        -------------------------------------.------------------------.C-G--------------------A-..------A-TA---.-----TA--AGA-GCA-GA-A-..----------------------T-------------TTCG  778
H2B.GH.86.D205.X61240      -------------------------------------.------A----A---T--------.C-G--------------------A-..------A-TA--..------A--AGA-GCA-GTAA...----------------------T---.---------CTCG  778
H2B.JP.01.KR020.AB100245   .........................................................................................................................................................................    0
H2AB.CM.03CM_510_03.NEW    .........................................................................................................................................................................    0
H2G.CI.92.ABT96.AF208027   ..................-------------------.---G-------------------Y-G-----------------C-AAGA------A-TA--G.------AA-----------A-GT---------------------T-------------C---  147
H2U.FR.96.12034.AY530889   -------------------------------------.--T----------A-----------------------------------A-..-----CA-TA--G.------G--AGA-GCA-GCTA-..----------------------T-------------TTCG  272
SMM.SL.92.SL92B.AF334679   GTA-GT-------------T--A---A-------A-C----ACA----AT-------T---.C-G-----------------G-.........----CA-TA--G----A-C--AGA-GCA-GT-A-..-A--------T-------------T-------------TTCG  217
MAC.US.x.MAC239            .........................................................................................................................................................................    0
MAC.US.x.251               ---------------T---------------------.-------------------------------------------------..-------------------------------.---------------------T-----------------  739
MAC.US.x.251_32H           -------------------------------------.-------------------------G---------------------A..------------------------------.---------------------T-----------------  740
MAC.US.x.251               -------------------------------------.-----------A-------------G---------------------A..-------------------------------C-.---------------------T-----------------  728
MAC.US.x.17EC1.AY033233    -------------------------------------.-----------------------------------------------..-------------------------------.-----------------------------------------  740
MAC.US.x.17EFR.AY033146    -------------------------------------.-----------------------------------------------..-------------------------------.-----------------------------------------  740
MAC.US.x.1937.AY611495     .........................................................................................................................................................................    0
MAC.US.x.2065.AY611493     .........................................................................................................................................................................    0
MAC.US.x.MM142.Y00277      -------------------------------------.-------------------------G---------------------A..-----------------------.----------.---------------------T-----------------  221
MAC.US.x.SMM142B.BD131285--------------------------------------.-----------------------C--G---------------------A..-----------------------.----------.---------------------T-----------------  221
MNE.US.82.MNE              -------------------------------------.-------------------------G---------------------A..-----------------T---.----------C-.-------C--------------T-----------------  221
MNE.US.x.MNE027.U79412     -------------------------------------.------------------T------G---------------------A..-----------------------.----------C-..-C--------------------T-----------------  221
SMM.US.x.F236              -------------------------------------.-------------------------G---------------------A..-----------------A--.----------A-..----------------------T-------------T---  736
SMM.US.x.H9.M80194         --------------------------G----------.------------------------RG--------------------AGA.-----------------G--.----------C-..----------------------M-------------T---  223
SMM.US.x.PBJ14             -------------------------------------.-------------------------G---------------------A..-----------------------.----------C-..------------------------------------T---  516
SMM.US.x.PBJA.M31325       -------------------------------------.-------------------------G---------------------A..-----------------------.----------C-..------------------------------------T---  516
SMM.US.x.PBJ               -------------------------------------.------------------------RG--------------------AGA.-----------------------.----------C-..------------------M----Y-------------T---  223
SMM.US.x.PBJ               -------------------------------------.-------------------------G---------------------A..----------------------..----------C-..------------------------------------T---  708
SMM.US.x.PGM53.AF077017    ----C--------------------------------.-----------------------C-G---------------------A..-----------------GC--.----------A-..----------------------T-------------T---  665
SMM.US.x.SME543.U72748     -------------------------------------.-------------------------G---------------------A..-----------------------.----------A-..----------------------T-------------T---  752
STM.US.x.STM.M83293        ------------A------------------------.-----------A-------T-----G-----------------T-TA..T-----------------A------------A-..----------------------T-----------------  413
MAC.US.x.80035.AY611486    .........................................................................................................................................................................    0
MAC.US.x.81035.AY599200    .........................................................................................................................................................................    0
MAC.US.x.85013.AY611490    .........................................................................................................................................................................    0
MAC.US.x.92050.AY603959    .........................................................................................................................................................................    0
MAC.US.x.92077.AY599201    .........................................................................................................................................................................    0
MAC.US.x.93057.AY611492    .........................................................................................................................................................................    0
MAC.US.x.93062.AY607704    .........................................................................................................................................................................    0
MAC.US.x.95058.AY611494    .........................................................................................................................................................................    0
MAC.US.x.95086.AY607703    .........................................................................................................................................................................    0
MAC.US.x.95112.AY588946    .........................................................................................................................................................................    0
MAC.US.x.96016.AY607701    .........................................................................................................................................................................    0
MAC.US.x.96020.AY611488    .........................................................................................................................................................................    0
MAC.US.x.96072.AY611491    .........................................................................................................................................................................    0
MAC.US.x.96081.AY597209    .........................................................................................................................................................................    0
MAC.US.x.96093.AY611489    .........................................................................................................................................................................    0
MAC.US.x.96114.AY588945    .........................................................................................................................................................................    0
MAC.US.x.96123.AY611487    .........................................................................................................................................................................    0
MAC.US.x.96135.AY607702    .........................................................................................................................................................................    0
MAC.US.x.97009.AY599199    .........................................................................................................................................................................    0
MAC.US.x.97074.AY599198    .........................................................................................................................................................................    0
MAC.US.x.r80025.AY576480   .........................................................................................................................................................................    0
MAC.US.x.r90131.AY576481   .........................................................................................................................................................................    0
```

```
                                                                                             5' LTR U5 end  Lys tRNA primer binding site
MAC.US.x.239                 GGTACTCAATA.AT.AAGAAGACCCTGGTC.TGTTAGGAC................CCTTTCTGCTTTGGGAAACCG.AAGCAG...GAAAATCCCTAGCAGATTGGCGCCTGAACAGGGACTTGA.AGGAGAGTGAGAGACTCCTGA.GTACGGCTGAGTGAAGGCAGT  885
H2A.CI.88.UC2.U38293         -TGTTCATC-GAG.T--C------------.---------................----C-C----------T--A.----G-...---------------G--------C-------------.--AG--C-----AG-C-TG--.AC-------------------  921
H2A.DE.x.BEN.M30502          -TGTTCATC-GAG.T--C------------.---------................----CTC---------T--A.-G---...---------------G--------C-------------.--AG--C-----AG-C-TG--.ACT------------------  922
H2A.DE.x.PEI2.U22047         -TGTTCACC--AG.TG-C------------.---------................----CT--------G----.----G-...------A-----------------C-------ACT-GA-............G--GA-TG--.AC-------------------  890
H2A.GH.x.GH1.M30895          -TGTTCATC-GA-.T--C------------.---------................----C.C---------T--A.-G---...---------------G--------C.------------.-A-G--C-----AG-C-TG--.AC------AGTGA-G-TAGTA  364
H2A.IN.NNVA.NEW              -TGTTCACC-GAG.T--C------------.---------................----CT----------T--A.-G--T-...---------------G--------C-------------.--A--GC-----AG-C-TG--.AC-------------------  916
H2A.IN.95.CRIK               -TGTTCATC-GA-.T--C------------.---------................----CT------------CG-G---...-----------T---G--------C-------------.------C-----AGTCTTG--.AC-------------------  668
H2A.GM.x.MCR35.AY509260      -TGTTCATC-GA-.T--C------------.---------................-------A-----------A.------...---------------G--------C-------------.--A---C---A-AG-CTTG--.AC-------------------  367
H2A.GM.87.D194.J04542        -TGTTCATC-GAG.T--C------------.---------................----C-C-----A---T--A.-G---...---------------G--------C------------A--AG--C-----AG-C-TG--.AC-------------------  367
H2A.GM.x.ISY.J04498          -TGTTCTCC-GAG.T--C------------.---------................----CT--------------.-G---...---------------G--------C-------------.G-A---C-----AG-CTTG--.AC-------------------  367
H2A.GM.x.MCN13.AY509259      -TGTTCATC-GA-.T--C------------.---------................-------------------A.------...---------------G--------C-----------AG.--A---C---A-AG-CTTG--.AC-------------------  367
H2A.GW.86.FG.J03654          -TGTTCACC-GT-.T--C------------.---------................----CT--------------.-G---GCG---------------G--------C------.------.--A---C-----AG-CTTG--ACA-G-----G-----------  368
H2A.GW.87.CAM2CG.D00835      -TGTTCACC-GA-.T--C------------.---------................----CT-------------A.-G---...---------------G--------C-------------.--A---C---A-AG-CTTG--.AC-----------G------  932
H2A.GW.x.ALI.AF082339        -TGTTCACC-GAG.T--C------------.---------................-----------------G-.-G---...---------------G--------C-------------.--A---C-----AGTCTAG--.AC-------------------  917
H2A.GW.x.MDS.Z48731          -TGTTCACC-GA-.T--C------------.---------................----CT--------------.-G---...---------------G--------C-------------.--A---C-----AGTCTTGA-.AC-------------------  366
H2A.SN.85.ROD.M15390         -TGTTCACC-GAG.T--C------------.---------................----CT--------------.-G---...---------------G---------------------.--A---C-----AGTCTTG--.AC-------------------  366
H2AB.CI.90.7312A.L36874      -TGTTCAGC-GA-.T--.-G----------.---------................----CT------------T.------...---------------G--------C---------CC--..-A---------T--T-G--.AC--A---------------  927
H2B.CI.88.UC1.L07625         ATGCTCT-C--A-AT--A..----------.---------................------------------A.-G---...------------------------C---------A---.--A---------T--T-G--.-C-------------------  922
H2B.CI.x.EHO.U27200          -TGCTC--C--G-A.......---------.---------................--C---------------A.-G---...------------------------C---------C---.G-A---------T--T-G-G.AC-------------------  918
H2B.GH.86.D205.X61240        -TGCTC--C-GA......T-----------.---------................----CT--------------.----G-...------------------------C---------C--G..-A---------T--T-AAGA-C--A---------------  919
H2B.JP.01.KR020.AB100245     .........................................................................................................------C------------.G-A---------T--T-G--.-C-----------G------   64
H2AB.CM.03CM_510_03.NEW      ...............................................................................................................C--GA-CAG--AC--G--....----------T--T-A-GA-C--A---------------   58
H2G.CI.92.ABT96.AF208027     ---GT--GCGT....-GC------------.---------................------------RG---A.G--G-...------------------------C-------------.GAAG---------T-CTT---.AC-----A-------------  290
H2U.FR.96.12034.AY530889     -AG-T---CCT.....TAT-A--------AC---------................----CT-------------A.G-----...---------------G--------C-------------....----------T--T----AACT-----------G------  413
SMM.SL.92.SL92B.AF334679     --GT-C-GGAT.....-ACGAG-------AC---C--A--................---C--ACTC-----GGCAAC.TCCTGAGTGA------T--G----T--------C-------------................AGAC-G.ACT-----------------  348
MAC.US.x.MAC239              .........................................................................................................................................................................    0
MAC.US.x.251                 -------GG--.--.---------------.---------................------------------.------..------------------------C-------------.---------------------.---------------------  884
MAC.US.x.251_32H             -------GG--.--.---G-----------.---------................-----------A------.------..----------------------------------------.---------------------.---------------------  885
MAC.US.x.251                 -------GG--.--.---------------.---------................-----------A------.------..------------------------C--------.----.---------------------.---------------------  872
MAC.US.x.17EC1.AY033233      ----------.--.----------------.---------................------------------.------..-----------------------------------------.---------------------.---------------------  885
MAC.US.x.17EFR.AY033146      ----------.--.----------------.---------................------------------.------..-----------------------------------------.---------------------.---------------------  885
MAC.US.x.1937.AY611495       .........................................................................................................................................................................    0
MAC.US.x.2065.AY611493       .........................................................................................................................................................................    0
MAC.US.x.MM142.Y00277        -------GG--.--.--A------------.---------CCTGGTCTGTTAGGAC-----------------------.------..------------------------C-------------.---------------------.---------------------  382
MAC.US.x.SMM142B.BD131285    -------GG--.--.--A------------.---------CCTGGTCTGTTAGGAC-----------------------.------..------------------------C-------------.---------------------.---------------------  382
MNE.US.82.MNE                -------G---.--.---------------.---------................------------------.------..------------------------C-------------.---------------------.---------------------  366
MNE.US.x.MNE027.U79412       ----------.--.----------------.---------................------------------.------..------------------------C-------------.---------------------.---------------------  366
SMM.US.x.F236                -------G-C-C--.---------------.---------................------------A-----.-G----..---------------G--------C-------------.-------------.-------.---------------------  881
SMM.US.x.H9.M80194           -------G-.-C--.------S--------.---------...............Y------------------.-G----..---------------G--------C-------------.-------------.-------.--------------------C  367
SMM.US.x.PBJ14               -------GG.-C--.------C--------.---------................------------------.-G----..---------------G--------C-------------.-------------.-------.--------------------C  660
SMM.US.x.PBJA.M31325         -------G-.-C--.------C--------.---------................------------------.-G----..---------------G--------C-------------.-------------.-------.--------------------C  660
SMM.US.x.PBJ                 -------G-.-C--.------S--------.---------...............Y------------------.-G----..---------------G--------C-------------.-------------.-------.--------------------C  367
SMM.US.x.PBJ                 -------G-.-C--.------C--------.---------................------------------.-G----..---------------G--------C-------------.-------------.-------.-------C-----------C  852
SMM.US.x.PGM53.AF077017      -------G.A-C--A---------------.---------................-------------C-----.-G----..------------------------C-------------.-------------.-------.---------------------  810
SMM.US.x.SME543.U72748       -------G-C-C--A---------------.---------................-------------------.-G----..---------------G--------C-------------.-----GTGAGAGC..------.---------------------  897
STM.US.x.STM.M83293          ---------G..........----------.---------................-------------CT--T-.-G----...A-----------------------C-------------.---------------------.---------------------  549
MAC.US.x.80035.AY611486      .........................................................................................................................................................................    0
MAC.US.x.81035.AY599200      .........................................................................................................................................................................    0
MAC.US.x.85013.AY611490      .........................................................................................................................................................................    0
MAC.US.x.92050.AY603959      .........................................................................................................................................................................    0
MAC.US.x.92077.AY599201      .........................................................................................................................................................................    0
MAC.US.x.93057.AY611492      .........................................................................................................................................................................    0
MAC.US.x.93062.AY607704      .........................................................................................................................................................................    0
MAC.US.x.95058.AY611494      .........................................................................................................................................................................    0
MAC.US.x.95086.AY607703      .........................................................................................................................................................................    0
MAC.US.x.95112.AY588946      .........................................................................................................................................................................    0
MAC.US.x.96016.AY607701      .........................................................................................................................................................................    0
MAC.US.x.96020.AY611488      .........................................................................................................................................................................    0
MAC.US.x.96072.AY611491      .........................................................................................................................................................................    0
MAC.US.x.96081.AY597209      .........................................................................................................................................................................    0
MAC.US.x.96093.AY611489      .........................................................................................................................................................................    0
MAC.US.x.96114.AY588945      .........................................................................................................................................................................    0
MAC.US.x.96123.AY611487      .........................................................................................................................................................................    0
MAC.US.x.96135.AY607702      .........................................................................................................................................................................    0
MAC.US.x.97009.AY599199      .........................................................................................................................................................................    0
MAC.US.x.97074.AY599198      .........................................................................................................................................................................    0
MAC.US.x.r80025.AY576480     .........................................................................................................................................................................    0
MAC.US.x.r90131.AY576481     .........................................................................................................................................................................    0
```

```
MAC.US.x.239              AAGGGCGGCAGGAACCAACCACGACGGAGTGCTCCTATAAAGGCGCGGGTC...GGTACCAGA..CGGCGTGAGGAGCGG.GAGAGGAAGAGGCCTCCG..GTTGCA..GGTAAGTGCAACACAAAAAAG.............AAATAGCTGTCT.TTTATCCAGGAAGG 1031
H2A.CI.88.UC2.U38293      ---------------A--------------------G---A-------C-AAG-TAC-A-A-GG-------T-------.---T-A------------..-G--A-..---------CTACACC----A.CTGTAGCCAG.A---GGCT---TATCC--CT..TTT--A 1082
H2A.DE.x.BEN.M30502       ---------------A--------------------G-----------C-GAG-TAC-A-AGG.-------T-------.---T-A------------..-G--A-..-------A-CTACACC----A.CTGTAGCCAG.A---GGCT---TATCC--C-..TTT--A 1082
H2A.DE.x.PEI2.U22047      ---------------A--------------------G-----------C-GAG-TAC-A-AGG.-------T-------.---T.A------------..-G--A-..-------A-CTACACC----A.CTGTAGCCAGAA---GGCT---TATCC--C-..TTT--A 1050
H2A.GH.x.GH1.M30895       -G-C..---------A--------------------G---A-------C-GAG-TAC-A-AGG.-------T-------.---T-A------------..-G--A-..---------CTACACCG---A.CCGTAGCCAG.A---GGCTA--TATCC----.TTTT--A  523
H2A.IN.NNVA.NEW           ---------------A--------------------G---A-------C-AAG-TAC-A-AGG.-------T-------.---T-A------------..-G--A-..-------A-CT.CACC----A.CTGTAGCCAG.A---GGCT---TATCC--C-..TTT--A 1075
H2A.IN.95.CRIK            ---------------A--------------------G-----------C-GAG-TAC-A--G..-------T-------.---TC.------------..-G--A-..-------A-CTACACC----A.CTGTAGCCAG.A--GGGCT--CTATCC--C-..TTT--A  826
H2A.GM.x.MCR35.AY509260   ---------------A--------------------G-------A--C-AAG-TAC-A-AGG.-------T-------.---TA.------------..-G--A-..-------A-CTACACC----AATTGTAGCCAG.A--GGGCT---TATCC--C-..TTT--A  527
H2A.GM.87.D194.J04542     ---------------A--------------------G---A-------C-GAG-TAC-G-AG..-------T-------.---T-A------------..-G--A-..-------A-CTACACCG---A.CTGTAGCCAG.A---GGCT---TATCC--C-..TTT--A  526
H2A.GM.x.ISY.J04498       ---------------A--------------------G-------A--C-AAG-TAC-A-AGG.-------T-------.---TC.------------..-G--A-..-------A-CTACACC----A.CTGTAGCCGG.A---GGCT---TATCC--C-..TTT--A  526
H2A.GM.x.MCN13.AY509259   ---------------A--------------------G-------A--C-AAG-TAC-A-AGG.-------T-------.---TA.------------..-G--A-..-------A-CTACACC----AATTGTAGCCAG.G--GGGCT---TATCC--C-..TTT--A  527
H2A.GW.86.FG.J03654       ---------------A-----GC-------------G-----------C-GAG-TAC-A-AG..-------T---.---.---TC.G-----------..-G--A-G.-------A-CTACACC----A.CTGTAGCCAA.AG-GGGCT---TATCC--C-..TTT--A  526
H2A.GW.87.CAM2CG.D00835   ---------------A------------------AG--------A--C-GAG-TAC-A-AGG.-------T-------.---TC.------------..-G--A-..-------A-CTACACC----A.CTGTAGCCTG.A--GGGCT---TATCC--C-..TTT--A 1091
H2A.GW.x.ALI.AF082339     ---------------A--------------------G-------A--C-AAG-TAC-A-AGGC-------T-------.---T-.------------..-G--A-..-------A-CTACACC----T..TGTAGCCG..A--GGGCT---TATCC--C-..TTT--A 1075
H2A.GW.x.MDS.Z48731       ---------------A--------------------G---A-------C-AAG-TAC-A-AGG.-------T-------.---T-A------------..-G--A-..-------A-CTACACCG---A.CTGTAGCCAA.A--GGGCT---TATCC--C-..TTT--A  526
H2A.SN.85.ROD.M15390      ---------------A--------------------G-----------C-GAG-TAC-A-AGG.-A-----T-------.---GA-------------..-G--A-..-------A-CTACACC----A.CTGTAGCCG..A--GGGCT--CTATCC--C-..TTT--A  525
H2AB.CI.90.7312A.L36874   --------------T-------------A------G---------A-C...---------G..-A--------------.---GA------------GGAG-AA..-------A.CT-CA--TTTT-..................-A-------....C--CT-TA--G-A 1064
H2B.CI.88.UC1.L07625      --------------T-------------A------G---------A--C-..---------G..-A--------------.---GAA------------.G-AG-A-A.---G---A.CT-CA--TTCT-..................-AT------....C-GCT-TAAG-AA 1062
H2B.CI.x.EHO.U27200       --------------T--------------------G---AA---A--C-..---------G..-A--------------.---GA-------TC-G-.GAG-AA-A.-------A.CT-CA--TTTGC................A-TTC--....C--CT-T.AGGAC 1058
H2B.GH.86.D205.X61240     --------------T-------------A------G---A---A--C-..---------G..-A--------------.---GA-------TC-G-.GAG-AA..-------A.CT-CA--TTTT-..................CAT------....C--CT-TAA-GA- 1057
H2B.JP.01.KR020.AB100245  --------------T-------------A------G---A---A--C-..---------G..-A--------------.---GA------------GGAG-AA..---G---A.CT-CA--TTTTA..................-AT------....C--CT-TAAG-AA  203
H2AB.CM.03CM_510_03.NEW   --------------T-------------A------G---------A--CT..---------G..-A--------------.---GAA-----------GGAG-A....--GT-AGTACT-CA--TTTT-..................CAT------....C-GCT-TAA-GA-  197
H2G.CI.92.ABT96.AF208027  -------------T----------------T-----G--------------..-----------..-A---------A---.----.T--------------..-G---GCA---------C--A---TTT-......TAGCCAA---GG-----GA.A-CC-A-CTTG-AT  444
H2U.FR.96.12034.AY530889  ---------------A--------------------G--------A--CT..-------AG..-A--------------.---T-A..----T-----.------.-------------A----T--.........AACC---GG-T-----.GCCTCTG--A----  561
SMM.SL.92.SL92B.AF334679  --------------T-----G---------TTGA-CCGGTGCCA-T---CTGAGC-GGAGTAC..A--TAA---AG--CT.-T-G-AC--TAAGG---A..AAA.....---GT-AAT-T-CT-CCT-C-..................................G-A  471
MAC.US.x.MAC239           ..................................-------------..------------..----------------.-----------------..------..--------------------------............-----------.-------------  108
MAC.US.x.251              -----------------------------------------------..------------..----------------.-----------------..------..--------------------------............-----------.------------- 1030
MAC.US.x.251_32H          -----------------------------------------------..------------..----------------.-----------------..------..------------------------AG............-----------.------------- 1032
MAC.US.x.251              -----------------------------------------------..------------..----------------.-------G---------..------..--------------------------............A--TAGCTGTCT-.G------------ 1019
MAC.US.x.17EC1.AY033233   -----------------------------------------------..------------..----------------.-----------------..------..--------------------------............-----------.------------- 1031
MAC.US.x.17EFR.AY033146   -----------------------------------------------..------------..----------------.-----------------..------..--------------------------............-----------.------------- 1031
MAC.US.x.1937.AY611495    .........----------------------------A---------..------------..----------------.-----------------..------..--------------------------............-----------.-------------  137
MAC.US.x.2065.AY611493    .........----------------------------A---------..------------..----------------.-----------------..------..--------------------------............-----------.-------------  137
MAC.US.x.MM142.Y00277     ------------------------------------G----------..------------..---------------.-GAGA-----------TG.------..-----------------------G-............-----------.-------------  529
MAC.US.x.SMM142B.BD131285------------------------------------G----------..------------..---------------.-GAGA-----------TG.------..-----------------------G-............-----------.-------------  529
MNE.US.82.MNE             -----------------------G-------------G---------..------------..----------------.-----.------------..------..--------------------------............G----------.-------------  511
MNE.US.x.MNE027.U79412    -----------------------G-------------G---------..------------..----------------.-----.------------..------..--------------------------............-----------.-------------  511
SMM.US.x.F236             ---------------A------------A-------G-----------C-..--------G..----------------.---TC-G-----------..------..---------------G---GTC.............-T-GGA---AG-.-CCC-A-TTTTGA- 1027
SMM.US.x.H9.M80194        ---------------A--------------------------------C-..--------G..----------------.---TC-G-----------..------..---------------A--G-GTC.............-T-GGA---AG-.-CCC-A-TTTTGA-  513
SMM.US.x.PBJ14            ---------------A--------------------------------C-..--------G..---A------------.---TC-G-----------..------..---------------A---G-GTC............-T-GGA---AG-.-CCC-A-TTTTGA-  807
SMM.US.x.PBJA.M31325      ---------------A--------------------------------C-..--------G..----------------.---TC-G-----------..------..---------------A--G-GTC.............-T-GGA---AG-.-CCC-A-TTTTGA-  806
SMM.US.x.PBJ              ---------------A--------------------------------C-..--------G..----------------.---TC-G-----------..------..---------------A--G-GTC.............-T-GGA---AG-.-CCC-A-TTTTGA-  513
SMM.US.x.PBJ              ---------------A--------------------------------C-..--------G..----------------.---TC-G-----------..------..---------------A---G-GTC............-T-GGA---AG-.-CCC-A-TTTTGA-  999
SMM.US.x.PGM53.AF077017   ---------------A--------------------------------C-..--------G..----------------.---TC-G-----------..-----G..---------------AG--G-GTC............-T-GGA---AG-.-CCC-A-TTTTGA-  957
SMM.US.x.SME543.U72748    ---------------A------------A-------G-----------C-..--------G..------------A-G---TC-G-----------..------..---------------A---GTC.............-T-GGA---AG-.-CCC-A-TTTTGA- 1043
STM.US.x.STM.M83293       ---------------A-------------------G------------CG.........AG..----------------.---GC------A.-----..-------..A--G---------AGT-GCC..............-G-GG----AT-.G--C-A-TCTTGA-  687
MAC.US.x.80035.AY611486   .........----------------------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  137
MAC.US.x.81035.AY599200   ........-----------------------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  138
MAC.US.x.85013.AY611490   ..................................................................................-----------------..------..----------------..............-----------.-------------   71
MAC.US.x.92050.AY603959   ...............----------------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  131
MAC.US.x.92077.AY599201   ........-----------------------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  138
MAC.US.x.93057.AY611492   ..........---------------------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  136
MAC.US.x.93062.AY607704   ........................-------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  123
MAC.US.x.95058.AY611494   ......................---------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  125
MAC.US.x.95086.AY607703   ............................................................................................................................................................................    0
MAC.US.x.95112.AY588946   ....................-----------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  127
MAC.US.x.96016.AY607701   ........-----------------------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  138
MAC.US.x.96020.AY611488   .................--------------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  130
MAC.US.x.96072.AY611491   ....--------------------------------G---A--------..------------..----------------.-----------------..------..----------------..............-----------.-------------  142
MAC.US.x.96081.AY597209   ......-------------------------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  140
MAC.US.x.96093.AY611489   ........-----------------------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  138
MAC.US.x.96114.AY588945   .............------------------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  133
MAC.US.x.96123.AY611487   ..............................................................................................................-...............-----------.-------------   27
MAC.US.x.96135.AY607702   ........................-------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  123
MAC.US.x.97009.AY599199   ......................................................................................................................................................------    6
MAC.US.x.97074.AY599198   ......................................................................................................................................................------    6
MAC.US.x.r80025.AY576480  ............................................................................................................................................................................    0
MAC.US.x.r90131.AY576481  .........----------------------------------------..------------..----------------.-----------------..------..----------------..............-----------.-------------  137
```

```
                                        Gag and Gag-Pol start
MAC.US.x.239               GGTAATAAGATA.GAGTGGGAGATGGGCGTGAGAAACTCCGTCTTGTCAGGGAAGAAAGCAGATGAATTAGAAAAAATTAGGCTACGACCCAACGGAAAGAAAAAGTACATGTTGAAGCATGTAGTATGGGCAGCAAATGAATTAGATAGATTTGGATTAGCAGAAAGCC 1200
Gag                                                 _M__G__V__R__N__S__V__L__S__G__K__K__A__D__E__L__E__K__I__R__L__R__P__N__G__K__K__K__Y__M__L__K__H__V__V__W__A__A__N__E__L__D__R__F__G__L__A__E__S_
H2A.CI.88.UC2.U38293       CAGGTAG-AGA..TT--------------C-------------C--AG------A--G-----C------------------T----G---GG-----------------G---A--A---A-T--G--------G--C-----G--C-----CA------------A-- 1250
H2A.DE.x.BEN.M30502        CAGGTAG-AGA..TT--------------C----------------AG------A--------C------------G-----T----G---GG-----------------G---A--A---A-T--G--------G--------G----A---C-----G-----G---- 1250
H2A.DE.x.PEI2.U22047       CAGGTAG-AGA..TT--------------C-----G----------AG------A----T----------------------T----G---GG---G--------A--T-GAC-A------A-T--G--------G--------G-GC-A---C-----G---------- 1218
H2A.GH.x.GH1.M30895        CAGGTAG-AGA..TT--------------C----------------AG------A--------------------G------T----G----G-----------------G-C-A--A---A-T--G--------G--C-----G--C-A---C-----G-----G---- 691
H2A.IN.NNVA.NEW            CAGGTAG-AGA-.TT--------------C--AG---------C--AG------A-----T-------------G-------T----G---GG------------A----GAG-A--A-----T--G--------G--C-----G--C-----C------A--------- 1244
H2A.IN.95.CRIK             CAGGTAG-AGA-.TT--------------C--AG-G-------C--AG------A-----T-------------G-------T----G---GG------------A----GAC-A--A-----T--G--------G--C-----G--C-A---C------A--------- 995
H2A.GM.x.MCR35.AY509260    CAGGTAG-AGA..TT--------------C----------------AA------A--------C----------C-------T----G---GG------------A----G-C-A------A-T--G--------G--------G--C-----C-----------G---- 695
H2A.GM.87.D194.J04542      CAGGTAG-AGA..TT--------------C----------------AG------A--------C------------G-----T----G----------------GA----G---A--A-----T--G--------G-----------C-----C-----G-----G---- 694
H2A.GM.x.ISY.J04498        CAGGTAG-AGA..TT--------------C--A-------------AG------A--G------------------------T----G---GG---G--------A----GA--A--A---A-T--G--------G--------G--C-----C------A----G---- 694
H2A.GM.x.MCN13.AY509259    CAGGTAG-AGA..TT--------------C----------------AA------A--------C----------C-------T----G---GG------------A----G-C-A------A-T--G--------G--------G--C-----C-----------G---- 695
H2A.GW.86.FG.J03654        CAGGTAG-AGA..TT--------------C----------------AG------A---------------------------T----G---GG-----------------A---A--A---A-T--G--------G--------G--C-----C-----------G---- 694
H2A.GW.87.CAM2CG.D00835    CAGGTAG-AGA..TT--------------C----------------AG------A--------C------------G-----T----G---GG------------A----A---A--A---A-T--G--------G--------G--C-----C-----G-----G---- 1259
H2A.GW.x.ALI.AF082339      CAGGTAG-AGA..TT--------------C----------------AG-----GA--------C----------G-------T----G---GG------------A--TCA-C-A--A---A-T--G--------G--------G--C-----C-----G---------- 1243
H2A.GW.x.MDS.Z48731        CAGGTAG-AGA..TT--------------C-----G----------AG------A--------C------------G-----T----G---GG------------A----G---A--A---A-T--G--------G--------G--C-----C-----------G---- 694
H2A.SN.85.ROD.M15390       CAGGTAG-AGA..TT--------------C----------------AG------A-------------------G---C---T----G---GG-----------------G-C-A--A---A-T--G--------G---A----G--C-----C-----------G---- 693
H2AB.CI.90.7312A.L36874    .-AGTAGG-CAT.A---------------C----GG-------C----------A--GA-----------------G-----T----G---GG---------------------A--A------A----------------------C-----C-----G-------A-- 1232
H2B.CI.88.UC1.L07625       --AGTAGG-CAT.A---------------C-----G----------------------A-----------------G-----T----G---GG-----------G---TTGTC-A--A---A--A---------TG--C-----G--C-----------------G---- 1231
H2B.CI.x.EHO.U27200        A-GGTACG--AT.A---------------C----GG---------------------GA----------------GG-----T----G---GG-----------GA------C-A--A---A------------TG-----------G-----------G-----G---- 1227
H2B.GH.86.D205.X61240      --AGTAGG-CAT.A---------------C----GG-------C--------------A-----------------G-----T----G---GG---------------T---C-A--A----------------TG--C--------C-----------------G---- 1226
H2B.JP.01.KR020.AB100245   --AGTAGG-CAT.A---------------C----GG----------------------A-----------------G-----T----G---GG------------A--TTGTC-A------A------------TG--C-----G--C----A------------G---- 372
H2AB.CM.03CM_510_03.NEW    .-AGTAGG-CAT.A---------------C-----G-------C--------------A-----------------G-----T----G---GG-----------G---T--------A------A---------TG--C--------------------------G-C-- 365
H2G.CI.92.ABT96.AF208027   A--C-GGTAGA-.-T--------------C-----G-------C----------A-----T---------------G-----T----G---GG------------A--------A--A-----T--G--------G--C-----G--C-----C-----------G---- 613
H2U.FR.96.12034.AY530889   A-AGG--GT---.TT--------------C---------.-------------------T----------------G-----T----G-----------------A--------A--A-----T--G--------G--C-----G--C---------------------- 729
SMM.SL.92.SL92B.AF334679   TAGGTAG-AG--.A--C------------C----GG------------------A---------------------G-----T----G---GG-----G-------------C-T------A-TA----------C-GA--------C----------CG---------- 640
MAC.US.x.MAC239            ------------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 277
MAC.US.x.251               -A----------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 1199
MAC.US.x.251_32H           -A----------.----------------C-------------------------------------------------------------GG----------------------------------------------------------------------------- 1201
MAC.US.x.251               -A----------.----------------C-------------------------------------------------------------GG----------------------------------------------------------------------------- 1188
MAC.US.x.17EC1.AY033233    ------------.------------------------------------------------------------------------------------------------------------------------------------------------------------- 1200
MAC.US.x.17EFR.AY033146    ------------.------------------------------------------------------------------------------------------------------------------------------------------------------------- 1200
MAC.US.x.1937.AY611495     ------------.----------------C---------------------------------------------------------------------------------------------------------------------------------R---------- 306
MAC.US.x.2065.AY611493     ------------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 306
MAC.US.x.MM142.Y00277      -A----------.----------------C---------------------------------------------------A---------GG----------------------------------------------------------------------------- 698
MAC.US.x.SMM142B.BD131285  -A----------.----------------C---------------------------------------------------A---------GG----------------------------------------------------------------------------- 698
MNE.US.82.MNE              -A----------.----------------C-------------------------------------------------------------GG---G------------------------------------------------------------------------- 680
MNE.US.x.MNE027.U79412     -A----------.----------------C-------------------------------------------------------------GG---G--------------------------------------------------------------------C---- 680
SMM.US.x.F236              -AA-GAGTAGG-.----------------C-------------------------G--------------------G-----T----G--------------------T-----------------------------------G--C---------------------- 1196
SMM.US.x.H9.M80194         AAA-GAGTAGG-.--------------T------------------------------------------------------T----G---GG-------------------C--------A----------------------G--C---------------------- 682
SMM.US.x.PBJ14             AAA-GAGTAGG-.----------------C-------------------------------------C--------------T----G---GG-----------G-----A-C--------A----------------------G--C---------------------- 976
SMM.US.x.PBJA.M31325       AAA-GAGTAGG-.----------------C----------------------------------------------------T----G---GG-----------G----CA-C--------A----------------------G--C---------------------- 975
SMM.US.x.PBJ               AAA-GAGTAGG-.--------------T------------------------------------------------------T----G---GG-------------------C--------A----------------------G--C---------------------- 682
SMM.US.x.PBJ               AAA-GAGTAGG-.----------------C----------------------------------------------------T----G---GG-----------G-----A-C--------A----------------------G--C---------------------- 1168
SMM.US.x.PGM53.AF077017    -AA-GAGTAGG-.----------------C----------------------------------------------------T----G--------------------------------------------------C-----G--C---------------------- 1126
SMM.US.x.SME543.U72748     -AA-GAGTAGG-.----------------C----------------------------------------------------T----G--------------------T-----------------------------------G------------------------- 1212
STM.US.x.STM.M83293        --A-GAGGT-G-.AG--------------C-----G------------------A---------------------G-----T----G---GG---G--------------------A--------G-----------C-----G--C-----------G---------- 856
MAC.US.x.80035.AY611486    ------------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 306
MAC.US.x.81035.AY599200    ------------.------------------------------------------------------------------------------------------------------------------------------------------------------------- 307
MAC.US.x.85013.AY611490    ------------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 240
MAC.US.x.92050.AY603959    ------------.----------------Y--------------------------------------------------------------------------G----------------------------------------------------------------- 300
MAC.US.x.92077.AY599201    ------------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 307
MAC.US.x.93057.AY611492    ------------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 305
MAC.US.x.93062.AY607704    ------------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 292
MAC.US.x.95058.AY611494    ------------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 294
MAC.US.x.95086.AY607703    ......------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 163
MAC.US.x.95112.AY588946    ------------.------------------------------------------------------------------------------------------------------------------------------------------------------------- 296
MAC.US.x.96016.AY607701    ------------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 307
MAC.US.x.96020.AY611488    ------------.------------------------------------------------------------------------------------------------------------------------------------------------------------- 299
MAC.US.x.96072.AY611491    ------------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 311
MAC.US.x.96081.AY597209    ------------.------------------------------------------------------------------------------------------------------------------------------------------------------------- 309
MAC.US.x.96093.AY611489    ------------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 307
MAC.US.x.96114.AY588945    ------------.------------------------------------------------------------------------------------------------------------------------------------------------------------- 302
MAC.US.x.96123.AY611487    ------------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 196
MAC.US.x.96135.AY607702    ------------.----------------Y-------------------------------------------------------------------------------------------------------------------------------------------- 292
MAC.US.x.97009.AY599199    ------------.----------------C----------------------------A--------------------------------------------------------------------------------------------------------------- 175
MAC.US.x.97074.AY599198    ------------.------------------------------------------------------------------------------------------------------------------------------------------------------------- 175
MAC.US.x.r80025.AY576480   ......------.----------------C-------------------------------------------------------------------------------------------------------------------------------------------- 163
MAC.US.x.r90131.AY576481   ------------.------------------------------------------------------------------------------------------------------------------------------------------------------------- 306
```

```
MAC.US.x.239                TGTTGGAGAACAAAGAAGGATGTCAAAAAATACTTTCGGTCTTAGCTCCATTAGTGCCAACAGGCTCAGAAAATTTAAAAAGCCTTTATAATACTGTCTGCGTCATCTGGTGCATTCACGCAGAAGAGAAAGTGAAACACACTGAGGAAGCAAAACAGATAGTGCAGAGA 1370
Gag                         L__L__E__N__K__E__G__C__Q__K__I__L__S__V__L__A__P__L__V__P__T__G__S__E__N__L__K__S__L__Y__N__T__V__C__V__I__W__C__I__H__A__E__E__K__V__K__H__T__E__E__A__K__Q__I__V__Q__R_
H2A.CI.88.UC2.U38293        --------TCA--------T--C-----G--T--CA-A--T----A---------A--------G----------------------T------C-----------T------T-G---------------------G-T-----A--G------A-AC----A------ 1420
H2A.DE.x.BEN.M30502         --------TCA--------T--C-----G--T--CAGA--T----A---------A--------G----------------------T------C-----------T------T-G---------------------G-T--------------GA-AC---CA------ 1420
H2A.DE.x.PEI2.U22047        --------TCA--------T--C--------TA--A-A--T----A------------------A----------------------T------------------T------T-G---------------------G-T-----A-G---------A-----A------ 1388
H2A.GH.x.GH1.M30895         --------TCA--------T--C-----G--T--CA-A--T----A---------A--------G----------------------T---C--A-----------T------T-G---------------------G-T-----A---------A-AC----A------ 861
H2A.IN.NNVA.NEW             -------ATCA-----G--T--C--------T---A-A--T----A------------G-----T----------------------T------------------T--------A---------------------G-T-----A---------A--T----A------ 1414
H2A.IN.95.CRIK              --------TCA-----G--T--C--------T---A-A--T----A------------G-----T----------------------T------------------T--------A---------------------G-T-----A---------A-------A------ 1165
H2A.GM.x.MCR35.AY509260     --------TCA--------T--C----G---T---A-A--T----A---------A--G-----T----------------------T------------------T--------A---------------------G-T-----A-G---------A-----A------ 865
H2A.GM.87.D194.J04542       -------ATCA--------T--C-----G--T---AAA--T----AA--------A--------G----------------------T------C-----------T------T-G---------------------G-T-----A--------GA-AC---CA------ 864
H2A.GM.x.ISY.J04498         --------TCA--------T--C--------TA----A--T----AA--------A--------G------------------------------AC---------T------T-G---------------------G-T-----A---------AGA-----AGG---- 864
H2A.GM.x.MCN13.AY509259     --------TCA--------T--C----G---T---A-A--T----G---------A--G-----T----------------------T------------------T--------A---------------------G-T-----A-G---------A-----A------ 865
H2A.GW.86.FG.J03654         --------TCA--------T--C--------T---A-A--T----A---------A--G-----T----------------------T---------------------------A---------------------G-T-----A-G---------A-----A------ 864
H2A.GW.87.CAM2CG.D00835     -------ATCA--------T--C----G---T---AAA--T----A---------A--G-----T----------------------T------------------A--------A---------------------G-T-----A---------AGA----CA-T---- 1429
H2A.GW.x.ALI.AF082339       --------TCA--------T--C----G---T---AAA--T----AA-----------------A----------------------T---------------AG-T------G-G---------------------G-T-----A-G---------A---A-A------ 1413
H2A.GW.x.MDS.Z48731         -------ATCA-----G--T--C--------T---A-A-------A---------A--G-----T----------------------T------------------T-TT-----A--------G------------G-T---------------A--T-C-CA-G---- 864
H2A.SN.85.ROD.M15390        --------TCA-----G--T-----------T---A-A--T----A----A-G--A--G-----T-----------------T----T------------------T--------A---------------------G-T-----A-G---------A-------G---- 863
H2AB.CI.90.7312A.L36874     -------ATCA--------------G-----T---A-----------------A----------G----------------------T------------------T------T-A---------A-----------G-T--C------------A--C---CA-----C 1402
H2B.CI.88.UC1.L07625        -------ATCA--------------C-----T---A-------------------A--------A----------------------T------------------T-AC---T-G---------------------G-T---------------A------CA------ 1401
H2B.CI.x.EHO.U27200         G-----GATCA-----------C-GG-----T-GGAAA-------GA--------A--------T-----------------T-----------C-------------TT---C-G---------------------G-T---------------A------CA---C-- 1397
H2B.GH.86.D205.X61240       -------ATCA--------------------T---AAA-----------------A--------G----------------------T-----TC-----------T-TT---C-G---------------------G-T--A------------A------CA------ 1396
H2B.JP.01.KR020.AB100245    --------TCA--------------C-----T---A---------AA--------A--------A----------------------T------------------T-AC--TC-G----------T----------G-T--------T--T---A------CA-----C 542
H2AB.CM.03CM_510_03.NEW     -------ATCA-----------C--------T---AAA-----------------A--------A----------------------T------------------T-TC---C-G---------------------G-T--A-----G------A------CA---C-- 535
H2G.CI.92.ABT96.AF208027    -------ATCA--------T--C--------T-----A--A--------------A--------T--------------G--------------------------T-----TT-G-----C---C-------A-----T-----A-----------AG----------- 783
H2U.FR.96.12034.AY530889    --------GC---------------------TT-A-----G--------------T--------T--------------G--------------------------T-----TT-G-----------------------T--------------G--------A------ 899
SMM.SL.92.SL92B.AF334679    -------A-G-----------------G----T-AG-A--AC-----------A------------G----------------T-G-T--G------------AG-T-----TC-G---------AT----------G----A--------T--GA-A-C---A-----C 810
MAC.US.x.MAC239             ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 447
MAC.US.x.251                -------------------------------------------------------------------------------------------------------------------------------------------------------------------A------ 1369
MAC.US.x.251_32H            ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1371
MAC.US.x.251                ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1358
MAC.US.x.17EC1.AY033233     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1370
MAC.US.x.17EFR.AY033146     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1370
MAC.US.x.1937.AY611495      -------R---------------------G-----------------------A-------------------------------------------------------------------------------------------------------------------- 476
MAC.US.x.2065.AY611493      -----------------------------------------------------M-----------------------------------------R-------------------------------------------------------------------------- 476
MAC.US.x.MM142.Y00277       ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 868
MAC.US.x.SMM142B.BD131285-  ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 868
MNE.US.82.MNE               -------------------------------------------------------------------------------G------------------------------------------------------------------------------------------ 850
MNE.US.x.MNE027.U79412      ---------------------------------------C---------------------------------------G------------------------------------------------------------------------------------------ 850
SMM.US.x.F236               -------T-----------T-----------T-----A--T--------------T--G-----T--------C-----G--------------------------T--------------------------------T-----------------A------------ 1366
SMM.US.x.H9.M80194          -------------------T-----------T-----A--T--------------T--G-----T--------------G-----------C------C-----C-T--------------------------------T-----A-----------A------------ 852
SMM.US.x.PBJ14              -------------------T-----------T-----A--T--------------T--G-----T--------------G-----------C------------C-T--------------------------------T-----A-----------A------------ 1146
SMM.US.x.PBJA.M31325        -------------------T-----------T-----A--T--------------T--G-----T--------------G-----------C------------C-T--------------------------------------A-----------A------------ 1145
SMM.US.x.PBJ                -------------------T-----------T-----A--T--------------T--G-----T--------------G-----------C------C-----C-T--------------------------------T-----A-----------A------------ 852
SMM.US.x.PBJ                -------------------T-----------T-----A--T--------------T--G-----T--------------G-----------C------------C-T--------------------------------T-----A-----------A------------ 1338
SMM.US.x.PGM53.AF077017     ---------C---------T-----------TA----A--T--------------T--GG----T--------------G--------------------------T--------------------------------T-----------------A------------ 1296
SMM.US.x.SME543.U72748      -------T-----------T-----------T--------T--------------T--G-----T--------------G--------------------------T--------------------------------T-----------------A------------ 1382
STM.US.x.STM.M83293         --------TCA--------G-----------TA-AA---------AA-----------------T--------------G--T----T------------------T-----T--C-----------------------T--------------G--AG-G--AA----- 1026
MAC.US.x.80035.AY611486     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 476
MAC.US.x.81035.AY599200     -----------------------------------------------------R-------------------------------------------------------------------------------------------------------------------- 477
MAC.US.x.85013.AY611490     -----------------------------------------------------R-------------------------------------------------------------------------------------------------------------------- 410
MAC.US.x.92050.AY603959     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 470
MAC.US.x.92077.AY599201     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 477
MAC.US.x.93057.AY611492     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 475
MAC.US.x.93062.AY607704     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 462
MAC.US.x.95058.AY611494     -----------------------------------------------------R-------------------------------------------------------------------------------------------------------------------- 464
MAC.US.x.95086.AY607703     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 333
MAC.US.x.95112.AY588946     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 466
MAC.US.x.96016.AY607701     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 477
MAC.US.x.96020.AY611488     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 469
MAC.US.x.96072.AY611491     -----------------------------------------------------T-------------------------------------------------------------------------------------------------------------------- 481
MAC.US.x.96081.AY597209     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 479
MAC.US.x.96093.AY611489     -----------------------------------------------------A-------------------------------------------------------------------------------------------------------------------- 477
MAC.US.x.96114.AY588945     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 472
MAC.US.x.96123.AY611487     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 366
MAC.US.x.96135.AY607702     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 462
MAC.US.x.97009.AY599199     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 345
MAC.US.x.97074.AY599198     -------------------------------------------------------------------------------------------------T------------------------------------------------------------------------ 345
MAC.US.x.r80025.AY576480    -----------------------------------------------------D-------------------------------------------------------------------------------------------------------------------- 333
MAC.US.x.r90131.AY576481    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 476
```

Gag p17 Matrix end Gag p24 Capsid start

```
MAC.US.x.239               CACCTAGTGGTGGAAACAGGAACAACAGAAACTATGCCAAAAACAAGTAGACCAACAGCACCATCTAGCGGCAGAGGAGGAAATTACCCAGTACAACAA...ATAGGTGGTAACTATGTCCACCTGCCATTAAGCCCGAGAACATTAAATGCCTGGGTAAAATTGATAGA 1537
Gag                        _H__L__V__V__E__T__G__T__T__E__T__M__P__K__T__S__R__P__T__A__P__S__S__G__R__G__G__N__Y__P__V__Q__Q__.__I__G__G__N__Y__V__H__L__P__L__S__P__R__T__L__N__A__W__V__K__L__I__E
H2A.CI.88.UC2.U38293       --T-------CA----------TG---G-AA--------T--------------------C----T--G-A---------C-T---C---------G...GC------C-----------TG-----C-G----CC----TC-------T-----------AG---- 1587
H2A.DE.x.BEN.M30502        --T-------CA-----T-----TG---G-AA--------T--------------------C----T--G-A-A-------C-----C--G------...GCG-----C-----------TG-----C-G----CC----TC-------A-----------AG-G-- 1587
H2A.DE.x.PEI2.U22047       --T-------CA----------TG---C-AA-------GC-----------G--------C------G----G-----------C--G------...-----C---C------TC---TG----G--G----CC----CC-------T--------G--AG-G-- 1555
H2A.GH.x.GH1.M30895        --T----G--CA----------TG-----AA-------GT--------------------C------G-----A----C-T---T-------GACAGGC-----C------A-----G-----C-G----CC----TC-------T-----------AG---- 1031
H2A.IN.NNVA.NEW            --T-------CA----------TG---G-AA--------T-------------------CC-----AGG---G-----C-TT--C-------GC...-----C--C-----CAC----G----GC-G-----C-----TC------TT-----------AG---- 1581
H2A.IN.95.CRIK             --T-------CA-----T-----TG---G-AA--------T--------------------C-------GG----------C-T---C--------C...-----C--C--T---AC---TG----GC-G--T--C-----TC---------------------AG---- 1332
H2A.GM.x.MCR35.AY509260    --T----C--CA----------TG---G-AA--------T--------------------C-------G-A---------C-T---C----------...G----C--C--T---AC---TG----GC-G--T--TC----CC-------T-----------AG---- 1032
H2A.GM.87.D194.J04542      --T-------CA----------TG---G-AA--------T-T------------------C----T--G-A---------C-T---C--G-----G...GC---C--C------A----TG----GC-G----CC----TC-------T-----------AG---- 1031
H2A.GM.x.ISY.J04498        --T-------CA-------A---TG---G-AA--------T-T------------------C-------G-A--G----C-T---C--G------...-----C--C-----------T-----GC-G--T--CC----CC-------T--------G--AG---- 1031
H2A.GM.x.MCN13.AY509259    --T----C--CA----------TG---G-AA--------T--------------------C-------G-AG--------C-T---C----------...G----C--C--T---AC---TG----GC-G--T--TC----CC-------T-----------AG---- 1032
H2A.GW.86.FG.J03654        --T-------CA----------TG---G-AA--------T--------------------C-------G-A.......--C-T---C--G--G----...G---C---C------AC---TA-A---C-G--T--CG----TC-------T-----------AG-G-- 1025
H2A.GW.87.CAM2CG.D00835    --T----C--CA----------TG---G-AA------G-T--------------------C----T--G-AG---------------C--G---AGC...------C--C------ACT--TG----GC----T--CC----TC-------T-----------AG---- 1596
H2A.GW.x.ALI.AF082339      --T----C--CA----T--A----G---G-AA-------GC--------------------C----T-AACAG--G-----C-T---C----------...G---CC--C------CAC---TG----GC-G-----CC----C---G----T-----------AG---- 1580
H2A.GW.x.MDS.Z48731        --T-------CA----------TG---G-AA--------T--------------------C-------G-----------C-----C--G--GA-T...-----C--C------AC---TG----G--G--T--CC----TC-------T-----------AG---- 1031
H2A.SN.85.ROD.M15390       --T-------CA----------TG---G-AA-------GC----------------------------AG-AG----------------G-----T...G----C--C-----CAC---TA-A--GC-G--T--CC----CC---------------------AG---- 1030
H2AB.CI.90.7312A.L36874    --T--------A--C--T.........----AA-CA--T-CTG--GAC-A----G----GA-------T------......----C------G-------...G---C---C--T---------G-------------CC----CC-------------------AG-G-- 1554
H2B.CI.88.UC1.L07625       --T----C--C---C---.........--G-AA------GCT--------------------C--------......---------------G--G--G...----C---C-----C------A----GC-------TC----C-----------------G--AG---- 1553
H2B.CI.x.EHO.U27200        --T----CA-C---C---.........----AA------GCT-TG----A-----GTAA----A-----C--CT-......GC---T-----G--G---...----C---C--T---TC---------GC-------CC----C--------------------AG---- 1549
H2B.GH.86.D205.X61240      --T----C--C---C---.........----AA------GCT----A--A--------------C--------......---------T-----G--G---...C-G-C---C-----C-----------GC-------CC----C-------T--------G--AG---- 1548
H2B.JP.01.KR020.AB100245   --T----C--C-A-C--G.........----AAGC----GCT-------A--------------C----T------......------T-----G--G--G...G-G-CC-----T--CAC---------GC-------CC----C---------------------AG---- 694
H2AB.CM.03CM_510_03.NEW    -------C-...--C.........G------AA---A--GCT----A--A--------------C--------......---------T-----G--G---...----C---C------------G----GC-------CC----C-------T-----------AG---- 684
H2G.CI.92.ABT96.AF208027   --T--------A-----T-----T-------AAG----TGCC-----C-------T-------C----T--------------C--------G--G---...G----C--C--------T--T-----------T--TC----T--------A-------------G---- 950
H2U.FR.96.12034.AY530889   --T----C---------T-A---TG------AA------GCC-----------------YT---C----T---G-CA-------------G--G--G--G...---------C--T----C---------C------A--TC----T--T------C-C---------------C--G---- 1066
SMM.SL.92.SL92B.AF334679   --T--------A----GT-----TG------AAT-----GCTCA---C-----------T---C----T--A......-----C--T----G--G---...G----AAA---T-----G---ACT---C-GTC---AC-G---C-C-----G-----C---C--G---- 971
MAC.US.x.MAC239            ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 614
MAC.US.x.251               ---------------------G------T--------------C--------T----------------------------------------------...-------------------------------------------------------------------- 1536
MAC.US.x.251_32H           ---------------------G---------------------------------------------------------------------------...-------------------------------------------------------------------- 1538
MAC.US.x.251               ---------------------G---------------------------------------------------------------------------...-------------------------------------------------------------------- 1525
MAC.US.x.17EC1.AY033233    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 1537
MAC.US.x.17EFR.AY033146    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 1537
MAC.US.x.1937.AY611495     ---------------------------------------------------------------------------------------------------...-----------------------------------A---------------------------G---- 643
MAC.US.x.2065.AY611493     ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 643
MAC.US.x.MM142.Y00277      ---------A-----------G------------------------------------------T----------------------------------...---------------AC------A---------------------------------------A---- 1035
MAC.US.x.SMM142B.BD131285  ---------A-----------G------------------------------------------T----------------------------------...---------------AC------A---------------------------------------A---- 1035
MNE.US.82.MNE              ---------------------G----------------------------------------------T------------------------------...G--------------AC------A-------------------------------------------- 1017
MNE.US.x.MNE027.U79412     ---------------------G----------------------------------------------T---A-A------------------------...---------------AC------A-------------------------------------------- 1017
SMM.US.x.F236              --T------------------G----C-GA------GC------C--------------C---T-------------------G--G---...G-------C--T-----------A--C-----T--A-------------T-----------AG---- 1533
SMM.US.x.H9.M80194         --T------------------G----C-AA------GC-----------------------C---T-------------------G--G---...G-------C--T--AC------A--T-----T--A-------------T------------G---- 1019
SMM.US.x.PBJ14             --T------------------G----C-AA------GC-----------------G---C---T---A---------------G--G---...---------C--T--AC------A--T-----T--A-------------T----------------- 1313
SMM.US.x.PBJA.M31325       --T------------------G----C-AA------GC-----------------------C---T---A---------------G--G---...---------C--T--AC------A--T-----T--A-------------T----------------- 1312
SMM.US.x.PBJ               --T------------------G----C-AA------GC-----------------------C---T-------------------G--G---...G-------C--T--AC------A--T-----T--A-------------T------------G---- 1019
SMM.US.x.PBJ               --T------------------G----C-AA------GC-----------------------C---T---A---------------G--G---...---------C--T--AC------A--T-----T--A-------------T----------------- 1505
SMM.US.x.PGM53.AF077017    --T------------------G----C-AA-------T------C---------------C---T-------------------G--G---...---------C--T--CAC-----A--T-----T--A-------------T------------G---- 1463
SMM.US.x.SME543.U72748     --T------------------G----C-AA------GC------C---------------C---T-------------------G--G---...G-------C--T-----------A--C-----T--A-------------T------------G---- 1549
STM.US.x.STM.M83293        --T--T-----A--G--T-----TG--A-C-AA-----TGCC---------------------C----T----------------------------G---...G----C--C--T---------------A-----------A-----C--G-----A--------G--AG-G-- 1193
MAC.US.x.80035.AY611486    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 643
MAC.US.x.81035.AY599200    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 644
MAC.US.x.85013.AY611490    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 577
MAC.US.x.92050.AY603959    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 637
MAC.US.x.92077.AY599201    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 644
MAC.US.x.93057.AY611492    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 642
MAC.US.x.93062.AY607704    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 629
MAC.US.x.95058.AY611494    ---------------------------------------------------------------------------------------------------...G------------------------------------------------------------------- 631
MAC.US.x.95086.AY607703    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 500
MAC.US.x.95112.AY588946    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 633
MAC.US.x.96016.AY607701    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 644
MAC.US.x.96020.AY611488    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 636
MAC.US.x.96072.AY611491    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 648
MAC.US.x.96081.AY597209    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 646
MAC.US.x.96093.AY611489    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 644
MAC.US.x.96114.AY588945    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 639
MAC.US.x.96123.AY611487    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 533
MAC.US.x.96135.AY607702    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 629
MAC.US.x.97009.AY599199    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 512
MAC.US.x.97074.AY599198    ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 512
MAC.US.x.r80025.AY576480   ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 500
MAC.US.x.r90131.AY576481   ---------------------------------------------------------------------------------------------------...-------------------------------------------------------------------- 643
```

```
MAC.US.x.239                GGAAAAGAAATTTGGAGCAGAAGTAGTGCCAGGATTTCAGGCACTGTCAGAAGGTTGCACCCCCTATGACATTAATCAGATGTTAAATTGTGTGGGAGACCATCAAGCGGCTATGCAGATTATCAGAGATATTATAAACGAGGAGGCTGCAGATTGGGACTTGCAGCACC 1707
Gag                         __E__K__K__F__G__A__E__V__V__P__G__F__Q__A__L__S__E__G__C__T__P__Y__D__I__N__Q__M__L__N__C__V__G__D__H__Q__A__A__M__Q__I__I__R__D__I__I__N__E__E__A__A__D__W__D__L__Q__H__
H2A.CI.88.UC2.U38293        ---------G--C--G-----------------------------C--------C-----G--------T--C-----A---C-T-----------C--T--A-----A--C--------A--------A-----C--T--A--A--A--G--------TGC-------- 1757
H2A.DE.x.BEN.M30502         ---------G--C--G-----------------------------C--------C-----G--------T--------A---C-T-----------C--T--C-----A--------A--A--------G-----T--T--A--A--A-----C-----T-C-------- 1757
H2A.DE.x.PEI2.U22047        A--------G--C--G-----------------G-----------C--------C-----G--------T--------A---C-T-----------C-----------A--------A--A--------G-----T--T--A--A--A-----------TG----A---- 1725
H2A.GH.x.GH1.M30895         ---C-----G--C--G-----------------------A-----C--------C-----G--------T--C--C--A---C-T-----------C--T--C-----A--------A--A--------G-----C--T--C--A--A-----------TGCA------- 1201
H2A.IN.NNVA.NEW             A--G-----G--C--G--------------------------G--C--------C-----G--------T--------A---C-T-----------C-----------A--C-----A--A-----G--A-----T--T--A--A--A-----------TG----A--T- 1751
H2A.IN.95.CRIK              ---------G--C--G---------------A-------------C--------C-----G--------T--C-----A---C-T-----------C-----------A--C-----A--A-----G--G-----T--T--A--A--A-----------TGC---A--T- 1502
H2A.GM.x.MCR35.AY509260     ---------G--C--G-----G-----------------------C--------C-----G--------T--C--C--A---C-T-----------C-----------A--------A--A-----G--G--C--T--T--A--A--A-----------TG----A--T- 1202
H2A.GM.87.D194.J04542       ---------G--C--G-----------------------------C--------C-----G--------T--C-----A---C-T-----------C--T--C-----A--------A--A--------A-----T--T-----A--A-----------TGC-------- 1201
H2A.GM.x.ISY.J04498         ---G--A--G--C--G--------------G--------------C--------C-----G--------T--------A---C-T-----------C-----------A--G-----A--A--------A-----T--T--A--A--A-----C-----TG-A--A--T- 1201
H2A.GM.x.MCN13.AY509259     ---------G--C--G-----G-----------------------C--------C-----G--------T--C--C--A---C-T-----------C-----------A--------A--A-----G--G--CG-T--T--A--A--A-----------TG----A--T- 1202
H2A.GW.86.FG.J03654         ------------C--G-----------------------------C--------C-----G--------T------------C-T-----------C-----------A--------A--A-----G--A-----T--T-----A--A-----C-----TG-TGCA--T- 1195
H2A.GW.87.CAM2CG.D00835     ---------G--C--G-----------------------------C--------C-----G-----------------A---C-T-----------C-----------A--------A--A-----G--G-----T--T--A--A--A-----C-----TGCCA-T--T- 1766
H2A.GW.x.ALI.AF082339       A--------G--C--G-----------------------------C--------C-----G--------T--------A---C-T-----------C-----------A--C-----A--A-----G--G-----C--T--A--A--A-----C-----TG-TGCA--T- 1750
H2A.GW.x.MDS.Z48731         A--------G--C--G-----------------------------C--------C-----G--------T--------A---C-T--C--------C--T--------A--------A--A-----G--A--C--T--T--A--A--A-----------TGC------T- 1201
H2A.SN.85.ROD.M15390        ------A--G--C--G-----------------------------C--------C-----G--------T--C--C--A---C-T-----------C-----------A--C--------A-----G--G-----C--T-----A--A-----A-----TG----A--T- 1200
H2AB.CI.90.7312A.L36874     ------------A--G-----------------------------A-----G--A--T-----T-----T------------C----------A-----A-----G--A--C-----AG-C--T--G--A--A--C--T--------G-----C------CA-------- 1724
H2B.CI.88.UC1.L07625        A--------G--C--G-----------A--------C--------A--------A-----T--T-----T--A---------C----------A-----------G--A--C-----A-----T--G--A--A--C--T--------G-----C------CA-------- 1723
H2B.CI.x.EHO.U27200         A--G-----G--C--G-----------A-----------------A--------G--------T-----T-----------------------A-----A-----G--A--C-----A-----T--G--G--A--T--T-----A--A-----C------CAA-----T- 1719
H2B.GH.86.D205.X61240       A--------G--C--G-----------A-----------------A--------A--------T-----T--A---------C----------A-----A-----G--A--C-----A-----T-----A--A--C--T-----A--A-----C------CA---A---- 1718
H2B.JP.01.KR020.AB100245    ------------C--G-----------A--------C--------A--------A--------T-----T------------C----------A-----A--------A--C-----A-----T--G--A--A--T--T--T--A--A-----C------CA-------- 864
H2AB.CM.03CM_510_03.NEW     A--------G--C--G-----G-----A--------C--A-----A--------A-----T--T-----T--A---------C----------A--G--------G--A--C-----A-----T-----A--A--T--T--T--A--A-----C------CAA--A---- 854
H2G.CI.92.ABT96.AF208027    A--------G--C--G--------G-----------------G-----------C---CTA--A-----T--C-----A---C----------A-----A--------A--C-----A-----T-----A--A--T--T--------A-----K-----TC-T--A--T- 1120
H2U.FR.96.12034.AY530889    ---G--A--------------------------------------C-----G--C-----G--------T--------A-----------C--------A--------T--------A--------G--A--A--T--T--T--A--A------------ACA------- 1236
SMM.SL.92.SL92B.AF334679    A--------------------------------------------A--------C--T--A---------------------C-T--------A--T--A-----G--A--C-----A-----------A--A--T--T-----A--A-----C------C-A--A---- 1141
MAC.US.x.MAC239             -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 784
MAC.US.x.251                ---------------------------------------------------------------------------------------------------------------------------------C--C------------------------------------- 1706
MAC.US.x.251_32H            ---------------------------------------A--------------C--------------------------------------------------------------------------C--------T------------------------------- 1708
MAC.US.x.251                ------------------------------------------------------C-----------------------------------------------------------------------------------T------------------------------- 1695
MAC.US.x.17EC1.AY033233     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1707
MAC.US.x.17EFR.AY033146     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1707
MAC.US.x.1937.AY611495      ----------------------------------------------------------TG----------------------------------------------------------------------G--------------------------------------- 813
MAC.US.x.2065.AY611493      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 813
MAC.US.x.MM142.Y00277       ---G------------------------T-------------------------C---CT------------------------------------------------------------C-----------------T------------------------------- 1205
MAC.US.x.SMM142B.BD131285---G------------------------T-------------------------C---CT------------------------------------------------------------C-----------------T------------------------------- 1205
MNE.US.82.MNE               ---G--------------------------------------------------C--------------------------------------------A--------A--------------------A---------------------------------------- 1187
MNE.US.x.MNE027.U79412      ---G--------------------------------------------------C----T---------------------------------------A--------A--------------------A------------------------------C--------- 1187
SMM.US.x.F236               A--G--A--------G-----G-----------------A--G--A--------C-----T--------T--------A---C----------A-----A-----G--A--C-----A-----T-----G--------T--A--A-----C--------T--A--A---- 1703
SMM.US.x.H9.M80194          A--G--A--------G-----G-----------------A--G--A--------C-----T-----Y--T--C-----A---C----------A-----A-----G--A--C-----A-----T-----A--------T--A--------T--------T--A--A---- 1189
SMM.US.x.PBJ14              A--G--A--------G-----G-----------------A--G--A--------C-----T-----------C-----A---C----------A-----G-----G--A--C-----A-----T-----A--------T--A--A-----T--------T--A--A---- 1483
SMM.US.x.PBJA.M31325        A--G--A--------G-----G-----------------A--G--A--------C-----T-----------C-----A---C----------A-----G-----G--A--C-----A-----T-----A--------T--A--A-----T--------T--A--A---- 1482
SMM.US.x.PBJ                A--G--A--------G-----G-----------------A--G--A--------C-----T-----Y-----C-----A---C----------A-----G-----G--A--C-----A-----T-----A--------T--A--A-----T--------T--A--A---- 1189
SMM.US.x.PBJ                A--G--A--------G-----G-----------------A--G--A--------C-----T-----------C-----A---C----------A-----G-----G--A--C-----A-----T-----A--------T--A--A-----T--------T--A--A---- 1675
SMM.US.x.PGM53.AF077017     A--G--A--------G-----G-----------------A--G--A--------C-----T-----------C-----A---C----------A-----A-----G--A--C-----A-----T-----A--------T--A--A-----T--------T-----A---- 1633
SMM.US.x.SME543.U72748      A--G--A--------G-----G-----------------A--G--A--------C-----T-----------C-----A---C----------A-----A-----G--A--C-----A-----T-----G--------T--A--A-----C--------T--A--A---- 1719
STM.US.x.STM.M83293         ---------------T--------G--------------------A--------C--T--------------C-----AC----------------G--G--------A--C-----A-----A-A---G--C-----T-----A-----T---------A-------T- 1363
MAC.US.x.80035.AY611486     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 813
MAC.US.x.81035.AY599200     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 814
MAC.US.x.85013.AY611490     ----------------------------------------------------------G-----------------------------------------------------------------------G--------------------------------------- 747
MAC.US.x.92050.AY603959     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 807
MAC.US.x.92077.AY599201     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 814
MAC.US.x.93057.AY611492     ---------------------------------------------------------T-T----------------------------------------------------------------------G--------------------------------------- 812
MAC.US.x.93062.AY607704     ---------------------------------------------------------------Y---------------------------------------------------------------------------------------------------------- 799
MAC.US.x.95058.AY611494     -----------------------------------------------------------T----------------------------------------------------------------------R--------------------------------------- 801
MAC.US.x.95086.AY607703     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 670
MAC.US.x.95112.AY588946     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 803
MAC.US.x.96016.AY607701     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 814
MAC.US.x.96020.AY611488     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 806
MAC.US.x.96072.AY611491     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 818
MAC.US.x.96081.AY597209     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 816
MAC.US.x.96093.AY611489     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 814
MAC.US.x.96114.AY588945     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 809
MAC.US.x.96123.AY611487     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 703
MAC.US.x.96135.AY607702     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 799
MAC.US.x.97009.AY599199     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 682
MAC.US.x.97074.AY599198     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 682
MAC.US.x.r80025.AY576480    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 670
MAC.US.x.r90131.AY576481    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 813
```

```
MAC.US.x.239            CACAACCAGCTCCA.........CAACAAGGACAACTTAGGGAGCCGTCAGGATCAGATATTGCAGGAACAACTAGTTCAGTAGATGAACAAATCCAGTGGATGTACAGACAACAGAACCCCATACCAGTAGGCAACATTTACAGGAGATGGATCCAACTGGGGTTGCAA 1868
Gag                     P__Q__P__A__P__.__.__.__Q__Q__G__Q__L__R__E__P__S__G__S__D__I__A__G__T__T__S__S__V__D__E__Q__I__Q__W__M__Y__R__Q__Q__N__P__I__P__V__G__N__I__Y__R__R__W__I__Q__L__G__L__Q_
H2A.CI.88.UC2.U38293    --AT-----GC--CTTA......C-GC---G--G-----A--C--AAGG--G--T--C--A-----------A--CA----------G--G-------------T--G--G--A--T---G-G--G-----G-----C-----A-----------GA-A---C----G 1921
H2A.DE.x.BEN.M30502     --AT-----GC--CTTA......C-GC------G--C--A--C--AAG---G--T--C--A-----------A--CA-------------G-------------T--G-C---A--T---G----G-----G-----C-----A------------A-A---C----- 1921
H2A.DE.x.PEI2.U22047    --AT-----GC--CTTG......C-GCG--G--G-----A--A--AAG---G--T-----A-----G-----A--CA-------A-----G------------TT---GC---A--T--T-----------G-----C--T-------------GA-A--AC---G 1889
H2A.GH.x.GH1.M30895     --AT-----GC--CTTA......C-GC---G--G-----A--C--AAGG--G--T--C--A-----------A--CA-------A-----G-------------T--G-C---A--T---G-G--G-----G-----C-----A-----------GA-A---C-A--G 1365
H2A.IN.NNVA.NEW         --AT-----GC--CTTA......C-GC---G--G--C--A--A--AAG---G--T--C--A-----------A--CA-------A-----G---G---------TT---GC---A--T--TG-------G--A-----C--T--A-----------GA-A---C----- 1915
H2A.IN.95.CRIK          --AT-----GC--CTTA......C-GC------G-----A-----AAG------T--C--A-----G-----A--CA-------G-----G------------TT---GC---A--T--TG-------G--A-----C--T--A-----------GA-A--AC---G 1666
H2A.GM.x.MCR35.AY509260 --AT-----G---CTTA......C-GCG--G--G-----A--A--AAG---G--T--C--A-----G-----A--CA-------------G------------TT--G-C---A--T---G-------G--A-----C--T--------------GA-A--AC---G 1366
H2A.GM.87.D194.J04542   --AT-----GC--CTTA......C-GC---G--G-----A--C--AAGG--G--T--C--A-----------A--CA-------------G-------------T--G----CA--T---G-G--G-----G-----C----------------GA-A---C-A--G 1365
H2A.GM.x.ISY.J04498     --AT-----GC--CTTA......C-GCG--G--G--C--A--T--ACG------T--C--A-----G-----A--CA-------G-----G---G-A--------T--G---G-A--T--TG----------A-----C--T-------------GA-A--AC---G 1365
H2A.GM.x.MCN13.AY509259 --AT-----G---CTTA......C-GCG--G--G-----A--A--AAG---G--T--C--A-----G-----A--CA-------------G------------TT--G-C---A--T---G-------G--A-----C--T-------------GA-A--AC---G 1366
H2A.GW.86.FG.J03654     -CAT-----GC--CTTA......C-GCG--G--G-----A-----AAG---G--T--C--A-----G-----A--CA-------A--G--G------------TT----C---A--T--TG----------------C--T-----------T--GA-A--AC----- 1359
H2A.GW.87.CAM2CG.D00835 --AT-----GC--CTTA......C-GCG--G--G--C--A--C--AAG---G--T--C--A-----------A--CA-------A-----G-----A-------TT--GGC---A--T---G----------G-----C--T--A-----------GA-A---C----G 1930
H2A.GW.x.ALI.AF082339   -CAT-----GC--CTTA......C-GC---G--G-----A--A--AAG---G--T--C--A-----------A--CA-------A-----G-------------T---G-C--G---T--TG-G--------G-----C--T--A-----------GA-A---C-A--G 1914
H2A.GW.x.MDS.Z48731     --AT-----GC--CTTA......C-GCG--G--G-----A--A--AAG---G--T--C--A-----G-----A--CA----------------------------T---GGC------T--TG-------G--A-----C--T--A-----------GA-A---C----G 1365
H2A.SN.85.ROD.M15390    --AT-----GC--CTTA......C-GCG--G--G-----A-----AAGG-----T--C--A-----G-----A--CA-------A-----G------------TT--G-C---A--T--TG----------A-----C--T--A-----------GA-A--A-----G 1364
H2AB.CI.90.7312A.L36874 -GGTG----G----CTG......C-GC-------------C--CAG-----------A--G-----C--C--CA-------G-----G--A-------------GGGT--A-GT--T--C---------G--------T--A---------T--G--A--A-----G 1888
H2B.CI.88.UC1.L07625    -GAT-----G----CTG......C-GC-------------C--CAG---G--------A--------C--C--CA-------G--------A-------------GGC---A--T--TG-T--G-----G-----------A---------T---A-A---C----G 1887
H2B.CI.x.EHO.U27200     --TCG----GC---ATG......C-GC---G-----C-----A--AAG---G-----C--A--G-----C--C--CA-------A-----G--A---------------CC--A--T--TG-C-----G--A-----C--T--A--G-----T--GT-A---C-A--G 1883
H2B.GH.86.D205.X61240   -GTC-----GC---ATG......CGGC-------------C--AAG---G--------A--------C--C--CA-------G-----G--A-------------GGCC--A--T--TG-C-----G--A--------T--A---------T---T-A--A-----G 1882
H2B.JP.01.KR020.AB100245 -GTC---G-GC---ATG......C-GC------G--------C--AAG---G-----C--A--------C-----CA----G--G--G--G--A---------------G-CC--A-----T--C---------A--------T--A---------T--GT-A---C----G 1028
H2AB.CM.03CM_510_03.NEW -GTC-----GC---CTG......CGGC------G-----A--C--AAGG-----------A--G-----C--C--CA-------G-----G--A---------------GGGC--A--T--T--C-----G--A--------T--A---------T---A-A--------G 1018
H2G.CI.92.ABT96.AF208027 ------A--GG--GCCG......C-GC-------M-C-----A---A-------------A-----G--C------A--A----A--------T------CAC----G--G--A-----T-----------A--T--Y--T--A-----------RT-A--AC-T--- 1284
H2U.FR.96.12034.AY530889 --A-C-A--GC---CCA......C-GC---G--------A-----AAG---T--T-------------------CA----G--A--G--G--A--------------G-CG--A--T--A-----G--G--T-----C--TC--------------T----CC-A--- 1400
SMM.SL.92.SL92B.AF334679 --AG-GG-CAA-AGCCGGCGCAA-CGGCT---GGG--A-----A--A-----------C--A-----G--------CA-CCC-AG------G--AG-A--------T--GGC---A-----AG-C-----G--AG----C--T--A--G---------G-----AC-C--- 1311
MAC.US.x.MAC239         --------------.........------------------------------------------------------------------------------------------------------------------------------------------------  945
MAC.US.x.251            --------------.........------------G-------------------------------------------------------------------------------------------------------------------------------------- 1867
MAC.US.x.251_32H        --------------.........------------G-------------------------------------------------------------------------------------------------------------------------------------- 1869
MAC.US.x.251            --------------.........------------G-------------------------------------------------------------------------------------------------------------------------------------- 1856
MAC.US.x.17EC1.AY033233 --------------.........------------------------------------------------------------------------------------------------------------------------------------------------- 1868
MAC.US.x.17EFR.AY033146 --------------.........------------------------------------------------------------------------------------------------------------------------------------------------- 1868
MAC.US.x.1937.AY611495  --------------.........-----------------------------------------A--------------------------------------------------------------------R--------------------------  974
MAC.US.x.2065.AY611493  --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  974
MAC.US.x.MM142.Y00277   ------A-------.........------------G-----------------------------A-------A------------------------------------------------------------------------------------------- 1366
MAC.US.x.SMM142B.BD131285------A-------.........------------G-----------------------------A-------A------------------------------------------------------------------------------------------- 1366
MNE.US.82.MNE           ------A-------.........------------G-------------------C---------A------------------------------------------------------------------------------------------------- 1348
MNE.US.x.MNE027.U79412  ------A-------.........------------G-------------------C---------A---------------------------------------------------------T--------------------------------------- 1348
SMM.US.x.F236           -G-------G----CTA......C-GC---G--------A-----AAG---------C-----------T------A------------------A-----------G-----A------------------------T--A--G---------T-A---C----G 1867
SMM.US.x.H9.M80194      -R--G----G----ATA......C-GC-----------A-----AAGR-----C--C-----------T-----CA--S-----C--------T--A--------Y--G--G--A-----T-----------------C-----R--G---------T-A---C----G 1353
SMM.US.x.PBJ14          -G--G----G---GATA......C--C-----------A-----AAG------C--C-----------T-----CA--------C--------T--A--------T--G--G--A-----T-----------T-----C-----------G---------T-A---C----G 1647
SMM.US.x.PBJA.M31325    -G--G----G---GATA......C--C-----------A-----AAG------C--C-----------T-----CA--------C--------T--A--------T--G--G--A-----T-----------T-----C-----A--G-----------A---C----G 1646
SMM.US.x.PBJ            -R--G----G----ATA......C-GC-----------A-----AAGR-----C--C-----------T-----CA--S-----C--------T--A--------Y--G--G--A-----T-----------T-----C-----R--G---------T-A---C----G 1353
SMM.US.x.PBJ            -G--G----G---GATA......C--C-----------A-----AAG------C--C-----------T-----CA--------C--------T--A--------T--G--G--A-----T-----------T-----C-----A--G---------T-A---C----G 1839
SMM.US.x.PGM53.AF077017 -G--G----G----ATA......C-GC---------C--A--T--AAG----------------------T------A----G--A--G-----T-----------T--G--G-------------------G--A-------------------------T-----A--- 1797
SMM.US.x.SME543.U72748  -G--G----G----CTA......C-GC---G--------A-----AAG---------C-----------T------A-------G----------A-----------G-----A-------------------T-----G---------T-A---C----G 1883
STM.US.x.STM.M83293     --------C-GGGTCCGCTG...C-GC---G--------A--A--AAGC--G---------A-----G--T-----C---CC---G--------T--A----------C-G--G--A--T--T-----G-----A-----C---------------T-----A---C-A--- 1530
MAC.US.x.80035.AY611486 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  974
MAC.US.x.81035.AY599200 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  975
MAC.US.x.85013.AY611490 --------------.........-----------------------------------------W--R---------------------------------------------------------------------------------------------------  908
MAC.US.x.92050.AY603959 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  968
MAC.US.x.92077.AY599201 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  975
MAC.US.x.93057.AY611492 --------------.........--------------------------------------------C--------------------------------------------------------------------------------------------------  973
MAC.US.x.93062.AY607704 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  960
MAC.US.x.95058.AY611494 --------------.........--------------------------------------------C---G-----------------C-------------------------------------------Y-------------------------------  962
MAC.US.x.95086.AY607703 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  831
MAC.US.x.95112.AY588946 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  964
MAC.US.x.96016.AY607701 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  975
MAC.US.x.96020.AY611488 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  967
MAC.US.x.96072.AY611491 --------------.........----------------------------------------------------------------------------------------------------------------------------M-R--------------  979
MAC.US.x.96081.AY597209 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  977
MAC.US.x.96093.AY611489 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  975
MAC.US.x.96114.AY588945 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  970
MAC.US.x.96123.AY611487 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  864
MAC.US.x.96135.AY607702 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  960
MAC.US.x.97009.AY599199 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  843
MAC.US.x.97074.AY599198 --------------.........-------------------------------------------------------------------------------------------------------------------------------------------------  843
MAC.US.x.r80025.AY576480 --------------.........------------------------------------------------------------------------------------------------------------------------------------------------  831
MAC.US.x.r90131.AY576481 --------------.........------------------------------------------------------------------------------------------------------------------------------------------------  974
```

```
MAC.US.x.239                AAATGTGTCAGAATGTATAACCCAACAAACATTCTAGATGTAAAACAAGGGCCAAAAGAGCCATTTCAGAGCTATGTAGACAGGTTCTACAAAAGTTTAAGAGCAGAACAGACAGATGCAGCAGTAAAGAATTGGATGACTCAAACACTGCTGATTCAAAATGCTAACCC 2038
Gag                         _K__C__V__R__M__Y__N__P__T__N__I__L__D__V__K__Q__G__P__K__E__P__F__Q__S__Y__V__D__R__F__Y__K__S__L__R__A__E__Q__T__D__A__A__V__K__N__W__M__T__Q__T__L__L__I__Q__N__A__N__P
H2A.CI.88.UC2.U38293        --G--------------C--------C---G-CT----C-----G--G--A--------AT----C-----------------------T-----C-----G-----G--AG----CC-G--------A-----------C------------G-A--G-----C------ 2091
H2A.DE.x.BEN.M30502         --G----------A---C--------T-----CT----CA-------G--A--------A--G--C--A--------------------------C--G--G--------A-----CC----------A-----------C-----G-----A--A--G-----C------ 2091
H2A.DE.x.PEI2.U22047        --G--C-----G-----C--T-----C-----C-----C--------G--A-----G-----G--C--A-----------T--A-----------CC----G--------A-----CC----------A-----------C------------G-A--G-----C------ 2059
H2A.GH.x.GH1.M30895         --G--------G-----C--------T-----CT----C-----G--G--A-----G--A--G--C-----------G-----------T-----C--G--G--------A------C-G-----------C--------C-----G-----A--A--G-----C------ 1535
H2A.IN.NNVA.NEW             --G--------G-----C--T--G--C-----C-----CA-------G--A--------A--G--C--A-----------T--A--------G--C--G------------------C----------------------C------T-A-----A--------C------ 2085
H2A.IN.95.CRIK              --G--------G-----C--T--G--T-----C-----CA-------G--A--------A--G-----A-----------T--A--------G--C--G-----------------CC----------------------C------T-A---G-A--------C------ 1836
H2A.GM.x.MCR35.AY509260     --G--C-----G-----C-----G--C-----C-----CA----------A--------A-----C--A--T--------T--A-----------C--G--G--------A------C--A-------------------C--G-----A--AG-A--G-----C------ 1536
H2A.GM.87.D194.J04542       -----------G-----C--------T-----CT-------G--G--G--A--------AT-G--C-----------------A-----------CC----G--------A-----CC-G--T-----A-----------C-----G-----A--A--G-----C------ 1535
H2A.GM.x.ISY.J04498         --G--------------C--T-----C-----------CA----------A---------T-G--C--A--------G--T--A-----------C-----G-----------------------G--------------C--G--G-----AG-G---...T-G------ 1532
H2A.GM.x.MCN13.AY509259     --G--C-----G-----C-----G--C-----C-----CA----------A-----G--A-----C--A--T--------T--A-----------C--G--G--------A------C-------G--------------C--G-----A--AG-A--G-----C------ 1536
H2A.GW.86.FG.J03654         --G--C-----G-----C--T-----C-----C-----CA----T-----A-----G--A-----C--A-----------T--------------C--G--G--------A------C----------A-----------C-----------A--A--G-----C------ 1529
H2A.GW.87.CAM2CG.D00835     --G--------G-----C-----G--C-----C-----CA-------G--A-----G---T----C--A-----------T--A-----------C--G--G--------A------C----------A-----------C-----------AG-G--G-----C------ 2100
H2A.GW.x.ALI.AF082339       --G--------G-----C--------C-----C-----CA----------A-----G--------C--A-----------T--A-----------C-----G--------A------C-----------------------------------G-A--G-----C------ 2084
H2A.GW.x.MDS.Z48731         --G--------G-----C--T-----C-----C-----C--------G--A-----G-----G--C--A-----------T--A-----------C--G--G--------A--------------G--A--------------G--------A--A--------C------ 1535
H2A.SN.85.ROD.M15390        --G--------G-----C-----G--C-----C-----CA-------G--A-----G-----G--C--A-----------T--A-----------C--G--G--------A------C-------G--------------C-----------AG-A--------C------ 1534
H2AB.CI.90.7312A.L36874     --G------C-G-----C--T--T--T--T--A--------------------------A--C--C--A--------G--T--A--------G--C---C-G--------AG-G--CA-G--C--G-GAGCG--------AG-G-----A---G-C--G------------ 2058
H2B.CI.88.UC1.L07625        ---------C-G-----C--T--T--C-----AT----CA----G--G--A--------A--T--C--A-----C-----T--A-----------C---C-G-----G--A-----CC-------G--A-----------A-----------A-----G------------ 2057
H2B.CI.x.EHO.U27200         ---------C-G-----C--T--T--T--T--A--G--CA----G--------------A--C--C--A-----------T--A-----------C---C-G-----G--A-----CC----------A-----------A-----------------G------------ 2053
H2B.GH.86.D205.X61240       ---------C-------C--T--T--C-----AT----CA----G--G--A-----G-----C--C--A-----------T--A-----------C---C-G--------A-----CC-------G--A-----------A-----------------G------------ 2052
H2B.JP.01.KR020.AB100245    ---------C-G-----------T--C--T--A--G--CA----G--G--A-----G--A--C--C--A--------------A---------------C-G--------A-----CC-G--------A------G----G-----------------G------------ 1198
H2AB.CM.03CM_510_03.NEW     ---------C-G-----C--T--T--C-----AT-G--CA----G--G--A--------A--C--C--A--T--------T--A-----------C---C-G-----G--A-----CC-------G--A-G---------A--------A---G---G------------ 1188
H2G.CI.92.ABT96.AF208027    -----------G-----Y--T-----------C--------------------------A--C------TCT--------T--A--------G---C-R-----------A--T---C----------A-----------A--G-----A--------------C------ 1454
H2U.FR.96.12034.AY530889    --------T--------C--T-----T--T--CT-----A----G-----------G-----------A--T--------T--A--------------G-----------A------C-------G--A--------------G--G-----------G------------ 1570
SMM.SL.92.SL92B.AF334679    -----C-----G--------------------C--G-----T--------T-----G-----C-----A--T--------T-----------------G--------G---------C-G-----G--------------A-----------------G--C--C------ 1481
MAC.US.x.MAC239             --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1115
MAC.US.x.251                --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2037
MAC.US.x.251_32H            --------------------------------------------------A--------------------------------------------C--------------A------------------------------------------------------------ 2039
MAC.US.x.251                -----------------------------------------------------------------------------------------------C--------------A------------------------------------------------------------ 2026
MAC.US.x.17EC1.AY033233     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2038
MAC.US.x.17EFR.AY033146     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2038
MAC.US.x.1937.AY611495      --------------------------------------------R------------------------------------------------------------------------C----------------------------------------------------- 1144
MAC.US.x.2065.AY611493      --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1144
MAC.US.x.MM142.Y00277       --------------------------------------------------------------------------------------------------------------A------C----------------------------------------------------- 1536
MAC.US.x.SMM142B.BD131285   --------------------------------------------------------------------------------------------------------------A------C----------------------------------------------------- 1536
MNE.US.82.MNE               ----------------------------G------------------------------------------------------------------C--------------A------C----------------------------------------------------- 1518
MNE.US.x.MNE027.U79412      ----------A----------------------------A---G---------------------------------------------------C--------------A------C----------------------------------------------------- 1518
SMM.US.x.F236               --------A------------------------T-------G--------A-----------------A-----C-----T--A------------C----------G--A------C-C--------------------C------------------------------ 2037
SMM.US.x.H9.M80194          -----C--A-----------T------------T-G-----G--------A---------T-------A-----------T--A--------G---C-------------A-----CC-G---S-G--A-----------------------------------C------ 1523
SMM.US.x.PBJ14              --------A-----------T------------T-G-----G--------A-----------------A-----------T--A--------G--C-C------T-----A-----CC-G-----G--A-----------------------------------C------ 1817
SMM.US.x.PBJA.M31325        --------A-----------T------------T-G-----G--------A-----------------A-----------T--A--------G---C-------------A-----CC-G-----G--A-----------------------------------C------ 1816
SMM.US.x.PBJ                --------A-----------T------------T-G-----G--------A-----------------A-----------T--A--------G---C-------------A-----CC-G---S-G--A-----------------------------------C------ 1523
SMM.US.x.PBJ                --------A-----------T------------T-G-----G--------A-----------------A-----------T--A--------G--C-C------T-----A-----CC-G-----G--A-G---------------------------------C------ 2009
SMM.US.x.PGM53.AF077017     --G--------------C--T--------------G--C--------------G--------G--C--------C-----T--------------C--------------A------C-------G--------------------------------------------- 1967
SMM.US.x.SME543.U72748      --------A------------------------T-------G--------A-----------------A-----C-----T--A------------C----------G--A-----CC-G--------------------C------------------------------ 2053
STM.US.x.STM.M83293         -----------------C--T---GTT------T----C-----G--------------AA-G--C--A--T--------------------G-----G-----------AG-----C--T----G--A-------------GG--------A--------------T--- 1700
MAC.US.x.80035.AY611486     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1144
MAC.US.x.81035.AY599200     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1145
MAC.US.x.85013.AY611490     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1078
MAC.US.x.92050.AY603959     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1138
MAC.US.x.92077.AY599201     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1145
MAC.US.x.93057.AY611492     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1143
MAC.US.x.93062.AY607704     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1130
MAC.US.x.95058.AY611494     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1132
MAC.US.x.95086.AY607703     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1001
MAC.US.x.95112.AY588946     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1134
MAC.US.x.96016.AY607701     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1145
MAC.US.x.96020.AY611488     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1137
MAC.US.x.96072.AY611491     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1149
MAC.US.x.96081.AY597209     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1147
MAC.US.x.96093.AY611489     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1145
MAC.US.x.96114.AY588945     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1140
MAC.US.x.96123.AY611487     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1034
MAC.US.x.96135.AY607702     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1130
MAC.US.x.97009.AY599199     -------------------------------------------------------------------------------T------------------------------------------------------------------------------------------- 1013
MAC.US.x.97074.AY599198     --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1013
MAC.US.x.r80025.AY576480    --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1001
MAC.US.x.r90131.AY576481    --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1144
```

```
                                                
```

```
                                                                                                                          Gag-Pol -1 ribosomal slip site
                                                                                                                                   Pol start
                                                                                                              Gag p8 Nucleocapsid end Gag p1 Spacer start
MAC.US.x.239               .....GGACCAAGAAAG...CCAATTAAGTGTTGGAATTGTGGGAAAGAGGGACACTCTGCAAGGCAATGCAGAGCCCCAAGAAGACAGGGATGCTGGAAATGTGGAAAAATGGACCATGTTATGGCCAAATGCCCAGACAGACAGGCGGGTTTTTTAGGCCTTGGTCCA 2369
Pol                                                                                                                                                                           F  F  R  P  W
Gag                        .__.__G__P__R__K__.__P__I__K__C__W__N__C__G__K__E__G__H__S__A__R__Q__C__R__A__P__R__R__Q__G__C__W__K__C__G__K__M__D__H__V__M__A__K__C__P__D__R__Q__A__G__F__L__G__L__G__P_
H2A.CI.88.UC2.U38293       ..........--G-G-...A-----G--------C-------G----------G--G-AA--G--C---A--T----------C-------------C--GCCA-GA--CA-C-----A--T-----G---------------------GA-G--CT-- 2413
H2A.DE.x.BEN.M30502        ..........-------...G-----G--A-----C-----A--G----------G----A--G--C---A--T----------C-------G-----C--GCCA-GA--CA-C-----A--C-----G--A--------A----------GT-G--C--- 2413
H2A.DE.x.PEI2.U22047       ..........----G-...A---------C----------A--G--T--G----G----A-----C---A--T----------C-------------C---TCA-GA----C-----A--C--------A--------T----------GA---C--- 2381
H2A.GH.x.GH1.M30895        ..........-------...GT----G--------C--C--A--G----------G--G--A--G--C---G--T----------C-------------C--G-CA-GA--C--C-----A-----------A--------A----------GA-G------ 1857
H2A.IN.NNVA.NEW            ..........-------...A-----G--C----------A--G--A-------A----A--G--C---A--T---A-----C--T----G-----C--GCCA-GA---A-C-----A--C--T-----T------------------GA-G--C--C 2407
H2A.IN.95.CRIK             ..........-------...A-----G--C----------A--G--A-------A----A-----C---A--T--G-------C--T----G-----C--G-CA-GA-----C-----A--C--------A------------------GA-AA-C--C 2158
H2A.GM.x.MCR35.AY509260    ..........-A--G-...A-------A--C----------A--G--A-------G----A-----C---A--T----------C-------G-----T---CCA-GA-----C--A--A--T--------T-------T-----------GA-G--C--C 1858
H2A.GM.87.D194.J04542      ..........----G-...G-----G---------------A--G----------G--G-AA--G--C---A--C----------C-------G-----C--GTCA-GA--CA-C-----A--C-----G--A--------A----------GA-G--C--- 1857
H2A.GM.x.ISY.J04498        ..........-A--G-...G---------------------A--G--A--G----G----A-----C---G--T----------C-------------C--GTCA-GA--CA-C-----A--C--------T--------T----------G-----A--- 1854
H2A.GM.x.MCN13.AY509259    ..........-A--G-...A------A--C----------A--G--A-------G----A-----C---A--T----------C-------G-----T---CCA-GA-----C--A--A--T--------T-------T-----------GA-G--C--C 1858
H2A.GW.86.FG.J03654        ..........-------...A--T-C----C----------A--G--A-------G----A----G--C--G--T----------C-------G-----T--GTCA-GA-----C-----A--T--------T--------T----------GA----C-AC 1851
H2A.GW.87.CAM2CG.D00835    ..........----G-...A-G-----A--C--------------G----G----G----A-----C---A--T----G----C-------G-----C--GCCA-GA--CA-C---A-A--T-----T--T------------------GA----C--C 2422
H2A.GW.x.ALI.AF082339      ..........----G-...A---------C----------A--G--A--G----G----A-----C---A--C---A-----C-------------C--GCCA-GA---C-------A--C--------A--------T----------G-----C--- 2406
H2A.GW.x.MDS.Z48731        ..........----G-...A---------C----------A-------G----G----A-----C---A--T----------C-------------C--G-CA-GA---A-C-----A--G------------------------------GA----C--C 1857
H2A.SN.85.ROD.M15390       ..........-------...G--T----A--C-----C-----A--G--A--G----G----A-----C---A--T----G----C-------G-----T--GCCA-GA--CA-C---A-A--C--------T--------A----------A--G--C--T 1856
H2AB.CI.90.7312A.L36874    .....-C-GG-----GAGGGA--G-G-CA--C-----C-----C----G------TA-A--C-A--------G--T--T--------------------------CCA-GA--CAAC-----A--T--T-----A--------T----------G-----A--C 2392
H2B.CI.88.UC1.L07625       .....-C-GG--AG-GAGGGA-GG-G-CA--C-----C-----C--G-TA------A-A--C-A--------G-----T--------------------------CAA-GA--CA-C---T-A-----------A--------T----------GT-A--A--C 2391
H2B.CI.x.EHO.U27200        .....-C-GGG-AG-G-...A--G-G-CA--C-----C--C--C--G-C------TA-A--C-----G---AG-----T----G--------------------C-GCAA-GA--CA-C---T-A-----------A--------------------GT-C--C--C 2384
H2B.GH.86.D205.X61240      .....-C-GGG-AG-GAGGGA--G-G-CA--C-----C-----C---C--------A-A--C----------G-----T--------------------------CA-GA--CA-C---T-A-----------A--------------------GT-A--A--C 2386
H2B.JP.01.KR020.AB100245   .....-C-GGG-AG-GAGGAG--G-G-CA--C-----C-----C---ATA------A-A--C--------T-A------T----G---------------C------CAA-GA--CA-C---T-A-----------A--------------------AT----C--C 1532
H2AB.CM.03CM_510_03.NEW    .....AC-GGG-AG-GAGGGA--G-G-CA--C-----C-----C----G-------A-A--C--A--G---AG-----T--------------------------CAA-GA--CA-C---T-A-----------A--------TT-----------GT-A--A--C 1522
H2G.CI.92.ABT96.AF208027   ..ACA---GG--A--G-AGCA----A--A-------------C--G--------TA-A-TT--A--G-----------C-----------G--------G-----G--GCCA-G---CA-------------T-----G--------A-----------GT----A--T 1791
H2U.FR.96.12034.AY530889   ..GGC--C-----G-G-...---G-C-GA--------C-----C--G--A------A-----------G--T-A------T------------------------A---GCCA-GA--CAAC-----T-GC--------A-CCA---T-------------G-----C--T 1904
SMM.SL.92.SL92B.AF334679   AAGGCCA-GGG-AT---CCTATC--C-G-----TT--C-----A--GACA-----------G--A--G--------T--C-----GA------------------GG--GAA-GA-G-A--CA------C------A--CAGA---------------------C--- 1821
MAC.US.x.MAC239            ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1446
MAC.US.x.251               ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 2368
MAC.US.x.251_32H           ....-------------...----------------------------A------------------------------------------------------------------------------------------------------------------- 2370
MAC.US.x.251               ....-------------...-------------------------G--------------------------------------------------------------------------------------------------------------------- 2357
MAC.US.x.17EC1.AY033233    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 2369
MAC.US.x.17EFR.AY033146    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 2369
MAC.US.x.1937.AY611495     ....-------------...--------------------------------------------------------------------------------------------------------------------------------M----Y--- 1475
MAC.US.x.2065.AY611493     ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1475
MAC.US.x.MM142.Y00277      ....-------------...-------------------------G--A-------------------------------------------------------------------A-------------------------C--T 1867
MAC.US.x.SMM142B.BD131285  ....-------------...-------------------------G--A-------------------------------------------------------------------A-------------------------C--T 1867
MNE.US.82.MNE              ....-------------...-----------------------------------------------A------------C--------------C-----G-----------------------A-----------T----C--- 1849
MNE.US.x.MNE027.U79412     ....-------------...------------------------------------------------------------C--------------C-----G-----------------------A-----------T----C--- 1849
SMM.US.x.F236              ....----A--G---...A---C---------------A--G---------------AA-------G----T----------C-------G---------C--GT-----------------T--A--------------------G-----C--- 2368
SMM.US.x.H9.M80194         ....----A-R-----...AT----W------R--C--R-A--G--A-------------A---T---------T-------R--C-------G--------GCA-G-----------T-------C--A---------------------G-----C--- 1854
SMM.US.x.PBJ14             ....----A--------...AT---------------C----A--G--A-------------A-------------T----------C-------G--------GCA-G-----------T-------C--A----A----------------G-----C--- 2148
SMM.US.x.PBJA.M31325       ....----A--------...AT---------------C----A--G--A-------------A-------------T----------C-------G--------GCA-G-----------T-------C--A----A----------------G-----C--- 2147
SMM.US.x.PBJ               ....----A-R-----...AT----W------R--C--R-A--G--A-------------A-------------T-------R--C-------G--------GCA-G-----------T-------C--A---------------------G-----C--- 1854
SMM.US.x.PBJ               ....----A--------...AT---------------C----A--G--A-------------A-------------T----------C-------G--------GCA-G-----------T-------C--A----A----------------G-----C--- 2340
SMM.US.x.PGM53.AF077017    ....-------------...A---------C----------A--G----C---------AA-------------T----------C-----------------CA-G-----C-----------------C--A---------------------G-----C--- 2298
SMM.US.x.SME543.U72748     ....----A--G-G-...A---C---------------A--G---------------A-------G----T----------C-----GGG--C------C--GT-----------------T--A--------------------GT----C--- 2384
STM.US.x.STM.M83293        .....CAGGG-C---G-...A--G-A--A--C---------------CA--G--A-----AA--G--T-A--G---G-------A--T--T-----------C--CA--G----CAG---------------A-------T-----------T----C--- 2028
MAC.US.x.80035.AY611486    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1475
MAC.US.x.81035.AY599200    ....-------------...------------------------------------------------------------------------------------------------------Y------------------------------------ 1476
MAC.US.x.85013.AY611490    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1409
MAC.US.x.92050.AY603959    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1469
MAC.US.x.92077.AY599201    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1476
MAC.US.x.93057.AY611492    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1474
MAC.US.x.93062.AY607704    ....-------------...A-------------------------------------------------------------------------------------------------------------------------------------------- 1461
MAC.US.x.95058.AY611494    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1463
MAC.US.x.95086.AY607703    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1332
MAC.US.x.95112.AY588946    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1465
MAC.US.x.96016.AY607701    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1476
MAC.US.x.96020.AY611488    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1468
MAC.US.x.96072.AY611491    ....-------------...------------------------------------------------------------------------------------------------------T-----------------C--- 1480
MAC.US.x.96081.AY597209    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1478
MAC.US.x.96093.AY611489    ....-------------...-H------------------------------------------------------------------------------------------------------------------------------------------ 1476
MAC.US.x.96114.AY588945    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1471
MAC.US.x.96123.AY611487    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1365
MAC.US.x.96135.AY607702    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1461
MAC.US.x.97009.AY599199    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1344
MAC.US.x.97074.AY599198    ....-------------...--------------------------------------------------------------------------------------------------------------------------------------Y--- 1344
MAC.US.x.r80025.AY576480   ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1332
MAC.US.x.r90131.AY576481   ....-------------...--------------------------------------------------------------------------------------------------------------------------------------------- 1475
```

```
                              Gag p1 spacer end Gag p6 start
MAC.US.x.239                TGGGGAAAGAAGCCCCGCAATTTCCCCATGGCTCAAGTGCATCAGGGGCTGATGCCAACTGCTCCCCCAGAGGACCCAGCTGTGGATCTGCTAAAGAACTACATGCAGTTGGGCAAGCAGCAGAGAGAAAAGCAGAGAGAA............................ 2510
Pol                         S  M  G  K  E  A  P  Q  F  P  H  G  S  S  A  S  G  A  D  A  N  C  S  P  R  G  P  S  C  G  S  A  K  E  L  H  A  V  G  Q  A  A  E  R  K  A  E  .  .  .  .  .  .  .  .  .
Gag                          W  G  K  K  P  R  N  F  P  M  A  Q  V  H  Q  G  L  M  P  T  A  P  P  E  D  P  A  V  D  L  L  K  N  Y  M  Q  L  G  K  Q  Q  R  E  K  Q  R  E  .  .  .  .  .  .  .  .  .
H2A.CI.88.UC2.U38293        C-----------------C-----G----C----CT-C-----------A-----A--A-------CA--T--G--A-C---C---T--G----A--TT-----CA---G-GAA-------G--GC--AG-GAGAG-CCATACAAGGAGGTGACGGAA...GACTT 2580
H2A.DE.x.BEN.M30502         C------------T-----C-----G--A-C----CC-C-----------A-----A--A--T--G-CA--T-----A-C---A---T-GG---GA--T-----CAA--G-GAA-------G--GC--AG-GAGAG-CCATACAAAGAGGTGACGGAG...GACTT 2580
H2A.DE.x.PEI2.U22047        -------------T-----C-----G--A-C-G---C-CG---------CA-----A--A-------CA--------A-CA--C------G----G---T-----CAA--G-G-A------A---GC--A-A-TGAG-CCATACAAGGAGGTGACAGAG...GACTT 2548
H2A.GH.x.GH1.M30895         ------------------C-----G----C----CT-C--C---------A-----A--A-------CA--T-----A--------AT-GG-A-GA--T------CA---G-GAG-------G--GC--AG-GAGAG-CCATACAAAGAGGTGACGGAA...GACTT 2024
H2A.IN.NNVA.NEW             C-----------------C-----G----C---A-C-CG--------A-CA-----A--A------AT---T-----A-----C--AT-GG----A--T------CAA--G-GAA-------G--GC--AGAGAGAG-CCATACAAGGAGGTGACAGAG...GACCT 2574
H2A.IN.95.CRIK              C------------T-----C-----G----C-----T-CA--------A-CA-----A--A--T---AT---T-----A-----C--A--GG----A--T------CAA--G-GAA-------G--GC--AGAGAGAG-CCATACAAGGAGGTGACAGAG...GACCT 2325
H2A.GM.x.MCR35.AY509260     C-----------------C-----G----C-----C-CG--------A-CA-----A--A-------TA--T-----A-----C--A--GG----A--T------CAA--A--AAGA--A------C--AGAGAGAG-CCATACAAAGAAGTGACAGAG...GACTT 2025
H2A.GM.87.D194.J04542       C--------C--------C-----GC---C----CT-C-----------A-----A--A------ATA--T-----A-----C---T-GG----A--T------CAA--G-GAA----------GC--AG-GAGAG-CCATACAAGGAGGTGACGGAG...GACTT 2024
H2A.GM.x.ISY.J04498         ------------------C-----G---TC-C-AGTTCG--------A-CA-----A--A------AT---T-----A-----C--A--GG----G---------CAA--G-GAA-A--------GC--AGACA-AG-CCATACAAAGAAGTGACAGAG...GACTT 2021
H2A.GM.x.MCN13.AY509259     C-----------------C-----G----C-----C-CG--------A-CA-----A--A-------TA--T-----A-----C--A--GG----T--T------CAA--A--AAGA--A------C--AGAGAGAG-CCATACAAAGAAGTGACAGAG...GACTT 2025
H2A.GW.86.FG.J03654         -----------G------C-----GC---C-----C-CG--------A-CA-----A--A------TT---T-----A-----C--A--GG----A--T------CAA--A--AAG---A--G--GC--AGAGAGAG-CCGTACAAGGAAGTGACAGAG...GACTT 2018
H2A.GW.87.CAM2CG.D00835     ------------------------G----C-----T-CA--------A-CA-----A--A------TT---T-----A---------A--GG----A--T------CAA--G--AAG------G--GC--AGACAGAG-CCATACAAAGAGGTGACAGAG...GACTT 2589
H2A.GW.x.ALI.AF082339       -------------T-----C-----G-AA-C-G---T-CG-----AT-A-CA-----A--A-------CA--G-----A-C---C-----GG--C-G--T------CA---G-GAA----A-----GC--AGAGAGAGGCCATACAAAGAGGTGACGGAG...GACTT 2573
H2A.GW.x.MDS.Z48731         ------------------C-----G----C--G--T-CG--------A-CA-----A--A-------CA--T-----A-----C--A--GG----A--T------CAA--G-GAA-------G--GC--AGAGAGAG-CCATACAAAGAGGTGACAGAG...GACTT 2024
H2A.SN.85.ROD.M15390        ------------------C-----G----C-----T-CG---------CA-----A--A-------T---T-----A---------A--GG----A--T------CAA--G--AAGA--------GC--AGAGAGAG-CCATACAAGGAAGTGACAGAG...GACTT 2023
H2AB.CI.90.7312A.L36874     -------------T-----C-----------C-------C-----A-A-AG-----T----G-----GAT-A--A----A-A--GCAA-AC-C-TC-GGGAGC-ATACCATCTGC--CC-CTGC---TCC-GCAGTG--GATGCTGAAAAGCTACATGCAACTAGGGAA 2562
H2B.CI.88.UC1.L07625        -------------T-----C---------A-C-------C------AG---CA---T----A-----GAT---------A-A--GCA--AC-CCTCGGGGGGC-ACACCATCTGC--CC-CTGC---TCCAGCAGTG--GATGCTGAAAAGCTACATGAAGATGGGGAG 2561
H2B.CI.x.EHO.U27200         -------------T-----------G-C...--G-CA-CC------A-AG-----T----G-----GAT-A-------AT-C-GCA--AC-CCTC-GGGTGC-ATTCCATCTGCA-CC-CTGC---TCCAGCAGAG--GATGCTAAAGAACTACATGCAACTAGGGAA 2551
H2B.GH.86.D205.X61240       -------------T-----C---------A-C-------C------AG---CA---T----A-----GAT-A-------A-A--GCA--AC-CCTCGGGGGGC-ACACCATCTGC--CC-CTGC---TCCAGCAGTG--GATGCTGAAAAGTTACATGCAGATGGGGAG 2556
H2B.JP.01.KR020.AB100245    -------------T-----C-----TG-A--C----CA-CC------A-AC-----T---A--A---AT-A-------A-A-A-CA--AC-CCTC-GGGTGCAAT-CCATCTGC--CC-CTGC---TCCAGCAGTG--GATGTTGAAGGACTACATGCAGCTAGGGAG 1702
H2AB.CM.03CM_510_03.NEW     -------------T-----C---------A-C--G----C------AG---CA---T----A-----GAT-A-------A-A-RGCA--AC-CCTC-GGGGG--ACACCATCTGC--CC-CTGC---TCCAGCAGTG--GATGCTGAAAAGCTACATGCAGATGGGGAA 1692
H2G.CI.92.ABT96.AF208027    -------------T-----C---------A-C-----C-C------T----CT---T-A--A--A--GAT---------A--A--C--A--G-----T--------RC----G-G-A----A-AG--G........................................ 1920
H2U.FR.96.12034.AY530889    ----A---A-----T-----C-----------C-----C-CG-----A----CC-----A--C--T----CA--A-----A----------G-CCCCGACAGCC-CTCCA-CGG-C-CAGCAGTG--TCT--TA-AGAGCTACATGCAACAGGGCAAGAAA........ 2066
SMM.SL.92.SL92B.AF334679    -------------T-----C----------...--GACA...ACATCAT---CA--GT-A--A--------T......C-A-CA-CC-G-A--GTC--GG-GTAT-TAGA-AAAGCA---AG-GA-A-G-CA.................................... 1941
MAC.US.x.MAC239             ----------------------------------------------------------------------------------------------------------------------------...............................  1587
MAC.US.x.251                ------------------------------------------C-----------------------------------------------------------------------A---------............................. 2509
MAC.US.x.251_32H            ------------------------------------------C--------------------------------------C-----------------------------------------............................. 2511
MAC.US.x.251                ------------------------------------------C-------------------------------------------------------------------------.................................... 2486
MAC.US.x.17EC1.AY033233     ----------------------------------------------------------------------------------------------------------------------------............................. 2510
MAC.US.x.17EFR.AY033146     ----------------------------------------------------------------------------------------------------------------------------............................. 2510
MAC.US.x.1937.AY611495      -------------------------------------------------------------------------------------------------------------------------R--G............................. 1616
MAC.US.x.2065.AY611493      ------------------------------------------Y------------------------------------------------------------------------------------............................. 1616
MAC.US.x.MM142.Y00277       ------------------------------------------C----------------A--A------------------------------C------------------.................................... 1996
MAC.US.x.SMM142B.BD131285   ------------------------------------------C----------------A--A------------------------------C------------------.................................... 1996
MNE.US.82.MNE               -------------------------------C---A--------------C-------------------------------------------A------------------A------------.............................. 1978
MNE.US.x.MNE027.U79412      -------------------------------C---A--------------CA------------------------------------------A------------------A------------.............................. 1978
SMM.US.x.F236               -------------------------------C--GA---C-----------A------------------T-------------A--G-----T------A--A-----GAA-------G--G..................................... 2497
SMM.US.x.H9.M80194          -------------------------------C--GA---C-----------CA----------Y-------T-------------A--G-----T------A--G-----GAAG------G--G..................................... 1983
SMM.US.x.PBJ14              -------------------------------C--GA---C-----------CA------------------T-------------A--G-----T------A--A-----GAAG------G--G..................................... 2277
SMM.US.x.PBJA.M31325        -------------------------------C--GA---C-----------CA------------------T-------------A--G-----T------A--A-----GAAG------G--G..................................... 2276
SMM.US.x.PBJ                -------------------------------C--GA---C-----------CA----------Y-------T-------------A--G-----T------A--G-----GAAG------G--G..................................... 1983
SMM.US.x.PBJ                -------------------------------C--GA---C-----------CA------------------T-------------A--G-----T------A--A-----GAAG------G--G..................................... 2469
SMM.US.x.PGM53.AF077017     -------------------------------C--GA---C-----------C-------------------T-------------A--G-G---T------A--A-----GAA-------G--G..................................... 2427
SMM.US.x.SME543.U72748      -------------------------------C--GA---C-----------CA------------------T-------------A--G-----T------A--A-----GAA-------G--G..................................... 2513
STM.US.x.STM.M83293         ---------------A---C-----------C---A-A-C-----------C---------C--T-----A----------C---------G-GA-GT---------C---------A----------G..................................... 2157
MAC.US.x.80035.AY611486     ----------------------------------------------------------------------------------------------------------------------------............................. 1616
MAC.US.x.81035.AY599200     ----------------------------------------------------------------------------------------------------------------------------............................. 1617
MAC.US.x.85013.AY611490     ----------------------------------------------------------------------------------------------------------------------------............................. 1550
MAC.US.x.92050.AY603959     ----------------------------------------------------------------------------------------------------------------------------............................. 1610
MAC.US.x.92077.AY599201     ----------------------------------------------------------------------------------------------------------------------------............................. 1617
MAC.US.x.93057.AY611492     ----------------------------------------------------------------------------------------------------------------------------............................. 1615
MAC.US.x.93062.AY607704     ----------------------------------------------------------------------------------------------------------------------------............................. 1602
MAC.US.x.95058.AY611494     ----------------------------------------------------------------------------------------------------------------------------............................. 1604
MAC.US.x.95086.AY607703     ----------------------------------------------------------------------------------------------------------------------------............................. 1473
MAC.US.x.95112.AY588946     ----------------------------------------------------------------------------------------------------------------------------............................. 1606
MAC.US.x.96016.AY607701     ----------------------------------------------------------------------------------------------------------------------------............................. 1617
MAC.US.x.96020.AY611488     ----------------------------------------------------------------------------------------------------------------------------............................. 1609
MAC.US.x.96072.AY611491     ----------------------------------------------------------------------------------------------------------------------------............................. 1621
MAC.US.x.96081.AY597209     ----------------------------------------------------------------------------------------------------------------------------............................. 1619
MAC.US.x.96093.AY611489     ------------------------------------------------------------------------------------------------------------------------A----............................. 1617
MAC.US.x.96114.AY588945     ----------------------------------------------------------------------------------------------------------------------------............................. 1612
MAC.US.x.96123.AY611487     ----------------------------------------------------------------------------------------------------------------------------............................. 1506
MAC.US.x.96135.AY607702     ----------------------------------------------------------------------------------------------------------------------------............................. 1602
MAC.US.x.97009.AY599199     ----------------------------------------------------------------------------------------------------------------------------............................. 1485
MAC.US.x.97074.AY599198     ----------------------------------------------------------------------------------------------------------------------------............................. 1485
MAC.US.x.r80025.AY576480    ----------------------------------------------------------------------------------------------------------------------------............................. 1473
MAC.US.x.r90131.AY576481    ----------------------------------------------------------------------------------------------------------A------------------............................. 1616
```

```
                                                
```

```
MAC.US.x.239          AGTTAGGTCCACATTATACCCCAAAAATAGTAGGAGGAATAGGAGGTTTTATTAATACTAAAGAATACAAAAATGTAGAAATAGAAGTTTTAGGCAAAAGGATTAAAGGGACAATCATGACAGGGGACACCCCGATTAACATTTTTGGTAGAAATTTGCTAACAGCTCTG 2833
Pol                   I__E__L__G__P__H__Y__T__P__K__I__V__G__G__I__G__G__F__I__N__T__K__E__Y__K__N__V__E__I__E__V__L__G__K__R__I__K__G__T__I__M__T__G__D__T__P__I__N__I__F__G__R__N__L__L__T__A_
H2A.CI.88.UC2.U38293  -------GGACA----C--T------G----G--G--------G--A--C--A--C--C--------T--G-----------A----AC--AAT-----AG-A-G--CC--C--A---------A--T-----A--C-----C-----C------A-TT-G-----CT-A 2916
H2A.DE.x.BEN.M30502   -A-----GGACA----C--T---------------G--------G--A-----A--C--C-------------------------A----AC--AAT-----AG-A-G--CC--C--A---------A--T-----A--C-----C-----C------A-T--G-----CT-A 2916
H2A.DE.x.PEI2.U22047  ----G--GAGCA------GT---------------G--------G--A--C--A--C--C--G-----T----------------A----AC--AAT----A-G-A---CC--C--A---------T--T-----A--C-----------C-----CA-T--G-----CT-A 2884
H2A.GH.x.GH1.M30895   -------GGACA----CGTT-----------G--G--------G--A--C--A--C--C-----GATT------A--------A----AC--AAT-----AG-A-G--CC--C--A---------A--T-----A--C-----C-----C------A-C--G-----CT-A 2360
H2A.IN.NNVA.NEW       -------AAGTG------AT---------------G--------G--A--C--A--------------T---------------------CC----T----A-G-A-GG-CC--C--A---------T--T-----A--C-----------C-----CA-TT-G-----CT-A 2910
H2A.IN.95.CRIK        -A-----AAGTA------GT---------------G--------G--A--C--A-----C--------T--C------------------CC----T----A-G-A-GG-CC--C--A---------T--T-----A--C-----------C-----CA-TT-G----TCT-A 2661
H2A.GM.x.MCR35.AY509260 -------GAGCA------GT-----G---------------T--G--A--C--A-----C--G-----T---------A--------------T----A-G-A-GG-CC--C--A---------T--------A--C-----------C------A-T--G-----CT-A 2361
H2A.GM.87.D194.J04542 -------GGACA----C--T-----------G--G--------G--A--C--A-----C--------T---------------A-G--AC--AAT-----AG-A-G--CC--C--A---------A--T-----A--C-----------C------A-T--GG--A-CT-A 2360
H2A.GM.x.ISY.J04498   -------GAGCA------GT---------------G-----------A--C--A-----C--------T---G-----------AG---GC--AAT----A-G-A-G--CC--C--A---------T--T-----A--C-----------C------A-C--G-----CT-- 2357
H2A.GM.x.MCN13.AY509259 -------GAGCA------GT-----G------------------G--A--C--A-----C--G-----T---------A---------------T----A-G-A-GG-CC--C--A---------T--------A--C-----------C------A-T--G-----CT-A 2361
H2A.GW.86.FG.J03654   -------GAGCA------GT------------------------G--A--C--A-----C--G-----T-------------------CC----T------G-A---CC--C--A---------T--T--------C-----------C------G-T--G------T-A 2354
H2A.GW.87.CAM2CG.D00835 -A-----GAGCA------G-----------G--G-----------A--C--A-----C--------T--------G------------C----T------G-A-GG-CC--T--A---------C--------A--C-----------C------A-T--G-----CT-A 2925
H2A.GW.x.ALI.AF082339 -------GAGCA-C------------------G--------G--A--C--A-----C-------TG--G-----------A----AC--AAT-----AG-A---CC--C--A---------T--------A--C--T--------C-----CA-TT-G-----CT-A 2909
H2A.GW.x.MDS.Z48731   -A-----AAGCA------GT---------------G--------G--A--C--A-----C--G-----T-------------------CC----T------G-A-GG-CC--C--A---------T--T-----A--C-----------C------A-TT-G-----CT-A 2360
H2A.SN.85.ROD.M15390  -------GAACA------G----------------G--------G--A--C--A-----C--G-----T---------------------C--AAT----A-G-ACGG-CC--C--A---------C--------A--C-----------C------A-T--G-----CT-A 2362
H2AB.CI.90.7312A.L36874 -A-----AAGCA----C------------------G--------G--C--A--------G-----------CA-------------AG----A----A-G-A-GGTCA--T--A---------A--------A--A--T--------C------A-CT---ACA-CT-A 2898
H2B.CI.88.UC1.L07625  -A------AGCA----C---------------G--G--------G--C--A-----C------------G----------------AG-G--A-----AG-A-GG-CA--T--A---------A--------A--A-----------C------A-TT---ATA-AT-A 2897
H2B.CI.x.EHO.U27200   -A-----CAGCA----C---------------T--G--------A-----A-----C--T-------------------------AG----A-----AG-A-G--CA---G-A------------------A--A-----------C------A-TT---ATAGCT-A 2887
H2B.GH.86.D205.X61240 -A------AGCA----C------------------G--------G--C--A-----C------------G----------------AG-G--A-----AG-A-GG-CA--T--A---------A--T-----A--A-----------C------A-TT---ATA-CT-- 2892
H2B.JP.01.KR020.AB100245 -A------AGTG----C---------------------------A--C--A-----------G-----------------------AG----A----A---A-G--CA--T--A---------A--T-----A--A-----------C------A-T----ATA-CT-- 2038
H2AB.CM.03CM_510_03.NEW -AY-----AGCA----C---------------G--R--------G-----A-----C-----------G----------------AG-G--A----CAG-A-GG-CA--T--A---------A--------A--A-----------C------A-TT---ATA--T-R 2028
H2G.CI.92.ABT96.AF208027 -A-----AAGTA------GT-------------------------A----------C--------------------TA---GG---A-----A--AC-G-CT--TG-T-----------T-----T--A--T--C-----A--G--CA-TT--GTTAAG--A 2243
H2U.FR.96.12034.AY530889 -A----AA---T-------T-------CT-----G-A-----A---G--C--------A---A---TTC-C------A-C---A----A--G-------CA----G---T--C--A---------A--------A--C--T-----------G---C-TT-G----AA--- 2398
SMM.SL.92.SL92B.AF334679 -----------GGG--C--A----G---T-----G--G--T-----A--C--A-----A--------T-G--CA---A----------------A--GGTA--A-G--AA---C-A---------A--T---------A--T--------C------A-T---G----AT-A 2264
MAC.US.x.MAC239       ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1910
MAC.US.x.251          -----------------------------------------------------------------------------------------A---G----------------------------------------------------A----------- 2832
MAC.US.x.251_32H      ---------T--------------------------------------------------------------------------------------------------------------------------------------------A----------- 2834
MAC.US.x.251          ------------------------------------------------------------A-----------------------------------------T-----------------G------------------ 2809
MAC.US.x.17EC1.AY033233 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2833
MAC.US.x.17EFR.AY033146 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2833
MAC.US.x.1937.AY611495 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1939
MAC.US.x.2065.AY611493 ---------Y---------------------------------------------------------------------------------------------------------------------------------------------------------- 1939
MAC.US.x.MM142.Y00277 --------------------------------------------------------------------------------------------------------------------------------------------------A----------- 2319
MAC.US.x.SMM142B.BD131285-------------------------------------------------------------------------------------------------------------------------------------------------A----------- 2319
MNE.US.82.MNE         ------------------------------------------------------------A----------------------------------------------------------------------------------------------- 2301
MNE.US.x.MNE027.U79412 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2301
SMM.US.x.F236         -A--G------A----------T---------------------C--C-----C--C--------T---G-----A-----A----C--------GGTA-----G--A-----T-----G--A--T-----A-----T---------C------------------A-- 2820
SMM.US.x.H9.M80194    -A--G------A----------T---------------------C--C-----C--C--------T---G-----A-----A----A---------GTA-----G--A-----T---------A--T-----A-----------------C------C-----------A-- 2306
SMM.US.x.PBJ14        -A--G------A----------T---------------------C--C-----C--C--------T---G-----A-----A----A---------GTA-----G--A-----T---------A--T-----A-----T-----------C------C-----------A-- 2600
SMM.US.x.PBJA.M31325  -A--G------A----------T---------------------C--C-----C--C--------T---G-----A-----A----A---------GTA-----G--A-----T---------A--T-----A-----T-----------C------C-----------A-- 2599
SMM.US.x.PBJ          -A--G------A----------T---------------------C--C-----C--C--------T---G-----A-----A----A---------GTA-----G--A-----T---------A--T-----A-----T-----------C------C-----------A-- 2306
SMM.US.x.PBJ          -A--G------A----------T---------------------C--C-----C--C--------T---G-----A-----A----A---------GTA-----G--A-----T---------A--T-----A-----T-----------C------C-----------A-- 2792
SMM.US.x.PGM53.AF077017 -AC--------A----------T---------------------C--C-----C--C-----G--T---G-----A-----A----C--------GGTA-----G--A-----T---------A--T-----A-----------------C------C--T--------A-- 2750
SMM.US.x.SME543.U72748 -A--G------A----------T---------------------C--C---------C--------T---G----GA-----A----C--------GGTA-----G--A-----T-----G--A--T-----A-----T-----------C------------------A-- 2836
STM.US.x.STM.M83293   --C----G-T---A--C--------GG-------T-----------A-----A--------G----TT--G------A-C-----G--A--------G-AA--A-----A--T--T-----T--A--------C--------------C--G--CC-AT-G------T-- 2480
MAC.US.x.80035.AY611486 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1939
MAC.US.x.81035.AY599200 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1940
MAC.US.x.85013.AY611490 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1873
MAC.US.x.92050.AY603959 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1933
MAC.US.x.92077.AY599201 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1940
MAC.US.x.93057.AY611492 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1938
MAC.US.x.93062.AY607704 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1925
MAC.US.x.95058.AY611494 -----------Y-------------------------------------------------------------------------------------------------------------------------------------------------------- 1927
MAC.US.x.95086.AY607703 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1796
MAC.US.x.95112.AY588946 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1929
MAC.US.x.96016.AY607701 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1940
MAC.US.x.96020.AY611488 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1932
MAC.US.x.96072.AY611491 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1944
MAC.US.x.96081.AY597209 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1942
MAC.US.x.96093.AY611489 -----------Y-------------------------------------------------------------------------------------------------------------------------------------------------------- 1940
MAC.US.x.96114.AY588945 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1935
MAC.US.x.96123.AY611487 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1829
MAC.US.x.96135.AY607702 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1925
MAC.US.x.97009.AY599199 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1808
MAC.US.x.97074.AY599198 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1808
MAC.US.x.r80025.AY576480 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1796
MAC.US.x.r90131.AY576481 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1939
```

HIV-2/SIV Com-
plete Genomes

```
                               Pol Protease end_Pol p66 and p51 RT start
MAC.US.x.239             GGGATGTCTCTAAATTTTCCCATAGCTAAAGTAGAGCCTGTAAAAGTCGCCTTAAAGCCAGGAAAGGATGGACCAAAATTGAAGCAGTGGCCATTATCAAAAGAAAAGATAGTTGCATTAAGAGAAATCTGTGAAAAG...ATGGAAAAGGATGGTCAGTTGGAGGAAGC 3000
Pol                      _L__G__M__S__L__N__F__P__I__A__K__V__E__P__V__K__V__A__L__K__P__G__K__D__G__P__K__L__K__Q__W__P__L__S__K__E__K__I__V__A__L__R__E__I__C__E_.._K__M__E__K__D__G__Q__L__E__E_
H2A.CI.88.UC2.U38293     --C-----AT-------A--AG-C----GGA-------AA-----A-AA-A--G-----------A--C-------GGC----A--A-----CC--A----------AG---AA-------AG--G-----------A...--------A--G--C--AC-A--A--G-- 3083
H2A.DE.x.BEN.M30502      --C-----AT-------A--AG-T--C--GA-------AA-------AA-A--G--------G--A----------GGC----A--A-----CC--A-------G--A----AA---C---A---G-----------A...--------A--G--C---C-A--A--G-- 3083
H2A.DE.x.PEI2.U22047     --C-----AT-----C-A--AG-C--C--G-----C--GA-------AATAC-G--A--------A------------G-A-GA--A-----TC--A---------------AG---C---A----------------...-------GA--A--C---C-A-------- 3051
H2A.GH.x.GH1.M30895      --C-----AT-------A--A--C--C--GA----A--AA-------AA-G--G--------G--A----------G-C---GA--A-----CC--A-------G--A----AA---C-------G-----C-----A...-----------G--C--AC-A--A--G-- 2527
H2A.IN.NNVA.NEW          --C------T----CC-A--AG-C--C--GA----A--AA-------AAAGC----A---------------------C---GA--------C---A----------A----AG-------AG---------------...-------GA--G--C---C-A-------- 3077
H2A.IN.95.CRIK           --C-----AT-----C-A---G-C--C--GA----A--AA-----A-AATGC----A---------------------C---GA--------CC--A-C--------A----AA-------A---G-----------A...-----------G--A---C-A-------- 2828
H2A.GM.x.MCR35.AY509260  --C-----AT-------A--AG-C--C---A----A--AA-----A-AATG--------------A------------C---G---A-----C---A----------A----AG---C---A---------------A...-------GA--A--C---C-A-------- 2528
H2A.GM.87.D194.J04542    --C-----AT----CC-A--AG-C--C--GT----C--AA-------AA-A--G--------G--A----------GGC----A--A-----CC--A----------A----AA---C---A------T--------A...-------G---G--C--AC-A--A----- 2527
H2A.GM.x.ISY.J04498      --C-----AT-------A--AG-C--C---A----A--A--------AA-A--------------A-----G------CAA-GA--A-----C---A---G------A----AA---C-------------------A...-------GA--A------C-A--A----- 2524
H2A.GM.x.MCN13.AY509259  --C-----AT-------A--AG-C--C---A----A--AA-----A-AATG--------------A------------C---G---A-----C---A----------A----AG---C---A---------------A...-------GA--A--C---C-A-------- 2528
H2A.GW.86.FG.J03654      --C-----AT----CC-G--AG-T--C--GA----A--AA-----A-AATGC----------G--A----------G-C-A--A--A-----T---A----------A----AA---C---A---------------A...--------A--A--C---C-A--A----- 2521
H2A.GW.87.CAM2CG.D00835  --C-----AT-----C-A--AG-C--C--GA----A--AA-----A-AATGC-------------A----------G-----GA--A-----C--GA----------A----AA---C---A---G-----------A...--------A--A--C---C-A-------- 3092
H2A.GW.x.ALI.AF082339    --C-----AT----CC-A--AG-T--C--GA-------AA--G-G--AAGA--------------A--C--G--------A-GA--A-----C---A----------A----AG---C---A---------------A...-C-----GA--A--C--A--A-----G-- 3076
H2A.GW.x.MDS.Z48731      --C-----AT-----C-A--AG-C--C--GA----A--AA-----A-AATGC----------G---------------C---GA--------T---A----------A----AA---C---A---G-----------A...-------GA--A--C---C-A-------- 2527
H2A.SN.85.ROD.M15390     --C-----AT-----C-A--AG-C--C-----------AA-----A-AATGC----------G--A------------C---GA--A-----C---A----------A----AA---C---A---------------A...--------A--A--C---C-A-------- 2529

H2AB.CI.90.7312A.L36874  --C---A----------C--AG-G--A---A-------A--G--G---CAG--------T------------------A-C-GA--A-----TC----T------------C-G--CC-T-----G-----------A...-----------G--G---C-A--A----- 3065

H2B.CI.88.UC1.L07625     --C---A----------C--AG----A--GA----A--A---------AAAC-------T--G--A------------A-C-GA--A-----TC----C------------T-G--CC-C-AG---------------...-----------G--A---C-A--A----- 3064
H2B.CI.x.EHO.U27200      --C---A----------C--AG----A-GGA----A--A---------CAG--------T-A---A-----G------A-C-GA--A-----CC----C-----G--A---C-A--CC-C-A---------------A...--------A--G--A-----A--A--G-- 3054
H2B.GH.86.D205.X61240    --C---A--T-------C--AG-G--A--G-----A--A--------T-AG-----A--T-----A-----G-----GA-C-GA--A-----TC----C-GG---------C-A--CC-C-A---------------A...-----------G--A--A--A--A----- 3059
H2B.JP.01.KR020.AB100245 --C---A----------C--AG----A---A----A--A-----G---AAAC----A--T-----A------------A-C-GA--A-----TC----T------------C-G--CC-T-A---G-----------A..A--------A--G--A--AC-A-------- 2206

H2AB.CM.03CM_510_03.NEW  --C---A----------C--AG-G--A-RGR----A--A--------T-AG--R--A--T-----R-----R-----GA-C-GA--A-----TC----C-GG--R------C-A--CC-Y-A---------------A...-----R-----G--G--A--A--A----- 2195

H2G.CI.92.ABT96.AF208027 --C-----------------A--T--A--G--------A--------AAGGC----A--T--C-T------G------A-C-G---A-----CC----T--G-----A---CAA---C-C--G-----T-----C--A...------C----A--G--A-----A----- 2410

H2U.FR.96.12034.AY530889 --C-----------C-----A---GA-----A--CA-TA--C-TT-GG-GAC----A-----C-----------------A-GA--------T-----T--G-----AT---AA-------A---G--A-----G--A...-----------G--A--A-----A----- 2565

SMM.SL.92.SL92B.AF334679 ---G-C---T------AC--AG-G--A--G-----ATA-AC------AAAG--G--GAG-----T------G----G---A--A--A-----C-----T--G---------CAA-------C---G-----C--G--A...-----------G--A--A-----ACG--- 2431

MAC.US.x.MAC239          -----------------------------------------------Y------------------------------------------------------------------------------------------...----------------------------- 2077
MAC.US.x.251             ---------------C-------------------------------------------------A-T----------------------------------------------------------------------...----------------------------- 2999
MAC.US.x.251_32H         ---------T----CC------------------------------------------------------------------------------------------------------------------------...----------------------------- 3001
MAC.US.x.251             ---------------C-------------G------------------A------------------T----------------------------------------------------------------------...----------------------------- 2976
MAC.US.x.17EC1.AY033233  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 3000
MAC.US.x.17EFR.AY033146  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 3000
MAC.US.x.1937.AY611495   ------------------------------------------------------------------A-----------------------------------------------------------------------...----------------------------- 2106
MAC.US.x.2065.AY611493   ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 2106
MAC.US.x.MM142.Y00277    ---------------C-------------G--------------GTCGC-----------------------------------------------------------------------------------------...--------A-------------------- 2486
MAC.US.x.SMM142B.BD131285---------------C-------------G--------------GTCGC-----------------------------------------------------------------------------------------...--------A-------------------- 2486

MNE.US.82.MNE            -A---------------C-----------G------------------A----------------A----------------G-------------------------------------------------------...----------------------------- 2468
MNE.US.x.MNE027.U79412   --A--------------C-----------G------------------A----------------A----------------G-------------------------------------------------------...----------------------------- 2468

SMM.US.x.F236            --C------T-----C-C-----------G--G------A-------AA-AC----A--------A--A-----------A-GA--------GC-----------------A-------------------------A...--------A-----C-----A-------- 2987
SMM.US.x.H9.M80194       --C------------C-C-----------G--G--R---A-------AA-AC----A--------A--------------A-GA--------GC----G------------A-------------------------A...-----G--A-----C-----A----R--- 2473
SMM.US.x.PBJ14           --T------------C-C-----------G---------A-------AA-AC----A--------A--------------A-GA--------GC-----------------A-------G-----------------A...-----G--A-----C-----A-------- 2767
SMM.US.x.PBJA.M31325     --T------------C-C-----------G---------A-------AA-AC----A--------A--------------A-GA--------GC-----------------A-------G-----------------A...-----G--A-----C-----A-------- 2766
SMM.US.x.PBJ             --C------------C-C-----------G--G--R---A-------AA-AC----A--------A--------------A-GA--------GC-----------------A-------------------------A...-----G--A-----C-----A----R--- 2473
SMM.US.x.PBJ             --T------------C-C-----------G---------A-------AA-AC----A--------A--------------A-GA--------GC-------G---------A-------G-----------------A...-----G--A-----C-----A-------- 2959
SMM.US.x.PGM53.AF077017  --C-----CT-----C-C---G-------G--G------A-------AA-AC----A--------A--C-------------GA--------GC-------------A-----------------------------A...--------A-----C-----A-------- 2917
SMM.US.x.SME543.U72748   --C------T-------C-----------G--G------A-------AA-AC----A-----G--A--A-------------GA--------GC-----------------A-------------------------A...--------A-----C-----A-------- 3003

STM.US.x.STM.M83293      --------AT----------AG-------------A--A-----G--GA-A-----A--------A------------A-A--A--A-------------------------AG---C---A---G-----------A...-----G--------G---C-A--A----- 2647

MAC.US.x.80035.AY611486  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 2106
MAC.US.x.81035.AY599200  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 2107
MAC.US.x.85013.AY611490  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 2040
MAC.US.x.92050.AY603959  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 2100
MAC.US.x.92077.AY599201  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 2107
MAC.US.x.93057.AY611492  --------------------------------------------G-----------------------G---------------------------------------------------------------------...----------------------------- 2105
MAC.US.x.93062.AY607704  ---------------Y---Y------------------------------------------------K---------K-----------------------------------------------------------...----------------------------- 2092
MAC.US.x.95058.AY611494  --------------------------------------C-----------------------------A---------------------------------------------------------------------...----------------------------- 2094
MAC.US.x.95086.AY607703  --------------------------------------------------------------------G---------------------------------------------------------------------...----------------------------- 1963
MAC.US.x.95112.AY588946  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 2096
MAC.US.x.96016.AY607701  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 2107
MAC.US.x.96020.AY611488  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 2099
MAC.US.x.96072.AY611491  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 2111
MAC.US.x.96081.AY597209  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 2109
MAC.US.x.96093.AY611489  ---------------C--------------------------------------------------------------------------------------------------------------------------...----------------------------- 2107
MAC.US.x.96114.AY588945  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 2102
MAC.US.x.96123.AY611487  --------------------------------------------------------------------G---------------------------------------------------------------------...----------------------------- 1996
MAC.US.x.96135.AY607702  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 2092
MAC.US.x.97009.AY599199  ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 1975
MAC.US.x.97074.AY599198  ---------------C--------------------------------------------------------------------------------------------------------------------------...----------------------------- 1975
MAC.US.x.r80025.AY576480 --------------------------------------------------------------------G---------------------------------------------------------------------...----------------------------- 1963
MAC.US.x.r90131.AY576481 ------------------------------------------------------------------------------------------------------------------------------------------...----------------------------- 2106
```

```
MAC.US.x.239                 TCCCCCGACCAATCCATACAACACCCCCACATTTGCTATAAAGAAAAAGGATAAGAACAAATGGAGAATGCTGATAGATTTTAGGGAACTAAATAGGGTCACTCAGGACTTTACGGAAGTCCAATTAGGAATACCACACCCTGCAGGACTAGCAAAAAGGAAAAGAATTA 3170
Pol                          __A__P__P__T__N__P__Y__N__T__P__T__F__A__I__K__K__K__D__K__N__K__W__R__M__L__I__D__F__R__E__L__N__R__V__T__Q__D__F__T__E__V__Q__L__G__I__P__H__P__A__G__L__A__K__R__K__R__
H2A.CI.88.UC2.U38293         A--T--A--T-----T--T--T--------------A--T-----------C--------------G-----A-----------A--------C-----G-----A--T--C--A---A-T--GC-------T--------A------T----C----A------G---- 3253
H2A.DE.x.BEN.M30502          A--T--A--T-----T--T--T--------------A--T-----------C--------------G-----------------A----------A---G-----A--T--C--A---A-T--GC-------T--------G-----------C----A------G--CT 3253
H2A.DE.x.PEI2.U22047         ------A--T-----T--T--T--------------A--T-----------C--A-----------------A-----------A----------A---A-----A--G--C--A---A-T--G--------T--------A------T----C--G-AA-G-------- 3221
H2A.GH.x.GH1.M30895          A--T--A--T-----T--T--T--------------A--T-----------C--------------G-----A-----C-----A--------C-----G-----A--T--C--A---A-T--GC-------T--------G------T----C----A------G--C- 2697
H2A.IN.NNVA.NEW              A--T--A--T-----T--T--T--------------A--C-GA--------C-----T--------G-----A--------C--A--------C-A---T--------G--C--A---A-T--GC-------T--------G------T-G--C--G-AA-G------C- 3247
H2A.IN.95.CRIK               A--T--A--T--C--T--T--T--------------A--C-----------C-----T--------G-----A--------C-----------C-A---A--------A--C--A---A-T--G--------C--------A------T-G--C--G-A--G------C- 2998
H2A.GM.x.MCR35.AY509260      G--T--A--T-----T--T-----------------A--C-----------C--A--T--------G-----A-----------A--------C-A---A-----A--T--C--A---A-T--G--------T--------A------T-G--C--G-AA---------- 2698
H2A.GM.87.D194.J04542        A--T--A--T-----T--T--T--------------A--T-----------C--------------------A-----------A--------C-----G-----A--T--C--A---A-T--GC-------T--------G------T----C----A------G---- 2697
H2A.GM.x.ISY.J04498          G--T--A--T-----C--T--T-----T--------A--T-----------C--A-----------G-----------------A--------C-A---A-----A--T--C--A--G--T--G--------T--------A------T----C--G-AA-G-------- 2694
H2A.GM.x.MCN13.AY509259      G--T--A--T-----T--T-----------------A--C-----------C--A--T--------G-----A-----------A--------C-A---A-----A--T--C--A---A-T--G--------T--------A------T-G--C--G-AA---------- 2698
H2A.GW.86.FG.J03654          A--T--A--T-----T--T--T--------------A--C-----------C--A-----------------A-----C--------G-----C-A---A--------T--C--A---A-T--G--------T--------A------T-G--C----A--GG------- 2691
H2A.GW.87.CAM2CG.D00835      A--T--A--T-----T--T--T--------------A--C-G---------C--------------G-----A-----------A--G-----C-AA--A--C--A--T--C--A---A----G--------T--------A------T----C----A--G-------- 3262
H2A.GW.x.ALI.AF082339        A--T--A--T--C--T--T--T--------------A--------G-----C--A-----------------A-----------A--T-----C-A---A-----A--T--C--A--GA-T--G-----G--T-----T--A------T----C--G-AA-G------C- 3246
H2A.GW.x.MDS.Z48731          A--T--A--T-----T--T--T--------------A--C-G---------C--A--T--------G-----A-----------A-----T--C-A---A--------T--C--A---A-T--G--------T--------A--G--------C--G-A--G-------- 2697
H2A.SN.85.ROD.M15390         A--T--A--T-----T--T--T--------------A--C-----------C--A-----------G-----A--------C--A--------C-A---A-----A--T--C--A---A-T--G--------T--------A-----GT-G--C--G-A--G-------- 2699
H2AB.CI.90.7312A.L36874      AT-T--C--T--------------A-----C-----------------------A-----------------A-----C--C------T------A---A--C--A--A-----A---------C----T--T--C-----G-----G--G---G---AA-GG-----A- 3235
H2B.CI.88.UC1.L07625         G-----T--T--------------A-----C-----C-----A--G-GA-----A-----------------A-----------A---T----C-A---G--C--A--------A---------C-G--T--T--C-----G-----GT-G---G---AA-GG-----A- 3234
H2B.CI.x.EHO.U27200          G--T--T--T-----------TT-G-----C--C--C-----A--G--A--C--A-----------G-----A--------C--A--------C-A---A--C--A--A-----A--G-----GC-G--T--T--T----A--------G---TC---A---------A- 3224
H2B.GH.86.D205.X61240        A-----T--T--------------A-----C-----C--------G--A-----------------------A-----C-----A--GT----C-A---A--C--------C--A------A-C-GG-T-T-C--C...A-CCGGCAGG-G---G---AA-GG-----A- 3226
H2B.JP.01.KR020.AB100245     G-----T--T-----C--------A-----C--C--C-----A--G-GA--C--A--------------T--------------A---T----C-A---A--C--A--------A--------GC-G--C--T--C-----G-----G--G---G----A--------A- 2376
H2AB.CM.03CM_510_03.NEW      A-----T--T--------------A-----C-----C--------G--A--C--A-----------------A-----C-----R---T----C-A---A--C--------C--A--------R--G--T--C--C-----G-----GT-G---G----A-GG-----A- 2365
H2G.CI.92.ABT96.AF208027     A-----C--------G--T-----A-----C---A----C-----G-----C--A---------------T-A-----C--C--A-----T--C-----A--C-----------T--------G--G--C-----T-----------TT-G---G---AA-GG-----A- 2580
H2U.FR.96.12034.AY530889     ------C--T--------------T-----C-----C--------G-----C--A---------------T-A-----C-----A--GT----C--A--A-----A--T----GT--------G-----T-----C-----A-----------TG-G-AA--G--G--A- 2735
SMM.SL.92.SL92B.AF334679     C--T--T--A-----T--T--T--A--G--C--------T---------A-----G----G-----------C--------C--A---T----C-AAA-G--A--A--A--C--A-----A--GC-G--T--C-----T--G-----C---AAGG---T-G----G--A- 2601
MAC.US.x.MAC239              ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2247
MAC.US.x.251                 ------------------------------------------------A-----------------------------------------------------------------A--------------------------------------------------G---- 3169
MAC.US.x.251_32H             ------------------------------------------------A-----------------------------------------------------------------A--------------------------------------------------G---- 3171
MAC.US.x.251                 ------------------------------------C-----------A-----------------------------------------------------------------A--------------------------------------------------G---- 3146
MAC.US.x.17EC1.AY033233      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3170
MAC.US.x.17EFR.AY033146      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3170
MAC.US.x.1937.AY611495       ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2276
MAC.US.x.2065.AY611493       ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2276
MAC.US.x.MM142.Y00277        ------------------T-----------------------------------A--------------------------------------------------A-----------------------------------------------------------G---- 2656
MAC.US.x.SMM142B.BD131285    ------------------T-----------------------------------A--------------------------------------------------A-----------------------------------------------------------G---- 2656
MNE.US.82.MNE                ------------------------------------C--------------C-----------------A-------------------------A------------------A-----------G-----------------------------------G--G--C- 2638
MNE.US.x.MNE027.U79412       ---------------------------------------------------C-------------------------------------------A------------------A-----------------------------------------------G--G--C- 2638
SMM.US.x.F236                C--T--A--------G--T-----------T--------------G--A--C--A--T--------G-----A-----------A---T-G----A---------A--------A-----A--G--------------------------------G--A-GG--G--C- 3157
SMM.US.x.H9.M80194           ---T--A-----------T------------------------R-GR-A--CGGA--T--------G-----A-----------A---T----C-A---------A--T-----A-----A--G-----G-----------R--------------G--A-G---G--C- 2643
SMM.US.x.PBJ14               C--T--A-----------T--------T-----------------G--A--C-----T--------G-----A-----------A---T----C-A---------A--T-----A-----A--G-----G--------------------------G--A-G---G--C- 2937
SMM.US.x.PBJA.M31325         C--T--A-----------T--------T-----------------G--A--C-----T--------G-----A-----------A---T----C-A---------A--T-----A-----A--G-----G--------------------------G--A-G---G--C- 2936
SMM.US.x.PBJ                 C--T--A-----------T-------TT---------------R-GR-A--C-GA--T--------G-----A-----------A---T----C-A---------A--T-----A-----A--G-----G-----------R--------------G--A-G---G--C- 2643
SMM.US.x.PBJ                 C--T--A-----------T--------T-----------------G--A--C-----T--------G-----A-----------A---T----C-A---------A--T-----A-----A--G-----G--------------------------G--A-G---G--C- 3129
SMM.US.x.PGM53.AF077017      C--T--A-----------T-----------------------A--G--A--C--A--T--------G-----A-----------A---T-G--------T-----A--T-----A-----A--G-----G--------------------------G----G------C- 3087
SMM.US.x.SME543.U72748       C--T--A--------G--T-----------T--------------G--A--C--A--T--------------A-----------A---T------A---------A--------A-----A--G--------------------------------G--A-GG--G--C- 3173
STM.US.x.STM.M83293          ---T--A--A-----T-----T--------T--------------------C--A-----G-----G--------------C--A----------AA--G--A--A--T-----A---A-T--GC----T--------------------------G----G------C- 2817
MAC.US.x.80035.AY611486      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2276
MAC.US.x.81035.AY599200      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2277
MAC.US.x.85013.AY611490      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2210
MAC.US.x.92050.AY603959      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2270
MAC.US.x.92077.AY599201      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2277
MAC.US.x.93057.AY611492      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2275
MAC.US.x.93062.AY607704      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2262
MAC.US.x.95058.AY611494      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2264
MAC.US.x.95086.AY607703      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2133
MAC.US.x.95112.AY588946      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2266
MAC.US.x.96016.AY607701      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2277
MAC.US.x.96020.AY611488      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2269
MAC.US.x.96072.AY611491      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2281
MAC.US.x.96081.AY597209      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2279
MAC.US.x.96093.AY611489      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2277
MAC.US.x.96114.AY588945      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2272
MAC.US.x.96123.AY611487      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2166
MAC.US.x.96135.AY607702      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2262
MAC.US.x.97009.AY599199      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2145
MAC.US.x.97074.AY599198      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2145
MAC.US.x.r80025.AY576480     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2133
MAC.US.x.r90131.AY576481     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2276
```

```
MAC.US.x.239                CAGTACTGGATATAGGTGATGCATATTTCTCCATACCTCTAGATGAAGAATTTAGGCAGTACACTGCCTTTACTTTACCATCAGTAAATAATGCAGAGCCAGGAAAACGATACATTTATAAGGTTCTGCCTCAGGGATGGAAGGGGTCACCAGCCATCTTCCAATACACT 3340
Pol                         I__T__V__L__D__I__G__D__A__Y__F__S__I__P__L__D__E__E__F__R__Q__Y__T__A__F__T__L__P__S__V__N__N__A__E__P__G__K__R__Y__I__Y__K__V__L__P__Q__G__W__K__G__S__P__A__I__F__Q__Y_
H2A.CI.88.UC2.U38293        -T-----A---G----G-----C--C--T--------A---T-------T-------C---T-----A-----CC-------------C----T---A---------A-------A--------CT-A--A--A--G--------A--------A-----T-----T--A 3423
H2A.DE.x.BEN.M30502         -TA--T-A---G----G-----C-----T--------A---C-------T-----------T-----A-----CC-----G-------C---ATG--A---------A----T--A-----A--CT----A--A-----------A--------A--T--T--------A 3423
H2A.DE.x.PEI2.U22047        -T-----A--------G-----C--C--T--------A---C----G--C-----A--A--T-----A------C-----A----G--C--------A--------GA----T--A-----A--C--A--A-----------A--A--G-----A--T--T---C----A 3391
H2A.GH.x.GH1.M30895         -T-----A---G----G-----C--C--T--------G---C-------T-----------T-----A-----CC-------------C--------A---------A----T--A--------CT-A--A--A-----------A--------A--T--T---C----A 2867
H2A.IN.NNVA.NEW             -T-----A---G----G-----T--C--T--------A--GC----G--C-----A-----------A-----C--G--G--------C--------A---------A----T--A--C-----CT----A-----------------------A--T--T--------A 3417
H2A.IN.95.CRIK              -T-----A---G----G-----T--C--T--------A--GC-------C-----A-----T-----A--------G--------G--C--------A---------A----T--A-----A--CT----A--------------A--------A--T--T--------A 3168
H2A.GM.x.MCR35.AY509260     -T-----A--------G-----T--C--T--------A---C-------C-----A-----T-----A---------------A----C--------A---------A----T--A--------C--------------------A--------A--T--T--------A 2868
H2A.GM.87.D194.J04542       -T-----A---G----G-----C--C--T--------A---C-------T--------A--T-----A-----CC-------------C-------------A----A----TG-A--------CT-A--A--A--------A--A--------A-----T-----T-TG 2867
H2A.GM.x.ISY.J04498         -T--GT-A---G----A-----C--C--T--------C---T----G--T-----A-----T-----A------C-G-----------C--------A---------A----T--A--C--A--CT-A--A-----------------------A--T--T--------A 2864
H2A.GM.x.MCN13.AY509259     -T-----A--------G-----T--C--T--------A---C-------C-----A-----T-----A---------------A----C--------A---------A----T--A--------C-----------------------------A--T--T--------A 2868
H2A.GW.86.FG.J03654         -T-----A---G----G-----T--C--T--------A---C----G--C-----A-----T-----A--C---C-------------C--------A---------A----T--A--C--A--C--A--A-----G-----------------A--T--T--------A 2861
H2A.GW.87.CAM2CG.D00835     -T-----A---G----G-----T-----T--------AT--C-------C-----A-----------A------C----------G--C--------A---------A-------A-----A--CT-A--G-----G-----A--A--------A--T--T--G-----A 3432
H2A.GW.x.ALI.AF082339       -T--G------G----G-----T--C--T--------A--GC----GAGC-----A-----T-----A------C-------------C--------A---------A----T--A-----A--CT-A--G--------------A--------A--T--T---C----A 3416
H2A.GW.x.MDS.Z48731         -T-----A---G----G-----T--C--T--------A--GC-------C-----A-----------A-----CC----------G--C--------A---------A----T--A-----A--CT----A--------------A--------A--T--T--------A 2867
H2A.SN.85.ROD.M15390        -T-----A---G----G-----T--C--T--------A---C----G--C-----A-CA--T-----A------C----------G--C--------A---------A-------A-----A--CT----A--------------A--------A--T--T---C----A 2869
H2AB.CI.90.7312A.L36874     ----GA-A---G----A-----C--C---AGTG-C--A-----CCC---C--C-----A--T--A--A--C-----G------A-----------------------A-G-----------A--C--A--A--------------A--C-----A--T--T------T-C 3405
H2B.CI.88.UC1.L07625        -------A---G-G--A-----C--C---AGT--C--A------CC-A-C--C--A-----T--A--A--C--C--G------A----C-----------------GA-------------A-----A--A--A--------------C-----A--T--T------T-C 3404
H2B.CI.x.EHO.U27200         -------A---G----A-----C--C---AGTG-C--A------CC---C--C--A--A--T--A--A--------G---G----------------A--------GA----TC----C--A--C--A--A--------------A--C-----A--T-----G-----C 3394
H2B.GH.86.D205.X61240       -----A-A---G-G--A-----C--C---AGT--C--A------CC-A-C--C--A-----T--A--A--C--C--G-----------C-----------------GA-------------A-----A--A--A--G-----------C-AGT-A--T-GT------T-C 3396
H2B.JP.01.KR020.AB100245    -------A--------A-----C--C---AGT--C--A------CC---T--C--A--A-----A--A-----C--G--------G-----------A---------A-------C--------C--A--A--A--G-----A-----C-----A--T--T------T-C 2546
H2AB.CM.03CM_510_03.NEW     -----A-A---G-G--A--C--C--C---AGT--C--A------CC-R-Y--C--A-----T--A--A--C--CY-G------AC---C--YA-------------GA-------------A-----A--A--R-----------R--C-----A--T--T------T-C 2535
H2G.CI.92.ABT96.AF208027    ----------------G-----C--C---AG---T----------T---T--C-----------A--T--C-----------------C--------A---------A-------C--C-----C--A--G--A-----------C--------T---------C-T--A 2750
H2U.FR.96.12034.AY530889    ----T--A---G----G-----C--C---AG---C--G--GTG------C--C-----A--T-----T--------------CA-T---------------------A-------C--C-----GT-------A--C-----A--A------A-A--T--T--G--T--A 2905
SMM.SL.92.SL92B.AF334679    ----G--A--------G-----C-----T--AG----C-----CCCG--------A--A-----A--A--C---A----------G--C--CCA---A--------GA-G--T-----C-----C-----A--A-----------A--T--T--A--T-----GGCA--C 2771
MAC.US.x.MAC239             ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2417
MAC.US.x.251                -------------------C-----------------------------------------------------------------------------------------------------------------A------------------------------------ 3339
MAC.US.x.251_32H            ----------------------------------------------------------------------------------------------------------G-----------------------------------A--------------------------- 3341
MAC.US.x.251                ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3316
MAC.US.x.17EC1.AY033233     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3340
MAC.US.x.17EFR.AY033146     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3340
MAC.US.x.1937.AY611495      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2446
MAC.US.x.2065.AY611493      ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2446
MAC.US.x.MM142.Y00277       -------------------C-----------T------------------------------------------------------------------------------------------------------------------------------------------ 2826
MAC.US.x.SMM142B.BD131285   -------------------C-----------T------------------------------------------------------------------------------------------------------------------------------------------ 2826
MNE.US.82.MNE               -----------G-------C-----------------------------------------------T-----------------------------A--------------------------------------G---------------------------C----- 2808
MNE.US.x.MNE027.U79412      -----------G-------C-----------------------------------------------T-----------------------------A--------------------------C-----------G---------------------------C----- 2808
SMM.US.x.F236               -----T-----G----------------------------------G-----C--------------------------------------------A---------A-------C--------AT-A-----A--G-----------------T--T--T--G--T--- 3327
SMM.US.x.H9.M80194          ----GT-----G--------------------------T-------------C-----A--------------C--------G-----C--------A---------A-------C--------AT-A-----A--G-----A--A--------T--T--T--RC-T--- 2813
SMM.US.x.PBJ14              ----GT-----G--------------------------T-------------C-----A--------------C--------G-----C--------A---------A-------C--------AT-A-----A--G-----A--A--------T--T--T--GC-T--- 3107
SMM.US.x.PBJA.M31325        ----GT-----G--------------------------T-------------C-----A--------------C--------G-----C--------A---------A-------C--------AT-A-----A--G-----A--A--------T--T--T--GC-T--- 3106
SMM.US.x.PBJ                ----GT-----G--------------------------T-------------C-----A--------------C--------G-----C--------A---------A-------C--------AT-A-----A--G-----A--A--------T--T--T--RC-T--- 2813
SMM.US.x.PBJ                ----GT-----G--------------------------T-------------C-----A--------------C--------G-----C--------A---------A-------C--------AT-A-----A--G-----A--A--------T--T--T--GC-T--- 3299
SMM.US.x.PGM53.AF077017     -----T-----G----------------------------------------C-----A--------T--------------------C--------A---------A-------C--------AT-A-----A--G-----A-----------T--T--T--GC-T--- 3257
SMM.US.x.SME543.U72748      -----T-----G----------------------T-----------------C--------------------------------------------A---------A-------C--------AT-A-----A--G-----------------T--T--T--G--T--- 3343
STM.US.x.STM.M83293         -----T-----G----G-----------T--------C--------G-GC-----------T-----T--C--C--------------C--------A---------A-------C--C-----------A--------------A--------A--T--T-----T--A 2987
MAC.US.x.80035.AY611486     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2446
MAC.US.x.81035.AY599200     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2447
MAC.US.x.85013.AY611490     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2380
MAC.US.x.92050.AY603959     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2440
MAC.US.x.92077.AY599201     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2447
MAC.US.x.93057.AY611492     -------------------------------------------------R------------------------------------------------------------------------------------------------------------------------ 2445
MAC.US.x.93062.AY607704     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2432
MAC.US.x.95058.AY611494     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2434
MAC.US.x.95086.AY607703     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2303
MAC.US.x.95112.AY588946     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2436
MAC.US.x.96016.AY607701     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2447
MAC.US.x.96020.AY611488     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2439
MAC.US.x.96072.AY611491     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2451
MAC.US.x.96081.AY597209     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2449
MAC.US.x.96093.AY611489     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2447
MAC.US.x.96114.AY588945     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2442
MAC.US.x.96123.AY611487     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2336
MAC.US.x.96135.AY607702     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2432
MAC.US.x.97009.AY599199     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2315
MAC.US.x.97074.AY599198     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2315
MAC.US.x.r80025.AY576480    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2303
MAC.US.x.r90131.AY576481    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2446
```

```
MAC.US.x.239               ATGAGACATGTGCTAGAACCCTTCAGGAAGGCAAATCCAGATGTGACCTTAGTCCAGTATATGGATGACATCTTAATAGCTAGTGACAGGACAGACCTGGAACATGACAGGGTAGTTTTACAGTCAAAGGAACTCTTGAATAGCATAGGGTTTTCTACCCCAGAAGAGAA 3510
Pol                        _T__M__R__H__V__L__E__P__F__R__K__A__N__P__D__V__T__L__V__Q__Y__M__D__D__I__L__I__A__S__D__R__T__D__L__E__H__D__R__V__V__L__Q__S__K__E__L__L__N__S__I__G__F__S__T__P__E__E
H2A.CI.88.UC2.U38293       -----G--GA-CT-------T-----A--A-----C--------C-T-C-CA-------C--------T-------------------A--G-GTT-A--G-------AA--G--CC-G---CT---A------C-A---G--C-------C-----------G----- 3593
H2A.DE.x.BEN.M30502        -----G--A--CT-------T-----A--A-----C--------C-TTC-CA-------C--------T-------------------------GTT-A--G-------AA--G--CC-G---CT---A-----TC-A---G--C-------------T-----T----- 3593
H2A.DE.x.PEI2.U22047       -----G--G--CT----G--A-----A--A-----C-----C--C-TTC-C-----A-----------T--------------C------------T-A--G--------AACG--CC-G----T---A-----T--A---G--C----A--C--C--------T----- 3561
H2A.GH.x.GH1.M30895        -----G--A--CT-------T-----A--A-----C--------C-TTC-CA----A--C--------T--T----------------------GTT-A--G-------AA-----CC-G---CT---A-----TC-----G--C-------C-----------T----- 3037
H2A.IN.NNVA.NEW            -----GA-GA-CT----G--A-----A--------C-AG-----C-TTA-CA-T--A--C--------T---C----------C------T----T-CA----------AA-----CC-----CT---------TC-A---G--C-------------------T----- 3587
H2A.IN.95.CRIK             -----G--GA-CT----G--A-----A--A-----C-AG-----C-TTA-CA-T-----C--------T-------------------------T--A----------A-----CC-----CT---------TC-A---G--C-------------------T----- 3338
H2A.GM.x.MCR35.AY509260    -----G--G--CT-------A-----A--A-----C-T------C-TTA-CA-T-----C--------T---C---------------------T--A----------A---G--CC-G---CT---------TC-A----A-C----A--------------T----- 3038
H2A.GM.87.D194.J04542      -----G--AA-CT-------T-----A--A-----C-----C--C-TTC-CA----A--C--------T--------------------G-GTT-A--G-------AA-----CC-G--ACT---A-----TC-----G--C-------C-----------T----- 3037
H2A.GM.x.ISY.J04498        -----G--A--CT-------A-----A--A-----C--------C-TTA-C--T-----C--------T-------------------------TT------------AA-----CC-----CT---------TC-A---G--C-G--A-----C--------C--A-- 3034
H2A.GM.x.MCN13.AY509259    -----G--G--CT-------A-----A--A-----C-T------C-TTA-CA-T-----C--------T---C---------------------T--A----------A---G--CC-G---CT---------TC-A----A-C----A--------------T----- 3038
H2A.GW.86.FG.J03654        -----G--GA-CT-------A-----A--A-----CGAG-----C-TTA-CA-T-----C--------T-----------C--C------------T-A----------AA--G--CC-G---CT---A-----TC-A---G-AC----A-----C--------T----- 3031
H2A.GW.87.CAM2CG.D00835    -----G--G--CT----G--A-----A--A-----CT-------C-TTA-CA-T-----C--------T---------------------------T-A----------AA--G---C-G---CT---------TC-A----A-C----A--C-----------T----- 3602
H2A.GW.x.ALI.AF082339      --------GA-CT----G--A-----A--------C-AG-----C-TTC-CA-T--A--C--------T-------------------------TT-A----------A---G--CC-G----T-----------C-A---G--C----A-----C--------T----- 3586
H2A.GW.x.MDS.Z48731        --------G--TT-------A-----A--A-----C-AG-----C-TTA-CA-T-----C--------T--------------C----------T--A------------A----CC-G---CT-----------C-A------C-G----C-----------T--A-- 3037
H2A.SN.85.ROD.M15390       --------G--AT-------A-----A--A-----CAAG-----C-TTA-CA-T-----C--------T-------------------------TT-A--------T--------CC-G---CTC--------TC-A---G--C----A--------------T----- 3039
H2AB.CI.90.7312A.L36874    ------A-G--A-----C---------A-G---C--CAAT-----C---A--A-------C-----------C-TG-G--A--C--T--A-GT--T--A--------T-A------G-CT--ACT---A--G--AC-A---GA---G--A-----C-----------A-- 3575
H2B.CI.88.UC1.L07625       -----GA-G--AT----T--T-----A-----C--CAGC-----C-TTA--A-T-----C-----------C-T-----A--------A-GT--T-----G-------------G-CC--ACT---A--G--A--A---GA---G--A--C--------------A-- 3574
H2B.CI.x.EHO.U27200        ---GC-A-G--A-----C--T-----A--A--C--CAAT-----C--TA--A-------C----------TC-CG-G--A---------GC--T-----G-------------G-CT--ACT---A--G--A--A----A---G--A--C-----T--------A-- 3564
H2B.GH.86.D205.X61240      ------A-G--AT----T--T--T--A-----C--CAGC-----C-TTA--A-T-----C-----------C-T-----A--------A-GT--T-----G--C----------G-CC--ACT---A--GT-A--A---GA---G--A--C--------------A-- 3566
H2B.JP.01.KR020.AB100245   G-AG-GA-A--A-----C---------A--A--C---GAT-----C---A----------C-----------C-TG-G--A--------A-GC--T--A-----------------GC-T--ACTG--A--G--A--A---GA------A--C--------------A-- 2716
H2AB.CM.03CM_510_03.NEW    -----GA-G--AT----T--T-----A--A--C--CAGC-----C--TA--A-T-----C-----------C-T-----A--------A-GT--T-----G-------------G-CC--ACT---A--G--A--A----A---G--A--C--------------A-- 2705
H2G.CI.92.ABT96.AF208027   -----GA-G--A-----C---------A--A--C---GAG---R-A-TTA-YA----A-------------TC-TG-G--C--------A--T--TT-A--------TCAA--------G---CT-------------A----AA-CR-----C--------T--T--A-- 2920
H2U.FR.96.12034.AY530889   -----GA-CC-A--G--G--A--T--A---------GA-----CTGT-A--A-T----------------TC-T-----C-----T--A-AT--TT-A--G-----TCA-T------C----CTC-----G--AC------AG-C---A---AGC--------C--A-- 3075
SMM.SL.92.SL92B.AF334679   -----G--G--AT-------------A--A--T--C---------CTTC-C----------------TT-AC----T-GC---A---AGGCCTAACA--G-------AAA-G--AAC------T--GA--CA-GC-C--C-ATC-G-----CAG------------C-- 2941
MAC.US.x.MAC239            ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2587
MAC.US.x.251               -----------------------------------------------------------------------------------A----------------------------------------------T------------A---------------------------- 3509
MAC.US.x.251_32H           ----------------------------------------------------------------------------------------------------------------------------------T-------------------------------------A-- 3511
MAC.US.x.251               ---------------------------------------------------------------------------------------------------------------------------------CT------------A-------------C-------------- 3486
MAC.US.x.17EC1.AY033233    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3510
MAC.US.x.17EFR.AY033146    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3510
MAC.US.x.1937.AY611495     ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2616
MAC.US.x.2065.AY611493     ----------------------------------------------------------------------------------------------------------------------------------T---------------------------Y------------- 2616
MAC.US.x.MM142.Y00277      ----------------------------------------------------------------------------------------------------------------------------------T---A--------A-----------------AT------------ 2996
MAC.US.x.SMM142B.BD131285- ----------------------------------------------------------------------------------------------------------------------------------T---A--------A-----------------AT------------ 2996
MNE.US.82.MNE              ------A-------G-----------------------------------------G-----------------------------------------------------------------------------T------------A---------------------------- 2978
MNE.US.x.MNE027.U79412     ------A-------G------------------------C-------------------------G-----------------------------------------------------------------------T------------A---------------------------- 2978
SMM.US.x.F236              ------A----AT-------T-----A--A-----------------C-GA----A--C--------------------------T--A-----TT-A--G---------------------T---------TC----CG-------A--C---------------- 3497
SMM.US.x.H9.M80194         R-----A----AT-------T--T--A--A-----------------C-GA----A--C--------------------------T--A-----TT-A------------------------A-T---------TC----C--T-----A-----C-------------- 2983
SMM.US.x.PBJ14             ------A----CT-----------T--A--A-----------------C-GA----A--C--------------------------T--A-----TT-A---------------------------T---------TC----C--T-----A-----C-G------------ 3277
SMM.US.x.PBJA.M31325       ------A----CT-----------T--A--A-----------------C-GA----A--C--------------------------T--A-----TT-A---------------------------TG--------TC----C--T-----A-----C-------------- 3276
SMM.US.x.PBJ               R-----A----CT-----------T--A--A-----------------C-GA----A--C--------------------------T--A-----TT-A---------------------------T---------TC----C--T-----A-----C-------------- 2983
SMM.US.x.PBJ               ------A----CT-----------T--A--A-----------------C-GA----A--C--------------------------T--A-----TT-A---------------------------T---------TC----C--T-----A-----C-------------- 3469
SMM.US.x.PGM53.AF077017    ------A----CT---G---T-----A--A-----------------C-GA----A--C------------C---------------A-----TT-A--G------------------------T---------TC----C--------A-----C-------------- 3427
SMM.US.x.SME543.U72748     ------A----AT-------T-----A--A-----------------C-GA----A--C--------------------------T--A-----TT-A--G------------------------T---------TC----CG-------A--C--C-------------- 3513
STM.US.x.STM.M83293        ------A--A-AT----G--A-------GA--C--C--------A--TC-GA----A--C--------T-----G-----C---------A-----T--A--G------------------------TG---------C-A----ATT----A-----C-----T-------- 3157
MAC.US.x.80035.AY611486    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2616
MAC.US.x.81035.AY599200    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2617
MAC.US.x.85013.AY611490    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2550
MAC.US.x.92050.AY603959    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2610
MAC.US.x.92077.AY599201    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2617
MAC.US.x.93057.AY611492    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2615
MAC.US.x.93062.AY607704    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2602
MAC.US.x.95058.AY611494    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2604
MAC.US.x.95086.AY607703    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2473
MAC.US.x.95112.AY588946    ----------------------------------------------------------------------------------------------------------------------------------Y----------------------------------------- 2606
MAC.US.x.96016.AY607701    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2617
MAC.US.x.96020.AY611488    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2609
MAC.US.x.96072.AY611491    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2621
MAC.US.x.96081.AY597209    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2619
MAC.US.x.96093.AY611489    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2617
MAC.US.x.96114.AY588945    ----------------------------------------------------------------------------------------------------------------------------------Y----------------------------------------- 2612
MAC.US.x.96123.AY611487    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2506
MAC.US.x.96135.AY607702    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2602
MAC.US.x.97009.AY599199    ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2485
MAC.US.x.97074.AY599198    --------------------------------------------------------------------------------------------------------------------------------R-Y----------------------------------------- 2485
MAC.US.x.r80025.AY576480   ----------------------------------------------------------------------------------------------------------------------------------T----------------------------------------- 2473
MAC.US.x.r90131.AY576481   ----------------------------------------------------------------------------------------------------------------------------------Y---------------------------Y------------- 2616
```

```
MAC.US.x.239                 ATTCCAAAAAGATCCCCCATTTCAATGGATGGGGTACGAATTGTGGCCAACAAAATGGAAGTTGCAAAAGATAGAGTTGCCACAAAGAGAGACCTGGACAGTGAATGATATACAGAAGTTAGTAGGAGTATTAAATTGGGCAGCTCAAATTTATCCAGGTATAAAAACCA 3680
Pol                          __K__F__Q__K__D__P__P__F__Q__W__M__G__Y__E__L__W__P__T__K__W__K__L__Q__K__I__E__L__P__Q__R__E__T__W__T__V__N__D__I__Q__K__L__V__G__V__L__N__W__A__A__Q__I__Y__P__G__I__K__
H2A.CI.88.UC2.U38293         G-----------C--T--G--------------C--T---C----------T--G------C----G--A---C-AC----C--G-A---A--A--------C-----C--C--A--AC----G---A-C-----------G--A-----C--------A---------- 3763
H2A.DE.x.BEN.M30502          G-----------C--T-----------------C-GT---C-A--------T---------C----G--AC--C-AC----C--G-A---C-TA--------C-----C--C--A---C----G-----C-----------G--A-----C---T----A---------- 3763
H2A.DE.x.PEI2.U22047         G-----------C-------ACA----------C--T---C-A--------C---------C-------A---C-A-----C----A---AGTA--------C-----C--C--A---C-------T--CC-------------A-----C--C-----G-----G---- 3731
H2A.GH.x.GH1.M30895          G-----------C--T--GC-------------C--T---C-A--------T---------C----G--AT--C-AC----C--G-A---A-TA--------C--C--C--C------C-G--G-----C-----------G--A-----C--------A---------- 3207
H2A.IN.NNVA.NEW              G--------------T----ACT-C--------C--T-----A--------T--------------G--A---C-------C--G-AG---GTA--------C-----C--C--A---C----G--C--CC----------G--A--------C-----A--------A- 3757
H2A.IN.95.CRIK               G--------------T----AC-----------T--T---C-A--------T--------------G--A---C-------C----AG---GTA--------C-----C--C--A---C-G--G--T--CC-------------A-----C--C---A-G-----G--A- 3508
H2A.GM.x.MCR35.AY509260      G--------G--C--T----AC--C--------C--T---C----------T--G------C----G------C-------C----A---TGTA--------A-----C--C--A--------G--T--C-----C--------A-----C--C-----G---------- 3208
H2A.GM.87.D194.J04542        G--------G--C--T--G--------------C--T--------------T--------AC----G--A---C-A--A--T--G-A---A-TA--------C-----C--C--A--AC----------T--G--C-----G--G--G--C--------G---------- 3207
H2A.GM.x.ISY.J04498          G--------G--C--T----AC-----------C--T---C----------C--------A--------A---C-A-----C----AG--AGTA--------T-----C--C-----AC----G--T--CC----C-----G--A-----C--C-----A-----G---- 3204
H2A.GM.x.MCN13.AY509259      G--------G--C--T----AC--C--------C--T---C----------T--G------C----G------C-------C----A---TGTA--------A-----C--C--A--------G--T--C-----C--------A-----C--C-----G---------- 3208
H2A.GW.86.FG.J03654          ------G-----C--T----A--GC--------C--T-----A--------T-----------------A---C--C----C----A---AGTA--------C-----C--C--A---C----G--T--CC-------------A-----C--C-----G-----G---- 3201
H2A.GW.87.CAM2CG.D00835      G-----------C--T----AC-GC--------C--T---C----------T--------------G--A---C-A-----C----AG---GTG--------C-----C--C--A---C-------T--CC-------------A-----C--C-----G-----G---- 3772
H2A.GW.x.ALI.AF082339        G-----------C--T----ACA----------C--T-G-C----------T---------C-------A---C-A-----C--G-A---AGTA--------C-----C--C--A--ACAT--G--T--CC-------------A-----C--C-----A-----G---- 3756
H2A.GW.x.MDS.Z48731          G-----------C--T----AC-GC--------C--T---C----------T---------C----G--A---C-------C----A---AGTA--------C-----C--C--A--AC----G--T--CC-------------A-----C--C-----A-----G--A- 3207
H2A.SN.85.ROD.M15390         G-----------C--T----AC--C--------C--T---C-A--------T--------------G--A---C-------C----A---A-TA--------C-----C--C------C----G--T--CC-------------A---C-C--C-----G-----G---- 3209

H2AB.CI.90.7312A.L36874      ---------G--C--T-----C--G--------T--T--GC-C-------AG--G------C-------A---C-AA-A---G----G---GTT--------A-----C--T-----AC----G-----------C------------C---TC--C--A--C-----A- 3745

H2B.CI.88.UC1.L07625         G-----------C--T--G--CA----------T--T--GC-C-------A--GG-----AC-------A---C-AC-----G---AG--AGTT--------------C--T--A--AC----G-----------C------------C-C-T---T--A--T--G--A- 3744
H2B.CI.x.EHO.U27200          G-----------C--T-----CA-------------T--GC-C-------AG--------AC-------A---C--C-A---G---A----GTT--------A-----C--T-----------G------------------------C---TC--G--G--T--G---- 3734
H2B.GH.86.D205.X61240        G-----------C--T--G--CA----------T--T--GC-C-------A---G-----AC-------A---C-AC-----G---A---AGTT-------------CA--T--A--AC-G--------------C------------C-C-T---T--A--T--G--A- 3736
H2B.JP.01.KR020.AB100245     G--------G--C--------CA-------------T--GC-C-------A---G-----AC-------A---C--C-----G-------A-GA--------A--------T-----AC----G------C----C------------C-C-T---T--G--T-----A- 2886

H2AB.CM.03CM_510_03.NEW      R-----R--R--C--T--G--C-G---------T--T--GC-Y-------A---G-----AC-------A---C-AC-----G---AG--AGTT--------------C--T--A--AC-G--------------C------------C-C-T---T--A--C--R--A- 2875

H2G.CI.92.ABT96.AF208027     G-----------C--T----A------------A--T-----R-------A---G-----AC-A--------TA-CC----MG---A------T--------A--------T--A--AC-------G---C-G--Y-----------GC-A--------A--C-----A- 3090

H2U.FR.96.12034.AY530889     ---T--G-----C--A---CACA----------C--T---C-A-----T-A---G-----A-----G--A--T---C-C---G-------TGTT--------A-----C--T------------------C-------------A---G----------A-----G--A- 3245

SMM.SL.92.SL92B.AF334679     G--T-----GA----A-----A--G--------A--TTTG--A-AC--C-AG--G-----A--A--G--A-----AC-C--TG---AG---CGA--------A-----C-----A--AC----G-----TC-------------A--G--A--------G--T-----A- 3111

MAC.US.x.MAC239              ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2757
MAC.US.x.251                 ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 3679
MAC.US.x.251_32H             ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 3681
MAC.US.x.251                 ------------------------------------------------G------------------------------------------------------------------------------------------------------------------------- 3656
MAC.US.x.17EC1.AY033233      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 3680
MAC.US.x.17EFR.AY033146      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 3680
MAC.US.x.1937.AY611495       -------------------------------------------------------------------------------------------R------------------------------------------------------------------------------ 2786
MAC.US.x.2065.AY611493       ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2786
MAC.US.x.MM142.Y00277        ------------------------------------------------G------------------------------------------------------------------------------------------------------------------------- 3166
MAC.US.x.SMM142B.BD131285    ------------------------------------------------G------------------------------------------------------------------------------------------------------------------------- 3166

MNE.US.82.MNE                G-----------------------------------T-------------------------------------------------A----------------------------------------------------------------------------------- 3148
MNE.US.x.MNE027.U79412       G-----------------------------------T-----A-------------------------------------------A----------------------------------------------------------------------------------- 3148

SMM.US.x.F236                G-----G--------------C--G--------A--T--------------C--------AC----G--A--------------------------------A-----C-----A--A-----------GC-------------A--------------A-----G--T- 3667
SMM.US.x.H9.M80194           G--------------------C--G--------A--T-----------C--C--------AC-------A-----------G--------A-----------A---A-------A--A-----------GC-------------A--------------A-----G--T- 3153
SMM.US.x.PBJ14               G-----G--------------C--G--------A--T-----------C--C--------AC-------A--------------------A-----------A-----------A--A-----------GC-------------A--------------A-----G--T- 3447
SMM.US.x.PBJA.M31325         G-----G--------------C--G--------A--T-----------C--C--------AC-------A--------------------A-----------A-----------A--A-----------GC-------------A--------------A-----G--T- 3446
SMM.US.x.PBJ                 G-----G--------------C--G--------A--T-----------C--C--------AC-------A--------------------A-----------A-----------A--A-----------GC-------------A--------------A-----G--T- 3153
SMM.US.x.PBJ                 G-----G--------------C--G--------A--T-----------C--C--------AC-------A--------------------A-----------A-----------A--A-----------GC-------------A--------------A-----G--T- 3639
SMM.US.x.PGM53.AF077017      G-----G--------------C--G--------A--T--------------C--------AC----G--A--------------------A-----------A-----C-----A--A------------C-------------A--------------A-----G--T- 3597
SMM.US.x.SME543.U72748       G-----G--------------C--G--------A--T--------------C--------AC----G--A--------------------------------A-----C-----A--A-----------GC-------------A--------------A-----G--T- 3683

STM.US.x.STM.M83293          ---T--G--G-----T-----------------A--T--G--A-----------G-----AC-------A-----A--A--T--G-----CGTT-----------------C--A--AC-G-----G--C-----C--------A--------------A--T-----T- 3327

MAC.US.x.80035.AY611486      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2786
MAC.US.x.81035.AY599200      -----------------------------------------------------------R--------------------------------Y----------------------------------------------------------------------------- 2787
MAC.US.x.85013.AY611490      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2720
MAC.US.x.92050.AY603959      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2780
MAC.US.x.92077.AY599201      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2787
MAC.US.x.93057.AY611492      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2785
MAC.US.x.93062.AY607704      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2772
MAC.US.x.95058.AY611494      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2774
MAC.US.x.95086.AY607703      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2643
MAC.US.x.95112.AY588946      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2776
MAC.US.x.96016.AY607701      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2787
MAC.US.x.96020.AY611488      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2779
MAC.US.x.96072.AY611491      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2791
MAC.US.x.96081.AY597209      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2789
MAC.US.x.96093.AY611489      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2787
MAC.US.x.96114.AY588945      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2782
MAC.US.x.96123.AY611487      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2676
MAC.US.x.96135.AY607702      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2772
MAC.US.x.97009.AY599199      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2655
MAC.US.x.97074.AY599198      ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2655
MAC.US.x.r80025.AY576480     ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2643
MAC.US.x.r90131.AY576481     ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 2786
```


HIV-2/SIV Complete Genomes

```
MAC.US.x.239               AACATCTCTGTAGGTTAATTAGAGGAAAAATGACTCTAACAGAGGAAGTTCAGTGGACTGAGATGGCAGAAGCAGAATATGAGGAAAATAAAATAATTCTCAGTCAGGAACAAGAAGGATGTTATTACCAAGAAGGCAAGCCATTAGAAGCCACGGTAATAAAGAGTCAG 3850
Pol                        T__K__H__L__C__R__L__I__R__G__K__M__T__L__T__E__E__V__Q__W__T__E__M__A__E__A__E__Y__E__E__N__K__I__I__L__S__Q__E__Q__E__G__C__Y__Y__Q__E__G__K__P__L__E__A__T__V__I__K__S_
H2A.CI.88.UC2.U38293       ----CT-G--C-----------------------A--C-----A-----G--------A---T-A-----------GCTA--A--G--C--G--T--C--AGAC--------G------CAC-------------AG--AGA---G--G--A---A-CCA---A-----A 3933
H2A.DE.x.BEN.M30502        ----CT-A-----AC-------------------A--C-----A-----G--------A--AC-A-----------GCTA--A-----C-----T--CT-G--C----------------A--------------AA--AGA------G--A--AA-CCA---A--C--A 3933
H2A.DE.x.PEI2.U22047       ----CT-A------C-------------------A--C--G--A-----A--------A--AC-A-----G------CTA--A--G--C-----T--CT-G--C--------G--G-----C-------------AA---GA---------A--A--CCA----GA---A 3901
H2A.GH.x.GH1.M30895        -----T-G------C----C-A------------A--C-----A-----G--------A---T-A-----------GCTA--A-----C-----T--CT-A--C----------------AC-------------AA--AGA------G--A--AA-CCA---A-A---A 3377
H2A.IN.NNVA.NEW            ----CT-A-----------C--G------T----A--C-----A-----G--------A---T-A--------G--GCTA--A-----C--G--C--AT-A--C-----------G---CAC------A----G-AA--AGA-C----G--A--A--CCA---AGAC--A 3927
H2A.IN.95.CRIK             -G--CT-A-----------C--G------T----A--C-----A-----A--A-----A---T-A--------G--GCTA--A-----C--G--C--CT-A--C-----------G---CAC-----------G-AA---GAGC----------A--CCA---AGAC--A 3678
H2A.GM.x.MCR35.AY509260    G---CT-A----A-C-------------------A--C-----A-----A--------A--AC-A-----G--G--GCTA--A--G--C--G--T--CT-A--C-----G-----G---CAC-------------AA--AGGG--------A--A--CCA----GA---A 3378
H2A.GM.87.D194.J04542      -G---T-A----AA--G-----------------A--C-----------A--------A---T-A-----G------CTA--A-----C-----T--CT-A--------------G---CC---C--T--G----AAG-AGA-C-------A--A--C--C--A--C--A 3377
H2A.GM.x.ISY.J04498        ----CT-A----A-C-------------G-----A-CC--G--A-----A--------A--AT-A-----------GCTG--------C-----T--CT-A--C--------G--G---CAC-----------G-AA--AGAG--------A--A--TCA----GA---A 3374
H2A.GM.x.MCN13.AY509259    G---CT-A----A-C-------------------A--C-----A-----A--------A--AC-A-----G--G--G-TA--A--G--C--G--T--CT-A--C-----G-----G---CAC-------------AA--AGAG--------A--A--CCA----GA---A 3378
H2A.GW.86.FG.J03654        ----C--A------C-------------------A--C-----A--G--A--------A--AT-A-----------GCTA--------C--G--T--CT-A--C--A------------CAC--------G----AA--AAAG--------A--A--CCAG--AGA---A 3371
H2A.GW.87.CAM2CG.D00835    ----CT-G-----A--------------------A--C-----A--G--A--A-----A--AT-A------------CTA--A-----C--G--T--CT-A--C--A--------G---CAC--C-----G----AA---GAG--------A--A--TCA----GA---A 3942
H2A.GW.x.ALI.AF082339      ----CT-A-----AC-------------------A--C-----A-G--G---------A--AC-A------------CTA-----G--C--G--T--CT-A--------------G--GCAC-------------AA---GAG--------A--A--CCA---AGA---A 3926
H2A.GW.x.MDS.Z48731        -----T-A--------------G-----------A--C--G--A---A-A--------A---T----------G---CTA--A-----C-G-G-T--CT-G--C-----------G---CAC-----------G-AA--AGA---------A--A--CCA----GA---A 3377
H2A.SN.85.ROD.M15390       ----CT-A-----------C--------------A--C-----A-----A--------A--AT-A-----------GCTA--A-----C--G--T--C--A--C-----------G---CAC-------------AA--AGAGC-------A--A--CCA----GA---A 3379
H2AB.CI.90.7312A.L36874    GG--CA-A--C--AC-G--C-AG-----G-----C--G--------G--G--------A--AC-A-----G-----GATGC-------------C--C--AG-A--------G------C---C---A--------GGTA--TC-------A--A--GCAG--A-ACTTA 3915
H2B.CI.88.UC1.L07625       GG--CA-A----AAC-------G-----G-----C--------A-----A--A-----A--AT-A-----G-----G-TAC----G--C-----C---T-AGAA---------------CC--C---A--G-----GGTA---C-------A--A---CAG--A-AC-TA 3914
H2B.CI.x.EHO.U27200        GG---A-A----AAC----A--G-----G-----C--------A--G--A--A--------AT-------G-------TCC-------C-----C--C--AGAA--A--G--G------CC------A--------GGTA--T--------A--A--GCAG--A-A--TA 3904
H2B.GH.86.D205.X61240      GG--CA-A--C-AAC-------G-----G-----C--------A-----A--------A--AC-A-----------GCTAC----G--------C--CT-AGAA---------------CC--C---A--G---A-GGTA--GC-------A--A---CAG--A-AC-TA 3906
H2B.JP.01.KR020.AB100245   GG---A-A--C--AC----C--G-----------CT-G-----A-----A--------A--AT-A-----G-------T-C-------C-----T--CT-AGAA--A-----G--G---CC--C--TA-----G--A-TA--TC-------A--A---CAG--A-AC-TA 3056
H2AB.CM.03CM_510_03.NEW    GG--CA-G--C-RAC-------G-----G-----C--------A---R-A--------A--AT-A-----G-----G-TAC----G--Y-----C--CT-AGAA---------------CC--C---A--G-----GGTA-----------A--A---CAG--A-AC-TA 3045
H2G.CI.92.ABT96.AF208027   --A-C--G--Y-AA------------T--R----CT-G--G--A-----A--------A--AT---------------TGC-A-----C--G--T--C--ARAA-----R--G-----CGCC-----A-------RAG----T--R--------T---CAG--A-A-TTA 3260
H2U.FR.96.12034.AY530889   -------T----AA-------AG--G--------A--C-----A--G--A--------A--A----------------T--CA--G-----G--T---T-G--------------G--GGC-------AGG----AAG-AAATC-------A--A--GC-T--A-A---A 3415
SMM.SL.92.SL92B.AF334679   --A-C--G--C-AAA-G--C-----G-----------G-----A--------------A--AT-A-------------TG-CA--G----G---T-----A-A---A--G--G------A-A------AG-----ATG-A--GC-------A-----TC------A---A 3281
MAC.US.x.MAC239            ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2927
MAC.US.x.251               --------------C-------------------------------------------------------------------------------------------------------------------------------------------------------A---- 3849
MAC.US.x.251_32H           -----------------------------------T-------------------------------------------------------------------------------------------------G------------------------------------ 3851
MAC.US.x.251               -------------------------------------------------------------------------------------------G------------------------------------------------------------------------------ 3826
MAC.US.x.17EC1.AY033233    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3850
MAC.US.x.17EFR.AY033146    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3850
MAC.US.x.1937.AY611495     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2956
MAC.US.x.2065.AY611493     ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2956
MAC.US.x.MM142.Y00277      --------------------------------------------------------------------------------------------------------------------------------------A----------------------G------------ 3336
MAC.US.x.SMM142B.BD131285  --------------------------------------------------------------------------------------------------------------------------------------A----------------------G------------ 3336
MNE.US.82.MNE              ----------------------------------------------------------------------G-----------------------------------------------------------------------------G----------------A---- 3318
MNE.US.x.MNE027.U79412     ----------------------------------------------------------------------G-----------------------------------------------------------------------------G----------------A---- 3318
SMM.US.x.F236              -------T--C--AC----C---------------T-------A--G-----------------------G-----------A-----C--G--------------A-----------------C-----------A--A---A----G--A--A--------------- 3837
SMM.US.x.H9.M80194         -------T--C--A-----C---------------T-------A--G-----------------------G-----------A-----C--G-------C------A-----C--G--------C--T--G--G--A------C----G--AC-A--------------- 3323
SMM.US.x.PBJ14             -------T--C--A-----C---------------T-------A--G--------------------G--G-----------A-----C--G--------------A-----------------C--T--G--G--A------C----G--A--T--------------- 3617
SMM.US.x.PBJA.M31325       -------T--C--A-----C---------------T-------A--G-----------------------G-----------A-----C--G--------------A-----------------C--T--G--G--A------C----G--A--A--------------- 3616
SMM.US.x.PBJ               -------T--C--A-----C---------------T-------A--G-----------------------G-----------A-----C--G-------C------A-----C--G--------C--T--G--G--A------C----G--AC-A--------------- 3323
SMM.US.x.PBJ               -------T--C--A-----C---------------T-------A--G-----------------------G-----------A-----C--G--------------A-----------------C--T--G--G--A------C----G--A--A--------------- 3809
SMM.US.x.PGM53.AF077017    -------T--C-AA-----C---------------T-------A--G-----------------------------------A-----C--G--------------A-----G--G--------C--------G--A-----------G--A--A--------------- 3767
SMM.US.x.SME543.U72748     -------T--C--AC----C---------------T-------A--G-----------------------G-----------A-----C--G--------------A-----------------C--------G--A--A---A----G--A--A--------------- 3853
STM.US.x.STM.M83293        -------G--C-AA-----C-----T--G---G-A--G--T-----G--A--------A--A--------G----------CA-----------T--------------G--G-----GCAA--C----G-----A------TC----G--A--A---G-G--AGA---- 3497
MAC.US.x.80035.AY611486    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2956
MAC.US.x.81035.AY599200    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2957
MAC.US.x.85013.AY611490    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2890
MAC.US.x.92050.AY603959    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2950
MAC.US.x.92077.AY599201    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2957
MAC.US.x.93057.AY611492    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2955
MAC.US.x.93062.AY607704    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2942
MAC.US.x.95058.AY611494    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2944
MAC.US.x.95086.AY607703    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2813
MAC.US.x.95112.AY588946    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2946
MAC.US.x.96016.AY607701    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2957
MAC.US.x.96020.AY611488    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2949
MAC.US.x.96072.AY611491    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2961
MAC.US.x.96081.AY597209    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2959
MAC.US.x.96093.AY611489    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2957
MAC.US.x.96114.AY588945    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2952
MAC.US.x.96123.AY611487    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2846
MAC.US.x.96135.AY607702    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2942
MAC.US.x.97009.AY599199    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2825
MAC.US.x.97074.AY599198    ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2825
MAC.US.x.r80025.AY576480   ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2813
MAC.US.x.r90131.AY576481   ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2956
```

```
MAC.US.x.239               GACAATCAGTGGTCTTATAAAATTCACCAAGAAGACAAAATACTGAAAGTAGGA...AAATTTGCAAAGATAAAGAATACACATACCAATGGAGTGAGACTATTAGCACATGTAATACAGAAAATAGGAAAGGAAGCAATAGTGATCTGGGGACAGGTCCCAAAATTCCA 4017
Pol                        _Q__D__N__Q__W__S__Y__K__I__H__Q__E__D__K__I__L__K__V__._G__K__F__A__K__I__K__N__T__H__T__N__G__V__R__L__L__A__H__V__I__Q__K__I__G__K__E__A__I__V__I__W__G__Q__V__P__K__F
H2A.CI.88.UC2.U38293       --------A---A-A--------A-----G-----G-----C--A---------...----A------A----A--C--C----------G--C---T--C-------G---G-T--------------------C----C--T-------G-A-A--------T-- 4100
H2A.DE.x.BEN.M30502        -GAC----A---A-A--C-----A-----G-----G-----C--A---------...--G-A-------------A-----C----------G--C---T--C-------G---G-T-------------A--G---C----C--T-------G-A-A--------T-- 4100
H2A.DE.x.PEI2.U22047       ------------A-A--------A-------G---G-----C--A---------...----A-------------A-----C----------G--C---T-G------------G-T--A-----------A------C----C--T-------GAA-A--------T-- 4068
H2A.GH.x.GH1.M30895        ------------A-A--C-----A-----G-----G-----C--A--------G...----A-------------A-----C----------G--C---T--C-------G---G-T-------------A------C----C----------G-A-A----------- 3544
H2A.IN.NNVA.NEW            ------------A-G--C--GG-A--T--G-G---A-----C--A---------...----A-------G----A---G-C--------C--G--C---T--C-------G---G-T--------------------C----C--------G-GAA-A--------T-- 4094
H2A.IN.95.CRIK             ------------A-A------G-A-----G-G---A-----C--A---------...----A-------G-G--A---G-C----------G--C---T-GC-------G---G-T----G--------A------C----C---------GAA-A--------T-- 3845
H2A.GM.x.MCR35.AY509260    ------------A-A------G-A-----G-G---G-----T--A--------G...----A-------------A-----C----------G--C---T-G--------A---G-T--A--G--------A------C----C--T-------GAA-A--------T-- 3545
H2A.GM.87.D194.J04542      ------------G-A--C-----A-----G-G---G-GGG-T--A---------...--G-A---G---------A-----T----------G--C------C-------A---G-C--A-----------------C-G--C--T-------GA--G--------T-- 3544
H2A.GM.x.ISY.J04498        ------------A-A------G-A-----G-G---A-----T--A---------...----A-------------A-----C--------C--G--C--GT-G--G-----G---G-T-------------A------C----C--T-------GAA-A--------T-- 3541
H2A.GM.x.MCN13.AY509259    ------------A-A------G-A-----G-G---G-----T--A--------G...----A-------------A-----C----------G--C---T-G--------A---G-T--A--G--------A------C--A-C--T------GAA-A--------T-- 3545
H2A.GW.86.FG.J03654        ------------A-A-----GG-A-----G-GG--G-----T--C----GT--G..A----A-------------A-----C--------C--G.-C---T-G--------G---G-T--A-----------A------C----C--T-------G-A-A--------T-- 3538
H2A.GW.87.CAM2CG.D00835    ------------A-A-----G--A-----G-----A-----T--A---------...----A-----A----AC----C---------------C-A-T----------G---G-T--A-----------A------C----C--T--A----GAA-A--T-----T-- 4109
H2A.GW.x.ALI.AF082339      --------A---A-A--------A-----G-G---A-----T--A-----G-A-...--G-A-----A--G--A-----C--------C--G--C---T-G--------G---G-T--A-----------A------C-G--C--T-------GAA-A----G---T-- 4093
H2A.GW.x.MDS.Z48731        ------------A-A------G-A-----G-G---A----CT--A---------...----A-------G----A--C--C--------C-----C---T-G--------G---G-T-------------A------C----C--T-------GAA-A--------T-- 3544
H2A.SN.85.ROD.M15390       --G---------A-A--------A-----G-----A-----T--A---------...----A-------G-G--A--C--C------------A-C---T-G--------G---G-T-------------A------C----C--T-------GAA-A--------T-- 3546
H2AB.CI.90.7312A.L36874    -CA--C------A-A--C--------T--G-G-A-T---G-C--A---------...----A-------G-T--A--C--C--C----------A---A--C-----------G-C--A-----------A------T-G--T--T--------AA-A----TG----- 4082
H2B.CI.88.UC1.L07625       -CA---------A-A--C--------T--G-G-A-T-G---TT-A---------...----A-------G-T--A--C--C--C-----C--G--A-----C-G---------G-T-------------A-----CC----C---------G--A-A---GT------ 4081
H2B.CI.x.EHO.U27200        -CA---------A-A--C--G-----T--G-G---T-----C--A---------...----A-------G-T--A--C--T--C----------A-------G--T------G-C--A------------------T-G--C---------G--A-A----TG----- 4071
H2B.GH.86.D205.X61240      -CA---------A-A--C--------T--G-G-A-T---G-C--A---------...----A-------G-T--A--C--G--C-----C--G--A-----C-G---------G-T----------C--A-----CC----C---------G--A-A---GTG----- 4073
H2B.JP.01.KR020.AB100245   -CA---------A-A--C-----C--T--G-G-----G---C--A-G-------...----A-------G-T--A--C--C--C-----C-----A-----C-G-----------C------------------T-G--C-C-------G-AA-A----TG----- 3223
H2AB.CM.03CM_510_03.NEW    -CA---------A-A--C--G-----T--G-G-A-Y-G---GT-A---------...----A-------G-T--A--C--C--C-----C--G--A--RA-RC-G---------G-T-------------A-----CC----C---------G--A-A----T------ 3212
H2G.CI.92.ABT96.AF208027   ------------A-A--C---G----T----G-A---G---C--A--------T...----A-------------A-----C-----A--------A---T----G--T-----GG-C--A--R---------------C----A--A------A-AT-A--CTTC--T-- 3427
H2U.FR.96.12034.AY530889   --T-----A---A-C--C--G--C--T--G-GT----GG-C-T-A---------...---------C---G----------T-----A--------TC-C--GC-----------GG-------G--------A----C-----A--A------A-AA-T----TG----- 3582
SMM.SL.92.SL92B.AF334679   ------------AG------G-----T----GT----GG--T--C--G-----G...--G--------A-----A----------A------A-A--GT--C-G--TA-----G-T--A--G--------A---AGTC--------A------A--ACA---TTT--T-- 3448
MAC.US.x.MAC239            ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3094
MAC.US.x.251               ------------------------------------------------------...------------------------------------T------------------------------------------------------------------------------ 4016
MAC.US.x.251_32H           ------------------------------------------------------...--G---------------------------------T---------------------------------------------------------T-------------------- 4018
MAC.US.x.251               ------------------------------------------------------...------------------------------------T------------------------------------------------------------------------------ 3993
MAC.US.x.17EC1.AY033233    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 4017
MAC.US.x.17EFR.AY033146    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 4017
MAC.US.x.1937.AY611495     ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3123
MAC.US.x.2065.AY611493     ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3123
MAC.US.x.MM142.Y00277      ------------------------------------------------------...------------------------------------T------------------------------------------------------------------------------ 3503
MAC.US.x.SMM142B.BD131285- ------------------------------------------------------...------------------------------------T------------------------------------------------------------------------------ 3503
MNE.US.82.MNE              -------------------G---------------------A------------...---------------A--------------------T------------------------------------------------------------------------------ 3485
MNE.US.x.MNE027.U79412     -------------------G----------------------------------...---------------A--------------------T------------------------------------------------------------------------------ 3485
SMM.US.x.F236              --T-----A-----A-------------------------G------------T...-----------G-T--A-----------A--------C---T--C-------C---G-G--------------A------C----A--T------G----G----------- 4004
SMM.US.x.H9.M80194         --T-----A-----A------------------G------------A-------C...-------------T--A-----------A--------C--CT---------------G-G--------------A---------A-A--T-----------G----G---T-- 3490
SMM.US.x.PBJ14             --T-----A-----A---------------------------------------C...-------------T--A-----------A--------C---T---------------G-G--------------A-----------A--T-----------G--G-G---T-- 3784
SMM.US.x.PBJA.M31325       --T-----A-----A---------------------------------------C...-------------T--A-----------A--------C---T---------------G-G--------------A-----------A--T-----------G--G-G---T-- 3783
SMM.US.x.PBJ               --T-----A-----A---------------------------------------C...-------------T--A-----------A--------C---T---------------G-G--------------A-----------A--T-----------G--G-G---T-- 3490
SMM.US.x.PBJ               --T-----A-----A---------------------------------------C...-------------T--A-----------A--------C---T---------------G-G--------------A-----------A--T-----------G--G-G---T-- 3976
SMM.US.x.PGM53.AF077017    --T-----A-----A--------C-------------G----------------C...-------------T--A-----------A--------C---T---------------G-CG----------G--A-----------A--T-----------G--------T-- 3934
SMM.US.x.SME543.U72748     --T-----------A------------------------G----A-----G--C...-----------G-T--A-----------A--------C---T--C------------G-G-----------------G---C----A--T-----G----G----------- 4020
STM.US.x.STM.M83293        --------A-----C-----G--A------------------T-----------C...--G--------A-----A--C---------A--------T---T----G---------G--T-----G-----G--A------------------------A------------ 3664
MAC.US.x.80035.AY611486    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3123
MAC.US.x.81035.AY599200    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3124
MAC.US.x.85013.AY611490    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3057
MAC.US.x.92050.AY603959    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3117
MAC.US.x.92077.AY599201    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3124
MAC.US.x.93057.AY611492    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3122
MAC.US.x.93062.AY607704    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3109
MAC.US.x.95058.AY611494    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3111
MAC.US.x.95086.AY607703    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 2980
MAC.US.x.95112.AY588946    ------------------------------------------------------...-------------------------------------------------------------------------R----------------------------------------- 3113
MAC.US.x.96016.AY607701    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3124
MAC.US.x.96020.AY611488    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3116
MAC.US.x.96072.AY611491    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3128
MAC.US.x.96081.AY597209    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3126
MAC.US.x.96093.AY611489    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3124
MAC.US.x.96114.AY588945    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3119
MAC.US.x.96123.AY611487    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3013
MAC.US.x.96135.AY607702    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3109
MAC.US.x.97009.AY599199    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 2992
MAC.US.x.97074.AY599198    ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 2992
MAC.US.x.r80025.AY576480   ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 2980
MAC.US.x.r90131.AY576481   ------------------------------------------------------...------------------------------------------------------------------------------------------------------------------- 3123
```

```
                                                
```

```
MAC.US.x.239                 ATAAACAGTCAAAAGAAGGGAAAGCAGGATATATCACAGATAGGGGCAAAGACAAAGTAAAAGTGTTAGAACAGACTACTAATCAACAAGCAGAATTGGAAGCATTTCTCATGGCATTGACAGACTCAGGGCCAAAGGCAAATATTATAGTAGATTCACAATATGTTATG 4357
Pol                          C__N__K__Q__S__K__E__G__K__A__G__Y__I__T__D__R__G__K__D__K__V__K__V__L__E__Q__T__T__N__Q__Q__A__E__L__E__A__F__L__M__A__L__T__D__S__G__P__K__A__N__I__I__V__D__S__Q__Y__V_
H2A.CI.88.UC2.U38293         ---G---------G-----A------------G-A--------A--A-G------G-----G--A--------A-----C-----G---------C-A--G--C---GCG------C--G----------T-------TT--C--C--------C-----G-----A--- 4440
H2A.DE.x.BEN.M30502          ---G------------G--A------------G-A--------A--A------------------A----------A-----C-----G--G--------A-----TC----GG------C--G----------C-------TT-----C--------------G-----A--- 4440
H2A.DE.x.PEI2.U22047         ---GG--A--------G--A--------------A--------A--G-G------G----GGA-A--G--G--A-----C-----G-----------A-----C--CGCA--------A-----------T-----A--C--------------C-----G-----A--- 4408
H2A.GH.x.GH1.M30895          ---G---------G-----A------C-------G-G--------A--A-G------G----G--AC-------GA-----C-----G-----------A-----C---GCG---A--C----------G--T-----A-TT-----C--------C-----G-----A--- 3884
H2A.IN.NNVA.NEW              ----G--------------A------------G-A--------A--G-G-----GG-----G-AC-----G--A-----C-----G-----------A-----C---GCA------C-A-----------T-----A-TC--------------C-----G-----A--- 4434
H2A.IN.95.CRIK               ---GG--------------A--G---------G-A-----C--A--G-G------G----G--AC-----G--A-----C-----G--T--------A-----C---GCA------C-A-----------T-----A--C-----------G--------------A--- 4185
H2A.GM.x.MCR35.AY509260      ---GG--A-----G-----A------------G-A--------A--G-G------G----GGA--C-------A-----C-----G-----------A-----C---GCA------C-A-----------T-----A--C--------------C-----G-----A--- 3885
H2A.GM.87.D194.J04542        ---G---------------A------------G-A--------A--G-G-----GG----G---A-----G--A--AT-C-----G---------C-A-----C---GCG------C--G----------T--C----TT-----C--------C-----G-----A--- 3884
H2A.GM.x.ISY.J04498          ---GG--A--------G--A--------------A--------A--A-------------GGA-A-----G--A-----C--C--------------A-----C---GCA------G-A-----------T-----A-TC------G-------C-----G-----A--- 3881
H2A.GM.x.MCN13.AY509259      ---GG--A-----G-----A------------G-A--------A--G-G------G----GGA--C-------A-----C-----G-----------A-----C---GCA------C-A-----------T-----A--C--------------C-----G-----A--- 3885
H2A.GW.86.FG.J03654          ---GG-----------------------------A--------A--G-G-----GG--------A-----G--A-----C-----G-----------A-----C---GCA------C-A-----------T-----A--T--------------C-----------A--- 3878
H2A.GW.87.CAM2CG.D00835      ---GG--A-----------A------------G-A--------A--A-G------G-----A-AC-----G--A-----C-----G--------G--A-----C---GCG------C-A-----------T-----A--T--------------C-----G-----A--- 4449
H2A.GW.x.ALI.AF082339        ---GG--A-----------A--------------A--------A--G-G------G--G-G--AC--------A-----C-----G-----------A-----C--CGCGC-----G-A-----------T-----A--C-----------------------G-----A--- 4433
H2A.GW.x.MDS.Z48731          ---GG--A-----------A------------G-A-------AA--G-G------G----G--AC-----G--A-----C-----G-----------A-----C---GCA------C-A-----------C-----A--T--------------C-----G-----A--- 3884
H2A.SN.85.ROD.M15390         ---GG--A-----------A------------G-A--------A--G--------G----GAAAC-----G--A-----C-----G---------C-A-----C---GCG------C-A----------G--T-----A-TT--------------C-----G-----A--- 3886
H2AB.CI.90.7312A.L36874      -C-G-ACC-----G-----A--G-----G---G-------C-----A-GG--------T---CCA-----G-----A--A---------------C-T---------GCAC-A---C-ACAG--T-----A---C-A-TC--C--C--------C-----------C--- 4422
H2B.CI.88.UC1.L07625         -C-G-GCC-----G--G--A--G---------G-------C-----A-----T--G--T------C----G-----A--A---------------C-T---------GCAC-A---C-ACAG--T-----A---C-A-TC--C--C--------C-----------C--- 4421
H2B.CI.x.EHO.U27200          -C---GCC------------------------G-------C-----A--G--T-----T---CCA--------A--A--A-----G--------GC-T---------GCAC-A---C-ACAG--------A---C---TC-----C--------------------C--- 4411
H2B.GH.86.D205.X61240        ---G-ACC-----G-----A------------G----T--C-----A-----T--G--T-----------------A--A--C-------------C-T---------GCAT-A-----A---------AA----C-A-TT--C--C--------------------C--- 4413
H2B.JP.01.KR020.AB100245     -C-GGACC--G--------A--G---------G-------C-----A--G--T-----T----CA-----G--A--A--A-----G---------C-T---------GCAC-A-----ACAG--T-----A---C---TC-----C-----G--C-----------C--- 3563
H2AB.CM.03CM_510_03.NEW      -----GCC-----------A------------G-------C-----A-----T--G--Y-----------------A--A---------------C-T---------GCAT-A-----RCAG--T--R--A---C-A-TT--C--C--------------------C--- 3552
H2G.CI.92.ABT96.AF208027     -----ACC--C--G--R--A--------C--CG-------C--A--A--------G-----R--C-----G-----A--A--C--G--G--R---C-T--G------GCA------C--CAG--------T-----A-TC-----C---------R-----C--G-----C--- 3767
H2U.FR.96.12034.AY530889     -C-GGGCT------C-G-----------C---G-G--------A--A-------G---------CCT-----------C--C--C--------------A-----C---GCT------C-AGA---T-----A-----A-T---C---G-GAC-------------A-A--- 3919
SMM.SL.92.SL92B.AF334679     ---G-A-CAGT---------------------G-------C-----A--------A--G--CCTGCCA--------G-----C-----G--------GC--C-G-CCGT-AT-------C---A---T--TCCAAGT----TC---G-AG--ACG--C-----G-----CT-- 3788
MAC.US.x.MAC239              -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3434
MAC.US.x.251                 ----------------------------------------------------------------------------------------G-----------------C-----------A--------------------------------------------------- 4356
MAC.US.x.251_32H             -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4358
MAC.US.x.251                 -------------------------------------------------------------------------------------------------A---------------------------------------A-------------------------------- 4333
MAC.US.x.17EC1.AY033233      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4357
MAC.US.x.17EFR.AY033146      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4357
MAC.US.x.1937.AY611495       -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3463
MAC.US.x.2065.AY611493       -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3463
MAC.US.x.MM142.Y00277        G------------------A--------------------C--------------G------------------------------------------------------------------------------------------------C--------------- 3843
MAC.US.x.SMM142B.BD131285    G------------------A--------------------C--------------G------------------------------------------------------------------------------------------------C--------------- 3843
MNE.US.82.MNE                ----------------------------------------------------------------C------------------------------------------------------G-------------------------------------------------- 3825
MNE.US.x.MNE027.U79412       ----------------------------------------------------------G-G---C--------------------------------------------------------------------------------------------------------- 3825
SMM.US.x.F236                ---GG--------G-----A--------C---G-G--G----------G--------C----C-T--------------C--C-----------G--------C--CTATC-A---C-AG----T--G--A-----A--------------------C------------- 4344
SMM.US.x.H9.M80194           ---GG--------G-----A--------C---G-G--------A----G--G---GAC----C--C-G-----------C--------------G--------C---TATC-A---C-AG----T-----A-----A--------------------C------------- 3830
SMM.US.x.PBJ14               ---GG--------G-----A--G-----C---G-G--------A----G------G-C----C--C-G-----------C--------------G--------C--CTC-C-A---C-AG----T-----A-----A--------------------C------------- 4124
SMM.US.x.PBJA.M31325         ---GG--------G-----A--G-----C---G-G--------A----G------G-C----C--C-G-----------C--------------G--------C--CTA-C-A---C-AG----T-----A-----A--------------------C------------- 4123
SMM.US.x.PBJ                 ---GG--------G-----A--G-----C---G-G--------A----G--G---GAC----C--C-G-----------C--------------G--------C--CTATC-A---C-AG----T-----A-----A--------------------C------------- 3830
SMM.US.x.PBJ                 ---GG--------G-----A--G-----C---G-G--------A----G------G-C----C--C-G-----------C--------------G--------C--CTA-C-A---C-AG----T-----A-----A--------------------C------------- 4316
SMM.US.x.PGM53.AF077017      ---GG--A-----G-----A--G-----C---G-A--------A----G------G-C----CCT--G-----------C--------------G--------C--CTATC-A---C-AG----T-----A--------C-----C-----------C------------- 4274
SMM.US.x.SME543.U72748       ---GG--------G-----A--------C---G-G--G----------G--------C----C-T--------------C--C-----------G--------C--CTATC-A---C-AG----T--G--A-----A--------------------C------------- 4360
STM.US.x.STM.M83293          -C-GG--A--T-----------------C--C--A--T-----A--A--A-------------CT--------------------------------A---------GC-------C--G-G--------T-----A---------G-G-----C-----------G--- 4004
MAC.US.x.80035.AY611486      ----------------------------------------------------------------------------------------------------------------------------------------------C--------------------------- 3463
MAC.US.x.81035.AY599200      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3464
MAC.US.x.85013.AY611490      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3397
MAC.US.x.92050.AY603959      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3457
MAC.US.x.92077.AY599201      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3464
MAC.US.x.93057.AY611492      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3462
MAC.US.x.93062.AY607704      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3449
MAC.US.x.95058.AY611494      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3451
MAC.US.x.95086.AY607703      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3320
MAC.US.x.95112.AY588946      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3453
MAC.US.x.96016.AY607701      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3464
MAC.US.x.96020.AY611488      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3456
MAC.US.x.96072.AY611491      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3468
MAC.US.x.96081.AY597209      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3466
MAC.US.x.96093.AY611489      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3464
MAC.US.x.96114.AY588945      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3459
MAC.US.x.96123.AY611487      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3353
MAC.US.x.96135.AY607702      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3449
MAC.US.x.97009.AY599199      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3332
MAC.US.x.97074.AY599198      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3332
MAC.US.x.r80025.AY576480     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3320
MAC.US.x.r90131.AY576481     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3463
```

```
MAC.US.x.239              GGAATAATAACAGGATGCCCTACAGAATCAGAGAGCAGGCTAGTTAATCAAATAATAGAAGAAATGATTAAAAAGTCAGAAATTTATGTAGCATGGGTACCAGCACACAAAGGTATAGGAGGAAACCAAGAAATAGACCACCTAGTTAGTCAAGGGATTAGACAAGTTCT 4527
Pol                       _M__G__I__I__T__G__C__P__T__E__S__E__S__R__L__V__N__Q__I__I__E__E__M__I__K__K__S__E__I__Y__V__A__W__V__P__A__H__K__G__I__G__G__N__Q__E__I__D__H__L__V__S__Q__G__I__R__Q__V
H2A.CI.88.UC2.U38293      ------G--G--A-CCAA--A-----G--------T--AA----A--C-----C--T-----C-----A--G--AGA--C-G-C-----T------A-C-----C--T-----C--------------G---G-------TT----A-----G--C--C--------A-- 4610
H2A.DE.x.BEN.M30502       --G---G--G----CCAG--A-----G-----A-AT--AA----G--C--G--C--------------A--G---GA--C-G-C-----T--------C-----C--T-----C--------------G---G-------TT----A--------C--C--------AT- 4610
H2A.DE.x.PEI2.U22047      ------G--G-G--CCAG--A--------------T-AA-----A--C-----C--------------A------GA-AC-C-C-----T--------C-----C--------C-----------T--G---G----T--TT----A-----G--C-----------AT- 4578
H2A.GH.x.GH1.M30895       ------G--GT---CCAA--A-----G-----A--T--AA----A--C-----C--T--G--C-----A--G--AGA--C-G-C-----C--G-----C-----C--T-----C--------------G---G-------TT----A-----G--C--C--------AT- 4054
H2A.IN.NNVA.NEW           ------G-GG----CCAA--A-----G-----A-AT--A-----A-----G--C--------------A------GA--C---C-----T--------C-----C--------C-----------T--G--GG----T--TT----A-----G--C--C--------AT- 4604
H2A.IN.95.CRIK            ------G-GG----CCAA--A-----G-----A-AT-A-A----G-----G--C--------------A------GA--CC--C-----T--------C-----C--------C-----------T--G--GG----T--TT----G-----G--C--C--------AT- 4355
H2A.GM.x.MCR35.AY509260   --G---G--G----CCAG--A--------------T--AA----A--------C-----G--G-----A------GA--C---C-----T--------C-----C--T-----C--------G--T--G--GG----T--GT----A-----G--C--C--------GT- 4055
H2A.GM.87.D194.J04542     --G---G--G----CCAA--A-----G-----A-AT--AA----A--C-----C--T--G--C-----A--G--AGA--C-G-C-----T--------C-----C--T-----C--------------G---G-------TT----A-----G--C--C--------AT- 4054
H2A.GM.x.ISY.J04498       ------G-------CCAA--GG-T-----------T--AA----A---A----T--------G-----A------GA--C---C-----T--------C--G--C--------C-----------T--------T------T----A-----G--C--C--------AT- 4051
H2A.GM.x.MCN13.AY509259   --G---G--G----CCAG--A--------------T--AA----A--------C-----G--G-----A------GA--C---C-----T--------C-----C--T-----C--------G--T--G--GG----T--GT----A-----G--C--C--------GT- 4055
H2A.GW.86.FG.J03654       --G---G--G----CCAA--A-----G-------AT--AA----A-----G--C--------------A--G--AGA--C---C-----T--G-----C-----C--------C--------------G---G----T--TT----A-----G--C-----------AT- 4048
H2A.GW.87.CAM2CG.D00835   ------G-GG----CCAA--A-----G-----A-AT--AA----A--------C--------------A------GA--C---C-----T--------C-----C--T-----C-----G--------G---G----T--TT----G--------C--C--------A-- 4619
H2A.GW.x.ALI.AF082339     ------G--G----CCAG--A-----------A-AT--AA----A--------C--------------A------GA--CC--C-----T--G-----C-----C--------C-----G-----T--G---G-------TT----A-----G--C--C--------AT- 4603
H2A.GW.x.MDS.Z48731       ------G-GGT---CCAA--A-----G-----A-AT---A----A-----G--C-----G--------A------GA--C---C-----T--------C-----C--------C-----------T--G---G----T--T-----A-----G--C--C--------AT- 4054
H2A.SN.85.ROD.M15390      --G--C-GTG--A-CCAA--A-----G-----A--T-AAA----G--C--G--C--------------A------GA--C---C-----T--------C-----C--------C-----G--------G---G----T--TT----G-----G--T--C--------GT- 4056
H2AB.CI.90.7312A.L36874   ------G--G-T---CAG--A------A-----TCACC-A----A---A-G-----T-----------C------GA--C---A-------G------G-----T--------AC----T--T--T--G---G-------------A--------A-----G--G----- 4592
H2B.CI.88.UC1.L07625      ------G--G-T---CAG--A------A-G---TCACCA-----A-----------T-----------C-----AGA--C---A-------G------G-----T----G---AC----T--T--T--G---G-------------A--------A-----G--G----- 4591
H2B.CI.x.EHO.U27200       ------G--G-T-C-CAA--A------A----ATCACC-A----A-GAG-------T-----------C------GA-A----A-------G------------T-----G--AC-G--T--T--T--G---G-------------G--C-----A---------A-C-- 4581
H2B.GH.86.D205.X61240     ---------G-T-C-CAG--A------A----ATCACCAA----AGCAA-------T-----------C-----AGAG-C-G-A-------G------------T-----G--AC-G--T--T--T--G---G-------------A--------A--C-----G--CT- 4583
H2B.JP.01.KR020.AB100245  ---------G-T-C-CAG--A------A----ATCACCAA----A--GAC----G-T-----------C------GA--C---A--C----G------G-----T--T-----AC-G--T--T--T------G-------------G--------A--C-----CA-C-- 3733
H2AB.CM.03CM_510_03.NEW   ---------G-T-C-CAR--A------A----ATCACCAA----R-CAAR------T-----------C-----RGAG-C---A-------G------------C-----G--AC----T--T---------G-------------A--------A--C-----G--CT- 3722
H2G.CI.92.ABT96.AF208027  --R--C---G----RCAA--A--G---A----ATC-CC-T----A---A-------T--------R--------AGA-RCTC-A-------G------------T--T-----A--------T--T-----GG------RT-----A--C--G--A-----------A-- 3937
H2U.FR.96.12034.AY530889  ---------G-----CAG--A------A-----TC-CCA-----A-G-A----C---------C-T--A-GG--AGA--C-G-G-----T-GC--------T--T--T--G-----T--T--C--T------G----T--TT-G--G-----G--A-----G-----A-- 4089
SMM.SL.92.SL92B.AF334679  AAC------------CAG--AT-------G--TTCAGATA----GGCA-----T--T---C--C-AG--C-G--AGA--C-G-A--CA---G------T-----T--T-----C--------G---A-T--GG------GTT-------------A--A-----G--C-- 3958
MAC.US.x.MAC239           -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3604
MAC.US.x.251              --------------------------------------A--------C---------A--------------------------------------------------------------------------G-------------------G----------------- 4526
MAC.US.x.251_32H          -----------------------------------------------C-------------------------------------------------------------------------------------------------------------------------- 4528
MAC.US.x.251              -----------------------------------------------C--------------------------------------------------------------------------------------------------------G----------------- 4503
MAC.US.x.17EC1.AY033233   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4527
MAC.US.x.17EFR.AY033146   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4527
MAC.US.x.1937.AY611495    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3633
MAC.US.x.2065.AY611493    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3633
MAC.US.x.MM142.Y00277     -----------------------------------------------C--------------------C------A-------------G-------------------------------------------------------------------------------- 4013
MAC.US.x.SMM142B.BD131285 -----------------------------------------------C--------------------C------A-------------G-------------------------------------------------------------------------------- 4013
MNE.US.82.MNE             -----------------------------------------------C---------------------------A----------------------G----------------------------------------------------------------------- 3995
MNE.US.x.MNE027.U79412    -----------------------------------------------C---------------------------A----------------------G--------------------------------------------------------------------C-- 3995
SMM.US.x.F236             --C------G----TCAG--C--T--------A--T--AT----A--C--G--------G---------------GA--C---------------------T-----T-----A--------C--T------G----T-----G-----C--G--A-----------C-- 4514
SMM.US.x.H9.M80194        --C---G--G----TCAA--C--T--------A--T--AT----A--C--G--------GA--------------GA--C---------------------T-----T---RRA-------A---T------G----T--T-------------AA-----------C-- 4000
SMM.US.x.PBJ14            --C---G--G----TCAA--C--T--------A--T--AT----A--C--G------------------------GA--C---------------------T-----T-----A-----------T------G----T--T--G-----------A-----------C-- 4294
SMM.US.x.PBJA.M31325      --C---G--G----TCAA--C--T--------A--T--AT----A--C--G------------------------GA--C---------------------T-----T-----A-----------T------G----T--T--G-----------A-----------C-- 4293
SMM.US.x.PBJ              --C---G--G----TCAA--C--T--------A--T--AT----A--C--G--------G---------------GA--C---------------------T-----T---RRA-----------T------G----T--T--G-----------A-----------C-- 4000
SMM.US.x.PBJ              --C---G--G----TCAA--C--T--------A--T--AT----A--C--G--------G---------------GA--C---------------------T-----T-----A-----------T------G----T--T--G-----------A-----------C-- 4486
SMM.US.x.PGM53.AF077017   --C------G----TCAA--C--T--------A--T--AT----A--C--G--------G---------------GA--C---------G-----------T--T--T-----A-----------T------G----T-----G-----------A-----------C-- 4444
SMM.US.x.SME543.U72748    --C---G--G----TCAG--C--T--------A--T---T----A--C--G-----------G------------GA--C---------------------T-----T-----A-----------T------G----T-----G-----C--G--A-----------C-- 4530
STM.US.x.STM.M83293       -----C---------CAG--C--------G-------A-T----A-----G------------------------GA--C---------------------------T-----A-----------T--G--GG----T--TT-G--A-----G--A--------G--C-- 4174
MAC.US.x.80035.AY611486   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3633
MAC.US.x.81035.AY599200   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3634
MAC.US.x.85013.AY611490   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3567
MAC.US.x.92050.AY603959   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3627
MAC.US.x.92077.AY599201   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3634
MAC.US.x.93057.AY611492   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3632
MAC.US.x.93062.AY607704   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3619
MAC.US.x.95058.AY611494   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3621
MAC.US.x.95086.AY607703   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3490
MAC.US.x.95112.AY588946   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3623
MAC.US.x.96016.AY607701   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3634
MAC.US.x.96020.AY611488   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3626
MAC.US.x.96072.AY611491   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3638
MAC.US.x.96081.AY597209   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3636
MAC.US.x.96093.AY611489   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3634
MAC.US.x.96114.AY588945   --------------------------------------------------------------------------------------------------------------------------------------------------------G----------------- 3629
MAC.US.x.96123.AY611487   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3523
MAC.US.x.96135.AY607702   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3619
MAC.US.x.97009.AY599199   --------------------------------------------------------------------------------G----------------------------------------------------------------------------------------- 3502
MAC.US.x.97074.AY599198   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3502
MAC.US.x.r80025.AY576480  -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3490
MAC.US.x.r90131.AY576481  -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3633
```

Pol RNAse H end Pol p31 Integrase start

```
MAC.US.x.239                CTTCTTG...GAAAAGATAGAGCCAGCACAAGAAGAACATGATAAATACCATAGTAATGTAAAAGAATTGGTATTCAAATTTGGATTACCCAGAATAGTGGCCAGACAGATAGTAGACACCTGTGATAAATGTCATCAGAAAGGAGAGGCTATACATGGGCAGGCAAATT 4694
Pol                         __L__F__.__L__E__K__I__E__P__A__Q__E__E__H__D__K__Y__H__S__N__V__K__E__L__V__F__K__F__G__L__P__R__I__V__A__R__Q__I__V__D__T__C__D__K__C__H__Q__K__G__E__A__I__H__G__Q__A__
H2A.CI.88.UC2.U38293        T---C--...--G--A--------T--C-----G---------A-----T-----C---A-------GC--A-CCAT---------A------CA-C----A--A------------A--------A----CCC----C--A--------------C-----------A-T----G 4777
H2A.DE.x.BEN.M30502         A---C--...--------------C--T-----G---------A-----T-----C-T-A-------C-AACCCAT---------A------CTTC----A--A------------A--T-A----CCC----C--A-----------A--C-----------A-T----G 4777
H2A.DE.x.PEI2.U22047        A---C-A...-----A--------C--T--G------------G-----T-----C-------------ATCCCAT---------C-G----A-C-------A-----A------A----A----CCC-------A--------G--A-----------------A-T-G--G 4745
H2A.GH.x.GH1.M30895         G---C-A...----G---------C--C---------------A-----T-----C---A-G------C-AACCCAT---------A------CA-C----A--A------------A----A----CCC----C--A--------------C------------T----G 4221
H2A.IN.NNVA.NEW             G---C-A...-----A--------C--T---------------A-----T----AC-----G------C-ATCCCA-------------------A-C-------A-----A------A----A----CCC-------A--------G--------------A--A-T----G 4771
H2A.IN.95.CRIK              G---C-A...-----A--------C--T--------G------A-----T-----C------------C-ATCCCAT----A-----------A-C-------A------------A----A----CCC-------A---------------------------A-T----G 4522
H2A.GM.x.MCR35.AY509260     G---C-A...-----A--------C--T--G------------G-----------C------------C-ATCCCAT-----------G----A-T----A--A-----A------A----A----CCC-------A--------G-----------------A-T-G--G 4222
H2A.GM.87.D194.J04542       A---C--...--------------C--T-----------C---A-----T-----C---A--------C-AACCCAT---------A------CA-C-------A------------A----A----CCC----C--A-----------A--C-----------A-T----G 4221
H2A.GM.x.ISY.J04498         A---C-A...--G-GA--------C--T--G-----------GA-----T-----C------------C-A-CCCAT--G-------------ACC-G-----A-----A------A----A----CCC-G--C--A--A-----G--A-----------------A-T----G 4218
H2A.GM.x.MCN13.AY509259     G---C--...-----A--------C--T--G------------G-----------C------------C-ATCCCAT-----------G----A-T----A--A-----A------A----A----CCC-------A--------G-----------------A-T-G--G 4222
H2A.GW.86.FG.J03654         A---C-A...--G--A-----A--C--G--G------------A-----T-----C---A--------C-TCCCAT---------A--------AGC-------A-----A------A----A----CCC-TGTG--A--------G-----------------A-T----G 4215
H2A.GW.87.CAM2CG.D00835     G---C-A...--------------T--T--G------------A-----T----CC-----------GC-ATGCCAT------ATA------CA-C-------A-----A------A----A----CCC---A---A--------G-----------------A-T----G 4786
H2A.GW.x.ALI.AF082339       G--TC-A...-----A-----------T--G------------A--G--T-----C-----G------C-ATCCCAT-------------------ATC-G-----A------------A----A----CCC-------G-----G---------------------A-TG---G 4770
H2A.GW.x.MDS.Z48731         G---C-A...-----A--------C--T---------------A-----T-----C------------C-ATCCCAT---------------CA-C-------A-----A------A----A----CCC-------A---------------------------A-T----G 4221
H2A.SN.85.ROD.M15390        G---C--...-----A--------C--T--G------------A-----T-----C------------C--TCTCAT---------A--------ATT-------A--G--A------A--T-A----CCC-------A--------G--A-----------------A-T----G 4223
H2AB.CI.90.7312A.L36874     A---C-A...-----------A-----C--G--G---------A-----T---G-C-----------GC----TCAT---------A-T--ACA-C-------A-A---A------A--T-----------------A--A--------A---G-T-----A----T----G 4759
H2B.CI.88.UC1.L07625        A---C-A...-----------A-------------------A--G------G-C-------------A--TCAT----------C-C--ACAGT-------A-A---A------A--T-------------C--A--A-----G--A---G-T-----A----T----G 4758
H2B.CI.x.EHO.U27200         A--TC-A...-----A-----A-----T---------------A-----T---A-------------C-A--CCAT---------GA-T--ACA-T-------A-----A------A--TT------------C--A--A-----G--A----T-----A----T----- 4748
H2B.GH.86.D205.X61240       G---C-A...-----A-----A-----C--G-----G------A-----T---G-C------------C---CCAT-----C---A-T--ACA-T-------A-A-----------A--TT------------C--A--A-----G--A----T-----A----T----G 4750
H2B.JP.01.KR020.AB100245    G--TC-A...-----T-----A-----C--------G------A-----T---G---------------A--TCAT---------A-T--ACA-T-------A-A---A------A--T-------C--------A-----------A-----C-----A----TG---G 3900
H2AB.CM.03CM_510_03.NEW     G---C-A...-----A-----A-----T--G-----G------R-----T---G-C-----------MC----CCAT-----Y---M-C--ACA-T----R--A-A-----------A--TT-Y----C-----C--A--R-----R--A----T-----A----T----G 3889
H2G.CI.92.ABT96.AF208027    A------...R----A-----A--T--T---------------A--R-TT----------------G--AACTCAT--G-----CA-T--TCA-T----A--A-A-----------A--TT-A-ACC-C-GT-C--A-----------A--C--T-----A----T----G 4104
H2U.FR.96.12034.AY530889    G--TC-A...-----A-----A-----T--G------------G-----------C-----T------GC-ACAGCAT----TA------ACA-T----A----A---A------A--T--T--C--GTT--C--G-----------A--C--G-----A----T----A 4256
SMM.SL.92.SL92B.AF334679    G---C-A...----GT--------------------C------C-----T-----C--------G--G--A--CTCAG---A-AACA-C--ACA-T----A--T-A-----------A--TG-T--A-C-----C--A------G-----A--G--------T---A-----G 4125
MAC.US.x.MAC239             -------...-----------------------------------------------------------------------------------------------------------------------------------------------------RY----- 3771
MAC.US.x.251                -------...--------------------------G-------------------------------------------------------------------------------------------------------A----------------T----- 4693
MAC.US.x.251_32H            -------...-----------------------G-----------G--------------------------------------------------------------------------A-----------------A--------A----------------T----- 4695
MAC.US.x.251                -------...----------------------------------------------------------------------------------------------------------------------------------A----------------T----- 4670
MAC.US.x.17EC1.AY033233     -------...----------------------------------------------------------------------------------------------------------------------------------------------------------- 4694
MAC.US.x.17EFR.AY033146     -------...----------------------------------------------------------------------------------------------------------------------------------------------------------- 4694
MAC.US.x.1937.AY611495      -------...----------------------------------------------------------------------------------------------------R-----------------------------R-------------------------T----- 3800
MAC.US.x.2065.AY611493      -------...------------------------------------------------------------------------------------------------------------------------------------------------------T----- 3800
MAC.US.x.MM142.Y00277       -------...------------------------------AG--------------CA-----------------------------------C---------A---------------A-----------------A--------A----------------T----- 4180
MAC.US.x.SMM142B.BD131285-------...------------------------------AG--------------CA-----------------------------------C---------A---------------A-----------------A--------A----------------T----- 4180
MNE.US.82.MNE               -------...---------------------------------------------------------------------------C---------A---------------A--------------------------A----------------T----- 4162
MNE.US.x.MNE027.U79412      -------...---------------------------------------------------------------------------C---------A---------------A--------------------------A----------------T----- 4162
SMM.US.x.F236               A-----A..AA----A-----A--------------G------A--G--------C-----------------------------T-----T--GC----A--A-A---------------A-----------C--C-----------A--C-----------A-T----G 4682
SMM.US.x.H9.M80194          A-----A...-----A-----A-----G--G-----G------A-----------C---------A-----------------R-T-----T--GC----A--A-A---A-----------A-----------C--C-T---------A--Y-----------A-T----G 4167
SMM.US.x.PBJ14              A-----A...-----A-----A-----G--G-----G------A-----------C-----------------------------T-----T--GC----A--A-A---------------A-----------C--C-----------A--C--------A--A-T----G 4461
SMM.US.x.PBJA.M31325        A-----A...-----A-----A-----G--G-----G------A-----------C-----------------------------T-----T--GC----A--A-A---------------A-----------C--C-----------A--C--------A--A-T----G 4460
SMM.US.x.PBJ                A-----A...-----A-----A-----G--G-----G------A-----------C-------------------------R-T-----T--GC----A--A-A---------------A-----------C--C-----------A--Y-----------A-T----G 4167
SMM.US.x.PBJ                A-----A...-----A-----A-----G--G-----G------A-----------C------------A----------------T-----T--GC----A--A-A---------------A-----------C--C-----------A--C--------A--A-T----G 4653
SMM.US.x.PGM53.AF077017     A-----A...-----A-----A--------------G------A--G--------C-----------------------------TA----T--GC----A--A-A---------------A---C-------C--C-----------A--C-----------A-T----G 4611
SMM.US.x.SME543.U72748      A-----A...-----A-----A--------------G------A--G--------C-----------------------------TA----T--GC----A--A-A---------------A--------G--C--C-----------A--C-----------A-T----G 4697
STM.US.x.STM83293           A---C-A...-----A-----A-----T--------G------A-----T-----C-----G-----GC-A--C-----G-----CA----A--GC----A--A-AG--A-----G--T--A-----C--G--C--C-----------A---------------A-T----G 4341
MAC.US.x.80035.AY611486     -------...------------------------------------------------------------------------------------------------------------------------------------------------------T----- 3800
MAC.US.x.81035.AY599200     -------...------------------------------------------------------------------------------------------------------------------------------------------------------Y----- 3801
MAC.US.x.85013.AY611490     -------...------------------------------------------------------------------------------------------------------------------------------------------------------T----- 3734
MAC.US.x.92050.AY603959     -------...------------------------------------------------------------------------------------------------------------------------------------------------------Y----- 3794
MAC.US.x.92077.AY599201     -------...------------------------------------------------------------------------------------------------------------------------------------------------------T----- 3801
MAC.US.x.93057.AY611492     -------...------------------------------------------------------------------------------------------------------------------------------------------------------T----- 3799
MAC.US.x.93062.AY607704     -------...------------------------------------------------------------------------------------------------------------------------------------------------------T----- 3786
MAC.US.x.95058.AY611494     -------...--------------------------------Y---------------------------------------------------------------------------------------------------------------------T----- 3788
MAC.US.x.95086.AY607703     -------...------------------------------------------------------------------------------------------------------------------------------------------------------T----- 3657
MAC.US.x.95112.AY588946     -------...------------------------------------------------------------------------------------------------------------------------------------------------------Y----- 3790
MAC.US.x.96016.AY607701     -------...------------------------------------------------------------------------------------------------------------------------------------------------------T----- 3801
MAC.US.x.96020.AY611488     -------...------------------------------------------------------------------------------------------------------------------------------------------------------------ 3793
MAC.US.x.96072.AY611491     -------...------------------------------------------------------------------------------------------------------------------------------------------------------T----- 3805
MAC.US.x.96081.AY597209     -------...------------------------------------------------------------------------------------------------------------------------------------------------------------ 3803
MAC.US.x.96093.AY611489     -------...------------------------------------------------------------------------------------------------------------------------------------------------------T----- 3801
MAC.US.x.96114.AY588945     -------...------------------------------------------------------------------------------------------------------------------------------------------------------------ 3796
MAC.US.x.96123.AY611487     -------...------------------------------------------------------------------------------------------------------------------------------------------------------T----- 3690
MAC.US.x.96135.AY607702     -------...-----------------------------------------------------------------------------------------------------------------------------------------------------RY----- 3786
MAC.US.x.97009.AY599199     -------...------------------------------------------------------------------------------------------------------------------------------------------------------T----- 3669
MAC.US.x.97074.AY599198     -------...-----------------------------------------------------------------------------------------------------------------------------------------------------R------ 3669
MAC.US.x.r80025.AY576480    -------...------------------------------------------------------------------------------------------------------------------------------------------------------T----- 3657
MAC.US.x.r90131.AY576481    -------...------------------------------------------------------------------------------------------------------------------------------------------------------Y----- 3800
```

```
MAC.US.x.239               CAGATCTAGGGACTTGGCAAATGGATTGTACCCATCTAGAGGGAAAAATAATCATAGTTGCAGTACATGTAGCTAGTGGATTCATAGAAGCAGAGGTAATTCCACAAGAGACAGGAAGACAGACAGCACTATTTCTGTTAAAATTGGCAGGCAGATGGCCTATTACACAT 4864
Pol                        N__S__D__L__G__T__W__Q__M__D__C__T__H__L__E__G__K__I__I__I__V__A__V__H__V__A__S__G__F__I__E__A__E__V__I__P__Q__E__T__G__R__Q__T__A__L__F__L__L__K__L__A__G__R__W__P__I__T_
H2A.CI.88.UC2.U38293       ----AA----CGT-----------C--C--A--CT----A--------C--T-----A-----------T------------------------C--C-----G--AT-------G-----------C--C--A------C----CA-T--G-----A--A-----C 4947
H2A.DE.x.BEN.M30502        ----AA----CGT-----------C-AC--A--CT----A--------C--T-----A-----------T--A--------------------A--C--C-----G--AT-------G-----------C--C--A------C----CA-T--G-----A--A--G--C 4947
H2A.DE.x.PEI2.U22047       ----A--G--C-------------C--C--A--CT---------------C--T-----A-----------T--A--C--G--T----------A--T--C-----G--A--G-----G--A--------C--C--A------C----CA-T--G-----A--A-----C 4915
H2A.GH.x.GH1.M30895        ----AA----TGTC----------C--C--A--CT----A-----G--C--T-----A-----------T--A--------------------A--T--C-----G--AT-------G--------G--C--C--A------C----CA-T--G-----A--A-----C 4391
H2A.IN.NNVA.NEW            ----AT----C--G----------C--C--A---T----A------G-C--------G-----------T--A--------T-----G-----A--C--C---------AT-------G--A--------C--C--A----GC-A--TA-T--------A--A-----C 4941
H2A.IN.95.CRIK             ----A-----T-------------C--C--A---T----A------G-C--T-----A-----G-----T--A--------T-----G-----A--C--C-----G--A--------G--A--------C--C---------C-A--TA-T--------A--A-----C 4692
H2A.GM.x.MCR35.AY509260    ----AT----C-------------C--C--A--CT----A-----G--C--T-----A-----------T--A--------------------A--C--C-----G--AT-------G-----------C--C--A------C----CA-T--G-----A--A-----C 4392
H2A.GM.87.D194.J04542      ----AA----CGT-----------C--C--A--CT----A--------C--T-----A-----G-----T--A--------------------A--C--C-----G--AT-------G-----------C--C--A------C----CA-T--G-----A--A-----C 4391
H2A.GM.x.ISY.J04498        ----A-----C--C----------C--C--A--CT----A--------C--T-----A-----------T--A--------T-----------A--C--C-----G--AT-------G--A--------C--C--A------C----CA-T--G-----A--A-----C 4388
H2A.GM.x.MCN13.AY509259    ----AT----C-------------C--C--A--CT----A-----G--C--T-----A-----------T--A--------------------A--C--C-----G--AT-------G-----------C--C--A------C----CA-T--G-----A--A-----C 4392
H2A.GW.86.FG.J03654        ----A-----C-------------C--C--A---T----A------G-C--T-----A-----------T--A--------T-----G-----A--T--C-----G--AT-------G--A--------G--CT-AC-G---C----CA-T--G-----A--A-----C 4385
H2A.GW.87.CAM2CG.D00835    ----AG-G--C--C----------C--C--A-------------------C--------A-----------T--A--------T-----------A--C--C-----G--AT-------G--A--------C--C--A-----------TA----G-----A--A-----C 4956
H2A.GW.x.ALI.AF082339      ----A-----C-------------C--C--A--CT----A------G-C------A-A-----G-----T--C-----------------------C--C--G--G--AT-----------A--------C--C--A------C----TA-T--------A--A-----C 4940
H2A.GW.x.MDS.Z48731        ----CT----C-------------C--C--T---T----A------G-C--T-----A-----------T--A--------T-----G-----A--C--C-----G--A--------G-----------C--C---------C----TA-T--G-----A--A-----C 4391
H2A.SN.85.ROD.M15390       ----A-----C-------------C--C--A---T----A-----G--C--T-----A-----------T--A--------T-----------A--C--C-----G--AT-----------A--------C--C--A--G---C-----A-T--G-----A--A-----C 4393
H2AB.CI.90.7312A.L36874    ----A--------A-----------C-----A---T----A--G---G-C--A-----A-----C--------C-----G--T--------------------A--C-----A-----------A-----T--C--C--------GC-----A-----------C---------C 4929
H2B.CI.88.UC1.L07625       ----A-----A--A-----G-----C-----A--CT----A-----GG-C--A-----G-----C--------C---------T-----G----------A--C-----A-----------A-----T--C--C--------------CA--------------C-----C 4928
H2B.CI.x.EHO.U27200        ----A--------A-----------C-----A---T----------GG-T--A-----G-----T--------C-----------------------A-----A--C-----A-----------------T--C--C--------GC----CA--------------C-----C 4918
H2B.GH.86.D205.X61240      ----C--------A-----G-----C-----A---T----A--------T--A-----G-----C--------C-----G--T-------------------A--C-----------------------T--C--C--AC----G-----CA--------------C-----C 4920
H2B.JP.01.KR020.AB100245   ----A--------A-----------------A--CT---------------T--A-----G-----C--------C---------T-----------A-----A--C-----A-----------------T--C--C------------C----CA--------------C-----C 4070
H2AB.CM.03CM_510_03.NEW    ----A--------A-----G-----C--C--A---T----A-----G--T--A-----A-----C--------C-----G--Y-----------AR----A--C-----AT----------------T--C--C---------G-----CA----R---------C-----C 4059
H2G.CI.92.ABT96.AF208027   ----AT-G--T-------------C--C------T----A------G-----------A-----T-----G--C-----C--TG-----------A-----C--------A---------G--A------T----C-------------A-----G------------C-----C 4274
H2U.FR.96.12034.AY530889   ----AG-------A-----------C--------CT-----------G-TG----C--G-----------G--C-----C--T-----------A--G--C--G--G--A--T--T--G--A----C--G--C--AC----GC-A-----A--G----------C--T--- 4426
SMM.SL.92.SL92B.AF334679   ----GG----C--A-----G-----C--C-----CT----A-----GG-T---------A-----------G---------C--T--------------------A--TAGG--A----------------T-G--C--C------A-A---TCA-----------A-AG--- 4295
MAC.US.x.MAC239            ------------------------------------------------------------------------------------------------------------------------------------------------------ 3941
MAC.US.x.251               ------------------------------------------------------------------------------------------------------------------------------------------------------ 4863
MAC.US.x.251_32H           -----------------------------------------------------------------------------------A----------------------------------------------------------G------- 4865
MAC.US.x.251               ----------------------C-------------A-------G--------------------------------------A------------------------------------------------------------------ 4840
MAC.US.x.17EC1.AY033233    ------------------------------------------------------------------------------------------------------------------------------------------------------ 4864
MAC.US.x.17EFR.AY033146    ------------------------------------------------------------------------------------------------------------------------------------------------------ 4864
MAC.US.x.1937.AY611495     ----Y----------------------------------------------------------------------------------------------------------------------------------------------------- 3970
MAC.US.x.2065.AY611493     ------------------------------------------------------------------------------------------------------------------------------------------------------ 3970
MAC.US.x.MM142.Y00277      ----C---------------------------------------G--------------------------------------A-------------A----------------------------------A----------------- 4350
MAC.US.x.SMM142B.BD131285  ----C---------------------------------------G--------------------------------------A-------------A----------------------------------A----------------- 4350
MNE.US.82.MNE              -----------------------------------------------------------------------------------A------------------------------------------------A--------------G-- 4332
MNE.US.x.MNE027.U79412     -----------------------------------------------------------------------------------A------------------------------------------------A--------------G-- 4332
SMM.US.x.F236              ----A-------------------C-----G--C-----A--C--------T-----------------G-----------T-----G-----A-----C--G--G--A-----------A--------G-----------G--A--T-----------C--C------ 4852
SMM.US.x.H9.M80194         -R-RA-------------------C-----A--C--------C--------T-----------------G-G----------------G-----A-----C--G-RG--A-----------A--------G------C-------A--CA----------C--C------ 4337
SMM.US.x.PBJ14             ----A-------------------C-----A--C--------C--------T-----------------G------------------G-----A-----C--G--G--A-----------A--------G------C-------A--CA----------C--C------ 4631
SMM.US.x.PBJA.M31325       ----A-------------------C-----A--C--------C--------T-----------------G------------------G-----A---------G--G--A-----------A--------G------C-------A--CA----------C--C------ 4630
SMM.US.x.PBJ               -R-RA-------------------C-----A--C--------C--------T-----------------G-G----------------G-----A-----C--GRGG--A-----------A--------G------C-------A--CA----------C--C------ 4337
SMM.US.x.PBJ               ----A-------------------C-----A--C--------C--------T-----------------G------------------G-----A-----C--G--G--A-----------A--------G------C-------A--CA----------C--C------ 4823
SMM.US.x.PGM53.AF077017    ----A-------------------C-----A--C--------C--------------A-----------G------------------G-----A---------G--G--A--------G--A----G--A--------G------C-------A--TA----------C--C------ 4781
SMM.US.x.SME543.U72748     ----A-------------------C-----G--C-----A--C--------T-----------------G-----------T-----G-----A-----C--G--G--A-----------A--------G-----------GC-A--T-----------C--C------ 4867
STM.US.x.STM.M83293        ----GT-------------G-----C-----A--------A--C--G-----T-----A------------------G--T-----G-----A-----C--------A-----G-----------T-G--------G--G--A---A-----------G-C-----C 4511
MAC.US.x.80035.AY611486    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3970
MAC.US.x.81035.AY599200    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3971
MAC.US.x.85013.AY611490    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3904
MAC.US.x.92050.AY603959    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3964
MAC.US.x.92077.AY599201    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3971
MAC.US.x.93057.AY611492    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3969
MAC.US.x.93062.AY607704    -----------------------------------------------------------------Y------------------------------------------------------------------------------------ 3956
MAC.US.x.95058.AY611494    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3958
MAC.US.x.95086.AY607703    ------------C------------------------------------------------------------------------------------------------------------------------------------------- 3827
MAC.US.x.95112.AY588946    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3960
MAC.US.x.96016.AY607701    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3971
MAC.US.x.96020.AY611488    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3963
MAC.US.x.96072.AY611491    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3975
MAC.US.x.96081.AY597209    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3973
MAC.US.x.96093.AY611489    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3971
MAC.US.x.96114.AY588945    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3966
MAC.US.x.96123.AY611487    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3860
MAC.US.x.96135.AY607702    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3956
MAC.US.x.97009.AY599199    ------------------------------------------------------------------------------------------------------------------------------------------------------ 3839
MAC.US.x.97074.AY599198    --------------------A--------------------------------------------------------------A------------------------------------------------------------------ 3839
MAC.US.x.r80025.AY576480   ------------C------------------------------------------------------------------------------------------------------------------------------------------- 3827
MAC.US.x.r90131.AY576481   ------------------------------------------------------------------------------------------------------------------------------------------------------ 3970
```

```
MAC.US.x.239               CTACACACAGATAATGGTGCTAACTTTGCTTCGCAAGAAGTAAAGATGGTTGCATGGTGGGCAGGGATAGAGCACACCTTTGGGGTACCATACAATCCACAGAGTCAGGGAGTAGTGGAAGCAATGAATCACCACCTGAAAAATCAAATAGATAGAATCAGGGAACAAGC 5034
Pol                        _H__L__H__T__D__N__G__A__N__F__A__S__Q__E__V__K__M__V__A__W__W__A__G__I__E__H__T__F__G__V__P__Y__N__P__Q__S__Q__G__V__V__E__A__M__N__H__H__L__K__N__Q__I__D__R__I__R__E__Q
H2A.CI.88.UC2.U38293       T-G--------C-----CC-C-----CA----A--G--------A-----G----------T---T--------AT-------A-----T-----C-----A--C-----------A--------------------A--G-----G---AG------T--A-----G-- 5117
H2A.DE.x.BEN.M30502        T-G--------C-----CC-C-----CA----A--G-----G--------G----------T---T-----A--AT-------A-----T-----C-----A--C-----------A--------------------A--G-----G---AG------T--A-----G-- 5117
H2A.DE.x.PEI2.U22047       --G-----------------C-----CA-C--A--G--------------A--G------A----T-----A--AT-------A-----T-----------A--C--A--------A-----------------T--A-----C--G---AGC-----T--A--G--G-- 5085
H2A.GH.x.GH1.M30895        T----------C-----CT-C-----CA----A--G-----G--------G---------AT---T--------AT-------A-----T-----C-----A--C-----------A-----------------T--A--G-----G---AG------T--A-----G-- 4561
H2A.IN.NNVA.NEW            --G--------------A--------CA----A--------G--------A----------T---T--------AT-------A-----T-----------A--C-----G-----A-----------------T--A-----C--------C--------A--G--G-- 5111
H2A.IN.95.CRIK             --G-----------------------CA----A--G--------------G---------AT---T--------AT-------A-----T-----------A--C-----------A-----------------T--A-----C------AG---------A--G--G-- 4862
H2A.GM.x.MCR35.AY509260    T-G-----------------C-----CA----A--G--------A-----A----------T---T-----A--AT-T--C--A-----T-----------A--C--A--------A--------------------A--------G---AG------T--A-----G-- 4562
H2A.GM.87.D194.J04542      T-G--------C-----CC-C-----CA----A--G-----G--------G---------AT---T--------AT-------A-----T-----------A--C-----------A--------------------A--------G---AG------T--A-----G-- 4561
H2A.GM.x.ISY.J04498        T-G-----------------C-----CA----A--G--G-----------A----------T---C-----A--AT-------A-----T-----------A--C-----------A-----------------------------G-------A---T--A--G--G-- 4558
H2A.GM.x.MCN13.AY509259    T-G-----------------C-----CA----A--G--------A-----A----------T---T-----A--AT-T--C--A-----T-----------A--C--A--------A--------------------A--------G---AG------T--A-----G-- 4562
H2A.GW.86.FG.J03654        --G-----------------C-----CA----A--G-----G--A-----A----------T---T-----A--A--------A--G--T-----------A--C--A--------A--------------T--T--A--------G-----C-----T--A--G--G-- 4555
H2A.GW.87.CAM2CG.D00835    T-------------------C-----CA----A--G--------------A----------T---C-----A--A--------A-----T-----C-----A--C--A--------A--------------------A--------G---AG------T--A-----G-- 5126
H2A.GW.x.ALI.AF082339      T-G----------G-----TC-----CA-C--A--G--------------A----------TT--T--------AT-------A-----T-----------A--C--A--------A--------------------A--------G---AGC-----T--A--G--G-- 5110
H2A.GW.x.MDS.Z48731        T-G-----------------C-----CA----A--G-----G--------A---------AT---T-----A--AT-------A-----T-----------A--C--A--------A-----C--------------A--------G---AG---------A--T--G-- 4561
H2A.SN.85.ROD.M15390       T-G--T--------------C-----CA----A--G--G--G--------A---------AT---T-----A--AT-------A-----T--------------C--A--------A-----------------T--A-----C------AG---------A-----G-- 4563
H2AB.CI.90.7312A.L36874    T-G--------C--------C------A----A-----T-----------G--T-------T---A-----A--A--------A-----C-----C-----A--------------A--------------T--T--------------------G-----A--T--G-- 5099
H2B.CI.88.UC1.L07625       --G--------C--C--C--C-----CA-C--A-----C--G-------CG--C------AT---------A--A-----C--------C--T--C-----A--------------A-----------C--T-----A-----C--------------------C--G-- 5098
H2B.CI.x.EHO.U27200        --G--------C--C-----C-----CA----A-----T--G--A----CA--C------AT---------A--A--A--C--A--G--C--T-----G--A--------------A-----------C--T--T-----------G-----C-----T--A--T--G-- 5088
H2B.GH.86.D205.X61240      -----------C--C-----C-----CA-C--A-C-AGT-----------A--C-------T---A-----A--A--T-----A-----C--T--C-----A-----A--------------------C--T--------------------C--C-----A--C----- 5090
H2B.JP.01.KR020.AB100245   -----------C--------C-----CA----A-----C--G--A----CA--C------AT---AG----A--G-----C--A-----T--T--------A--------------A-----------C--T--T--A-----------------C-----A--C--G-- 4240
H2AB.CM.03CM_510_03.NEW    --G--------C--C-----C-----CA-C--A-----T--G--------A--C-------T---A-----A--G--------A-----C--T--C-----A-----A--------A-----------C--T-----A--------G---AGC--C-----A--Y----- 4229
H2G.CI.92.ABT96.AF208027   -----T-----C--------C-----CA-C--------G--G--------A--C-------T---C--C-----AG-A-----A-----------C--C--------A--G-----A--T-----------T--T--A--G--A-----------A--T--A--T----- 4444
H2U.FR.96.12034.AY530889   T-G-----------------C-----CA----A--G--------------A----------T---T--------T--------A-----C--T--C--G--A--C--A-----G--A-----C-----------T--A--G--C--------G--------A-----G-- 4596
SMM.SL.92.SL92B.AF334679   --G--T-----C--C-----C--T--CA-C--C--G--G--G--------A---------TT---AG-G--A--GT-------A-----G-----C-----A--------------A-----C---G-CTTA--T--T--G--AA-C-------A---A----------- 4465
MAC.US.x.MAC239            -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4111
MAC.US.x.251               -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 5033
MAC.US.x.251_32H           -----------------------------C-------------------------------------------------------------------------------------------------------------------------------------------- 5035
MAC.US.x.251               -----------------------------C-------------------------------------------------------------------------------------------------------------------------------------------- 5010
MAC.US.x.17EC1.AY033233    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 5034
MAC.US.x.17EFR.AY033146    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 5034
MAC.US.x.1937.AY611495     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4140
MAC.US.x.2065.AY611493     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4140
MAC.US.x.MM142.Y00277      --G----------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4520
MAC.US.x.SMM142B.BD131285  --G----------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4520
MNE.US.82.MNE              --------------------C-----------------------------------------------------------------------------------------------------------------T--A-------------------------------- 4502
MNE.US.x.MNE027.U79412     --------------------C------A----------------------------------------------------------------------------------------------------------T--A-------------------------------- 4502
SMM.US.x.F236              --G--T--T-----------C--T--CA-A--A--------G--A--------C--------------T--A--G-----------G--T--T-----------C--A--------------------C--T--TT-A----CC--G-----------T--A-------- 5022
SMM.US.x.H9.M80194         --G--T--T-----------C--T--CA-A--A--------G--A--------T-----------R--T--A--G------R-------T--T-----------C--AR-G-----------G-----C--T--T--A----CC--------------T--A-----G-- 4507
SMM.US.x.PBJ14             --G--T--T-----------C--T--CA-A--A--------G--A--------T--------------T--A--G--------------T--T-----------C--A--G-----------G-----C--T--T--A----CC--------------T--A-----G-- 4801
SMM.US.x.PBJA.M31325       --G--T--T-----------C--T--CA-A--A--------G--A--------T--------------T--A--G--------------T--T-----------C--A--G-----------G-----C--T--T--A----CC--------------T--A-----G-- 4800
SMM.US.x.PBJ               --G--T--T-----------C--T--CA-A--A--------G--A--------T-----------R--T--A--G------R-------T--T-----------C--AR-G-----------G-----C--T--T--A----CC--------------T--A-----G-- 4507
SMM.US.x.PBJ               --G--T--T-----------C--T--CA-A--A--------G--A--------T--------------T--A--G--------------T--T-----------C--A--G-----------G-----C--T--T--A----CC--------------T--A-----G-- 4993
SMM.US.x.PGM53.AF077017    --G--T--T-----------C--T--CA-A--A--------G--A--------C-----------A--T-----G-----C--------T--T-----------C--A--------------------C--T--T--A----CC--G-----------T--A-----G-- 4951
SMM.US.x.SME543.U72748     --G--T--T-----------C--T--CA-A--A--------G--A--------C--------------T--A--G--------A--G--T--T-----------C--A--------------------C--T--TT-A----CC--G-----------T--A-------- 5037
STM.US.x.STM.M83293        --G-----G--------C--------CA-C-----G--G-----------A--C--------G--A-----A--A--------A-----T--------------C--A--------------C--------T--TT-A--G-CC--G-----C-----T--A--T----- 4681
MAC.US.x.80035.AY611486    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4140
MAC.US.x.81035.AY599200    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4141
MAC.US.x.85013.AY611490    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4074
MAC.US.x.92050.AY603959    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4134
MAC.US.x.92077.AY599201    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4141
MAC.US.x.93057.AY611492    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4139
MAC.US.x.93062.AY607704    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4126
MAC.US.x.95058.AY611494    -----Y-------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4128
MAC.US.x.95086.AY607703    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3997
MAC.US.x.95112.AY588946    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4130
MAC.US.x.96016.AY607701    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4141
MAC.US.x.96020.AY611488    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4133
MAC.US.x.96072.AY611491    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4145
MAC.US.x.96081.AY597209    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4143
MAC.US.x.96093.AY611489    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4141
MAC.US.x.96114.AY588945    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4136
MAC.US.x.96123.AY611487    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4030
MAC.US.x.96135.AY607702    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4126
MAC.US.x.97009.AY599199    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4009
MAC.US.x.97074.AY599198    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4009
MAC.US.x.r80025.AY576480   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3997
MAC.US.x.r90131.AY576481   -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4140
```

```
MAC.US.x.239      AAATTCAGTAGAAACCATAGTATTAATGGCAGTTCATTGCATGAATTTTAAAAGAAGGGGAGGAATAGGGGATATGACTCCAGCAGAAAGATTAATTAACATGATCACTACAGAACAAGAGATACAATTTCAACAATCAAAAAACTCAAAATTTAAAAATTTTCGGGTCT 5204
Pol               _A_N_S_V_E_T_I_V_L_M_A_V_H_C_M_N_F_K_R_R_G_G_I_G_D_M_T_P_A_E_R_L_I_N_M_I_T_T_E_Q_E_I_Q_F_Q_Q_S_K_N_S_K_F_K_N_F_R_
H2A.CI.88.UC2.U38293  T---A----G-----A------C-------------------------------------------------------C--G---------C----C--------T--C-----------A--------C-TC---AG------T-----T--C-----A--C-A----- 5287
H2A.DE.x.BEN.M30502   ----A--A-------A------C-G-----------------------------------------------------C-----------C----C--------T--C-----------A--------C-TC---AG------T-----T-----------C-A----- 5287
H2A.DE.x.PEI2.U22047  ----A--A-G-----A-------------------------------------------------------------C-----------C----C--T-------C-----------A--------C-TC--CG------T--------A--------C------- 5255
H2A.GH.x.GH1.M30895   ----A--A--------------C-----------------------------------------------------C-----------C----C--------T--C-----------A--------C-TC---AG------T-----T--------C--C-A----- 4731
H2A.IN.NNVA.NEW       ----A---------A-------------------------------------------------------------C---T---G---C----C-----------CG------------------C-TC---G-C-----T--------A--------------- 5281
H2A.IN.95.CRIK        ----A---------A------C------------------------------------------------------C---T---G---C----C--T-------C-----------A--------C-TC---G-C-----T--------A--------------- 5032
H2A.GM.x.MCR35.AY509260 ----A---------A------C-G-----ACA--C-----------------------------------------C-----------C----C--T-------C-----------A--------C-TC--CG-C-----T--------A--------------- 4732
H2A.GM.87.D194.J04542 ----A--A-------A------C-----------------------------------------------------C-----------C----------------C-----------A--------C-TC---AG------T-----T--------A--C-A----- 4731
H2A.GM.x.ISY.J04498   ----A--A-G-----A------C-------------------------------------------------------C----T------C--G-C--T-------C-----------A--------C-TC---G-------T--------A--------------- 4728
H2A.GM.x.MCN13.AY509259 ----A---------A------C-G-----ACA--C-----------------------------------------C-----------C----C--T-------C-----------A--------C-TC--CG-C-----T--------A--------------- 4732
H2A.GW.86.FG.J03654   ----A---------A-------------------------------------------------------------C-----------A----C--T-------C-----------A--------C-TC---G-------T--------A--------------- 4725
H2A.GW.87.CAM2CG.D00835 ----A---------A------C-----------------------------------------------------T-------C----C--T-------C-----------A--------C-TC---G-C-----T--------A--------------- 5296
H2A.GW.x.ALI.AF082339 ----A----G-----A------C-------------------------------------------------------C-----------C-C--C--T------T-C-----------A--------C-TC---A-------T-TG-------------C-C----- 5280
H2A.GW.x.MDS.Z48731   ----A---------A------C-----------------------------------------------------C---T-------C----C--T-------C-----------A--------C-TC---G-C-----T--------A--------------- 4731
H2A.SN.85.ROD.M15390  ----A--A------A------C-------A-----------------------------------G-----------T-----------C--T-------C-----------C-TC---G-C-----T--------A---G----------- 4733
H2AB.CI.90.7312A.L36874 -GTA---A------AG-T--GC---------C-----------------------------------------C--T-----G---C--G---------A--C-----------A--------CTTC---G-------T-T-------C-------C-A----- 5269
H2B.CI.88.UC1.L07625  -GTA---A----G--AG-T---C---------C---C---------------------------------------C--T-----G---C--G---------A--C-----------A-----G--CTTC---G-------T-T-------C-------C-A----- 5268
H2B.CI.x.EHO.U27200   -GTA---A----G--AG-T--G--------AC---C---------------------------------------C--T--------A--G-C--------A-------------A--------C-TC---A-------T-T------CC-------C------- 5258
H2B.GH.86.D205.X61240 -GTA---A----G--AG-T---C-------AC---C---------------------------------------C--T--------C--G---------A--C-----G-----A-----G--CTTC---G-------T-T-------C-------C-A----- 5260
H2B.JP.01.KR020.AB100245 -GTA---A----G--AG-T---C---------C-----------------------------------------C--T--------A--G---------A---G------G-----------C-TC---A-------G-T-------C-------C------- 4410
H2AB.CM.03CM_510_03.NEW -GTR---A----G--AG-T---C-R------AC---C---------------------------------------C--T--------C--G---------A--C-----G-----A-----G--CTTC---G-------T-T-------C-------C-A---Y- 4399
H2G.CI.92.ABT96.AF208027 -------A------------G----------A-----------------------------------------C--T--------A------------T-------------A--------------A-------T------------------------- 4614
H2U.FR.96.12034.AY530889 ----A--A-G--G--------CC---------C---C-------------------------------------T---------A-C----------AT-C----------A--------C-TC-----------T------------------------- 4766
SMM.SL.92.SL92B.AF334679 -G-G--------G---C-----C---------CA----------------------------------------T--------A--G-A--T--------C-------T---A-CC---ACTT-A-T---C----T---------C------------T- 4635
MAC.US.x.MAC239       ----------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4281
MAC.US.x.251          ------------------------------------CG----------------------C-----------------------------------------------------A-----------------------------------------------------------------------------------------------------------------------
```

```
                                                
```

Pol, Gag-Pol, and p31 Integrase end

```
MAC.US.x.239              ATGGAGGATACCGGAGAGGCTAGA...GAGGTGGCAT..AGCCTCATAAAATATCTGAAATATAAAACTAAAGATCTACAAAAGGTTTGCTATGTGCCCCATTTTAAGGTCGGATGGGCATGGTGGACCTGCAGCAGAGTAATCTTCCCACTACAGGAAGGAAGCCATTT 5539
Pol                       _H__M__E__D__T__G__E__A_...__R__E__V__##A
Vif                       __W__R__I__P__E__R__L__E__.__R__W__H_##_S__L__I__K__Y__L__K__Y__K__T__K__D__L__Q__K__V__C__Y__V__P__H__F__K__V__G__W__A__W__W__T__C__S__R__V__I__F__P__L__Q__E__G__S__H__L
H2A.CI.88.UC2.U38293      C------G-G--A-G---A-G--...----------..-----TG-C-----C--A-----C-G---A-----C---G-GGG---GC-------T-----CCAC-----G-----------------T--------G-----A---------A-A-G--A---T---C- 5622
H2A.DE.x.BEN.M30502       C------G-G--A-G---A-G--...--AA-------..GC---TG-C--G--C--------C-G---A-----C---G-GG---GC-------T-----CCAC-----G-----------------T--------G-----A----------A-G-AA---T---C- 5622
H2A.DE.x.PEI2.U22047      T------G----A-G---A-G--...--A-------..-----TG-C--G-----A---C-C-G---A--------GG--GG---G--------T-----CCA------G-----------------T--------G-----A------T----A-G-AAT--T--CC- 5590
H2A.GH.x.GH1.M30895       GA--GT-CC-GG-AG---A-G--...----------..--T--TG-C--G--C--A-----C-G---A-G---C---G-GG---GC-------T-----CCAC-----G--------G--------T--------G-----A-----T---A-A-G--A---T---C- 5066
H2A.IN.NNVA.NEW           T------G----A-G---A-G--...--A-------..-----TG-C--G-----A-----C-----A--------GG---G---G--------C-----CCA------G--G-------------T--------G-----A------T--A-CA-CAAT---A-CC- 5616
H2A.IN.95.CRIK            T----A-G----A-G---A-G--...--A-------..-----TG-C--G-----A-----C-----A---------G-------G--------C-----CCA------G--G-------------T--------G-----A------T--A-A--CAAT--T--CA- 5367
H2A.GM.x.MCR35.AY509260   T------G-G--A-G---A-G--...--AA------..-----TG-C--------A-----C-G---A---------G--G-C--G--------T-----CCA---A--A--------------T--------G-----A------T--A---G-AAC--T---C- 5067
H2A.GM.87.D194.J04542     C------G-G--A-G---A-G--...----------..-----TG-C--GC-C-----G--C-G---A-----CT--G-GG---GC-------T------CAC----A-----------------T--------G-----A---------G-A-G--A---T---C- 5066
H2A.GM.x.ISY.J04498       C------G-G--A-G---A-G--...--AA------..-----T--C--G-----A-----C-G---A---------G--C----GC-------T-----CCA------G--G-------------T--------G-----A------T--A-A-G-AAC--T---C- 5063
H2A.GM.x.MCN13.AY509259   T------G-G--A-G---A-G--...--AA------..-----TG-C--------A-----C-G---A---------G--G-C--G--------T-----CCA---A--A--------------T--------G-----A------T--A---G-AAC--T---C- 5067
H2A.GW.86.FG.J03654       T------G-G--A-G---A-G--...---A------..---T-GG-C--G--CT-------C-G---A-----CT--G-G-----A--------T------CA------G--G-------------T--------G-----A--------GA-A--GAAC--T---C- 5060
H2A.GW.87.CAM2CG.D00835   C------G-G--A-G--AA-G--...--A-------..-----TG-C--G-----A-----C-----AGG-------G--C----G--------T-----CCA------A-----------------T--------G-----A------T--AGA-G--AC--T-G-C- 5631
H2A.GW.x.ALI.AF082339     T------G-G--A-G---A-G--...--A-------..-----TG-C--G--C--A--G--C-G---A---------G-------G--------T-----CCA------G-----------------T--------G-----A------T----A-G-A----T---C- 5615
H2A.GW.x.MDS.Z48731       T------G----A-G---A-G--...--A-------..-----TG-C--G---T-A-----C-----A---------G--G---GC-------C-----CCA------G-----------------T--------G-----A------T-GA-A-G-AAC--T---C- 5066
H2A.SN.85.ROD.M15390      C------G-G--A-G---A-G--...--AA------..-----TG-C--G-----A-----C-----A--G-----G-------G--------T-----CCA------G-----------------T--------G-----A------T--A-A-G-AAC--T---C- 5068
H2AB.CI.90.7312A.L36874   G----A------A-GC--------...--AA-----C..--T--AG-C--G--------GC---G---AGG---AT-G---C----C-CT-----C-----CCAC-----G--------C-----C--------T-----T--------A-----C---A-A-G----GCATGCC- 5604
H2B.CI.88.UC1.L07625      ------------A-GC--------...--AA-----C..--T--GG-C--G--C-----GC---G---A-----GT-G---C----C-CT-----C--T--CCAC-----G--------C-----C--------T-----T--------A-----C---A-A----A-GCAT-C-- 5603
H2B.CI.x.EHO.U27200       G-----------ATGC--------...---------C..--T--A--T-----C-----------GG--A-----CT-G---C----C-CT-----T--T--CCA------A--------T-----T--------T-----T--------A--T--C--GA-A------GCA---C- 5593
H2B.GH.86.D205.X61240     ------------A-GC--------...----A----C..--T--G--T--G-----T--G----G---AGG---GT-G---C----C-CT-----C--T--CCAC-----A--------T-----T--------T-----T---A----A--T--C----A-CA-----GCATGGC- 5595
H2B.JP.01.KR020.AB100245  --------------GC--------...--AA-----C..-----AG-C------------C---G---A-----AT-GG--C----CACT-----T--T--CCA------A--------T-----T--------------T--------A-----C---A-A------GCAT--C- 4745
H2AB.CM.03CM_510_03.NEW   -----A------A-GC-------C...---A-----C..--T--G--C-----C--T--G--------AGG---GT-G---C----C-CT-----C--T--CCAC-----R--------T-----T--------T-----T--------A--T--C---G-CA---R-GCAY-CC- 4734
H2G.CI.92.ABT96.AF208027  ------------A-GC--A--G--...--A--------..--T--G-------TCT-A--G-T------A-----C-----G----CGGTT-----T--T---CAC--A--T--------T--------T--------G-----A------T--ACAA---A-GCA----- 4949
H2U.FR.96.12034.AY530889  T------G-G---AGA---T----...---A-----C..TC---G--C---C-C----.G-----C--A-----G-----G--A--G--T-----------CAC-----A--------T--------T--------G-----A--T--T---A-AA-----TCA----- 5100
SMM.SL.92.SL92B.AF334679  ---------CAGTAGA--AA---TGGA-C-C-------..---TG---T----TC-AT--G--------A-G---G---G-G--A-CC-----------A---CAC--A--A------------AC--AGC-TCT--G--G--T-----CT-GG----G-----T--C-- 4973
MAC.US.x.MAC239           ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4616
MAC.US.x.251              ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 5538
MAC.US.x.251_32H          ------------------------...----------..----------------------------------------------------------------A------------------------------------------------------------------ 5540
MAC.US.x.251              ------------------------...----------..------------------------------------------------------------------------------------------------------------C---------------- 5515
MAC.US.x.17EC1.AY033233   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 5539
MAC.US.x.17EFR.AY033146   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 5539
MAC.US.x.1937.AY611495    ------------------------...----------..-------------------------------------------------------------------------------------------------------------T-------------K-- 4645
MAC.US.x.2065.AY611493    ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4645
MAC.US.x.MM142.Y00277     ------------------------...----------..--------------T----------------------------C-------------------CA------------------------------------------------------------------ 5025
MAC.US.x.SMM142B.BD131285-------------------------...----------..--------------T----------------------------C-------------------CA------------------------------------------------------------------ 5025
MNE.US.82.MNE             ------------------------...----------..-----------------------------------------------------------------CA--------------------------------------------------------A---AA-----A-- 5007
MNE.US.x.MNE027.U79412    ------------------------...----------..-----------------------------------------------------------------CA--------------------------------------------------------A---A------G-- 5007
SMM.US.x.F236             T-----------------------...----------..--------C---C-C-----------C--------C-----G-T--C---T-----------CA---A--T-----------------T-------------T------T--AG---T-AG-CT----- 5527
SMM.US.x.H9.M80194        T------------G----------...--A-------..--------C---C-C-----------C--------C-----G---C---T-----A------CA---A--T-----------------T-------------------GT--AG---T-AGGCT----- 5012
SMM.US.x.PBJ14            T------------G----------...--A-------..--------C---C-C-----------C--------C-----G---C---T-----A------CA---A--T-----------------T--------------------T--A----T-A-GCT----- 5306
SMM.US.x.PBJA.M31325      T------------G----------...--A-------..--------C---C-C-----------C--------C-----G---C---T-----A------CA---A--T-----------------T--------------------T--A----T-A-GCT----- 5305
SMM.US.x.PBJ              T------------G----------...--A-------..--------C---C-C-----------C--------C-----G---C---T-----A------CA---A--T-----------------T-------------------GT--AG---T-AGGCT----- 5012
SMM.US.x.PBJ              T------------G----------...--A-------..--------C---C-C-----------C--------C-----G---C---T-----A------CA---A--T-----------------T--------------------T--A----T-A-GCT----- 5498
SMM.US.x.PGM53.AF077017   T----A-------G----------...--A-------..--------C-----CT------T--G-----G--C-----G---C---T-----A------CA---A--T-----------------T------------------------AGA--T-A-TCT----- 5456
SMM.US.x.SME543.U72748    T-----------------------...----------..--------C---C-C-----------C-----G--C-----G-T--C-----------------CA---A--T-----------------T-------------T------T--AG---TAAG-CT----- 5542
STM.US.x.STM.M83293       C-----------A-G------G--...----------..------------C--T-A--G-----C--------G---AGC----CA--------------CCA------T--G-----------------T-----T-----G--T--T--CT-G--A-G--A-GCA--CC- 5186
MAC.US.x.80035.AY611486   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4645
MAC.US.x.81035.AY599200   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4646
MAC.US.x.85013.AY611490   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4579
MAC.US.x.92050.AY603959   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4639
MAC.US.x.92077.AY599201   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4646
MAC.US.x.93057.AY611492   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4644
MAC.US.x.93062.AY607704   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4631
MAC.US.x.95058.AY611494   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4633
MAC.US.x.95086.AY607703   ------------------------...----------..--------------------------------------------------------------C-------------------------------------------------------------------- 4502
MAC.US.x.95112.AY588946   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4635
MAC.US.x.96016.AY607701   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4646
MAC.US.x.96020.AY611488   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4638
MAC.US.x.96072.AY611491   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4650
MAC.US.x.96081.AY597209   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4648
MAC.US.x.96093.AY611489   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4646
MAC.US.x.96114.AY588945   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4641
MAC.US.x.96123.AY611487   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4535
MAC.US.x.96135.AY607702   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4631
MAC.US.x.97009.AY599199   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4514
MAC.US.x.97074.AY599198   ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4514
MAC.US.x.r80025.AY576480  ------------------------...----------..--------------------------------------------------------------C-------------------------------------------------------------------- 4502
MAC.US.x.r90131.AY576481  ------------------------...----------..----------------------------------------------------------------------------------------------------------------------------------- 4645
```

```
MAC.US.x.239      AGAAGTACAAGGGTATTGGCATTTGACACCAGAAAAAGGGTGGCTCAGTACTTATGCAGTGAGGATAACCTGGTACTCAAAGAACTTTTGGACAGATGTAACACCAAACTATGCAGACATTTTACTGCATAGCACTTATTTCCCTTGCTTTACAGCGGGAGAAGTGAGAA 5709
Vif               _E_V_Q_G_Y_W_H_L_T_P_E_K_G_W_L_S_T_Y_A_V_R_I_T_W_Y_S_K_N_F_W_T_D_V_T_P_N_Y_A_D_I_L_L_H_S_T_Y_F_P_C_F_T_A_G_E_V_R_
H2A.CI.88.UC2.U38293  ---GA----G-CA-----A-CC-A--G----------A------TCCT--C--T----A--AC----------TA--G-A--G--C----------T--C---G---G---------CC--A-A--------C-------TG-------G-G--A--T-----A---- 5792
H2A.DE.x.BEN.M30502   ----A----G-CA-----A-CC-A-------------A------TCCT--C-------A--AT----------TA--G-A--G--C----------T--C---G---G---------CC--A-A----------------T---------G--A--T-----A---- 5792
H2A.DE.x.PEI2.U22047  ---GA----G-CA-----A-CC-A-------------A------TCCT----------A--A----------TA--G---GG--C----------T--C---G---G-------TCCC--A-A----------------T----T-----G----T-----A---- 5760
H2A.GH.x.GH1.M30895   ---GA----G-CA-----A-CC-A--G----------A------TCCT--C--T-------A----------TA--G-A-GG--C----------T--C---G------------CC--A-A----------------T---------G--A--T-----A---- 5236
H2A.IN.NNVA.NEW       ---GA----G-CA-----A--C-A-------------A-----TTCCT----------A--------T-----TA--G-A--A--C----------T--C---G------------TCCC--A-A---G----------G------------A--T-----A---- 5786
H2A.IN.95.CRIK        ---GA----G-CA-----A--C-A--------G----A-----TTCCT----------A--------------A--G----T--C----------T--C---G---G-A-----TCCC--A-A----------------G------------A--T-----A---- 5537
H2A.GM.x.MCR35.AY509260  ----A----G-CA-----A-CC-A--G----------A------TCCT-----T---A--A--G--T-----TA-GG-A-GG--C----------T--C---G---G-------TCCC--A-A----------------T------------A--T-----A---- 5237
H2A.GM.87.D194.J04542 ---GA----G-CA-----A-CC-A-------------A------TCCT--C--T---A---T----------TA--G-A--G--C----------T--C---G---G-------TCCC--A-A--C-----------T---------G--A--T-----A---- 5236
H2A.GM.x.ISY.J04498   ---GA----G-CA-----A-CC-A-------------A------TCCT-----T---A--A--G--T---------G-AGGG--C----------T--C---G---G-------CCC--A-A--C-----------T---------G--A--T-----A---- 5233
H2A.GM.x.MCN13.AY509259  ----A----G-CA-----A-CC-A--G----------A------TCCT-----T---A--A--G--T-----TA-GG-A-GG--C----------T--C---G---G-------TCCC--A-A----------------T------------A--T-----A---- 5237
H2A.GW.86.FG.J03654   G--GA-----CA-----A-C--A-------------A------TCCT--C-CT-------A-----T-----TA--G-A--G--C----------T--C---G---G------T-CCC--A-A----------------T---------G--A--T-----A---- 5230
H2A.GW.87.CAM2CG.D00835  ---GA----G-CA-----A-C--A-------------A------TCCT----------A--A--G--T------A--G-A--G--C----------C--C---G---G-------CCC--A-A-----T---------T------------A--G-----A---- 5801
H2A.GW.x.ALI.AF082339 ---GA----G-CA-----A-CC-A-------------A------TCCT----------A--A-----T-----TA--G-A--G--C----------T--C---G---G-------TCCC--A-A---G----------T---------G--A--T-----A---- 5785
H2A.GW.x.MDS.Z48731   ---GA----G-CA-----A-CC-A-------------A-----TTCCT-A--------A---C-G--T------A--G-A--A--C----------T--C---G---G-------TCCC--A-A----------------A------------A--T----CA---- 5236
H2A.SN.85.ROD.M15390  ---GA----G-CA-----A-C--A-------------A------TCCT-----T---A--A-----T------A--G-A--G--C----------T--C---G---G------TG-CC--A-A----------------------------A--T-----A---- 5238
H2AB.CI.90.7312A.L36874  ---G--C-----------A--C----C--G----GG--A-TCT-G---T-C-----T-----AT----A-----TGA-G---G-----AC-------------TG-TGTG------CAGC--T-----G-G---------T-------------CAAT-----A--G- 5774
H2B.CI.88.UC1.L07625  ---G--C-----------A--C----C--G----GG--A-TCT-G--CT-C-----T--A--AC----A-----TAA--G--G-----AT-------G-----TG--GTG------CAGC--------G-GT-------T------------CAAT-----A--G- 5773
H2B.CI.x.EHO.U27200   ------C-----A--C---A-CC----C-------GG--A-TCT-G---T-C-----T--A--AC----A-----TGAG-G--G-----AT----------T--TG-TGTA------CGA---------G-GT--------T---CG--------TAAT-----A--G- 5763
H2B.GH.86.D205.X61240 ------C-----A------A-CC-A--C-------GG--A-TCT-G--CT-C-----T--A--AC----A-----TGAG-G--------AT-------------TG-TGTG------CAGC--------G-GT-------T--------T---CAAT-----A--G- 5765
H2B.JP.01.KR020.AB100245  ---GA-C-----A------A-CC----C--G----GG--A-TC--G---G-C-----T--A--AC----A-----TAAG--A-GT--C-AT--------G-----TG-TGTA------CACC--T-----G-GT--------TT------------TAAT-----A--G- 4915
H2AB.CM.03CM_510_03.NEW  ---R--C-----A------A-CC-A--C--G----GG--ACTCT-G--C--C-----T--A--AC----A-----YGAG-G--------AT-------G-----TG-TGTG------CAGC--------G-GT-------T------------CAAT-----A--G- 4904
H2G.CI.92.ABT96.AF208027  G---A-------------A-C-----C--T-----G--A---T-A--CCAG--------A--AT----------A---GA--A----AT--------G------G-GACA-----TCAA---T-----G-AT-------TGA------------T--T-----A---- 5119
H2U.FR.96.12034.AY530889  G--G--G----CC------A------------G--G--A-----ATCCT----------A---T----T------A---GAGGT--------------------TG-T-GC------CAA---------G-G-G-------T----------------T--------G- 5270
SMM.SL.92.SL92B.AF334679  G--G-------T---C---A--C----C-----------A---T-ATCA-GC--------T-----T--T---------G-A--A----------------C---G-TGTA----TCAG---GCA-----T-----------------G-----CAC-CG------C 5143
MAC.US.x.MAC239       ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 4786
MAC.US.x.251          ------------------------------------------------------------------------------G-------------------------------------------------------------------------------------------------------------------
```

```
                              Vpx start
MAC.US.x.239              GGGCCATCAGGGGAGAACAACTGCTGTCTTGCTGCAGGTTCCCGAGAGCTCATAAGTACCAGGTACCAAGCCTACAGTACTTAGCACTGAAAGTAGTAAGC......GATGTCAGATCCCAGGGAGAGAATCCCACCTGGAAACAGTGGAGAAGAGACAATAGGAGAGGC 5873
Vpx                                                                                                                    M  S  D  P  R  E  R  I  P  P  G  N  S  G  E  E  T  I  G  E  A
Vif                       R__A__I__R__G__E__Q__L__L__S__C__C__R__F__P__R__A__H__K__Y__Q__V__P__S__L__Q__Y__L__A__L__K__V__V__S__.__.__D__V__R__S__Q__G__E__N__P__T__W__K__Q__W__R__R__D__N__R__R__G_
H2A.CI.88.UC2.U38293      -A--------A--G---A-GT-AT----C-------AC-AT--CCA---------AGCA---------TCA--T--A---C----C--AGTG-----GCAACAA...A---A----C-----A--A--GG---G--A-------------------C--T-------- 5959
H2A.DE.x.BEN.M30502       -A--------A--G---A-GT-AT----C-------AC-AT--CCA---------AGCA---------TCA--T--A---C----C--AGT-------CAACAA...A---A----C-----A--A--GG-A--G--A-------------------C--T-------- 5959
H2A.DE.x.PEI2.U22047      -A--------A--G---A-GT-A-----C-------AT-A---CCA---C----GA-CTA-------GTTA--C--A-TTC-G--CT-AGT------GCAACAA...A---G----C-----AA-A-C-G-A-----A------G----C----TA--T--T-------- 5927
H2A.GH.x.GH1.M30895       -A--------A--G---A-GT-AT----C-------AC-AT--CCA---------AGTA---------TCA--T--A---C----C--AGTG-----GCAACAA...A---A----C-----A--A--GG-A--G--A-------------------C--T-------- 5403
H2A.IN.NNVA.NEW           -A--------A--G---A--G-AT----C-------AT-AT--CCAG--C------AC---------G-CA--T--A-TT--G--CT-AGT-------CAACAA...A---A-------------A----C-G-A-----A---------C--C--------T--C-------- 5953
H2A.IN.95.CRIK            -A--------A--G---A-GA-AT----C-------AT-AT--CCAG--C----GAAC---------G-CA--T--A-TT--G--CT-AGT------GCAACAA...A---A-------------A----CGG-A-----A---------C--C--------T--C-------- 5704
H2A.GM.x.MCR35.AY509260   -A--------A--G---A-GT-AT----C-------AT-AT--CCA---C----GAGC---------GTCA--C--A-TT--G--CT-AGTG-----GCAGCAA...A---A----C-----A--A-CGG-A--C--A------------C--------T--C-A------ 5404
H2A.GM.87.D194.J04542     -A--------A--G---A-GT-AT----C-------AC-A---CCA---------AGCA--------TCA--T--A---C----C--AGTG-----GCAACAA...A---G----C-----A--A--GG-G--G--A-------------------C--T-------- 5403
H2A.GM.x.ISY.J04498       -A---------------A-GTCAT----C-------AT-AT--CCA---C--------C-A-------GTCA--C--A-TTC-G--CT-AGT-------CAGCAA...A---A--A-C-----A----C---A-----A------------C-----A--T--C-A------ 5400
H2A.GM.x.MCN13.AY509259   -A--------A--G---A-GT-AT----C-------AT-AT--CCA---C----GAGC---------GTCA--C--A-TT--G--CT-AGTG-----GCAGCAA...A---A----C-----A--A-CGG-A--C--A------------C--------T--C-A------ 5404
H2A.GW.86.FG.J03654       -A--------A--G---A--T-AT----C-------AA-A---CC----G----GA-C---------GTCA--T--A-TTC-G--CT-AGTG-----GCAACAA...A---A---------------A-C-G-G-----A---------C--C--G-----T--C-A------ 5397
H2A.GW.87.CAM2CG.D00835   -A--------A--G---A-GT-AT----C-------AA-AT--CC----C----GA-C----------TCA--T--A-TTC-G--CT-AGTG-----GCAACAA...A---A----C-----A----CCG-A-----A------------C--------T--C-A------ 5968
H2A.GW.x.ALI.AF082339     -A--------A--G---A-GT-AT----C-------AT-AT--CCAG--C--------CA--------GTCA--C--A-TTC-G--CT-AGTG-----GCAACAA...A---G--A-C-----A--A-C-G-A-----A------------C--------T--C-------- 5952
H2A.GW.x.MDS.Z48731       -A--------A--G---A-GT-AT----C-------AT-AT--CCAG--C-----A-CT--------G-CA--T--A-TTC-G--CT-AGTG-----GCAACAA...A---G----C-----A----C-G-A-----A---------C--C--------T--C-AGC---- 5403
H2A.SN.85.ROD.M15390      -A--------A--G---A-GT-AT----C-------AT-AT--CC---------GAGC---------GTCA--T--A-TTC-G--CT-AGTG-----GCAACAA...A---A----C-----A----C-G-A-----A---------C--C--------T--C-------- 5405
H2AB.CI.90.7312A.L36874   -A---------------A-GA-AT----C-A-----AC-A---ATC------CG--GGG-----------TT------TTC----CT-A-G---T---CAAGAAGGAA-AAATG---------------G-G-----A---------C-AC----------G----------- 5944
H2B.CI.88.UC1.L07625      -A---------------A-GA-AT----C-A-----AC-AT--ATC------CG-AGGG------------T------TTC----C--A-GG--CA--CAGGAAGGAA-A-ATG---------------G-G-----A---------C-AC-----A-----------A-- 5943
H2B.CI.x.EHO.U27200       -A---------------A-GA-AT----CCA-----AC-A---ATC--------CAGG-----------TT------TTC----C--A-G---T---CAAGAAGGAA-A-ATG---------------G-A-----A---------C--C-----A--G---------- 5933
H2B.GH.86.D205.X61240     -A---------------A-GA-AT----C-A-----AC-AT--ATC------CG-AGGG------------T------TTC----C--A-GG--C---CAGGAAGGAA-AAATG---------------G-G-----A---------C-AC-----A--G-------A-- 5935
H2B.JP.01.KR020.AB100245  -A---------------A-GA-AT----C-A-----AC-A---ATC---C----GAGGA-----------TT------TTC----C--A-GG--T--GCAAGAAGGA-GAAATG--C-----------G-G--G-A---------C--C-----A-----------A-- 5085
H2AB.CM.03CM_510_03.NEW   -A---------------A-GA-AT----C-A-----AC-AT--ATC------CG-AGGG------------T------TTC----C--A-GG--CA--CAGGAAGGAA-AAATG-----------A-----A-----A--G-----C-AC-----A-----------A-- 5074
H2G.CI.92.ABT96.AF208027  -A--------------G--GA-AT----C-------AC-AT--A-C---------AAGG-----G--T---T-G--A-TTC----CT-AC-------GCAAAAA...-GACATG-----A---------G-M-----A-A------C-AC--G--------C-------- 5286
H2U.FR.96.12034.AY530889  -A--------A---------GT-AT----C-------AC--T--ATCG-------GAC-G---------TCTT----A-TTC-------C-G--C-T-GCAA......----GG------A--A---------------A-----TTCA-A------A--TG---------- 5434
SMM.SL.92.SL92B.AF334679  AA--------A-----G--GG--T-A--C-A---TG-C-A-G--GT---C---C-C-CTAGT--C-A----------CTG--G--C--A--G------TTACAA...A---A----C--A-----A-----------A----------------A-------------- 5310
MAC.US.x.MAC239           -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4950
MAC.US.x.251              -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 5872
MAC.US.x.251_32H          ---T-------------------C-------A-----------------G---------------T-------------------G---------T......T--------------------------------------------------G-------A-- 5874
MAC.US.x.251              -------------------------------A-----------------G-------------------------------A-------------------......--------------------------------------------------------------- 5849
MAC.US.x.17EC1.AY033233   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 5873
MAC.US.x.17EFR.AY033146   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 5873
MAC.US.x.1937.AY611495    -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4979
MAC.US.x.2065.AY611493    -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4979
MAC.US.x.MM142.Y00277     ------------------G-------------------------A------------C------------T-------------------G---------T......C----------------------------------------------------------A-- 5359
MAC.US.x.SMM142B.BD131285 ------------------G-------------------------A------------C------------T-------------------G---------T......C----------------------------------------------------------A-- 5359
MNE.US.82.MNE             -----------------------------------------------------AC-----------T-------------------G---------T......T-------------A-------------------------------------------A-- 5341
MNE.US.x.MNE027.U79412    -----------------------------------------------------A------------T-------------------G---------T......T-------------A-------------------------------------------A-- 5341
SMM.US.x.F236             A--------------GA--T---------------A--------A---------AA-T-------------------TC-------A-C---------T......C------------------G---------------------------A---------A---- 5861
SMM.US.x.H9.M80194        -----------------A--T---------------A--------A----------A-T-------------------TC-------A-C---------T......C-----RSR-----------R---------------------------A--RR----C----- 5346
SMM.US.x.PBJ14            -----------------A--T---------------A--------A----------A-T-------------------TC-------A-C-----G--......C----------------------------------------------A-------C----- 5640
SMM.US.x.PBJA.M31325      -----------------A--T---------------A--------A----------A-T-------------------TC-------A-C-----G--......C----------------------------------------------A-------C----- 5639
SMM.US.x.PBJ              -----------------A--T---------------A--------A----------A-T-------------------TC-------A-C-----G--......C-----RSR-----------R---------------------------A--RR----C----- 5346
SMM.US.x.PBJ              -----------------A--T---------------A--------A----------A-T-------------------TC-------A-C-----G--......C----------------------------------------------A-------C----- 5832
SMM.US.x.PGM53.AF077017   ----------------GA--T---------------A--------A----------A----------------------TC-------A-C-----G--T......C-------------A-----G------------------------A--------------- 5790
SMM.US.x.SME543.U72748    -A--------------GA--T---------------A--------A---------AA-T-------------------TC-------A-C---------T......C-------------A-----G--------------------G--A-------A------ 5876
STM.US.x.STM.M83293       -A---------------A-GT-A------------------A---A----------A-------------T------------TT-A--------GGAA......C-------------A--------A--G--C-----------------G------AG--A-- 5520
MAC.US.x.80035.AY611486   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4979
MAC.US.x.81035.AY599200   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4980
MAC.US.x.85013.AY611490   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4913
MAC.US.x.92050.AY603959   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4973
MAC.US.x.92077.AY599201   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4980
MAC.US.x.93057.AY611492   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4978
MAC.US.x.93062.AY607704   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4965
MAC.US.x.95058.AY611494   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4967
MAC.US.x.95086.AY607703   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4836
MAC.US.x.95112.AY588946   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4969
MAC.US.x.96016.AY607701   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4980
MAC.US.x.96020.AY611488   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4972
MAC.US.x.96072.AY611491   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4984
MAC.US.x.96081.AY597209   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4982
MAC.US.x.96093.AY611489   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4980
MAC.US.x.96114.AY588945   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4975
MAC.US.x.96123.AY611487   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4869
MAC.US.x.96135.AY607702   -----------------------------------------------------------------------------------------------------......----Y---------------------------------------------------------- 4965
MAC.US.x.97009.AY599199   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4848
MAC.US.x.97074.AY599198   -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4848
MAC.US.x.r80025.AY576480  -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4836
MAC.US.x.r90131.AY576481  -----------------------------------------------------------------------------------------------------......--------------------------------------------------------------- 4979
```

```
                                                                                                                                                  Vif end
MAC.US.x.239                 CTTCGAATGGCTAAACAGAACAGTAGAGGAGATAAACAGAGAGGCGGTAAACCACCTACCAAGGGAGCTAATTTTCCAGGTTTGGCAAAGGTCTTGGGAATACTGGCAT..GATGAACAAGGGATGTCACCAAGCTATGTAAAATACAGATACTTGTGTTTAATACAAAA 6041
Vpx                          __F__E__W__L__N__R__T__V__E__E__I__N__R__E__A__V__N__H__L__P__R__E__L__I__F__Q__V__W__Q__R__S__W__E__Y__W__H_##_D__E__Q__G__M__S__P__S__Y__V__K__Y__R__Y__L__C__L__I__Q__K_
Vif                          _L__R__M__A__K__Q__N__S__R__G__D__K__Q__R__G__G__K__P__P__T__K__G__A__N__F__P__G__L__A__K__V__L__G__I__L__A__##*_
H2A.CI.88.UC2.U38293         ------G------G----G--TA----A-CCT-G-----------------------G--CC-A-----C-----------G-----------C---AG----------..------------C-----T-------CAC---G--T-----T-----CC----G----- 6127
H2A.DE.x.BEN.M30502          ------G------G-G--G--CA----A-CCT-------G-----A--G-----T--G--CC-A-----C-----------G-----------C---AG---T------..------------------G-------CAC---G--T-----T-----CC----G----- 6127
H2A.DE.x.PEI2.U22047         T----CT------G-A--G-TG-------CC--------------A--G--------G--TC-------T-----------G-----------C---AG----------..-----C-T-----------A---T--CAC---G--T-----T---C-C-----G--GT- 6095
H2A.GH.x.GH1.M30895          ------G------G----G--TA----A-CCT-------------A--G-----T--G--CC-A-----C-----------G-----------C---AG----------..-----C--------------------CAC---G--T-----T-----CC-G--G--G-- 5571
H2A.IN.NNVA.NEW              -----CT------G-G--G--G-----A-CC--------------A--G--T-----G--C--------C-----------G-----------C----G----------..------------------G----T--CAC---G--T--------A--C--G--G--G-- 6121
H2A.IN.95.CRIK               -----CT------G----G--G-----A-CC-----------------G--T-----G--C-----A--C-----------C-----G-----C----G----------..------------------G----T--CAC------T------C----C--G--G--G-- 5872
H2A.GM.x.MCR35.AY509260      ----A-T------G----G--GA----A-CC--------------A--G--T-----G--CC-A-----T-----------G-----G-----C---AG----------..-------------------A---T--CAC---G--T-----T-----C--------G-- 5572
H2A.GM.87.D194.J04542        ------G------G----G--TA----A-CCT-------G-----A--G-----T--G--CC-A-----C-----------G-----------C----C---T------..------------------A----T--CAC---G--T-----T-----CA----G--G-- 5571
H2A.GM.x.ISY.J04498          ------T------G----G--G-----A-CC--------------A--G--------G--C--------T-----------G-----------C---AG----------..-----G-------------G------CAC---G--T-----T-----C-----G--G-- 5568
H2A.GM.x.MCN13.AY509259      ----A-T------G----G--G-----A-CC--------------A--G--T-----G--CC-A-----T-----------G-----G-----C---AG----------..-------------------A---T--CAC---G--T-----T-----C--------G-- 5572
H2A.GW.86.FG.J03654          -----C-------G----G----------CT--------------A--G--T----AC---C-A-----C-----------C-----G-----C---AG----------..------------------A-------CAC------T-----T-----C--------GT 5565
H2A.GW.87.CAM2CG.D00835      -----CC------G----G--------A-CT--------------A--G--T-----G--CC-------T-----------G-----------C---AG----------..-------------------A---T--CACC--G--T-----T-----C-----G--G-- 6136
H2A.GW.x.ALI.AF082339        ------G------G----G----------CCT----------------G--T-----G--CC-------C-----------G-----------C---AG----------..-----------A-------A---T--CAC------T-----------C-----G--G-- 6120
H2A.GW.x.MDS.Z48731          -----CC------G----G--------A-CC-----------------G--T-----G--CC-A-----T-----------G-----------C---AG----------..------------------A----T--CAC---G--T-----T-----C-----G--G-- 5571
H2A.SN.85.ROD.M15390         -----CC-----------G--------A-CC-----------A--A--G--T--------CC-A--A--T-----------G-----G-----C---AG----------..------------------GA---T--CAC---G--T-----T-----CA-------G-- 5573
H2AB.CI.90.7312A.L36874      A----C-------G-A------A---TA--AC-C-----G--A--A--C-----TT-G--CC-A--A--T-----------C----------------C---T----G-..-----T--G--C------ATT-----CACC--G--T--G------CT-C-G-------- 6112
H2B.CI.88.UC1.L07625         A-----T------G-A------A--ACA---C-C-----G-TA--A--C------T-G--CC-A--A--C-----------C-----G----G-----C--------G-..--G-----G--C------AGT--T---ACC-----T----------TGC----G--G-- 6111
H2B.CI.x.EHO.U27200          A------------G-G-C----T----AC-TC-C-------TA--A--C--T---T-G--C-----A--T-----------C-----G-A--------C---T----G-..--G-----G--C------ATT--T---ACC-----T--G--T------C----G----- 6101
H2B.GH.86.D205.X61240        A----CT------G-A------A--ACA---C-C-----G-TA-----C-----TT-G--CC-A--A--T-----------C-----G----------C--------G-..--G-----G--C------ATT------ACC-----T----------TGC----G--G-- 6103
H2B.JP.01.KR020.AB100245     A------------G-A------A--ACA---C-C-------TA--T--C--T--T--G--C-----A--T-----------C-----G----------C---T-----C..--G-----G--C---A--ATT-----CACC--G--T------------C-G--G--G-- 5253
H2AB.CM.03CM_510_03.NEW      A-----K------G-A------A--ACA---C-C-----G-TA--A--C------T-G--CC-A--A--C-----------C----------G-----C--------G-..--------G--C------AGT--T--CACC-----T-----TC---TGC----G--G-- 5242
H2G.CI.92.ABT96.AF208027     A------------G-A--------MACA-----C-----GATA--A--C--T--TT----C--------T----------------G------C--------T------..--------G--C------G-T--T--CACT--G--T-----T----TGC-G--G----- 5454
H2U.FR.96.12034.AY530889     T---AGT------G-A-----T--GAGA--T--C-----G-CA--A--GC-A--T--T--C-----A--T-----------C-----------C---C---------G-..--------G--C------GA---T---ACT--G--T-----TC-C--C--G-----G-- 5602
SMM.SL.92.SL92B.AF334679     ----------T--C---AT-----G----CCC-G--TCA-ACA--A---C-A-----------A---T-------------C---AG-----G-------------GTA..-----G-----ATAC-------------C------T-----TG-ACAG-----G----- 5478
MAC.US.x.MAC239              -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5118
MAC.US.x.251                 -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 6040
MAC.US.x.251_32H             -------------------------------------------------------------------------------------------------------------..-------------------A--------------------------------------- 6042
MAC.US.x.251                 ---------------------------------------------A-----------------------G---------------------------------------..-------------------A---------------------------------G----- 6017
MAC.US.x.17EC1.AY033233      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 6041
MAC.US.x.17EFR.AY033146      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 6041
MAC.US.x.1937.AY611495       ---------------------------------------------------------R-----A---------------------------------------------..----------------------------------------------------------- 5147
MAC.US.x.2065.AY611493       -------------------------------------------------------------------T-----------------------------------------..----------------------------------------------------------- 5147
MAC.US.x.MM142.Y00277        ------G-----------------------------------------------------G------------------------------------------------..-------------------A-------AC------------------------------ 5527
MAC.US.x.SMM142B.BD131285------G-----------------------------------------------------G------------------------------------------------..-------------------A-------AC------------------------------ 5527
MNE.US.82.MNE                ---------------------------------------------A---------------------T-G---------------------------------------..-----------------G-A------------G-------------------------- 5509
MNE.US.x.MNE027.U79412       ---------------------------------------------A-----------------------G---------------------------------------..--------------------------------G-------------------------- 5509
SMM.US.x.F236                ---------------------------A--A--------G-CA--A-----T-----G--G-------------------------G------C---------------..------ATG--A------GA------CAC------T-----------C--G-----G-- 6029
SMM.US.x.H9.M80194           Y-----C------C--------S----A--A--------G-CA--A--G--T--TT-G--R--------------------C----G----K-C---------------..------ATG---------GTC-----CACT----R------------C--G-----G-- 5514
SMM.US.x.PBJ14               ------C------G-------------A--A--------G-CA--A--G--T--TT-G--G-----A--------------C----G------C------------T--..------ATG---------GTC-----CACT-----------------C--G-----G-- 5808
SMM.US.x.PBJA.M31325         ------C------G-------------A--A--------G-CA--A--G--T--TT-G--G-----A--------------C----G------C---------------..------ATG---------GTC-----CACT-----------------C--G-----G-- 5807
SMM.US.x.PBJ                 Y-----C------C--------S----A--A--------G-CA--A--G--T--TT-G--R--------------------C----G----K-C---------------..------ATG---------GTC-----CACT----R------------C--G-----G-- 5514
SMM.US.x.PBJ                 ------C------G-------------A--A--------G-CA--A--G--T--TT-G--G-----A--------------C----G------C---------------..------ATG---------GTC-----CACT-----------------C--G-----G-- 6000
SMM.US.x.PGM53.AF077017      A--------------------------A--A--------G-CA--A--G--T---T-G--G-------------------------G--------------------G-..------ATG---------GAG-----CAC------------------C--------G-- 5958
SMM.US.x.SME543.U72748       ---------------------------A-GA--------G-CA--A-----T-----G--G--------------------------------C---------------..------ATG--A------GA------CAC------T-----------C--G-----G-- 6044
STM.US.x.STM.M83293          A-----G------C-------G--------C------C-------A--------T-----G-----A--G-----------C----G----------------------..-----G---------C-CGGG--------T-----T-----------C--G-----G-- 5688
MAC.US.x.80035.AY611486      ---------------------------------R---------------------------------------------------------------------------..----------------------------------------------------------- 5147
MAC.US.x.81035.AY599200      -----------------------------------R-------------------------------------------------------------------------..-----G--------------M-------------------------------------- 5148
MAC.US.x.85013.AY611490      ---------------------------------------------------------------------G---------------------------------------..----------------------------------------------------------- 5081
MAC.US.x.92050.AY603959      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5141
MAC.US.x.92077.AY599201      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5148
MAC.US.x.93057.AY611492      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5146
MAC.US.x.93062.AY607704      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5133
MAC.US.x.95058.AY611494      ---------------------------------------------------------------------G---------------------------------------..----------------------------------------------------------- 5135
MAC.US.x.95086.AY607703      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5004
MAC.US.x.95112.AY588946      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5137
MAC.US.x.96016.AY607701      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5148
MAC.US.x.96020.AY611488      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5140
MAC.US.x.96072.AY611491      ---------------------------------------------------------------A---------------------------------------------..----------------------------------------------------------- 5152
MAC.US.x.96081.AY597209      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5150
MAC.US.x.96093.AY611489      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5148
MAC.US.x.96114.AY588945      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5143
MAC.US.x.96123.AY611487      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5037
MAC.US.x.96135.AY607702      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5133
MAC.US.x.97009.AY599199      -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5016
MAC.US.x.97074.AY599198      ---------------------------------------------------------------A-----G---------------------------------------..----------------------------------------------------------- 5016
MAC.US.x.r80025.AY576480     -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5004
MAC.US.x.r90131.AY576481     -------------------------------------------------------------------------------------------------------------..----------------------------------------------------------- 5147
```

```
                                                
```

```
                                                                                                            Tat exon 1 start
MAC.US.x.239              CATGGGATGAATGGGTAGTGGAGGTTCTGGAAGAACTGAAAGAAGAAGCTTTAAAACATTTTGATCCTCGCTTGCTAACTGCACTTGGTAATCATATCTATAATAGACATGGAGACACCCTTGAGGGAGCAGGAGAACTCATTAGAATCCTCCAACGAGCGCTCTTCATG 6363
Vpr                       P__W__D__E__W__V__V__E__V__L__E__E__L__K__E__E__A__L__K__H__F__D__P__R__L__L__T__A__L__G__N__H__I__Y__N__R__H__G__D__T__L__E__G__A__G__E__L__I__R__I__L__Q__R__A__L__F__M
Tat exon 1                                                                                                             _M__E__T__P__L__R__E__Q__E__N__S__L__E__S__S__N__E__R__S__S__C
H2A.CI.88.UC2.U38293      ....T----G-A--AGCTA----A-GA-T-TGT--A-A----------C-A--GG----------C---------T---T----AC--CT------------------A---------A--C--CA---G------A-----A-------C---C--- 6402
H2A.DE.x.BEN.M30502       T-G--AG---C------A-A--AAC----AGG---A-A--G-------C----G----------C---------T---T----CT-CT-----C-----------------------------A--C--CA---G------A--C---A-------C-----G-- 6463
H2A.DE.x.PEI2.U22047      --G-------G------A-A-GAA-C---AG----T-A-G--------------G--------C--------------A-T-----C--CT-----G-GC----------------C--AA-C--CA---G------ATG---G-------C-----G-- 6425
H2A.GH.x.GH1.M30895       T-G---G---C------A-AAGAA------G----A-A--G-------C-----G---------C---------TC--T-----C--CT-----C---G---------------A-C---A--C--CA---G-------------A-------C-----G-- 5904
H2A.IN.NNVA.NEW           --G-------G------A-A--AA-C--AAG----A-A---------------GG-------C--C--------G-----T-----C-GAT------C-C----T-----------C---A--T--CA---G--------G--T-G-------C--T----CA 6455
H2A.IN.95.CRIK            --G-------G------A-A--AA-C--AAG----A-A---------------GG-------C-----------------T-----C-TAT-C------G-----T--------------A--C--CA---G--------G--T-G-------C--T--T-CA 6205
H2A.GM.x.MCR35.AY509260   --G-------G---A--A-A--AA-C---AG-A--A-A--GA------------G-------C-----------------T-----C--CT----C---C------------------A--C--CA---G------ATG---A-------C--------- 5905
H2A.GM.87.D194.J04542     T-G--AG-ACC------A-A--AAC----A-G---A-C--G-------C---------------CT--------T---T-----C--CT-------------------------------A-----CA---G--------G---A-------C-----G-- 5904
H2A.GM.x.ISY.J04498       --G-------G------A-A--AA-----AG----A-A--G-------------G-------C-----------------T-----CT-CT-------C------------------A--C--CA---G-------GG---A-------C------C- 5901
H2A.GM.x.MCN13.AY509259   --G-------G---A--A-A--AA-C---AG-A--A-A--GA------------G-------C-----------------T-----C--CT----C---C------------------A--C--CA---G------ATG---A-------C--------- 5905
H2A.GW.86.FG.J03654       --G------C-------A-A--AA-C---AG----A-AG--------------GG-------C-----------------T-----C-GAT--------C---G--------------A--C--CA---G-------------G-------C-----GC- 5898
H2A.GW.87.CAM2CG.D00835   --G--------------A-A--AA-C---AG---CA-A----------CC----G--C----C-----------------T-----CGGA--------GC------------------AA----CA---G-------GG---G-------C-------C- 6469
H2A.GW.x.ALI.AF082339     AGG-AT..--G------A-A--A--C---AG----A-A---------------GG-------C-----A------T---T-----CGGCT--------C------------------AA----CA-G--G------ATGC--A-------C--------- 6453
H2A.GW.x.MDS.Z48731       --G-------G------A-A--AA-----AG---GA-A-C---------C-T--G-------C-GCA------------T-----C-GA---------C-----------T-------A--C--CA----------------T-A-------C-------CA 5904
H2A.SN.85.ROD.M15390      --G-------G---A--A-A--AA-CT--AG----A-A---------------G-------C-----------T---T-----C--AT--------C------------------A--C--CA---G------A-G---G-------C--T----C- 5906
H2AB.CI.90.7312A.L36874   -------A--G-------A---C--C-----G---A-A--GC------C----G-----------C-------------G-------CTT---T--C-G---G-------T------CA--G---------------AG--T-----G-----------C-C 6445
H2B.CI.88.UC1.L07625      -------A--G-------A---C--CA----G---A-A--GC------C----GG----------C-------------G--------TT------C-G---G-------T-------CA----------G------A--------G-----C------C-C 6444
H2B.CI.x.EHO.U27200       -G-----AC-G-----------C--C-----G---A-A---C------C----G-------C--C----AT-----------C---TT------C-----G-----A-T-------A----------G------A-C------G--G--A------C-C 6434
H2B.GH.86.D205.X61240     -G-----A--G---A---G----A-C-----G---A-A--GC------C----G-------------------------G-------CTT-----C-G---G-------T------CA----------G------A---------------C-----C-C 6436
H2B.JP.01.KR020.AB100245  -------A--G-------A---C--C-----G---T-A--GC------C---GG----------C-------------G--------TT-----CGG--G-----G--T---A---CA----------------------T-----G-----------C-C 5586
H2AB.CM.03CM_510_03.NEW   AG-----A--GA-AR---G----A-C-----G---A-A--GC------CY----G-------------------Y---G--------CTT-----C-G---GT-------T------CA--G--------G---G---A--------G-----C------C-C 5575
H2G.CI.92.ABT96.AF208027  ------------------A--A------------A-C----------C--------------C-------------G-------CT-------G----G---------------C-------------------A------G--GA----T------C-- 5779
H2U.FR.96.12034.AY530889  -TG----C--G-----GAGA--AA-A-----G-----A-----------G---GG-------------------------G-------CT--------------G-------T--------A----------G------A------G---A----AA-----C-C 5923
SMM.SL.92.SL92B.AF334679  ----------G---A--AGA--T--A---------T---------G--AC-G-----------C-----------T-----TT-A--A---T-CG-A---G---CTT-C--G--T--AA-A--A--G-----G---A-A----AG--AT-A---AAG--AT-G---T-A 5796
MAC.US.x.MAC239           ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5440
MAC.US.x.251              ------G-------------------------------G----G-------------------------------------------------------------------------------------------------G--G--A--------- 6362
MAC.US.x.251_32H          ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 6364
MAC.US.x.251              -----------------------------------T------------------------------------------------------------------------------------------------------------------------------ 6339
MAC.US.x.17EC1.AY033233   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 6363
MAC.US.x.17EFR.AY033146   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 6363
MAC.US.x.1937.AY611495    --------------------------Y---------R----------------------------------------------------------------------------------------------------------------------------- 5469
MAC.US.x.2065.AY611493    ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5469
MAC.US.x.MM142.Y00277     ----------G----------A-----A----------------------G------------GC-T-----C-------------------------------------------------------------------------------------A 5849
MAC.US.x.SMM142B.BD131285 ----------G----------A-----A----------------------G------------GC-T-----C-------------------------------------------------------------------------------------A 5849
MNE.US.82.MNE             -------------------------------G------------------------------------------G-------------------------------------------------A-----------G----------- 5831
MNE.US.x.MNE027.U79412    -------------------------------G------------------------------------------G--------T------G--------------------------------------A-----------G----------- 5831
SMM.US.x.F236             -------------------------A-----A--G---T-A----------CC-G---------------------------G-------CT----T--G-------------------A------------------------G--GA----A--------C 6352
SMM.US.x.H9.M80194        -G--------------R--A--A--------G----KA----------CC-G--C--K------------------TY--G-------CT-------G-------------------A----------G-----------------GG----A--------C 5837
SMM.US.x.PBJ14            -G-----------------A--A--------G---A-A----------CC-G--C---------------------------G-------CT-------G-------------------A----------G-----------------GA----A--------C 6131
SMM.US.x.PBJA.M31325      -G-----------------A--A--------G---A-A----------CC-G--C---------------------------G-------CT-------G-------------------A----------G-----------------GA----A--------C 6130
SMM.US.x.PBJ              -G--------------R--A--A--------G---AGA----------CC-G--C--K------------------Y--G-------CT-------G-------------------A----------G-----------------GA----A--------C 5837
SMM.US.x.PBJ              -G-----------------A--A--------G---A-A----------CC-G--C---------------------------G-------CT-------G-------------------A----------G-----------------GA----A--------C 6323
SMM.US.x.PGM53.AF077017   -G-----------------AA-A--------G---G-A-----------C-G-----C---------------------G-------CT-------G-------------------A-------------------G------------------------A 6282
SMM.US.x.SME543.U72748    -----------------------A--------G---A-A----------CC-G------------------------G-------CT----T--G-------------------A-------------------------A--GA----A--------C 6367
STM.US.x.STM.M83293       ----------G-----------A--------G---A-A---C-------CC-GCGG--------------------GC--G-------CT-------G-------------------A----------G------AG--------G-----C---------- 6011
MAC.US.x.80035.AY611486   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5469
MAC.US.x.81035.AY599200   -----------------------------------------------------------------------------------K---------------------G------------------------------------------------------ 5470
MAC.US.x.85013.AY611490   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5403
MAC.US.x.92050.AY603959   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5463
MAC.US.x.92077.AY599201   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5470
MAC.US.x.93057.AY611492   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5468
MAC.US.x.93062.AY607704   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5455
MAC.US.x.95058.AY611494   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5457
MAC.US.x.95086.AY607703   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5326
MAC.US.x.95112.AY588946   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5459
MAC.US.x.96016.AY607701   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5470
MAC.US.x.96020.AY611488   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5462
MAC.US.x.96072.AY611491   -----------------------------------------------------------------------------------------------------------------------------------------------R-------------- 5474
MAC.US.x.96081.AY597209   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5472
MAC.US.x.96093.AY611489   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5470
MAC.US.x.96114.AY588945   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5465
MAC.US.x.96123.AY611487   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5359
MAC.US.x.96135.AY607702   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5455
MAC.US.x.97009.AY599199   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5338
MAC.US.x.97074.AY599198   ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5338
MAC.US.x.r80025.AY576480  ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5326
MAC.US.x.r90131.AY576481  ------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5469
```

```
                                                                                                                      Vpr end                                                                            Rev exon 1 start
MAC.US.x.239                CATTTCAGAGGCGGATGCATCCACTCCAGAATCGGCCAACCTGGGGGAGGAAATCCTCTCTCAGCTATACCGCCCTCTAGAAGCATGCTAT..AACACATGCTATTGTAAAAAGTGTTGCTACCATTGCCAGTTTTGTTTTCTTAAAAAAGGCTTGGGGATATGTTATGA 6531
Vpr                         _H__F__R__G__G__C__I__H__S__R__I__G__Q__P__G__G__G__N__P__L__S__A__I__P__P__S__R__S__M__L_##*
Tat exon 1                  __I__S__E__A__D__A__S__T__P__E__S__A__N__L__G__E__E__I__L__S__Q__L__Y__R__P__L__E__A__C__Y_##_N__T__C__Y__C__K__K__C__C__Y__H__C__Q__F__C__F__L__K__K__G__L__G__I__C__Y__E
Rev exon 1                                                                                                                                                                                        _M__
H2A.CI.88.UC2.U38293        --CC------TG---G-T-G--G---A-----T---T--A-AA--A--A---C----TG-C------C----A--C-----G------AC-..---C-------------G-----CA-T-T----------C-G-----CT-G---------GC-C--------G---T 6570
H2A.DE.x.BEN.M30502         --C-------CG-----T-A--G---A-----T---T--A-AA--A--A---C----TG-C-----GC----A--C-----GC-----AC-..----A------------GCGA--CA-T-T----------C-G-----CTCG-----G--GC-C--A----CA----- 6631
H2A.DE.x.PEI2.U22047        --C-------CA-----T-AAAT---A-----T------A-AA------AG-C----T-------------AA--C-----G------A--..---C---T---------G--A-----T------------C-A------T-A------G--AC-C------------- 6593
H2A.GH.x.GH1.M30895         --CC------CA--G--T-A--G---A-----TA-----A-AA--A--A---C----T--C-----GC----A--C-----G-----TAC-..---T-------------GC----CA---T----------C-G-----CT-G---------GC-C--------G---C 6072
H2A.IN.NNVA.NEW             --C-------CA-----TGG------A--------T-----AA-A-----------------------T--AA--C------------G--..---T-G-----------G--A--C----T---C---G----G------T-AC-----G--AC-C------------- 6623
H2A.IN.95.CRIK              --C-------CA-----TGG------A--G-----------AA--------------------A--------A--C------------GC-..-G---T------------CTA--C-----------------G--C---T-AC-----G--GC-C------------- 6373
H2A.GM.x.MCR35.AY509260     --C-------CG-----T-GG-T---A-----T------A-A------A---C----T--C------C-T--A--C------C-----A--..----A-------------GGA--C---T---C-------C-G------T-A--C--G---GC-C--A---------- 6073
H2A.GM.87.D194.J04542       --CA------CG-----TGA--G---A----AG------A-AA--A--A--GC----TG-C-----GC----A--C-----G------AC-..---T-------------GC----CA-T------------C-G-----CT-G---------GC-C--------G---C 6072
H2A.GM.x.ISY.J04498         --C-------CA-----TGG------A-----T--------AA-----A-----------------------A--C------A-----A--..----A----T-------GGG---C----T----------C-G------T-A--C--G---GC-C------------- 6069
H2A.GM.x.MCN13.AY509259     --C-------CG-----T-GG-T---A-----T------A-A------A---C----T--C------C-T--A--C------C-----A--..----A-------------GGA--C---T---C-------C-G------T-A--C--G---GC-C------------- 6073
H2A.GW.86.FG.J03654         --C-------CA-----TGG------A-----T-----GA-AA-----------------------------A--C-----G------AC-..---T-C--------C--G--A--C-----G---------C-G------T-AC-----G--GC-C--------G---- 6066
H2A.GW.87.CAM2CG.D00835     --C-----G-CA-------A------A-----T------A--A----------------------------AA--C-------A---TG--..---T-------------GCGA--C-----T---------C-A--------A------G--GC-C------------- 6637
H2A.GW.x.ALI.AF082339       --C-------CA-----TGG--GT--A---G-T------A-AA-----A-----------------------A--C------A-----A--..-----------------GG-A--C---------------C-G------T-A--T--G---AC-C--------G---- 6621
H2A.GW.x.MDS.Z48731         --C-------CA-----TGG------A-----T-----GA-AA-----------------------------A--C------------A--..----AT-----------GCGA--C---------------A-A------T-A--C--G---GC-C--------G---- 6072
H2A.SN.85.ROD.M15390        --C-------CA-----TGG------A-----T-----GA-AA-----------------------------A--C------A-----A--..---T-------------GCGA--C-----------T---A-G--------A--C--G---GC-C------------- 6074
H2AB.CI.90.7312A.L36874     --C-----------T--TCAA-----A--G--T--A-----A-------------------------C----------GAGG------A--..--T--G-----------G--A------------------C-----------------G---C-T------------- 6613
H2B.CI.88.UC1.L07625        --C-------C---T--TCAA-----A--G-----A-----A------------------------------------T---------G--..--T-A------C--------A--C--T------------C-------C---------G---C-T--------G---- 6612
H2B.CI.x.EHO.U27200         --C--------G--T---CAA-----A--G--T--A----G-------------------------------------TA-G------AG-..--T--------C--------A----C-------------C----C------------G---C--------------- 6602
H2B.GH.86.D205.X61240       --C-------C---T--TCAA-----A--G--T--A---T-A---------------------A-----------C--TA-G------G--..--T--------C-----G--A--C---------------C-----------------G--TC-T------------- 6604
H2B.JP.01.KR020.AB100245    --C--------A--G---CAA-----G--G--T--A-----A----------------------T-------------T---------AGA..--T-AG-----C-----G--A--C--T------------C-------C---------G---C-T------------- 5754
H2AB.CM.03CM_510_03.NEW     --C-------C---T--TCGA-----A--G-----A--G--A---------------------R-----------C--TA-G------A--..--T-A------C-------RA--C---------------C------------G-------TC-T--------G---- 5743
H2G.CI.92.ABT96.AF208027    --C--TC----A-AC--T-GATG------G--T--AGGGAAC-CA--C-----------A---A-----------------G--G--T-C-..--T-A--T-----C------A--C-----T-----T-----A--C--------C--G---AC-C--AG--------- 5947
H2U.FR.96.12034.AY530889    --C--------G-----TCGG-----A--G------AGCAGCA-----------------A--A--G-----------T--G--G----C-..--T-A---------------A--------T---------C-A----------------G---C-T------------- 6091
SMM.SL.92.SL92B.AF334679    --C--T--GCAT-------C---T--A--G--T--A--G--AA-----CC------C--GGG-T--------T--G-----GA-G--T---..----GC--T-T------------------T---------C-------------GG--G---C-T--C--TAC----C 5964
MAC.US.x.MAC239             -------------------------------------------------------------------------------------------..----------------------------------------------------------------------------- 5608
MAC.US.x.251                --C-----------C-----------------T---------------A------------------------------------------..----------------------------------------------------------------------------- 6530
MAC.US.x.251_32H            -------------------A-------------------------------------------A-----------------G--G------..----------------------------------------------------------G------------------- 6532
MAC.US.x.251                -----T-------------A-------------------------------------------A--------------T--G--G------..----------------------------------------------------------G--A---------------- 6507
MAC.US.x.17EC1.AY033233     -------------------------------------------------------------------------------------------..----------------------------------------------------------------------------- 6531
MAC.US.x.17EFR.AY033146     -------------------------------------------------------------------------------------------..----------------------------------------------------------------------------- 6531
MAC.US.x.1937.AY611495      -------------------C-------------------------------------------------------------------K--..----------------------------------------------------------------------------- 5637
MAC.US.x.2065.AY611493      -------------------C-----------------------------------------------------------------------..----------------------------------------------------------------------------- 5637
MAC.US.x.MM142.Y00277       ---------A----C----G---T---------------------------------------A---------------------------..--------------C-------------------------------------------G-----------A------- 6017
MAC.US.x.SMM142B.BD131285   ---------A----C----G---T---------------------------------------A---------------------------..--------------C-------------------------------------------G-----------A------- 6017
MNE.US.82.MNE               --C-----------C----C------T------------T---A-------------------A-------------------A-------..--------------C-------------------------------------------G------------------- 5999
MNE.US.x.MNE027.U79412      --C-----------C----A------T------------T---A-------------------A---------------------------..-----G--------C------------------------------------------------------------------ 5999
SMM.US.x.F236               ---------A-T-----TGC-------------------T--....................................---G--G------..----AG-----C-----G-G---------T-----T---CA----------------G--T-----A-----C---- 6484
SMM.US.x.H9.M80194          -----------T-------GA-----------------GT----A---R--------------A-----------KG----G--G--T---..----AG--Y--C-----G--A--------T-----T---CA---------Y------G--------A-----C---- 6005
SMM.US.x.PBJ14              -----------T-------GA-----------------GT----A------------------A-----------------G--G--T---..----AG-----C-------G---------T-----T---CA----------------G--------A-----C-C-- 6299
SMM.US.x.PBJA.M31325        -----------T-------GA-----------------GT----A------------------A-----------------G--G--T---..----AG-----C-------G---------T-----T---CA----------------G--------A-----C---- 6298
SMM.US.x.PBJ                -----------T-------GA-----------------GT----A---R--------------A-----------K-----G--G--T---..----AG--Y--C-------G---------T-----T---CA---------Y------G--------A-----C---- 6005
SMM.US.x.PBJ                -----------T-------GA-----------------GT----A------------------A-----------------G--G--T---..----AG-----C-------G---------T-----T---CA----------------G--------A-----C---- 6491
SMM.US.x.PGM53.AF077017     --C--------A-----TGCT------------------AG-A-A--------------G---A-----------------G--G----C-..--T-----T--C--------A--C--T------------C-----------------G------------------- 6450
SMM.US.x.SME543.U72748      ---------A-T-----TGC-------------------T--A-A------------------A-----------------GC--------..----AG-----C-----G-G---------T-----T---CA----------------G--T-----A-----C---- 6535
STM.US.x.STM.M83293         --C-----------T---CG---------G--T--------A--A------------T-GG--A-----------A-----G-AG--T-G-..----AG----T---C-----A--------T--C------C-G------G----C------TC-T--A---ACC---- 6179
MAC.US.x.80035.AY611486     -------------------C-----------------------------------------------------------------------..----------------------------------------------------------------------------- 5637
MAC.US.x.81035.AY599200     -------------------------------------------------------------------------------------------..----------------------------------------------------------------------------- 5638
MAC.US.x.85013.AY611490     -------------------C-----------------------------------------------------------------------..----------------------------------------------------------------------------- 5571
MAC.US.x.92050.AY603959     -------------------------------------------------------------------------------------------..----------------------------------------------------------------------------- 5631
MAC.US.x.92077.AY599201     -----------------------Y--------T----------------------------------------------------------..----------------------------------------------------------------------------- 5638
MAC.US.x.93057.AY611492     -------------------C------------Y----------------------------------------------------------..----------------------------------------------------------------------------- 5636
MAC.US.x.93062.AY607704     --------------------G-----------------------------------------------------R----------------..----------------------------------------------------------------------------- 5623
MAC.US.x.95058.AY611494     -----------------Y-C---------R-------------------------------------------------------------..----------------------------------------------------------------------------- 5625
MAC.US.x.95086.AY607703     -------------------C----------------------C------------------------------------------------..----------------------------------------------------------------------------- 5494
MAC.US.x.95112.AY588946     -------------------------------------------------------------------------------------------..----------------------------------------------------------------------------- 5627
MAC.US.x.96016.AY607701     -------------------Y------------Y----------------------------------------------------------..----------------------------------------------------------------------------- 5638
MAC.US.x.96020.AY611488     -------------------C-----------------------------------------------------------------------..----------------------------------------------------------------------------- 5630
MAC.US.x.96072.AY611491     ------------------T------------------------------------------------------------------------..----------------------------------------------------------------------------- 5642
MAC.US.x.96081.AY597209     -------------------------------------------------------------------------------------------..----------------------------------------------------------------------------- 5640
MAC.US.x.96093.AY611489     ---------------------------------------T----------------------------------------M----------..----------------------------------------------------------------------------- 5638
MAC.US.x.96114.AY588945     --------------------------------T----------------------------------------------------------..----------------------------------------------------------------------------- 5633
MAC.US.x.96123.AY611487     -------------------C------------------R--K-------------------------------------------------..----------------------------------------------------------------------------- 5527
MAC.US.x.96135.AY607702     -------------------C-----------------------------------------------------------------------..----------------------------------------------------------------------------- 5623
MAC.US.x.97009.AY599199     -------------------------------------------------------------------------------------------..----------------------------------------------------------------------------- 5506
MAC.US.x.97074.AY599198     -------------------C-----------------------------------------------------------------------..----------------------------------------------------------------------------- 5506
MAC.US.x.r80025.AY576480    -------------------C----------------------C------------------------------------------------..----------------------------------------------------------------------------- 5494
MAC.US.x.r90131.AY576481    -------------------------------------------------------------------------------------------..----------------------------------------------------------------------------- 5637
```

```
                                                                                         Env signal peptide start
                                                                      Tat Rev exon 1 end  Tat Rev intron start                                    Env signal peptide end Env gp120 start
MAC.US.x.239               GCAATCACGAAAGAGAAGAAGAACTCCGAAAAAGGCTAAGGCTAATACATCTTCTGCATCAAACAA.GTAAGTATGGGATGTCTTGGGAATCAGCTGCTTATCGCCATCTTGCTTTTAAGTGTCTATGGGATCTATTGTACT......CTATATGTCACAGTCTTTTATG 6694
Env                                                                                                    _M__G__C__L__G__N__Q__L__L__I__A__I__L__L__L__S__V__Y__G__I__Y__C__T__.__.__L__Y__V__T__V__F__Y__
Tat exon 1                 __Q__S__R__K__R__R__R__T__P__K__K__A__K__A__N__T__S__S__A__S__N__K
Rev exon 1                 S__N__H__E__R__E__E__E__L__R__K__R__L__R__L__I__H__L__L__H__Q__T__
H2A.CI.88.UC2.U38293       --G-AA-G-CGGAC------------A--G-GAA-----A--C--C-GC-----A--C--G----.--G-------AGCC-GG-A-------------GCT-T---T--AT-AAC-----CT-GCTTA--A-----C-AA......-A------G--T-----C---- 6733
H2A.DE.x.BEN.M30502        --G-AAGG-C-GAC-------G-----A-GG--AA-----A--CC-T-GC--------C--G----.--G-------AGCC-GG-A----------T--G-T-T---T--A--AAC-----CT-GCTTAG-A-------GC......-AG-----G--T--T--C---- 6794
H2A.DE.x.PEI2.U22047       A-GGAAGG-C-GAC-------G-----A--G--A--A------C--T-G-----------G----.--G-----AATGGA-AG-A-A--------AA--G-T-----T--A--AACT----CT-GCTTA--A-----CG-C......-A------G--T--T--C---- 6756
H2A.GH.x.GH1.M30895        A-G-AA-A-C-GAC-G-----------A-G----A-----A--C--T-G-----------G----.--G----A-ATG---GG-AA--G--TA--ATG-G-T----G------AGCT----CT--CTTAG-A-----C--C......-A------G--T--T--C---- 6235
H2A.IN.NNVA.NEW            C-G-AAGG-C-G-C-------G--C-GA--G--AA--------C--C-G-----A--C--G----.--G------AC-C--GAAA--GCC---------G-T-T---T-----AACT---ACT-GCTTA--A-----C-AACAA...-A------A--T--T--C---- 6789
H2A.IN.95.CRIK             A-G-AAGG-C-GAC-------G-----A--G--AA-----A--C--C-GA--CT-A-----G----.--G-------CGCA-GAA----C------A-----T---T-T--AT-AACT----CT-GCTTA--A-----C-AA......-A------A--T--T--C---- 6536
H2A.GM.x.MCR35.AY509260    C-GGAAGG-C-GAC----------------G--AA--------C--T--------------G----.--G------ATGG--GG-A-A-----------G-T-----T-----AGCT---ACT-GCTT---A-----C--C.....AAT-----G--T--T--C---- 6236
H2A.GM.87.D194.J04542      --G-CAGG-C-GAC-------G-----A-G---AA-----A--C--C-GC--C--------G-T--.-----------AGCC-GG-A-------------G-T-----T--AT-AACT----CT-GCTTA--A-----C-AA......-A------G--T--T--C---- 6235
H2A.GM.x.ISY.J04498        C-G-AAGG-C-GAC-------G-G------G--AA--------C--T-G---C--------G----.--G----CAATGA--GG-AAA-T----------G-T---T-TC----AACT----CT-GCTTA--A-----C--C.....AA------G--T--T--C---- 6232
H2A.GM.x.MCN13.AY509259    C-GGAAGG-C-GAC----------------G--AA--------C--T--------------G----.--G------ATGG--GG-A-A-----------G-T-----T-----AACT---ACT-GCTT---A-----C--C.....AAT-----G--T--T--C---- 6236
H2A.GW.86.FG.J03654        C-G-AAGG-C-GAC-------G-----A--G--AA--------C--C-G------------G----.--G------AAGG--AG-AA------A-----G--T--T--TG-A--AGCT----CT--CCTA--AC----C-AG......-A--T---G--T--T--C--C- 6229
H2A.GW.87.CAM2CG.D00835    C-G-AAGG-C-GAC-------G-----A--G--A---------C--T-G------------G----.--G-------AGC--GGCA-A-------------T-----TC-A--AGCT----CT-GCTTA--A-----C-GA...CAA-A------G--T--T--C---- 6803
H2A.GW.x.ALI.AF082339      C-G-AAGG-C-G-C--------T----A--G--AAT-------C--T-G------------G----.--G------ATG-C-AG-A-A-----------G-TA-T-----A--AGCT----CT-GCTTAG-A-------AA......-A------G--T--T------- 6784
H2A.GW.x.MDS.Z48731        A-G-AAGG-C-GAC-------G-----A--G--A------A--C--C-G------------G----.--G------ACGC-CAAAAT-C-------------T-----T--A--AACT----CT-GCTTA--A-----C-AACCACAA-A------A--T--T--C---- 6241
H2A.SN.85.ROD.M15390       A-G-AAGG-C-GAC-------G-----A--G--AA-----A--C--C-G---C--A--C--G----.--G------.........AT---------------T-----T--AT-AGCT----CT-GCTTAG-A-----C--C......-A------A--T--T--C---- 6228

H2AB.CI.90.7312A.L36874    C-GC--GA----A---......T--T-A----G-AA----A--GC----------------G----.--G----A-ATG---GG-AA-----TA--AT--G-T----G------AGCT----CT--CTTA--A-----C--C.....AA------G--T--T--C---- 6770

H2B.CI.88.UC1.L07625       C--C--GA----A---......T--T-A----G--------T--C-G-------------G--G-.--G-------C-CACAC-A-C-----C---T----TTTGC--C-A---A--------------T-TCTGG--CA-AAGAAAAAC--------T-----C---- 6775
H2B.CI.x.EHO.U27200        A-GC--GA--------......T--T-A----G-------A---C-------------C-----G-.--G-------C-CA-G--AAT---T-C--A---G-TA-AC--C-----A-----A--------TATATGG-C-AG.....AAC-T------T-----C---- 6759
H2B.GH.86.D205.X61240      C-GC--GA---G----......T--G-A----GA------A---C-G--C--------C--G----.--G-------C--A-T--A-C-GC-GC----C---T--GC--C-----A--G--A--AG----T-TGTA----AA......-A------T--T-----C---- 6761
H2B.JP.01.KR020.AB100245   C--C--GA----A---......T--T-A----G--------T--C-G-GC--A-----------G-.--G-------C-CA-A-CAAT--GT-C---T--GCT--AC--C-----A-----A-------CT-TATGG-CGAG.....GAT-TC-----T-----C---- 5911

H2AB.CM.03CM_510_03.NEW    C--C--GA--------..............----GA------A---C-G-----C-----C--G----.--G-------C--A-TC-A-C-GCTGC----C----A-GC--C-----C-----A---G----T-TACA--Y-AA......-A----A-Y--T-----C---- 5894

H2G.CI.92.ABT96.AF208027   ----C-TA---GA------..-------------A-----------T---T-------------G-.----------C--A------C-----------------A-C---R--C------ACT------T-----------C......AA---------T--T------- 6107

H2U.FR.96.12034.AY530889   ----C-CA---GA--G......-G---A--G----T-------T--T-----G--C----------.----------C---------C-----------------TC-A---TG-----A--CA-GC-T---A--C------......-AG--------T----------- 6248

SMM.SL.92.SL92B.AF334679   AA--C-CA---GA---......GT-AA---G---------------C---T-GA-A--------C-.----------CG----C---ACT---C--------A-A-------T----------G-TA----C--GG---G-A......-AG-----A---A--------- 6121

MAC.US.x.MAC239            ------------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 5771
MAC.US.x.251               ---G--------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 6693
MAC.US.x.251_32H           ---G--------------------------------------------------------------.-----------------------------------------------------------C---------------......-A-------------------- 6695
MAC.US.x.251               ---G--------------------------------------------------------------.---------------------------------------------------------------------------......-A-------------------- 6670
MAC.US.x.17EC1.AY033233    A-----------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 6694
MAC.US.x.17EFR.AY033146    A-----------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 6694
MAC.US.x.1937.AY611495     ----------RRR----------------G------------------------------------.--------------------------------------------R--------------Y-----------Y---......---------------------- 5800
MAC.US.x.2065.AY611493     ------------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 5800
MAC.US.x.MM142.Y00277      -A-G----AC-G-----------------G----------------------------------G-.-----------T--------A---------------------------C----------------T-------T-......-A-------------------- 6180
MAC.US.x.SMM142B.BD131285  -A-G----AC-G-----------------G----------------------------------G-.-----------T--------A----------------------GCTC-AAGT--C-A-GG...--TAT-----T-......-A-------------------- 6177

MNE.US.82.MNE              ---G-----C-GA----------------G------------------------------------.---------------------------------------------T--C------C--------------C-T-......-A-------------------- 6162
MNE.US.x.MNE027.U79412     ---G-----C-GA----------------G----T-------------------------------.---------------------------------------------T--C------C--------------C-T-......-A-------------------- 6162

SMM.US.x.F236              ---GCAC----GA---......--------G--A-----------C---T-C--------------.--------------------------------------GC------AG-------T-TA-A---T-G----GT-......-A------A-----A--C---- 6641
SMM.US.x.H9.M80194         ---GCA-----GA---......--------G--A-----------T----------G----------.--------------------------------------GC---C---A-Y-----CT-CA-----T------GT-......-A------A-----A--C---- 6162
SMM.US.x.PBJ14             ---GCA---G-GA---......--------G--A----------------------G----------.--------------------------------------GC-------A-------CT-CA-----T------GT-......-A------A-----A--C---- 6456
SMM.US.x.PBJA.M31325       ---GCA-----GA---......--------G--A----------------------G----------.--------------------------------------GC-------A-------CT-CA-----T------GT-......-A------A-----A--C---- 6455
SMM.US.x.PBJ               ---GCA-----GA---......--------G--A----------------------G----------.--------------------------------------GC-------A-Y-----CT-CA-----T------GT-......-A------A-----A--C---- 6162
SMM.US.x.PBJ               ---GCA-----GA---......--------G--A----------------------G----------.--------------------------------------GC-------A-------CT-CA-----T------GT-......-A------A-----A--C---- 6648
SMM.US.x.PGM53.AF077017    ---GCAG--C-G-----------------G-----------------T-------C-----G-.----------------C----------------------------------------------CT---------------C---......-A------------T------- 6613
SMM.US.x.SME543.U72748     ---G...-ACCGA-------..-----------A-----A-----C--CT-C--------------.--------------------------------------GC-------AG-------T-TA-A---T-G----GT-......-A------A-----A--C---- 6692

STM.US.x.STM.M83293        --G----A---GA---......GT-AA------------A--T--C--AT----------------.-----------CC--C-C---A-----A-----------T--------A-------CT-GCTT--C------C---......-AG--------G--A------- 6336

MAC.US.x.80035.AY611486    ------------------------------------------------------------------.--------------------------R------------------------------------------------......---------------------- 5800
MAC.US.x.81035.AY599200    ------------------------------------------------------------------.---------------------------------------------------------------------------......-A-------------------- 5801
MAC.US.x.85013.AY611490    ---R--------------------------------------------------------------.-------------------------------------------------------------------------Y-......-W-------------------- 5734
MAC.US.x.92050.AY603959    ------------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 5794
MAC.US.x.92077.AY599201    ------------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 5801
MAC.US.x.93057.AY611492    ------------------------------------------------------------------.---------------------------------------------------------------------------......-A-------------------- 5799
MAC.US.x.93062.AY607704    ------------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 5786
MAC.US.x.95058.AY611494    ------------------------------------------------------------------.--------------------A-------------------------------------------Y----------......---------------------- 5788
MAC.US.x.95086.AY607703    ------------------------------------------------------------------.---------------------------------------------------------------------------......-A-------------------- 5657
MAC.US.x.95112.AY588946    ------------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 5790
MAC.US.x.96016.AY607701    ------------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 5801
MAC.US.x.96020.AY611488    ------------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 5793
MAC.US.x.96072.AY611491    ------R-----------------------------------------------------------.----------------C----------------------------Y-----------------------------......-A-------------------- 5805
MAC.US.x.96081.AY597209    ------------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 5803
MAC.US.x.96093.AY611489    ------------------------------------------------------------------.---------------------------------------------------------------------------......--------------------Y- 5801
MAC.US.x.96114.AY588945    ------------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 5796
MAC.US.x.96123.AY611487    ------------------------------------------------------------------.---------------------------------------------------------------------------......-W-----------Y------- 5690
MAC.US.x.96135.AY607702    ------------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 5786
MAC.US.x.97009.AY599199    ------------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 5669
MAC.US.x.97074.AY599198    ------------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 5669
MAC.US.x.r80025.AY576480   ------------------------------------------------------------------.---------------------------------------------------------------------------......-A-------------------- 5657
MAC.US.x.r90131.AY576481   ------------------------------------------------------------------.---------------------------------------------------------------------------......---------------------- 5800
```

```
                                                     V1 loop start                    V1 loop end
MAC.US.x.239                  GTGTACCAGCTTGGAGGAATGCGACAATTCCCCTCTTTTGTGCAACCAAGAATAGGGATACTTGGGGAACAACTCAGTGCCTACCAGATAATGGTGATTATTCAGAAGTGGCCCTTAATGTTACAGAAAGCTTTGATGCCTGGAATAATACAGTCACAGAACAGGCAATA 6864
Env                           G__V__P__A__W__R__N__A__T__I__P__L__F__C__A__T__K__N__R__D__T__W__G__T__T__Q__C__L__P__D__N__G__D__Y__S__E__V__A__L__N__V__T__E__S__F__D__A__W__N__N__T__V__T__E__Q__A__I_
H2A.CI.88.UC2.U38293          -C-----C-TG-----A-----AT-T--C-----A--------------A-----A-----------G--C-TA------T-G-----C----A-------CAG---A-AC-TT-A-----G------GCT--------A---G----------A--------A------ 6903
H2A.DE.x.BEN.M30502           -CA----C--G----AA-----AT-T------T-A-----------T--A-----A--C--------G--C-TA------T-G-----C----A-------CAG---A-AATTT-A-----G-----GGCT--------A--------------G--------A---G-- 6964
H2A.DE.x.PEI2.U22047          -CA----C--G----A------AT-C----------------------GA-----A--------------C-TA------T-G-----C----A-------CAG---A-AC-TT-A-----G-----GGCT-----C--A-----C--------A--------A---G-- 6926
H2A.GH.x.GH1.M30895           -C--G--C-TG-----A-----AT-C--------------------T--A--C--A--------------C-TA------AAG-----C----A-------CA----A-AA--T-G-----G-----GGCT--C--C--A---G----------A--------A---G-- 6405
H2A.IN.NNVA.NEW               -CA----C--G----A------AT-C--C--T-----------------A-----A--------------C--A--A-----G----------A-------CAG--GA-A--TT-A-----G-----GGCT--C-----A---G-C--------G--------A---G-- 6959
H2A.IN.95.CRIK                -CA----C--G----A------AT-C----------------------GA--------------------C-TA------T-------C----A---C---CAG---A-A-TTT-G-----GG----GGCT--C-----A---G-C--------A--------A------ 6706
H2A.GM.x.MCR35.AY509260       -CA----C--G-----A-----AT-C--------------------------------------------C-TA------T-G-----C----A---A---CAG--GA-AA-TT-G-----G-----GGCT--C-----A---G----------A--------A------ 6406
H2A.GM.87.D194.J04542         -CA----C--G-----A-----AT-T--------A--------------A-----A-----------G--C-TC------T-G-----C----A-------CAG---A-AA--T-A-----G------GCT--------A---G----------A--------A------ 6405
H2A.GM.x.ISY.J04498           -A-----C-TG----AA-----AT-C--------------------T--A-----A--------------C-TA------T-G-----C----A-------CA---GA-AC-TT-G-----A-----GGCT-----C--A---G------T---A--------A---G-- 6402
H2A.GM.x.MCN13.AY509259       -CA----C--G-----A-----AT-C--------------------------------------------C-TA------T-G-----C----A-------CAG--GA-AA-TT-G-----G-----GGCT--C-----A---G----------A--------A------ 6406
H2A.GW.86.FG.J03654           -CA----C--G-----------AT-C--------G--------------A-----A----------------TA------T-------C----A-------CAG---A-AA-T--A-----G-----GGCT--C-----A--------------A--------A---G-- 6399
H2A.GW.87.CAM2CG.D00835       -C-----C--G----AA-----AT-C--------------------T--A--------------------T-TA------T-------C----A-------CAG--GA-AC-TT-A-----G-----GGCT--C-----A---G--------A-A--------A------ 6973
H2A.GW.x.ALI.AF082339         -C--G-----A----AA-----AT-C-----------------------A-----A--------------C-TA------T-------C----A-------CAG---A-A--TT-G-----G-----GGCT--C-----A---G----------A--------A---G-- 6954
H2A.GW.x.MDS.Z48731           -CA----C--G-----------AT-C-----------C-----------A--------------------C-TA------T-G-----C----A-------GA----A-A--TT-G-----G-----GGCT--------A---G----------A--------A---G-- 6411
H2A.SN.85.ROD.M15390          -C-----CA-G----AA-----A--C----------------------GA--------------------C-TA------T-G--T--C----A-------CAG---A-AA-TT-G-----A-----GGCT--------A--------------A--------A------ 6398
H2AB.CI.90.7312A.L36874       -C--G--C-TG-----A-----AT-C--------------------T-----C--A--------------C-TA------T-G-----C----A-------CA----A-A-----G-----G-----GGCT--C-----A--------------A------------G-- 6940
H2B.CI.88.UC1.L07625          -CA----T--A--------C--A--GG----T-----C----------CA--C--------C--------TGTA--------C-----C-----G-----CA-C--GA-CAGTG-A---A-A-----GGCT--------A--------------A------------G-- 6945
H2B.CI.x.EHO.U27200           --A----C--A----AA-----AT-------------------T----GA--C--A-----C--------TGTA--A-----C----------A---C---A-C---A-CCAAT-A---A-A-----GGCT--------A---G----------G-----T-------C- 6929
H2B.GH.86.D205.X61240         -CA----C--A--------C--A---G--------A------------CA--C--A--C--C--------TGTA-----T--C-----C--------C--CA-T--GA-CAGG--A--CA-A-----GGCT--------A---G----------G---C--------G-- 6931
H2B.JP.01.KR020.AB100245      -CA-------A----A------A---G----------------C-----A-----A---T----------TGTG--A-----C-----C----A---C---GAC---A-C--G--G---A-A-----GGCT--------A---G----C-----G------------G-- 6081
H2AB.CM.03CM_510_03.NEW       -CA----C--G----RA-----AT-Y--Y-----G--Y--------T--A-----A-----------G--C-TA------T-G-----C----A-------CAG---A-AA-GT-A-----G------GCT--------A---R----------A--------A------ 6064
H2G.CI.92.ABT96.AF208027      -CA----C--A--------C----GCG-A--T-----C-----------A--------C--A-----C--T--A--------T--T-------A------CAGT---C-A--AT-G-----A------GCT------------G----C-----A------------G-- 6277
H2U.FR.96.12034.AY530889      -CA----T--A--------C-----------------------------C-----A-----A--------C--A--------T--T--------------------GT-A--AT-A-----C-----GGC------------------C-----A-----G--A---G-G 6418
SMM.SL.92.SL92B.AF334679      --A-T--T--A--------C--T--G--C--------C-----G---C-------A--C--A-----G--CGT-------T-G-----------A--C---------T-------------------GGC---C-----T---G----------A--T-----A------ 6291
MAC.US.x.MAC239               -----------------------------------------------------------------------------------------------------------A-------------------------------------------------------------- 5941
MAC.US.x.251                  -------------------------G-----------C---------------------------------------------------------------------T-------------------------------T---G-G------------------------ 6863
MAC.US.x.251_32H              -------------------------------------C---------G-----------------------------------------------------------T----------C------------------------G-G------------------------ 6865
MAC.US.x.251                  -------------------------------------C---------------------------------------------------------------------T-------------------------------T---G-G------------------------ 6840
MAC.US.x.17EC1.AY033233       ---------T------------------------T------------------------------------------------------------------------A-------------------------------------------------------------- 6864
MAC.US.x.17EFR.AY033146       ---------T------------------------T------------------------------------------------------------------------A-------------------------------------------------------------- 6864
MAC.US.x.1937.AY611495        -----------------------------------------------------------------------------------------------------------A-------------------------------------------------------------- 5970
MAC.US.x.2065.AY611493        ---------------------------------------------------------------------------------------------R-------------A-----------------------------------R-----------------------R-- 5970
MAC.US.x.MM142.Y00277         -------------------------------------C-------------------------------------------------------A-------------T-------------------------------T---G-G------------------------ 6350
MAC.US.x.SMM142B.BD131285     -------------------------------------C-------------------------------------------------------A-------------T-------------------------------T---G-G------------------------ 6347
MNE.US.82.MNE                 -------------------------------------C----T-----G--------------------------------------------A-------------T-----------A-------------------T---G-G------------------------ 6332
MNE.US.x.MNE027.U79412        -------------------------------------C----T-----G--------------------------------------------A-------------T-----------A-------------------T---G-G------------------------ 6332
SMM.US.x.F236                 ----------A----A---------------------C--------------------C--------------A--A---T-G----------A------C------T----AA-C-----C-----GGCT--------T---G-------------------A------ 6811
SMM.US.x.H9.M80194            ----------A---------------G----------C----------GA--------------------C--A--A---T-G-----------------C------T----AA-------C-----GGCT--------T---G-------------------A------ 6332
SMM.US.x.PBJ14                --A-------A---------------G----------C--------------------C--------------A--A---T-G-----------------C------T----AA-------C-----GGCT--------T---G-------------------A------ 6626
SMM.US.x.PBJA.M31325          --A-------A---------------G----------C--------------------C--------------A--A---T-G-----------------C------T----AA-------C-----GGCT--------T---G-------------------A------ 6625
SMM.US.x.PBJ                  --A-------A---------------G----------C----------GA--------------------C--A--A---T-G-----------------C------T----AA-------C-----GGCT--------T---G-------------------A------ 6332
SMM.US.x.PBJ                  --A-------A---------------G----------C--------------------C--------------A--A---T-G-----------------C------T----AA-------C-----GGCT--------T---G-------------------A------ 6818
SMM.US.x.PGM53.AF077017       ----------A----A---------------------C--------------------C-----------------A---T-G----------A-------------T-----A-A--------------A--------------C------------------------ 6783
SMM.US.x.SME543.U72748        ----------A----A---------------------C----------G---------C--------------A--A---T-G--T-------A------C------T----AG-C---A-C-----GGCT--------T-----------------------A------ 6862
STM.US.x.STM.M83293           ----------A--------------G-----------C--------------------------------------A--T--------C-----------C------T----AA-A-----C-----GGCT------------G-C-----------------A------ 6506
MAC.US.x.80035.AY611486       -----------------------------------------------------------------------------------------------------------A-------------------------------------------------------------- 5970
MAC.US.x.81035.AY599200       -----------------------------------------------------------------------------------------------------------A-------------------------------------------------------------- 5971
MAC.US.x.85013.AY611490       -----------------------------------------------------------------------------------------------------------W-------------------------------------------------------------- 5904
MAC.US.x.92050.AY603959       -----------------------------------------------------------------------------------------------------------A-------------------------------------------------------------- 5964
MAC.US.x.92077.AY599201       -----------------------------------------------------------------------------------------------------------A-------------------------------------------------------------- 5971
MAC.US.x.93057.AY611492       -----------------------------------------------------------------------------------------------------------A-------------------------------------------------------------- 5969
MAC.US.x.93062.AY607704       -------------------------G---------------------------------------------------------------------------------A-------------------------------------------------------------- 5956
MAC.US.x.95058.AY611494       -----------------------------------------------------------------------------------------------------------A-------------------------------------------------------------- 5958
MAC.US.x.95086.AY607703       -----------------------------------------------------------------------------------------------------------T-------------------------------------------------------------- 5827
MAC.US.x.95112.AY588946       -----------------------------------------------------------------------------------------------------------A-------------------------------------------------------------- 5960
MAC.US.x.96016.AY607701       -----------------------------------------------------------------------------------------------------------A-------------------------------------------------------------- 5971
MAC.US.x.96020.AY611488       -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 5963
MAC.US.x.96072.AY611491       -----------------------------------------------------------------------------------------------------------C-------------------------------------------------------------- 5975
MAC.US.x.96081.AY597209       -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 5973
MAC.US.x.96093.AY611489       -------------------------G---------------------------------------------------------------------------------T-------------------------------------------------------------- 5971
MAC.US.x.96114.AY588945       -------------------------G---------------------------------------------------------------------------------A-------------------------------------------------------------- 5966
MAC.US.x.96123.AY611487       -----------------------------------------------------------------------------------------------------------C-------------------------------------------------------------- 5860
MAC.US.x.96135.AY607702       -----------------------------------------------------------------------------------------------------------A-------------------------------------------------------------- 5956
MAC.US.x.97009.AY599199       -----------------------------------------------------------------------------------------------------------A-------------------------------------------------------------- 5839
MAC.US.x.97074.AY599198       ----------------------------------------------------------------------------A------------------------------A-------------------------------------------------------------- 5839
MAC.US.x.r80025.AY576480      -----------------------------------------------------------------------------------------------------------T-------------------------------------------------------------- 5827
MAC.US.x.r90131.AY576481      -------------------------------------------------T---------------------------------------------------------A-------------------------------------------------------------- 5970
```

```
                                                                              V2 loop start
MAC.US.x.239                GAGGATGTATGGCAACTCTTTGAGACCTCAATAAAGCCTTGTGTAAAATTATCCCCATTATGCATTACTATGAGATGCAATAAAAGTGAGACAGATAGATGGGGATTGACAAAATCAATAACAACAACAGCATCAACAACATCAACGACAGCATCAGCA............ 7023
Env                         _E__D__V__W__Q__L__F__E__T__S__I__K__P__C__V__K__L__S__P__L__C__I__T__M__R__C__N__K__S__E__T__D__R__W__G__L__T__K__S__I__T__T__T__A__S__T__T__S__T__T__A__S__A__.__.__.__.
H2A.CI.88.UC2.U38293        --A-----C---AGG--A--------A--G-----A--A-----C--G--GA-A--TC----TG-GG-A----AT--T---CC-GT-ACAGGGA---ACACCAATGCC---GCCAA-CC---CG--G-CAG-C----C--CAACC-T........................ 7050
H2A.DE.x.BEN.M30502         --A-----C-----T--A--------A--------A--A-----C--GC--A-A--TC----TG-GG-A----AT--T-GC-GGGT-C-AGGGA---CCAC-ACCCC--ATCCCAGG-CCT-G-GTT-CA--A-CT-G-G-C--C-C---..................... 7114
H2A.DE.x.PEI2.U22047        --A-----C---A-T--A--------A---G----A--A-----C------A-A--C-----TG-GCAA---GA---T--C-GC-CAAGT-----G--CA-TAACAGC----GTGAGGGG-GC---GTCC--.................................... 7058
H2A.GH.x.GH1.M30895         --A-----C---AGT--A--------A--------A--A--C--C---C--A-A--TC----TG-AG-A-----T--T--C-GC-CCACT-ACA-C-C-ACCAC-ACAGGG-GCA---C-GGG-TG.......................................... 6531
H2A.IN.NNVA.NEW             --A-----C---A-T--A--C-----A----C---A--A-----C------A-A--C-----TG-AG-A---GAG-----G--C--A-------GG-ACGCAACCACA--CGC-C---C---C---C---AGATT-ATGAGA--......................... 7103
H2A.IN.95.CRIK              --A-----C---A-T--A--------A----C---A--A-----C------A-A--C-----TG-A--A-----------C-C----ACA--CACA-CCACAACCGCAC-C-C-AGC-C--GTG--GGGAGCA--G-C---C-C........................ 6850
H2A.GM.x.MCR35.AY509260     --A-----C---A-T--A--------A--------A--A-----C------A-G--T-----TG-AG-A--------T--C--C-CA--TG--AGG-ACACAACCACAC-C-C-A--GC-T-CC-GCGTA--...................................... 6538
H2A.GM.87.D194.J04542       --A-----C---AG---G--------A--------A--A-----C--G--GA-G--CC----TG-GG-G----AT--T----T--C-TCAGGGAC--CCGC-ACCCC--GTCC-C--.............................................. 6522
H2A.GM.x.ISY.J04498         --A-----C---A-T--A--------A--------A--A-----C------A-G--T-----TG-A--A----AC--T--CGC----ACAGAGAGCGC-GTT-C-ACT----GCC--TCTGG-C-T.......................................... 6528
H2A.GM.x.MCN13.AY509259     --A-----C---A-T--A--------A--------A--A-----C------A-G--T-----TG-AG-A--------T--C--C-CA--TG--AGG-ACACAACCACAC-C-C-A--GC-T-CC-GCGTA--...................................... 6538
H2A.GW.86.FG.J03654         --------C---A-T--A--------A--------A--A-----C------A-A--C-----TG-AG-A----AC--T-CA-GG-ACAT---CACAT-GACA--GAG----G-CA-CCA-........................................... 6519
H2A.GW.87.CAM2CG.D00835     --------C---A-T--A--C-----A--------A--A-----C------A-A--T-----TG-AG-A----A---T--C-T----ACA-GT--C-C-ACTATGA-A-GG-C-A-C-C-C-C-GC-----T.................................. 7105
H2A.GW.x.ALI.AF082339       --A-----C---AG---A--------A--------A--A-----C--G---A-A--T-----T--AG-A----AG--T-GC--C-TAAGC-----G--CACAACCACAT-CCCGAGCCC-GGG-GC.......................................... 7080
H2A.GW.x.MDS.Z48731         --A-----C---AGT--A--C-----A--------A--A-----C--G---A-A--C-----TG-AG-A----AC--T-GC-G-CTCAGTGACAGCGC-A-CAC-AG--AC-C-A-C-C--ACG-C--C---G-C-----CA---TC-G---CA--A-CATAGCACCA.. 6579
H2A.SN.85.ROD.M15390        --A-----C-----T--A--C-----A--------A--A-----C---C--A-A--T-----TG-AG-A----A-----GC-GC-CA----GCAGC-C-G--AACAAC----CC----AG-GC----GCA--A-C-----CA--C-C----AC................ 6551
H2AB.CI.90.7312A.L36874     --A-----C---AGT--A--------A--------A--A--C--C---C--A-A--TC----TG-AG-A-----T--T--C-GC-CCACCG-CACC-CTACACC-CCA-GC-CCA-T-AC-AC-----CA--A--GAGC-CA----A...................... 7087
H2B.CI.88.UC1.L07625        --T-----G---AGT--T-----A-----C-----A--A-----T--GC-GA-G--GC-G---G-GG-A--------T-----C-CA-GA--CA---CTACTACTAAAC-T-TTA-T-C-C-T-TT--TA-TA-T-A-C--............................ 7086
H2B.CI.x.EHO.U27200         A-------G---AGT--------------------A--A-----T------A-----C-G--TG-G--A----AG--T-----G...AC-TGGAGCTC-GCTA-CAAAGAG-CCA-T-CGT-CT-CG-GT-CCTC-G-T-T--T--T....................... 7073
H2B.GH.86.D205.X61240       --T-----G---AG---------A-----C-----A--A-----C---C--A-----C-G--TG-GG-A----AC--T-G-----CC--A---A-CCC-G--AATGCC-GT-GTA-T-CC--C--T-AGC-TA-T--C--C--TCGT...................... 7078
H2B.JP.01.KR020.AB100245    -C------G---A-T--A--------T--------A--A-----T------A-----C-----TG-A--A----AT--T-G---CCT-A-C--CACAGCCAACACCACC---GCCAAC-AC-TC---G-C....................................... 6210
H2AB.CM.03CM_510_03.NEW     --------C---AG---A--------A--------A--A-----C--G--GA-A--CC-R--YG-GG-A----AT--T-CA-CC-CG--AG-CACACCCCA-AATACC----C-CGG-AT.............................................. 6184
H2G.CI.92.ABT96.AF208027    --A-----T---A-T--G--------A--T-C---A--------T------A-A--TC-G------G-C---------------C-------A-C-A------TC-A---GGGA-C-CC--TGT---CC-TG-----GTTC-C--A--TAA-ATG.............. 6433
H2U.FR.96.12034.AY530889    --A-----C---AGT--T--------A--T-----A--A---------G--GA-A--CC-T--T--AG-A--------T-G-----C---A---A-------------A---GG-A--GC-G--------A-TAGCC--G--ATGCAAGA-T--AGTAA-........... 6577
SMM.SL.92.SL92B.AF334679    --A-----T---A-T--------A--A--T--T--A-----------G---A-T-------T---G------A---T-------A---------C-------G-------G-G--GCT--T--T--TAGC---C----TA-T--T-GCC-C-T-A-TGCTGCTAGCCC 6461
MAC.US.x.MAC239             ----------------------------------G-----------------------------------------------------------------------------------------------------------------------......................... 6094
MAC.US.x.251                --A-------------------------------------------------A----------------------------------------------------------------TC---------CA--G--G--C--AAGG-A-TGT--GA-............... 7019
MAC.US.x.251_32H            -------------------------------------------------------------------------------------------------A-------------------T--------------C-----G--C-----G--------............... 7021
MAC.US.x.251                -----C---------------------------------------------------------------------------------------------------------------TC--------T-A--A--G--G--C----AT-----C---T-TCAGAA..... 7005
MAC.US.x.17EC1.AY033233     ---------------------------------------------------------------------------------------------------------------------------------------------------------------AGAGTA..... 7029
MAC.US.x.17EFR.AY033146     ---------------------------------------------------------------------------------------------------------------------------------------------------------------AGAGTA..... 7029
MAC.US.x.1937.AY611495      ------------------------------------------------------------------------------------------------------------------------------------Y-----R---------------................. 6123
MAC.US.x.2065.AY611493      ------------------------------------------------------------------------------------------------------------------------------------Y---------------------................. 6123
MAC.US.x.MM142.Y00277       -----C-------------------------------------------------------------------------------------------A-------------------TC----------------------A----A---A--G--AA-............ 6509
MAC.US.x.SMM142B.BD131285   -----C-------------------------------------------------------------------------------------------A-------------------TC----------------------A----A---A--G--AA-............ 6506
MNE.US.82.MNE               --------------T-------------------------------------A--------------------A------C----------------A-------------------TC-------------C-----G--AT-C-A---..................... 6482
MNE.US.x.MNE027.U79412      --------------T-------------------------------------A--------------------A------C--------A-------A-------------------TC-------------G--C-----A---AA---A--A--A--............ 6491
SMM.US.x.F236               --------G---A-C--------A--A--C--T-----C---------C-CA-----C----T--AG-A--------T-------C-------------G-----T------GG-AACGC-GGG------A--A--G---T-A----A------...............AC 6966
SMM.US.x.H9.M80194          ------------A-C--------A--A--C--T-----C---------C-CA-----C----T--A--A-----G--T------------M---MC--------T--A---GG-GS-CC-G--C-----A--A--CA---------A---..............CCACC 6487
SMM.US.x.PBJ14              ------------A-C--------A--A--C--T-----C---------C-CA-----C----T--A--A-----G--T------------------C--------T--A---GG-A-CCC-G--C-----A--A--CA---A----A---CA-G--T--ACAACACCAAC 6796
SMM.US.x.PBJA.M31325        ------------A-C--------A--A--C--T-----C---------C-CA-----C----T--A--A-----G--T------------------C--------T--A---GG-A--CC-G--C-----A--A--CA---A----A---CA-G--T--ACAACACCAAC 6795
SMM.US.x.PBJ                ------------A-C--------A--A--C--T-----C---------C-CA-----C----T--A--A-----G--T------------M---MC--------T--A---GG-AS-CC-G--C-----A--A--CA---A----A---CA-G--T--ACAACACCAAC 6502
SMM.US.x.PBJ                ------------A-C--------A--A--C--T-----C---------C-CA-----C----T--A--A-----G--T------------------C--------T--A---GG-A--CC-G--C-----A--A--CA---A----A---CA-G--T--ACAACACCAAC 6988
SMM.US.x.PGM53.AF077017     ------------A-T--------------T-----A--C-------------A-----C-----------C--------------G-CA--A-----C---------G--A-----GAATGA--T-C-----A--A---------A----A---AA-......AGCCCAAAAGC 6947
SMM.US.x.SME543.U72748      --------G---A-C--------A--A--C--T-----C---------C-TA-----C----T--AG-A--------T-------C-------------G-----T------GG-AG-GC-GAG------A--A--G-G-A------A---T--.........ACAACAAC 7023
STM.US.x.STM.M83293         --------C---A-C--T-----A--A--C--T--A------------GA-T---C----T---G-A-----G--T-----G-A---A-----C-A-------------GGGAA--C-GT--------TGA--C-----G--G-AG-----G--.............. 6662
MAC.US.x.80035.AY611486     -----------------------------------------------R----------------------------------------------------------------------------------------------------------................. 6123
MAC.US.x.81035.AY599200     ------------------------------------------------------------------------------------------------------------------------------------------------------A---................. 6124
MAC.US.x.85013.AY611490     ----------------------------------------------------------------------------------------------------------------------------------------------------------................. 6057
MAC.US.x.92050.AY603959     ----------------------------------------------------------------------------------------------------------------------------------------------------------................. 6117
MAC.US.x.92077.AY599201     ----------------------------------------------------------------------------------------------------------------------------------------------------------................. 6124
MAC.US.x.93057.AY611492     -------------------------------------------------------------------------------------R-------------------------------------------------------C---Y--------................. 6122
MAC.US.x.93062.AY607704     ----------------------------------------------------------------------------------------------------------------------------------------------------------................. 6109
MAC.US.x.95058.AY611494     ----------------------------------------------------------------------------------------------------------------------------------------------------------................. 6111
MAC.US.x.95086.AY607703     ----------------------------------------------------------------------------------------------------------------------------------------------------------................. 5980
MAC.US.x.95112.AY588946     ----------------------------------------------------------------------------------------------------------------------------------------------------------................. 6113
MAC.US.x.96016.AY607701     ------------------------------------------------------------------------------------------------------------------------R--G------------------------------................. 6124
MAC.US.x.96020.AY611488     ----------------------------------------------------------------------------------------------------------------------------------------------------------................. 6116
MAC.US.x.96072.AY611491     -----------------------------------------------G------------------------------------------------------------------------------------Y-----------------A---................. 6128
MAC.US.x.96081.AY597209     ----------------------------------------------------------------------------------------------------------------------------------------------------------................. 6126
MAC.US.x.96093.AY611489     -----------------------------------------------------------------------------------------K------------------------------------------------------S---------................. 6124
MAC.US.x.96114.AY588945     ----------------------------------------------------------------------------------------------------------------------------------------------------------................. 6119
MAC.US.x.96123.AY611487     -------------------------------------------------------------------------------------R----------------------------------------------------------S-----R---................. 6013
MAC.US.x.96135.AY607702     -------------------------------------------------------------------------------------R----------------------------------------------------------R---------................. 6109
MAC.US.x.97009.AY599199     ----------------------------------------------------------------------------------------------------------------------------------------------------------................. 5992
MAC.US.x.97074.AY599198     ---------------------------------------------------------------------------------------------------------------R------------------------------------------................. 5992
MAC.US.x.r80025.AY576480    ----------------------------------------------------------------------------------------------------------------------------------------------------------................. 5980
MAC.US.x.r90131.AY576481    ----------------------------------------------------------------------------------------------------------------------------------------------------------................. 6123
```

```
                              V2 loop end
MAC.US.x.239                  ..........AAAGTAGACATGGTCAATGAGACTAGTTCTTGTATAGCCCAGGATAATTGCACA...GGCTTGGAACAAGAGCAAATGATAAGCTGTAAATTCAACATGACAGGGTTAAAAAGAGACAAGAAAAAAGAGTACAATGAAACTTGGTACTCTGCAGATTTGG 7180
Env                           __.__.__.__K__V__D__M__V__N__E__T__S__S__C__I__A__Q__D__N__C__T__.__G__L__E__Q__E__Q__M__I__S__C__K__F__N__M__T__G__L__K__R__D__K__K__K__E__Y__N__E__T__W__Y__S__A__D__L__
H2A.CI.88.UC2.U38293          .......TCAT-CC-TAC---AA-A-----A-G-TC-A-C---G---G-GCA--C---------...--A----GGG-C--AGGG---G-T-AT-----G-----T-----G--AC-GG--CA---T----T---G-G---T-C---C--A--------A-AT---G--- 7210
H2A.DE.x.BEN.M30502           .......TCCGC--CCTC---AA-A-----A---TC-AAC--C----AAA-CA-C-CA---G--...--A--A-GGT-T---G-G-----GCAA---G-G-----T----AG------G--CA---T-------GGAG---T--G--C--A-----T-TA-A----G--- 7274
H2A.DE.x.PEI2.U22047          ................--G--AT-A--C--A----C---A--C---A--A-CA-C-GC---T--...-AT--A-GGAGT--AG-GG-AG-CGAT---CGG-----T--------AC--C--CT---T---CC-C-GC-A--T-G------A--------AAAG---G-A- 7209
H2A.GH.x.GH1.M30895           ................TCAGA-A-A-----A...-C----CCATC-TA-TCA--C--C--------...--A----GGA-G--AG-G--AG-C-AT---C-G---T-T--------A---G-G-----T---------C----T--------A--------AAA----G-A- 6679
H2A.IN.NNVA.NEW               ................ACTGA-A-A-G-----AC-C---A--CG--CG-GGA--C--C---T--...--A--AAG-G-G---G-G--AG-C-TA---C-G-----T--------A----T------T---CC----AG---T--------A--------AAGC---G--- 7254
H2A.IN.95.CRIK                ............---CCCATA-----A-------AC-C---A--C--GCA-GCAA-C--C---T--...--A----G-G-G--AG-GG-AG-C-AT---G-G-----T--------A---GTG-----T--ACC---GACA--T--------A--------ACGT---G--- 7007
H2A.GM.x.MCR35.AY509260       ....ATAAAACCC--GACAGA-A-A-G-----A-TCCA-A--C---CG-GCAA-C--C---T--...--A----G-A----AG-GG--G-C-AT---C-------T--------A---G-G-----T--------GC----T-----GG-A--------GAAG---G-A- 6701
H2A.GM.87.D194.J04542         ...........--CA-TACA--AA-AG---A-A-TC-A-C-------G-G-CA-C--C--------...--A--A-GGA----AG-GG--G-TGAG---G-G-----T-----G---C--G--CA---T-------GGA----T-----CG-A--------AAG----G--- 6679
H2A.GM.x.ISY.J04498           ................--T---A-A-----T---GA-C-A--C---CAATT-A-C------T--...--AC--AGGG-G--AG-C---G-CGAG---C-G-----T--------AC--G-GTT---T---------C----T-G------C--------AAA----G--- 6679
H2A.GM.x.MCN13.AY509259       ....ATAAAACCC--GACAGA-A-A-G-----A-TCC--A--C---CG-GCAA-C--C---T--...--A----G-G----AG-GG--G-C-AT---C-------T--------A---G-G-----T--------GC-A--T-G---G--A--------GAAG---G-A- 6701
H2A.GW.86.FG.J03654           .............AATAT--CAA-T-TAA-TGAC-C---CCA-GCGCGTGCA--C--C--------...--A--AA-GG-G--AG-------CGA----C-G--T-GT--------A---G-G----------G----C----T-C----G-A--------AAA----G--- 6673
H2A.GW.87.CAM2CG.D00835       ............----AG-CACCAA-A-G---C-A-TC-C----C---CG-ACAA-C--C---T--...--A--A--GG---AA-G--AG-C-AG---C-C-----T--------A---G--------T---------C------------A--------AAGC---G--- 7262
H2A.GW.x.ALI.AF082339         .......ACACTCAA-CC-C--A-A-------GCGA-C-A--C---AAGGCA--C--C---C-CAGG--AC-A-GGG-T--AG-G---G-C-AT---CGG--------------A---C-G-----T---CC----C----T--------A--------AAA----G--- 7243
H2A.GW.x.MDS.Z48731           ..........-GGAACACA-CCA-A-G-----A-TC-C-A--C---CG-GCA--C--C---T--...--AC-A--GG-G--AG-G---G-C-AG---C-------T--------A---C-G-T---T---------C----T--C-----A--------AAGC---G--- 6736
H2A.SN.85.ROD.M15390          ..........C-G-AGC-AGA-A-A-G----GA--C-C-A--CGC-CG-GCA--C--C---T--...--A----G-G-G--AG---C---C-AT--CC-G-----T--------A---G--------T---------C----T--------A--------AAA----G--- 6708
H2AB.CI.90.7312A.L36874       ..........GG--GGCCAGA-A-A-----A---TT-C-A--C--GCG-ACA--C--C-------...--A----G-G-G--AG-G---G-CGAT---C-G--------------A---G-G-----T----C----C-A--T-G------A--------AAA----G-A- 7244
H2B.CI.88.UC1.L07625          ..........TC--A-A--T-AC----C--C--A---C-A-----TAAAA-T--C-CC---C-G...--AA-A-G--T---AA-C-CAG-TGA----T-C--T-----------AC---G------TG---------C-A--T--G--C--C------GAAAA------A- 7243
H2B.CI.x.EHO.U27200           .............CAGAC-C-AC----C--AGA---CAAA-----TCAAA-T--C-GC--TG--...--GA-A-G--T---AG--------GA----C-------A--------A-----------TG--TC----C-A--T--G--C--C-----TAAACA------A- 7227
H2B.GH.86.D205.X61240         ..........GGGC-GA-A-C-A-T--C--A--AGACC-A--C---AAAA-T--C-GC-------...--AC-A-G-G-----G----A--GCAA-----T--T-GT-----G--AC---G------TG--CT----C-A--T--A--C--C--------A-A------A- 7235
H2B.JP.01.KR020.AB100245      ....TCTAAG--CA--AGT---C------GA-A-GAC--A-----CCGAG-T--C--C-------...--TA-A---TTT---A-C---G---CA---CG-----T--------A------GT---TG-ACCG---AC---T-----C--C-----TG-A-A---G-AC 6373
H2AB.CM.03CM_510_03.NEW       .............ACCACA--AA-AG-G--T-A-GACC---------AGTAGAA-C--C-------...--A--A---A-T--AG-G---G-T-CA-----C----GT------------G--CA---T---------GAG---T-----G--A-----T--ACAG---G--- 6338
H2G.CI.92.ABT96.AF208027      .......GTG-C--CT--AC-A--G---AGC-----CCAG--CT--ATGT-C--------T---...--TA-AC-GTC---ATCC---G-TG----C-----T------------C-------G--TC---G---G--A--------G--------T--ACA-------C 6593
H2U.FR.96.12034.AY530889      ..................G-AA-A-----A--AGAGC--------CGAA-TA---GC---T--...--A--A--------A-C-T--G-G--T----G---T--T--------C--------G--T------C-------T--A-----A-----T---CAT--CA-A- 6725
SMM.SL.92.SL92B.AF334679      A......TCAGG--AG--A--C--T--C--C----TG-------C-AAGA-CA-C------T-T...---A-A--G--G--A-C-------G-T---C----T-----------AC-----------C------GGC-------------C--------CAG----C-A- 6622
MAC.US.x.MAC239               ....TCAGCA------------------------------------------------------...---------------------------------------------------R--------------------------------------------------- 6257
MAC.US.x.251                  ..........---A--A-----A------------------------TT--T------------...------------------------G------------------------------------------------------------------------------ 7176
MAC.US.x.251_32H              ..........---A--------------------------------A-T--T------------...------------------------G----------------------------------------C---G------------------------A-------- 7178
MAC.US.x.251                  ..........---A----------------------------------T---A-----------...-------------------------------------C---------------------------C---G------------------------A-------- 7162
MAC.US.x.17EC1.AY033233       ................---------------------------------------------G--...-----------------------------------------------------T------------------------------C------------------ 7180
MAC.US.x.17EFR.AY033146       ................---------------------------------------------G--...-----------------------------------------------------T------------------------------C------------------ 7180
MAC.US.x.1937.AY611495        ....TCAGCAR--------------------------Y--------------------------...------------------------------------------------------------------------------------------------------- 6286
MAC.US.x.2065.AY611493        ....TCAGCA------------------------------------------------------...---------------------R--------------------------------------------------------------------------------- 6286
MAC.US.x.MM142.Y00277         .TCAGTAGAG-C-AG------A---------------C-----G---TT--T------------...----------------C----------------------------------------------------G--------------------------------- 6675
MAC.US.x.SMM142B.BD131285     .TCAGTAGAG-C-AG------A---------------C-----G---TT--T------------...----------------C----------------------------------------------------G------------------------------C-- 6672
MNE.US.82.MNE                 .AAAGCAGAGGC-A--A-AG------------A----C--------AAT--T------------...----------------C------------------------------------------------G--G---------------------------------- 6648
MNE.US.x.MNE027.U79412        .......AAGG--A---AG-----------A-A----A-----G--AAT-GT------------...----------------C---------------------------------------------------G---------------------------------- 6651
SMM.US.x.F236                 ACCAAGTGTAGC--A-A-TG-TA-A-----A-G--A-C---C---AAAA-TA---G---TG--...----------G---CC-----G-T-------T-----------------T--G-----A--G-----A--T--------A-----T--AAG------AA 7133
SMM.US.x.H9.M80194            ATCACCAATA-T--C-A-GG-T--A-----C-G-GA-C--------AGAAGTA-------T---...-----------G---CC--G---T-------T--------------------------A--G-G---A--T--------A-----T--AAG------A- 6654
SMM.US.x.PBJ14                ATCACCAATA-C--C-A-GG-T--A-----C-G-GA-C--------AAAATTA-------T---...-----------G---CC--G---T-------T--------------------------A--G-G---A--T--------A-----T--AAG------A- 6963
SMM.US.x.PBJA.M31325          ATCACCAATA-C--C-A-GG-T--A-----C-G-GA-C--------AAAATTA-------T---...-----------G---CC--G---T-------T--------------------------A--G-G---A--T--------A-----T--AAG------A- 6962
SMM.US.x.PBJ                  ATCACCAATA-C--C-A-GG-T--A-----C-G-GA-C--------AGAAGTA-------T---...-----------G---CC--G---T-------T--------------------------A--G-G---A--T--------A-----T--AAG------A- 6669
SMM.US.x.PBJ                  ATCACCAATA-C--C-A-GG-T--A-----C-G-GA-C--------AAAATTA-------T---...-----------G---CC--G---T-------T--------------------------A--G-G---A--T--------A-----T--AAG------A- 7155
SMM.US.x.PGM53.AF077017       AGAAGCAATA-C--C-A-AG-TA-A-----A-G-GA-C--------AG-A-CA-------T---...-----A--G--G--A-C----G--------G---T-------------------AG-----A--G-----A--T--------A-----T--AAG---A-A- 7114
SMM.US.x.SME543.U72748        AACAACAGTA-C-CC-A-GG-TA-A-----AGG-GA------C---AAAA-TA---G---TG--...-----------G----CC-----G-T-------T-----------------G--G-----A--G-T---A--T--------A-----T--AAG------AA 7190
STM.US.x.STM.M83293           .......ACA--GCC---GT-A--A-----A--A--C--C---G--AGTA-CA-C--C-------...---------G---ATC-C-AG--G-T-----------T--------AC----------T--A--G-G---------------G---------AGC---C-AA 6822
MAC.US.x.80035.AY611486       ....TCAGCA------------------------------------------------------...------------------------------------------------------------------------------------------------------- 6286
MAC.US.x.81035.AY599200       ....TCAGCAR-----------------------------------------------------...------------------------------------------------------------------------------------------------------- 6287
MAC.US.x.85013.AY611490       ....TCAGCAG-----------------------------------------------------...----------------------------------------------------------------------------------------------R-------- 6220
MAC.US.x.92050.AY603959       ....TCAGCA------------------------------------------------------...------------------------------------------------------------------------------------------------------- 6280
MAC.US.x.92077.AY599201       ....TCAGCA------------------------------------------------------...------------------------------------------------------------------------------------------------------- 6287
MAC.US.x.93057.AY611492       ....TCAGCAR-------------------------------Y---------------------...-----------------------------------------------------------Y----M-------------------------------------- 6285
MAC.US.x.93062.AY607704       ....TCAGCA------------------------------------------------------...------------------------------------------------------------------------------------------------------- 6272
MAC.US.x.95058.AY611494       ....TCAGCA-----------------------Y------------------------------...----------------------------------------------------------------------------------------------R-------- 6274
MAC.US.x.95086.AY607703       ....TCAGCAG-----------------------------------------------------...----------------------------------------------------------------------------------------------R-------- 6143
MAC.US.x.95112.AY588946       ....TCAGCA------------------------------------------------------...------------------------------------------------------------------------------------------------------- 6276
MAC.US.x.96016.AY607701       ....TCAGCAR-----------------------------------------------------...-------------------------------------------------------------------MR---------------------------------- 6287
MAC.US.x.96020.AY611488       ....TCAGCA------------------------------------------------------...------------------------------------------------------------------------------------------------------- 6279
MAC.US.x.96072.AY611491       ....TCAGCA-----------------------------------------R------------...-------------------------------------------------------------------MM---------------------------------- 6291
MAC.US.x.96081.AY597209       ....TCAGCA------------------------------------------------------...------------------------------------------------------------------------------------------------------- 6289
MAC.US.x.96093.AY611489       ....TCAGCAG-----------------------------------------------------...------------------------------------------------------------------------------------------------------- 6287
MAC.US.x.96114.AY588945       ....TCAGCA------------------------------------------------------...------------------------------------------------------------------------------------------------------- 6282
MAC.US.x.96123.AY611487       ....TCAGCA------------------------------------------------------...------------------------------------------------------------------------------------------------------- 6176
MAC.US.x.96135.AY607702       ....TCAGCA------------------------------------------------------...------------------------------------------------------------------------------------------------------- 6272
MAC.US.x.97009.AY599199       ....TCAGCA---------------------------------------------------G--...---------------------------------------------------------R--------------------------------------------R 6155
MAC.US.x.97074.AY599198       ....TCAGCA------------------------------------------------------...------------------------------------------------------------------------------------Y------------------ 6155
MAC.US.x.r80025.AY576480      ....TCAGCAG-----------------------------------------------------...----------------------------------------------------------------------------------------------R-------- 6143
MAC.US.x.r90131.AY576481      ....TCAGCA------------------------------------------------------...---------------------------------------------------------R--------------------------------------------- 6286
```

```
                              V3 loop start
MAC.US.x.239                 TATGTGAACAAGGGAATAAC...ACTGGTAATGAAAGTAGATGTTACATGAACCACTGTAACACTTCTGTTATCCAAGAGTCTTGTGACAAACATTATTGGGATGCTATTAGATTTAGGTATTGTGCACCTCCAGGTTATGCTTTGCTTAGATGTAATGACACAAATTAT 7347
Env                          V__C__E__Q__G__N__N__.__T__G__N__E__S__R__C__Y__M__N__H__C__N__T__S__V__I__Q__E__S__C__D__K__H__Y__W__D__A__I__R__F__R__Y__C__A__P__P__G__Y__A__L__L__R__C__N__D__T__N__Y_
H2A.CI.88.UC2.U38293         -T-----CAGCACA--C--G...--A--C-CAA-C-CG-C-----------GA--T--------A--A--C---A-G--A--A-----T---------------T----G-AG-----A--C--CA----A--G--------C--A--A-----C-----T--C--C--C 7377
H2A.DE.x.BEN.M30502          -T-----CA-CACA-CA................-CTG-C-C-----------GA--T--C-----A--AA-C---A-------A-----T--G--C--------------G--G-----A--C--------A--G--C-T---CC-AT-A-----------T--C--C--- 7429
H2A.DE.x.PEI2.U22047         -T-----CAC-ACT---GGG.........-CCAGCC-C-A---------------T--C-----A--A--C---AC------A-----T--G--C--------------G--G-----A--C--------A--G----TATGC-----A-----C-----T--C------ 7370
H2A.GH.x.GH1.M30895          -T------TC-AAT----C-...-AA-A-GGGA---AC-----------------T--C-----A--A--C---AC------A-----T--G--C--------------A-AG-----A--C--------A--G--------CC-A--A-----C-----T--C------ 6846
H2A.IN.NNVA.NEW              -T-------C--AT-GC-C-ACT-GCAC-CGCA-C-A---C-CC-CAGGCC-G---ATGCG---G-GGTAC--GAG---T--A-----T--G-------------A---G--G-----A-----C-----A--G--------C-----A-----------T--C-----C 7424
H2A.IN.95.CRIK               AT-------C--AT-GC-C-......ACA--CTCG--A-AG--------------T--C-----A--A--C---AC------A-----T--G--C----------A---A--G-----A-----------A--G--------C-----A-----C-----T--C-----C 7171
H2A.GM.x.MCR35.AY509260      -T------GG-AATGGC-C-............AC---C-C---------------T--C-----A--G--C--TAC------A--------G--C--------------G--G-----A--C--------A--G---T---CC-A--A-----C-----T--C------ 6859
H2A.GM.87.D194.J04542        -T-----CA-GACA--CGGA...........AC-G-C-C-----------GA--T--C-----A--A--C---A-------A--------G--C--------------G-AG-----A--C--------A--G---T---CC-A--A-----C-----T--C--C--- 6837
H2A.GM.x.ISY.J04498          -T------TC--AT--C-G-......ACAG-CCG--AA-----------------T--C-----A--A--C---AC------A--------G--C--------------G-----A--C--------A--G---T--TC-----A--G--C-----T--C-----C 6843
H2A.GM.x.MCN13.AY509259      -T------GG-AATGGC-C-............AC-GA--C---------------T--C-----A--G--C---AC------A--------G--C--------------G--G-----A--C--------A-------T---CC-A--A-----C-----T--C------ 6859
H2A.GW.86.FG.J03654          -T-----CA-TAAC..........-CTCA-G-C-G--C-AG--------------T--C-----A--A--C---AC------A-----T--G--C--------------G--G-----A--C--------A--G---T---C--A--A-----C-----T--C------ 6834
H2A.GW.87.CAM2CG.D00835      -T-----TA-TA-C-CAG-TCAG--CACA-----G-CC-CG--------------T--C-----A--A--C---AC------A-----T--G--C--------------G--G-----A--C--------A--G---T---CA-A--A-----C-----T--C--G--C 7432
H2A.GW.x.ALI.AF082339        -T-------C-TTT--C-C-...--CACA--CC-G-CC--G--------------T--C-----A--A--C---AC------A-----T--G--C--------------A--G-----A--C--------A--T-----C--CC-A--A-----CG----T-TC------ 7410
H2A.GW.x.MDS.Z48731          -T----G-G--AT-------...--CACA--CC----C--G--C-------G---T--C-----A--A--C---A-------A-----T--G--C---------A---G--G-----A--C--------A--G---T---C-----A-----C--------C------ 6903
H2A.SN.85.ROD.M15390         -T-----GAC-...-----T...-GCACA---C-G-CCCAG--------------T--C-----A--A--C---AC--A--A--------G--C--------------A--G-----A--C--------A--G--------CC-AT-A-----------T--C------ 6872
H2AB.CI.90.7312A.L36874      -T-----GTC-AAT--CGC-...-G--A-GGGAG-GAC-----C--------T--T--C-----A--A--C---AC---A--A--C-----G--C--C-----------A--G-----A--C--------A--G----T----C----A---------------C--C--- 7411
H2B.CI.88.UC1.L07625         AG--CA-CGGCAAC-GC.............-CCAGC-CC-T-----------GAACT-----T--C--A--G-----------A-----T--------C-----CAGCT-A-----------C-------T--C-----A-----C----A---------------C--C--- 7401
H2B.CI.x.EHO.U27200          -G------A-G----CA.......GGA----------C-A------T--A--AAC------T--A--AA----------A--A-------------------AG-T-A-----------------T--C-----A-T-----------A-----C-----T--T--A--- 7391
H2B.GH.86.D205.X61240        AG---A-TA-TACC-GG.........AAGT--ACC--C-----C--T--A-GAAC---C-----AA-AA------------A-------------------CAGCT-A--G-----------------T--C--G--G-T-TT-C-A--A-----------T--C--C--- 7396
H2B.JP.01.KR020.AB100245     AG--.........................----GG--C-----CC-T--A--AAC--------A--AA------------A-------------------CCAAT-A---------------------C--C-----C-T------TA-G-----C-----TT-T--C--- 6519
H2AB.CM.03CM_510_03.NEW      -T-----CA--AAC---TCAAGCGAAACATTAA-C-AA-C-----------GA--T--------A--A--C---A----A--A-----T--G--C---------A----G--------A--C---A----A--G--------C--A--A-----C-----T--C--C--- 6508
H2G.CI.92.ABT96.AF208027     A---Y-----GTCC-----G...-G--AG-----G--------C--T----GG--------T--A--A--C--T--G--A--A-----T-------------------A--------A--------T--G-----G-T-----------A-----------T--T--C--- 6760
H2U.FR.96.12034.AY530889     -T--------GAAT-C-.......-G--AA-G--------A-----------------C-----C-----------------------------------------C--A------------------------G-----C------C--T-A-----------T--T-----C 6889
SMM.SL.92.SL92B.AF334679     -C-----G-----AGGA--T.........G-AAGC----------------T--T--C--T--AAG------A--------A-------G--C-----------C--C----A------C--------A-----C-----------------------T--------- 6783
MAC.US.x.MAC239              ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6424
MAC.US.x.251                 ------------------G-...---------------------------T-------------------------------------------------------------------------------------------------------------------- 7343
MAC.US.x.251_32H             -T----------------G-...----A-------------C--------T---------------A------------------G-----------A------------------------------------------------------------------- 7345
MAC.US.x.251                 -T----------------G-...----A---------C-----C--------T--------------------------A----------------------A---------------------------------------------------------------- 7329
MAC.US.x.17EC1.AY033233      ---------------------...-------------------------------------------------A------------------------------------------------------------------------------------------------ 7347
MAC.US.x.17EFR.AY033146      ---------------------...-------------------------------------------------A------------------------------------------------------------------------------------------------ 7347
MAC.US.x.1937.AY611495       ---------------------...---------------------------------------------------------------------------------------------------------W-------------------------------------- 6453
MAC.US.x.2065.AY611493       ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6453
MAC.US.x.MM142.Y00277        -T----------------G-...----------------------------T--------T-------------------G----------G--------------------G---A------------------------------------------------- 6842
MAC.US.x.SMM142B.BD131285    -T----------------G-...----------------------------T--------T-------------------G----------G--------------------G---A------------------------------------------------- 6839
MNE.US.82.MNE                -T-----------T----G-...----AA----------------------T-------------------T-----A------------------------------------------------------------------------------------------ 6815
MNE.US.x.MNE027.U79412       -T-----------T----G-...----AAG---------------------T-------------------T-----A---------------------------------------------------------------------------------------G--- 6818
SMM.US.x.F236                -C-----G--GTCAGCG--T......-AA-G---G----A-----------C-T--T--------CAG------T-----A--C--------G---------------------------A--C--------G--------------------G--------TT--------- 7297
SMM.US.x.H9.M80194           -G---------AAT-GC--TGGA-A--AA-C---C----A---C--T--------T--------CAG------T-----A--C--------G------------------------Y--A-----------G--------------------G--------T--C------ 6824
SMM.US.x.PBJ14               -T---------AAT-GC--T......-AA-C---C----A---C--T--------T--------CAG------T-----A--C------------------------------------A-----------G--------------------G--------TT--------- 7127
SMM.US.x.PBJA.M31325         -T---------AAT-GC--T......-AA-C---C----A---C--T--------T--------CAG------T-----A--C------------------------------------A-----------G--------------------G--------TT--------- 7126
SMM.US.x.PBJ                 -T---------AAT-GC--TGAA-A--AA-C---C----A---C--T--------T--------CAG------T-----A--C--------G------------------------Y--A-----------G--------------------G--------T--C------ 6839
SMM.US.x.PBJ                 -T---------AAT-GC--T......-AA-C---C----A---C--T--------T--------CAG------T-----A--C------------------------------------A-----------G--------------------G--------TT--------- 7319
SMM.US.x.PGM53.AF077017      -----------A-T-C--CT...-G-A-A----------A-----------T--T--------CAG------T-----A--C--------G---------------------------A--C--------G--------------------G--------T---------- 7281
SMM.US.x.SME543.U72748       -C-----G--GCCAGCA--T.......--A-G---G----A----------C-G--T--------CAG------T--G--A--C--------G---------------------------A--C--------G--------------------G--------TT--------- 7354
STM.US.x.STM.M83293          -C---------AATGTC-CT.......--CG-A--G--C-----C-------GA--T--C--T--CAG---A--T-----A--C-------------------------CT-A--G-----------------A--T-------------C-----C-----T--------- 6986
MAC.US.x.80035.AY611486      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6453
MAC.US.x.81035.AY599200      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6454
MAC.US.x.85013.AY611490      ---------------------...----A----------------------------------------------------------------R-------------------------------------------------------------------------- 6387
MAC.US.x.92050.AY603959      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6447
MAC.US.x.92077.AY599201      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6454
MAC.US.x.93057.AY611492      ---------------------...------R---------------------------------------------------------------------------------------------------------------------------------------- 6452
MAC.US.x.93062.AY607704      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6439
MAC.US.x.95058.AY611494      ---------------------...----R------------------------------------------------------------------------------------------------------------------------------------------ 6441
MAC.US.x.95086.AY607703      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6310
MAC.US.x.95112.AY588946      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6443
MAC.US.x.96016.AY607701      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6454
MAC.US.x.96020.AY611488      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6446
MAC.US.x.96072.AY611491      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6458
MAC.US.x.96081.AY597209      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6456
MAC.US.x.96093.AY611489      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6454
MAC.US.x.96114.AY588945      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6449
MAC.US.x.96123.AY611487      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6343
MAC.US.x.96135.AY607702      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6439
MAC.US.x.97009.AY599199      ---------------------...------------------------R---------------------------------------------------------------------------------------------------------------------- 6322
MAC.US.x.97074.AY599198      ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6322
MAC.US.x.r80025.AY576480     ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6310
MAC.US.x.r90131.AY576481     ---------------------...----------------------------------------------------------------------------------------------------------------------------------------------- 6453
```

V3 loop end

```
MAC.US.x.239                TCAGGCTTTATGCCTAAATGTTCTAAGGTGGTGGTCTCTTCATGCACAAGGATGATGGAGACACAGACTTCTACTTGGTTTGGCTTTAATGGAACTAGAGCAGAAAATAGAACTTATATTTACTGGCATGGTAGGGATAATAGGACTATAATTAGTTTAAATAAGTATTA 7517
Env                         _S__G__F__M__P__K__C__S__K__V__V__V__S__S__C__T__R__M__M__E__T__Q__T__S__T__W__F__G__F__N__G__T__R__A__E__N__R__T__Y__I__Y__W__H__G__R__D__N__R__T__I__I__S__L__N__K__Y__Y
H2A.CI.88.UC2.U38293        --G------GCA-----T--CC----A--A--A-CTG----------T----------A--G--A-----------------------C------------------A-----C--T----------A--C-----------C-----C-----C--AC---- 7547
H2A.DE.x.BEN.M30502         ---------GAA-----G--CA----A--A--A-CTG---------------------A--G--A-----------------------C------------------A-----C--T--------C--A-------------C-----C-----C-------- 7599
H2A.DE.x.PEI2.U22047        ---------GA---C--G---C----A--A--A-CTG--A---------A--------A--G--A-----------------------C-----G------------A-----C--T----------A-------------T--C--C-------CAC---- 7540
H2A.GH.x.GH1.M30895         --------CGA---A--G--------A--A--A-CT---A------------------A--G--A--C--C-----------------C-----G-----G-------A-----C--T--------C--A-------------C--C--C-----C-------- 7016
H2A.IN.NNVA.NEW             ---------GA---C--T---A----A--A--A-CTG--------C------------A--G--A----C--G---------------C------------------A-----C--T------A---A----C--A--------C--C-----C--AC---- 7594
H2A.IN.95.CRIK              ---------GCA--C--T--------A--A--A-CT--CA-T-----C----------A--G--A-------A---------------C------------------A--CC----T------A---AAAG-G-C--A--------C--C-----C--A----- 7341
H2A.GM.x.MCR35.AY509260     ---------GC---C--T--C--------A--A-CTG--A-------C--A-------A--G--A-------A---------------C------------------A-T---C--T-----------------C--A----C--C--C-----C-------- 7029
H2A.GM.87.D194.J04542       ---------GAA-----G--C-----A--A--A-CTG---------------------A--G--A-----------------------C------------------A-----C--T------------A-------------C-----C-----C-------- 7007
H2A.GM.x.ISY.J04498         ---------GA---C--T--C-----A--A--A-CT---A----T----A--------A--G--AC---------C------------C-----G------------A-----C--T-----------------C--A----T--C--C-----C--A----- 7013
H2A.GM.x.MCN13.AY509259     ---------GC---C--T--C--------A--A-CTG--A-------C--A-------A--G--A-------A---------------C------------------A-T---C--T-----------------C--A----C--C--C-----C--A----- 7029
H2A.GW.86.FG.J03654         ---------GCA-----C--C-----A--A--A-CTG--A-------C--A-------A--G--A-------A---------------C------------------A-----C--T----------AA-----C--A----T--C--C--------C-T--- 7004
H2A.GW.87.CAM2CG.D00835     ---------GCA--C--C-----------A--A-CT---A-------C--A-------A--G----------G---------------C--C-------G-------A-----C--T----------AA-----C--A----C--C--------C---C---- 7602
H2A.GW.x.ALI.AF082339       ---------GCA--C--T--C-----A--A--A-CTG--A------------------G--A--------------------------C-----G------------A-----C--T-----------A-------A----C--C--C-----C--AC---- 7580
H2A.GW.x.MDS.Z48731         -----G---CA---C--T--------A--A--A-CT---A----T--C----------A-----A-------G---------------C------------------A-----C--T----------AA-----C--A----C--C--C------GCAAG--- 7073
H2A.SN.85.ROD.M15390        ---------GCA--C--C--------A--A--A-CT---A-------C----------A--G--A----C--A---------------C-----------G------A-----C--T--------C--A---------A----C--C--C-----C--A----- 7042
H2AB.CI.90.7312A.L36874     --------------C--C---AG------A--A--G--C--T--------A-------A-------C-----A-----------C-----T--A--G----------G--A-----G--T------A---AA-------------T--A--C--G---------- 7581
H2B.CI.88.UC1.L07625        --------------A------AG---A--A--A--G--C--T--------A---------------C-----A-----------C-----T--G--GA---------G--A-----G--T------A---AA-----------C--C--A--C--G---------- 7571
H2B.CI.x.EHO.U27200         --------C-----C--C---AG------A--A--A---...CTGTAC--A-------A-------C-----A-----------C-----T--A--G-------C--G--A-----C--T-------------AA--C-----------C--A--C------TCA----- 7558
H2B.GH.86.D205.X61240       --------C-----C--C--CAG------A--A-CG--C--C--------A-------A-----T-C-----A-----------C-----T--A--G-----G--C--G--A-----A--T--------AA-AA--C--------C--C--A--C-------CA--C-- 7566
H2B.JP.01.KR020.AB100245    --------------C--C---AGC-----A--A--A--C--C--T-----A-------A-------C-----G-----------C-----T--A--G----------G--A-----A-------------AAAGC--------CG-C--A--C--------AA-A-- 6689
H2AB.CM.03CM_510_03.NEW     ---------CAA-----T--C-----A--A--A-CTG----------------A----A--G--A-------A---------------C------------------A-----C--T-----------A-------------C-----C-----C-------- 6678
H2G.CI.92.ABT96.AF208027    -----A---GCA-----T--C-----A--A----------A----T-----A-------A-----A--C-----A-----C--G-----------A--G-------------A-----C--T--------G--A--C---------------A---------------- 6930
H2U.FR.96.12034.AY530889    -----A-----------T--C-----A--A--A------A---------A-----A----G--GA-----A--C--------------A--------C------T--------------A--------------A--C-----A--A--------C-------------- 7059
SMM.SL.92.SL92B.AF334679    ---------GCA-----T---AG------A--A--A--A-----T-------------A-------------A---------------T---------G-------A-----A--T--------CG-AAG------A--A--------C-------------- 6953
MAC.US.x.MAC239             ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6594
MAC.US.x.251                -------------------------------------------------------------------------------------------------------------------------------A-------------------------------- 7513
MAC.US.x.251_32H            ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 7515
MAC.US.x.251                ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 7499
MAC.US.x.17EC1.AY033233     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 7517
MAC.US.x.17EFR.AY033146     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 7517
MAC.US.x.1937.AY611495      ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6623
MAC.US.x.2065.AY611493      ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6623
MAC.US.x.MM142.Y00277       ----------------C----------A----------------------------------------------C-G--------------------------C-------------------A--------------------C-------C---- 7012
MAC.US.x.SMM142B.BD131285-  ----------------C----------A----------------------------------------------C-G--------------------------C-------------------A--------------------C-------C---- 7009
MNE.US.82.MNE               ----------------C--------------------------------A--------------------------------------------------------------A-C-AA---------------------G---------- 6985
MNE.US.x.MNE027.U79412      ----------------C--------------------------------A--------------------------------------------------------------A-C-AA---------------------G---------A- 6988
SMM.US.x.F236               -T-------GCT----C----------A-----T---------------A----------G--A--C---------------C-----T-----G-------------A--C-----T--------C-AAAG------A--C---------C--G---------- 7467
SMM.US.x.H9.M80194          ---------GCT----C---A------A--A--T---------------A-------A-----A--C------------T--C-----T-------------------A--C-----T------R-C--AAGC-----A--C---------C-------------R 6994
SMM.US.x.PBJ14              ---------GCT----C---A------A--A--T---------------A-------A-----A--C------------T--C-----T-------------------A--C-----T--------C--AAGC-----A--C---------C-------------- 7297
SMM.US.x.PBJA.M31325        ---------GCT----C---A------A--A--T---------------A-------A-----A--C------------T--C-----T-------------------A--C-----T--------C--AAGC-----A--C---------C-------------- 7296
SMM.US.x.PBJ                ---------GCT----C---A------A--A--T---------------A-------A-----A--C------------T--C-----T-------------------A--C-----T------R-C--AAGC-----A--C---------C-------------R 7009
SMM.US.x.PBJ                ---------GCT----C---A------A--A--T---------------A-------A-----A--C------------T--C-----T-------------------A--C-----T--------C--AAGC-----A--C---------C-------------- 7489
SMM.US.x.PGM53.AF077017     ---------GCT----C----------A-----T---------------A----------G--A--C------------T--C-----T-----G-------------A--C-----T------A----AAG------A--C---------C-------------- 7451
SMM.US.x.SME543.U72748      ---------GCT---------------A-----T---------------A----------G--A--C---------------C-----T-----G-------------A--C-----T--------C-ATAG------A--C---------C-------------- 7524
STM.US.x.STM.M83293         A--------GCA-----T--C--------A--A-----C--G--T--T--A-------------A-----A--------------------------------------A-----C--T--------C-----------------T-----C--G--C-------- 7156
MAC.US.x.80035.AY611486     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6623
MAC.US.x.81035.AY599200     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6624
MAC.US.x.85013.AY611490     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6557
MAC.US.x.92050.AY603959     -------------------------------------------------------------------------------------------------------------------------------------------------------YW--- 6617
MAC.US.x.92077.AY599201     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6624
MAC.US.x.93057.AY611492     ---------------R------------------------------------------------------------------------------------------------------------------------------------------------ 6622
MAC.US.x.93062.AY607704     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6609
MAC.US.x.95058.AY611494     ----------------T--Y-------------------------------------------------------------------------------------------------------------------------------------------- 6611
MAC.US.x.95086.AY607703     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6480
MAC.US.x.95112.AY588946     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6613
MAC.US.x.96016.AY607701     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6624
MAC.US.x.96020.AY611488     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6616
MAC.US.x.96072.AY611491     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6628
MAC.US.x.96081.AY597209     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6626
MAC.US.x.96093.AY611489     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6624
MAC.US.x.96114.AY588945     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6619
MAC.US.x.96123.AY611487     ------------------------------------------------------------------------------------------------------------------------------------------------------C---- 6513
MAC.US.x.96135.AY607702     ------------------------------------------------------------------------------------------------------------------------------------------------------C---- 6609
MAC.US.x.97009.AY599199     ---------------------------------------------------------------R------------------------------------------------------------------------------------------------ 6492
MAC.US.x.97074.AY599198     ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6492
MAC.US.x.r80025.AY576480    ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6480
MAC.US.x.r90131.AY576481    ---------------------------------------------------------------------------------------------------------------------------------------------------------------- 6623
```

```
                                           V4 loop start                                                                            V4 loop end
MAC.US.x.239                 TAATCTAACAATGAAATGTAGAAGACCAGGAAATAAGACAGTTTTACCAGTCACCATTATGTCTGGATTGGTTTTCCACTCA......CAACCA...ATCAATGATAGGCCAAAGCAGGCATGGTGTTGGTTTGGAGGAAAATGGAAGGATGCAATAAAAGAGGTGAAGC 7678
Env                          __N__L__T__M__K__C__R__R__P__G__N__K__T__V__L__P__V__T__I__M__S__G__L__V__F__H__S__.__.__Q__P__.__I__N__D__R__P__K__Q__A__W__C__W__F__G__G__K__W__K__D__A__I__K__E__V__K__
H2A.CI.88.UC2.U38293         ------T------C-T--C-AG---------------------G-----A-A--AC-------A---CATAGG--T--T--T......--GG--GTC-----CA-A-AA--T-GA-----------C------AA---C--C-----A-GA--C--GC-G---------- 7711
H2A.DE.x.BEN.M30502          ------C------CGT----AG---------------------------A-A--AC-------A-----A--G--T-----T......--G---...-----CACA-----T-G------------CC----------C-G-----G---A--C--GC-G---------- 7760
H2A.DE.x.PEI2.U22047         ------C------C-T----AG--G------------T-------G---A-A--AC---G---A--GAGA--G--T-----C......-G---GATC--------A--A--C--------------C-----C-----TG-T------A-A--C--GC-G--------A- 7704
H2A.GH.x.GH1.M30895          ---C--TT-T--AC-T----AG-----------C---------G-----A-A--AC-C-----A--G--A--G--T-----C......--G---...------ACA--A--T-G---A--------C-----CAA---C--G----G---A--C--GC-G-----A---- 7177
H2A.IN.NNVA.NEW              ------CT-T--AT-T----AG--------------A-----AG----TA-A--ACA----C-A--C--A--G--T--T--C......--G---...------ACC--A--T-GA--A--------C-----CAA---C-G---------A--C--GC-G---------- 7755
H2A.IN.95.CRIK               ------CT-C--AC-T----AG--G-----------A-----AAC----A-A--AC-------A----ATAAG--T--T--C......-GG---GTA------ACA--A--T--A--A--------C-----CAA---CCGG--------C--C--GC-G-----A---A 7505
H2A.GM.x.MCR35.AY509260      ------C--T--AC-T----AG--G-----------------GG-----A-A--AC-------A--G--AAGG--T-----C......--G--GGTC------A-A--A--C-GA--A--------------CAA---TG---------GA--C--GC-G---------G 7193
H2A.GM.87.D194.J04542        ------C------C-T----AG---------------------G-----A-A--AC-------A--GCGAAGG--T-----T......-GG---GTCTA---CA-A-AA--TGG-------------------CA---C--C----TA--A--C--GCGG---------- 7171
H2A.GM.x.ISY.J04498          ------C--C--ACTT-----G-------A------A------G-----A-A--AC-C-----A--CCGCAGA--T-----C......--GAAGATC------A-A-AA--C-G---A--------CC----CAA---CG-G----G---A--C--GC-G--------A- 7177
H2A.GM.x.MCN13.AY509259      ------C--T--AC-T----AG--G-----------------GG-----A-A--AC-------A--G--AAGG--T-----C......--G--GGTC------A-A--A--C-GA--A--------------CAA---TG---------GA--C--GC-G---------G 7193
H2A.GW.86.FG.J03654          ------C--T---C-T----AG--G--G---------------G-----A-A--AT-------A--G--TAAG--T-----C......--G---GTC------A-A-AA--C-G---A--------------C-A---CC----------A--C--GC-G---------G 7168
H2A.GW.87.CAM2CG.D00835      ---C--CT-T---T-T-----G--G-----------------GG-G---A-A--AC-------A--GCAAAGA--T-----C......-GG---ATC--T---A-A-----C-G---A--------C-----CAA---C--C----CA--A--C--GC-G-----A---- 7766
H2A.GW.x.ALI.AF082339        ------T--T---C-T----AG--G------------------G-----A-A--AC-------A--G--AA-A--T-----C......--G---...------A-A--A--C-GA--A--------C-----CAA---CG------G-A-A--C--GC-G---------G 7741
H2A.GW.x.MDS.Z48731          ------CT-CT--C-T----AG--G-----------A-----GG-----A-A--AC-------A-----A--G--T--T--C......--G--G...------AGA--A--T-G---A--------C-----CAAG--C--C----G-AGA--C--GC-G-----A---G 7234
H2A.SN.85.ROD.M15390         ------C-GTT--C-T----AG--G-----G-------T---GAA--A-A-A-TGC-------A---CAT--G--T-----CCACTAC--G--G...------A-A--A--C-GA--A--------C-----CAA---C--------A--C--C--GC-G---------G 7209
H2AB.CI.90.7312A.L36874      ----T-------AC-T----AG--G--------C---------G-----A-A--AC-C-----A--G--A--G-----T--C......--G--T...------A-A--A--T-G---A--------C-----CAA---CG-G----G---A--C--GC-G--------A- 7742
H2B.CI.88.UC1.L07625         -------------C-T--C-----G-----G--C--------CA-----A-A--A--C-----A---C--AA------T--G......--G--T...C-----ACC--A--C-GA-----T-----C------AAG-----C----TA--A--C----GG-----A---G 7732
H2B.CI.x.EHO.U27200          ------G------C-C----A---G--------C--A-TG---G-G---A-A--GA-CCG---A--TA-TC-C-----T---......--G--T...------A-A--A--T--A--A--T-----C------AA------C----CA--A--C---C-G--------AG 7719
H2B.GH.86.D205.X61240        ----T-GT----AC-C----AG--G--------C-----G---G-----A-A--GA-CCG---A---C-AC-------T---......--G--T...------A-G--A--C-GA--A--T-----C------AAG-----C----CA--A--C-----G--------AA 7727
H2B.JP.01.KR020.AB100245     ------G-----AC-C--------G--------C--A-----CG-G-----A--A--------A--TC-AA-C-----T---......--G--T...--T---C-A--A--C-G---A--C-----C--------------C----CAC-G--T-----G---------G 6850
H2AB.CM.03CM_510_03.NEW      ------C--T---C-C----AG---------------------G-----A-A--AC-----C-A---C-A--G--T-----T......--G---...-----CCGC-----T-GA-----------C------A----C--C----CA--A--C--G--G---------- 6839
H2G.CI.92.ABT96.AF208027     ----T-------A-R------G--------G-----A-----A------A-T--T--------A--GC-A--C--------T......------...--------A--A--C-GA-----T-----------------C-----------A--C--GC-G-----A---G 7091
H2U.FR.96.12034.AY530889     ----T--------T-C--------G-----G-----------C--------A-----------A---C-A--G-----T--T......---G-T...------A-C-AT-----A--A--T--------------T-----T-----A-GG-----CC-G---------G 7220
SMM.SL.92.SL92B.AF334679     ----T-G-----A--G--C------------------------CC-------------------A--T-----C--T--T--G......------...--A-----G-------GA--A--------C-------------------GA--A--T--GC-G-----T---A 7114
MAC.US.x.MAC239              ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6755
MAC.US.x.251                 ------------A---------------------------------------------------CA----------------......------...------------------------------------A------------------------------------ 7674
MAC.US.x.251_32H             ----------------------------------------------------------------------------------......------...--------------------------------------------T---------------------A------ 7676
MAC.US.x.251                 ----------------------------------------------------------------------------------......------...-----------------------------------------------------------------------A- 7660
MAC.US.x.17EC1.AY033233      ------------A---------------------------------------------------------------------......----G-...----------------G-------------------------------------------------------- 7678
MAC.US.x.17EFR.AY033146      ------------A---------------------------------------------------------------------......----G-...----------------G-------------------------------------------------------- 7678
MAC.US.x.1937.AY611495       -------------------------------------------------------------------W--------------......------...------------------------------C------M----G------------------------------ 6784
MAC.US.x.2065.AY611493       --------------------R-------------------------------------------------------------......------...------------------------------------------------------------------------- 6784
MAC.US.x.MM142.Y00277        -----------------------------------------------------------------C----------------......------...G-------G---------------------A-------------T--------G------------------- 7173
MAC.US.x.SMM142B.BD131285    -----------------------------------------------------------------C----------------......------...G-------G---------------------A-------------T--------G------------------- 7170
MNE.US.82.MNE                ---------------------------------------------------------C------------------------......------...-----------------A------------A------A------T--------G--------------A---- 7146
MNE.US.x.MNE027.U79412       ----------------------------------------------------------------------------------......------...--------A--------A------------A------A------T--------G------------------- 7149
SMM.US.x.F236                --------------G--------------A---------------------------------A--G-----C-----T--G......--G--C...--A-----G--A-----A-----C-----C-------A----GC------AA-G--C--CC-G--A------G 7628
SMM.US.x.H9.M80194           --------------G---------------------------C--------------------A--G-----C-----T--G......-----C...--A-----G--A-----A-----C-----C------------G-------A--G--C--CCGG--A------G 7155
SMM.US.x.PBJ14               --------------G---------------------------C--------------------A--G-----C-----T--G......-----C...--A-----G--A-----A-----C-----C--------------------AA-G--C--CC-G--A------G 7458
SMM.US.x.PBJA.M31325         --------------G---------------------------C--------------------A--G-----C-----T--G......-----C...--A-----G--A-----A-----C-----C------------G-------AA-G--C--CC-G--A------G 7457
SMM.US.x.PBJ                 --------------G---------------------------C--------------------A--G-----C-----T--G......-----C...--A-----G--A-----A-----C-----C------------G-------A--G--C--CCGG--A------G 7170
SMM.US.x.PBJ                 --------------G---------------------------C--------------------A--G-----C-----T--G......-----C...--A-----G--A-----A-----C-----C------------G-------AA-G--C--CC-G--A------G 7650
SMM.US.x.PGM53.AF077017      ------------A-G---------G-----------------C--------------------A--G---A-C-----T--G......-----C...--A-----G--A-----A--A--C-----C-------------GT----GA--G--C--CC-G---------G 7612
SMM.US.x.SME543.U72748       ---------------------------------------------------------------A--G-----C-----T--G......-----C...--A-----G--A-----A-----C-----CC-------------C----GC--G--C--CC-G--A------G 7685
STM.US.x.STM.M83293          C---T---------GC--------------G--C--A-----C--------------------A--C-----C-----T--G......-----C...--T-----G--A-----A--A--T-----C------------G------GA-GA--------------T---G 7317
MAC.US.x.80035.AY611486      ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6784
MAC.US.x.81035.AY599200      ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6785
MAC.US.x.85013.AY611490      ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6718
MAC.US.x.92050.AY603959      ------------R---------------------------------------------------------------------......------...----R----R------R-------------------------------------------------------- 6778
MAC.US.x.92077.AY599201      ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6785
MAC.US.x.93057.AY611492      ----------------------------------------------------------------------------------......------...------------------------------------------R------------------------------ 6783
MAC.US.x.93062.AY607704      ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6770
MAC.US.x.95058.AY611494      ----------------------------------------------------------------------------------......------...------------------------------C------A----R------------------------------ 6772
MAC.US.x.95086.AY607703      ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6641
MAC.US.x.95112.AY588946      ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6774
MAC.US.x.96016.AY607701      ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6785
MAC.US.x.96020.AY611488      ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6777
MAC.US.x.96072.AY611491      ----------------------------------------------------------------------------------......------...------------------------------------------G------------------------------ 6789
MAC.US.x.96081.AY597209      ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6787
MAC.US.x.96093.AY611489      ----------------------------------------------------------------------------------......------...--------A---------------------C-----------G----------G------------------- 6785
MAC.US.x.96114.AY588945      ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6780
MAC.US.x.96123.AY611487      ----------------------------------------------------------------------------------......------...------------------------------------------R----------R------------------- 6674
MAC.US.x.96135.AY607702      ----------------------------------------------------------------------------------......------...------------------------------------------R----------R------------------- 6770
MAC.US.x.97009.AY599199      ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6653
MAC.US.x.97074.AY599198      ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6653
MAC.US.x.r80025.AY576480     ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6641
MAC.US.x.r90131.AY576481     ----------------------------------------------------------------------------------......------...------------------------------------------------------------------------- 6784
```

```
                          ⌐V5 loop start
MAC.US.x.239             AGACCATTGTCAAACATCCCAGGTATACTGGA...ACTAACAATACTGATAAAATCAATTTGACG...GCTCCTGGA...GGAGGAGATCCGGAAGTTACCTTCATGTGGACAAATTGCAGAGGAGAGTTCCTCTACTGTAAAATGAATTGGTTTCTAAATTGGGTAGAA 7839
Env                      Q__T__I__V__K__H__P__R__Y__T__G__.__T__N__N__T__D__K__I__N__L__T__.__A__P__G__.__G__G__D__P__E__V__T__F__M__W__T__N__C__R__G__E__F__L__Y__C__K__M__N__W__F__L__N__W__V__E_
H2A.CI.88.UC2.U38293     -A---C---CAGG------------C-AA---...--C--TG----GAG---G--T--C--TGT-...AAA--A---GTA--CTC---C--A--G--G--A-A---------T--C-----------A--TT----------C----C------C--C--------G--G 7875
H2A.DE.x.BEN.M30502      -A---C----AC----------A--C-AA---...-TC--TG----A-GG-----T--C--T----...AAA--G---GCA--CTC---C--------GG-A--T---------T--C-----------A--T-----------C----C------C--C------------ 7924
H2A.DE.x.PEI2.U22047     -A---C----G----------------GA---...--C---G-C--AC-G-----T--C--T--A...CAA--A---AAA--TTC----G-A-----GGTA-A---------T--C-----------A--T--A-----C--C----C-C----C--C------A----- 7868
H2A.GH.x.GH1.M30895      -A---C--A-A----------------AA---...--C---G-C--AA-G--T--T--C--T--A...AAA--A---AGA--CTC---C--A-----GG-A-A---------T--C-----------A--T--------C--C----C------C--C------------ 7341
H2A.IN.NNVA.NEW          -A-------CAG----C----------GG---...--C--T-T---A--A-----T--A--T--A...--A--G---AAA--TTC---C--------GG-A-A---------T--C-----G-----A--T--------C--T----C------C--C------A----G 7919
H2A.IN.95.CRIK           -A---G---CAG-----------GCGGTA......--C--GG--C-AA-G--C--T-CC--TG-A...--A--G---AAA--CTC---C--------AGAA-A---------T--C-----G-----A--T--------CG-C----C------C--TG-----A----G 7666
H2A.GM.x.MCR35.AY509260  -A---C---CA-----------------AA---...--C--TG-A--AA-G--T--T--C--T--A...--A--A---AAG--CTC---C--A--G--GG-A-A---------T--C-----------A--T--------C--C----C------C--C------A----- 7357
H2A.GM.87.D194.J04542    -A---C---CA-------------CGGA---...--A--TG----A-GA-----T--C--T----...AAG--A---ATA--TTC---C--A-----G--A-A---------T--C-----------A--T-----------T----C------C--C------------ 7335
H2A.GM.x.ISY.J04498      -A---C----A-----------------AA---...--C--TG-C--AA-------T--C--T--A...--A--A-A-AAA-ACTC---C--A-----AG-A-AT--------T--C-----------A--------T--C--C----C------C--T------------ 7341
H2A.GM.x.MCN13.AY509259  -A---C---CA-----------------AA---...--C--TG-A--AA-G--T--T--C--T--A...--A--A---AAG--CTC---C--A--G--GG-A-A---------T--C-----------A--T--------C--C----C------C--C------A----- 7357
H2A.GW.86.FG.J03654      -----C---CG-----------A----AA--GAAC-GG-G-CGC--A--G--T--T--A--T-AA...--A--A--GAGA--CTC---C--A-----A--A-A---------T--C-----------A-CT-------C--C----C------C--C-----------G 7335
H2A.GW.87.CAM2CG.D00835  -A---C---CGG-G--------------AA---...--C--A--C-TAACGG-T--T-CC--C-AA...--A--A-A-AGA--TTC---C--A-----G--A-A-------T-T--C-----------A--TT-------C--C----CC-----C--C------------ 7930
H2A.GW.x.ALI.AF082339    -A---C----A-----------------AA---...--C--TG-C--AA-CC----T--C--T--A...AAA--A---AGA--CTC----G-A-----GGTA-AT--------T--C-----------A--T---C-T--C--C----C------C--C--------G--- 7905
H2A.GW.x.MDS.Z48731      -A------A-A-----C-----------AG---...--C-GT--C-TAACA--T--T-CC--C---...--A--A---AAA--CTC---C--A-----AG-A-A---------T--C-----------A--T--------C--C----C------C--C------------ 7398
H2A.SN.85.ROD.M15390     -A---C---CA-----------------GA---...--C--TG-C--AAGG--T--T-GC--TG-A...--G--A---AAA--CTC---C--A-----AG-A-A---------T--C--------------T--------C--C----C------C--C------A----G 7373
H2AB.CI.90.7312A.L36874  -A---C--A-A----------------AA---...--C--TG-C--AAGG--T--T-CC--T--A...AAA--A---ACA--CTC---C--A-----GG-A-A---------T--C-----------A--T--------C--C----C------C--C------------ 7906
H2B.CI.88.UC1.L07625     -------CA-A-----------------AA---...--A--T--C--C--G-GG--A-GA---GTA...-GG--CTCGGCA--TTC---C--A--G--G-GACAT--------T--C--T--G-----A---T----T-----T----CC------T-----C-----G--G 7896
H2B.CI.x.EHO.U27200      --------AAA--T--------A---T-A---...--A-CA----TCTCAC-G--A-GG--AG-A...-AG-AC-CGAGAA-CTC------A-----A-GA-AT--------T--C--T--G---------T-----------T----C--TT--CT----C--------- 7883
H2B.GH.86.D205.X61240    G------CA-A-----------------AA---GGTG-A--A----TCACA-GCG-A--G--AGTATCA-AA-A----AAA--TTC------A---ACA--A-AT--------C-----T--------A---T-A--------T----CC-----CT----C--------G 7897
H2B.JP.01.KR020.AB100245 ---------CA--G--------A-----A---...--A---G----C-CA-----A--G--AG-A...-AG-ACTCGCAA--CTC------A-----A-GA-AT--------T--C--T------------TT-G--------T----C------C-----C--------- 7014
H2AB.CM.03CM_510_03.NEW  -A---C----A--G--------A--C-AG---...--C--TG----GACA-----T--C--T----...AAA--A---GAA--CTC---C--A-----GG-A-A---------T--C-----------A--T--------C--C----C------C--C-----------G 7003
H2G.CI.92.ABT96.AF208027 -A---G-G-----G--C-----C----AA---...--A--TG-C--AA-G--M--A-CC--T--A...A-A--A---GAA--TTC------A-----A-AG-----------T-----------------A---T-G--T----------C-----------------A----G 7255
H2U.FR.96.12034.AY530889 ---------C---C------------CA---T...--C--G----T-AG-C-G--A---C--G-A...-AA--A-C-...-----T---T-A--G--A-AA--T--------T--C-----------A--------------G----CA-----C--C------------ 7381
SMM.SL.92.SL92B.AF334679 -A-----A--------C----------------...--C--TG-----AGG-----T---C-A---...--C--G---...-----G-----------C--A----------------------------A-----T-------------------C--C-----------G 7275
MAC.US.x.MAC239          --------------------------------...------------------------------...------A--...------------------------------------------------------------------------------------------ 6916
MAC.US.x.251             --------------------------------...------------------------------...---------...--------------------------------------------------T-----------------------T--------------- 7835
MAC.US.x.251_32H         --------------------------------...------------------------------...------A--...------------------------------------------------------------------------------------------ 7837
MAC.US.x.251             --------------------------------...-------------------------A----...---------...-----------------------------------------------------------------------------------------G 7821
MAC.US.x.17EC1.AY033233  --------------------------------...------------------------------...------A--...------------------------------------------------------------------------------------------ 7839
MAC.US.x.17EFR.AY033146  --------------------------------...------------------------------...------A--...------------------------------------------------------------------------------------------ 7839
MAC.US.x.1937.AY611495   --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6945
MAC.US.x.2065.AY611493   --------------------------------...------R-----------------------...------AR-...------------------------------------------------------------------------------------------ 6945
MAC.US.x.MM142.Y00277    --------------------------------...------------------------------...------A--...--------------------------------------------------T--------------------------------------- 7334
MAC.US.x.SMM142B.BD131285--------------------------------...------------------------------...------A--...--------------------------------------------------T--------------------------------------- 7331
MNE.US.82.MNE            --------------------------------...------------------------------...---------...--------------------------------------------------T--------------------------------------- 7307
MNE.US.x.MNE027.U79412   --------------------------------...------------------------------...------C--...--------------------------------------------------T--------------------------------------- 7310
SMM.US.x.F236            -A---T-G-------------------G---...-----TG-----AGG-----T---C-A--A...-----A-C-...-----------A-----C--T--T-------------TC-------A---T-A--T--C-----------------T-----------G 7789
SMM.US.x.H9.M80194       MAM--T-G------------------AG---...-----TR-K-Y-R-AC----T--GC-A--A...-----A---...-----------A-------T-----------------TC-------A---T-A--T--C-----------------T------------ 7316
SMM.US.x.PBJ14           -A---T-G-------------------G---...-----T--G-----AC----T--GC-A--A...-----A---...-----------A-------T-----------------TC-------A---T-A--T--C-----------------T------------ 7619
SMM.US.x.PBJA.M31325     -A---T-G-------------------G---...-----T--G-----AC----T--GC-A--A...-----A---...-----------A-------T-----------------TC-------A---T-A--T--C-----------------T------------ 7618
SMM.US.x.PBJ             MAM--T-G------------------AG---...-----TR-K-Y-R-AC----T--GC-A--A...-----A---...-----------A-------T-----------------TC-------A---T-A--T--C-----------------T------------ 7331
SMM.US.x.PBJ             -A---T-G-------------------G---...-----T--G-----AC----T--GC-A--A...-----A---...-----------A-------T-----------------TC-------A---T-A--T--C-----------------T------------ 7811
SMM.US.x.PGM53.AF077017  -A------------------A-----G---...-----TG-G---A-G-----T---C----A...-----A---...-----------A-----C--T--T-------------T-----------T-----T--C--------------C--T------------ 7773
SMM.US.x.SME543.U72748   -A---T-G-------------------G---...-----TG-----AGG-----T---C-A--A...-----A-C-...-----------A-----C--T--T-------------TC-------A---T-A-----C-----------------C-----------G 7846
STM.US.x.STM.M83293      -A---T-G-------------------A---...--C--TG-C--G-CA-----A-GGA-AGT-...---------G...-----------A--G--C---------------------C-T-----A--T--T--T--C---------------CT--------A----- 7478
MAC.US.x.80035.AY611486  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6945
MAC.US.x.81035.AY599200  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6946
MAC.US.x.85013.AY611490  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6879
MAC.US.x.92050.AY603959  ------------R-------------------...------------------------------...------R--...------------------------------------------------------------------------------------------ 6939
MAC.US.x.92077.AY599201  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6946
MAC.US.x.93057.AY611492  --------------------------------...------G-----------------------...---------...------------------------------------------------------------------------------------------ 6944
MAC.US.x.93062.AY607704  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6931
MAC.US.x.95058.AY611494  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6933
MAC.US.x.95086.AY607703  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6802
MAC.US.x.95112.AY588946  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6935
MAC.US.x.96016.AY607701  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6946
MAC.US.x.96020.AY611488  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6938
MAC.US.x.96072.AY611491  --------------------------------...------G-----------------------...---------...------------------------------------------------------------------------------------------ 6950
MAC.US.x.96081.AY597209  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6948
MAC.US.x.96093.AY611489  --------------------------------...------G-----------------------...---------...------------------------------------------------------------------------------------------ 6946
MAC.US.x.96114.AY588945  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6941
MAC.US.x.96123.AY611487  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6835
MAC.US.x.96135.AY607702  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6931
MAC.US.x.97009.AY599199  --------------------------------...------------------------------...---------...R-----A-----------------------------------------------------------------------------------4 6814
MAC.US.x.97074.AY599198  --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6814
MAC.US.x.r80025.AY576480 --------------------------------...------------------------------...---------...------------------------------------------------------------------------------------------ 6802
MAC.US.x.r90131.AY576481 --------------------------------...------------------------------...---------...R-----A------------------------------------------------------------------------------------ 6945
```

```
                                                
```

```
                                                                                                                                                     Rev Responsive Element (RRE) start
MAC.US.x.239               CAGTCTCATAGCAAACATAGATTGG.........ATTGATGGAAACCAAACTAATATCACCATGAGTGCAGAGGTGGCAGAACTGTATCGATTGGAATTGGGAGATTATAAATTAGTAGAGATCACTCCAATTGGCTTGGCCCCCACAGATGTGAAGAGGTACACTACTG 8155
Env                        __S__L__I__A__N__I__D__W__.__.__.__I__D__G__N__Q__T__N__I__T__M__S__A__E__V__A__E__L__Y__R__L__E__L__G__D__Y__K__L__V__E__I__T__P__I__G__L__A__P__T__D__V__K__R__Y__T__T__
H2A.CI.88.UC2.U38293       ---CA-A--T--T-----T--C............-CA-------T--G--C--C--T---T-T-----------------------C-----A----------C--C-----GA----A--A--A-----------C--A--T---AG--AA--A--A--TT-CT-G- 8167
H2A.DE.x.BEN.M30502        ---CA-A--T--T-----T--CATA......GAT-AAA--C-G-CT--T--C--C--T---T-T--------A-------------C-----A---C-------C--C------A----A--A--A-----------C--A--T------CA--GA------T-CT-AA 8213
H2A.DE.x.PEI2.U22047       ---CA-A--T--T-----T--CAT-............TT--AT-----G--A-GC--T---T-T-------------------A--C---------------------C-----------A--A--A-----------C--A--T---TCA-AA--A-----TT-CT--- 8157
H2A.GH.x.GH1.M30895        ---CA-A--T--T-----T--CGTA............A--A-T-----G--A-----T---T-T-------------------A--C-----A-------------------A----AG-A--A-----------C--A--T--GAGA-A---A--A--TT-CT-G- 7633
H2A.IN.NNVA.NEW            ---CA-A--T--T-----T--AGA-...................--T-----A---G-T---T-T-------G---------C--A--C-----A--------G-----C-------------A--A--A-----------C--A--T---CCA-AA--AC-A---T-CTT-- 8205
H2A.IN.95.CRIK             ---CA-A--T--T-----T---GTA......AACGAAATA-AT--A-GG--A-----T---T-T-------------------A--C---G-A--------G----------G-----AG-A--A-----------C--A--T---TCA-AA--AC-----T-CT--- 7967
H2A.GM.x.MCR35.AY509260    T--CA-A--T--T-----T---GCA............A---A---TA-T--AG----T---T-T-------------------A--C-----A--G----------------G-----A--A--A-----------C--A--T----CA-AA--A--A---T-CT--A 7640
H2A.GM.87.D194.J04542      ---CA-A--T--T-----T--C-CA..........--------T--G--C--C--T---T-T--------A-------------C-----A-------G--C--C-----GA----AG-A--A------CCG--C--A--T---A-A-A---A--A--TT-CT-G- 7627
H2A.GM.x.ISY.J04498        ----A----T--T-----T---GTT.........GA--GA-AT----GG--A-----T---T-T-------------------A--C----------------G-------------------AG-A--A-----------C-----T----CA-AA--A--A---T-CT--- 7633
H2A.GM.x.MCN13.AY509259    T--CA-A--T--T-----T---GCA............A-------TA-T--A-----T---T-T-------------------A--C-----A--G----------------G-----A--A--A-----------C--A--T----CA-AA--A--A---T-CT--A 7640
H2A.GW.86.FG.J03654        ---CA-A--T--C-----T---GC-...............---G-T-----A-----T---T-T---------C---------A--C---------------G-----C------------A-----A-----------C--A--T---TCA--A-----A---T-CT--- 7624
H2A.GW.87.CAM2CG.D00835    ---CA-A--T--T-----T--CGA-..........-GG---AATC-GAC---A-----T---T-T--------T-------------C---------------G-----------G-----A--A--A-----------C--A--T---TCACAA----A---T-CC--- 8225
H2A.GW.x.ALI.AF082339      ---CT-G--T--T-----T--CAC-...................-----C-----G--A-----T---T-T-------------------A--C-----A-------G-------------------------A--A-----------C--A--T---TCA-AA-G---A---T-CT--A 8197
H2A.GW.x.MDS.Z48731        ---CA-A--T--T-----T--CGAC............A-G-AT--T--G--A-----T---T-T--------T----------A--C-----A-------------------------A--A--A--A-----------C--A--T---CCA-AA--AC-A---T-CT--- 7690
H2A.SN.85.ROD.M15390       ---CA-A--T--T-----T--C---..........CAAA-CAAT--T--G--A--C--T---T-T-------------------A--CA------G---------------G-----A--A--A-----------C--A--T---A-A-AA--A--A---T-CT--- 7659
H2AB.CI.90.7312A.L36874    ---CT-G--T--T-----T--CGTA......GACG-A-G-AAT----GG--A-----T---T-T-----------------------C-----A---------------------A----AG-G--A--------T--C--A--T---TCA-AA--A--A---T-CT--A 8204
H2B.CI.88.UC1.L07625       ------T-----C-----T--CGT-TATTATGATGGCA---AT-C-A-G--C-----T-------------A----G---------CA-------C---G--C--C-----------A--A--A--G--------T--A--T-----GA-A--A--A--TT-CT-AA 8200
H2B.CI.x.EHO.U27200        ---C---------C-----T--C---..........--A---AAG----TT--------T--TG----------A---T-----------AA-------C---G--C--C-----------------A--A-------------T--A--T---AG-A-A--A--A--TT-CT-A- 8175
H2B.GH.86.D205.X61240      T--C---------C-----CA-C-CA.............--CAAC-GTACC--C-----TT-TG----------A---T----------A-------C---G--C-------------A--A--A--A-----------C--A--T--------A-GA--A---T-CT-A- 8189
H2B.JP.01.KR020.AB100245   ----A-------C--T--T-----..........-CAA--AAT--TAGG--------T------------A---T----------A-------C---G--C--C---C------A--A--A--A-----------T--A--T---AG---A--A-----TT-CT-A- 7306
H2AB.CM.03CM_510_03.NEW    ---CA-A--T--T-----T--CACA............A-A-----T--G--C--C--T---T-C-----------------------C-----A--G---------C--C-----GA----A--A--A-----------C--A--T---AC--AAGG------TT-CT-AA 7292
H2G.CI.92.ABT96.AF208027   ---C--A-----G--T-----AGA-......AAT-GGTC-AAC-----T--A-------TTT-C---------------------------A---------G---------------A--A-------A--A--------------T--A------A----A-----A--TT-CT--- 7574
H2U.FR.96.12034.AY530889   T--CA-G--T-----T--T---C--.............---A-CAA---T-----A------T-------C-----A-------------A---A--G-----G-----------G-----A--A-GA-----------T--T------CCA-----A--A--TT-CT-A- 7697
SMM.SL.92.SL92B.AF334679   ---C--A--------T--T------..............-----CAAC--TG-G--C-----T-------------A--------T-----------------G-----------G-------------------A----T------C-------A--------C--AA 7606
MAC.US.x.MAC239            -------------------------.........---------------------------------------------------------------------------------------------------------------R-------------------- 7232
MAC.US.x.251               -------------------------.........-C-------------------------------------------------------------------------------------------------------------A------G-------------- 8151
MAC.US.x.251_32H           -------------------------.........-C---------------------------------------------------------------------------------------------------------------------------------- 8153
MAC.US.x.251               -------------------------.........-C----------------G---------------------------------G-----------------------------G----------------------------------------------- 8137
MAC.US.x.17EC1.AY033233    -------------------------.........----------------------------------------------------------------------------------------------------------------------------------- 8155
MAC.US.x.17EFR.AY033146    -------------------------.........----------------------------------------------------------------------------------------------------------------------------------- 8155
MAC.US.x.1937.AY611495     -------------------------.........---------------------------------------------------------------------------------------------------------------AR------------------- 7263
MAC.US.x.2065.AY611493     -------------------------.........---C---A---------------------------------------------------------------------------------------------------------G------------------- 7261
MAC.US.x.MM142.Y00277      -----------------A-------.........-C----------------G------------------------------------------------------A------A----------------------------A-------------------- 7650
MAC.US.x.SMM142B.BD131285- -----------------A-------.........-C----------------G------------------------------------------------------A------A----------------------------A-------------------- 7647
MNE.US.82.MNE              -------------------------.........---------------------------------------------------------------------------------------------------------------A-------------------- 7629
MNE.US.x.MNE027.U79412     -------------------------.........---------------------------------------------------------------------------------------------------------------A-------------------- 7632
SMM.US.x.F236              T------------G-G---------.........--CA--A-C--TG-G--C-------------------------------G----------C-----A------T--------------------------AG---A-GA--------C--AA 8120
SMM.US.x.H9.M80194         T--C---------G-A---------.........--CA---C--TA-G--C-------------------------------G----------C------------T--------------C---------AG---A--A--------C--AA 7644
SMM.US.x.PBJ14             T--C---------G-A---------.........--CA---C--TG-G--C-------------------------------G----------C------------T----------C---C---------AG---A--A--------C--AA 7947
SMM.US.x.PBJA.M31325       T--C---------G-A---------.........--CA---C--TG-G--C-------------------------------G----------C------------T----------C---C---------AG---A--A--------C--AA 7946
SMM.US.x.PBJ               T--C---------G-A---------.........--CA---C--TA-G--C-------------------------------G----------C------------T--------------C---------AG---A--A--------C--AA 7659
SMM.US.x.PBJ               T--C---------G-A---------.........--CA---C--TG-G--C-------------------------------G----------C------------T--------------C---------AG---A--A--------C--AA 8139
SMM.US.x.PGM53.AF077017    T--C---------G-T-----C---.........--------C-----G----C---------------------CA-------G-------------------------T-----------------------AG---A-G---------A---. 8103
SMM.US.x.SME543.U72748     T------------G-G---------.........--CA--AAC--TG-G--C-------------------------------G----------C------------T-------------------------------A-GA--------C--AA 8177
STM.US.x.STM.M83293        ----A-A----------T--C---..........-C-A-CAAT--TG-G-----C-----TGCA--------------------------------C----------A--A------A--A------------------A--GA----------A-----A---A 7809
MAC.US.x.80035.AY611486    -------------------------.........----------------------------------------------------------------------------------------------------------------R------------------- 7261
MAC.US.x.81035.AY599200    -------------------------.........---------------------------------------------------------------------------------------------------------------A-------------------- 7262
MAC.US.x.85013.AY611490    -------------------------.........---------------------------------------------------------------------------------------------------------------RR-----------R------ 7195
MAC.US.x.92050.AY603959    -------------------------.........----------------------------------------------------------------------------------------------------------------------------------- 7255
MAC.US.x.92077.AY599201    -------------------------.........---------------------------------------------------------------------------------------------------------------R-------------------- 7262
MAC.US.x.93057.AY611492    -------------------------.........---------------------------------------------------------------------------------------------------------------C-------------------- 7260
MAC.US.x.93062.AY607704    -------------------------.........----------------------------------------------------------------------------------------------------------------G------------------- 7229
MAC.US.x.95058.AY611494    -------------------------.........---------------------------------------------------------------------------------------------------------------AR------------------- 7249
MAC.US.x.95086.AY607703    -------------------------.........---------------------------------------------------------------------------------------------------------------A-------------------- 7118
MAC.US.x.95112.AY588946    -------------------------.........----------------------------------------------------------------------------------------------------------------------------------- 7251
MAC.US.x.96016.AY607701    -------------------------.........---------------------------------------------------------------------------------------------------------------CR-----------R------ 7262
MAC.US.x.96020.AY611488    -------------------------.........----------------------------------------------------------------------------------------------------------------------------------- 7254
MAC.US.x.96072.AY611491    -------------------------.........---------------------------------------------------------------------------------------------------------------C------------G------ 7266
MAC.US.x.96081.AY597209    -------------------------.........----------------------------------------------------------------------------------------------------------------------------------- 7264
MAC.US.x.96093.AY611489    -------------------------.........---------------------------------------------------------------------------------------------------------------AR------------------- 7262
MAC.US.x.96114.AY588945    -------------------------.........----------------------------------------------------------------------------------------------------------------------------------- 7257
MAC.US.x.96123.AY611487    -------------------------.........----------------------------------------------------------------------------------------------------------------G------------------- 7151
MAC.US.x.96135.AY607702    -------------------------.........----------------------------------------------------------------------------------------------------------------G------------------- 7247
MAC.US.x.97009.AY599199    -------------------------.........---------------------------------------------------------------------------------------------------------------R-------------------- 7130
MAC.US.x.97074.AY599198    -------------------------.........----------------------------------------------------------------------------------------------------------------G------------------- 7124
MAC.US.x.r80025.AY576480   -------------------------.........---------------------------------------------------------------------------------------------------------------A-------------------- 7118
MAC.US.x.r90131.AY576481   -------------------------.........----------------------------------------------------------------------------------------------------------------------------------- 7261
```


HIV-2/SIV Complete Genomes

```
                                   Env gp120 end Env gp41 start
MAC.US.x.239                       GTGGCACCTCAAGAAATAAAAGAGGGGTCTTTGTGCTAGGGTTCTTGGGTTTTCTCGCAACGGCAGGTTCTGCAATGGGCGCGGCGTCGTTGACGCTGACCGCTCAGTCCCGAACTTTATTGGCTGGGATAGTGCAGCAACAGCAACAGCTGTTGGACGTGGTCAAGAGA 8325
Env                                G  G  T  S  R  N  K  R  G  V  F  V  L  G  F  L  G  F  L  A  T  A  G  S  A  M  G  A  A  S  L  T  L  T  A  Q  S  R  T  L  L  A  G  I  V  Q  Q  Q  Q  Q  L  L  D  V  V  K  R
H2A.CI.88.UC2.U38293               CC...C-AG-G--G--C-------T--G--C---------C----------------G--A-----------------------CC--------T-A----------G-----C----C-----------------------------A-A--------- 8334
H2A.DE.x.BEN.M30502                C---.C-AGTG--G--C-------T--G--C---------------------------G--A---------------------CG---CC--------T-A--C--------G-----C----C-----------------------------A--------- 8380
H2A.DE.x.PEI2.U22047               C---.C-ACAG--G----------T--G----------AG-----------------A-----------------------C--------T-G-----TC--G--CT--AC-G------------------------------------------ 8324
H2A.GH.x.GH1.M30895                C---.C-AGTG--G--C-------T--G----------------------------G--A-----------------------CC--------T-G-------G--G-----C----C-----------------------------A--------- 7800
H2A.IN.NNVA.NEW                    CC...T--AGG---C-A-C-----T--G--C-------------------------C--A----A-------------------C--------T--------C----G-----C----C-----------------------------------G 8372
H2A.IN.95.CRIK                     -C.----GGG---C---C------T--GC-C-------AA----------------A-------------------A-----C--------T--------C----G-----C----C------------------------------------ 8135
H2A.GM.x.MCR35.AY509260            C---.C-AATG--G--C--G-----T--G--C-------------------------A-----C-------------------C--A-----T-G-------T--G-----C----C------------------------------------ 7807
H2A.GM.87.D194.J04542              C---.C-AGTG--G--C-------T--G--C-------------------------G--A-------------------GC-----C--------T-G---------G-----C----C------------------------------------ 7794
H2A.GM.x.ISY.J04498                C---.C-AGGG---C---G-----T--GC-------------C-A--------A-G--A-----G---------G--------TC--------T-G-------T--G--------CCG-------------------------------------- 7800
H2A.GM.x.MCN13.AY509259            C---.C-AATG--G--C--G-----T--G--C-------------------------A-----C-------------------C--A-----T-G-------T--G-----C----C------------------------------------ 7807
H2A.GW.86.FG.J03654                C---.CA-CAG---C---C------T--G--------------C-A-------------A----------C-------------C--------T-A-------T--G----------C-------------------------------T----------- 7791
H2A.GW.87.CAM2CG.D00835            C---.CA-GGG---CC--------T-CG-------------------------A----A-----AGT---------A------C--------T-A--C----T--G----------C-------------------------A--------------- 8392
H2A.GW.x.ALI.AF082339              C---.C-AAGG--G----------T--G--C-------------A----------G--A-----------------A---AG-T--A-----T-T-------T--G----------C-------------------------------------- 8364
H2A.GW.x.MDS.Z48731                C---.CA-GGG---C----------T--G-------------------------A----A-----G-----------------C--------T---------T--G-----C----C-------------------------------------- 7857
H2A.SN.85.ROD.M15390               C---.CA-GGG---C---C------T--G--C-------------------------A-------------------------CC----CG--T-G----------G-----C----C-------------------------------------- 7826
H2AB.CI.90.7312A.L36874            C---.C-GGGG---C----------T--A--C-----------------------A-G--A-----AG------------------C--------T-G-------T--G-----C----C-----------------------------A--------- 8371
H2B.CI.88.UC1.L07625               CG...--AC-G--G----------T--AA-G----------------AC---T---T-------------------AA---C--------T-A----------G-----C-----------------G-----------C------------A--- 8367
H2B.CI.x.EHO.U27200                TG...--AC-G--G----------T--AC------------T-----A--C--T--G-----------------------C-----T--T-G----------G-----C-----------------G---------CG------------A--- 8342
H2B.GH.86.D205.X61240              TA...-AAC-G--G----------T--AA-G--------A-------A--C--T---T-------------------AA---T--------T-A----------G-----C----C----------G---------CTG------------A--- 8356
H2B.JP.01.KR020.AB100245           CG...--AC-G--G-----G-----T--ACCG----------T-----A-----T-----------------------------C-----T--T-G--C-------G-----------------G-----------C------------A--- 7473
H2AB.CM.03CM_510_03.NEW            C---.C-AGGG--G----------T--A----------A-------A--C--T--G-----------------------A-----C--------T-A----------G-----C-----------------G-----------C------------A--- 7459
H2G.CI.92.ABT96.AF208027           TG...--AC-G-A-----------T--A-------------T-----A-----T---T------------------A-A--------------T-G----------G-----G-----------------------------------T------------ 7741
H2U.FR.96.12034.AY530889           TG...--AC-G--G--------G--A--A--------T-------------------C-------------------T-----------------T-A----------G-----C-----------------------------------A--- 7864
SMM.SL.92.SL92B.AF334679           ----.-------AG----------A-----------------------------G-----------------------A-----C-A-------T----------G-----------------------A-------------------------- 7773
MAC.US.x.MAC239                    ------------------------------------------------------------------------------------------------------------------------------------------------------------ 7402
MAC.US.x.251                       ------------------------------------------------------------------------------------------------------------------------------------G-----------------------T---------- 8321
MAC.US.x.251_32H                   ------------------------------------------------------------------------------------------------------------------A----G-----------------------------------------------------------------------------------------------------------------------
```

```
                                                
```

```
MAC.US.x.239              TCTAACACCAAAGTGGAACAATGAGACTTGGCAAGAGTGGGAGCGAAAGGTTGACTTCTTGGAAGAAAATATAACAGCCCTCCTAGAGGAGGCACAAATTCAACAAGAGAAGAACATGTATGAATTACAAAAGTTGAATAGCTGGGATGTGTTTGGCAATTGGTTTGACC 8653
Env                       __L__T__P__K__W__N__N__E__T__W__Q__E__W__E__R__K__V__D__F__L__E__E__N__I__T__A__L__L__E__E__A__Q__I__Q__Q__E__K__N__M__Y__E__L__Q__K__L__N__S__W__D__V__F__G__N__W__F__D__
H2A.CI.88.UC2.U38293      -T-------T-GA---------AT---A-----G--A-----AAA-C-A--CCG--A-C-A--G-C------C-GTCAAAG-T----A--------------------A-----T-----------------A--A--------------C--------C--------TT 8662
H2A.DE.x.BEN.M30502       -T--T-G--TG-C-----A---AT---A-----G---------AA-C-A--CCG--A-C-A--G-C------C-GTCAAAGTT----A--A--C--------------A-----T-----------------A--A------------A-TC-------C---------T 8708
H2A.DE.x.PEI2.U22047      CA--GT---TG-T--------CAT---G-----G--A-----A-A-C-AACCCG-GA-C-A--G-C------C-GTAGATCGT-----C----------C-----------A--T-------GC--------A--A--------------T--------C--------TT 8652
H2A.GH.x.GH1.M30895       -T--T-G--TG-C---------AT---A-----G--A-----AAA-C-A--CCG--A-C-A--G-C------C-GTCAAAGTT----AC-------------------A-----T------------------C-A--------------C--------C--------TT 8128
H2A.IN.NNVA.NEW           CT--GT---T--T-----A---AT---G--CA----A-----A-A------CCG-CA-C----G-C------C-GTCAAAGTT----AC----------C--G--G--A-----T--------C--------A--A------------A-T--------C---------T 8700
H2A.IN.95.CRIK            CT-G-----TG-T-----A---AT---A---T----A-----A-A---A--CCG--A-C----G-C------C-GTCAAAGTT----AC----------C-----G--A----G--------GC--T-----A--A--------------T--------C---------T 8463
H2A.GM.x.MCR35.AY509260   CT----G--TG-----------AT---G--------A-----AG-C--AA-CCG-GA-C----G-C------C-GTCAA-AAT----AC-A-----------G-----------T--------C-----------A--------------T---A-T--C---------T 8135
H2A.GM.87.D194.J04542     CT-------TG-C---------AT---A-----G--A-----AAA-CGA--CC---A-C-A--G-C------C-GTCAAAGTT----AC-------------------A-----T--------C--------AC-A--------------C--------C--------TT 8122
H2A.GM.x.ISY.J04498       CT-------TG----------CAT---A--------A-----A-AC--AA-CCG----C-A--G-C------C-GT-AGAGTT----AC----------C--G-----A-----T-------GC--G------C-A--------------T------------------T 8128
H2A.GM.x.MCN13.AY509259   CT----G--TG-----------AT---G--------A-----AG-C--AA-CCG-GA-C----G-C------C-GTCAA-AAT----AC-A-----------G-----------T--------C-----------A--------------T-----T--C---------T 8135
H2A.GW.86.FG.J03654       CT-G-----TG-T---------AT---G-----G--A-----A-A---A--CCG--A-C----G-C------C-GTCAAAGT-----AC-A-----------G-----A-----T-------GC--------A--A--------------T---AC-------C-----T 8119
H2A.GW.87.CAM2CG.D00835   -T-------TG-T---------AT---A-----G--A-----A-A------CCG--A-C----G-C------C-GTCAAAGTT----A---------C----G-----A-----T--------C--------A--A--C-A---------T---AC---C---------T 8720
H2A.GW.x.ALI.AF082339     CT---A---TG-T---G----CAT---G--------------A-A-C-A--CCGT-A-C-A--G-C------C-GT-AA-AGT----ACG------------G-----A-----T-C------C--------A--A--------------T---AC---C---C-----T 8692
H2A.GW.x.MDS.Z48731       CT--GA---TG-T---G-A---AT---G-----G--A-----A-A------CCG--A-C----G-C------C-GTCAAAAGT----A----------C--G--G--AC------------GC--------A--A--C---------A-T------------------T 8185
H2A.SN.85.ROD.M15390      CT--G----TG-C---G-----AT---G-----G--A-----AAA-C-A--CCG--A-C----G-C------C-GTAAAAGTT----AC-------------G--------A--T--------C--------A--A------------A-T------------------T 8154

H2AB.CI.90.7312A.L36874   CT-G-----TG-T---G----CAT---G------C-A-----AAA-C-AA-CCG-GA-C----G-C------C-GT-AAAGT-----AC----------C--G-----A-----------------------A--A--------------T--------C--------TT 8699

H2B.CI.88.UC1.L07625      ---C------G-C---G-A--CAT---A------C-------AAAGCGA--CA-T--------T-C---------------GT----A-----T--G--A--------A-G------------------G--A--A--------------T-----T------------T 8695
H2B.CI.x.EHO.U27200       C--T-AG---G-C-----T--CAT---A------C-A------A-GC-A--CCG---------T-C----------AAAT-A-----A--------G--A--------A-------------G------G--A--A---CAA------A-T--CA-T------------T 8670
H2B.GH.86.D205.X61240     C--C--------T-----T--CAT---A------C-------AAAGC-A--CC-T--------T-C---------G--T--GT----A-----T--G--A-----G--A--------------------G--AA-A--------------T-----T------------- 8684
H2B.JP.01.KR020.AB100245  -T-C--T----C----G----CAT---A-----GC--------AACC----CCG---------T------------AAAT-AT----A-T------G--A--------AG---------CA-G------G--AC-A--CCAA--------T---A-T------------T 7801

H2AB.CM.03CM_510_03.NEW   CT----G--T--C---G-----AT---A-----GC-A------GA------CCG-GA-C-A--G-C------C-GTCAAA-TT----AC----G--------------A-----T-----------------AC-A--------------C---A----C--------TT 7787

H2G.CI.92.ABT96.AF208027  AT-GGAG---C-A-----T---AT---A--------A-----AAAGC--A-CA----T-----G--C--C--C---CGG--AT----A--A-----------G-----A--A--T-------G------G--A--A--------------A------------------- 8081

H2U.FR.96.12034.AY530889  AT-------T--T---G-T---AT---G--------A------GA------GA--A----------------C---CAG--AT----A-CA--T-----A-----------A-----------------G--A--A----A-------A-C-----T--C--------T- 8192

SMM.SL.92.SL92B.AF334679  ---TGTC---G-C---------AT---A--------A------AA---------A---C-A--G-C---------TCAAA-GT-G-----A----G-C-A--G--------A-----------------G--A-----------------------A------------- 8101

MAC.US.x.MAC239           ---------------------------------------------------------------R---------------------------------------------------------------------------------------------------------- 7730
MAC.US.x.251              ---------------------------------------------------------------G-----------------------------------------------------------------------A---------------------------------- 8649
MAC.US.x.251_32H          ----------G-C-----------T--------------------------------------G-----------------------A---------------------------------------------------------------------------------- 8651
MAC.US.x.251              ----------G-C-----------T--------------------------------------G-----------------------A---------------------------------------------------------------------------------- 8635
MAC.US.x.17EC1.AY033233   -----T---------------------------------------------------------------------------------------------------------------------------------------------A------------------------ 8653
MAC.US.x.17EFR.AY033146   -----T---------------------------------------------------------------------------------------------------------------------------------------------A------------------------ 8653
MAC.US.x.1937.AY611495    -M---------G----G----------------------------------------------G---------------------------------------------------------------------------------------------------------- 7761
MAC.US.x.2065.AY611493    ----------------G----------------------------------------------G---------------------------------------------------------------------------------------------------------- 7759
MAC.US.x.MM142.Y00277     ----------G-T--------------------------------G-----------------G-C---------G-----------A---------------------------------------------------------------------------------- 8148
MAC.US.x.SMM142B.BD131285 ----------G-T--------------------------------G-----------------G-C---------G-----------A---------------------------------------------------------------------------------- 8145

MNE.US.82.MNE             ------------T--------------------------------------------------G-----------G-----TT----A--------------------A------------------------------------------------------------- 8127
MNE.US.x.MNE027.U79412    ------------T--------------------------------------------------G-----------G------T----A--------------------A------------------------------------------------------------- 8130

SMM.US.x.F236             AT-GGTG--T--T---------AT------------------AA--C-----------C-A--G-C---------TCAAT-AT----A--A-----------G-----A-----T-----------G-----A--A------------A-C------------------- 8618
SMM.US.x.H9.M80194        AT-G-----T-RC---------AT--Y---------------AAA-C----GA-----C-A--G-C---------TSAATYAT-G--A--A-----------G-----AM----T-----------G-----A--A----SK------N-T--Y---------K------ 8142
SMM.US.x.PBJ14            AT-G-----T--C---------AT------------------AAA-C----GA-----C-A--G-C---------TCAATCAT-G--A--A-----------G-----A-----T-C---------G-----A--A------------A-T------------------- 8445
SMM.US.x.PBJA.M31325      AT-G-----T--C---------AT------------------AAA-C----GA-----C-A--G-C---------TCAATCAT-G--A--A-----------G-----A-----T-C---------G-----A--A------------A-T------------------- 8444
SMM.US.x.PBJ              AT-G-----T-RC---------AT--Y---------------AAA-C----GA-----C-A--G-C---------TSAATYAT-G--A--A-----------G-----AM----T-----------G-----A--A----SK------N-T--Y---------K------ 8157
SMM.US.x.PBJ              AT-G-----T--C---------AT------------------AAA-C----GA-----C-A--G-C---------TCAATCAT-G--A--A-----------G-----A-----T-C---------G-----A--A------------A-T------------------- 8637
SMM.US.x.PGM53.AF077017   AT-GGTG--T--T---------AT------------A-----AA--C--------GA-C-A--G-C---------TCAGGCAT-G--A--A-----------G-----A-----T-----------G-----A--A------------A-T------------------- 8599
SMM.US.x.SME543.U72748    AT-GGTG--T--T---G-----AT------------------AG--------------C-A--G-C---------TCAAT-AT----A--A-----------G-----A-----T-----------G-----AC-A------------A-C------------------- 8675

STM.US.x.STM.M83293       -T-GGT---GG-T---------AT---A-----G--A------A-------------TC-T--G-C----------CAA--GT----A--A--T---G-------G--A-----T-------G---G-----AC-A--------------A------------------- 8304

MAC.US.x.80035.AY611486   ---------------------------------------------------------------R---------------------------------------------------------------------------------------------------------- 7759
MAC.US.x.81035.AY599200   ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 7760
MAC.US.x.85013.AY611490   ----------R----------------------------------------------------R---------------------------------------------------------------------------------------------------------- 7693
MAC.US.x.92050.AY603959   ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 7753
MAC.US.x.92077.AY599201   -----T---------------------------------------------------------------------------------------------------------------------------------------------------------------------- 7760
MAC.US.x.93057.AY611492   ----------------R----------------------------------------------R---------------------------------------------------------------------------------------------------------- 7758
MAC.US.x.93062.AY607704   ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 7727
MAC.US.x.95058.AY611494   -----Y----------G----------------------------------------------G---------------------------------------------------------------------------------------------------------- 7747
MAC.US.x.95086.AY607703   -----Y---------------------------------------------------------------------------------------------------------------------------------------------------------------------- 7616
MAC.US.x.95112.AY588946   ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 7749
MAC.US.x.96016.AY607701   ----------R-----R----------------------------------------------G---------------------------------------------------------------------------------------------------------- 7760
MAC.US.x.96020.AY611488   ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 7752
MAC.US.x.96072.AY611491   ----------------G----------------------------------------------G---------------------------------------------------------------------------------------------------------- 7764
MAC.US.x.96081.AY597209   ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 7762
MAC.US.x.96093.AY611489   ----------G----------------------------------------------------G---------------------------------------------------------------------------------------------------------- 7760
MAC.US.x.96114.AY588945   ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 7755
MAC.US.x.96123.AY611487   -----Y---------------------------------------------------------R---------------------------------------------------------------------------------------------------------- 7649
MAC.US.x.96135.AY607702   ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 7745
MAC.US.x.97009.AY599199   -----------------G---------------------------------------------------------------------------------------------------------------------------------------------------------- 7628
MAC.US.x.97074.AY599198   ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 7622
MAC.US.x.r80025.AY576480  -----Y---------------------------------------------------------------------------------------------------------------------------------------------------------------------- 7616
MAC.US.x.r90131.AY576481  ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 7759
```

```
                                                
```

```
                                                                                    Tat end
MAC.US.x.239                   CCGGCACTGCCAACCAGAGAAGGCAAAGAAAGAGACGGTGGAGAAGGCGGTGGCAACAGCTCCTGGCCTTGGCAGATAGAATATATTCATTTCCTGATCCGCCAACTGATACGCCTCTTGACTTGGCTATTCAGCAACTGCAGAACCTTGCTATCGAGAGTATACCAGAT 8993
Env                            _P__A__L__P__T__R__E__G__K__E__R__D__G__G__E__G__G__G__N__S__S__W__P__W__Q__I__E__Y__I__H__F__L__I__R__Q__L__I__R__L__L__T__W__L__F__S__N__C__R__T__L__L__S__R__V__Y__Q__I_
Tat exon 2                     __R__H__C__Q__P__E__K__A__K__K__E__T__V__E__K__A__V__A__T__A__P__G__L__G__R__*
Rev exon 2                     P__G__T__A__N__Q__R__R__Q__R__K__R__R__W__R__R__R__W__Q__Q__L__L__A__L__A__D__R__I__Y__S__F__P__D__P__P__T__D__T__P__L__D__L__A__I__Q__Q__L__Q__N__L__A__I__E__S__I__P__D_
H2A.CI.88.UC2.U38293           -G--G--A---GG---AC----AA-C----G------CC-----C-C-A-----T--GA--T-----------C----A-C--C--A--GC---------A--TG----CT--T---------G------A-----T------GGA---A-----CGCGAAC-C--C--C 9002
H2A.DE.x.BEN.M30502            -G--G--A----G---AC-----A-C----GA-----TC----GC-A-A----TT--GA--TG----------CA---A-C---G-G--G----------A--T-----CT----------TCG------A--A--T-------GA---A-----C-AGAAC-C--C--C 9048
H2A.DE.x.PEI2.U22047           TG--A--A----GA--------AA-C---CGA-----CC---A-CA--AT---AG------G-----------CA-----C-----A-----------------G-----T-----------CG------A-----GT----AGGA---A-----C--GAGC-T--C--C 8992
H2A.GH.x.GH1.M30895            -G--G--A----G---AC-----A-C----GA-----ACA----C-A--A---TT--GA--TG-A--------C----A-C-----A-------------A--T-----CT-----------CG------A--AG-T-------GA---A-----C-CGAAC-C--C--C 8468
H2A.IN.NNVA.NEW                -A--A--A---GA----------A-C----GA-----T-------CAA-----AT----A--T-----C----C-----C------A-------------T--TG-----T-----G-----CA-A----A--A--G-----AGGA---A-----C--GAGC-T--T--C 9040
H2A.IN.95.CRIK                 -T--A--A---G-----------A-T----GA-----T-------CAA-----TT----A--T-A---C----C-----C------A-------------A--TG--A--T-----G-----CA-A----A--A--G-----AGGA---A-----C--GAGC-T--T--C 8803
H2A.GM.x.MCR35.AY509260        TG--A--A----G---------AA-C----GA-----T-----A-CAA--T--AG------TG-A--------C----AG------A-------------A--G------T-----------CG-A----A--A--T------G-A---A-----C--GA-C-C--T--C 8475
H2A.GM.87.D194.J04542          AG--G--A----G---AC----AA-C----GA-----CC-----C-A-A----TTT-G---TG----------CAC--A-C--C--A--A----------A--T-----CT-----------CG------A--A--G------GGG---A-----C-AGAAC-C--C--C 8462
H2A.GM.x.ISY.J04498            TG--A--A----GA--------AA-C----GA-----T---GA-C-A--T---A-G--A--------------C------------A----------------TG----C------------CA-A----A--A--G------GGA---A-----C---C-C----T--- 8468
H2A.GM.x.MCN13.AY509259        TG--A--A----G---------AA-C----GA-----T-----A-CAA--T--AG------G-----------C----AG------A-------------A--G------T---------G-CG-A----A--A--T------G-A---A-----C--GA-C-C--T--C 8475
H2A.GW.86.FG.J03654            -A--A--A----G---------AA-C----GA-----T-----AGCAA-----AG---A--T-----------C-----C------A----------------TG----T------------CG------A--A--T------GGA---A-----C--GA-C-C--C--- 8459
H2A.GW.87.CAM2CG.D00835        ----A--A----G----C----AA-C----GA-----T------GCAA-----AG---A---------------------------A-----------------G----C------------CG-A----A--A--T------G-A---A-----C--GA-C-T-----C 9060
H2A.GW.x.ALI.AF082339          -A--A--A-----------G--AA-C----GA-----T-------CAA--T--GG----A-TG-----C----C---C-C----T-A-------------A--TG--AGCT----------TCG----G-A-----T------GGA---A-----C--GA-C-C--C--- 9032
H2A.GW.x.MDS.Z48731            -G--A--A----GA--------AA-C----GA-----T------GCAA-----AG---A--T------C----C-----C------A----------------TG----T------------CA-A--G-A--A--G------GGA---A-----C--GA-C-T--T--C 8525
H2A.SN.85.ROD.M15390           -G--G--A----G---AC----AA-C----GA-----------AGCAA-----AG----A-A------C----C-----C------A-----------------G-----T-----------CA-A----A-----T------GGA---A-----C--GAGC-T--T--C 8494
H2AB.CI.90.7312A.L36874        -A--A--A----G---AC----AA-C----GA--GA------A-C-A---G---T----A--T-----C--------C-----C--C--C-----A--T-----G----GGAA--------T--------A-GA-GG------------A--G-T--AGACC-T---A-C 9039
H2B.CI.88.UC1.L07625           -G--G--A----G---AC----AA-C----GAC--A------A-C-AA--G-CTT----A--T-----C--------------CGC---C--T-----T----------GGAA--------T------G-A--A-GG------G-A---A--G-T--AGACC-C---A-- 9035
H2B.CI.x.EHO.U27200            -G--G--A----G---AC----AA-C----GA--GA------A-CAA--AG---T----A--T-----C--------C-----C--C--C----C---T----------GGGA--------T--------A----GG------------AT----A-AGACC-T---A-C 9010
H2B.GH.86.D205.X61240          -G--G--A----G---AC----AA-C----GA--GA------A-C-A---G-A-T----A--------C--------------C--------T--AC-G-----G----GGAA------A-T--------A--A-GG------------A--G-T--AGACC-T---A-- 9024
H2B.JP.01.KR020.AB100245       -A--G--G----G----C----AA-C----GA--GA--------C-A--AG-A-T----A------------------CAG--C--C--C--------T----------GGAA------A-T--------A--A--G---------G--A--G-T--AGA-C-T---ACC 8141
H2AB.CM.03CM_510_03.NEW        -A--G--A---------C----AA-C----GA--GA------A-C-A---G-A-T----A--T-----C--------------CG-------T--C-G------G----GGAA------A-T--------A--A-GG------------A--G-T--AGACC-C---A-- 8127
H2G.CI.92.ABT96.AF208027       -A--A--W--------A-----AA-T----G----AA-----A-CAAA------T----A--T---------------------------------M-------G--AGGGAA-----------------A-------------G----A-----C--GA-CYT-----C 8421
H2U.FR.96.12034.AY530889       -A--A--A--------A--G--AA-T----GA--G--C------CA--------T----A--T---------------AG------A-------------A---------T-----G-------------A-----G------GGA---------C--GA-T-G-----G 8532
SMM.SL.92.SL92B.AF334679       -A--A-------G---A------A-C----G----A------A-C--A-----AT----A-TG-----------------------------------------G--------GA-A-------------A--A---T-TG-CTCG---AGC---C----CC-------A 8441
MAC.US.x.MAC239                ------------------------------G---------------------------------------------------------------------------------------------------------------------------Y--------------- 8070
MAC.US.x.251                   -A----------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 8989
MAC.US.x.251_32H               ------------------------------G---------------------------------------------------------------------------------------------------------------------------------C--------- 8991
MAC.US.x.251                   ------------------------------G---------------------------------------------------------------------------------------------------------------------------------C--------- 8975
MAC.US.x.17EC1.AY033233        -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 8993
MAC.US.x.17EFR.AY033146        ------------------------------G----------------------------------------A-------------------------------------------------------------------------------------------------- 8993
MAC.US.x.1937.AY611495         ------------------------------G------------------------------------------------------------------------------------------------------------------R-------------------------- 8101
MAC.US.x.2065.AY611493         ------------------------------G----------------------------M-------------------------------------------------------------------------------------------------------------- 8099
MAC.US.x.MM142.Y00277          -----T----------A----------A--G------------GCA------------------------------------------------------------------------------------------------------------------C--------- 8488
MAC.US.x.SMM142B.BD131285      -----T----------A----------A--G------------GCA------------------------------------------------------------------------------------------------------------------C--------- 8485
MNE.US.82.MNE                  -A--------------A------A-C----G------------GCA-------------------------------------------------------------A------------------T---------------------------------C--------- 8467
MNE.US.x.MNE027.U79412         -AA-------------A------A-C----G-------------CA-------------T--------------------G-------------------------------------------------------------------------------C--------- 8470
SMM.US.x.F236                  -A--A--C--------A-----AG------G-------------CA-A-------G---A--T-----------------------------------------GT------------------------------G------GGAT-G-----T---GAGC-------- 8958
SMM.US.x.H9.M80194             -A--A-----------A-----AG------G------------RCA-A-----------A--T-------R---------------------------------G-------------------------------GT-----GGAT-G-----T---GANC-G---A-- 8482
SMM.US.x.PBJ14                 -A--A-----------A-----AG------G------------GCA-A-----------A--T-----------------------------------------GT------------------------------GT------GAT-G-----T---GAAC-G---A-C 8785
SMM.US.x.PBJA.M31325           -A--A-----------A-----AG------G------------GCA-A-----------A--T-----------------------------------------GT------------------------------GT------GAT-G-----T---GAAC-G---A-C 8784
SMM.US.x.PBJ                   -A--A-----------A-----AG------G------------RCA-A-----------A--T-------R---------------------------------G-------------------------------GT-----GGAT-G-----T---GANC-G---A-- 8497
SMM.US.x.PBJ                   -A--A-----------A-----AG------G------------GCA-A-----------A--T-----------------------------------------GT------------------------------GT------GAT-G-----T---GAAC-G---A-C 8977
SMM.US.x.PGM53.AF077017        -A--A--A--------A-----AG------G-------------CA-A-----------A--T-----------------------------------------G-----T-----------------G----A---------G---AT---C-T----A---------- 8939
SMM.US.x.SME543.U72748         -A--A--C--------A-----AG------G----A--------CA-A-------G---A--T--------------------------------A----------------------------------------G------GGAT-G-----T---GA-C-------- 9015
STM.US.x.STM.M83293            -A--A--A--------A------A-C----GA--GA--------CA-A-----A-T--A---------------A----------C------------------G--AG-------------------G-A--A----T----GG-T-GT-----C--GA-C-------C 8644
MAC.US.x.80035.AY611486        ------------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 8099
MAC.US.x.81035.AY599200        ------------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 8100
MAC.US.x.85013.AY611490        ------------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 8033
MAC.US.x.92050.AY603959        ------------------------------G---------------------------------------------------------------------------------------------------------------------------T--------------- 8093
MAC.US.x.92077.AY599201        ------------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 8100
MAC.US.x.93057.AY611492        ------------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 8098
MAC.US.x.93062.AY607704        ------------------------------G-------------------------------------------------------------------------------------------------G----------------------------------------- 8067
MAC.US.x.95058.AY611494        ------------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 8087
MAC.US.x.95086.AY607703        ------------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 7956
MAC.US.x.95112.AY588946        ------------------------------R------------------------------------------------------------------------------------------------------------------------------------------- 8089
MAC.US.x.96016.AY607701        ------------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 8100
MAC.US.x.96020.AY611488        ------------------------------R------------------------------------------------------------------------------------------------------------------------------------------- 8092
MAC.US.x.96072.AY611491        ------------------------------G---------------------------------------------------------------------------------------------------------------------------T--------------- 8104
MAC.US.x.96081.AY597209        ------------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 8102
MAC.US.x.96093.AY611489        ------------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 8100
MAC.US.x.96114.AY588945        ------------------------------G-------------------------------------------------------------------------------------------------R----------------------------------------- 8095
MAC.US.x.96123.AY611487        ------------------------------G-------------------------------------------------------------------------------------------------G----------------------------------------- 7989
MAC.US.x.96135.AY607702        ------------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 8085
MAC.US.x.97009.AY599199        ------------------------------G-----------------------------------------------------------------R------------------------------------------------------------------------- 7968
MAC.US.x.97074.AY599198        --R---------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 7962
MAC.US.x.r80025.AY576480       ------------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 7956
MAC.US.x.r90131.AY576481       ------------------------------G------------------------------------------------------------------------------------------------------------------------------------------- 8099
```

```
                              Rev end                      Nef start
MAC.US.x.239              CCTCCAACCAATACTCCAGAGGCTCTCTGCGACCCTACAGAGGATTCGAGAAGTCCTCAGGACTGAACTGACCTACCTACAATATGGGTGGAGCTATTTCCATGAGGCGGTCCAGGCCGTCTGGAGATCTGCGACAGAGACTCTTGCGGGCGCGTGGGGAGAC...TTAT 9160
Env                       __L__Q__P__I__L__Q__R__L__S__A__T__L__Q__R__I__R__E__V__L__R__T__E__L__T__Y__L__Q__Y__G__W__S__Y__F__H__E__A__V__Q__A__V__W__R__S__A__T__E__T__L__A__G__A__W__G__D__.__L_
Nef                                                                                      _M__G__G__A__I__S__M__R__R__S__R__P__S__G__D__L__R__Q__R__L__L__R__A__R__G__E__T__.__Y_
Rev exon 2                _P__P__T__N__T__P__E__A__L__C__D__P__T__E__D__S__R__S__P__Q__D__*_
H2A.CI.88.UC2.U38293      --G--G--TG--CTC-----AT--G...........ACAGCA--CA-G--CTGG--G--ACT-A-GGC-G-------G-----------CGAG-GGA---A--A--TT----A--AA--GCA--GA------G--------------------A-G-GG...---- 9157
H2A.DE.x.BEN.M30502       --G--G--TG--CTC------T--A...........ACAGCA--CA-G--CTGG--G--ACT-A-GGC-G--C-A--G-----------CGAG-GGA---A--A--TT----A--AT--GC---GA--A---G-------------------------TGG...---- 9203
H2A.DE.x.PEI2.U22047      --------T---CT-------T--TCAGAGAG-A---ACA-CA--CA-G--CTGG--G--ACT-AC-A-AG------G-----------CGAG-GGA---A--A-T-C------T-C-TGCA--GA--A---G-------------A--------A------...--G- 9159
H2A.GH.x.GH1.M30895       --A--G--TG--CTC-----AT--A...........ACAGCA--CA-G--CTGG--G--ACT-A--GC-G-------G---------G-CGAG-GGA---A--A--AT----A--AT--GCA-AGA--A---G-------------A----A-----G-GG...---- 8623
H2A.IN.NNVA.NEW           ---------C--C------AT---......................A---CTGG--G---CT-AG-ACAG---T---G-----------CGAG-GGA---A--A--AT---------C-GC---GA--A--GG-------------A-------CA-GAG-...--G- 9186
H2A.IN.95.CRIK            ---------C--C------AT---......................A---CTGG--G--ACT-A--ACAG---T---G-----------CGAG-GGA---A--A--AT----A---T--GC---GA----GG-A-----------A---AG-CA--A--...--G- 8949
H2A.GM.x.MCR35.AY509260   -----G----GCTT-------T--TCAGAG-G-A--GACAGCA--CA----CTGG--A--A-----CGCAG-----T-G--G-------CGAG-GGA---A-GA---T--------T--GCA--GG--A---G---------------A-----A-------...--G- 8642
H2A.GM.87.D194.J04542     --G--G--TG--CTC------T--A...........ACAGCA--CA-G--CTGG--G--ACT-A-GGC-G------G-----------CGAG-GGA---A--A---T---GA--AT--GCA--GA------G-------A-------------A-G-GG...---- 8617
H2A.GM.x.ISY.J04498       ----------.................................C-CA---CTGG--G--ACTCA-GGCAG------G--G-------CGAG-GGA---A--A---T--------C--GC---GGT-A-A-G----------A--A-----G--A-GAG-...--G- 8602
H2A.GM.x.MCN13.AY509259   -----G----G-TT-------T--TCAGAG-G-A--GACAGCA--CA----CTGG--A--A-----CGCAG-----T-G--G-------CGAG-GGA---A-GA---T--------T--GCA--GG--A---G---------------A-----A-------...--G- 8642
H2A.GW.86.FG.J03654       ------------CT-------T---CAGAGAG-A---ACAGCA--CA----CTGG--G---CT-A--GCAG------G--G-------CGAG-GGA---A--A---T----A---C-TGCA--GA--A-A-G-------------A------G--A-------...--G- 8626
H2A.GW.87.CAM2CG.D00835   ---------C--C------AT---......................A---CTGG--G---C--A--G-AG---T---G-----------CGAG-GGA---A--A---T---------CTGCA--GG----A-G--------------------CA-G---...G-G- 9206
H2A.GW.x.ALI.AF082339     ---------G--CT---G---T--TCAGAGAG-G--GACA-CA--CA-G--CTGG--G--ACT-A--GCAG------G--G-------CGAG-GGA---A--A---T---G----T-TGCA--GAT-----G----------A-AAA-A-C---A-------...--G- 9199
H2A.GW.x.MDS.Z48731       ---------C--CT------AT---.....................A---CCAG--GC-ACT-AG-A-AG---T-----------------CGAG-GGA---A--A---T--------T--GC---GGT-A-A-G----------A--A-------CA-G-G-...--G- 8671
H2A.SN.85.ROD.M15390      --------TC--CTA-----AT---.....................A---CTGG--G--ACT-AG-ACAG---T-T-G-----------CGAG-GGA---A--A--AT---------C-GC---GG--A-A-G---------------------CA-G-G-...--G- 8640
H2AB.CI.90.7312A.L36874   ----------GCT---.................................C--CCA-----CTCCTGT-TG-G-----C---------ATCG---GG-----A--A--A-----A--A-CAGC-G-GG--A--GG-------------A--A-AG--A-GAC-...---- 9173
H2B.CI.88.UC1.L07625      ----------GC----.................................C--CCG------CT-TC---TG-A-----C--G------ATC----GG-----A--A--AA----A--A-CAAC---GG------G-------------AA-A-AG-AA---C-...C--- 9169
H2B.CI.x.EHO.U27200       ----------G-----.................................C--CCG------CT-CC--CTG-A-----C-G------ATC----GG-----A--A--AA----A--A-CAGCC--GG-----GG-------------A-----GC-A-GAC-...-CG- 9144
H2B.GH.86.D205.X61240     ------C-A---CTCAACC-AT---........................C--CCA-----CTCCCGG-TG-G-----C--G-----ATC----GG-----A--A---C---GA--G-CAGCA--GG--A--GG---A---------A-T--CG-A-AGAC-...---- 9167
H2B.JP.01.KR020.AB100245  --------T-G--T--.................................C--CAG-----ACT-CC---TG-A-----C---------ATCG---GG-----A--A--AA----A--A-CGAC---GG-----GG----------A-AAA-A-AG-AA-G-C-...C--- 8275
H2AB.CM.03CM_510_03.NEW   ----------------A...................................CCA------CTCCCGT-TG-A-----C--G------ATCG---GG-----A--A---C---GA--A-CAGC---GG--A---G-------------A-T---G-A-AGAC-...---- 8261
H2G.CI.92.ABT96.AF208027  T--------GC-GT----A--AA----CAG------C--AGCA---A-----CAT---C--CTA---GCAG-T--TT-CAGC-------TCC---GGC----A--A---TG-ACC--A-CAACT--GG------CA----------A-AA--A-----A---C-...---- 8588
H2U.FR.96.12034.AY530889  ------G--GT-GT-------CA--.....................A-----CGT--GCACCT----A-AG-T-------------------CAG------A-A--A--AT----A--GT--G---A-G-----CG----------AGCC--A--G--A----A...C--- 8678
SMM.SL.92.SL92B.AF334679  --------A-C-GTG------A-----A-A--T-AGT---CC----A----GC-TG----A-GA---GCAGG---TA-C-GT----------AT--C---ATC--A--CTG-----AA-CA-------------CA---AG--A---TC------------CT-...A--- 8608
MAC.US.x.MAC239           -----------------------------------------------------------------------G-------------------------------------------------------------------------------------------...---- 8237
MAC.US.x.251              -------------------------------------GA-----------------------------------------------------A-----------------------T--A-------------------------------------------...---- 9156
MAC.US.x.251_32H          --------------------------------G---------A-----------------------------------------------------------G-----------A-T--G-------------------------------------------...---- 9158
MAC.US.x.251              -------------------------------------GA---G-----------------------------------------------------------------------A----G-------------------A------------------A----...---- 9142
MAC.US.x.17EC1.AY033233   -------------------------------------------------------------------------------------------------------------------------------------------------------------------...---- 9160
MAC.US.x.17EFR.AY033146   --------------------------------------------------------------------------------------------------------------------T----------------------------------------------...---- 9160
MAC.US.x.1937.AY611495    -------------T-------------Y------------------------------------------------------------------------T--------------------------------------------------------------...---- 8268
MAC.US.x.2065.AY611493    -------------T---------------------------------R----------------------R-----------------------------YW--------------------------------------------------------R----...W--- 8266
MAC.US.x.MM142.Y00277     -------------T------------------------G---A------------------CT---------------------------------------A--A--------A----C-CA---------------------------------------A...---- 8655
MAC.US.x.SMM142B.BD131285-------------T--------------------TACGGAGAAT-CG--------------CT---------------------------------------A--A--------A----C-A-.--------------------------------------A..A---- 8652
MNE.US.82.MNE             -------------T-------AT----CA-------------A--C-----------------------A--------------------------C-----A-----------A-T--C-------------------------------------------...---- 8634
MNE.US.x.MNE027.U79412    -------------T-------AT----CA-------------AG-C-----------------------A--------------------------C-----A--A--------A-T--C-------------------------------------------...---- 8637
SMM.US.x.F236             ----------G-G--------C-----AA----GT-G--A--AG-C--T------A----A-T----A-AG-------------------C----------A--A--A--A--A--GTGG----A-CT----CG------------AA----------------...A--- 9125
SMM.US.x.H9.M80194        ------G---G-G--------C-----AAG---G--G-----AGCC--T------A----AGT---GA-A----------G---------C-T--CC----A--A--A-CG--RRG-TGG----A--T--YRCG------------AA--------A-----...---- 8649
SMM.US.x.PBJ14            ------G---G-G--------C-----AAG---G--G-----AGCC--T------A----AGT-C-GA-AG----------G---------C-T--CC----A--A--A-CG-----GTGG----A-CT--TACG------------AA--------A--C--...---- 8952
SMM.US.x.PBJA.M31325      ------G---G-G--------C-----AAG---G--G-----AGCC--T------A----AGT-C-GA-AG----------G---------C-T--CC----A--A--A-CG-----GTGG----A--T--TACG------------AA--------A-----...---- 8951
SMM.US.x.PBJ              ------G---G-G--------C-----AAG---G--G-----AGCC--T------A----AGT---GA-AG----------G---------C-T--CC----A--A--A-CG--RRG-TGG----A--T--YRCG------------AA--------A-----...---- 8664
SMM.US.x.PBJ              ------G---G-G--------C-----AAG---G--G-----AGCC--T------A----AGT-C-GA-AG----------G---------C-T--CC----A--A--A-CG-----GTGG----A--T--TACG------------AA--------A-----...---- 9144
SMM.US.x.PGM53.AF077017   ----------G-GT-----------.........T-------A--C---------G---------GA-AT-------G-----------CC---GG-----A--A--A--------A-CG-----GG--CG-GG---A--------AA--------A-----...-C-- 9097
SMM.US.x.SME543.U72748    ----------G-G--------A-----AAG---G--G--A--AG----T------A----A-T----A-A------------------------------------A--A--A-CA-----GTGG----A--T-----CG-------------A--------A------...A--- 9182
STM.US.x.STM.M83293       ------C----CCT----A---A----CAG--TA--G-----A---A--------G----ACTA-G-GCAG----TT----G---------C-T--GGA----A--A--A-CG--A--A-CA------G-----GG-------------A-----G--A------...---- 8811
MAC.US.x.80035.AY611486   -------------------------------------------------------------------------------------------------------------------------------------------------------------------...---- 8266
MAC.US.x.81035.AY599200   -----------------------------------------------R-------------------------------------------------------------------------------------------------------------------...---- 8267
MAC.US.x.85013.AY611490   ----------------------------------------------------------------------------------------------------------------------K-----------------------------R--------------...---- 8200
MAC.US.x.92050.AY603959   -------------T--------------------------------------------------------G--------------------------------------------------------------------------------------------...---- 8260
MAC.US.x.92077.AY599201   -------------------------------------------------------------------------------------------------------------------------------------------------------------------...---- 8267
MAC.US.x.93057.AY611492   -------------------------------------------------------------------------------------------------------------------------------------------------------------------...---- 8265
MAC.US.x.93062.AY607704   ------------------------------------------------------------------------------------------C-----------G------------------------------------------------------------...--G- 8234
MAC.US.x.95058.AY611494   -------------------------------------------------------------------------------------------------------------------------------------------------------------------...---- 8254
MAC.US.x.95086.AY607703   -------------------------------------------------------------------------------------------------------------------------------------------------------------------...---- 8123
MAC.US.x.95112.AY588946   -------------------------------------------------------------------------------------------------------------------------------------------------------------------...---- 8256
MAC.US.x.96016.AY607701   -------------------------------------------------------------------------------------------------------------------------------------------------------------------...---- 8267
MAC.US.x.96020.AY611488   -------------------------------------------------------------------------------------------------------------------------------------------------------------------...---- 8259
MAC.US.x.96072.AY611491   ---------------------------Y------------R---------------------T-------------------------------------T-----------------T--------------------------------------------...---- 8271
MAC.US.x.96081.AY597209   -------------------------------------------------------------------------------------------------------------------------------------------------------------------...---- 8269
MAC.US.x.96093.AY611489   ------------------------------------------------------------------------------------------------------C----------------K-------------------------------------------...--G- 8267
MAC.US.x.96114.AY588945   ------------------------------------------------------------------------------------------------------K------------------------------------------------------------...--R- 8262
MAC.US.x.96123.AY611487   -------------------------------------------------------------------------------------------------------------------------------------------------------------------...---- 8156
MAC.US.x.96135.AY607702   -------------------------------------------------------------------------------------------------------------------------------------------------------------------...---- 8252
MAC.US.x.97009.AY599199   --------------------------------------------------------------------------------------------------------------------------------------------------------------R----...---- 8135
MAC.US.x.97074.AY599198   ---------------------------------------------------------------------------------------------Y-------R-------------------------------------------------------------...---- 8129
MAC.US.x.r80025.AY576480  -------------------------------------------------------------------------------------------------------------------------------------------------------------------...---- 8123
MAC.US.x.r90131.AY576481  -------------------------------------------------------------------------------------------------------------------------------------------------------------------...---- 8266
```

```
                              Env gp41, gp160 end
MAC.US.x.239              GGGAGACTCTTAGGAGAGGTGGAAGATGGATACTCGCAATCCCCAGGAGGATTAGACAAGGGCTTGAGCTCACTCTCTTGT..GAGGGACAGAAA......TACAATCAGGGACAGTATATGAATACTCCATGGAGAAAC...CCAGCTGAAGAGAGAGAAAAATTAGCA 9319
Env                       W  E  T  L  R  R  G  G  R  W  I  L  A  I  P  R  R  I  R  Q  G  L  E  L  T  L  L  ##*
Nef                       _G__R__L__L__G__E__V__E__D__G__Y__S__Q__S__P__G__G__L__D__K__G__L__S__S__L__S__C_##_E__G__Q__K__.__.__Y__N__Q__G__Q__Y__M__N__T__P__W__R__N__.__P__A__E__E__R__E__K__L__A_
H2A.CI.88.UC2.U38293      -TA--G-AG-GCA-C-CATC--G--GG-A--C------G---A--------C--G--G--AGCA--AA--G-C---C---..---------C-G.......-TCGG------G-T--C-----C--C----------C-...-----AATG--AG--C-G---GAGT-- 9316
H2A.DE.x.BEN.M30502       ----AG-AGCGC-AC-CATC--G--GG-A---------G-T--A--A--A--C--G--G--AGCA--A---G-C---C---..---------C-G.......-TC-G------G----C-----C-GC-------------...-----AAC---A---C-G---GATTTG 9362
H2A.DE.x.PEI2.U22047      ---G-G-AA-GG-AC-GATC--C--GG-A-----T---G-T--A--A-----C--G--G---GCA--A--TG-C---C---..-----G---CGG.......-TC-A------G-C-T------C--C----------CT...-----A-C--GA--G--GGG-AC-TTG 9318
H2A.GH.x.GH1.M30895       -T-CAG-AG-GCAAC-C-TC--G--GG-A---------G---A--A-----C--G------GCA--AA--G-C---C---..---------C-G.......-TCGG------G-T-TC-----C--C----------C-...--G--AAT---AG--C-G---AAGTT- 8782
H2A.IN.NNVA.NEW           ---G---GT-GC-AC-CAT---G--GG-A--------GG-T--A--A--A--C--G--G---GCA------G-C---C---..-------G-C-GAATCACC-GGGAG--TACGC-G-CCGCCCATGGAGGAATC..CCATCAA-G-AGAG-CC--AC--TTGG-AT-GG 9352
H2A.IN.95.CRIK            -T-G---AG-GGAA---AT---G--GG-A------------A--A-----C---------GCA---G--G-C---C---..---------C--.......-TC-A--A--GG-C-TC-------------------C-...-----AAC---A-A--GG---GACTT- 9108
H2A.GM.x.MCR35.AY509260   ---G-G-A--GCA-C-GATC--G--GG-A-----T---G----A--A--A--C--G--G--AGCA---A--G-C---C-A-..--A------CGG.......-TC-G------G-C-T---------C----A-----C-...-----AA----AG-G--G---GA-TTG 8801
H2A.GM.87.D194.J04542     -T--AG-AGCGCAAC-CATC--G--GG-A--C------G----A--A-----C--G--G--AGCA--AA--G-C---C---..---------C-G.......-TC-G------G---TC-----C--C----------C-...-----A-C-ATAG-GC-G---AATT-- 8776
H2A.GM.x.ISY.J04498       ---G-G----GG-AC--ATC-----GG-----------G-T--AC-A-----C--G--G--AGCA--AA-TG-C---C---..----------G-.......-TC-G--A---G-C-T---------C----------C-...-----AAC---A-AG------GA-T-G 8761
H2A.GM.x.MCN13.AY509259   ---G-G-A--GCA-C-GATC--G--GG-A-----T---G----A--A--A--C--G--G--AGCA---A--G-C---C-A-..---------CGG.......-TC-G------G-C-T---------C----------C-...-----AA----AG-G--G---GA-TTG 8801
H2A.GW.86.FG.J03654       --AG-G-G--GCAAC--ATC--G--GG-------T---G----AC-A-----C--G--G--AGCA--A---G-C---C---..---------CG-.......-TC-G------G-T-T-G-A-----C----------C-...-----A-C---A--G--G---GA-TTG 8785
H2A.GW.87.CAM2CG.D00835   ---GA-TGT-ACAAC-CAT------GG-A--C--T--GG---A--A-----C-----G--AGCA--A---G-----C---..---------CG-.......-TC-G-----GG-C-T---------C---------GC-...-----A-C----G-GA-G---AAT--- 9365
H2A.GW.x.ALI.AF082339     ---G-G-AG-GCA-T-G-TC--G--GA-----------G---A--------C--G--G---GCA--AA-TG-C---C---..--A------CGG.......-TC-G------G-T-T---------C----------C-...-----AAC---A-----G---GA-TTG 9358
H2A.GW.x.MDS.Z48731       ---G---AT-GG-AC--AT---G--GG-A--------GG---A--A-----C---------GCA---A--G-C---C---..---------C-G.......-TC-A--A--GG-C-TC-------------------C-...-----AAC---AG-G--G---GATT-- 8830
H2A.SN.85.ROD.M15390      --AG-GTAT-GGAAC--ATC--G--GG-A--------GG-T--A--A-----C-----G--AGCA--AA--G-C---C---..-------G-C-G.......-TC-G------G-C-T----------------AGG--...-----A-C---A--G--G---AATTTG 8799
H2AB.CI.90.7312A.L36874   -----G----C-----GACG-CG--GG-A--CA-----G-------A--A--C-----G-------A---G-C---C---..--------A.......................-CCGC-...----GA-C---AG-G-G-GGG...... 9290
H2B.CI.88.UC1.L07625      --A-AG----C--A--GACA-CGGA-GCA--CA----G-------------C--------A-----A---G-C-------..-------GC.......................-ATGC-...---AGA-C---AG-G-G-GGG...... 9286
H2B.CI.x.EHO.U27200       ---GAGTC--C--------CG-C-G--GA---CA-T-----A-----------C---------GC--A---G-C-------..--T-----A.......................-AG-CT...TTG-GA-C---AG-G-G-GG-...... 9261
H2B.GH.86.D205.X61240     -----G----C--------CG-CG---GCA--CA----G-----------A--C--------A-----A----C-------..--------A.......................-ATGC-...-GG-GG-C---AG-G-G-GGG...... 9284
H2B.JP.01.KR020.AB100245  -----G----C--A---ACA-CG--GGCA--CG----G-------------C--------A--C--A---G-C-------..-------GC.......................--TGCT...----GA-C-C-AG-G-G-GGG...... 8392
H2AB.CM.03CM_510_03.NEW   ----AG----C---G-G-CG-CG---GC---C-----G-------A--A--C--------A-----A---G-C-------............................-ATGC-...-TG-GG-C---AG-G-G-GGG...... 8375
H2G.CI.92.ABT96.AF208027  --A-A-----GG-A--G-T-------G------------A--G--A--A--C--G-------------------------..--A--G--AC--.......----CG-----T--A-T---------C-------------...-----AAC---A----C---G------ 8747
H2U.FR.96.12034.AY530889  ----------GG--C-G-T----------C-T-GA--T--A--A--------C--G-----AT--------G-G--T-----..--A-CT---GGCCAAGGC--T-G-G-A---G---TC--------A--T-------C-...-----AAC--G--A---T------CA- 8843
SMM.SL.92.SL92B.AF334679  ----------GG-AC-G-T-------G----CGCA--------A--------C-----G--C-----A----TG--TAAC-..AG--ATGGAT-TAAA..-------GTA--GG-A-T------C--C------------...--T---AC---A---C----GCA-CAG 8770
MAC.US.x.MAC239           ---------------------------------------------------------------------------------..------------......---------------------------------------...--------------------------- 8396
MAC.US.x.251              -A---G-------------------------C-----------T-----A-------------------------------..------------......----------A------C---------------------...--------------------------- 9315
MAC.US.x.251_32H          -------------------------------------------T-------------------------------------..------------......----------G----------------------------...--------------------------- 9317
MAC.US.x.251              -------------------------------C-----------T-----------G------------------G------..------------......----------G----------------------------...-----A-A------------------- 9301
MAC.US.x.17EC1.AY033233   ---------------------------------------------------------------------------------..------------......---------------------------------------...--------------------------- 9319
MAC.US.x.17EFR.AY033146   ---------------------------------------------------------------------------------..--------A---......---------------------------------------...------A-------------------- 9319
MAC.US.x.1937.AY611495    ----------------------G----------------------------------------------------------..------------......---------------------------------------...--------------------------- 8427
MAC.US.x.2065.AY611493    -----------------------R---------------------------------------------------------..------------......---------------------------------------...--------------------------- 8425
MAC.US.x.MM142.Y00277     -----G-----CAA--G--------------C-----------T-------------------------------------..-----C--A---......------------G-A--C---------------------...--------------GA--------C-- 8814
MAC.US.x.SMM142B.BD131285-----G-----CAA--G--------------C-----------T-------------------------------------..-----C--A---......------------G-A--C---------------------...--------------GA--------C-- 8811
MNE.US.82.MNE             ----------GG-A----T------------------------T--------C-----------C-----T----------..------------......----------------T-----------------A----...--------G-----G------------ 8793
MNE.US.x.MNE027.U79412    ----------GG-A--G-T------------------------T--------C-------A---------T----------..------------......----------------T-----------------A----...--------G-----G------------ 8796
SMM.US.x.F236             ----------GG-A--G-T-------G-----------------T---C-C--C--G------------------------..---CCT------.......-T-G-G-------A----------C--C-----G---...-----AAC---A----C--------GT 9284
SMM.US.x.H9.M80194        ----------GG-A--G-T------GG-A-------------MR-AC-C--C--G--------------------------..---CCT------.......--TTG-G----T--A-TC--------C--T-----G---...-----AAC---AG---C--------GT 8808
SMM.US.x.PBJ14            ----------GG-A--G-T------GG-A-------------A--AC-C--C--G--------------------------..---CCT------.......--TTG-G----T--A-T---------C--T-----G---...-----AAC---A----C--------AT 9111
SMM.US.x.PBJA.M31325      ----------GG-A--G-T------GG-A-------------A--AC-C--C--G--------------------------..---CCT------.......--TTG-G----T--A-T---------C--T-----G---...-----AAC---A----C--------AT 9110
SMM.US.x.PBJ              ----------GG-A--G-T------GG-A-------------MR-AC-C--C--G--------------------------..---CCT------.......--TTG-G----T--A-TC--------C--T-----G---...-----AAC---A----C--------AT 8823
SMM.US.x.PBJ              ----------GG-A--G-T------GG-A-------------A--AC-C--C--G--------------------------..---CCT------.......--TTG-G----T--A-T---------C--T-----G---...-----AAC---A----C--------AT 9303
SMM.US.x.PGM53.AF077017   ----A-----GG-A--G-T----------G------------G----A-----G--G--A--G--A--TG-A---C-T-..---------GG-.......---G-G-------A-TC-----C--C-----------T...-------AG------------G---AA- 9256
SMM.US.x.SME543.U72748    ----------GG-A--G-T-------G-----------------T---C-CG-C--G--------------G---------..---CCT------.......-T-G-G-------A----------C--C-----G---...---A-A-C---A-A--C--------GT 9341
STM.US.x.STM.M83293       ----A------G-A--G-T---G---A----CGGG--C-----A---C-C--C---------------A-----A--------..---CCT----G-.......--T---G-A--T--A-T---------C--T----A----...--G--A-C---A--C-CT--GC---A- 8970
MAC.US.x.80035.AY611486   ---------------------------------------------------------------------------------..------------......----------R----------------------------...--------------------------- 8425
MAC.US.x.81035.AY599200   ---------------------------------------------------------------------------------..------------......---------------------------------------...--------------------------- 8426
MAC.US.x.85013.AY611490   ---------------------------------------------------------------------------------..------------......---------------------------------------...--------T------------------ 8359
MAC.US.x.92050.AY603959   -------------------------------------------------A-------------------------------..------------......---------------------------------------...--------------------------- 8419
MAC.US.x.92077.AY599201   ---------------------------------------------------------------------------------..------------......---------------------------------------...--------------------------- 8426
MAC.US.x.93057.AY611492   -----R-------------------------------------------------R--------Y----------------..------------......----------A----------------------------...--------T------------------ 8424
MAC.US.x.93062.AY607704   ---------------------------------------------------------------------------------..------------......---------------------------------------...--------T------------------ 8393
MAC.US.x.95058.AY611494   ---------------------------------------------------------------------------------..------------......-------------------------------------R-...--------------------------- 8413
MAC.US.x.95086.AY607703   ---------------------------------------------------------------------------------..------------......---------------------------------------...--------------------------- 8282
MAC.US.x.95112.AY588946   ---------------------------------------------------------------------------------..------------......---------------------------------------...--------------------------- 8415
MAC.US.x.96016.AY607701   ---------------------------------------------------------------------------------..------------......---------------------------------------...--------------------------- 8426
MAC.US.x.96020.AY611488   ---------------------------------------------------------------------------------..------------......---------------------------------------...--------------------------- 8418
MAC.US.x.96072.AY611491   ---------------------------------------------------------------------------------..------------......------------G--------------------------...----------RR--------------- 8430
MAC.US.x.96081.AY597209   -------------------------------------------------R-------------------------------..------------......---------------------------------------...--------------------------- 8428
MAC.US.x.96093.AY611489   -------------------------------------------------------R-----------------W-------..------------......---------------------------------------...--------------------------- 8426
MAC.US.x.96114.AY588945   ---------------------------------------------------------------------------------..------------......---------------------------------------...--------------------------- 8421
MAC.US.x.96123.AY611487   ---------------------------------------------------------------------------------..------------......---------------------------------------...--------------------------- 8315
MAC.US.x.96135.AY607702   ---------------------------------------------------------------------------------..------------......---------------------------------------...---------------R----------- 8411
MAC.US.x.97009.AY599199   ---------------------------------------------------------------------------------..------------......---------------------------------------...--------------------------- 8294
MAC.US.x.97074.AY599198   ---------------------------------------------------------------------------------..------------......---------------------------------------...--------------------------- 8288
MAC.US.x.r80025.AY576480  ---------------------------------------------------------------------------------..------------......---------------------------------------...--------------------------- 8282
MAC.US.x.r90131.AY576481  ---------------------------------------------------------------------------------..------------......---------------------------------------...--------------------------- 8425
```

```
                                                 Premature stop in SMM239                                                                                                           3' LTR U3 start
MAC.US.x.239             TACAGAAAACAAAATATGGATGATATAGATGAGTAAGATGATGAC...TTGGTAGGGGTATCAGTGAGG...CCAAAAGTTCCCCTAAGAACAATGAGTTACAAATTGGCAATAGACATGTCTCATTTTATAAAAGAAAAGGGGGGACTGGAAGGGATTTATTACAGTGC 9483
Nef                      _Y__R__K__Q__N__M__D__D__I__D__E__*__D__D__D__.__L__V__G__V__S__V__R__.__P__K__V__P__L__R__T__M__S__Y__K__L__A__I__D__M__S__H__F__I__K__E__K__G__G__L__E__G__I__Y__Y__S__A
H2A.CI.88.UC2.U38293     ----AGC-G--------------G-----TCTG-T---A------...CAA--G--A--T--T--T-CA...T---G---A--A------C------CC-----G------G----T-----A-----------------A----------------G-T-------AG 9480
H2A.DE.x.BEN.M30502      --T--GC-G--------------G-----TCTG-T----------...C-AA----A--TC-T--T-CA...----G---A--A-GG---GA-----CC--T-------------------T-----A-----------------A--------C-------G-T-------AG 9526
H2A.DE.x.PEI2.U22047     ----AGC----------------G------CAG-TA-----A--...C-AA-------CC-T--C-CA...----G---A--AT----GG-----CA--T--G------G----T--A--A------C----T----------------T-----G---------A 9482
H2A.GH.x.GH1.M30895      ----AGC-G--------------------TCTAGT----------...C-A--G--A--TC-T--T-CA...----G---A--A------G------CC--T---------G----T-----A-----------GA---G-...---------T-----G-T-------AG 8943
H2A.IN.NNVA.NEW          G-GCAG--CA-GG.GGAT-TA---TC------TG-T--C.........C-A-----A--CC-T--C-CA...T---GGACA-AA--G--GGA-----CC--T-----T------------------A-------T--------------------G---------A 9509
H2A.IN.95.CRIK           -----GC----------------G-----TCAG-T----------...C-A-----A--T--T--C-CA...----GG--A--A--------------CA--T-G-A----------------A-----A-------------------------G-T--------A 9272
H2A.GM.x.MCR35.AY509260  ----AGC---------GA------G-----TCAG-T----------...C-A---AA---C--T--C-CA...----G---A--A------GA-----CAC-T-G---A--------T-----A-----------------A----------------G---------A 8965
H2A.GM.87.D194.J04542    --T-AGC-G--------------G-----TCTG-T----------...C-A--G--A--TC-T--T-T-...----G---A--G--G---GA-----CC--T---C---------T-----A-----------------A--A-------------A-T-------AG 8940
H2A.GM.x.ISY.J04498      -----GC----------------G-----TCAG-T----------...C-A--------C--T-AC-CA...T---G---A--AT-G---G------CA--T-G-A-----G---------AG---A---------T--------------------G---------A 8925
H2A.GM.x.MCN13.AY509259  ----AGC---------GA------G-----TC-G-T----------...C-A--------C--T--C-CA...----G---A--A------GA-T---CAC-T-G---A--------TG---A-----------------A----------------G---------A 8965
H2A.GW.86.FG.J03654      ----A-C-G--------------G-----CTAG-T----------...CAA----AT-CC-T--C-CA...----G---A--AT-----C------CA-T----------G----T---------------------A---------------C-G-T---T---CA 8949
H2A.GW.87.CAM2CG.D00835  -----GC-G--------------------TCAG-T----------...C-A--G-----TC-T-CC-CA...----G---A--A--------------CA--T---------G---------A-----------------A---------------C-G-T--------A 9529
H2A.GW.x.ALI.AF082339    ----AGC----G-----------G-----TTAG-T--------TAGCC-A--------C--T--C-CA...----G---A-AAT-------------CA------------G----T-----A-----A----------G-------------------G-T--------A 9525
H2A.GW.x.MDS.Z48731      --T---C----------------G-----TCAG-T----------...C-A-----ATAC--T--C-CA...----G-ACA--A------G------CA--T---------G----T-----A-----------------A----------------G-T--------A 8994
H2A.SN.85.ROD.M15390     -----GC----------------G-----TCAG-T----------...CAA---A-A--T--T--C-CA...---------A--A------C------CAC-T-G-----------T-----A-----A-----AC--G-------------------G-T--------A 8963
H2AB.CI.90.7312A.L36874  ......C-----G-C--A--G--A......--TG-T--G--CA-T...GAA----------AT--C----...--CC-G--G--AT-G---C------CA-----------------------------------------------------------T---T----A 9442
H2B.CI.88.UC1.L07625     ......C-----G---CA-----......AGTG-T--G---A-T...GAA--G-----C-AT--A--A...--C--TAGG--A--G---T------CA-----GA-------------------------------------------------C--T----A 9438
H2B.CI.x.EHO.U27200      ......------G--TCA-----G......--TG-T--G--CA-T...GAA--------CCGT--A--A...--CGGG--C--A--G--GC------CA-T----C-A---G-----------------T--------------A------------TC--T----A 9413
H2B.GH.86.D205.X61240    ......C-----G--GCA-----A......AGTG-T--G--C--T...GAA--G--A-CCAT-TGT-A-ACC----T---A--A--G---C------CA------C-A---G------------------------C-----------------G-----T----A 9439
H2B.JP.01.KR020.AB100245 ......C----GGG---A--A--......AGTG-T--A---A-T...GAA--G--A---C-AT--A--A...--C--TAGG--A------C------CA--T-------------------------GG----------------------------C-------A 8544
H2AB.CM.03CM_510_03.NEW  ......C--------GCA-----A-GT------G-C-----A..........--G--A-CC-AT--A--A...--C-GTAGG--A--G---C------CA------ACA---------------------C---------------------T----------------A 8527
H2G.CI.92.ABT96.AF208027 -----GY-G-------AT------G-----AGTG-T---A-----...--A-----A---C----CTAC...--C-----C--G------GT----------------------------------A-------------------------------------------A 8911
H2U.FR.96.12034.AY530889 --T-AGC-------C---------G-------AG-----A----T...--A-----A---G----A--A...----G---AG-A--G----------C--------------G----T-----A-----------------A-----------T-----A-T---TTC-CA 9007
SMM.SL.92.SL92B.AF334679 --T--GC-G--GTG-------------------G-T--------G...C---CT--A---GTT--CT--...---------A---------G-----TCA---------------------A-----------------------T----------------A 8934
MAC.US.x.MAC239          ---------------------------------G-----------...-----------------------...-------G--------------------------------------------------------------------------------------- 8560
MAC.US.x.251             ---------------------------------G----------T...-----------------------...------------------------------------------------------------------------------------T------ 9479
MAC.US.x.251_32H         -----------------------G---------G-----------...-------------C------T-...---CG--------------------------------G----T-------------------------------------------------- 9481
MAC.US.x.251             ---------------------------------G-----------...-----------------------...-------------------G-----C------------------T-------------------------------------------------- 9465
MAC.US.x.17EC1.AY033233  ---------------------------------------------...-----------------------...------------------------------------------------------------------------------------------------ 9483
MAC.US.x.17EFR.AY033146  ---------------------------------C-----------...-----------------------...------------------------------------------A----------------------------------------------------- 9483
MAC.US.x.1937.AY611495   ---------------------------------G----Y------...-----------------------...A-----------------G------------------------------------------------------------------------------ 8591
MAC.US.x.2065.AY611493   ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8589
MAC.US.x.MM142.Y00277    ----------------A----------------G-----------...---------A--C----TGA-...G-C-G-----------------------------------T-------------------------------------------------------- 8978
MAC.US.x.SMM142B.BD131285----------------A----------------G-----------...---------A--C----TGA-...G-C-G-----------------------------------T-------------------------------------------------------- 8975
MNE.US.82.MNE            ----------------A---------------AG-----A-----...-------------C---------...---CG------T-----T--A-------------G---T--------------------------------------------------------A 8957
MNE.US.x.MNE027.U79412   ---------------------------------G---------G-...-------------C---------...---CG------T----G-----C-----------G---T--------------------------------------------------------A 8960
SMM.US.x.F236            --T---C-------C---------G-G-----TG-----------...--AA---T--C------CAC...----G---C---T----GG-C----CA------------------------------------------------------------------A--A 9448
SMM.US.x.H9.M80194       --T---C-------C---------G-G---A-TG--.........--T...--A-----TTGTC-------C...----G---C---G------TC----CA-------------------------------------------------G-----------AT 8966
SMM.US.x.PBJ14           --T---C-------C---------G-G---AGTGCT.......---...--A-----TTGTC-------C...----G---C---G------TC----CA-----------------------------------T-----------------------------A 9269
SMM.US.x.PBJA.M31325     --T---C-------C---------G-G---AGTGCT.......---...--A-----TTGTC-------C...----G---C---G------TC----CA------------------------------------------------------------------A 9268
SMM.US.x.PBJ             --T---C-------C---------G-G---AGTGCT.......---...--A-----TTGTC-------C...----G---C---G------TC----CA-------------------------------------------------G-----------AT 8981
SMM.US.x.PBJ             --T---C-------C---------G-----AGTGCT.......---...--A-----TTGTC-------C...----G---C---G------TC----CA------------------------------------------------------------------A 9461
SMM.US.x.PGM53.AF077017  --T---C-------C---------G-------TG-T--------G...C--A----A--------ACAC...-----G--G--AT-----G-C---TCA-----------G-------------------------------------------------------A 9420
SMM.US.x.SME543.U72748   --T-A-C-------C---------G-G-----TG-----------...--A-----T--C------CAC...---------C---T----GG-C----CA-----------------------------------------------------------------A--A 9505
STM.US.x.STM.M83293      --T---C-G-----C---------G--------G--------A-T...C-A-----A---G----ACAT...----G---C--A--T--GGA-----C------G------------C----A-----------AGC------------------------------A 9134
MAC.US.x.80035.AY611486  ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8589
MAC.US.x.81035.AY599200  ---------------------------------G-----------...-----------------------...--------------------------------A--------------------------------------------------------------- 8590
MAC.US.x.85013.AY611490  ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8523
MAC.US.x.92050.AY603959  ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8583
MAC.US.x.92077.AY599201  ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8590
MAC.US.x.93057.AY611492  ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8588
MAC.US.x.93062.AY607704  ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8557
MAC.US.x.95058.AY611494  ---------------------------------G-----------...-----------------------...-------------------G---------------------------------------------------------------------------- 8577
MAC.US.x.95086.AY607703  ---------------------------------G-----------...-----------------------...--------------------------------------R------------------------------------------------------- 8446
MAC.US.x.95112.AY588946  ---------------------------------S-----------...-----------------------...------------------------------------------------------------------------------------------------ 8579
MAC.US.x.96016.AY607701  ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8590
MAC.US.x.96020.AY611488  ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8582
MAC.US.x.96072.AY611491  ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8594
MAC.US.x.96081.AY597209  ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8592
MAC.US.x.96093.AY611489  ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8590
MAC.US.x.96114.AY588945  ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8585
MAC.US.x.96123.AY611487  ---------------------------------G-----------...--------------------A-...------------------------------------------------------------------------------------------------ 8479
MAC.US.x.96135.AY607702  ---------------------------------G-----------...--------------------R-...----R------------------------------------------------------------------------------------------- 8575
MAC.US.x.97009.AY599199  ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8458
MAC.US.x.97074.AY599198  ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8452
MAC.US.x.r80025.AY576480 ---------------------------------G-----------...-----------------------...--------------------------------------R------------------------------------------------------- 8446
MAC.US.x.r90131.AY576481 ---------------------------------G-----------...-----------------------...------------------------------------------------------------------------------------------------ 8589
```

```
MAC.US.x.239                 AAGAAGACATAGAATCTTAGACATATACTTAGAAAAGGAAGAAGGCATCATACCAGATTGGCAGGATTACACCTCAGGACCAGGAATTAGATACCCAAAGACATTTGGCTGGCTATGGAAATTAGTCCCTGTAAATGTATCAGATGAGGCACAG............GAGG 9641
Nef                          __R__R__H__R__I__L__D__I__Y__L__E__K__E__E__G__I__I__P__D__W__Q__D__Y__T__S__G__P__G__I__R__Y__P__K__T__F__G__W__L__W__K__L__V__P__V__N__V__S__D__E__A__Q__.__.__.__.__E__
H2A.CI.88.UC2.U38293         GGAT------------C-----T-G---C----------G-----GG-A---------------A----T--TCAT--G------G-G--G-------T-TGC--C--G-----G------C----A--A-----CA-G---C-A------G--............--T- 9638
H2A.DE.x.BEN.M30502          G--G------------C-----------C-------A--G-----G--A---------------A----T--TCAT--G------G-A--G-------T-TAC--C--G-----G-----GC----AT-A---G-AC-C---C-A------G--............--A- 9684
H2A.DE.x.PEI2.U22047         G---------------------------A-G--------------G--A--T------------A-C--T--TCAT--G------G-A--G---------TTC-----G-----------GC----A--A---G-C--CC--C-A-GT.....................- 9631
H2A.GH.x.GH1.M30895          GGAT------------C-----T-G---C----------G-----G--A---------------A----T--TCAT--G------G-A--G-------T-TGT--C--G-----------GC----A--A---G----C---C-A------G--............--C- 9101
H2A.IN.NNVA.NEW              G-------GAGA----A-----TT------------------G--G--A--T-----C------A-C--T---AGT--G--------A-----T----T-TTCCA---G-----------GC----A--A--------CC--C-A----G-G-A............--CA 9667
H2A.IN.95.CRIK               ----TT----G............................................................................................................................................................... 9283
H2A.GM.x.MCR35.AY509260      G---------------------------C-TA-------------G--A--T------------A-C--T--TCAT--G--------A----------T-TTC-----G------....................................................... 9080
H2A.GM.87.D194.J04542        GGAG------------C-----T-G-T-C----------G-----G--A---------------A----T--TCAT--G-------CA--G-------T-TAC--C--G-----G------C----A--A---G-CA-C---C-A------G--............--A- 9098
H2A.GM.x.ISY.J04498          G---------------C----------------------------G--A--T------------A-C--T--TCAT--G-T----G-A--G-------T-TTC-----G-----------GC----A--A----C---CC--C-A--A-GGG--............--CA 9083
H2A.GM.x.MCN13.AY509259      G---------------------------C-T--------------G--A--TG-----------A-C--T--TCAT--G--------A----------T-TTC-----G-----------GC----A--A---G----CA--CGAC---AGG--............--C- 9123
H2A.GW.86.FG.J03654          G---------------------T----------C---..................................................................................................................................... 8986
H2A.GW.87.CAM2CG.D00835      ----------------C-G--T--------------A--------G--A--TG-----------A-C--T--TAGT--G------G-A----------T-TTC-----G-----------GC----A--A---G--ACC---C-A----G-G--GACACTGAGACT--CA 9699
H2A.GW.x.ALI.AF082339        G--------------------------------------------G--A--T------------A-C--T--TCAT--G--------A--G-----G-T-TTC-----G-----G-----GC----A--A---G----CC--C-A--A-GGG--............--CA 9683
H2A.GW.x.MDS.Z48731          ----------------C----T-----------------------G--A--T------------A----T--TCAT--G-----G-CA--------C---TGC-----G-----------GC----A--A---G--C-CC--C-A--T-AGA-C...............A 9149
H2A.SN.85.ROD.M15390         -----------A-------A-T-----------------------G--A--TG-----------A-C-----TCAT--G------G-A----------T-TTC-----G-----------GC----A--A---G----CC--C-A--A-GGG--............--CA 9121
H2AB.CI.90.7312A.L36874      ------------G--A-----T-C---T--T--G--T--GAG---T--AG-TGGT-GA------A-C-----A-AT--G-----G--A-----------ACAC-----------G------C----G--A--GG-A--GG---CA-TAA-C-G-.........GAG--A- 9603
H2B.CI.88.UC1.L07625         G--------------AC------C---TC----G--T--G--------TG--T-T-GA-----AA-C--T--A-AT--G-----G--A--G--T----G---C-----G-----G-----GC----G--A---G--A--C----G----A-AG-............-GA- 9596
H2B.CI.x.EHO.U27200          G-----G----A---AC------C------------T-----------TG-GT-T-GA------A-C-----ACAT--G------G-A-----T--C---TTC-----------------GC----G--AA-----A-GAT--CA--AC--G--............--C- 9571
H2B.GH.86.D205.X61240        G--------C-----A-------C---T--T--G--T-----------TG-GT-T-GC-----AA-C--T--TCAT--G-----G--A--G--T--C--ATAC-----T-----G-----GC-G--A--A---G-G--GC---CA-C-A-C-GA.........GAG---- 9600
H2B.JP.01.KR020.AB100245     G--G---------T-AC------C---TC-------T--G--------TG--T-T-GA-----AA-C--T--ACAT--G------G-A--G--T--C--A--C-----G-----G-----GC----G--A------A--C----G--A-ACA--............--A- 8702
H2AB.CM.03CM_510_03.NEW      G--------------A-------C---T--T--G--T-----------TG-GT-C-GC-----AA-C--T--TCAT--G-----G--A-----T--C--ATAC-----------G-----GC----A------G-GAC-C--ACA-----C-G-.........GAG-GA- 8688
H2G.CI.92.ABT96.AF208027     ----------------C-----------A-G--------GC----T--A--T------------A-------AG----------T--A-----T---RT---G------------------C----A--A--C-CG--C--T------------............---- 9069
H2U.FR.96.12034.AY530889     G--------------------T--C-----------T--G--G--A--A--C------------A---------------------C-----GG----T-TGC-----A-----T-----GC---AA--A---G---C-AAT---A-A------............---- 9165
SMM.SL.92.SL92B.AF334679     G--G------------C----T------C-------A--G-----T-----C------------A----T--AAGT--G--------A--G---------TTC-----G-----C---C-GC-G--G--A---G-C--C--G---------A-C............A-T- 9092
MAC.US.x.MAC239              --------------------------------------------------------------------------T-----------------W-------------------------------------------------------------............---- 8718
MAC.US.x.251                 ----------------------------------------------------------------------------------------------------------------------------------------------------------............---- 9637
MAC.US.x.251_32H             ------------------------G--------------G-----------------------------------------------------------------------------------------C------------------------............---- 9639
MAC.US.x.251                 ------------------------G---------------------------------------------------------------------------------------------------------------------------------............---- 9623
MAC.US.x.17EC1.AY033233      ----------------------------------------------------------------------------------------------------------------------------------------------------------............---- 9641
MAC.US.x.17EFR.AY033146      ----------------------------------------A---------------------------------------------------T-------------------------------------------------------------............---- 9641
MAC.US.x.1937.AY611495       --------------------------------------------------------------------------------------C-------------CAT------------------------------G--------------------............---- 8749
MAC.US.x.2065.AY611493       --------------------------------------------------------------------------------------C-------------------------------------------------------------------............---- 8747
MAC.US.x.MM142.Y00277        ----------------------------------------------------------------ATA...CA---C-------------------T-----TG---------------------A-----------------------------............---- 9133
MAC.US.x.SMM142B.BD131285    ----------------------------------------------------------------ATA...CA---C-------------------T-----TG---------------------A-----------------------------............---- 9130
MNE.US.82.MNE                -----------A------------G--------------------------G------------A----------G---------CC-------------------------------------------------------------------............---- 9115
MNE.US.x.MNE027.U79412       -----------A-------------------------------------G-G-----------------------G---------CC-------------------------------------------------------------------............---- 9118
SMM.US.x.F236                G--------------A-----T--G-----G--------------A--A--------------AA-------A-----G-----------------T-T-CAC-A---G-----C--------------A---G----C--------A--T---............--A- 9606
SMM.US.x.H9.M80194           T----------A---A-----TC-R-----------R--------A--A--------------AA-------AG----G--------C--------T-T-TTC-----A-----C----------W---G--------C--------A--T---............--A- 9124
SMM.US.x.PBJ14               T----------A---A-----TC-G--------------------A--A--------------AA-------AG----G--------C--------T-T-TTC-----A-----C--------------G--------C--------A--T---............--A- 9427
SMM.US.x.PBJA.M31325         T----------A---A-----TC-G--------------------A--A--------------AA-------AG----G--------C--------T-T-TTC-----A-----C--------------G--------C--------A--T---............--A- 9426
SMM.US.x.PBJ                 T----------A---A-----TC-R-----------R--------A--A--------------AA-------AG----G--------C--------T-T-TTC-----A-----C----------W---G--------C--------A--T---............--A- 9139
SMM.US.x.PBJ                 T----------A---A-----TC-G--------------------A--A--------------AA-------AG----G--------C--------T-T-TTC-----A-----C--------------G--------C--------A--T---............--A- 9619
SMM.US.x.PGM53.AF077017      G-------------------------------------------------CC------------A-------A--------------C----------T-TTC--------------------G-----A----G--CC--------A--T--A............--A- 9578
SMM.US.x.SME543.U72748       G-A------------A-----T--G---A-G--------------A--A--------------AA-------A-T---G-------C---------T-T-TAC-----G-----C--------------A---G----C--------A--C---............--A- 9663
STM.US.x.STM.M83293          G--------------A--G-----G-----------A--------G--AG-T------------A-------AG----------G--A------------CAG-----A------------C----A--A---G--A-G---A-----------.........GAG--T- 9295
MAC.US.x.80035.AY611486      ----------------------------------------------------------------------------------------------------------------------------------------------------------............R--- 8747
MAC.US.x.81035.AY599200      ----------------------------------------------------------------------------------------------------------------------------------------------------------............A--- 8748
MAC.US.x.85013.AY611490      ----------------------------------------------------------------------------------------------------------------------------------------------------------............---- 8681
MAC.US.x.92050.AY603959      --------------------------------------------------------------------------------------------T-------------------------------------------------------------............---- 8741
MAC.US.x.92077.AY599201      ----------------------------------------------------------------------------------------------------------------------------------------------------------............---- 8748
MAC.US.x.93057.AY611492      -------------------------------------------------R--------------------------------------------------------------------------------------------------------............R--- 8746
MAC.US.x.93062.AY607704      -------------------------------------------------G--------------------------------------------------------------------------------------------------------............---- 8715
MAC.US.x.95058.AY611494      -------------------------------------------------R--------------------------------------------------------------------------------------------------------............---- 8735
MAC.US.x.95086.AY607703      ----------------------------------------------------------------------------------------------------------------------------------------------------------............R--- 8604
MAC.US.x.95112.AY588946      ----------------------------------------------------------------------------------------------------------------------------------------------------------............---- 8737
MAC.US.x.96016.AY607701      ----------------------------------------------------------------------------------------------------------------------------------------------------------............R--- 8748
MAC.US.x.96020.AY611488      ----------------------------------------------------------------------------------------------------------------------------------------------------------............---- 8740
MAC.US.x.96072.AY611491      ------------------------------------------------------------------G------------------CC-------------------------------------------------------------------............---- 8752
MAC.US.x.96081.AY597209      ----------------------------------------------------------------------------------------------------------------------------------------------------------............---- 8750
MAC.US.x.96093.AY611489      -------------------------------------------------G----------------A---------------------------------------------------------------------------------------............---- 8748
MAC.US.x.96114.AY588945      ----------------------------------------------------------------------------------------------------------------------------------------------------------............R--- 8743
MAC.US.x.96123.AY611487      ----------------------------------------------------------------------------------------------------------------------------------------------------------............---- 8637
MAC.US.x.96135.AY607702      ----------------------------------------------------------------------------------------------------------------------------------------------------------............---- 8733
MAC.US.x.97009.AY599199      ----------------------------------------------------------------------------------------------------------------------------------------------------------............---- 8616
MAC.US.x.97074.AY599198      --------------------------------------------------------------------------------------C-------------------------------------------------------------------............---- 8610
MAC.US.x.r80025.AY576480     ----------------------------------------------------------------------------------------------------------------------------------------------------------............R--- 8604
MAC.US.x.r90131.AY576481     ----------------------------------------------------------------------------------------------------------------------------------------------------------............---- 8747
```

```
                                                
```

```
                                                
```

```
                                                                                  TAR element start
                                                     TATA box          3' LTR U3 end 3' LTR repeat start
MAC.US.x.239               ACTGGGGAGGAGCCGGTCGGGAACGCCCA.CTTTCTTGATGTATAAATATCACTGCA.TTTCGCTCTGTA.TTCAGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCTCCAGCACTAG.CAGGTAGAGCCTGGGTGTTCCCTGCTAGACTCTCAC.C10073
H2A.CI.88.UC2.U38293       --AT---------T---G---------T.-A-A---AC---------G-AC-C--T.-CC---A-----.--T--C----------------------C--------A-------------------.-C---------C-----------------------.-10107
H2A.DE.x.BEN.M30502        --AT---------T---G---------T.-A-A---AC---------G-AC-C--T.-C-T--AT----.----------------------------C--------A-------------------.-------------------------G---------.-10153
H2A.DE.x.PEI2.U22047       --GT-------A-----G---------T.-A-A---.C---------G-AC-C--T.GC-T--AT----.----------------------------C-------------------------------.-----------------------------------.-10088
H2A.GH.x.GH1.M30895        --AT---------T---G---------T.-A-A---AC---------G-ACTC--T.-C-TA-AT----.--............................................................................................ 9480
H2A.IN.NNVA.NEW            --AT-------------G---------T.-A-ATC-TC---------G-AC-C--T.GC-T--AT----C----------------------------C---------AG.....--CT----.T---.------------------...TA------------.-10130
H2A.IN.95.CRIK             --AT---------T---G---------T.-A-ATC-TC---------G-AC-C--T.---GCA-TGTAC.-------------------------------------------------------.-------------------------G---------.- 9645
H2A.GM.x.MCR35.AY509260    G-AT------G------G---------T.-A-A---TC--------G--AC-C--T.GC-T--AT-A--C-------------------------------------------------------.-----------------------------------.- 9466
H2A.GM.87.D194.J04542      --AT---------T---G---------T.-A-AT-CTC---------G-AC-C--T.-C-T--AT----.---........................................................................................... 9472
H2A.GM.x.ISY.J04498        --AT---------T---G---------T.-A-A---TC---------G-AC-C--T.GC----AT----.-------------------------------------------------------.-------------------------G---------.- 9551
H2A.GM.x.MCN13.AY509259    --AT-------------G---------T.-A-A---TC--------G--AC-C--T.GC-T--AT----C-------------------------------------------------------.-----------------------------------.- 9591
H2A.GW.86.FG.J03654        --AT---------T---G---------TT-A-A---TC---------G-AC-C--T.GC-T--AT----.-------------.-C----------------------------------------.-----------------------------------.- 9347
H2A.GW.87.CAM2CG.D00835    -TAT---------T---G---------T.-A-A-C-TC---------TCAC-C--T.GC-T--AT----C-------------------------------------------------------.-----------------------------------.-10167
H2A.GW.x.ALI.AF082339      --AT---------T---G---------T.-A-A--C.C---------G-AC-C--T.GC-T--AT----.-------------------------------------------------------.-----------------------------------.-10148
H2A.GW.x.MDS.Z48731        --AT---------T---G---------T.-A-A-C-TC---------G-AC-C--T.GC-T--AT----C---.......................................................................................... 9525
H2A.SN.85.ROD.M15390       --AT---------T---G---------T.-A-AT-CTC-----------AC-C--TAGC-T--AT----C---G------------------------------------------G---.-----------------------------------.- 9591
H2AB.CI.90.7312A.L36874    --AT--------TT---G-----------.T-AAGCCTC---------G-AC-C--T.-C-T--AT----.-------------------------------------------------------.-----------------------------------.-10078
H2B.CI.88.UC1.L07625       --AT--------TT---T-----------.T-AAA-CTC---------G-AC-C--T.-C-T--AT----.----------------------------C--------------------------.-----------------------------------.-10068
H2B.CI.x.EHO.U27200        --AT---------A--.GA----------.TCAAGACTC---------G-AC-C--T.--GT--AT----.----------------------------C--------------------------.-----------------------------------.-10043
H2B.GH.86.D205.X61240      G-AA------G-TT---T-----------.T-AAA-CTC---------G-AC-C--T.-C----AT----.-------------------------------------------------------.-----------------------------------.-10072
H2B.JP.01.KR020.AB100245   --AT---------T---G-----------.T-AAACCTC---------G-AC-C--T.--CTA-AT----.AC--------------------------C--------------------------.-----------------------------------.- 9173
H2AB.CM.03CM_510_03.NEW    --AT--A---G-TT---T-----------.T-AAA-CTC---------G-AC-C--T.-C-T--AT----.AC-----------------------.................................................................. 9089
H2G.CI.92.ABT96.AF208027   CTAT------GA-T---T---------C..AAAA-CTC-----------AG-----.-------------.---------------------------------C---------------------.-----------------------------------.- 9501
H2U.FR.96.12034.AY530889   G-AA----------.A-G---------T.-AAATGATC---------G-AC-C--T.C--T--AT----.-------------------------------------------------------.-----------------------------------.- 9602
SMM.SL.92.SL92B.AF334679   T-----CG--G-GAA.CG--A-CGC---T.AAACTC--TA----------AC-C--TT-GC---------.------------A----------C-----A------C-A---------G------TAGT----.-------------T--A---A-------AC- 9533
MAC.US.x.MAC239            ----------------------------.------------.................................................................................................................. 9028
MAC.US.x.251               ------------------------A----.--------------------------.---T---A----.G------------------------------------------------------.-----------------------------------.-10068
MAC.US.x.251_32H           ------------------------A----.--------------------------.-------------.-------------------------------------------------------.-----------------------------------.-10071
MAC.US.x.251               ..........--------------A----.--------------------------.-------------.-------------------------------------------------------.-----------------------------------.-10043
MAC.US.x.17EC1.AY033233    -----------------------------.--------------------------.-------------.-------------------------------------------------------.-----------------------------------.-10073
MAC.US.x.17EFR.AY033146    -----------------------------.--------------------------.-------------.-------------------------------------------------------.-------------------------G---------.-10073
MAC.US.x.1937.AY611495     -----------------------------.---------.................................................................................................................. 9055
MAC.US.x.2065.AY611493     -----------------------------.----------.................................................................................................................. 9054
MAC.US.x.MM142.Y00277      ----------------T------A----.--------------------------.-------------.-------------------------------------------------------.-----------------------------------.- 9565
MAC.US.x.SMM142B.BD131285-----------------T------A----.--------------------------.-------------.-------------------------------------------------------.-----------------------------------.- 9562
MNE.US.82.MNE              ------------------------A----.--------------------------.-------------.-------------------------------------------------------.-----------------------------------.- 9547
MNE.US.x.MNE027.U79412     -------------T----------A----.--------------------------.-------------.-------------------------------------------------------.-----------------------------------.- 9550
SMM.US.x.F236              -------------T--CT--.---------.---AT-CTC----------CA------.-------------.-------------------------------------------------------.-----------------------------------.-10035
SMM.US.x.H9.M80194         T------------T--CT--.---------.---AT-CTC----------CA------.-------------.-------A-----------------------------------------------.-------------------------G---------.- 9556
SMM.US.x.PBJ14             -------------T--CT--.---------.-G-AT-CTC----------CA------.-------------.-------A-----------------------------------------------.-----------------------------------.- 9881
SMM.US.x.PBJA.M31325       -------------T--CT--.---------.-G-AT-CTC----------CA------.-------------.-------A-----------------------------------------------.-----------------------------------.- 9880
SMM.US.x.PBJ               -------------T--CT--.---------.---AT-CTC----------CA------.-------------.-------A-----------------------------------------------.-----------------------------------.- 9593
SMM.US.x.PBJ               -------------T--CT--.---------.-G-AT-CTC----------CA------.-------------.-------A----A------------------------------------------.-----------------------------------.-10073
SMM.US.x.PGM53.AF077017    -------------T--CT--.---------.---AT--TC----------CA------.-------------.-------------------------------------------------------.-----------------------------------.-10010
SMM.US.x.SME543.U72748     -------------T--CT--.---------.---AT-CTC----------CA------.-------------.G------------------------------------------------------.-----------------------------------.-10108
STM.US.x.STM.M83293        ------A.---A-T---G---------T.T--A---TC----------CA--C--T.-A-A--------.-------A-----------------------------------------------.-----------------------------------.- 9727
MAC.US.x.80035.AY611486    -----------------------------.----------.................................................................................................................. 9054
MAC.US.x.81035.AY599200    -----------------------------.----------.................................................................................................................. 9055
MAC.US.x.85013.AY611490    -----------------------------.---------.................................................................................................................. 8987
MAC.US.x.92050.AY603959    -----------------------------.----------.................................................................................................................. 9048
MAC.US.x.92077.AY599201    -----------------------------.----------.................................................................................................................. 9055
MAC.US.x.93057.AY611492    -----------------------------.---------.................................................................................................................. 9052
MAC.US.x.93062.AY607704    -----------------------------.---------.................................................................................................................. 9021
MAC.US.x.95058.AY611494    -----------------------------.----------.................................................................................................................. 9042
MAC.US.x.95086.AY607703    -----------------------------.---------.................................................................................................................. 8910
MAC.US.x.95112.AY588946    -----------------------------.----------.................................................................................................................. 9044
MAC.US.x.96016.AY607701    -----------------------------.---------.................................................................................................................. 9054
MAC.US.x.96020.AY611488    -----------------------------.----------.................................................................................................................. 9047
MAC.US.x.96072.AY611491    -----------------------------.--------.................................................................................................................. 9057
MAC.US.x.96081.AY597209    ----------------------------.............................................................................................................................. 9045
MAC.US.x.96093.AY611489    -----------------------------.---------.................................................................................................................. 9054
MAC.US.x.96114.AY588945    -----------------------------.---------.................................................................................................................. 9049
MAC.US.x.96123.AY611487    -----------------------------.----------.................................................................................................................. 8944
MAC.US.x.96135.AY607702    -----------------------------.---------.................................................................................................................. 9039
MAC.US.x.97009.AY599199    .......................................................................................................................................................... 8867
MAC.US.x.97074.AY599198    .......................................................................................................................................................... 8859
MAC.US.x.r80025.AY576480   -----------------------------.---------.................................................................................................................. 8910
MAC.US.x.r90131.AY576481   -----------------------------.---------.................................................................................................................. 9053
```

```
                                                    TAR element end                          Poly-A signal  3' LTR R repeat end  3' LTR U5 start
MAC.US.x.239                AGCACTTGGCCGGTGCTGGGCAGAGTGACTCCACGCTTGCTTGCTTAAAG..CCCTCTTCAATAAAG.CTGCCATTTTAGAAGTAAGCTAGTGTGTGTTCCCATCTCTCCTAGCCGCCGCCTGGTCAACTCGGTACTCAATAA..TAAGAAGACCCTGGTCTGTTAGGAC10238
H2A.CI.88.UC2.U38293        --TG--C-----CA---------.C-G-----------------------..A------.-------.------A--AGA-GCA-GTTA------------------------T-------------TTCG-TGTTCATC-G-G.---C---------------------10271
H2A.DE.x.BEN.M30502         --T----------A---------.C-G-----------------------..A------.-------.------G--AGA-GCA-GTTA------------------------T-------------TTCG-TGTTCATC-G-G.---C---------------------10318
H2A.DE.x.PEI2.U22047        --T----------A---------.C-G-----------------------..A------.-------.------G--AGA-GCA-GTTA.................................................................................10172
H2A.GH.x.GH1.M30895         .......................................................................................................................................................................... 9480
H2A.IN.NNVA.NEW             --TG--------C----------.C-G-----------------------..A------.-------.------G--AGA-GCA-GT-A------------------------T-------------GTCG-TGTTCACC-G-.G---C---------------------10295
H2A.IN.95.CRIK              --TG--------CA---------.C-------------------------..A------.-------.------G--AGA-GCA-GTTA-------------------T----T-------------TTCG-TGTTCATC-G-A.---C--------------------- 9809
H2A.GM.x.MCR35.AY509260     G-TG--------CA---------.C-G----------------------AG.A------.-------.------G--AGA-GCA-GTTA------------------------T------------.................................................... 9588
H2A.GM.87.D194.J04542       .......................................................................................................................................................................... 9472
H2A.GM.x.ISY.J04498         --TG--------C----------.C-G----------------------AG.A------.-------.------G--AGA-GCA-GTTA................................................................................. 9636
H2A.GM.x.MCN13.AY509259     G-TG--------CA---------.C-G----------------------AG.A------.-------.------G--AGA-GCA-GTTA------------------------T------------.................................................... 9713
H2A.GW.86.FG.J03654         --TG--------C----------.C-G-----------------------..A-----C.-------.------A--AGA-GCA-GTTA................................................................................. 9431
H2A.GW.87.CAM2CG.D00835     --TG--------CA---------.C-G-----------------------..A------.-------.------G--AGA-GCA-GTTA------------------------T-------------TTCG-TGTTCACC-G-A.---C---------------------10331
H2A.GW.x.ALI.AF082339       --TG--------CA---------.C-G-----------------------..A------.-------.------A--AGA-GCAGGTTA-AG---------------------T-------------TTCG-TGTTCACC-G-G.---C---------------------10312
H2A.GW.x.MDS.Z48731         .......................................................................................................................................................................... 9525
H2A.SN.85.ROD.M15390        -----------------------.C-G-C--------------------A.A--TC---.-------.------G--AGA-GCA...................................................................................... 9671
H2AB.CI.90.7312A.L36874     -----------------------.C-G----------------------A..-------.-------.------A--AGA-GCA-GT-A------------------------T-------------.------GT---GCTGAA-.-A-G-------------------10241
H2B.CI.88.UC1.L07625        -----------------------.C-G----------------------A..-------.------..------A--AGA-GCA-GTTG------------------------T-------------.T--AA-G---T-CT-AAAT.A---------------------10230
H2B.CI.x.EHO.U27200         -----------------------.C-G----------------------A..-------.------A.-----TA--AGA-GCA-GA-A------------------------T-------------.T-----G---C-CT-GA......A------------------10201
H2B.GH.86.D205.X61240       ------A---A---T--------.C-G----------------------A..-------.------..------A--AGA-GCA-GT.A------------------------T--.----------.------G---C-CTGATA.....-------------------10228
H2B.JP.01.KR020.AB100245    -----------------------.C-G----------------------A..-------.------..-----TA--AGA-GCA-GTTG------------------------T-------------.------G-----CTGAGAT.A--------------------- 9335
H2AB.CM.03CM_510_03.NEW     .......................................................................................................................................................................... 9089
H2G.CI.92.ABT96.AF208027    ------------------------C-G--------------------C-ACGA------.-------.------AA------------A---------GT....................................................................... 9599
H2U.FR.96.12034.AY530889    -AA-G..................................................................................................................................................................... 9607
SMM.SL.92.SL92B.AF334679    ----ACA---AT-------T---.C-G-------------------G......A----.--------G----A-C--.-----C---TC................................................................................ 9613
MAC.US.x.MAC239             .......................................................................................................................................................................... 9028
MAC.US.x.251                --------------------------------------------------..----------------.---------------------------------------------T------------------------GG---..-------------------------10233
MAC.US.x.251_32H            --------------------------G-----------------------..A--------------.---------------------------------------------T------------------------GG---..-------------------------10236
MAC.US.x.251                ----------A---------------G-----------------------..A--------------.--------------------C------------------------T------------------------GG---..-------------------------10208
MAC.US.x.17EC1.AY033233     --------------------------------------------------..----------------.---------------------------------------------------------------------------..-------------------------10238
MAC.US.x.17EFR.AY033146     --------------------------------------------------..----------------.---------------------------------------------------------------------------..-------------------------10238
MAC.US.x.1937.AY611495      .......................................................................................................................................................................... 9055
MAC.US.x.2065.AY611493      .......................................................................................................................................................................... 9054
MAC.US.x.MM142.Y00277       --------------------------G-----------------------..A--------------.---------.-------............................................................................... 9646
MAC.US.x.SMM142B.BD131285---------------------------G-----------...................................................................................................................................... 9600
MNE.US.82.MNE               --------------------------G-----------------------..A--------------.-----T---.-------............................................................................... 9628
MNE.US.x.MNE027.U79412      -----------------T--------G-----------------------..A--------------.---------.----------C--------C---------------T----................................................... 9664
SMM.US.x.F236               --------------------------G-----------------------..A--------------.------A--.----------A------------------------T-------------T----------G-C-.CA-------------------------10200
SMM.US.x.H9.M80194          -------------------------RG-----------------------..A--------------G------G--.-------............................................................................... 9638
SMM.US.x.PBJ14              --------------------------G-----------------------..A--------------.---------.----------C------------------------T-----................................................... 9996
SMM.US.x.PBJA.M31325        --------------------------G-----------------------..A-------------AG---------.----------C------------------------T-----................................................... 9996
SMM.US.x.PBJ                -------------------------RG-----------------------..A-------------AG---------.-------............................................................................... 9675
SMM.US.x.PBJ                --------------------------G-----------------------..A--------------.---------.----------C--------------------------------------T----------G-A.CAT.------C-----------------10237
SMM.US.x.PGM53.AF077017     --------------------------G-----------------------..A--------------.-----GC--.----------A------------------------T-------------T----------G-ATCAT.------------------------10175
SMM.US.x.SME543.U72748      --------------------------G-----------------------..A--------------.---------.----------A------------------------T-------------T----------G-C-CATA------------------------10274
STM.US.x.STM.M83293         ----------A--------T------G--------------------T-T..AT-------------.------A-------------A----G-------------------T------------------------...........--------------------- 9883
MAC.US.x.80035.AY611486     .......................................................................................................................................................................... 9054
MAC.US.x.81035.AY599200     .......................................................................................................................................................................... 9055
MAC.US.x.85013.AY611490     .......................................................................................................................................................................... 8987
MAC.US.x.92050.AY603959     .......................................................................................................................................................................... 9048
MAC.US.x.92077.AY599201     .......................................................................................................................................................................... 9055
MAC.US.x.93057.AY611492     .......................................................................................................................................................................... 9052
MAC.US.x.93062.AY607704     .......................................................................................................................................................................... 9021
MAC.US.x.95058.AY611494     .......................................................................................................................................................................... 9042
MAC.US.x.95086.AY607703     .......................................................................................................................................................................... 8910
MAC.US.x.95112.AY588946     .......................................................................................................................................................................... 9044
MAC.US.x.96016.AY607701     .......................................................................................................................................................................... 9054
MAC.US.x.96020.AY611488     .......................................................................................................................................................................... 9047
MAC.US.x.96072.AY611491     .......................................................................................................................................................................... 9057
MAC.US.x.96081.AY597209     .......................................................................................................................................................................... 9045
MAC.US.x.96093.AY611489     .......................................................................................................................................................................... 9054
MAC.US.x.96114.AY588945     .......................................................................................................................................................................... 9049
MAC.US.x.96123.AY611487     .......................................................................................................................................................................... 8944
MAC.US.x.96135.AY607702     .......................................................................................................................................................................... 9039
MAC.US.x.97009.AY599199     .......................................................................................................................................................................... 8867
MAC.US.x.97074.AY599198     .......................................................................................................................................................................... 8859
MAC.US.x.r80025.AY576480    .......................................................................................................................................................................... 8910
MAC.US.x.r90131.AY576481    .......................................................................................................................................................................... 9053
```

```
                                                 3' LTR U5 end
MAC.US.x.239                CCTTTCTGCTTTGGGAAACC.GAAGCAGGAAAATCCCTAGCA  10279
H2A.CI.88.UC2.U38293        ----C-C----------T--.A----G---------------  10312
H2A.DE.x.BEN.M30502         ----CTC----------T--.A-G-----------------  10359
H2A.DE.x.PEI2.U22047        ........................................  10172
H2A.GH.x.GH1.M30895         ........................................   9480
H2A.IN.NNVA.NEW             ----CT----------T--.A-G--T--------------  10336
H2A.IN.95.CRIK              ----CT--------------C--G--------------T--   9851
H2A.GM.x.MCR35.AY509260     ........................................   9588
H2A.GM.87.D194.J04542       ........................................   9472
H2A.GM.x.ISY.J04498         ........................................   9636
H2A.GM.x.MCN13.AY509259     ........................................   9713
H2A.GW.86.FG.J03654         ........................................   9431
H2A.GW.87.CAM2CG.D00835     ----CT--------------.A-G-----------------  10372
H2A.GW.x.ALI.AF082339       ------------------G.--G-----------------  10353
H2A.GW.x.MDS.Z48731         ........................................   9525
H2A.SN.85.ROD.M15390        ........................................   9671
H2AB.CI.90.7312A.L36874     ----CT--------------.T-------------------  10282
H2B.CI.88.UC1.L07625        --------------------.A-G-----------------  10271
H2B.CI.x.EHO.U27200         --C-----------------.A-G-----------------  10242
H2B.GH.86.D205.X61240       ----CT--------------.-----G--------------  10269
H2B.JP.01.KR020.AB100245    ----....................................   9339
H2AB.CM.03CM_510_03.NEW     ........................................   9089
H2G.CI.92.ABT96.AF208027    ........................................   9599
H2U.FR.96.12034.AY530889    ........................................   9607
SMM.SL.92.SL92B.AF334679    ........................................   9613
MAC.US.x.MAC239             ........................................   9028
MAC.US.x.251                --------------------.--------------------  10274
MAC.US.x.251_32H            --------------------.--------------------  10277
MAC.US.x.251                -------------A------.--------------------  10249
MAC.US.x.17EC1.AY033233     --------------------.--------------------  10279
MAC.US.x.17EFR.AY033146     --------------------.--------------------  10279
MAC.US.x.1937.AY611495      ........................................   9055
MAC.US.x.2065.AY611493      ........................................   9054
MAC.US.x.MM142.Y00277       ........................................   9646
MAC.US.x.SMM142B.BD131285.  ........................................   9600
MNE.US.82.MNE               ........................................   9628
MNE.US.x.MNE027.U79412      ........................................   9664
SMM.US.x.F236               --------------A-----.--G-----------------  10241
SMM.US.x.H9.M80194          ........................................   9638
SMM.US.x.PBJ14              ........................................   9996
SMM.US.x.PBJA.M31325        ........................................   9996
SMM.US.x.PBJ                ........................................   9675
SMM.US.x.PBJ                ---------...............................  10246
SMM.US.x.PGM53.AF077017     ---------...............................  10184
SMM.US.x.SME543.U72748      --------------------.--G-----------------  10315
STM.US.x.STM.M83293         ---------...............................   9892
MAC.US.x.80035.AY611486     ........................................   9054
MAC.US.x.81035.AY599200     ........................................   9055
MAC.US.x.85013.AY611490     ........................................   8987
MAC.US.x.92050.AY603959     ........................................   9048
MAC.US.x.92077.AY599201     ........................................   9055
MAC.US.x.93057.AY611492     ........................................   9052
MAC.US.x.93062.AY607704     ........................................   9021
MAC.US.x.95058.AY611494     ........................................   9042
MAC.US.x.95086.AY607703     ........................................   8910
MAC.US.x.95112.AY588946     ........................................   9044
MAC.US.x.96016.AY607701     ........................................   9054
MAC.US.x.96020.AY611488     ........................................   9047
MAC.US.x.96072.AY611491     ........................................   9057
MAC.US.x.96081.AY597209     ........................................   9045
MAC.US.x.96093.AY611489     ........................................   9054
MAC.US.x.96114.AY588945     ........................................   9049
MAC.US.x.96123.AY611487     ........................................   8944
MAC.US.x.96135.AY607702     ........................................   9039
MAC.US.x.97009.AY599199     ........................................   8867
MAC.US.x.97074.AY599198     ........................................   8859
MAC.US.x.r80025.AY576480    ........................................   8910
MAC.US.x.r90131.AY576481    ........................................   9053
```


HIV-2/SIV Complete Genomes

IV

PLV Complete Genomes

Contents

IV-1	Introduction	**219**
IV-2	Sequences	**220**
IV-3	Alignments	**222**

IV-1 Introduction

Complete genomes of HIV-1, HIV-2, and SIV sequences are represented in the alignment. For the selection, one of each of the major subtypes and a few CRF s of the HIV-1 M group were chosen, plus one N group and 2 O group sequences. Most of the SIV-CPZ s were included, including new sequences from wild caught chimpanzees. Three HIV-2 A group and 2 HIV-2 B group sequences were included along with the single genomes sequenced to date for HIV-2 G and HIV-2 U. For the rest of the SIVs, all complete genomes were included. The principle in deciding which sequences to include in this alignment was to provide a representative sampling of all primate lentiviruses which have had complete or nearly complete genomes sequenced. Many HIV-1, HIV-2, SIVsmm and SIVmac genomes were not included, but they can be found in the HIV-1/SIVcpz and HIV-2/SIVsmm alignments. Also, many shorter sequences from PLVs are included in the PLV protein sequence alignments later in this compendium. Because of the complicated phylogeny, most likely muddled by recombination and many cross-species transmissions, we do not show a tree for the PLV alignment this year. Please refer to the Sequence Compendium of 2005 for a more comprehensive evaluation of the PLV phylogeny.

The alignment was based on the previous version published last year, produced by a hidden Markov model using HMMER and manual editing using the programs BioEdit and Se-Al. It is codon-aligned, meaning that the correct translation reading frame has been maintained as much as possible; in the case of overlapping frames, obviously one will not translate correctly. We have tried to "reset" the alignment so that each new gene starts in the first frame; this means some empty columns have been inserted. The HIV-2/SIVsmm *vpx* gene is postulated to be a duplication of the *vpr* gene and thus there may be two alternative alignments of this region of the genome, as there are for the duplicated stem-loops of the TAR element.

The annotation is based on known protein coding regions in HIV-1 and on annotations found in SIV sequence database entries. The protein cleavage sites that create Gag p17, Gag p24 and other mature peptides from the Gag and Gag-Pol precursor polyproteins have been experimentally determined for HIV-1 and at least one strain of HIV-2, but the study of analogous cleavages in SIV polyproteins have not been published. Two representative genomes have been translated: HIV-1 subtype B strain HXB2, and MAC239. The translations are provided as a visual aid for finding landmarks in the genomes.

IV-2 Sequences

Primate lentivirus (PLV) Sequences in the complete genome alignments.

Name	Accession	Country	Author	Reference
H1B.FR.83.HXB2	K03455	France	van Beveren, C.P.	*(in) Weiss,RL, Teich,N, Varmus,H and Coffin,J(Eds); RNA TUMOR VIRUSES, SECOND EDITION,* **2**, Vol 2: 1102-1123; Cold Spring Harbor Laboratory, Cold Spring Harbor (1985)
H1A1.UG.85.U455	M62320	Uganda	Oram, J.D.	*ARHR* **6**(9):1073-1078 (1990)
H1B.US.90.WEAU160	U21135	United States	Wei, X.	*Nature* **422**(6929):307-312 (2003)
H1C.ET.86.ETH2220	U46016	Ethiopia	Salminen, M.O.	*ARHR* **12**(14):1329-1339 (1996)
H1D.CD.84.84ZR085	U88822	D.R.C.	Gao, F.	*J Virol* **72**(7):5680-5698 (1998)
H1F1.BE.93.VI850	AF077336	Belgium	Laukkanen, T.	*Virology* **269**(1):95-104 (2000)
H1G.SE.93.SE6165	AF061642	Sweden	Carr, J.K.	*Virology* **247**(1):22-31 (1998)
H1H.CF.90.056	AF005496	C.A.R.	Gao, F.	*J Virol* **72**(7):5680-5698 (1998)
H1J.SE.93.SE7887	AF082394	Sweden	Laukkanen, T.	*ARHR* **15**(3):293-297 (1999)
H1K.CM.96.MP535	AJ249239	Cameroon	Triques, K.	*ARHR* **16**(2):139-151 (2000)
H101_AE.TH.90.CM240	U54771	Thailand	Carr, J.K.	*J Virol* **70**(9):5935-5943 (1996)
H102_AG.NG.x.IBNG	L39106	Nigeria	Howard, T.M.	*ARHR* **10**(12):1755-1757 (1994)
H1N.CM.95.YBF30	AJ006022	Cameroon	Simon, F.	*Nat Med* **4**(9):1032-1037 (1998)
H1O.BE.87.ANT70	L20587	Belgium	Vanden Haesevelde, M.	*J Virol* **68**(3):1586-1596 (1994)
H1O.CM.91.MVP5180	L20571	Cameroon	Gurtler, L.G.	*J Virol* **68**(3):1581-1585 (1994)
CPZ.CD.90.ANT	U42720	D.R.C.	Vanden Haesevelde, M.M.	*Virology* **221**(2):346-350 (1996)
CPZ.CM.01.SIVcpzCAM13	AY169968	Cameroon	Nerrienet, E.	*J Virol* **79**(2):1312-9 (2005)
CPZ.CM.05.SIVcpzEK505	DQ373065	Cameroon	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.CM.05.SIVcpzMB66	DQ373063	Cameroon	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.CM.98.CAM3	AF115393	Cameroon	Corbet, S.	*J Virol* **74**(1):529-534 (2000)
CPZ.GA.88.GAB2	AF382828	Gabon	Bibollet-Ruche, F.	*ARHR* **20**(12):1377-1381 (2004)
CPZ.CM.98.CAM5	AJ271369	Cameroon	Muller-Trutwin, M.C.	*J Med Primatol* **29**(3-4); 166-72 (2000)
CPZ.CM.05.SIVcpzMT145	DQ373066	Cameroon	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.TZ.00.TAN1	AF447763	Tanzania	Santiago, M.L.	*J Virol* **77**(3):2233-2242 (2003)
CPZ.CM.05.SIVcpzLB7	DQ373064	Cameroon	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.US.85.CPZUS	AF103818	United States	Gao, F.	*Nature* **397**(6718):436-441 (1999)
CPZ.GA.88.GAB1	X52154	Gabon	Huet, T.	*Nature* **345**(6273):356-359 (1990)
MON.CM.99.L1	AY340701	Cameroon	Courgnaud, V.	*J Virol* **77**(23):12523-12534 (2003)
MON.NG.x.NG1	AJ549283	Nigeria	Barlow, K.L.	*J Virol* **77**(12):6879-6888 (2003)
GSN.CM.99.CN166	AF468659	Cameroon	Courgnaud, V.	*J Virol* **76**(16):8298-8309 (2002)
GSN.CM.99.CN71	AF468658	Cameroon	Courgnaud, V.	*J Virol* **76**(16):8298-8309 (2002)
MUS_1.CM.01.1085	AY340700	Cameroon	Courgnaud, V.	*J Virol* **77**(23):12523-12534 (2003)
MUS_1.CM.01.CM1239	EF070330	Cameroon	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
MUS_2.CM.01.CM1246	EF070329	Cameroon	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
MUS_2.CM.01.CM2500	EF070331	Cameroon	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
RCM.NG.x.NG411	AF349680	Nigeria	Beer, B.E.	*J Virol* **75**(24):12014-12027 (2001)
RCM.GA.x.GAB1	AF382829	Gabon	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
DEN.CD.x.CD1	AJ580407	D.R.C.	Dazza, M.C.	*J Virol* **79**(13):8560-8571 (2005)
MAC.US.x.239	M33262	United States	Kestler, H.	*Science* **248**(4959):1109-1112 (1990)

Name	Accession	Country	Author	Reference
H2A.GW.x.ALI	AF082339	Guinea-Bissau	Azevedo-Pereira, J.M.	Unpublished
H2A.DE.x.BEN	M30502	Germany	Kirchhoff, F.	*Virology* **177**(1):305-311 (1990)
H2A.SN.x.ST	M31113	Senegal	Kumar, P.	*J Virol* **64**(2):890-901 (1990)
H2B.GH.86.D205	X61240	Ghana	Dietrich, U.	*Nature* **342**(6252):948-950 (1989)
H2B.CI.x.EHO	U27200	Cote D'Ivoire	Rey-Cuille, M.A.	*Virology* **202**(1):471-476 (1994)
H2G.CI.92.ABT96	AF208027	Cote D'Ivoire	Brennan, C.A.	*ARHR* **13**(5):401-404 (1997)
H2U.FR.96.12034	AY530889	France	Damond, F.	*ARHR* **20**(6):666-672 (2004)
MAC.US.x.EMBL_3	Y00295	United States	Franchini, G.	*Nature* **328**(6130):539-543 (1987)
SMM.US.x.H9	M80194	United States	Courgnaud, V.	*J Virol* **66**(1):414-419 (1992)
SMM.US.x.PGM53	AF077017	United States	Novembre, F.J.	*J Virol* **72**(11):8841-8851 (1998)
SMM.US.x.SIVsmH635F_L3	DQ201172	United States	Kuwata, T.	*J Virol* **80**(3):1463-75 (2006)
SMM.SL.92.SL92B	AF334679	Sierra Leone	Chen, Z.	*J Virol* **70**(6):3617-3627 (1996)
STM.US.x.STM	M83293	United States	Novembre, F.J.	*Virology* **186**(2):783-787 (1992)
MNE.US.x.MNE027	U79412	United States	Kimata, J.T.	*J Virol* **72**(1):245-256 (1998)
DRL.x.x.FAO	AY159321		Hu, J.	*J Virol* **77**(8):4867-4880 (2003)
MND_2.x.x.5440	AY159322		Hu, J.	*J Virol* **77**(8):4867-4880 (2003)
MND_2.CM.98.CM16	AF367411	Cameroon	Takehisa, J.	*ARHR* **17**(12):1143-1154 (2001)
MND_2.GA.x.M14	AF328295	Gabon	Souquiere, S.	*J Virol* **75**(15):7086-7096 (2001)
OLC.CI.97.97CI12	FM165200	Cote D'Ivoire	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
GRV.ET.x.GRI_677	M66437	Ethiopia	Fomsgaard, A.	*Virology* **182**(1):397-402 (1991)
VER.KE.x.TYO1_patent	DJ048201	Kenya	Omori, T.	Patent: WO 2007049749-A 13 03-MAY-2007; DNAVEC CORPORATION
VER.KE.x.9063	L40990	Kenya	Hirsch, V.M.	*J Virol* **69**(2):955-967 (1995)
VER.KE.x.AGM155	M29975	Kenya	Johnson, P.R.	*J Virol* **64**(3):1086-1092 (1990)
VER.DE.x.AGM3	M30931	Germany	Baier, M.	*Virology* **176**(1):216-221 (1990)
TAN.UG.x.TAN1	U58991	Uganda	Soares, M.A.	*Virology* **228**(2):394-399 (1997)
SAB.SN.x.SAB1C	U04005	Senegal	Jin, M.J.	*EMBO J* **13**(12):2935-2947 (1994)
MND_1.GA.x.MNDGB1	M27470	Gabon	Tsujimoto, H.	*Nature* **341**(6242):539-541 (1989)
DEB.CM.99.CM40	AY523865	Cameroon	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
DEB.CM.99.CM5	AY523866	Cameroon	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
TAL.CM.01.8023	AM182197	Cameroon	Liegeois, F.	*Virology* **349**(1):55-65 (2006)
TAL.CM.00.266	AY655744	Cameroon	Liegeois, F.	*Virology* **349**(1):55-65 (2006)
SUN.GA.98.L14	AF131870	Gabon	Beer, B.E.	*J Virol* **73**(9):7734-7744 (1999)
WRC.CI.97.97CI14	AM745105	Cote D'Ivoire	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
WRC.CI.98.98CI04	AM713177	Cote D'Ivoire	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
WRC.GM.05.Pbt_05GM_X02	AM937062	Gambia	Locatelli, S.	*Virology* **376**(1):90-100 (2008)
LST.CD.88.524	AF188116	D.R.C.	Beer, B.E.	*J Virol* **74**(8):3892-3898 (2000)
LST.CD.88.447	AF188114	D.R.C.	Beer, B.E.	*J Virol* **74**(8):3892-3898 (2000)
LST.KE.x.lho7	AF075269	Kenya	Hirsch, V.M.	*J Virol* **73**(2):1036-1045 (1999)
SYK.KE.x.SYK173	L06042	Kenya	Hirsch, V.M.	*J Virol* **67**(3):1517-1528 (1993)
SYK.KE.x.KE51	AY523867	Kenya	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
COL.CM.x.CGU1	AF301156	Cameroon	Courgnaud, V.	*J Virol* **75**(2):857-866 (2001)

PLV Complete Genomes

```
                         5' LTR U3 start
H1B.FR.83.HXB2           .TGGAAGGGCTAATTCACTCCCAACGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCA...GGGCCAGGG...ATCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGA 163
H1A1.UG.85.U455          .......................................................................................................................................................................... 0
H1B.US.90.WEAU160        .--------------T--------AA------------------------G----------------T----------G---------------...---------...-CT-----C------TG-------------T--------------G--A----- 163
H1C.ET.86.ETH2220        .......................................................................................................................................................................... 0
H1D.CD.84.84ZR085        .......................................................................................................................................................................... 0
H1F1.BE.93.VI850         .......................................................................................................................................................................... 0
H1G.SE.93.SE6165         .......................................................................................................................................................................... 0
H1H.CF.90.056            .......................................................................................................................................................................... 0
H1J.SE.93.SE7887         .......................................................................................................................................................................... 0
H1K.CM.96.MP535          .......................................................................................................................................................................... 0
H101_AE.TH.90.CM240      .......................................................................................................................................................................... 0
H102_AG.NG.x.IBNG        .......................................................................................................................................................................... 0
H1N.CM.95.YBF30          .......................................................................................................................................................................... 0
H1O.BE.87.ANT70          C-------T-----T-------TAA---GC---A-----G-----T---G-G--TA---T--G--A-T-----------G----------------...--A-----A...-C---G-TC--------A----------TG-TT--A-----------GTCAGA--- 164
H1O.CM.91.MVP5180        C----T---T-----T-------TAAG---GC---A-----G-----C-----A--T-----T--G--A-T-----------G----TGT--------G...--A-----A...CCT----TC--------A----------TG-TT--A--G--------GTCAG---- 164
CPZ.CD.90.ANT            .......................................................................................................................................................................... 0
CPZ.CM.01.SIVcpzCAM13    .......................................................................................................................................................................... 0
CPZ.CM.05.SIVcpzEK505    .......................................................................................................................................................................... 0
CPZ.CM.05.SIVcpzMB66     .......................................................................................................................................................................... 0
CPZ.CM.98.CAM3           .......................................................................................................................................................................... 0
CPZ.GA.88.GAB2           .......................................................................................................................................................................... 0
CPZ.CM.98.CAM5           .......................................................................................................................................................................... 0
CPZ.CM.05.SIVcpzMT145    .......................................................................................................................................................................... 0
CPZ.TZ.00.TAN1           .......................................................................................................................................................................... 0
CPZ.CM.05.SIVcpzLB7      .......................................................................................................................................................................... 0
CPZ.US.85.CPZUS          .--------T--G--T-T---AG-A----------G-----A-----C---G-T-----T--------AAT---------C-G------T-----C--...-----T-A...G-A--------C--A--A-A---GC------TT---T-----C--TC-CACAGA--- 163
CPZ.GA.88.GAB1           .--------T--G--T-----AGGA----------G--------C--C---G----T-------------T---------C-G-------------A--...--A-----A...-CA----TC-----CTGT-------------T---A-----G--CC-GACAGAG-- 163
MON.CM.99.L1             .......................................................................................................................................................................... 0
MON.NG.x.NG1             .......................................................................................................................................................................... 0
GSN.CM.99.CN166          .......................................................................................................................................................................... 0
GSN.CM.99.CN71           .......................................................................................................................................................................... 0
MUS_1.CM.01.1085         .......................................................................................................................................................................... 0
MUS_1.CM.01.CM1239       .......................................................................................................................................................................... 0
MUS_2.CM.01.CM1246       .......................................................................................................................................................................... 0
MUS_2.CM.01.CM2500       .......................................................................................................................................................................... 0
RCM.NG.x.NG411           .......................................................................................................................................................................... 0
RCM.GA.x.GAB1            .......................................................................................................................................................................... 0
DEN.CD.x.CD1             .......................................................................................................................................................................... 0
                         5' LTR U3 start
MAC.US.x.239             .--------A-TTA-T--AGTGC-A-------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-A-A--A----G----G-T-----CT--...--A-----A...--T-----C---AA---A-----C---CTA-GG--AT----C--T--AA-T..... 158
H2A.GW.x.ALI             .--------A-GT--T--AGTG-GA-------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G----GATAA-T--A----G---------T--T-AT...--------A...--A--G--C--GA--TT------G---CTG-GG--------------A--T..... 158
H2A.DE.x.BEN             .--C-----A-GT--T--AGTAGGA-G-----TAGA-----A--CA-A-ACC-AG-AA-AGAGG----GATAA-A--A----G------T--T--T-AT...--------A...G-A--G--C---A--TA---C--G---CTG-GG---------T----A--A..... 158
H2A.SN.x.ST              .......................................................................................................................................................................... 0
H2B.GH.86.D205           .--------A-GTA-T-TAGTG-GA-------CAGA--AT-A--CACA-ATT-TG-GA-TGA-G-----ATTG-GT---GC-G---A-----T--T-AT...----------...--A--GC----CAAATA------T---CTG-GG-----G--------A---...-T 160
H2B.CI.x.EHO             .--------A-TT-CT-TAGTG-GA----G--TA-A--A--A--CACA-ACT-AG-AA-TGA-G-----ATTG-GT---GA-G-------------AT...--------A...G-A--------CAA-TT------C---CTA-GG--------G---A-AA-T..... 158
H2G.CI.92.ABT96          .......................................................................................................................................................................... 0
H2U.FR.96.12034          .......................................................................................................................................................................... 0
MAC.US.x.EMBL_3          .......................................................................................................................................................................... 0
SMM.US.x.H9              .......................................................................................................................................................................... 0
SMM.US.x.PGM53           ..........................................................................--A---G--A--T-----T--...--------A...--------C--TA--TT------C---CTA-GG--AT-G--C-----AAGT..... 84
SMM.US.x.SIVsmH635F_L3   .--------A-TTA-T--AATG-GAA------TAGA--AT-A---A---AC--GG-AA-GGA-G---AATAA-A--A----G--A--T------TT-...--------A...-CT-----C--TA--TA------G---CT--GG--AT----C-----A--T..... 158
SMM.SL.92.SL92B          .......................................................................................................................................................................... 0
STM.US.x.STM             .--------A-TTA-T--AGTG-................................................................................................................................................... 22
MNE.US.x.MNE027          .......................................................................................................................................................................... 0
DRL.x.x.FAO              .......................................................................................................................................................................... 0
MND_2.x.x.5440           .......................................................................................................................................................................... 0
MND_2.CM.98.CM16         .......................................................................................................................................................................... 0
MND_2.GA.x.M14           .......................................................................................................................................................................... 0
OLC.CI.97.97CI12         ......................................................................AA-AGG-A-C-TC---GCAACAGAGGAGGTA-AAGAG--CACGT-A--GAGA--TAA---GCCA--G----TA-GGG-T--TAGGAG---G--A..... 94
GRV.ET.x.GRI_677         .----T---A--TA-T----TG--A-----G--A-G-----GA--T---ATGC--TGA--GAGTGG--AATAA-AGA-----G---AGCT---T-----...--C--G---...--A--G--C--GAGAGT------C-TC----TT--------C-----G--C..... 158
VER.KE.x.TYO1_patent     .----T---A-TTA-T-----G-TA-G---A-TA-G-----GA------ATGCTCTTA-TGA-TGG--GATAA-TGA-----G-A-TGC--GGT-GAAG...--A-----A...--A----TC--TAAATG------G-TC----TT--------G-----G--C..... 158
VER.KE.x.9063            .----T---A-TTA-T-----G--A-G--GA-TAGA-----CA-C----ATGCTCTTA-TGA-TGG--AAT-A-TGA-----G-A-CGC--GGT--A--...--A-----C...G-A--------C-GCTG------C-TT----T------------G--A---..... 158
VER.KE.x.AGM155          .----T---A-TTA-T-----G-TA-G---A-TA-G------A------ATGC-CT-A-TGA-TGG--AAT-A-TGA-----G-A-CGCA-GGT--AA-...--A--T---...--A-----C--GAG-TG------C-TC----T------------G----CC..... 158
VER.DE.x.AGM3            .......................................................................................................................................................................... 0
TAN.UG.x.TAN1            .----G--CA--TA-TGG--A-C-AAG---G-GC-A-----GA----C-ATGC-CTAA-TGA-TGG---ATTA-AGA-----G---AGCA--TT-C--T...--------C...-C---GA-G--C-GATG---C--C-TT----TTG--T----G--G--A--T..... 158
SAB.SN.x.SAB1C           .--------A-TTA-T--AGTG--A-G---A-GA-G---T-G-----C-ATGCTCTTA-TGAGTGG---ATAG-AGA--GA-G------T--T---GAT...--------T...-C---G--C---AA-TG--------------TT--A--G-----T--G--C..... 158
MND_1.GA.x.MNDGB1        .......................................................................................................................................................................... 0
DEB.CM.99.CM40           .......................................................................................................................................................................... 0
DEB.CM.99.CM5            .......................................................................................................................................................................... 0
TAL.CM.01.8023           .......................................................................................................................................................................... 0
TAL.CM.00.266            .......................................................................................................................................................................... 0
SUN.GA.98.L14            .---------CTGGT-----AG-ACC---G----A---T-G---T---ATGCAG-AA-TGAGTGG----T-A-AA---GC-G---AG-T-----CAA-...-----T--A...G-G---------AA-G-------C---CTA-GG-------C------CACC...AT 160
WRC.CI.97.97CI14         .----G-ATA--T--TGG--AGCTA-----G----A--AG-GA-C--C-ATGCAC--A-TGA-TGG--GCTGA-AGA--GC-G---AGCT------GA-...--A------...--A----GG---AA---A-----G-------GG--AT-G--T-----AC-C..... 158
WRC.CI.98.98CI04         .----GTATA--T--TGG--T-CTA-----G----A--AG-GA-C--T-ATGC--T-A-TGA-TGG--ACTGA-AGAA-G--G----GCA------GAG...--A-----C...--AC-C-GG---AA---A-----C-----T-GG--AT----G--T--G--A..... 158
WRC.GM.05.Pbt_05GM_X02   .......................................................................................................................................................................... 0
LST.CD.88.524            .......................................................................................................................................................................... 0
LST.CD.88.447            .......................................................................................................................................................................... 0
LST.KE.x.lho7            .-----------TA-TGG-----GA-------G--C---T-AAT---C-ACTGTG-GA-TGAGTGG--TCT-A-TGGA--C-TCAT----------GAT...--A-----A...-CA-----C--T-----A----------TG-GGC--T---A------G-CT...TG 160
SYK.KE.x.SYK173          ...........................................................................................................---A...-C-C-G--C--GT-A-T-----------TA-GGG-A-----T---AA----..... 55
SYK.KE.x.KE51            .......................................................................................................................................................................... 0
COL.CM.x.CGU1            .......................................................................................................................................................................... 0
```

```
H1B.FR.83.HXB2        GAAGTTAGAAGAA...........GCCAACAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGC............CTGCATGGAATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCACATGGCCCGAGAGCTGCATC  309
H1A1.UG.85.U455       ........................................................................................................................................  0
H1B.US.90.WEAU160     ----G-------------------T--TG------------A---------------A-----------------------G----------AT----A------A---AT-----A----------AAA--------------G-A----------------  309
H1C.ET.86.ETH2220     ........................................................................................................................................  0
H1D.CD.84.84ZR085     ........................................................................................................................................  0
H1F1.BE.93.VI850      ........................................................................................................................................  0
H1G.SE.93.SE6165      ........................................................................................................................................  0
H1H.CF.90.056         ........................................................................................................................................  0
H1J.SE.93.SE7887      ........................................................................................................................................  0
H1K.CM.96.MP535       ........................................................................................................................................  0
H101_AE.TH.90.CM240   ........................................................................................................................................  0
H102_AG.NG.x.IBNG     ........................................................................................................................................  0
H1N.CM.95.YBF30       ........................................................................................................................................  0
H1O.BE.87.ANT70       AG--GC---AG-........CTA-GAG-T-C-T-T----GGG-T-ATC-CC-G--T--A-CAT-T............GCC-----T-T--A--TA-AC-T-A----A-AC-GAT-----A------T--ATCT---GCAACACC--TG-T--TATGATAACT--C-  313
H1O.CM.91.MVP5180     AG--GC---GAG-........CTG-GT--T-C-AAT--AG-TG-T--TC-TC----T--A-CTT-T............AAT-----GCT--G--TG-AC-CG-G--GA-AC--A-A---CA-----T--ATCAT---GC--AACA--T--A----TGC-AAA---C-  313
CPZ.CD.90.ANT         ........................................................................................................................................  0
CPZ.CM.01.SIVcpzCAM13 ........................................................................................................................................  0
CPZ.CM.05.SIVcpzEK505 ........................................................................................................................................  0
CPZ.CM.05.SIVcpzMB66  ........................................................................................................................................  0
CPZ.CM.98.CAM3        ........................................................................................................................................  0
CPZ.GA.88.GAB2        ........................................................................................................................................  0
CPZ.CM.98.CAM5        ........................................................................................................................................  0
CPZ.CM.05.SIVcpzMT145 ........................................................................................................................................  0
CPZ.TZ.00.TAN1        ........................................................................................................................................  0
CPZ.CM.05.SIVcpzLB7   ........................................................................................................................................  0
CPZ.US.85.CPZUS       -G-AG----GCGG...............-----T---------C--A--TAC--C-G----CA--T--CAGCATGGAATGGAAG---A-CAT...........G-C-----C---ATC---CA--------CTGAA--G--TCGGAGA----GA--TAA-----------  309
CPZ.GA.88.GAB1        -C--G-----C--...............-----TG-------T----A-T--C----G--T--CA-AT-TCAGCATGGGATGGAAG---AGGAC............-A---G--C--G-TC---C-C---------A-G--G--C--AAGA--T--T---A----A-AA----  309
MON.CM.99.L1          ........................................................................................................................................  0
MON.NG.x.NG1          ........................................................................................................................................  0
GSN.CM.99.CN166       ........................................................................................................................................  0
GSN.CM.99.CN71        ........................................................................................................................................  0
MUS_1.CM.01.1085      ........................................................................................................................................  0
MUS_1.CM.01.CM1239    ........................................................................................................................................  0
MUS_2.CM.01.CM1246    ........................................................................................................................................  0
MUS_2.CM.01.CM2500    ........................................................................................................................................  0
RCM.NG.x.NG411        ........................................................................................................................................  0
RCM.GA.x.GAB1         ........................................................................................................................................  0
DEN.CD.x.CD1          ........................................................................................................................................  0

MAC.US.x.239          .GTA-C---T--G............--AC-GG-G-AT---G-GCATTAT--AA-G--T--A-CTCAAACTTCCCAGTGGGATG-CCCTTG-............G----G--TC---CA----A------TCCAACT--G--C-ACAC-T-TGA---ATAT-TTAGAT-C-  303
H2A.GW.x.ALI          .GTCCC-C-----............-GGG-GG-CACT----CTCA-T--C--C-------A--ACAAACAAGCAGGCATGATG-CACGCAT...........G-G--GACA----TT-----A------CCTAAG--G--TCA-G--T---AA---TTTATT--A--C-  303
H2A.DE.x.BEN          .CTC-C-C----G............--AG-GG---AT---GC--A-T----AG-------A-CACAAACAAGCAGACATGATG---AGCAT...........G-G--GACA----T----CA------TC-ATG--G--C-A-A-CT---A----TTCACT---T-C-  303
H2A.SN.x.ST           ........................................................................................................................................  0
H2B.GH.86.D205        -CCAGC--CGACC...........CGAG-GG-G-AG--A-C-CATT--C-AA-G----G-CACAGATCTCCTCATGGGATG-CATCCAT...........G-G--GACTC-TATC---CA------TTC--T---G----A-G--T-TG----TTTCA-TAG-TT--  306
H2B.CI.x.EHO          .-T-A---C----............C-AG-GG-C-AG--A-C-CATT-TC--G-G--T--A-CACAGACCTCCGCATGGGATG-CCCCCAC...........GA---GACCC-T-TC---CA------TC--T--------A-G-CT-TG-----TTCAGCAG-TTC-  303
H2G.CI.92.ABT96       ........................................................................................................................................  0
H2U.FR.96.12034       ........................................................................................................................................  0
MAC.US.x.EMBL_3       ........................................................................................................................................  0
SMM.US.x.H9           ........................................................................................................................................  0
SMM.US.x.PGM53        .GTC-C---T---............--TC-AG---AC----CACATTA---AG-G--T--A-CTCAGATATCCCAGTGGGATG--CCCTG...........G----G--TC-G-CA----A------TTCA-AAT----C-ACAGAT-TGA---TTTTATTAGA--C-  229
SMM.US.x.SIVsmH635F_L3 .GTC-C---T---............--C-GG---AC----CACATT--C--G-G--T--A-CACAGACTCATCAGTGGGATG-CCCCTG...........G----G--AC-G-CA----A------TCCAGAAT-G--T-A-AGCT-T-A---ATTTATTAA-T-C-  303
SMM.SL.92.SL92B       ........................................................................................................................................  0
STM.US.x.STM          ........................................................................................................................................  22
MNE.US.x.MNE027       ........................................................................................................................................  0
DRL.x.x.FAO           ........................................................................................................................................  0
MND_2.x.x.5440        ........................................................................................................................................  0
MND_2.CM.98.CM16      ........................................................................................................................................  0
MND_2.GA.x.M14        ........................................................................................................................................  0
OLC.CI.97.97CI12      .CCA--GCCT--G...........A--CTTG---AT----GGGATGC------G--T--A-TCAGACCCCTCAAACA-CTG-G-AGT-A-GG...............C-GG-T---GTC--CA-TCCA-ATAG----..................A-TAC--C--  218
GRV.ET.x.GRI_677      .CT-CAT--G--G...........--ACG---CT-T----GACA-T-TC--A-G--T--A-CACAGATGGGG......GAAG--CCTGAT-GAATAGATC-TG-------C--G-TC----A-------CCGAAGT-G--G..................T-GA-T-C-  291
VER.KE.x.TYO1_patent  .TTACAT--G---.............--AC-A-C-T-T--A-GACATT--C-AG-C--T--A-C-CAGATGGGA......GAAG--CC-GAT-G-AT-AGCC-TG----GA-C--G-T-----A-----TCCTATGT-G---..................ATA-A-T-CG  291
VER.KE.x.9063         .-T-CAT----G............--AG-G-CCT-T----GACATTA----GCG--T----CACAAATAAAA......GA-G-CCCCGAT-GCAT-AGTC-TG----GACC--G-TC----A-------CCTATGT-----..................-T--A-T--G  291
VER.KE.x.AGM155       .CT-CAT--G---...........--AG-A-C-T-T--A-GGCATT-----G-------A-CACAACTGCAT......GAAG-CCCTGAT-G-ATAAATC-TG-----A-A--G-CA----A------TCCAATGT-G--T..................-TT-A-T-CG  291
VER.DE.x.AGM3         ........................................................................................................................................  0
TAN.UG.x.TAN1         .GTC-CCC-G---...........--C-GG-T-AG...-GACA-T-T---C-C-------CCCAGATTGAA......TG-G-AA-CGACCC-TGG.......A---GACA--G-T-----AA---A-TCCA-T-------..................-TT-A-T-CA  282
SAB.SN.x.SAB1C        .CT--C---G--G...........--T--G--TA-T------CA-T--C--C-G--T----CTCAAGTGGCC......TATG-G-ATGAC-CCTGG.......A---GACCC-G-T-----A-------CCA-TA------..................-T-GACT--G  285
MND_1.GA.x.MNDGB1     ........................................................................................................................................  0
DEB.CM.99.CM40        ........................................................................................................................................  0
DEB.CM.99.CM5         ........................................................................................................................................  0
TAL.CM.01.8023        ........................................................................................................................................  0
TAL.CM.00.266         ........................................................................................................................................  0
SUN.GA.98.L14         TG-TGA---CAGG...........-ATCCT--TCATCCTTGTCAAGC-C-AC-G--TT-CAGCCAGCAAGGGGTTAATGAAG-CCCTTG...........G----GAGAC-TATC----CA-----TCC-ACT--G--C..................T-TGATTT-A  288
WRC.CI.97.97CI14      .TTCAGT------TCAGATAATACA-A-CCAT-CTAT----GA-AT-AG--AC-G--T--A-CAT-TGCTGAG...CCAGA-G--CCCTG...........G-----CATC-G-TC----A-------CCTAAG--G--C......TGTGATTTTGT--CAGGAAGA-  306
WRC.CI.98.98CI04      .TTTAGT-----GTCTAATGACAGCT-TC-GT-CTAC-----G-A--AGC-------T--A-GCT--CAAGAA...TCAGA-G--CCTTG...........G-----CAC--G-TC----A-------CCT--A-----C......TGTGACTTTGTG-CAAGTAGA-  306
WRC.GM.05.Pbt_05GM_X02 ........................................................................................................................................  0
LST.CD.88.524         ........................................................................................................................................  0
LST.CD.88.447         ........................................................................................................................................  0
LST.KE.x.lho7         TG-TGAGT-TA-............-ATCC-TC--AT----CACAAT--C-T--G--TT-AAGCCAGCTGGGGGTCCTGGAAG-CCCCTG...........G----GAG-C-TATC---CAC--CA-TCCAATGT-G---..................-T-GACTT-A  288
SYK.KE.x.SYK173       .-TAGAG-G-TGC...........TTGG-GT-T-AG-----T--ACT-C-CC-C--T--A-CCTCT...GGGACAGGGGA-TTC-A-CAT...........G-G--GAACC-CAT-----AC--CA--CC--A---G--C..................T-TACA-CAG  179
SYK.KE.x.KE51         ........................................................................................................................................  0
COL.CM.x.CGU1         ........................................................................................................................................  0
```

```
                                                      TCF-1 alpha
H1B.FR.83.HXB2              CGGAGTACTTCAAGAAC..........................................TGCTGACA..................................... 334
H1A1.UG.85.U455             ..........................................................................................................   0
H1B.US.90.WEAU160           ----------T--AG--..........................................--------...................................... 334
H1C.ET.86.ETH2220           ..........................................................................................................   0
H1D.CD.84.84ZR085           ..........................................................................................................   0
H1F1.BE.93.VI850            ..........................................................................................................   0
H1G.SE.93.SE6165            ..........................................................................................................   0
H1H.CF.90.056               ..........................................................................................................   0
H1J.SE.93.SE7887            ..........................................................................................................   0
H1K.CM.96.MP535             ..........................................................................................................   0
H101_AE.TH.90.CM240         ..........................................................................................................   0
H102_AG.NG.x.IBNG           ..........................................................................................................   0
H1N.CM.95.YBF30             ..........................................................................................................   0
H1O.BE.87.ANT70             -A---CT----C----G..........................................GA--A-A-ACTGCTG.............................. 345
H1O.CM.91.MVP5180           -A---CT----CCC--G.............................................-A--TGACACTGC............................ 343
CPZ.CD.90.ANT               ..........................................................................................................   0
CPZ.CM.01.SIVcpzCAM13       ..........................................................................................................   0
CPZ.CM.05.SIVcpzEK505       ..........................................................................................................   0
CPZ.CM.05.SIVcpzMB66        ..........................................................................................................   0
CPZ.CM.98.CAM3              ..........................................................................................................   0
CPZ.GA.88.GAB2              ..........................................................................................................   0
CPZ.CM.98.CAM5              ..........................................................................................................   0
CPZ.CM.05.SIVcpzMT145       ..........................................................................................................   0
CPZ.TZ.00.TAN1              ..........................................................................................................   0
CPZ.CM.05.SIVcpzLB7         ..........................................................................................................   0
CPZ.US.85.CPZUS             -A---------CG----.............................................---TG...................................... 331
CPZ.GA.88.GAB1              ---------A---AG--.............................................----T...................................... 331
MON.CM.99.L1                ..........................................................................................................   0
MON.NG.x.NG1                ..........................................................................................................   0
GSN.CM.99.CN166             ..........................................................................................................   0
GSN.CM.99.CN71              ..........................................................................................................   0
MUS_1.CM.01.1085            ..........................................................................................................   0
MUS_1.CM.01.CM1239          ..........................................................................................................   0
MUS_2.CM.01.CM1246          ..........................................................................................................   0
MUS_2.CM.01.CM2500          ..........................................................................................................   0
RCM.NG.x.NG411              ..........................................................................................................   0
RCM.GA.x.GAB1               ..........................................................................................................   0
DEN.CD.x.CD1                ..........................................................................................................   0

                                TCF-1 alpha
MAC.US.x.239                -A--AG-G--TGGA-G-AAG...TCAGGCCTGTCAGAGGAAGAG...GTTAGAAGAAGGCTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAAGGAAACTCGC...---.................................... 407
H2A.GW.x.ALI                -A---G-A--TGG-T--AAG...TCAGGCCTGCCAGAAGATGAG...TGGAAGGCAAGACTGAAAGCAAGAGGGATACCATTTAGTAAGAACAGGAACAGC.........---TTTGGTCA........................... 409
H2A.DE.x.BEN                -A--AG-G--TGG-C--AAG...TCAGGATTGCCAGAGAAAGAA...TGGAAGGCAAAACTGAAAGCAAGAGGGATACCATATAGTGAA....................-A---GGAACAACCATACTTGGTCA.............. 411
H2A.SN.x.ST                 ..........................................................................................................   0
H2B.GH.86.D205              -A--AG-G---GG-T--CAG...TCAGGGCTACCAGAGGAGGAG...TGGAAGGCTAGACTAAAAGCAAGAGGGATACCTACAGAT.......................-AGGCAAGAGACAGCAGCATAAACAGGAACT........... 417
H2B.CI.x.EHO                -A---G-G--TGG-T-TCAG...TCAGGAATGCCAGAGAAAGAG...TGGAAGGCTAAACTGAGAGCAAGAGGAATACCTACAGAG.......................-AGACAGGAAATAACAGCACAAAGAGGAACT........... 414
H2G.CI.92.ABT96            ..........................................................................................................   0
H2U.FR.96.12034             ..........................................................................................................   0
MAC.US.x.EMBL_3             ..........................................................................................................   0
SMM.US.x.H9                 ..........................................................................................................   0
SMM.US.x.PGM53              -A--AG-G--TGGT-GTAAG...TCAGGCTTGTCAGGGGAAGAG...GTAAAGAGAAGGCTAACCGCAAGAGGCCTTTTAAAAATGGCTGACAAGAAGGAAACGAGC...---.................................... 333
SMM.US.x.SIVsmH635F_L3      -A--AG-G--TGGT-GTAAG...TCAGGCTTGTCAGAGGAAGAG...GTAAAGAGAAGGCTAACCGCAAGAGGCATTTATAAAATGGCTGACAAGAGGGAAACAAGC...---.................................... 407
SMM.SL.92.SL92B             ..........................................................................................................   0
STM.US.x.STM                ......................................................GAGAAGGCTAACCGCAAGAGGCCTTCTAAAGATGGCTGACAAGAAGGAAACAAGC...---.................................  80
MNE.US.x.MNE027             ..........................................................................................................   0
DRL.x.x.FAO                 ..........................................................................................................   0
MND_2.x.x.5440              ..........................................................................................................   0
MND_2.CM.98.CM16            ..........................................................................................................   0
MND_2.GA.x.M14              ..........................................................................................................   0
OLC.CI.97.97CI12            TTA----AAGG...-CATGTTCAGACTTGCATAGGATG............TGGTTTGTGTGGTTAAGC...........................-AGA-GACTTGGTTGCCTGGGAGACGGGGGTTGG....ACATGGGCGGGCT.......... 318
GRV.ET.x.GRI_677            GCCC-......G-C-TGTTTAAGGACATGCACGAACAT.........GCAAAGCGC......................................-AGTGTCAG................................................ 341
VER.KE.x.TYO1_patent        ACCCCA-TCGGG--T--TTTACTGACATGCATGGGCTG.........GTGAAGAGGAAG...................................-AG-CAGACCGCAAGCCTGCGGTTAGA.............................. 368
VER.KE.x.9063               ACCCAA--AGAC--T-TTTTGAAGACATGCATGCAATA.........GTGAAGAGGAAG...................................-AG-TAACCGCAGGCTTGTGGTTAAG............................... 367
VER.KE.x.AGM155             ACCCC-CAAGGG--T--TTTACAGACTTATATTCAACA.........GTTGGTACAGGAAAC................................-AG-CGACCACAGGCTTGCGGTTTCC............................... 370
VER.DE.x.AGM3               ..........................................................................................................   0
TAN.UG.x.TAN1               ACCC-......G-C-G-TTTAAGGACATGCATGGTCTT.........GTTAAGAGGAAG...................................-AG-CTAACCGCAGGCTTGCGGTTAAGCTAACTAGGCCGTTGCCTAGGAGATGACATAAA 386
SAB.SN.x.SAB1C              TT-CT-GGAGGCT-C--CCAGAGCAGGTGCCGTCAGCA.........CAAGGC.........................................-A--CACAAAACCACATCCTATG.................................. 352
MND_1.GA.x.MNDGB1           ..........................................................................................................   0
DEB.CM.99.CM40              ..........................................................................................................   0
DEB.CM.99.CM5               ..........................................................................................................   0
TAL.CM.01.8023              ..........................................................................................................   0
TAL.CM.00.266               ..........................................................................................................   0
SUN.GA.98.L14               G--CCATACAG---C--CCTGAAGAGTTCAAGCATGTT.........ACTAGTCTGCAGTGGGAGGTG..........................-AG-TTAACCGCAAACCGCATCCTCTTGCATCGCCT...................... 383
WRC.CI.97.97CI14            TCCCTGGACAGTCA-T-GCCACTGGGAAAGAGGCCATTGAG...AGTCTCTATAAGAACAAGAAGCAGTGCAAG....................-AGAGACATGTTGCAGTAGGAAGTGGCTTTGCGGTTAAGCTACAT............. 422
WRC.CI.98.98CI04            TCCCTGGACAT-TA-TAGCCCAAGGAGAAGATGCTATTGAGACCCTCCATAAAGAGAAGCTGAAGCGGCACAAG....................-AGAGACATGTTGCAGTAGGATGTGGTTTTGTGGTTACGCTACAT............. 425
WRC.GM.05.Pbt_05GM_X02      ..........................................................................................................   0
LST.CD.88.524               ..........................................................................................................   0
LST.CD.88.447               ..........................................................................................................   0
LST.KE.x.lho7               TA-CCCT-AAA---C-GCCTGCGAAGATACAAAATACT......GCTTTCGCTTTTGACTGCAAGAGGAAG.......................-AG-CAAACCGCAAACCGCAGCACATCCTCTTGCAGCCCGGTTGCTAA............ 401
SYK.KE.x.SYK173             GCTG-G-GA-GGCCCG-TTGCAGTTAGAGAGACAAACA.........GGAAAACCACAAGAGCTGCAGTCAGCACTTAGTAAGAACATCTCC.................-AGG-GACTCCATGGTGACAAGCTCGG............................... 289
SYK.KE.x.KE51               ..........................................................................................................   0
COL.CM.x.CGU1               ..........................................................................................................   0
```

```
                         NF-κ-B-II                  NF-κ-B-I     Sp1-III          Sp1-II   Sp1-I                TATA Box
H1B.FR.83.HXB2          ....................TCGAGCTTGCT...ACA...AGGGACTTTCCG....................CTGGGGACT...TTCCAG............GG.AGGCGTGG.CCTGGGCGGGACTGGGGAGTGGCGA.......GCCCT..CAGATCCTGCATAT  429
H1A1.UG.85.U455         .......................................................................................................................................................................    0
H1B.US.90.WEAU160       ....................-----T--T--..---..-----------A....................---------..----G-............--.-------A.T-----------A-----------.......-----.-----G--------  429
H1C.ET.86.ETH2220       .......................................................................................................................................................................    0
H1D.CD.84.84ZR085       .......................................................................................................................................................................    0
H1F1.BE.93.VI850        .......................................................................................................................................................................    0
H1G.SE.93.SE6165        .......................................................................................................................................................................    0
H1H.CF.90.056           .......................................................................................................................................................................    0
H1J.SE.93.SE7887        .......................................................................................................................................................................    0
H1K.CM.96.MP535         .......................................................................................................................................................................    0
H101_AE.TH.90.CM240     .......................................................................................................................................................................    0
H102_AG.NG.x.IBNG       .......................................................................................................................................................................    0
H1N.CM.95.YBF30         .......................................................................................................................................................................    0
H1O.BE.87.ANT70         ACCTGAAGATTGCTGACACTG-..................--A-------A..GCAAAGACTGCTGACACTGC-------..------..............TG--A-G-ACAGG------TT-G---AGTG-C........TAA----.-----AG--------  464
H1O.CM.91.MVP5180       ...................................----------A.......GACTGCTGACACTGC-------..------.............C-TG--A-G-ATAAG------TT-G---AGTG-C........TAA----.-----G--------  437
CPZ.CD.90.ANT           .......................................................................................................................................................................    0
CPZ.CM.01.SIVcpzCAM13   .......................................................................................................................................................................    0
CPZ.CM.05.SIVcpzEK505   .......................................................................................................................................................................    0
CPZ.CM.05.SIVcpzMB66    .......................................................................................................................................................................    0
CPZ.CM.98.CAM3          .......................................................................................................................................................................    0
CPZ.GA.88.GAB2          .......................................................................................................................................................................    0
CPZ.CM.98.CAM5          .........GCGCCTGCGCAG-A--AAC----G..---A..---------TA.A...................------..-----A..............-GG--A-GCTCA..--------.ACA---C----TTTC......A----.----AG--------  105
CPZ.CM.05.SIVcpzMT145   .......................................................................................................................................................................    0
CPZ.TZ.00.TAN1          .......................................................................................................................................................................    0
CPZ.CM.05.SIVcpzLB7     .......................................................................................................................................................................    0
CPZ.US.85.CPZUS         GCTGTAACCGCGCAGGCGC.AATA-AAC----G..-..CTG.A--------TAA...GGGACTTTCCAA.........-----G..----..............AAGG--GTG--T-A.------A.ACA---C----TTTT......A----.--A-GA-------A  456
CPZ.GA.88.GAB1          TTAG.ACTGGCGCATGCGC.ACAAGAAC----G..--TCTGC----------AA...................------..----G-............--AGA-GTG-TCGGGA------GT-----T-----TTT.....-----.----.AG--------  449
MON.CM.99.L1            .......................................................................................................................................................................    0
MON.NG.x.NG1            .......................................................................................................................................................................    0
GSN.CM.99.CN166         .......................................................................................................................................................................    0
GSN.CM.99.CN71          .......................................................................................................................................................................    0
MUS_1.CM.01.1085        .......................................................................................................................................................................    0
MUS_1.CM.01.CM1239      .......................................................................................................................................................................    0
MUS_2.CM.01.CM1246      .......................................................................................................................................................................    0
MUS_2.CM.01.CM2500      .......................................................................................................................................................................    0
RCM.NG.x.NG411          .......................................................................................................................................................................    0
RCM.GA.x.GAB1           .......................................................................................................................................................................    0
DEN.CD.x.CD1            .......................................................................................................................................................................    0

                         NF-κ-B-II                           NF-κ-B-I                                                  TATA Box
MAC.US.x.239            ................................-ACAGC-----------A.....................CAA--GGA...-GTT-CGGG..GAGGTACT--GGA-GAGCCG.G-C--GAACG.......................---ACTTTCT-GA--T---A  492
H2A.GW.x.ALI            GGGCAGGAAGTAACTACTG..AAA-C.AGCTGA..GACTGC-----------A.....................GAA--GGC...-GTA-C............CAGG--AAG-AC.A----A--AG----T-G-GAA-.............-----CAT-CTC.---T---A  525
H2A.DE.x.BEN            AGGCAGGAAGTAGCTA.CTAAGAA-C.AGCTGA..GGCTGC-----------A.....................GAA--GGC...-GTA-C............CAAG--A-G-AC.A----A--AG----T-G-GAA-.............-----CAT-CT-A---T---A  529
H2A.SN.x.ST             .......................................................................................................................................................................    0
H2B.GH.86.D205          AGCTGACACTGCACAA.GAAGGAA-CTAGCAGA..CACTGC-----------A.....................AAA-AGGC...-GTA-C............A.TG--A-GA-C.AA---A---GT---TTG-GAA-.............----ATTA-ATCT---T---A  535
H2B.CI.x.EHO            AGCTAACACTGCATAGAGAAGGAA-CTAGCTGA..TACTGC---.-------A.....................GAA--GGC...-GTA-C............A-TG--A-G-AC.A----A--AG-A---AG-.AA-.............----ATCA---CT---T---A  532
H2G.CI.92.ABT96         .......................................................................................................................................................................    0
H2U.FR.96.12034         .........................................................................................................................GAA-.............-----CAA-TGAT---T---A   25
MAC.US.x.EMBL_3         .......................................................................................................................................................................    0
SMM.US.x.H9             .......................................................................................................................................................................    0
SMM.US.x.PGM53          ................................GACAGC-----------A.....................CAAA-GGGA..-GTT-T..............--GGA-GTACTG..G--AG-AACTG-TT-GACC..............----ACTT-TT-T---T---A  418
SMM.US.x.SIVsmH635F_L3  ................................GACAGC-----------A..CAAGGGACTTTCCA......CAA--GGC...-GT--TGGG..GAGGTACT--GGA-GAGCTG.GCT--.AACG.......................---ACTT-TTCT---T---A  505
SMM.SL.92.SL92B         .......................................................................................................................................................................    0
STM.US.x.STM            ................................GACAGC-----------A.....................CAA--GGC...-GTA-............CA--GGA-G..TAC.TG--AG-A-CTG-T--G-.AA-.............-----TTT-CT-T---T---A  165
MNE.US.x.MNE027         .......................................................................................................................................................................    0
DRL.x.x.FAO             .......................................................................................................................................................................    0
MND_2.x.x.5440          .......................................................................................................................................................................    0
MND_2.CM.98.CM16        .................................................................................-GTCGGGTGGGGGAGTG-TC-AC-CAACCTGGCT--ATATAA-CA-CTC-CTTTGCGCTTGAAGTGAGTCTC---CTG-G-G   82
MND_2.GA.x.M14          .................................................................................................................................................................G-G-G    5
OLC.CI.97.97CI12        A..............................T..CCTT--TTG-----G-TTTTGCCC.......................................................................................................T---A  349
GRV.ET.x.GRI_677        ..............................-ACTTTGCGGTT-----------..CCAGGGACTTTCCACA.....-T--GTGG.................ATCG-AGGC-GTACAGGGGC--TACTGGGA-T-GC-T.............T-CCCTCAGAG-------A  444
VER.KE.x.TYO1_patent    ACATCACCATGGAGATGACAT-AA-AAC----GAC.......----------AGCGAAGGGACTTTCCA..................................A--CG--ACAT-.GGC--T.CC-GGGA-T-GC-T.............TACCCTCAGAA-------A  481
VER.KE.x.9063           CCGTTGCCAGGGAGATGACAT-T--AAC----GAC-G.....----------A....AGGGACTTTCCA..................................G--CG---CAT-.GGC--T.AC-GGGA-T-GT-T.............TACCCTCAGAG-------A  478
VER.KE.x.AGM155         TGGTTGCCTAGGAGATGACAT-AAGAAC----GAC.......----------A.GCACGGGACTTTCCA..................................A--CG--ACAT-.GGC--T.AC-GGGA-T-GC-T.............TACCCTCAGAG-------A  482
VER.DE.x.AGM3           .......................................................................................................................................................................    0
TAN.UG.x.TAN1           GAAGTGC.......................-GA..TCAGC-----------A.......................................................G--TG-TGCGA-.GGC--TACTTGGGA-T-GC-T..............-CCCTCAGAG-------A  464
SAB.SN.x.SAB1C          GAGTTGTCATGGTGATGACAT-AAGAAC----GAC..TGAG-----------A...........................GGG...-GGAGA............CTGG-CG--ACTGGGA-T..........-GC-T.............G-CCCGT-C-G--CA----  454
MND_1.GA.x.MNDGB1        ......................................................................................................................................................GGAGTCTCTACTACAGA   17
DEB.CM.99.CM40          .......................................................................................................................................................................    0
DEB.CM.99.CM5           .......................................................................................................................................................................    0
TAL.CM.01.8023          .......................................................................................................................................................................    0
TAL.CM.00.266           .......................................................................................................................................................................    0
SUN.GA.98.L14           AGGCAACGGGGCTAGC.GCAT......GC--GC..TAGAGT-----------.......................................................GACAG--AGGG--A-GC-T...............ACC--CT-A--------  463
WRC.CI.97.97CI14        CCTGAGACTGTGTACCAACGGA-TTAGCAA-CATGCT-ATG-ACTGGGACATTGGGAGGAGACT.GGGAGGTGC--T-T................................................................GGTTT-G-CCT---A  515
WRC.CI.98.98CI04        CCTGGACCTGTGTACCAACGAA-TTAGCAA-CATGCT-ATGGACTGGGAC-T-G.GTGGAGACTTGGGAGGAGA--T-T................................................................GGTTT-G-CCT---A  518
WRC.GM.05.Pbt_05GM_X02  .......................................................................................................................................................................    0
LST.CD.88.524           .......................................................................................................................................................................    0
LST.CD.88.447           .......................................................................................................................................................................    0
LST.KE.x.lho7           GGCAACCGGGCTAGCGCATGCG-A-TGGCTTGCTG--GAGC-----------A...............GGACGGGC---GGAGG.--T....................................................GCC--CT-G--------  490
SYK.KE.x.SYK173         CCCAGGGGA...................------C-GGGC..................GGGG-AA-G--...CA--CT.............-CTGC.........................................................---A  335
SYK.KE.x.KE51           .......................................................................................................................................................................    0
COL.CM.x.CGU1           .......................................................................................................................................................................    0
```

```
                                               +1 mRNA start site
                         TAR element start                                                              TAR element end
            TATA Box                                      5' LTR U3 end 5' LTR R repeat begin
H1B.FR.83.HXB2          AAGCAGCTGCT..TTTTGCCTGTACT........................................................GGGTCTCTCTGGTTAGACCAGA.TCTGAGCCTGGGAG..CTCTCT..GGCTA.ACTA..GGG.AACCCACTGCTTA.....AGCCT..  525
H1A1.UG.85.U455         ..........................................................................................................................................................................    0
H1B.US.90.WEAU160       -----------..-------------........................................................----------------------.--------------..------..-----.G---..---.-------------.....-----..  525
H1C.ET.86.ETH2220       ..........................................................................................................................................................................    0
H1D.CD.84.84ZR085       ..................................................................................................................---..------..---CG.G---..---.-------------.....-----..   39
H1F1.BE.93.VI850        ..........................................................................................................................................................................    0
H1G.SE.93.SE6165        ..........................................................................................................................................................................    0
H1H.CF.90.056           ..........................................................................................................................................................................    0
H1J.SE.93.SE7887        ..........................................................................................................................................................................    0
H1K.CM.96.MP535         ..........................................................................................................................................................................    0
H101_AE.TH.90.CM240     ..................................................................................---------T---------G.--.-----C-----..------..-----.G-A-..---.-----------A-.....-----..   72
H102_AG.NG.x.IBNG       ...................................................................................................-----..------..------..-----.G-GG..A--.-------------.....-----..   54
H1N.CM.95.YBF30         ............--C-CG--TGTACT........................................................----------T-C-G-------.-TA-----------..------..-----.G-..---.-------------.....-----..   85
H1O.BE.87.ANT70         -------C---..--C---T-----C........................................................------GG-TAGAG-----G.--------C-----.---C--.---CT.CTAG..CT-.-----G---------.....-CG--..  560
H1O.CM.91.MVP5180       -----------..--CC--T--.ACC........................................................-----TAG-TAGAG-----G.--------C-----.---C--.---CT.CTAG..CT-.-----G---------.....-CG--..  533
CPZ.CD.90.ANT           ..........................................................................................................................................................................    0
CPZ.CM.01.SIVcpzCAM13   ....................................................................................------CAC-G-------.--------C-----..------..-C-AGTG....A--.-------------.....-----..   67
CPZ.CM.05.SIVcpzEK505   ....................................................................................-------T-----------.-TA-----------..------..-----GCA...A--.-------------.....-----..   69
CPZ.CM.05.SIVcpzMB66    ....................................................................................-------T----------T---------..------..-----GCT.-..A--.-------------.....-----..   70
CPZ.CM.98.CAM3          ..........................................................................................................................................................................    0
CPZ.GA.88.GAB2          ..........................................................................................................................................................................    0
CPZ.CM.98.CAM5          -------C---..--C---T-----G........................................................T--------T--G-------.-TA-----------..------..---.GCC-GC-..---.-------------.....-----..  200
CPZ.CM.05.SIVcpzMT145   ....................................................................................-------T-G---------.--------------..------..-----ACA.-..---.-------------.....-----..   70
CPZ.TZ.00.TAN1          ..............................GCTCTTGCCTAATCTG.........................................................-------.-------..------..--TAGTGGCTGGCTAG-GA--G-------.....-CG--..   75
CPZ.CM.05.SIVcpzLB7     ....................................................................................-------T----------T---------..------..-----GCT.-..A--.-----------A-.....-----..   71
CPZ.US.85.CPZUS         -------C---..C-C---T------........................................................--------TC-C-AGA.-----.-AA-----C-----..------..-----GTG--..---.-------------.....-----..  550
CPZ.GA.88.GAB1          -----------..--C----------........................................................--------T-CAC-G-------.-T------------..------..-A.-C-GTGA-..---.-------------.....-----..  544
MON.CM.99.L1            ....................................................................................................AG-G-CT-G-CTCCA.CAG-GGC.CT-GG-GC-CTTG-A-CT-GGT-AA--GGCCTGGG----CTT   64
MON.NG.x.NG1            ..........................................................................................................................................................................    0
GSN.CM.99.CN166         ......................................................................................................TGCTT-G-CTTTCACAG-GGC...............CTGGGAG-C--TG..GAGCTAG-TAAAG   47
GSN.CM.99.CN71          ......................................................................................................TGCCT--GTCTCCACAG-GGC...............CTGGGAG-C--TG..GAGCTAG-TAGAG   47
MUS_1.CM.01.1085        ......................................................................................................AGT-CTAGCTCC-.CAG-GGC...............CTGGGAG-C--CG..GAACTAG-TA-AG   46
MUS_1.CM.01.CM1239      .....................................................................................................CAGTTCTAGCTCC-.CAG-GGC...............CTGGGAG-C--CG..GAACTAG-TAAAG   47
MUS_2.CM.01.CM1246      ......................................................................................................TCTTCTG-CTCC-.CAG-GGC...............CTGGGGG-C--CA..GAACTAG-TAGAG   46
MUS_2.CM.01.CM2500      ...................................................................................................-TC-TCTCCTG-CTCC-.CAG-GGC...............CTGGGAG-C--CG..GAGCTG--TAGAG   50
RCM.NG.x.NG411          ..........................................................................................................................................................................    0
RCM.GA.x.GAB1           ..........................................................................................................................................................................    0
DEN.CD.x.CD1            ............................................................................................G-AC-T-CTTGCGAGTCTCT--A-C-GCCTGGGAGCCCT-AG-GG-CTTA-C...-G-GG-CTGGG-GCCCTCTAAGGT   76

                                                 +1 mRNA start site
                                 TAR element start
            TATA Box        5' LTR U3 end 5' LTR R repeat begin
MAC.US.x.239            -TATCA----A.T--CGCT-----T-C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TTGGCCG..GTGCTGG--AGAG  637
H2A.GW.x.ALI            -T-T-C-C---.GC--GCAT----T-C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGTG-TTGGCCG..GCACTGG--AGAC  670
H2A.DE.x.BEN            -T-T-C-C---.TC--GCAT----T-C...AGTCGCTCTGCGGAGAGGCTGGCAGATCGAGCCCTGAGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CTG-.A-----A..............CCAGT--TTGGCCG..GTACTGG--AGAC  674
H2A.SN.x.ST             ..............................AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGTG-TTGGCCG..GCACTGG--AGAC  119
H2B.GH.86.D205          -T-T-C-C---.TC-CGCAT----T-C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TAGGCCA..GTGTTGG--AGAC  680
H2B.CI.x.EHO            -T-T-C-C---.T-G-GCAT----T-C...AGTCGCTCTGCGGAGAGGCTGGCAGATCGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TTGGCCG..GTGCTGG--AGAC  677
H2G.CI.92.ABT96         .............................................................................................................---CT--.A-----A..............CCAGCG-TTGGCCG..GTGCTGG--AGAG   41
H2U.FR.96.12034         -T-T-C-C---CT--.GCAT----T-C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-G..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGT--TTGGCCG..GTACTGG--AGAG  170
MAC.US.x.EMBL_3         ..........................C...AGTCGCTCTGGC.AGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TTGGCCA..GTGCTGG.-AGAG  118
SMM.US.x.H9             ..........................C...AGTCACTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CTG-.A-----A..............CCAGC--TTGGCCG..GTGCTGG--AGAG  120
SMM.US.x.PGM53          -TA--A----A.T--CGCT-----T-C...AGTCGCTCTGGGCAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A--C.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TTGGCCG..GTGCTGG--AGAG  563
SMM.US.x.SIVsmH635F_L3  -TA--A----A.T--CGCT-----G-C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TTGGCCG..GTGCTGG--AGAG  650
SMM.SL.92.SL92B         ...............................GTCGCTCTACGGAGAGGCTGCCAGA.TAGAGCCCCGAGAGGTTCTCT.GCAGCAC-TA.GTA---.--G-CTGGGTGTTCT--ACT-A.A-----A.............ACCAGC-A-AGGCCA..TTGCTGG-TAGAC  119
STM.US.x.STM            -TA--A-C---.TA-AGCT-----T-C...AGTCACTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGAGTGTTC----CT--.A-----A..............CCAGC--TTGGCCA..GTGCTGG-TAGAG  310
MNE.US.x.MNE027         ..............................AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TTGGCCG..GTGCTGT--AGAG  119
DRL.x.x.FAO             ..........................................................................................................................................................................    0
MND_2.x.x.5440          ..........................................................................................................................................................................    0
MND_2.CM.98.CM16        GCTACCAGAT-.GAGCCTGGGTGTTCTCTGGTGAGTCTT.GGAAGGAGTGCCTGCTTGTAGCCCTGGGTGGTTCGCAGGCCCCT-G--TGTA-C-CT-GG---CTCG-C--GTGTTCCTGAAGAGGCTCAT-GGGGGCTCCCTTGCTTGGT--TGG--GACCTCTAGCAG  250
MND_2.GA.x.M14          GCTACCAGAT-GAGCC--GG-TGTTCTCTGGTGAGTCCT.TGAAGGAGTGCCTGCTTGTAGCCCTGGGCGGTTCGCAGGCCCCT-G--TGTA-C-CT-GGT--CTCG-C--GTGTTCT-GAAAG.G-CTC----AGTG-GCACTCTTG-TTGG..--TTGGTAGA---CT  171
OLC.CI.97.97CI12        ---TCTG----..C-GC--T---TAGA...GGTCT.TGCTCGGAGA..CCAGCC.................................................T-GA-------A-T-CTT--GGCAGCCTGCTTTG-GC-CCTGGGTGTGCTTGGCAGGTGGTTAAAG  461
GRV.ET.x.GRI_677        ------A----.CGC-G--T----ACTC..AGTCTCTTACTAGGAG....ACCAGCTAGAGCC.TGGGT.GTTCGC....................................---TTA.G-CTAAC.CC-G-.TGGCCA.CC-GGGGTA-GGA--..C...........  551
VER.KE.x.TYO1_patent    ------A----CGC-GGC.T----ACTC..AGTCTCTTACTAGGAGA.................................C-A-C--.--G-CTGGGTGTTCG---TTA.G-CTAAC.CT-G-.TGGCCA.CCAGGGGTA-GGA--..C...........  588
VER.KE.x.9063           ------A----.CGC-G--T----ACTC..AGTCTCTTACTAGGAGA.................................C-A-C--.--G-CTGGGTGTTCG---TTA.G-CTAAC.CT-G-.TGGCCA.CCAGGGGTA-GGA--..C...........  585
VER.KE.x.AGM155         ------A----.CGC-G--T----ACTC..AGTCTCTTACTAGGAGA.................................C-A-C--.--G-CTGGGTGTTCG---TTA.G-CTAAC.CT-G-.TGGCCA.CCAGGGGTA-GGA--..C...........  589
VER.DE.x.AGM3           ..........................C..AGTCTCTTACTAGGAGA.................................C-A-C--.--G-CTGGGTGTTCG---TTA.G-CTAAC.CT-G-.TGGCCA.CCAGGGGTA-GGA--..C...........   81
TAN.UG.x.TAN1           ------A----.CGC-G--T----ACTC..AGTCTCTTACTAGGAGA.................................C-A-C--.--G-CTGGGAGTTCG----TGAGC--AA-C.T--TC.TGGCCATCCAGGGGTA-GA--C-.C...........  574
SAB.SN.x.SAB1C          -----------.--GC-------AAAC..AGTCTCTGCTTGGAG.G.CTGCCAGATTGAGCC.TGGGT.GTTCTCTGGTAA-TC-GAAC-A-C--.--G-CTGGGTGTTCG-----GAAGGCAAGC.CTAAG-GCACTCT--TG-GGT-TTG-TA..GACCCTCG-CTG  612
MND_1.GA.x.MNDGB1       GG...CTAAGGGT-G-AT-TCTG-G.CAGATCCCCTTAGAGCAAGGACCAGAGTCCT.............GAGTGACTGGGTCT-AGCAC--CA.CTCGGGGCTGATCACCT-GA--TAGTGGAA--CCT..-GCTTG..............................  135
DEB.CM.99.CM40          .....................................................................CGGGAGTCGCGCCTTG-GAGTCCTGGT--GT---GCGGCCTGGGA-CCCTTG.---ACT--GATAG-CT.A-CTCC-G-GG-CTGGG-GCCCTGAG--GG   98
DEB.CM.99.CM5           .....................................................................CGGGAGTCGCGCCTAG-GAGTCCTGGTAGGT---GCGGCCTGGGA-CCCTTGC--GCTA-TGACT--CT.A-CTCC-GTGG-CTGGG-GCCCTGAG--GG   99
TAL.CM.01.8023          ..........................................................................................................................................................................    0
TAL.CM.00.266           ..........................................................................................................................T-TCTG-AGAG-GGTCAC-GCG-CCT---AG---TTGGC-ACCGGTAA---G-CC   55
SUN.GA.98.L14           -----------.C--CGC.----T-GGG..AGTCTCTGCTCAGAGG....CTAGGCGACGACCCT........................GAGA-G-CTCA-CT.AGAGCT-G--T-TC-.CTC--GA.............GT-AGC--TTG---A..GCACCTCA--CGA  584
WRC.CI.97.97CI14        ---T--A----..C-G-AG--A-CT................................GGTTGCAGTTTTCTGCCTGA--CTAGGTGCC-GC-TAGT-GTC--AC-TGGCCT--CA----GAGT-C--GAGTGGTG-TCCC--TTGGCT-AG-GACTC--GTAGG  645
WRC.CI.98.98CI04        ---T--A----..C-G-AG--A-C.................................CTTTGCAGTCTTCTGCCTGA--CTAGGTGCC-GC-TAGT-GTC--AC-TGGCCT--CT----GAGT-C--GAGTGGTG-TCCC--TTGGCT-AG-GACTC--GTAGG  647
WRC.GM.05.Pbt_05GM_X02  ..........................................................................................................................................................................    0
LST.CD.88.524           ..........................................................................................................................................................................    0
LST.CD.88.447           ..........................................................................................................................................................................    0
LST.KE.x.lho7           -----------..C-CC--T---T-GG...GAG.......................................TCCTT.GCTCT--GGA-CAGCAG-C-CT-CTGGGT.----ACACTG-.T-AGCTA.A-CAGTC-GCCCT---AGG-TACACTTA..GCACCCA---GG  610
SYK.KE.x.SYK173         ----C-A----..C-G-AG--A-C-GCGG.AGTTGCAACTAGCTCTCTCCAGCGGCCTGGGAGCCCTGAGAGAC.............................................TGG-AG.A-CGGCCCGGGAGCCCTTG--AGT......AAAGGCTAG-GA  448
SYK.KE.x.KE51           .........................CGA..AGTTGCAGCTAGCTCTCCTCAGCGGCCTGGGAGCCC.TAGGAGAC..............................................TGG-AG.A-CGGCCTGGGAGCCCTTG--AGT.....-AAAGGCTAG-AA   90
COL.CM.x.CGU1           ..............................TGGAGTTCTGACTCTT......................................G-CT-GCT-CG-CTGAG-CCT-AG---CTTGC-AGCTA.TCAGCCT.TGCCT-A-TCTGAGTAGCT-G.....CTTGG-TG   90
```

```
                                                                                                                              Extensive secondary structure
                                                                                      Poly-A signal        5' LTR R repeat end  5' LTR U5 start
H1B.FR.83.HXB2              ...............................................................................................C.AATAAAGC..TTGCCTTG..AGT....GCTTCA.AGTA...GTGTGTGCCCGTCTG.TT.GT........GT  576
H1A1.UG.85.U455             .....................................................................................................----..--------..---....----G-.----...---------------.--AT-.........-   46
H1B.US.90.WEAU160           ...............................................................................................-.--------..--------..---....------.----...---------------.--.--........--  576
H1C.ET.86.ETH2220           .........................................................................................................................................................................    0
H1D.CD.84.84ZR085           ...............................................................................................-.--------..--------..---....----T-.----...---------------.--TT........--   90
H1F1.BE.93.VI850            .........................................................................................................................................................................    0
H1G.SE.93.SE6165            .........................................................................................................................................................................    0
H1H.CF.90.056               .........................................................................................................................................................................    0
H1J.SE.93.SE7887            .........................................................................................................................................................................    0
H1K.CM.96.MP535             .........................................................................................................................................................................    0
H101_AE.TH.90.CM240         ...............................................................................................-.--------..--------..---....----A-.---G...---------------.-G.T-........AG  123
H102_AG.NG.x.IBNG           ...............................................................................................-.--------..--------..---....------.----...---------------.A-T--........--  106
H1N.CM.95.YBF30             ...............................................................................................-.--------..--------..---....---AA-..--G...----------A--CA.--C-........--  135
H1O.BE.87.ANT70             ...............................................................................................-.--------..--------..---GAGAAGCA.........--------T-A----.--CAACC........  610
H1O.CM.91.MVP5180           ...............................................................................................-.--------..--------..---GAGAAGCA.........--------T-A----.--CAACC........  583
CPZ.CD.90.ANT               .........................................................................................................................................................................    0
CPZ.CM.01.SIVcpzCAM13       ...............................................................................................-.--------..--------..---....------.----...---------TA----.--CA........--  118
CPZ.CM.05.SIVcpzEK505       ...............................................................................................-.--------..--------..---....----.-.----...----------A--CA.--C-........--  119
CPZ.CM.05.SIVcpzMB66        ...............................................................................................-.--------..--------..-------...AC-T-.----...--------T-A--CA.--GT........--  121
CPZ.CM.98.CAM3              .........................................................................................................................................................................    0
CPZ.GA.88.GAB2              .........................................................................................................................................................................    0
CPZ.CM.98.CAM5              ...............................................................................................-.--------..--------..-------...AC-TG.--C-...---------TA--CA.--CA........--  251
CPZ.CM.05.SIVcpzMT145       ...............................................................................................-.--------..--------..---....---G--.----...--------------A.--CA........--  121
CPZ.TZ.00.TAN1              ...............................................................................................-.---------.C----G...--A....-TG-T..-AC-...----------A-T-CA-ACCGCGT.....C-  129
CPZ.CM.05.SIVcpzLB7         ...............................................................................................-.--------..--------..---....---G--.----...---------A-TCA.C-TT........--  122
CPZ.US.85.CPZUS             ...............................................................................................-.--------..--------..---....----T-.----...---------TA--C.--CA........--  600
CPZ.GA.88.GAB1              ...............................................................................................-.--------..--------..---...GTA--TG.--C-...---------TA---A.--CA........-A  596
MON.CM.99.L1                G...................CCTAGAGAGTTGGTGAAGCTGACCCCGGCTCGCTTG..CT.T........TAGTA...AGGCCTTACAGCC..-.--------..........................................................................  130
MON.NG.x.NG1                .........................................................................................................................................................................    0
GSN.CM.99.CN166             CGGCCTGGGAGCCCTTT....GCTAG..........AAGACGGTAGAATTTCGCTTG..CT.T...GCCTAAGGCT...TATGACT..........-------.CT-CCAAAG-..-TAA.CGAG-CTGCTC-C....AT-CATTGG-.TC-AGGTCGGGCTTG....  168
GSN.CM.99.CN71              CGGCCTGGGAGCCCTTGC...CTAGAG...........GACGGTAGATTCTCGGCTTG..CT.T...GCCCTAGGGC...TTTAGTGGCT.....-.--------.CT-C-AAAG-..TACAGCAAG-CTGCTC-C....AT-CATTGG---G-GA-C-GGCTTGG....  176
MUS_1.CM.01.1085            CGGCCTGGGAGCCCTTG....CCTAG...........GAGTACTGATTCTCGGCTTG..CT.T..TGCTTCGAAGC...TTATTG.AAGCCCA..-.-------.G...........................................................  123
MUS_1.CM.01.CM1239          CGGCCTGAGAGCCCTTG....CCTAG...........GAGAACTGGTTCTCGGCTTG..CT.T..TGCTTCGAAGC...TTATTG..AGCCCA..-.--------TAG-A-AACC.AAG-CAAGTCT.......................................  144
MUS_2.CM.01.CM1246          CGGCCTGAGAGCCCTTG....CCTAG........GGGTTGAGACCCCGGCTCGCTTG..CT.T...GCAGAAA...GGCTTTTACATAGCC....-.--------TCC---AAAC.AACC-GCGA-TC-A..................----TCA-.C-GT--GGCGTCTAAG  172
MUS_2.CM.01.CM2500          CGGCCTGGGAGCCCTTG....CTCAG...........AAGGAATAGACCCCGGCTTG..CT.T...GCTTACTTCAGGCTTTATC..AGCC....-.--------.CT---AAC-.AACT-GCGA-TC-A..................----TCA-.CA-T-GCGGCGC....  169
RCM.NG.x.NG411              .........................................................................................................................................................................    0
RCM.GA.x.GAB1               .........................................................................................................................................................................    0
DEN.CD.x.CD1                ATTG.............................CAACAAGTACCCAGACAAGCTTG..CTCG.....TGCTGCTTTTGATGACCATA..........-.-------..TCA-TTAA.................C-GGCATCT---CTTT-CTCCTCATTAGTGGGTTCC.  178
         TAR element end                                                              Poly-A signal     5' LTR R repeat end 5' LTR U5 start
MAC.US.x.239                T..GACTC.CACGCTTG...................................CTTG..CT.T....AAAG........CCCTCTT........-.---------.-.---A-T..TTAGAAGTAAG-T.-.GT...------T---A---.C-CCTAGCCGCCGCC-  731
H2A.GW.x.ALI                ...GGCTC.CACGCTTG...................................CTTG..CT.T....AAAG........ACCTCTT........-.---------.-.---AAT...TAGAAGCAGGTT.-AAG...------T---A---.C-CCTAGTCGCCGCC-  762
H2A.DE.x.BEN                ...GGCTC.CACGCTTG...................................CTTG..CT.T....AAAG........ACCTCTT........-.---------.-.---A.-..TTAGAAGCAAGTT.-AGT...------T---A---.C-CCTAGTCGCCGCC-  767
H2A.SN.x.ST                 ...GGCTC.CACGCTTG...................................CTTG..CT.T...AAAAG........ACCTCTT........-.---------.-.---AGT...TAGAAGCAAGTT.-AGT...-----CT---A---.C-CCTAGTCGCCGCC-  212
H2B.GH.86.D205              ...GGCTC.CACGCTTG...................................CTTG..CT.T....AAAA........CCCTCTT........------.-.-.---A.A..TTAGAAGCAAGT..-AGT...------T---A---.C-CCTAGTCGC.GCC-  769
H2B.CI.x.EHO                ...GGCTC.CACGCTTG...................................CTTG..CT.T....AAAA........CCCTCTT........------A-..-.-----.A..TTAGAAGCAAGAC.-AGT...------T---A---.C-CCTAGTCGCCGCC-  769
H2G.CI.92.ABT96             Y..GGCTC.CACGCTTG...................................CTTG..CT.T..ACAAAG........ACCTCTT........---------..-GC-AAAT...TAGAAGTAAG--AGTGT...------T---A---.C-CCTAGTCGCCGCC-  138
H2U.FR.96.12034             T..GACTC.CACGCTTG...................................CTTG..CT.T....AAAA........CCCTCT.........-.---------.-.---AGT...TAGAAGCAAG-TA-.GT...------T---A---.C-CCTAGTCGCCGCC-  263
MAC.US.x.EMBL_3             T..GGCTC.CACGCTTG...................................CTTG..CT.T....AAAG........ACCTCTT........-.--------..C-.---CATTT.TAGAAGTAAG-C.-.GT...----C-T---A---.C-CCTAGTCGCCGCC-  213
SMM.US.x.H9                 T..RGCTC.CACGCTTG...................................CTTG..CT.T...AAAAG........ACCTCTT........-.--------..C-.---AGT...TAGAAGTAAG-C.-.GT...------T---A---.C-CCTAGMCGCCGCC-  214
SMM.US.x.PGM53              C..GGCTC.CACGCTTG...................................CTTG..CT.T....AAAG........ACCTCTT........-.---------.-.---GCT...TAGAAGTAAG.C.-AGT...------T---A---.C-CCTAGTCGCCGCC-  656
SMM.US.x.SIVsmH635F_L3      T..GGCTC.CACGCTTG...................................CTTG..CT.T....AAAG........ACCTCTT........-.---------.-.---A-T...TAGAAGTAAG--.-.GT...------T---A---.C-CCTAGTCGCCGCC-  743
SMM.SL.92.SL92B             ...GGCTC.CACGCTTG...................................CTTG..CT.T...............GACTCT.........-.-------G.C-.--AACT...TAGAAGCAAGTC.-AG-...------T-T-A---.C-CCTAGTCGCCGCC-  208
STM.US.x.STM                T..GGCTC.CACGCTTG...................................CTTG..CT.T....ATAT........ATCTCTT........-.---------.-.---AAT..TTAGAAGTAAG.C.-AGT...------T---A---.C-CCTAGTCGCCGCC-  404
MNE.US.x.MNE027             T..GGCTC.CACGCTTG...................................CTTG..CT.T....AAAG........ACCTCTT........-.---------.-.---A-T...TAGAAGTAAG-C.-.GC...------T---A---.C-CCTAGTCGCCGCC-  212
DRL.x.x.FAO                 .........................................................................................................................................................................    0
MND_2.x.x.5440              .........................................................................................................................................................................    0
MND_2.CM.98.CM16            TCTCGG.......T.GGCCAAGAGGCTGTGGGATTGACTACCG..........CTTG..CT.T...GCTTT......TAAGTGCT........-.------AG.C--A--CAAGTTAAGAA.AAGCA-.GTG-...----ACT-G-C-A-.T--GTCTTTGGT....  367
MND_2.GA.x.M14               AGCAATCTC....TAGGCCAGGAGATTGTGGGTTAGACTACCG..........CTTG..CT.T.....GCT......ATTGATGCT........-.-------G.C--A---GA..TTAGAA.TAGCA-.GTA-...----ACT-GTCCA-.T--GTCCTTGAT....  289
OLC.CI.97.97CI12            AGTAAGAACCTCAG......................................CTTTG.CT.GTGT.GAGAGA.........................-.------..................GCTGAG-AA-AGAAGACTC-C---CT-G-GTGTT--AC-CATCCAGGT-  548
GRV.ET.x.GRI_677            ....................................................CTTGG.CT.TCATATAGCT.......................-.------C-...---TCGCT.TAGTCGCTA-ATTGGAGTCAA---CTCATT...............GCTGC-C  622
VER.KE.x.TYO1_patent        ....................................................CTTGG.CT.T..AGAAAG........................-T------.-.------GC..-T-AGAGCT-ATC.T-AG...TCAA---T--TCA-.TGAC-CCTCACTCTC-  666
VER.KE.x.9063               ....................................................CTTGG.CT.T..GGAAAG........................-T------A-.T------GC..-T-AGAGCT-ATC.T-AG...TCAA------TCA-.TGAC-CCTCACTC..A  663
VER.KE.x.AGM155             ....................................................CTTGG.CT.T..AGAAAG........................-T------..C------GC..-T-AGAGCT-ATC.T-AG...TCAA------TCA-.TAAT-CCTCACTC..-  665
VER.DE.x.AGM3               ....................................................CTTGG.CT.T..AGAAAG........................-T------T..C--CG--GC..-T-AGAGCT-CTC.T-AG...TCAA------TCA-.TGAC-CCTCACTCTTG  160
TAN.UG.x.TAN1               ....................................................CTTG..CT.T...AGAA........GGCTATTG.......-T-------.-------GC..TTAAGAGT--ATC.T-AGCAA---CC.T-.AT-GC...GC-CCC.TTCCTC.  654
SAB.SN.x.SAB1C              GCGAC.TGGCCATTGCCAGTAGCAGAGACTCCG...................CTTG..CT.T...GCTT.GA.......TTGCCT.........-.-------T.C-------A..GT-TACAAG-AA-.GCA-...----------A-T-.A--CC-CAGGACAACC  728
MND_1.GA.x.MNDGB1            ....................................................CTTG..CTAT..................TGTCTT........-.-------TAAC-TAGAA...TTAGAGCAAG-G-.GTA-...----TAT--AT-G--CGCCTCTCTTCT....  211
DEB.CM.99.CM40              GG................................TAGGCACTCCTCGGCACGCTTG..CT.T.....GCTACGCTTCTATAGCCCAG.......-.-------..TG--GTAA..................C-GCAA------CTG-CCTCATCATTGACTCCAGG-.  196
DEB.CM.99.CM5               GG...............................TTAGGTACTCCTCGGCACGCTTG..CT.T.....GCTGCATTTT.ATTGCCCAG.......-.-------..TG--GTAA..................C-GCAA------CTG-CCTCATCATTGACTCCAGG-.  197
TAL.CM.01.8023              .........................................................................................................................................................................    0
TAL.CM.00.266               TGGGAGACCTTGCCGGTAAACGCTCAGTAGACATCCCGCTGCTTGCTTCGCGAAGCGCTTTAAACATAGCC..................-.--------..TG--TTCA.AAGAGCAGCAG-GTGTGCAGCCCA-C---TGGA-C-GCGGCG-GTGCTGGTCCC  196
SUN.GA.98.L14               G..CCTGA.TCAGCTTG..................GGGAGCTAGAGGCTC...CTTG..CT.T..TGCAGTACA...AG...CCTT........-.--------..CC------A..GT-TTACTGCAAG.CAA.......----T-----G-TC-C.TCTTTCTCCTTC  696
WRC.CI.97.97CI14            AGGCTGTCACTCGCTTG...................................CTTG..CT.A............GG.GTTT.GCAGTT.........-------AGTC-TTGACTTAGACAAGCAG-GAGTCAGTAGTC-TAT-TGGGCAAA-TAAG-AACCCTGGTTC  753
WRC.CI.98.98CI04            AGTCTGTC.CTCGCTTG...................................CTTG...TTA............GGAGTTTTGCA.TT.........-------AGCC-TTTGCTTG-CAAGCA-TGAGTTT-GTAGTT-CAT-TA-A--ACAG-AG-AACCCTGGTTA  755
WRC.GM.05.Pbt_05GM_X02      .........................................................................................................................................................................    0
LST.CD.88.524               .........................................................................................................................................................................    0
LST.CD.88.447               .........................................................................................................................................................................    0
LST.KE.x.lho7               GCCTGATCA..............GCCTAGGGAGC.TAAGGGCTCCTAG.....CTTG..CT.....ACTTTAAA..G...CCTT........-.------AA.G----TCAT..T--TAGAAAG-AC.-AGC...TA----CT-ATAG-.AG-CTCTTTCTCCTTC  727
SYK.KE.x.SYK173             G.GTCCGGGTTTG.....GCCGCCCGGTCTTCGGCAGCTCTCCGTTG.......CTTG..CT.T..TGCTTACGCT..GATCAGCCCA........-.-------T.G-G-TAAAC..GCAAGTAA-G--T.GC-C...TCA-C--TT-TG---GG-CACCTCCCGAA-G  581
SYK.KE.x.KE51               G..ACAAGGACTG.....ATCCACCTAGCCTCTACAATTCTCCGTTA.......CTTA..CT.T..TACTACGCG..CTGATCAGCC........-.-------T.G-GTTAACT..-CAAGTAA-GAGT...GC...TCTCA-CTGTTCTGCCGGC--CTCCCAGAAG  220
COL.CM.x.CGU1               AAGGGCTTGAGGGACTC.......CAGTCCCCTCGCGCTAAGGGTAGATAGGTGCTTTGCAT.T...............AT.AAGAG........-.-------T.GA-TTTGAA..GC-TCAA--AG-G.TT-G...AGTAC-AGTTTCAC-TCCGCCTTCTC.....  214
```

```
                               Extensive secondary structure                                    Lys tRNA primer binding site
                                                                                  5' LTR U5 end
H1B.FR.83.HXB2            GACTCT.GGTAACTAGAGATCCCTCAGACCCTTTTAGTCA.GTGTGG.................................AAAA.TCTCTAGCAG....TGGCGCCCGAACAGGG..ACCTGA............................................ 657
H1A1.UG.85.U455           ------.----------------------AC-A---A-T.----A...................................----.----------....----------------..--TC--............................................ 126
H1B.US.90.WEAU160         ------.----------------------A---------.------..................................----.----------....----------------..---C--............................................ 657
H1C.ET.86.ETH2220         ..............................................................................----.----------....----------------.G------............................................  36
H1D.CD.84.84ZR085         ------.--------------------AGC-------C.-G---....................................----.----------....----------------..------............................................ 171
H1F1.BE.93.VI850          ...............................................................................................................--..--T---............................................   8
H1G.SE.93.SE6165          ..........................................................................GGCTT.---.----------....----------------..--T---............................................  40
H1H.CF.90.056             .....................................................................................................................T---............................................   4
H1J.SE.93.SE7887          .............................................................................................................................................................   0
H1K.CM.96.MP535           .............................................................................................................................................................   0
H101_AE.TH.90.CM240       ------.----------------------AC-C---A-T.-A--A...................................----.------C---....----------------.C--TC--............................................ 204
H102_AG.NG.x.IBNG         ------.----------------------AC-C---A-T.----A...................................----.----------....---------------A..-T---............................................ 185
H1N.CM.95.YBF30           A-----.----C-----------------A-C-AGAC.T.-A---A..................................----.----------....----------------..--T---............................................ 215
H1O.BE.87.ANT70           ......---.---GT---------------T-AC--AGAC.T.-AAGCAG..............................----.----------....----------------..--T---............................................ 687
H1O.CM.91.MVP5180         ......---.---GT---------------T-AC--AGAC.T.-AAGCAG..............................----.----------....----------------..-.-GCG............................................ 659
CPZ.CD.90.ANT             ............................................................................................................---..--T---............................................   9
CPZ.CM.01.SIVcpzCAM13     T-----.----T-----------------TAAAA--TAGTC.CAG.............................G.----.----------....----------------..------............................................ 198
CPZ.CM.05.SIVcpzEK505     A-----.----T-----------------T.AA--CTA-TGAGTA...................................----.----------....----------------..--T---............................................ 198
CPZ.CM.05.SIVcpzMB66      ------.----T----------------TA..AC-GTAGTCTGTGC-GCC..............................----.----------....----------------..--T---............................................ 204
CPZ.CM.98.CAM3            .........................................................................................----....----------------..--T---............................................  26
CPZ.GA.88.GAB2            .....................................................................................................................T---............................................   4
CPZ.CM.98.CAM5            ------.----------------------A....CA-T-AT.AA--A-.........................GTG.----.----------....----------------..--T---............................................ 331
CPZ.CM.05.SIVcpzMT145     ------.----T----------------A-.A-C--TATTGA-AGA..................................----.----------....----------------..--T---............................................ 201
CPZ.TZ.00.TAN1            -C-....CTGGGG----------------T......................TTGTA..GTGGCTAAGTA.----.------C---....----------------..--T---............................................ 206
CPZ.CM.05.SIVcpzLB7       ------.----T----------------A-....--GT-GTCTGCGC-G...............................----.----------....----------------..--T---............................................ 202
CPZ.US.85.CPZUS           ------.----------------------A....C-CT-A-.AA--A..........................GGTG.----.----------....----------------..--T---............................................ 680
CPZ.GA.88.GAB1            C-----.----------------------T....-A-A-T-.TA--CA.........................AGG..----.----------....----------------..------............................................ 676
MON.CM.99.L1              ...-GCTT-A--GAGACTTG-GAGTCTG----CA-CTGTG.CG-ATCAGGAGCGGGCTTGAGCCGGGGAG..GGGACAGCAG-C-.G-C--TCTCCA..G----------------..--T---............................................ 243
MON.NG.x.NG1              .............................................................................................................................................................   0
GSN.CM.99.CN166           .............................................AGTCCG.GGTGGACACC-G-..--CA-C----....----------------..--T---............................................ 219
GSN.CM.99.CN71            .............................................AGTCGC..GGTGGATACC-G-..--CA-C----....----------------.G-T---............................................ 227
MUS_1.CM.01.1085          ..........TTG----AC-AAGG--AGT-TGC-CCTATT.CAT-CACACTGAGAGCCTATCTCAGCGTGGTGTGAGTCTGG---.CAC--C----....----------------..--T---............................................ 229
MUS_1.CM.01.CM1239        ................................GC-CCTGTT.TAT-CACACTGAGAGCCTACCCCGGCGTGGTGTGGTTCCAG---.CAC--C----....----------------..--T---............................................ 229
MUS_2.CM.01.CM1246        AGTAGAC---TG-CG-CACAG--CG-.....................ACCTCGCGGG..................GGGA--T-.----CTC---....----------------.G-T---............................................ 248
MUS_2.CM.01.CM2500        CTAGAAGT--.....................................ACC...................GATCCAGGGCT-C-.-AC--CC---....----------------..--T---............................................ 228
RCM.NG.x.NG411            ...........................................................................................................----------------..--T---............................................  22
RCM.GA.x.GAB1             ...........................................................................................................----------------..--T---............................................  22
DEN.CD.x.CD1              ...............................................AGGAGGTTCGGGTTGTGAGCCTTACGCAGGGTGA----.C-C--C----....----------------..--T---............................................ 248
                                                                                  5' LTR U5 end   Lys tRNA primer binding site
MAC.US.x.239              -GTCAA.CTCGGTACTCA--AATAAG...AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCGAAGCAGG----.--C-------A..T------T---------..--T---............................................ 843
H2A.GW.x.ALI              -GTCA-..TCGGTGTTCACCTG.AGTA--AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACGGAGGCAGG----.--C-------G..T----------------..--T---............................................ 875
H2A.DE.x.BEN              -GTCA-..TCGGTGTTCATCTGAGT-AC.AAGACCCTGGT.C---TAGGACCCTTCTCGCTTTGGGA..ATCCAAGGCAGG----.--C-------G..T----------------..--T---............................................ 880
H2A.SN.x.ST               -GTCA-..TCGGTGTTCATCTAAAGTA--AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCAAGGCAGG----.--C-------G..T----------------..--T---............................................ 326
H2B.GH.86.D205            -GTCAC..TCGGTGCTCC-CTGA......TAGACCCTGGT.C---TAGGACCCTTCTTGCTTTGGGA..AACCGAAGCGGG----.--C-------A..T----------------..-----G............................................ 877
H2B.CI.x.EHO              -GTCA-..TCGGTGCTCC-CTAG.......AAACCCTGGT.C---TAGGACCCCTTCTGCTTTGGGA..AACCAAGGCAGG----.--C-------A..T----------------..------............................................ 876
H2G.CI.92.ABT96           -GTCAC.CTCGGTGTTC-CGTAGC.....AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGRG..AACCAGAGCGGG----.--C-------A..T----------------..--T---............................................ 248
H2U.FR.96.12034           -GTCA-..TCGGAG-TTC-C-TTAT-.....AACCCTGGTAC---TAGGACCCTTCTTGCTTTGGGA..AACCAGAGCAGG----.--C-------G..T----------------..--T---............................................ 373
MAC.US.x.EMBL_3           -GTCAA.CTCGGTACTC..GGTAAT-AG.AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGAAAC..CGAAGCAGG----.--C-------TGAT----------------...--T---............................................ 326
SMM.US.x.H9               -GTCA-.CTCGGTACTC--A-ATAAG...AAGSCCCTGGT.C---TAGGACYCTTTCTGCTTTGGGAAAC..CGAGGCAGG----.--C-------G..T----------------..--T---............................................ 326
SMM.US.x.PGM53            -GTCA-.CTCGGTACTC--A-ATAA--..AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGC..AACCGAGGCAGG----.--C-------A..T----------------..--T---............................................ 769
SMM.US.x.SIVsmH635F_L3    -GTCA-.CTCGGTACTC--CA-A-A-AG.AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCGAGGCAGG----.--C-------G..T----------------..--T---............................................ 857
SMM.SL.92.SL92B           -GTCA-..TCGGGGTCCCGGATAA-....GAG.CCCTGGTAC---CAGAACCCTCTCACTCTTGGGGGCAACTCCTGAGTG----A-----G----T..T----------------..--T---............................................ 321
STM.US.x.STM              -GTCAA.CTCGGTACTC............AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGC..TACTGAGGCAGA----.--C-------A..T----------------..--T---............................................ 507
MNE.US.x.MNE027           -GTCAA.CTCGGTACTC..AATAAT-AG.AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCGAAGCAGG----.--C-------A..T----------------..--T---............................................ 324
DRL.x.x.FAO               ...........................................................................................................----------------..--T---............................................  22
MND_2.x.x.5440            ...........................................................................................................----------------..--T---............................................  22
MND_2.CM.98.CM16          .........AG-AACTCT.GGTTA-T-G.AGA-CCCTCAGATC-GT-GCAGAGATTTGATATCTAGTGGC..AGTAGAGAA----.--C-C-----....----------------..--T---............................................ 471
MND_2.GA.x.M14            .........AG-AACTCT.GGTTA-T-G.AGA-CCCTCAGATC-GT-GCAGAGATTTGATATCTAGTGAG..AGTAGAGAA----.----C-----....----------------..--T---............................................ 393
OLC.CI.97.97CI12          A..................................................................................--T.--C-A-----....----------------..--T---............................................ 584
GRV.ET.x.GRI_677          CGAG-CTCTAG-GGT--ACCT-TCTTACTG........................GGTT..CTCCTG..TACCCAGGTGGG-G--ACTC-.-----....----------------..--T---............................................ 710
VER.KE.x.TYO1_patent      TGAA-G.--A-T-.TTCCT-A-TG..............................GGTTCTCTCTCT..GACCCAGGCGAG-G--.A---C-----....----------------..--T---............................................ 748
VER.KE.x.9063             AG-AGG.--A-C-GTTCCT-A-TG..............................GGTTCTCTCTCT..GACCCAGGCGGG-G--.A---C-----....----------------..--T---............................................ 746
VER.KE.x.AGM155           TGAA-G.--AG-AGTTCCT-A-TG..............................GGTTCTCTCTCA..AACCCAGGCGAG-G--.A---C----.....----------------..--T-.............................................. 745
VER.DE.x.AGM3             A--GGG.TAA---.TTCCT-A-TG..............................GGTTCTCTCTCA...ACCCAGGCGAG-G--.A---C-----....----------------..--T---............................................ 241
TAN.UG.x.TAN1             .-GGGAACCCTC.GTT-CTGG.................................GGTT...TCTTA..T..CCAGGCTGAG---.----C-----....----------------..--T---............................................ 728
SAB.SN.x.SAB1C            CT....................................................GGTTACTAAGGA..TCCCTGATAGAA----.--CT------....----------------..--T---............................................ 790
MND_1.GA.x.MNDGB1         ..................................AAACCTGT--TGTTCTCATTTAGAGAACAGAAGGACTTCTAGTTA-CCCTAGAAGCCTTTCAG.-------------.--..--T---............................................ 295
DEB.CM.99.CM40            ...............................................AATTCGGGTCGAGCGGGCTAGGCTCAAG.CG-G-G.GTC--------...T----------------..--T---............................................ 263
DEB.CM.99.CM5             ...............................................GTCGGGAGCTCAAGCGGCTAGGCTCTGGGCAGG-G.AAC--------...T----------------..--T---............................................ 265
TAL.CM.01.8023            .............................................................................................................................................................   0
TAL.CM.00.266             -CGC-CTC-G.............................................GTCGAGGGAACTTGGTGGGTGAAGA--TTTA-C-ACCAGC....----------------..--T---GAACACCCAGGTTCTCTGTCCTGAGGGCCAAGCGGCTTCGGGGACC 314
SUN.GA.98.L14             T....A.AACCCTGTTC-...........AC-GCTC-T-.ACTA--GAGCAATTACTGGCCGCCTA..GTGATCCGGGGCTG-C.GGC.------....----------------..--A---............................................ 792
WRC.CI.97.97CI14          -CT-GAG-T-G-TCTTCAG.AGGAGCTGAAGACCCC.........................................-A-----T---....----------------..------............................................ 822
WRC.CI.98.98CI04          -CT-GAGCT-GGACTTCAGAAGGAGCTGAAGACCCC.........................................-A-----T---....----------------..--AA--............................................ 825
WRC.GM.05.Pbt_05GM_X02    .............................................................................................................................................................   0
LST.CD.88.524             .............................................................................................................................................................   0
LST.CD.88.447             .............................................................................................................................................................   0
LST.KE.x.lho7             -C-CTG.-T-C-G-GATCTCT-ACT---GAGA--GGAG-C.T---............................GATCCGGG-C-.GGG--TC---....----------------..--T---............................................ 813
SYK.KE.x.SYK173           TG---A.--G-G-GC-......................................................................-CTT---....---------.------..--T---............................................ 624
SYK.KE.x.KE51             .--G-AC--G--AC.....................................................................G-G-CTT---....----------------..--T---............................................ 265
COL.CM.x.CGU1             ..............................-GCCCT....C....AGAGCGGGGCCTTTTTGCTG.....TCCGGGAAGG---.GGC--T----....---------G------..--T---............................................ 286
```

```
H1B.FR.83.HXB2          ........................................................................................................ 657
H1A1.UG.85.U455         ........................................................................................................ 126
H1B.US.90.WEAU160       ........................................................................................................ 657
H1C.ET.86.ETH2220       ........................................................................................................ 36
H1D.CD.84.84ZR085       ........................................................................................................ 171
H1F1.BE.93.VI850        ........................................................................................................ 8
H1G.SE.93.SE6165        ...................................................................................AAGCGAAAGTTAACAGG 57
H1H.CF.90.056           ........................................................................................................ 4
H1J.SE.93.SE7887        ........................................................................................................ 0
H1K.CM.96.MP535         ........................................................................................................ 0
H101_AE.TH.90.CM240     ...................................................................................AAGCGAAAGTTAATAGG 221
H102_AG.NG.x.IBNG       ...................................................................................CGGTAATAGG 195
H1N.CM.95.YBF30         ........................................................................................................ 215
H1O.BE.87.ANT70         ........................................................................................................ 687
H1O.CM.91.MVP5180       ........................................................................................................ 659
CPZ.CD.90.ANT           ........................................................................................................ 9
CPZ.CM.01.SIVcpzCAM13   ........................................................................................................ 198
CPZ.CM.05.SIVcpzEK505   ........................................................................................................ 198
CPZ.CM.05.SIVcpzMB66    ........................................................................................................ 204
CPZ.CM.98.CAM3          ........................................................................................................ 26
CPZ.GA.88.GAB2          ........................................................................................................ 4
CPZ.CM.98.CAM5          ........................................................................................................ 331
CPZ.CM.05.SIVcpzMT145   ........................................................................................................ 201
CPZ.TZ.00.TAN1          ........................................................................................................ 206
CPZ.CM.05.SIVcpzLB7     ........................................................................................................ 202
CPZ.US.85.CPZUS         ........................................................................................................ 680
CPZ.GA.88.GAB1          ........................................................................................................ 676
MON.CM.99.L1            ...................................................................................GTAACGACCCAGCGACT 260
MON.NG.x.NG1            ........................................................................................................ 0
GSN.CM.99.CN166         ...............................................AACAGCCCAGGACTTCAAGCCCGAAGGCGGGACCCAAGGATTGCCCCT 267
GSN.CM.99.CN71          ...............................................AGCATCCCAGTGACTTCAAGCTCGCAGGCGGGACCCAAGGATTGCCCCT 276
MUS_1.CM.01.1085        ...............................................TTAAGCCCGCGAGGAGGTAGGCGGTGTGACGGAGAGAAGG...CCCC. 273
MUS_1.CM.01.CM1239      ...............................................TTAGCCCGCGAGGGGGTAGGCGGTGTGACAGAGAGAAGG...CCCC. 272
MUS_2.CM.01.CM1246      ...........CTACATATCAACCCGTTCGGAGGTTCGCGGCCGGCAGTCAACAGGCGGTGGAGCAGTCCCAGAAGGCTCCATCCCGAAGG..CGAGACGCCTGAG 341
MUS_2.CM.01.CM2500      ...........................CGAACGTCAGGTAGGATCGGTGAGCAGGCAGTAGAG.CCTCCCAGGGACTCTGCGCTGAAGGCGACAGGCGCCGAGAG 305
RCM.NG.x.NG411          ...................................................................................GAGTGACT 30
RCM.GA.x.GAB1           ...................................................................................GAGTGGCT 30
DEN.CD.x.CD1            ...............................................AAGAGAGTCCTGAAACAGGGACTTGAGTACCCAGAAGGCTCAGGCCCGA 297

MAC.US.x.239            ...................................................AGGAGAGTGAGAGA..CTCCT. 862
H2A.GW.x.ALI            ...................................................AGAAGACTGAGAA.G..TCTAG. 894
H2A.DE.x.BEN            ...................................................AGAGGACTGAGAA.G..CCCTG. 899
H2A.SN.x.ST             ...................................................AGAAGACTGAGAA.G..CCTTG. 345
H2B.GH.86.D205          ...................................................GAAGAGTGAGAG.T..CTTCAA 896
H2B.CI.x.EHO            ...................................................GGAAGAGTGAGAG.T..CTTCG. 895
H2G.CI.92.ABT96         ...................................................GAAGGAGTGAGAGT...CCTTT. 267
H2U.FR.96.12034         .....................................AGAGTGAGAGT...CTTCTG 390
MAC.US.x.EMBL_3         ...................................AGGAGAGTGAGAAGA.ACTCCT. 347
SMM.US.x.H9             ...................................AGGAGAGTGAGAG....CTCCT. 344
SMM.US.x.PGM53          ...................................AGGAGAGTGAGAG....CTCCT. 787
SMM.US.x.SIVsmH635F_L3  ...................................AGGAGGTGAGAG....CTCCT. 874
SMM.SL.92.SL92B         ...........................AGACG.......................... 326
STM.US.x.STM            ...................AGGAGAGTGAGAGA...CTCCT. 526
MNE.US.x.MNE027         ...................AGGAGAGTGAGAGA...CTCCT. 343
DRL.x.x.FAO             ...........AGACAGTCGACCCCACCTGTGAGGGACGAA.GGCGGCAGCCATCCGGACCGACCCACCCGGAAAAGAGAGTGC..TAAGAAGCCCCGACG 109
MND_2.x.x.5440          ...........CAAGGGTTCAGGTCGCACCCACCTGTAAGGGACGAGCGGCGGCAGC.AGTCGGACAGACCCACCCGGCG.AAAGAGTTAATTGAAGAGCCCCGACG 116
MND_2.CM.98.CM16        ...........CAAGGGTTCGAGTCATTCCAACCTGTAAGGGACGAAGGCGGCAGCCAGCCGGACCGACCCACCCGGCGAAGTGAGTTAACCAAG 555
MND_2.GA.x.M14          ...........CAAGGGTTCGAGTCGCTCCGACCTGTAAGGGACAAGCGGCGGCAGCCGGTCGGACAGACCCACCCGACGAAGAGAGTTAATCAAGGAGCCCCGACG 489
OLC.CI.97.97CI12        ...................................................................GAAGA 584
GRV.ET.x.GRI_677        ...................................................................GAAGA 715
VER.KE.x.TYO1_patent    ...................GTGA..GAGTGT 758
VER.KE.x.9063           ...................GGTCCGAGGACTT..GAGTGA 765
VER.KE.x.AGM155         ...................GAGTGA 751
VER.DE.x.AGM3           ...................CTTGAGTGA 250
TAN.UG.x.TAN1           ........................................ 728
SAB.SN.x.SAB1C          ...........CTAAAGACGTCCCGGCCGCAGAGTAGACCA 820
MND_1.GA.x.MNDGB1       ........................................ 295
DEB.CM.99.CM40          ...................GCTACCCAGAGCTCGGCCCGCAGGC 288
DEB.CM.99.CM5           ...................AGCATCCCAGTGTTTGAGCCCGAAGGC 292
TAL.CM.01.8023          ........................................ 0
TAL.CM.00.266           GCGGGTTTTGGGTCCCAAAGAAGAGAGAGAAACTCCGAGGGGTCGTGAGAGGACGACAGTTCCGGAGTCGAGCCCCGGATCGGCGAAGACGACGCGCGTCGCAAAGTGAAAGTAAGGTAAGGAGCCGATCGGTCACCCCGGGCGAGTAGCCGCCGTGAGCGAGTGCGGTG 484
SUN.GA.98.L14           ...................GGGAGAGAGAGTGACGAAGAGACGTGGAGTGAAACCTTCAGAAGAAGTGAGCAGCTGCTGAGAAGAAGGAAGGTTTCGACCTGGTGGAA 881
WRC.CI.97.97CI14        ...................GGGGCGTCTAGAGTGAAAGAAGACCGGCCGGCGA 856
WRC.CI.98.98CI04        ...................GGGGCGACTAGAGTGAAAGAAGACCGGCCGGCGA 859
WRC.GM.05.Pbt_05GM_X02  ........................................ 0
LST.CD.88.524           ........................................ 0
LST.CD.88.447           ........................................ 0
LST.KE.x.lho7           ...................AGGAGAAGAGGCGCTGGCAAGAACCTCTGGAGGGAGGAAAAGCGCGCAG........AAGAGGCCAGGAAGCTAGGACAGA 886
SYK.KE.x.SYK173         ........................................ 624
SYK.KE.x.KE51           ........................................ 265
COL.CM.x.CGU1           ...AAGAGGTCCCTGAGGGTCGCGGAAGAAGAAGAGGAGACTGATCGCCTCCTCAGAAGAGAGGTAAGTCGAAAAAGGGGACTTACGGGTGGAGTCAAGGAGGACTCTACGGTCTGCGCAAAAGTGA 410
```

```
                                                                Packaging loops begin
H1B.FR.83.HXB2          .......AAGCGAAAGGGAAA......CCAGA....GGAGC....TCTCTCGACGC..AGGACTCGGCTTGCTGAA...GCGCGCAC.GGCAAGAGGCGAGGGGCGGC.....GACTGGTGAGTACGCC...AAAAATTTTG.ACTAGCGGAGGC...............  776
H1A1.UG.85.U455         .......---------TT........-----.....-A--T...-----------.------------------G...-T--A---.A-----------A.-----....GACTG-TGAGTACGC-..TA---TT-----.------------........CTAG....  246
H1B.US.90.WEAU160       .......---------..........-----.....-----...-----------.-------------------G...-------.----------------.-----....---------------G...------.------------........CTAG....  778
H1C.ET.86.ETH2220       .......---------T--G-.....-----.....---A...-----------.------------------...-T--A-T-.-----------A.-----....---------------.AATTTTTA----.------------........CTAG....  159
H1D.CD.84.84ZR085       .......---------TAG--.....-----.....---A...-----------.------------------..A-----T-.--------------A--....--------------TG.A---TT-----.------------..............  293
H1F1.BE.93.VI850        .......---------TAG--.....-----.....-A--A...----------G.G-----------------..GTGCAA---.-----------A.-----....GACTG-TGAGTACGC-GAATTTTTT-----.------------..............  131
H1G.SE.93.SE6165        GACTCGA---------TT........-----.....-A--T...-----------.-----------------G...-T--A---.A-----------A.-----....---------------......---T-----.------------..............  179
H1H.CF.90.056           .......T--T-----TA--......-----.....-A--A...-----------.------------------...-T--A---.A-----------A.-----....---------------...-TTTTG---TG------------..............  123
H1J.SE.93.SE7887        ..........................................T...-----------.------------------...-T--A---.-------------------G....---------------AAA--T-T-----.------------........CTAG....  102
H1K.CM.96.MP535         .......................................................................................................................................................................    0
H101_AE.TH.90.CM240     GACTCGA---------TT........-----.....-A--T...-----------.------------------G...-T--A---.A-----------A.-----....---------------....-------.------------........CTAG....  344
H102_AG.NG.x.IBNG       GACTCGA---------TT........-----.....-A--A...----------A.------------------G...-T------.A-----------A.-----....GACTG-TGAGTACGC.-....---------.------------........CTAG....  319
H1N.CM.95.YBF30         .......--A-----TAG--......--G--.....--CTGAA..-----------.-----------C-T--.....-T--A---.A--G---------C-GC-GA.....-G--TGAGTACGC.........----.-----A.--G----T--..CC............  330
H1O.BE.87.ANT70         .......--T-----T---.......-----.....---AGAAAAC---C.-----AAC--G-------AGC-G-...-T--A-C-.-CT---------A--AACTC.ACA--GG-TGAGTA-..................-----.--G----T--..CCA...........  805
H1O.CM.91.MVP5180       ......A---T-----T-G--.....-----.....---AGAAAAC---C.-----AAC--G-------AGC-G-...-T--A-CT.-CT---------A--AACTCACAA--GG-TGAGTA........-.........-----.--G----T--..CCA...........  780
CPZ.CD.90.ANT           ...............--AG.......-T-C-.....-A-A.....C--G-.-----C.---------TC-----C-........GT.A---GA--------.----A......-GG-TGAGTA.-C.TTCATTTTG-----..T-GCTA-T-A..GCA...........  116
CPZ.CM.01.SIVcpzCAM13   .......---------TA-G-.....--G--.....--CTGAA..----C-----A.C-----------T--.-...-T--A---.AA-------------.----G....ACTG-TGAGTACGC........--.T-----.--------T--..CTA...........  315
CPZ.CM.05.SIVcpzEK505   ..........---T-----TA-C.....-G--.....--CTGAA..--C----G---.-----------C-T--.-...-T--A---.A--G----------.----G.....--AA------------...........-----A.--G----T--..............  310
CPZ.CM.05.SIVcpzMB66    .......--A------TA-G-.....-----.....-A--T...-----------.------------------...-T--A---.-------------------A---G...---------------....CT--T-----.-------CG-T..............  321
CPZ.CM.98.CAM3          .....A-GTT-----T-G--......-----.....--CTGAC..---TC-----ACG--------------.-...-T--A---.A--G----------C-AACTC...CAGAG-TGAGTG............T------.-C-----T--..CTG...........  143
CPZ.GA.88.GAB2          .........---T-----CA---.......--G--GGAA--CTGAC..----C-----G.C-------A-------T-...-T--A--ACA--G----------C.G-TAC.......G-TGAGTACA.......TTTTC----.-----A-T--..CTA...........  125
CPZ.CM.98.CAM5          ..............---T----GAAGAA---------CTGAC..---TC-----A.C-----------A---.-...-T--A-G-.A-T---------C-ACT-C....TGAG-TGAGTATTT.......-..T------.-C-----T--..CTG...........  449
CPZ.CM.05.SIVcpzMT145   ..............---T----GTAAAA--G--.....--CTGAA..-----------A.C---------------.-...-T--A-G-.A------------C-GC-AC.TCAAGTG-TGAGTACGC.......-..-----.-----------T--..CTA...........  322
CPZ.TZ.00.TAN1          ..............-AAGC-......GGGA-....C-C-G....C-C--G-----.-----------.---CA.----AAT-..A------------C-GACTC...C-GTG-TGAGTAC......A--TTT-G--.......T---T--..GCA...........  313
CPZ.CM.05.SIVcpzLB7     .......--A------TA-.........-T-G...A-C--T...----------G.G-----------------...-T--A---.------------------A---G....ACTG-TGAGTACGC.....CT--C-----.---------------.TA...........  318
CPZ.US.85.CPZUS         .......--A------TAC--.....-----.....--CTGAA..A--TC------..G---------C---.-...-T--A---.A--G---------C-AACTC...CAGAG-TGAGTACA..........------..-C---A-T--..CTG...........  795
CPZ.GA.88.GAB1          ..............--T----GTAA..--G--.....--CTGAA..----C-----A.C-----------.-...-T--A---.A------------C-GC-AC......TG-TGAGTACGC.......--.T-----.-------T--..CTA...........  790
MON.CM.99.L1            CGAGCCCGCAG-CG-ACAG-TCCGCGAGA--CCTTGCCGAGTGCGC-CGA-A-A--TGC-C-GCAA-AGGTAA-G-GAGATCG-GCG.-T-CCCTA-G-T--CAGACTA....-CG---GAC-GGAA--TATCC--GAGGGAAGG-TC........CAGAGGTGCTCCAA  417
MON.NG.x.NG1            .......................................................................................................................................................................    0
GSN.CM.99.CN166         GGCCGAGTG-TCCGG-CTCC-CGAAATTTGCTGCGCA-CC-CTTG...AAGAGAAGGT-A-GAGACCGGACGG-TCTA.-AAG-GTG.-CAG-CTTCTAG-CCAG-TG..GGTTG....AAGAAGG-TACCC-C-C-G..A-GGGAGAAACT-TG...............  409
GSN.CM.99.CN71          GGCCGAGTG-TCC.G-CTCC-C..AACTTGCTGCGCA-CA-CTAG...AAGAGAAGGT-A-GAGACCGGACGG-TCTG.-AAG-GTG.-CAG-CTTC-AG-CC-GT-G..GTT--....GAGAAGG-TACCC-C-C-GAGG-G-GA-A-T-T-A...............  416
MUS_1.CM.01.1085        ..AGCCG..AGTGCGT-CC--G..CCA--G--GAACAAGTGCTGCG-AGC-ATAAGGT---TGA-TTGGACGG-TCTGG-AAG-GTG.-CAG-CTTC-TCA-AT--AGCA...-G-ATAGAGAG-TC-TGCGGG-GGGGAGAA---C......................  410
MUS_1.CM.01.CM1239      ..AGCCG..AGTGCGT-CC--G.ACCA--G--GGACAAGTGCTGCG-AGC-A-GAGGT---TGG-TTGGACGG-TCTGG-AAG-GTG.-CAG-CTTC-TCA-AT--AGC....AGG-ACGAGAA--C-TGCGTG-G-GGAGAA---CC......................  410
MUS_2.CM.01.CM1246      GATCGCCT---CG-GT-..GCCCGGCA-GGAGAAAAAAGAGCTGCG-AGCAA-A-AGGTAAGG-GTCGGGCGA-CCTA.-A-G-GTG.-CAG-CCTCTAG-T--GC-GGT...AGAC-AG---.GTC--CGCG-GG-GGAGAA--CGCT.................  482
MUS_2.CM.01.CM2500      AACGCCCC---CG-GT-...-GCACCGA--CGAGGAAATCGCTGCAGC.GT--A-AGGTAAGG-GTCGGGCGA-TCTA.-A-G-GTG.-CAG-CCTCTAG-C--GC-GT....T-A-AGA--.GGA--GCG-G-G-GGAGAA--......................  438
RCM.NG.x.NG411          GAGAGACCTC---CGC--GTCCGGACCAG-AGAGTGACT-GTAGAA-G-GGAC-CAGTCCAGG-AA-AGCAGCCTGGTG--T--T-A.AT-T--C-AGA-AA-TGA-GG....CTAGTAGAGCA-G--GACCCGGTCAGG-ACTGGGC........CTCTAGGGAAGGGT  187
RCM.GA.x.GAB1           GAGAGACCTC---GGCTA-GGTTCGGCG-G-CAGGTCACCGCGGGAG-GGAAC-TTGACCAGG-AA-AGCTGCCTGGTG--TTCA-A.-TGCTAGA-AA--T-AGCTAG....CGAA--AAGCAGG--AACCCGGTCCGG-ACTGGGC........CTCTAGGGAAGGAG  187
DEN.CD.x.CD1            AGGCGGTTG-AACGTACC-CCCTCTTA--GAGCAAG--C-AGGAAGACGCT-CGCAGCGCA-AGGTAAGGATCTCGCCG-A--A---.ACT-GTG-CA-TA-T-T--GG....CTGGAAGATAGCAAGGAGGGCT-TC-GGAA--A-A........GGGCTAGATAAGGG  454

MAC.US.x.239            GAGTACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-C-AACCA-GACG--GTGCTCCT-TAAA-GCGCGG-TC..G-TA-C.....AG-C.G-CGTGA--AGC-GG....AGAG-AA---G.-CT-CGGTTGC...AG-T-AGT--.........AACACAAAAAAGA 1006
H2A.GW.x.ALI            GAACACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-CAAACCA-GACG--GTGCTCCT-GAAA-GCGCAG-CCAAG-TA-C..A.AAGGCCG-CGTGT--AGC-GG....AGTG.AA---G.-CT-CGGGTG-...AG-T-AGTA-........CTACACCAAAA.T 1041
H2A.DE.x.BEN            GAACTCG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-CAAACCA-GACG--GTGCTCCT-GAAA-GCGCGG-CCGAG-TA-C..A.AAGGC.G-CGTGT--AGC-GG....AGTGAAA---G.-CT-CGGGTG-...AG-T-AGTA-........CTACACCAAAAAC 1047
H2A.SN.x.ST             GAACACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-CAAACCA-GACG--GTGCTCCT-GAAAAGCGCAG-CCGAG-TA-C..A.A-GGC.G-CGTGT--AGC-GG....AGTGAAA---G.-CT-CGGGTG-...AG-T-AGT--........CTACACCAAATAC  493
H2B.GH.86.D205          GAGCACA..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-T-AACCA-GACG--GAGCTCCTGTAAAAGCGCAG-CC..G-TA-C.....AGGC.--CGTGA--AGC-GG....AGGA-AA---G..CT-CGGG-GT..AAG-T-AGTA-........TCCAAA....... 1033
H2B.CI.x.EHO            GGACACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-T-AACCA-GACG--GTGCTCCTGTAAAAACGCAG-CC..G-TA-C.....AGGC.--CGTGA--AGC-GG....AGGA-AA---G..CT-CGGG-GT-AAAG-T-AGTA-........TCCAAA....... 1034
H2G.CI.92.ABT96         GAACACG..--AG-GT-A-GGCAGTAAGGGCGGCAG-ATC-AACCA-GACG--GTGTTCCT-GAAA-GCGCGG-TC..G-TA-C.....AG-C.--CGTGA--AAC-GG....AGTG.AA---G.-CT-CGGGTGCGC.AG-T-AGT--.........ACCAAAAATTTGT  412
H2U.FR.96.12034         AAACTCG..--TG-GT-AGGGCAGTAAGGGCGGCAG-A-CAAACCA-GACG--GTGCTCCTGTAAA-GCGCAG-CT..G-TA-C.....AAGC.--CGTGA--AGC-GG....AGTG..A---G.TCT-CGGTTGC...AG-T-AGT--.........AACAAAAAATAGA  532
MAC.US.x.EMBL_3         GAGTACGG.C-TG-GT-A-GGCAGTAAGGGCGGCAG-A-C-AACCA-GACG--GTGCTCCT-TAAA-GCGCAG-TC..G-TA-C.....AG-C.G-CGTGA--AGC-GG....AGAG-AG---G.-CT-CGGTTGC...AG-T-AGT--.........AACACAAAAAAGA  492
SMM.US.x.H9             GAGTACG..--TG-GT-A-GGCAGCAAGGGCGGCAG-A-CAAACCA-GACG--GTGCTCCT-TAAA-GCGCGG-CC..G-TA-C.....AGGC.G-CGTGA--AGC-GG....AGTC--A---G.-CT-CGGTTGC...AG-T-AGT--.........AACAAAAGAGTCA  488
SMM.US.x.PGM53          GAGTACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-CAAACCA-GACG--GTGCTCCT-TAAA-GCGCGG-CC..G-TA-C.....AGGC.G-CGTGA--AGC-GG....AGTC--A---G.-CT-CGGTTGC...GG-T-AGT--.........AACAAGAAGAG..  929
SMM.US.x.SIVsmH635F_L3  GAGTACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-CAAACCA-GACG--GAGCTCCT-GAAA-GCGCGG-CC..G-TA-C.....AGGC.G-CGTGA--AGCAGG...GAGTC--A---G.-CT-CGGTTGC...AG-T-AGT--.........AACAAAAA..... 1014
SMM.SL.92.SL92B         GACT.CG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-T-AACCG-GACG--G....TTTGA....-CC-G--CC..A-T-G--...T-AGC-G-AGT-CA--TAAGA...GAGGCCTGTG-G--AGTAAGGTCC-AAAG-T..GT-A.......ATATCCTACCTACG  468
STM.US.x.STM            GAGTACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-CAAACCA-GACG--GTGCTCCTGTAAA-GCGCGG-C.........-.....AAGC.G-CGTGA--AGC-GG....AGGC-AA---A..CT-CGGTTGC...AA-TGAGT--.........AACAAGTAGCCAG  663
MNE.US.x.MNE027         GAGTACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-C-AACCA-GGCG--GTGCTCCT-GAAA-GCGCGG-TC..G-TA-C.....AG-C.G-CGTGA--AGC-GG....AGAG.AA---G.-CT-CGGTTGC...AG-T-AGT--.........AACACAAAAAAGA  486
DRL.x.x.FAO             CGAGACCCGC-TCGGTAAGT-AAAAGGAA--TAAAATC--TGCAC-G-GAG-CGAAGGGTCTAGT--GCC-TGTG-GCA---GCT-G.-CA---GCAACTCC-TA---A....A-AA-AAAGAA-AAGGGAG--GT-AGGCCTGT-GC........TGATAAGTGAGGCA  266
MND_2.x.x.5440          CGCAGGACTCAA-GGTAAGC-GTGCACTGTA....ATTGATTGAGAGAGAGTC-AGGG-TCCGCTT-AGCAGGCG-GATCGCGAAGG.CAACCCCA-TAG-CAAAA-AA....A-GAA-G-GAAGTAAGGCCG-CGCAAAGTG-AAGT........AAGGGCTTCCTCCG  269
MND_2.CM.98.CM16        GAGCCCCG-CGCGC---ACTCAAGGTAAGC-GTGCACCGAACAGT-G-G-GTGT-AGTCCAGGGTCCGC-AGA-C-AGC-A-AC-G-.C-AGGC-AC-CCA-TAG-CAA....A-GGAAAAGAGGG-AAGTA-GGCCGAGGCA-AGT-........AAAAGAGGTCCTTG  712
MND_2.GA.x.M14          CGCAGGACTC.A-GGTAAGCGGTGCAC-.......GTATTGTGGTAG-GAGTC-.AGG-TCCGCTT-AGCAGGCG-GATCGCTAAGG.CAACCCCA-TAG-T.AAAAGA....A-AGA-G-GAAGTAAGGCCG--GCAAAGTG-AAGT........AAAAGGGCCTTCCG  636
OLC.CI.97.97CI12        .......................................................................................-AAA-G-TTCGAG.TC-CT-AAGAAGAAGTTGGTC-T-AGG--AGTG-C........CGGGGAACACGAGG  646
GRV.ET.x.GRI_677        GGCATCGGCA-CG-CC-CTG-GTTGCTGAGCGTCGGA---GGACGA---AG-TA-GGTGA--GC-TA-GA-T-TTTTGCTAC-TAGT.CAGCGAGAAA-GCTA-GCCGC........---A--G-C-CGG-T-CCATT-GTGGCAACCAACC........CAGTTGGACGAAGG  872
VER.KE.x.TYO1_patent    AGGCACGT-CA-CTGA-A-GGCGTCGGA-GCGAAGGAAGCGCGGGGTG-GACG--ACC-A--AGGA-AC-TGGTG-GTA-GCTT-T-.-AGTGCC--GA-AAA--TCGA....-C--A--T--AGGA-TAGG-G-GGCCG-AGC-GTA........ACTACTCTGGGCAA  915
VER.KE.x.9063           AGGCACGT-CA-CTGA-A-G-CGTCGGA-GCGTAGGAACCGCGGGGTG-GACGT-ACC..A-GAA--GC-TGGTG-GTG-GCTT-T-.-AGTGCC-.AA--AA--TCGA....-C--A--T--AGGA-TAGG--GGGCC.-AGC-ATA........ACTACTCTGGGCAA  918
VER.KE.x.AGM155         AGGCACGT-CA-CTGA-A-G-CGTCGGA-GCGAAGGAACCGCGGGGTG-GACGT-ACC..--GAA--GC-CGGTG-GTA-GCTT-T-.-AGTGCC--GA-AAA--TCGA....-C--A--T--AGGA-TAGG--GGGCCG-AGC-GTA........ACTACTCTGGGCAA  906
VER.DE.x.AGM3           AGGCACGT-CA-CTGA-A-G-CGTCGGA-GCGAAGGAAG--GCGGGG-GCGACGTGAC.CA-GAA--GC-TGGTG-GTA-GCTT-T-.-AGTGCC--GA-AAA--TCGA....-C--A--T--AGGA-TAGG--GGGCCG-AGC-ATA........ACTACTCTGGGCAA  406
TAN.UG.x.TAN1           .......GTCA-----ACTTCACTGAAGGG--GGAA--GCATCACCAGTCGAGTT-CCC----G---TAA-TC-GGACGCG-G-AGG.C-A-GAG-ATCC-CCAG-AAA....A-AGTACTGCC-G-GAGGC-GGGG-GG--C-GGTA........GATCTAGGTGGGCG  878
SAB.SN.x.SAB1C          GCCTGAGG-CTCCTTT-ACTGAGTAAGA-C-GACGGAAGAAAGGCGGA-CGTG-CGGAGT-CAGGTAAGA-GC-CCTGC-GA-C-GA.-A--CCCTCG-GCT-TGT-TG....TGT........---A-GAGAGGA----GAGAA-TCAG-C........AGGGCTGACAAGGC  973
MND_1.GA.x.MNDGB1       ..............................AG---GGCACTGA-ACTTGAGGCA--GCACTCC-C--G-AAGAAGCAGGTTGAA-GAG-G.T-G-CTG-T-TGAA-A--C-A....-GAG-TGAGTC-GTGGGACTG-CT--ACAAGAATT........AGTTGTACCCTAGT  425
DEB.CM.99.CM40          GGTGTGCGGAGTCTG-AA-CCCTGAGC-G-TCTCTTCCC--CGGA-.AGCT-CGCAGC-AAGG-AA--GGA-GCCGGAGAATACAGG.-C-GC--ATAAG--C-G-TGT....-TG.-TCTCTGCA-G-AGAGT-CCAC-GTGGAGCA........AGAGAGGCTCCTGG  443
DEB.CM.99.CM5           GGAGTGTGGAGTCTG-AA-CCCTGAGC-G-TCTCTTCCC-TCGAG-GAGCT-CGCAGC-TAGG-AA--GGA-GCCGGAGAATACAGG.-C-GC---TG-G-AC-G-TGT....-TG.ACCTCTGCT-G-AGAGTGTCAC-GTGGAGCA........AGAGG..CTCCTGG  446
TAL.CM.01.8023          ..................................CAG--G-GAAGG--G--T--G-GACGC-CAGGCAGGG-TCTGCGGG-AAGT--G.C-AC-CTTC-AGACAAGCA-C....CGG-A-AA--GTGA-AGCTC-G-GGAGGAGCTC--........GGACCTGCGGGTGG  124
TAL.CM.00.266           CTAGGAAG--T-GC--A--GGCCTAGCTT-ACCCAG--G-GA.GGA-GGA--GA-GACTT-CAGGCAGGG-TCTGC.AG-AAGTTCGT-CGT-CCTC-AGACAAGCA-C....CGG-A-AA--G..AGACTGCTC-GGAGGAGCAC--........GGACCTGCGGGTGG  638
SUN.GA.98.L14           GGCGCCGGGAA-TT-A--T--GTGAGGG--CTTACCA-GCACTGTCAGGGGACA-GCC--A-AGTCCTAG----GGGGGAG-TC-CG.CTAGGTTA-TACT-TCAA-TA....-G-A-ACT----GTA-GTGG--T-AAA-AAGAAGT........GTGTGTATAACTCT 1038
WRC.CI.97.97CI14        GGGATTC-CT------AAGTCGTCC...AGA-.GGTAAGAGGGAGGACT--TGGT-CCT-TTGAGAC--CTAGA-GGAATTAGAGG-.TAGCCTT--AAC-TA-A---TGAGTAGT-CA-A--GGACTACTGGG-G-CAGGAG-AAG-ACC---GGGGAGGC....TATA 1017
WRC.CI.98.98CI04        GGGATTC-CT---G--A-GTCAGCCTC-AGA-AGGTAAG-AGGAGGGCT--TGGT-CCT-TTGAGAC--CTAGA-GGACTTAGAGG-.TAGCCTT-AGAC-CATA---TGAGTAGT-CA-A--G-GCTACTGGG-G-CAGGAG-.GTAACC---GGGGAAGC....TATA 1023
WRC.GM.05.Pbt_05GM_X02  .......................................................................................................................................................................    0
LST.CD.88.524           .......................................................................................................................................................................    0
LST.CD.88.447           .......................................................................................................................................................................    0
LST.KE.x.lho7           GAAAGTCTGCTCGGG-CAGG-AGAGAAGACCTGCTT-AGA-GAGGATAAGAAGGA-TTGACTGCTA-AAG--G-C-GAA-GTAAGGA.-AGT-AGAC-TCACC-G-AGA....-TGGCAGC--GC--GGGGC-GCCT-G-GAG-AGT-........TGTAATTCTGATCT 1043
SYK.KE.x.SYK173         ...........CT-G-CC-GGGACCCGAG--GAGGT-AC--CAGTCAAAGG--GT-CGGATC-AA-AGGGATC-GCTGTC--TC-TA.---CGACC-GTGCT--T---T....CGTG-TGT-CACACGAGACG-GGTAAG--AT-C--........CGCCAGACTGCCCA  770
SYK.KE.x.KE51           ..............CT--GCTAGGGA--C--GAGA----ACGCCAG--AAAGG-ACCGCAGAGGAA-G---GC-CGAA-AAGAG-A.-CGGT-TCCGCCCTATGCCGA....TCGGC-ACC-G-TTGAGTAGGTG-G--CTCCGGC-........TCAGACTGCACGCG  407
COL.CM.x.CGU1           AAGTAAA-GTGCGCGCTCGTCGCGCAGT---GAAGT--CAGACGAGAGAG-ACA-AGAG---G-GCT-GCCAGAG-AGGAGAA-AG-.TCATTAG-AG-TCTTCA---A....-GGA-TA-C-CC-CTGGGC-GGTCGAGG-GGGCCT........GCAAGTGGAGTAAG  567
```

```
                                                   Packaging loops end      Gag and Gag-Pol start
H1B.FR.83.HXB2     ..............................................TAGAAGGAGAGAG.....ATGGGTGCGAGAGCGTCAGTA...TTAAGCGGGGGAGAATTAGATCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGT  893
Gag                                                                                  M  G  A  R  A  S  V  .  L  S  G  G  E  L  D  R  W  E  K  I  R  L  R  P  G  G  K  K  K  Y  K  L  K  H  I  V
H1A1.UG.85.U455    ...............................................-AGGA---A.......-------------------------AA-A-------TC------G-------------------------C---------G-C-G------T---  360
H1B.US.90.WEAU160  ...............................................-AGGA-----.......-------------------G-----------------------A-G-----------------------T-----------C-----------  892
H1C.ET.86.ETH2220  ...............................................-AGGA-----.......-------------------A--.-------A--C-A-A--------GCC------------AA----------------------C-C----TGC-G-----CC----  273
H1D.CD.84.84ZR085  ...............................................................-------------------------------A---G---GCT-----------------------C-----------------G-----GC-------------  410
H1F1.BE.93.VI850   ...............................................................-------------------A--.-------------A--------GA-------------A--------G-------------G-------A-G------C--A-  248
H1G.SE.93.SE6165   ...............................G-----------------------------------------------.--.-C------A--------GCT-----------------G-------------G----TC-------A--------C----  296
H1H.CF.90.056      ...............................................................-----------------------------C---A--------GCT----G---------C-----------------------------GGC--------C----  240
H1J.SE.93.SE7887   ...............................................-AGGA-----.......-------------------A--.------T------A--------GAT-----------------G-------------GGA----G---C----  216
H1K.CM.96.MP535    ...............................................................-------------------------------A--------GC--------G---------C-----------------------C-G------C----  104
H101_AE.TH.90.CM240 ..............................................-AGGA-----.......-------------------.--.------T------A--------GC-----------------GC----------GA---------GGC-G------T----  458
H102_AG.NG.x.IBNG  ...............................................-AGGA-----.......-------------------.--.------T------A--------GC-----G-----------------------------------G-C--------C----  433
H1N.CM.95.YBF30    ..............................AGAAA...........GTAGGA-----G.......-------------------G..-----CA-----A--------A-------TC----TAT--G--A-------------------C-G-A-G-----T----  451
H1O.BE.87.ANT70    ........GACCTAGGGGAAGGGCGAAGTCTC..............-GG-AG--AG.......-----------T-----T--G...--G-CA--AA-CA----G---GC------C-----A------A------ATCT--A--G------G-C-------CT----  945
H1O.CM.91.MVP5180  ........GACCTAGGGGAAGGGCGAAGTCCC..............-GG-AG--AG.......-----------T--G...--G-CA---A-TA---G---GC------CG----A-------------ATCT--A--GGC----G-GGC-------T----  920
CPZ.CD.90.ANT      ........ATCCTAGGGGAAGGGTCGAAGTCTC..............-G-AC--GA-.......-A--.-G-G-----T--T...--G--G-A-AGA-GC------AC-------GT--CA--C-TC----C--T--C--------G--C-TGA-------C-G--  257
CPZ.CM.01.SIVcpzCAM13 ........GACCTAGAGGAAGGGCGAAGTCTC..............---GG---AG.......-----------------G...----CA----GA----G----C---------G----C-T-A--C-------G-------TGA-G------T----  455
CPZ.CM.05.SIVcpzEK505 ..........CCAGGA.......................G-AGGA-G--G.......-----------------.--.--CA--A---A--------A-------------TAT--G--A-------------------C-TGA-G-----CT----  432
CPZ.CM.05.SIVcpzMB66 ..........GG.CTAG.......................-AGGA-----G.......-----------------T..-----CA------A-------CA------G------TATC-G-----------------------TGA-G------T----  442
CPZ.CM.98.CAM3     ........GACCTAGGGGAAGGGCGAAGTCTC..............-GG-A----G.......-----------------.--.-----CA------CG------GCT-----------A--C-------------------TGA-G------T----  283
CPZ.GA.88.GAB2     ........GACCTAGAGGAAGGGCGAAGTCTC..............-GT-----G.......-----------------TA--.----C---C--GCG------GCC----G-----A--C-C--A----------G-----G---C-TGA-G----------  265
CPZ.CM.98.CAM5     ........GACCTAGAGGAAGGGCGAAGTCTC..............-GG-A----G.......-----------------.--.-----CG--A---CG---G---GCT----G------A--C----M----A-----------------TGA-G------T----  589
CPZ.CM.05.SIVcpzMT145 ........GACCTAGAGGAAGGGCGAAGTCTC..............-GG----AG.......-----------------A--.-----C-------AG------GCT-----------A--C----A-----------------C-TGA-G------T----  462
CPZ.TZ.00.TAN1     ........ACCCTAGAGGAAGGGCGAAGTCTCTAGG.........AC---G--A.......-----------------G..--G--G--A-ATA-GC-G---AC------TCC--AA--C-T-AAT-CA-A--C-G-------TT-A--------C----  456
CPZ.CM.05.SIVcpzLB7 ........GG.CTAG.......................-AGGA-----G.......-----------------A-C..-----CA--C---A-G------GAT--------G----C-------C-------GA---G-----TGA-G------T----  439
CPZ.US.85.CPZUS    ........GACCTAGGGGAAGGGCGAAGTCTC..............-GG-A----G.......-----------------T..C-G-CA------CG------GCT--------G-----C-T--A--T-------------G---TGA-G------T----  935
CPZ.GA.88.GAB1     ........GACCTAGAGGAAGGGCGAAGTCTC..............-GG-----AG.......-----------------T..C---CA------A---------C--------G-----C-T-----C-------GA---G-----TGA-G------T----  930
MON.CM.99.L1       GAGGGAGCATAAAATAAGGGCACG.....................................-----------------G----GCAC--G-CGATGC-T--T--TACTA-------CAAG-AT--G--GG-G---C-T--A--CG-C--T-----G--G--CCTGA-T--G--------  548
MON.NG.x.NG1       ..............................................................-----------------.--.--.--
```

```
H1B.FR.83.HXB2            ATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGA 1063
Gag                        W  A  S  R  E  L  E  R  F  A  V  N  P  G  L  L  E  T  S  E  G  C  R  Q  I  L  G  Q  L  Q  P  S  L  Q  T  G  S  E  E  L  R  S  L  Y  N  T  V  A  T  L  Y  C  V  H  Q  R
H1A1.UG.85.U455           -----------------G--AA---A--C----C--------T---------G-------A--CAG-------------AT-------G-T--C--------A-----------------------------------------GT-----------A---------  530
H1B.US.90.WEAU160         -------------------------------C---------------------------------------------AT----------------------------------------------------------------TGT------------------A-- 1062
H1C.ET.86.ETH2220         C-------A---------G--AA--T---C----C----A--T-----T------C--------A------A-TAA------------G-T-----------A---G-------A-------T--------G-----T-----------A-------A--  443
H1D.CD.84.84ZR085         ------------A-----G-------T---C-----------T---------------------A------A--A------------G-T--A----------------------------A---------------------------A---A------  580
H1F1.BE.93.VI850          -------------------------T---C--G---------TC------------------CA-A-----A-AA---------------------------------G---AG------T------------GT--------A---A---------G  418
H1G.SE.93.SE6165          -----------------G--GA---T---C----C----A--T---------G------T---CA-------A--A-----T-G-------T--C--------A---G--A--A-------------G----------------------A----------  466
H1H.CF.90.056             -----------------G---A---T---C----C--C-----T---------C-----------CT---G---A-A-A----A---G---G-TA--A--------A-----------A-------T----CT-------GT---------C--A----G--AA-  410
H1J.SE.93.SE7887          -----------------G--CA---T---C----C--------TC----GT--G--A--------CA----------A-T------C-----G-T--C---------A--C-----A---A-----G-------------------------C--A-----G----  386
H1K.CM.96.MP535           -------------------------T---C-C--C--------T-----G---A-----------C-G------A-AAC---AA----G------A----A---------------GA---A----C-----------A----GT--------T---A-------A--  274
H101_AE.TH.90.CM240       ------------A---T-----A--------C----C---A--T-T---------G-------A---CA------A-A-A----T----GT--A-T--CA---------T-----------A--------T-------------------GG--C--A--C-------  628
H102_AG.NG.x.IBNG         -----------------G---A--------C----C--------T---------G----G--A---CA----C--A---A----T-----T--G-T--CAG--------------------A-------T-------A----------T-GG--C-------------  603
H1N.CM.95.YBF30           -----------------G---A-------TTG---C--A--T--CA-G--C---G-G--C------GCCA-GT----AAAT--AT--G-----G-T--CA-------G-----------GC-C--T---------CG-TC-----GTT--T---------C---AGT----  621
H1O.BE.87.ANT70           ------------A-----G---A--------TG--------AG--AC-------TG---G--TAA-GAGA-GC-G--ACAG---T--G-G---G-T--CA-------G-----CAGC--GCAG---C-C-GG--CG--A---T-GTG----GG-----T--CA-C--AT 1115
H1O.CM.91.MVP5180         -----------------G---A---A----TG--------T--A--------TG-------TAC-GAG---C-G--ACAG---T--G-G---G-T--CA-------G-----G--C--G-A---TC-C-GG--CG--A-----GTA----GG--C--T--CA-C--AT 1090
CPZ.CD.90.ANT             T--------ATC------GC-G--T--T--GC-C-GCT-CTC---TC-----------------T---GA-A-GGCTA-CCAT--AT-GAGC--T---A-AG-A-T-A----CCCT---A-A-T---T--G-T---C--CA-TTGTGTT--G-GG--C--A---A--G--G  427
CPZ.CM.01.SIVcpzCAM13     ------------A-----G---A--------TG---C--C--G--AA-G----GTG-------A---CTTA--T-G--ACAG---T--G-G---G-T--GA-A--C--C--G----GG--GC----T----T---C--TC-----GTGT-A-GG-----T--CAGT-AC-  625
CPZ.CM.05.SIVcpzEK505     -----------------G---A-------GTG---C--A--T--CA-G--C---G-------T---GCC---T----CA---AT--G-----G-T--CA-A-----G-------G--GC-C--TC---T---C--CC-G---GTT--T--C-----T---A-T-AC-  602
CPZ.CM.05.SIVcpzMB66      -----------------G---A-------TTG---C--CTC---CA-G-----A-----------AG--GC-TT--CAG------G-----G-T--CA-A-----T------G---------T--G-T---C--CA---T-GT---T-GG--C---------G---  612
CPZ.CM.98.CAM3            ------------A-----G--CA------TTG---C-----T--AA-G----GG-------A---GAG--GT--T-AA-G--AT--G-----G-T--CA----------------GC--GC----GC-T-T------TC-----GTG--T-GG-----T---A-G--A-  453
CPZ.GA.88.GAB2            -----------------G---A-------TTG-G-C-----T--AA-G-----AA-----------TAT--GT--T--AA-------G-----G-T---A-A--------------G-T-AC-G--C----T------TT-G--TGTGT-A-GG-----T---GCT--A-  435
CPZ.CM.98.CAM5            -----------------G---A----T--TTG---C-----T--AA-G--G---G-T-----A---GA-A-TC---AAA----T--G-----G-T--CAG-----------G-GCT-GC----GC---T---C--TC----GGTA--T-GG-----T---A-G--A-  759
CPZ.CM.05.SIVcpzMT145     ------------A-----G---A------TTG---C-----T--AA-G---ATG-------T---GAGA-GT--T--AA-------G-----G-T---A------G-------G---GC-C--T----T---C--TT-G---GT---T-GG-----A---A-G--A-  632
CPZ.TZ.00.TAN1            ------CG-A--C--A---C-G--T-----GA-G-----C--T--CA-G--G-ACGT---------CT-GA----CA-CCTC-----G--G--T--GG-AG-C-TT--T--TCC----A-C-TT--T--G-T------CA-CTGTGTA-----C--C--A--CGC-G-AG  626
CPZ.CM.05.SIVcpzLB7       -----------------G--GA---T--TTG---C--C--G--TA-G-----A----------GAG---T-GA-TAC----T--G-----G-T---A------G-----A-C----A---T----T---C--CA---TGGT-----GG--C---------G--A-  609
CPZ.US.85.CPZUS           T----------------G---A---T--TTG---C--A--T--AA-G-----G----C-----CTC---T--T-AAA----T--G-----G-T--AA-A------G---GC--GC----GC-C-T---C--CT----GGT---T-GG-----C--CAGT---G 1105
CPZ.GA.88.GAB1            -----------------G---A--------TG-G-C--C--G--AA-G---GTAAG-----A----CTA--T-GT-ACA---AT--G-G---G-T--CA-A-----C-------G---GC-G--C--G-T---C--TC-G---GTA--G-GG--CA-A---AGTGAC- 1100
MON.CM.99.L1              ------TGC--A------T--TA------GTC-CTCGGACTCG--CC-G--G---CAG--C--------AGA-G--TT---A-GTTA-T-TT--TCTA-AG-C---------C--GTCTA-C-A---TC-C-T-GG--TC-CTT-GGTAT-G-----CA-A---GCCG-T-  718
MON.NG.x.NG1              ...........................................................................................................................................................    0
GSN.CM.99.CN166           ------CTC--AA-----T--GA---AT--GC-GTCCGAC-CT--TC-T----ATCA---G--G---C---GG---T-A-A-ATTG-CTTC--CCTAGAA-CA-----G-------CT--C-A---TC-C-T-GGA-TC--GT--GTG---------A-C---GC-GAC-  725
GSN.CM.99.CN71            ------CTC--A------T--GA---A---GC-ATCAGGCA-T--TC-T---T-TAA------G---CA---GG--T---A-GTCACCTCC----TAGAA-CA-AT--G--C---TCT--C-A-AGCC-C---GGA-TC--GT--GTG---------A-C---GC-GAC-  731
MUS_1.CM.01.1085          ------TTC---------T--TA----T-G-C-CTC-GA-TC---T--G---T--AGG--T--G---CA-A-G--CA-A-AGGTCA-C-TC--C-TAGAA-CAG----G--C--GTCC--C-A----C-C-T-GGA-TC--GT--GTG----GG---A-C--CGC-GA-G  723
MUS_1.CM.01.CM1239        ------TTC--A------T---A-G--T-G-C-CTCGGACTC---T--G--G---AA---G--A---CA-A-G--CT-A-A-GTT--CAT-----TAGAA-CA-----G--C---TCT--C-A----C-C-T-GGA-TCACGT--GTG----GG---A-C---GC-GA-G  725
MUS_2.CM.01.CM1246        ------TTC---A-----T--G--G----GTC-CTCAGA-ACG--CC----G-GCAAG--C--A--CCA--GC--C--A-A-GTAA-CTGC--C-TGGAAG-A--T--GAGC--G-CG--C-A----C-C-T-GGA-TC--GT--GTA----GG---A-C---GC-GACG  794
MUS_2.CM.01.CM2500        ------TTC--AA--A--T--CA------G-C-CTCAGACTC---T--------TAAG-----G--CCAG-GC--C--A-A-GT-A---T-----TGGAG-C---C--G--C--GTCT--C-A----C-C-T-GGA-TC--GT--GTA---------A-C--CGC-GA-G  753
RCM.NG.x.NG411            -------T-------A-----GA----T-GCT-G-G-GA-ACTT---------GTG-------AATCCAGA----TA-A--GGTTA-T-TG--TCTAG---C------G--T--G-GCT-A-A---CC-C-T----TTGTGCTGCGTA--T-GG--C-----CA-GGAAG  485
RCM.GA.x.GAB1             -------T---AAA-A---A-TAA---T-GCT-G-G-GA-CATT-----------G--AC---A---GA-A-----T-A---GTC--G-TG--TCTAG---C------G-----G-GG--A-A-AGCC-C-T----TTGTGCTGCGTA----GG--C--A--CA-GGAAG  485
DEN.CD.x.CD1              T---T---AG-A----A-G--G-----T-GCT-AGGAGAACAG---C-------C-AT-----A--CCA-ACT--CA-A-AGACCT--GCT----TAGAA-CACA---CAGT--GAGT--G-A-AGCC---TC-GA---TGCTGCGTTG---GG---TGC--CAGGG-AT  755

MAC.US.x.239              -------GCA-AT--AT----TA----T-G-T-AGCAGAAA-------G--G-ACAA------A---CA-A-------TTCGGTCT--GCT----TAG-G-CA-----C------A-TT-A-A-AGCC-T--------T--CTGCGT-A---GG--CA-T--CGC-GAAG 1326
Gag                        W  A  A  N  E  L  D  R  F  G  L  A  E  S  L  L  E  N  K  E  G  C  Q  K  I  L  S  V  L  A  P  L  V  P  T  G  S  E  N  L  K  S  L  Y  N  T  V  C  V  I  W  C  I  H  A  E
H2A.GW.x.ALI              G------GCG-AT--AT-G--CA------G-T-GGCAGAAA-------G--GT--AA------T--CCA-AG---T--TAA-GTTT--G------TAG-G-CA------------A-TT-A-A-AGCC-T-T------T--CTGCGTAG-T-GG--C-----CGC-GAAG 1369
H2A.DE.x.BEN              G------GCG-AT--AT-G--TAA-----G-T-GGCAGAGA-------G--GT--AA------T--CCA-A-G--T--CA--GTTT--G-T----TAG-A-CA-----G------A-TT-A-A-AGCC-T-T------C--CTGCGT-A-T-GG--CT----CGC-GAAG 1376
H2A.SN.x.ST               G------GCG-AT--AT-G--CA------G-T-GGCAGAGA-------G--GT--AA---G--T--CCA-A----T--TAC-GTTT--G-T----TAG-A-C------G------A-TT-A-A-AGCC-T-T------T--CTGCGT-A-T-GG---A-A--CGC-GAAG  821
H2B.GH.86.D205            -------GTG-AC--AT----CA----T-G-T-AGCAGAGA-------G---T--AA------A---CA-A----T--TAA-GTCT--GCT----TAG-A-CA-----G------A-TT-A-A-AGCC-T-T-----TC--CTGCGT-A-T-T---CC----CGC-GAAG 1352
H2B.CI.x.EHO              -------GTG-AT--AT----GA----T-G-T-GGCAGAGA---G---G-G-T--AA------A--CC-GA----T-G-AA-GTCT--GG-----TAG-A-CA-----T------A-TT-A-A-AGTC-T---------C--CTGCGT-A---T---CC----CGC-GAAG 1353
H2G.CI.92.ABT96           G------GCG-AC--AT-G--CA------G-T-AGCAGAGA-------G---T--AA------T--CCA-A----T--TTC-GTAT--GCT----TAG-A-CA-----T------A-TT-A-AGAGCC-T--------T--CTGCGT-A-T-GG---T----CGCCGAAC  739
H2U.FR.96.12034           G------GCG-AC--AT-G--CA----T-G-T-AGCAGAAA-------G--GG-CAA------A---CA-A----TT-ATCGGT-T--GCT----TAG---CA-----T------A-TT-A-AGAGCC-T--------T--CTGCGT-A-T-GG---T----CGC-GAAG  855
MAC.US.x.EMBL_3           -------GCA-AT--AT----TA----T-G-T-AGCAGAAA----C--G--G-ACAA------A---CA-A-------TTCGGTCT--GCT----TAG-G-CA-----C------A-TT-A-A-AGCC-T--------T--CTGCGT-A---GG--CA-T--CGC-GAAG  812
SMM.US.x.H9               -------GCA-AT--AT-G--CA----T-G-T-AGCAGAAA-------G--G-ACAA------T---CA-A----T--TTC-GTTT--GCT----TAG---C------T------A-TT-A-AGAGCC-T-----C--T--CCGCGT---T-GG--CA-T--CGC-GAAG  808
SMM.US.x.PGM53            -------GCA-AC--AT-G--CA----T-G-T-AGCAGAAA-------G--G--CAA------T---CA-A----TA-TTC-GTTT--GCT----TAG---C-G----T------A-TT-A-AGAGCC-T--------T--CTGCGT-A-T-GG--CA-T--CGC-GAAG 1252
SMM.US.x.SIVsmH635F_L3    -------GCA-AT--AT-G--TA----T-G-T-AGCAGAAA-------G--T-ACAA------T---CA-A----T--TTCGGTTT--GCT----TAG---C------T------A-TT-A-AGAGCC-T--------TA-CTGCGT-A-T-GG--CA-T--CGC-GAAG 1338
SMM.SL.92.SL92B           -------GC---A--AT----CA----T-G-TCGGCAGAAA-------G----GCAA------A---CA-AG----T-A-C-GTA---GCT----TAA-G-CA-----C--G---A-TT-A-A-AGC--G-T--G---T--CTGCGTAG-T-GG---C----CGC-GAA-  766
STM.US.x.STM              G------GCA-AC--AT-G--CA----T-G-T-GGCAGAAA-------G--GT--AA------G---CA-A----TA-AACGGTCT--G------TAG-G-CA-----T------A-TT-A-AGAGTC-T-T------T--CTGCGT-A-T-GG---A-C--CGC-GAAG  982
MNE.US.x.MNE027           -------GCA-AT--AT----TA----T-G-T-AGCAGACA-------G--G-ACAA------A---CA-A-------TTCGGCCT--GCT----TAG-G-CA-----C------A-TT-A-AGAGCC-T--------T--CTGCGT-A---GG--CA-T--CGC-GAAG  806
DRL.x.x.FAO               C----T-GCG-A---A-----T-------GCC-GC--GAGA-AT-A-------GTCAG-----G--CCA-A----TT-A---GTA--CTTT--TCTAG---CT-----G-----GA-TT-A-TT--TC-C-T------CTGCTGCGTAG-T-GG--C--T--CTCG-AAG  564
MND_2.x.x.5440            -----T-----A---A-----CA----T-GCC-AC--GAGAAGT---------GTCAG-----A--CGA-A----T--TA-CGTA--CTTT---CTAG---CT-----G------A-TT-A-TT--GC-G--C--C--CTGTTGTTG-A-T-GG--C--A---GCG-AAG  567
MND_2.CM.98.CM16          -----T-----A---A-----CA----T-GCC--C--GAAAAG----------GTAAG-----A--CGAGA----T--TA-CGTA--CTTT---CTAG---CT-----G------A-TT-A-TC--GC-G--C--C--CTGCTGCTG-G-A-GG--C--A---GCG-AAG 1010
MND_2.GA.x.M14            -----T-----A---A-----T-----T-GCC--C--GAAAAA--A--------TCAG-----A--CGAGA----T--TA-CGTA--CTTT---CTAG---CT-----G-----GA-T---CTTG-GC-G--C--C--CTGTTGTTG-A---GG--C--A---GCG-AAG  934
OLC.CI.97.97CI12          -------GTACAA---T-GA---ATC-ATGCA-AG----A-ATGATC---G---TAAG-----------A-AG---TG-A-A-A-AT---TG--T-TGGAAG--GG----A-TC-GA-GT-G-A-AGT--GCTG--C-TCTGTAT-GTTA-T-TG---T----AG--G-CC  979
GRV.ET.x.GRI_677          -------G-A-AA--AA-G--------T-GGT-AC--GAGAAA--T---------AA----------CA-A---CA-A-A-GTTT--ACC--G-TGGAA-C------C--C--G-GG--A-A-G-TC-G-T----TTGTGCTGCGT-A-T-GG--CA-T--CGC-GAAC 1170
VER.KE.x.TYO1_patent      -------G---A---A-G--G--C----GCC-CC--GAGA-GT----G--G---GAG--G--G----A-AG---CA-A-A-GTC--CT-C--CCTAGAA-CA--------G--G-GCT-A-A-AGTC-G-TC---CTT--GTGCGTA--A-----CT----CA-GGAAC 1212
VER.KE.x.9063             -------G-T-AA---A-G--G--C----GCC-CC--GA-A-GT----------GA---T--A----A-A-G--CA-A-A-GTC--TTTT--T-TAGAA-CA-----C--G--G-GCT-A-A-AGTC-G-TC---CTT--GTGCGTG--A-----CA-C--C-GCGAAC 1215
VER.KE.x.AGM155           -------G---A-A--A-G--C--C----GCC-CC--GAGAAGT-A--G--G---GAG-----T----A-A-G--CA-A-A-GTT--CTCT--CCTAGAA-CA-----G--G----G-A-G-A-AGTC-G------CTG--GTGCGTAT-GCT---C--C--C---GAA- 1204
VER.DE.x.AGM3             -------G---A--AA-G--G--C----GCC-CC--GAGA-AT-AC-----T--GA------A----AGA-G--CA-A-A-GTA--CT-C--GCTAGAA-CA-----G--G--G-GCT-A-A-AGTC-G-T---CCTT--GTGCGTAT-G-T---C--A--CA--GAT-  704
TAN.UG.x.TAN1             -------G---AAA-AA-G--G--C--T-GGC-CC-CGAGAAA--C----------AA---G------CA-A-G--CA-A-A-GTAT-GAGC--GCTAGAA-CA--------------G---A-A-AGTC-G-T----TTGTGCTGCGT-A-T-GG---A-T--CGC-GAAC 1175
SAB.SN.x.SAB1C            -----------A---A-----T-----TAGCC--TCAG-AAAT--C-----G---AAG-----GGTCGTCA----T--TA-TGTA--C-TT--CCTAG---CT-----G-----GA-TT-A-TTG-TC-C-T----TT-TGCTGCGTGT-AGC----A-T--CGC-GAA- 1290
MND_1.GA.x.MNDGB1         T---TGT-AAG-T--AT----TA----T-GCT-ATCGGA-AAA--CC-T------CAGC----A---GA-A----C--CTC-GTATGTTGG---TATA-G-CCA---------TA-T--A-A-G-T--GGTAGGG-----CTGTGTTG-AGCC--CA-A--CGC-G-T-  723
DEB.CM.99.CM40            C------G-T-A---A-----CA----T-GCC--GCAG-A-CAT-A-------ATAAG---------CAGG----TT--ACTAG-T----G--T-TAGAAGCA-----CAGT--GA-TT-G-A---TC-----GGC-T---CTGTGT-A---GGGCATGC---GCC-ACC  761
DEB.CM.99.CM5             C--------T-----A-----CA----T-GCC--GCAGGA-CAT-A--G-----CAA------A---CA-G-G--CT-AACTAGAT-------T-TAGAGGCAT---CAGT----GC----AG--CC----CGGC-T----TGTGT-G-T-GGGCATGC--CGCC-ACC  764
TAL.CM.01.8023            ------TTC--AA---A-G---A-G--T-GCC-G-G-GAC-C---T--G--G---CAA------G--CGTTA-G--CA-A-A-GCCA-TTCT--G-TAGAA-CA-AT--C-------G-A---AG--TC-C-T-GGA-TC--GTGTGTT--A--C--CA-T---GC-G-T-  425
TAL.CM.00.266             ------TTC--AA---A-G--GA-G--T-GCC-G-G-GAC-C----------G--CAA------G--CGTTA-G--CA-A-A-GCCA-TT-CATG-TAGAA-C--AT--C-------GGA-A-AG--TC-C-T-GGA-TC--GTGTGTT--A--C--CA-T---GCTG-T-  939
SUN.GA.98.L14             -------G-T-AA--AT----TA-G--T-G-C----CAAG-AA------AG----GTG---------CAGA-G--T---TC-GTCTGTTGG--T-TGTA-GCAT-T--T--TA-GA-C--G-AGG-C---GTAGG-------TGTGTG--TGCC---TGT------G-C- 1336
WRC.CI.97.97CI14          ----TT-GTG-A---AT----GAAG--GTGCT-AG--T-A-CAT-AA-CAGT---G-------GGTG-AGA-GC-TA--TC---T---GCT--T-TA--G-CTGAT-CC--TA-G---G-A-AGAGCA-T-GGGGA-TGT--TGTGTGT-AGCA---C-C---G-G--A- 1321
WRC.CI.98.98CI04          ----CT-GTG-A---------GAA---GTGTT-AG-CT--ACAT-AA---C----GA-AC---GAC--A-A--C-GA---A---TT--GC-----TAG---CAGAT------A-----G-G-A-AGCC-T-GGGGA-TGA-GTGTGTG--TGCA---T-----GC-C-T- 1327
WRC.GM.05.Pbt_05GM_X02    T---CT-GTA-AA---T----GAAG--ATGTC-AG--GGGAAAT--A--TCT-GCCA------AGTG-AGA--C-CTGT-AG--TA--GCT--T-TGG-A-CAGA---G---A--AC-A-A---AGTA-T-GGGGA-TGA--AGTGTG--AGCA--CT-AT--A-T-AC-  277
LST.CD.88.524             C-------CG-AA--A-----TA-G--T-GGC-AGGAG-AAA---TC--------G----G-------AGA-G--TT-----GTC-GTTGG--T-TATA--CT-----C--TA--A-T--G-A-G-TC--GTAGGG-----GTGTGT-A---G---CTGT--C-T-G-A-  274
LST.CD.88.447             G-------CA--A--AT----TA-G--T-G-T-AGGAG-ACA-T--C----G---G----T-----C-AGA----T--T---GTCTGTTGG---CTATACTCC-----C--GA-GA-TT-A-A-G--C-TGTAGGA--T---TGTGTG--G-G---CTGC---TT-G---  274
LST.KE.x.lho7             --------CG-AA--AT----CA----T-G-T-GGGAT-ACAA--C--G------G-------G----AGA----C--ATCTGTCTGCTGG--T-TATA-GCA--T--T--TA-GA-C--A-AGG-C---GTAGGG------TGTGT-A---GC--CTGT--C-T-G-AG 1349
SYK.KE.x.SYK173           T-------AG-A----T-G--CA------GCC--TCGGACCAA--CA-G--GT-TAA-----------GA-A----T--CAC-GTC----TT----TAGAAGCT-AT--G------A-TT-A-A-AGCC-T-T-GGC-T-A-TT-GGT-G-A-GGGCC--T---GCC-AA- 1071
SYK.KE.x.KE51             T------G-A-A------G--CA------GCC-CTCAGACCAG--TC-T----GTAA------T---GA------CA-AAA-ACCA-CTT-----TAGAAA-ACAT--C------A-TT-A-A--------T-GGC-T-AC----GTAG---GGGCC--C--CGC-GAA-  708
COL.CM.x.CGU1             ----ATGT---CA---G--AGC--C-GT-TTT-A-TATT--AG------A-GT--G--AC---AGTGGCT------T---A-A--G-GACG----TAG-GG-C-------------TGT-G--G--CC-----GGA...TGTT-TGTGTGT--C--CC----CAG--AAT  874
```

```
H1B.FR.83.HXB2     TAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAA.........................................................AACAAAAGTAAGAAAAAAGCACAGCAAGCAGCAGCTGACACAGGA 1158
Gag                I  E  I  K  D  T  K  E  A  L  D  K  I  E  E  E  Q  _  _  _
```

```
                                       Gag p17 Matrix end‿Gag p24 Capsid start
H1B.FR.83.HXB2          ...CACAGCAATCAGGTCAGCCAAAATTACCCTATAGTGCAGAACATCCAGGGGCAAATGGTACATCAGGCCATATCACCTAGAACTTTAAATGCATGGGTAAAAGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTGATACCCATGTTTTCAGCATTATCAGAAGGAGC 1325
Gag                      .  H  S  N  Q  V  S  Q  N  Y  P  I  V  Q  N  I  Q  G  Q  M  V  H  Q  A  I  S  P  R  T  L  N  A  W  V  K  V  V  E  E  K  A  F  S  P  E  V  I  P  M  F  S  A  L  S  E  G  A
H1A1.UG.85.U455         ...............AG---T----------C--------A--TGCA--A-----CCA----C-----T---------G--C--G----------G-------------C--------------------A-------------------------G-----  780
H1B.US.90.WEAU160       ...A----A--C-----------------------------C-A----A-----------------C-------------------------------------A------A--------------------A----------------------------------- 1324
H1C.ET.86.ETH2220       .........GGAA---A---T--------T-------------T--G-------------------C-------G----------G-------------------G--A-----------------G--A-----------A----T--------------  699
H1D.CD.84.84ZR085       ...A------GC-----------------T--------A------C-A--A-------------C-----C----------------G--C----------------A-------A------------A-----------A-------------------------  842
H1F1.BE.93.VI850        ...............-----T-----------------A-----TC-T-----A-----------C---T-TC---------------------------------G--GA-----------------T-----------A-----------------------G--  668
H1G.SE.93.SE6165        ...A-------C--A--------------T-------------TGCA--A------------C--------A-----------G--------------------------A-----C-----T------------A-------------------------------  728
H1H.CF.90.056           ...A-AGA---CA-------T--------T--------A-----TGCT--A-----G--------C-----------------G--C----------------------A------------T--------A-------------------------------  672
H1J.SE.93.SE7887        ...G---A--G---------T--------T-------------TC-G--A-----CC-----C-----C------------------------------GA------A--A------------------------------------------------------  648
H1K.CM.96.MP535         ......GA---AGG------T-----------------A-----TC-G---------------C-----C-----------------------G--G---A----G---------------------------A-----------A--------------------  524
H101_AE.TH.90.CM240     ...AG-----GCA-A----------------------A--TGCA--A---------C-------C-TT---------------G----------G--------------A----G--T-A--------A----------C---------------G-----  890
H102_AG.NG.x.IBNG       ............AGCAG-----------------------A--TGC-A-A--------AC-------T----G-------------------------G------A-------A----G-----------------A----------------------G-----  856
H1N.CM.95.YBF30         GCAACTGAT-GCA-TA----TAGG-----T---C----C----CTGCT--A--A---------------C-GC-GA---C-----C--------T-----G-----GA----G--------C--T--T--------A----A------ATG--C--G--------G--  889
H1O.BE.87.ANT70         ACAAG-GCA-GG----CGG-T-----------------ATCA--TGCG-----A-------------------C--C--C--G-----------------------G-C---------A--A--C--T-A---T---A-C--T--T-----CATG-----G-----G----- 1371
H1O.CM.91.MVP5180       ACAAG-CCT-GG---ACA--T-----------------AACA--TGCA-----A---------------A-----C--C--C--G---------------------G-C-----------------C--T-A---T---A-T--T--T------ATG--------------G-- 1346
CPZ.CD.90.ANT           CGA--TCTT-G-GGC-AAG-GAG---------C--CA-AGT-G-TGCAGGA--AAT-GCAAGG------C-AC-GA----A-----C--------C--------GTGT-----------AAA-----AT--------C--C--T--------T--T------------G--  737
CPZ.CM.01.SIVcpzCAM13   AGTGCTTC-TC-GGCACT--TGGG---------------A-----TGCT--A-----G---------ACTC----C--C--------C---------------------------G------------------T-AT--T--G--A-----A-----------TC-G-----------  902
CPZ.CM.05.SIVcpzEK505   GCA-CTGAA-GCAGTA----TAGG--C--T---C-G--A--A--TGCT--A-------------------C-GC-G-----C-----C--------T-----------G-------A-----C--T-AT--------A-----A-----ATG--C--G--------G--  873
CPZ.CM.05.SIVcpzMB66    .......-CT-GCACT-----TAGG--C-TT--------AGCA--TGCT--------G-----------AC----C--G---------------------------------------------A-----A--T-----------A--T--------CATG--TC-G-----------  868
CPZ.CM.98.CAM3          CAG--AGA--GCGCT-----TAG---C-----AG----A-----TGCA--A--A--GC-A--G------C----G-----C--G--C--------T-----T------A------A---AA---T-A------G--------A-----CATG--CC-G-----------  715
CPZ.GA.88.GAB2          TCAA---C--C-G--A-A---AG---------AG-G------TGCA-----A--G---A-------------G-----C------------G---C-----T----C--------A--A--C--T----------C--T--A-----CATG--T------------G--  703
CPZ.CM.98.CAM5          CAG--GGAA-GCAGCA-A--TAG---C-----A-----A-----TGCA--A--A------A-------------G-------G--CC------T-----C--G---A-------A---AAC--T-A---------A-----A-----ATG--CC-G--G--G----- 1018
CPZ.CM.05.SIVcpzMT145   GGTG-A-CACCATCA-CT---GCG--C-T----G---A-GA---GCT----A--GT--ACG---A-TC-A--------C-----CC-------C-----------C------A--A--------AT----G--C-----A-----CATG--------------G--  864
CPZ.TZ.00.TAN1          ACAG-G-CACC-AGT-GG--TAG-CTA-----AG-GA-AACTG-TGCA-----AGTTGCAAGG------C-T--T-------------C-------C-------GG---A-------A--A-GG--T-AT--------A-----A-----C---------G--T--G-----  957
CPZ.CM.05.SIVcpzLB7     AGGG---ATG-CAGT-GT---AG---C------------ACA--TGCA--A-----GCC------C------C-C-------------------C-----C--G-----------A-----A--T-----------A------------CG----C---G-----------  871
CPZ.US.85.CPZUS         GGC...--T--CATT-GT---AGC---------AG--A-A------GCT-----A--G---------------A--G--T--C-------------C------------CG---------A-----C--T-A---G-----C--C--G-----CATG--C-----------G-- 1373
CPZ.GA.88.GAB1          AGTGC-TCTGC-GGCA-T--TGG----------CC-C--A--A--TGCT--A-----G-----G------------C-----C-----C--G--------------------G-----G--A-----A--T-----T--G--A-----A-----------TC-------G----- 1386
MON.CM.99.L1            CCAACAG--GGCGTACC---TGG---C------G-C---AGA-CACAGGG---AGGCT-CCA---------AG-GGAG---C-GCTCC-C--GA--------C-G--CA-T--G------AAA--TGC-----------G-GG--C-C--CCAG--TC----------CAT  989
MON.NG.x.NG1            ...............................................................................................................................................................   0
GSN.CM.99.CN166         GGGGGACAGGCGGTA--T--T-----C------G-GA-TAG---TGC---A--A--GTATCA----------C-CAATAG-----T-C-T--GA-T-------GCA----G--G-----AAAG--TGCT-----GACAG-TG--T-----CAG--T-----------CTG 1002
GSN.CM.99.CN71          GGGGGACAGGCGGTA--T--T--G--C-----C-----TAGA--TGCT-----A--GT-TCA---C------C-CAACAGC---GTCC-T--GA-T-------GCA----G-----A---AAG--TGCT-----GACAG-CG-TC-----CAG--TC--A-------CTG 1008
MUS_1.CM.01.1085        CAAGCGCAGGCAGCT...--TGCC--C-----AG-CA-AAG----GCA--A-----GT-CCA----------C-CAATG------TA-----AA-------GTC-A----G--G--A---AAA--TGCTG-------TG-GG-A------CAG-----GG----G--G--  973
MUS_1.CM.01.CM1239      ACAGCACAG-----AACT--TGCT--------C--CA-AAG----CCA-----A---T-TCA---C---T--C-TAATG------TC-----GA-------GTC-A----G-----A---AAA--TGCAG-------AG-GG-A------CA----C--G--------A-  975
MUS_2.CM.01.CM1246      GCAA-AGC--CCACCAGTG----G--------AG-CA-AAGA-CAGCA-----A--GTATCAG-------T-TC-TAGC-----GCTA-----AA-C------TC-AC------G-----AAAA--TGC----------AG-GG--C----CCAG--CC-----------TG 1056
MUS_2.CM.01.CM2500      TCAGGAGCAGGGGCAAG---------C-----AG-GA-TAG--CAGCT--A---ACCTATCA-------T--C-CAGC--C----T-------GA-C-----GTCCAC-A-T----------AAA--TGC----------AG-GG-TT----CCA---C------------TG 1018
RCM.NG.x.NG411          GTGACATCT-G-GGACAA--AGG---------A-----AACA-TT-AT...CA---GCCT-A-------C-A-----T--------A--------T------------------G--A---AAA--TG-AG-------AG-G--T--------G---C-G---------CTG  753
RCM.GA.x.GAB1           GCG-CACCT-G-GGACAA--AGG------T------A-TACT-TA-AT...CA---GCCT-AG---A-TC-T--------A--G---C--------C-----C--G--G-----G----AAAA---TCAG-------AGCG--------C--G---C-----------CTG  753
DEN.CD.x.CD1            TCCTTT--AG---CAAATTCT--------T--A--T--CAG--CAGCT-----A---TATCAG-----A---T--AGTA-C--GGTG-----AA-----A-T----AG-----G--C-GAAAA--TTCG-----GA-AG-G--AC-A-----G-TCC-CG-C-----GCT 1092
                                       Gag p17 Matrix end‿Gag p24 Capsid start
MAC.US.x.239            GCA-CATCT-GCGGCAGAG-AGG---------AG--CAA--A-TAGGT...--TA-CTAT--C--C-T-C-AT--AGC--G-----A--------C---------T-GA----G--A---AAA--TG-AG-------AG-G--AGGA---CAG---C-G--------TTG 1594
Gag                      A  P  S  S  G  R  G  G  N  Y  P  V  Q  Q  I  G  .  G  N  Y  V  H  L  P  L  S  P  R  T  L  N  A  W  V  K  L  I  E  E  K  K  F  G  A  E  V  V  P  G  F  Q  A  L  S  E  G  C
H2A.GW.x.ALI            GCA-CACCT-G-G-ACAGG-GGG---C-T---CG--CAA--AGTAGC-...--CA-CTACACC---GT-C-GC-GAGC--CC----C---G----T---------T---------A---AAG---G-GG-------AG-G--AGGA---CAG---C-C--------CTG 1637
H2A.DE.x.BEN            GCA-CACCT-G-GG-AAA--AGG---C-----CG-GCAA--AGCGGGT...--CA-CTAT--C---GT-C-AC-GAGC--CC-----C-----------------T----G--G--A---AAG---G-GG-------AG-G--AGGA---CAG---C-C--------CTG 1644
H2A.SN.x.ST             GCA-CACCT-GCGG-AA-G-AGG---C-T---CG--CAA--AGTGGC-...--CA-CTATACC---GT-C-AC--AGT--CC----CC--------T---------C--------G--A---AAG---G-GG-------AG-G--AGGA---CAG---C-C--------CTG 1089
H2B.GH.86.D205          GCA-CACCT-GCGGC-GA.........-----T--AG-GCA---ACTGGCT...--CA-CTAC--C--C-T-C-GC--AGC--CC----C--------T--------GT-----------A---AAG---G-GG-------AG----AGGA---CAG---C-----------TG 1605
H2B.CI.x.EHO            AAA-CA-CT-GC-GCC-A......GC---T--AG-GCA---A-TAGCT...--CA-TTATTCC--C-T-C-GC--AGC--CC----C--------C---------T--------------AAG---G-GG-------AG----AGGA---CAG---C----------GTG 1606
H2G.CI.92.ABT96         GCA-CACCT-G-GGCAGAG-AGG---C-----AG-GCA---AGTAGG-...--CA-CTAT--T----T-C-AT--AGT---C-------------------------T-G--------A---AAG---G-GG--------G-G--AGGA---CAG--GC-G--------CTG 1007
H2U.FR.96.12034         GYT-CACCT-G-GGC-G---AGG---------GG-GCA-----TAGGT...--CA-TTAT-CC--C-TAC-TC--AGT--------AC-C-----C---------C-G-----G-----AAAA--TG-AG-------AG-G--AGGA---CAG---C-C-----G--CTG 1123
MAC.US.x.EMBL_3         GCA-CATCT-GCGGCAGAG-AGG---------AG--CAA--A-TAGGT...--TA-CTAT--C--C-T-C-AT--AGC--G-----A--------C---------T-GA----G--A---AAA--TG-AG-------AG-G--AGGA---CAG---C-G--------CTG 1080
SMM.US.x.H9             GCA-CACCT-G-GGCAGAG-AGG---------AG-GCA---AGTAGGT...--CA-TTATACC--C-TAC-TT--AGT--A-----A--------T---------T-G-----------AAAA--TG-GG---G--AG-G--AGGA---CA---GC----------CTG 1076
SMM.US.x.PGM53          GCA-CACCT-G-GGCAGAG-AGG---------AG-GCA---A-TAGGT...--CA-TTACACC--C-TAC-TT--AGT--A-----A--------T---------T-G-----------AAAA--TG-GG---G--AG-G--AGGA---CA---GC----------CTG 1520
SMM.US.x.SIVsmH635F_L3  GCA-CACCT-G-GGCAGAG-AGG---------AG-GCA---AGTAGGT...--CA-TTAT--C--C-TAC--T--AGT--A-----A--------T---------T-G-----------AAAA--TG-GG---G--AG-G--AGGA---CA---GC----------CTG 1606
SMM.SL.92.SL92B         GCT-CACCT-G-GGA-GA......--C--T--AG-GCA---AGTAGGA...AATA-TTAT--G--CACTC-AC-G--C--AC-G--AC-C-----G-----C---C-G--------A---AAA--TG-AG-------AG-G--AGGA---CAG---C----------CTG 1028
STM.US.x.STM            GCA-CACCT-G-GGCAGAG-AGG---------AG--CA---AGTAGG-...--CA-TTAT--C--C-TAC-AT--AGC--A-----C--G--------------GT----G--G--A---AAA--TG-TG--------G-G--AGGA---CAG---C----------CTG 1250
MNE.US.x.MNE027         GCA-CATCT-G-GGCAAA--AGG---------AG--CAA--A-TAGGT...--TA-CTATACC--C-TAC-AT--AGC--G-----A--------C---------T-GA----G-----AAA--TG-AG-------AG-G--AGGA---CAG---C-G--------CTG 1074
DRL.x.x.FAO             GCGACATCT-GCGGCAGGTCAAG---C-----C--TCA-GTTGTT-AT...CAA--GGCA-----C-----T--T-----A-----C--------C-----G--G---A-T--G--A---AAG-----TG--------G-G--T-----C-TG--CC-G--------CTG  835
MND_2.x.x.5440          GCGACACCTGC-GTTCG-TCAA----------C--TCA-GTA-TA-AT...CAAAC-CCA----------GA--T-----A-----C--G---------------TGTA----G-----AAAG------------A-AG-G--T------ATT--C--G---------TG  838
MND_2.CM.98.CM16        GTGACACCT-G-GGCCG-TCAA----------A---CA-ATA-TA-AT...CA-ACCCCA--C--------GA--T-----GC-C--GC-G-----G---------TGCA-------------AAG--T----------A-AG-G--------CAT---C--G--T-----GTG 1281
MND_2.GA.x.M14          GCGACACCT-G-GGCCG-TCAA-------T-----TCA-ATA-TA-AT...CAAAC-CCA--G-----A-GA--T-----A---------------C-----C---TGTA-T--G--A---AAA-----T-------A-AG----T-----CAT---CC-----------TG 1205
OLC.CI.97.97CI12        .......-AG--AG-A-AGCAGTT-GCCATG--A---A-CAGA--TGCT-----AATCT-T-AG---GT-C-GC-GG-TA-------CC-GC-AAGT----------C-T-G-----C-G--G---T-AAT-----A-AG-G--TC-A-----T--G--TG-T--G---TT 1193
GRV.ET.x.GRI_677        GCG-CACCTGG-GGC-AATCAAG---------AG----AA-TC-G-AT...AATGCCTG------C---C-TT-G--T--GC-C--G--------G-----C---TGC--G--G--A--AAGG-GGG-AG-------AG-C--------CCA----C-C-----G---TG 1438
VER.KE.x.TYO1_patent    GCG-CACCTGG-GGC...--T--G----TT--AGCGCAA--AC-AGGA...AATGCCTG-------GTAC--T-G----GC-C--C--------G----------C------G-----AAAA--TG-AG-----A-AG-----------CA---CC----------CTG 1498
VER.KE.x.9063           GCG-CACCTGG-GGC...--T-------T---AGCGCAA--AC-AGGA...AATGCCTG-------GTTC-AC----C---C-C--A--------G----------C------G--A---AAA--TG-AG---GA-AG-G--A-----CCA----C-C--------GTG 1501
VER.KE.x.AGM155         GCGACACCTGG-GGC...-----------TT--CGC-CAA---C-AGGG...AATGC-TG---G---GTAC-AC-T------C-C--CC------------------C-------------AAAA--TG-GG-----A-AG----------CCA---CC-C--------CTG 1490
VER.DE.x.AGM3           ACAGTGCCACC-GGT-G---T-------T---AGC-CAA--AC-GGGA...AATGC-TG-A-----GT-C--T-G-----AC-C--C--------G----------C------G-----AAAA---G-AG-----A-AG-G--------CCAG--T-----------GTG 1002
TAN.UG.x.TAN1           ACA-CT--TGGCAGA...CATGGT---------G----A--AC-G-AT...AAT---TG-------ACTC-TT-G-----AC-T--CC-------------T---AC---T--G--A--AAGG---G-GG-----A-AG-----------CCA---CC-C--------GTG 1461
SAB.SN.x.SAB1C          GTGACACCT-G-GGT-G-TCG-G---------C--T--CAGTGTT-AT...AAT---TG---C--C---C-AT-G--C--A--G--AC-T-----T-----------GA-T--G--A---AAG-----TG---G--AG-----------AGC--CC--G---------- 1591
MND_1.GA.x.MNDGB1      ....................--TA-G----T----G--CAAAG-G-TGCAGCA--A--GTATCAGT--ACTC-A---AGT-----G-T-A--C-GA-----------AC---G-----A---AAG-GG-AA--G--G--C--C--TC-A--C---------GA----------  952
DEB.CM.99.CM40          GCA-C-CAG-CAACACCATCAGTT--C---------CT-AAGAC-GGGG...CAA--GT-T--G---AT-C-AT--AGC--C--G-TAG----AA-C---A----T-CT-----G--A---AAA--TTCT-----GA-AG-T--TC-T---CA--TC--GG----G--CTG 1095
DEB.CM.99.CM5           CAG--A-AA-CAA-CCAACCAGCT----TT--A--CC--AGAC-AGGA...CA----TAT------AT-C-TT--AGT-----G-T-G-G--AA-C---A----T-C---G--G--A---AAA--TTC------GA-AG-G---C-A---CA--T---GG-------GTG 1086
TAL.CM.01.8023          GTGACATCT-G-GGCCAT--TGGG--------C--TA-AAGA--TGCA--ACAA--GTATCAG---------T-G-----C---CTGC-G--AA-----A--GCCACTA-------A---AAA--TGC------GACAG-GG-GC-T--C--T---C--G----------TG  711
TAL.CM.00.266           GTGACATCT-G-GGCCAA-CAGG---------C---A-AAGA--TGCA---CA---GTATCA---------AC-G---------GCTGC----GA-----A-TGCCACTA---G-----AAAA--TGC---T---ACAG-GG-GC-C--C--G---C--G-G-----GTG 1213
SUN.GA.98.L14           ..................TCT-------TT--AG-GCA-AGAG-GGGA...CAAA-TTACA-T-------C-TC-C--G---------AG-GC-AA-------G--GA----G-----------GA-GG-AA------ACAG-GG-A------G----T---ATGAC------ 1559
WRC.CI.97.97CI14        ..................TCT--G---ATG--AG-GC--AGA-CAGCA...CAAGGGCCTCA-TGGG-CC--T-G-----------AA--C-AA-------G---CATA-T--G--T-G--GG--TG-AG-----AGTG-T--AT-C---G----TC-TA----CA-G-- 1529
WRC.CI.98.98CI04        ..................TCT-----CATG--AG--T--AG--CAGCT...CAAGG-CCTCA-TGGG-TC-TC-G--T--A-----CA--C-GA-T-----G---CACA---G--T-GA-GC--TG-AG-----TGTG-T--AT-C-----T--T--GA----TA---- 1547
WRC.GM.05.Pbt_05GM_X02  ..................TCAGTG---ATG--AG-CC-TAGAGGACCA...CAAGG-CCACA-TGGG--C-TC----C--G--G--CA-TC-AA-----------CATA-------T-G--GA--TG-AG-----TGTG-T--AT-C-----T--T--GA----TA----  491
LST.CD.88.524           ......CCTGTG-CA-AA--TA------TT---T-GA-TAGAG-A-AT...CAAAG-TG---C---ACCC-TC-------C-----AA-TC-AA-------G---A----------T-G--GA-GG-AG------ACAG-GG-T-----------TC-GA-----AA---  491
LST.CD.88.447           ...G-A--GC-G-CTAAA--TGGC-----T--AC-C--CAGAG-A-AT...CA-AG-TG------CACAC-AT-G-----------AA--C-AA-T----------A----------T-GA-GA-GG-AG------ACAG-GG-A--------T--CC-CA-----AA---  497
LST.KE.x.lho7           .........C-A---ACAG-GGG------T--CC-GA-AAGAG-A-AT...CA-AG-TG-------ACAC-TT----G--C-----AA--C-AA-------G---A----G-----C-GA-G--GG-AA------ACAG--G-A-----------C---A-----AA--- 1581
SYK.KE.x.SYK173         GCAGCA-CACCG-CACCT--AGG---C--T--AC-GC-CAGA--TCCT--AAAT--GTG-A-T---ACA-GAG-CC--GT---G---C----AA-------GG-G-CT--GA-CTCC--AAAA--TGATG--TCCA-AG-G--AT-A---CAG--TC--A--------TT 1405
SYK.KE.x.KE51           ATG-CATCT-G-GGC-G-TCAGG----------C--A-TAGA---CCA--AAAT---TG-A-T---GTA-GGG-CAACA-C--G---------AA-C-------G-G-C---GA-CAGC--AAAA--TGATG-CTC-C--G-C--AC-A--CCA-ATT---A-------TTT 1033
COL.CM.x.CGU1           ......-AGG-AGCA-AG-AATTGG-CATG--A-----AACAGGGCCT...CAAGGGCCA--G-----AC--C-GAGC--C--G--G---GGA--C-----T--GTGT-----G...GGA-GAA-TGCT--TTCCT--GCT--------C-T---T-AT--TACG--G-- 1133
```

```
H1B.FR.83.HXB2            CACCCCACAAGATTTAAACACCATGCTAAACACAGTGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGCTGCAGAATGGGATAGAGTGCATCCAGTGCAT...GCAGGGCCTATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTG 1492
Gag                       __T__P__Q__D__L__N__T__M__L__N__T__V__G__G__H__Q__A__A__M__Q__M__L__K__E__T__I__N__E__E__A__A__E__W__D__R__V__H__P__V__H__.__A__G__P__I__A__P__G__Q__M__R__E__P__R__G__S__
H1A1.UG.85.U455           ------------------T-TG-----G--TGT------------C--G-----T-----------------T-----------------------G-----C--GT-A------------...------------C--------------------------------- 947
H1B.US.90.WEAU160         ---------------------------------------------------------------------------------------------------------T-A-------------...---------------------------------------------- 1491
H1C.ET.86.ETH2220         ------------------------------T-----------------------------------------C--------------G--------------C--GT-A------------...---------G-------------A--------C------------- 866
H1D.CD.84.84ZR085         ------------------------------------------------------------------------------------------------------C--GC-A-----------G...---------G------------------------------------ 1009
H1F1.BE.93.VI850          ---T--CAC---------------------T-----------------------------------------C-----------------------------C---T-A------------...---------GCCC----------------G-----T-----G---- 835
H1G.SE.93.SE6165          ------------------TCT----T----------------G-----------T---------C----G--T--T--------------------G-----C--GA--------CAA--G...---------T--C-------------A------------------- 895
H1H.CF.90.056             ------------C-----TG-T--------T-----------------------------G-----------T--A--------------------------C--G--A------------...------------C----------A--------------------C- 839
H1J.SE.93.SE7887          ------G-----------T-------------------A-A--------C--------T-----------------T--T--------------------------C--G--A--------A---...------------------------G--------------------- 815
H1K.CM.96.MP535           ---T---------C----------------------------------------------------------T--T--------T-----------------C--GT-A--C---------...-----------CC----------A-----------G-----G---- 691
H101_AE.TH.90.CM240       ------------------T-TG--------T-T------------C--G-----A-----------------A---------------C----------------G--A--C-----A---...------------C----------------G---------------- 1057
H102_AG.NG.x.IBNG         ------G-----------T-TG----------T------------C--G-----A-----G-----------T--------------------------------G--A--------A---...------------C----------------G---------------- 1023
H1N.CM.95.YBF30           A--G--CTC----C----T--T---T----T-----A-----------G-----A-----G---C-G--G--AGT---------------A-----C--------GACA--------C-C-...-TG--A--AC-AC-C-----G--AC-------C--T--A------- 1056
H1O.BE.87.ANT70           A-TTT-CT-T---A-T--T--T--------TG-CA-A--A-----------G--TT-A---G--C----G--AGTA--------------A-T--G----------ACT--C---CCA-CG...-T------GT-GC-------G-----A--G-------CA------- 1538
H1O.CM.91.MVP5180         TGT---CT-T---A-C--T--------G--TG-CA-A--------C----GG--TT-A---G----G--G--AGTA--------------A---------------ACT------CCAGCA...ATG-----GT-AC-------G-----A--G-------CA------- 1513
CPZ.CD.90.ANT             A--T--T--T-----------------T--TG---T----AC-------G---------GG--C--------AGTA--------------A--T--G--------GT-A--C--CACT---...-----A--AG-ACAGG----A--AT-A--G-------CA------- 904
CPZ.CM.01.SIVcpzCAM13     TCTA--TTTG---G-T-----A-----G--TG-T--A-----------G-G---------G-TC----G---GT----------A------------------C-GA-------T--T---...--C--C--C--A--T-----A---C-A---------C----T---- 1069
CPZ.CM.05.SIVcpzEK505     A--G--C-----C------------T----T-----A-----------G-----A-----G---C-G--G--AGTT--------------A-----G-----C--G-GA--------A---...ATG-----A--AC-------A--AG-A--G--G-----A------- 1040
CPZ.CM.05.SIVcpzMB66      ------T-----C-----------------------------------------C--G-----A---------C-------AGTA--------------------G-----C--GC-C--------T---...-----A--G--C--------A--A---------------------- 1035
CPZ.CM.98.CAM3            A--T---------G-T--T------T----TG-TA-A--A--G--C--G-G---T----GG--C--G--G--AGT---------------A---------------ACA---C--T------...-----T--AG-A--C-----A---C-A----GG--C-A-------- 882
CPZ.GA.88.GAB2            A--A--C------G-G-----A-----G--TG-TA-A--A-----C---G----T----GG-AC---------GTA--------A-----A--G--G---------T-A--C---T-A---...-----A--AG-G--------T--A--------G-----A--C---- 870
CPZ.CM.98.CAM5            T--T--C--G---G-C--T--T-----T--TG-C--A--A--G--C--G-G---T----GG-T------G--AGT---------A-----A-----G--------GC-A--C--T------...---------A--A--C----A--AC-------GG---A--G----- 1185
CPZ.CM.05.SIVcpzMT145     A--G--T-----CC-C--T--T--------T---A-A--A--------G---------------TC------AGTA--------T--G-----T--G-----C--GACA--C--TACC---...-----------A--T-----A--A-----G--------A--C---- 1031
CPZ.TZ.00.TAN1            A-----TT-T---C----T-GT-----C--TG-T--T---A-------------A-----------G--G--AGT---------------A-----G-----C----CA----C-CT----...-----A--CCAGCA-G----GAT-C-A-----G--C-CA--GGCA- 1124
CPZ.CM.05.SIVcpzLB7       ------T-----------T-----------T-----A--------C--G-----------G--C---------GT--T------------ACA------------GC-------TA-C---...ATG-----A--A-----------A---------------------- 1038
CPZ.US.85.CPZUS           G--T--C------G-C-----T-----C--TG-TA-A-----------G-G---A------G--T----G--AG----------------A----T--A-------ACA-----TC-----...-----C----A---------G--------------C---------- 1540
CPZ.GA.88.GAB1            -TTA--T--G--G-T----------T----TG----A-----------G------------G--C----G--AGT---T-----------C--C--G-----C-CT-A------CACT---...--T--T--G--A--C-----A---C-A--G--G--------T---- 1553
MON.CM.99.L1              G-TT--CT-T--CA-C--TCAGC-A------G-CA-A--T-AT-----G-GG--AC-C-----TA-T--G--TGT---------------G--CA-T-----CTT------C---CA--CG...C--CAA--GCAAC-CAAT-CAGG-C--G----C----C---GGCC- 1156
MON.NG.x.NG1              ........................................................................................................................................................................... 0
GSN.CM.99.CN166           T-TA--CT-T---A-G---CAA-----C--TG-TA-A--A-ATT------GG--AG-A--G--TA-T-----TGT---------AC-G--A-----G-----CCTTT-A-----G--A-CA...----CA---CAAC-GGTG-CAGG-C-T-----C--T--C---GCA- 1169
GSN.CM.99.CN71            T-TA--TT-T---A-G---CAA-----C---G-T--T--A-ACT-------G--AG-A--G--TA-T-----TGT---------AC-G--A-----G------CTCT-A-----GCA--CC...--G-CA--CCAAC--GT--CAGG-C-T-----C--T--C---GCA- 1175
MUS_1.CM.01.1085          A-TA---T-T--CA-C--TCAAC----G--TG-CA-A--A-A------G-G--T--A------TA-------TGTA--T-----AC----A-----------CCTGT--A-C---CA--CC...CA-CA---GCAGC-CAAT-CAGGCC-C--GA-C----CA---TCA- 1140
MUS_1.CM.01.CM1239        A-TA--CT-T--CA-G--TCAGT-------TG-CA-A---AT------G-GG--T--A--G--CA-------TGT--A------C-----A-----------CTGC-A---C---CA--C-...CAGCA---GCAGC-CAAT-CAGGCC--C-GAGC---TCA---TCA- 1142
MUS_2.CM.01.CM1246        --T---CT-T---G----TCAG-----C---G-CA-A--A-AC--C--G-GG--A--A---C-AA-------TGTGG--------C-G--A-----CA-A--CGTTT-A-----CCAA-CA...A--CA---ACAAC--AAT-CAGG-C-A--GT-C---TCA--GGC-- 1223
MUS_2.CM.01.CM2500        T-TT--TT-T---G-G--TCAA----T---TG------A-AT---C----GT--AG----GC-AA-T--G--TGTGG--------C-G----T---G-----CTTT-----C--CCAA-CA...---CA---CCAAC--AAT-CAGG-C-TC--T-T--CTCA---GCA- 1185
RCM.NG.x.NG411            --TA--CT-T---G----TCAG---T----TG-CA-A---A---------GT---T-A-----AG----G--AGTA--------T--G--A-----T-----CTTGAGA-----CAAT-CA......--A--AC-AC--G----T--AT-A--G-------CA-----C- 917
RCM.GA.x.GAB1             --TA--CT-T---A----TCAA--------TG-CA-A---A----C--G-GT--GC----G--AG----G--AGTG--------------A-----C------GCTAGA--------A-CA......--C--G--AC--G----G--AC-T----------CA------- 917
DEN.CD.x.CD1              G-TA--TT-T---A----TCAAC-C-----TG---A-T--A----C----G----------G----------TGT---T-----A-----CAT--T------TTG-CT------G-CTA--TTACA-CA---GC--C------CAGCACAA--G--C-----A--TTCA- 1262

MAC.US.x.239              ------CT-T--CA-T--TCAG---T----TTGT----A-AC--------G---T-----G--TA-C--G--T-TT--A--C-----G--------T-----CTTGCA---C---CAA-CA......-CT--A...CA--A---A--AC-T--G--G--GTCA---TCA- 1755
Gag                       __T__P__Y__D__I__N__Q__M__L__N__C__V__G__D__H__Q__A__A__M__Q__I__I__R__D__I__I__N__E__E__A__A__D__W__D__L__Q__H__P__Q__P__.__.__.__A__P__.__Q__Q__G__Q__L__R__E__P__S__G__S__
H2A.GW.x.ALI              ---G--CT-T---A-T--TCAA----T--TTGT-----C-AC------------------AA-C-GG---TT-------A----A----C------GTT-CA----CA-A-CA......--C--CT-AC--G----G---C-T-----------A--GTC-- 1801
H2A.DE.x.BEN              ---G--CT-T---A-T--TCAA-----T--TTGT-----C-AT--C-------T-------AA-C-G----TT--T----A----A----C------TCGCA---C---A-A-CA......--C--CT-AC--G----A---C-C-----C-----A--GTC-- 1808
H2A.SN.x.ST               ---G--CT-T---A-T--TCAA-----T--TTGT-----C-AC---------T-------AA-C-GG-A-TT--T----A----A----T----CGC-CAA--C---A-A-CA......--C--CT-AC--G-G--G---C-C--G--G---------TC-- 1253
H2B.GH.86.D205            ------TT-T---A----TCAG--------TTGT--A--A-A-----G-------------TA-T-G---A-TA--------------A----C-----CCAGCAA--C--GTCA-CA......--C--A--GC-GG----A--AC-T--G--C-----A--GTCA- 1769
H2B.CI.x.EHO              ------TT-T---A-T--TCAG---T----TTGT--A--A-A-----G-------------TA-T-GG---TA--T-----------A----C-----CCA-CA-------TC--CA......--C--A--GC--G----G--AC-C--G--------A--GTCA- 1770
H2G.CI.92.ABT96           -CTA---T-T---A-C--TCAA--------TTGT--A--A-A------------------TA-T-G---A-TA--T-------G--A----K-----CTTCAA------CAA--A......-----GCCGC--G----A--AM-C--G-----G-CA---TCA- 1171
H2U.FR.96.12034           ---G--CT-T---A-T--TCAA---T----TTGC-----A-A--------T--T-------TA-C-GG-A-TA--T-----T----A----T-----C-C-CA---C---AAC--A......--C--ACCAC--G----G--AC-T-----G-----A--TTC-- 1287
MAC.US.x.EMBL_3           ------CT-T--CA-T--TCAG---T----TTGT--A--A-AC-------G--T-----G--TA-C-G---T-TT--A-------G-------T-----CTTGCA---C---CAA-CA......-CT--A...CA--A---A---C-T--G--G--GTCA---TCA- 1241
SMM.US.x.H9               ---T--CT-Y---A-C--TCAA--------TTGT--A--A-A-----G-------------TA-T-G---A-TT--A-----A--G-----T--T-----TT-CAA--C--RCA--CA......--T--A--AC--G----A--AC-T-----G-----R---TCC- 1240
SMM.US.x.PGM53            ---T--CT-T--CA-C--TCAA--------TTGT--A--A-A-----G-------------TA-T-G---A-TT--A-----A--------T--T-----TTGCAA--C--GCA--CA......--T--A--AC--G----A--AC-C-----T-----A---TCA- 1684
SMM.US.x.SIVsmH635F_L3    ---T--CT-T--CA-C--TCAA--------TTGT--A--A-A-----G-------------TA-T-G-----TT--A-----A--------C--T-----TT-CAA--C--GCA--CA......--T--AC-AC--G----G--AC-T-----G-----A---TCA- 1770
SMM.SL.92.SL92B           T--A--CT-T--CA-T--TCAG----T--TTGT--A--T-A-----G-------------TA-C-G---A-TA--T-------------A----C-----CCT-CAA--C---AGAGGACAACAGCC-G-GCAAC-GG-T--AGG-C-A--G------TCA---TCA- 1198
STM.US.x.STM              T-----CT-T--CA-C--TCAAC--T----TTGT-------AG-------------------TA---------T---A--------------T--T-----C-TGCA-------CAA-CA...C-G--T--GC-GC--G----G--AC-T-----------C--GTCA- 1417
MNE.US.x.MNE027           --T---CT-T--CA-T--TCAG---T----TTGT-----A-A-----------T-----G--TA-C-G---A-TT--A--C-----G-------T-----CCTGCA---C---CAA--A......-CT--A...CA--A---A---C-T--G--G--GTCA---TCA- 1235
DRL.x.x.FAO               --T-AGCT-T--CA-C---CAG---T----TG-TA-T----AT--C----GT---C-C-----AG-------C-T--------------A--T--T---------GAGA-----TCAAGTA......--A---C-CC---A---GGT-C-T---A-C--CTCA--GTCA- 999
MND_2.x.x.5440            --T---GT-T--CC-C--TGG------C--TG-CA-T--A-AT-----G-G---TC-C-----AG-G--G--TGT-------------------C------CTTAGA--C--CCCTGTG......--A-----GC-C-A---AGT-C-A---A-C----CA---A-- 1002
MND_2.CM.98.CM16          -CT---CT-T---C-C--TGG------C--TG-TA-T--A-AC-----G-G---GC-T-----AG-G--G--TGT-------------------C---A-CTTGAGA--C---CAAGTG......--T-----GC-C-A---TGTCT-----A-C----C---C---- 1445
MND_2.GA.x.M14            -CT----T-T--CC-C--TGGT---T----TG-TA-T----A------G-G---TC-C-----AG-G--G--TGT-------------------C------CTTAGA--C--CCAAGTA......--A---T-GC-T-A---GGT-C-----A-C----CA------- 1369
OLC.CI.97.97CI12          GCTG---G----GA-C--T---C--T-G--TGT---T----ACT------GG--T------G----G--G--AG--------------A-A---G-AT--C--GCA-------CA--CA...AT------AT-GC-GGTG--A--A--TCC-CCG-------CT---- 1360
GRV.ET.x.GRI_677          TCT-T-CT-T---G----TCAG-----C--TGT-A-A--A-AC-----G-GG--AT-A-----TC-T--G--AGT---T-----A-----A-----G-----C--GACA--CAG-CCA-CA..--T--C--GT-AC--G----G---C-A-----C--G-CA--GTCA- 1605
VER.KE.x.TYO1_patent      ---A--CT-T--CA-T--TCAG-----T--TGTGC-A--A-AT------GG--AT-A-----AG-G---------T---T-----A-----A--CC-G------GT-ACA--C---C-A-CC..-----A--CC-AC--G----A---C-C--G--C--TC-C--CTCA- 1665
VER.KE.x.9063             T-----CT-T--CA-C--TCAA-----T--TGTGT-A--A-AC------GG--AC-A-----AG--------A-T-----------------A--CC-G------GT-ACA--C---CCA-CA..-----A---T-AC-------A---C-T-----C--T-----TTCA- 1668
VER.KE.x.AGM155           -------T-T--CA-C--TCAG-----T--TGTCT-A--A-AT-----G-GG---T-------AG-G-----A-TA--T-------------A--CC-G------GT-ACC--C---CC--CG..-----C--CT-GC--G-G--A---C-C--G--T--GG-----TCA- 1657
VER.DE.x.AGM3             ---A--CT-T--CA-C--TCAA-----T--TGTCC----A-AC-------GG--GC-A-----AG--------A-T-----------------A--CC-G--------T--CT--C---CCA-CA..-----A--AT-AC--G----A--AC-C-----C--T--A--CTC-- 1169
TAN.UG.x.TAN1             -TTAAGTT-T--CA-T--TCAG---T----TGTCA-A--A-AC-------GG--A-----G--TA-C--G--AGTA--------T-----A--CC-G--------T-ACA--C--CCCA-CA..-----A--GC-AC--G----G---C-A--G--C-----A--TTCA- 1628
SAB.SN.x.SAB1C            T-T---TT-T---A-C--TCAA-----C--TG----A--A-A---C--G-GG--AC-A-----AG-G-----TGTT--------A-----G-----T-----CCTTAGA--C--GCC--CG...CA-CA---CCCA--T-AG--AGT-C-A--G--C---CAA--GTCA- 1758
MND_1.GA.x.MNDGB1         A-T-AGT--T-----G--T-T------G--TG----A--A-AT-----G-G---A------G-C--------TGTA--T------C----A-------------CT-ACA-----TCAA--A...CA-CCAG-ACAAC--GG---AGGAT-A--GACC--TTCA--CTC-- 1119
DEB.CM.99.CM40            T-----CT-T--CA-T--TGGTC-A--T--TG--A-T--A-AT-T---G-GG--T------G-CA-T-----TGTA--T------------A------G------TT-CAA--C---CAA--A...CAGCC-......C---AG--G---C-T--G------TCA---GCA- 1256
DEB.CM.99.CM5             T--T--CT-T---A-T--TGGGC-------TG-CA-A--A-AT-TA---GG--A-----------TA------TGTG--T------------A------G-----CTT-CAA--C---CA---A...CAGCCA......C---A---T---C-T-----G---TCA---GC-- 1247
TAL.CM.01.8023            --TT--TT-T---A-C--TCAG-----T--TG----A--A-A------G-GG----------G--CA-T-----T-T---A-------C-G-----C---------CAGA----C--CCA-......AT---T--AT-AC--G----A---C----G--G---GCA--C---- 875
TAL.CM.00.266             --T---CT-T--CA-C--TCAA-----T--TG----A--A-AG-----G-GG--A-----G--CA----G--C-T---T-------C-G--AT----G-------CAGA--------CCA-......AT---T--AC-AC--G----A---T-A--G------GCA--G---- 1377
SUN.GA.98.L14             A-TT---G----CA-C--TGTG-----G--TG-CA-A----AG-------G---T-----G---A-------TCAT--AGTG-------AG-------------C---A-----C--TCA---G...CAGCCTG-ACAGC--GG---AGGCC-A--GACT--T-AT--T-CA- 1726
WRC.CI.97.97CI14          ATTAT-T--T---A----T---C--T------ACA-A--A-AT-------T--A---------AA-------C-TAG-A-----A---ATACA---G---------ACA-----CA-GCA..------------A--------A--A------C-G----CA---TCA- 1696
WRC.CI.98.98CI04          TTTGG-T--T--CA-T------C--T----T-ATA-A--T-AT----G--T---------TA-------A-TAG-A---------AT-CA---C---------ACT--C---CAA-CA..---------A--A--------A--A-----GC-G----CA--GTCA- 1714
WRC.GM.05.Pbt_05GM_X02    ATTGA----T--CA-C--T--TC-AT-G--T-ATA-A--A-AT-----------------G--AG-T----A-TAG-G-----A---ATAA-----------C---ACA-----CCA-CA...-T---A---C-GC-C-----GGCT------A-T----CA---TCA- 658
LST.CD.88.524             -TTG---G----C-----TGT----T----TG-TA-A--A-AC------GG--A-----G--AA------TCAT--AGCG--A----GA--T------------CA-----CCA---G...CA-CCAG-ACAAC--GG---AGG-T-A---ACT---AT-C--CA- 658
LST.CD.88.447             TTTG---G-T--C-----TGT----T----TG-CA-A--A-AC------GG--A-----G--AA-T----TCAT--TGTG--A----GG-------------C--GAAA---GCAA--A...CA-CCTG-ACAAC--GG---AGG-C-T---ACC--T-AT-C--CA- 664
LST.KE.x.lho7             TTTG--TG-T--------TGTT---T----TG--A-T--A-AT-----G-GC-----------AA-------TCA---AGTG-------G------G---------CA---C--GCAA--A...CAGCCAG--CAAC--GG---GGGTT-A---ACT--T-AT-C--CA- 1748
SYK.KE.x.SYK173           T-TT--TT-T---C-C--TGG------T--TG-T--A--A-AT-----G-G---------GG-AA-T-----C-T-------------GG--G-C-------CTT-AA-----CCAA-CG...CA-CAA---GC-C---A--CAGGAT----G--T--CTCC-CCTCA- 1572
SYK.KE.x.KE51             T-TA--TT-T--CC-C--TGA------T--TG-CA-T--A-AT-------G---A------G--A-C-----TGTG--T-----------GGT-T--------CCT-CAA--C---CAA-C-...CA-CA--AGCAGC-TGT--CAGG-C-A-----T--T--T-CCTCA- 1200
COL.CM.x.CGU1             A-TAG-TT-T---A-G---TTG-----C--T-TTT---ATACT-------G-TT-CCT--GG-C--G--G--TGAG--T---A-AA-G--A-A---G-AT--CCT-C-A-----G--C......CAGCA---ACAACAG-A---AGCAT-A--GC-G--T-CA-C-TCA- 1297
```

```
H1B.FR.83.HXB2           ACATAGCAGGAACTACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAAT..............AATCCACCTATCCCAGTAGGAGAAATTTATAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGAATGTATAGCCCTACCAGCATTCTGGACATAAGA 1647
Gag                      D  I  A  G  T  T  S  T  L  Q  E  Q  I  G  W  M  T  N  .  .  .  .  .  N  P  P  I  P  V  G  E  I  Y  K  R  W  I  I  L  G  L  N  K  I  V  R  M  Y  S  P  T  S  I  L  D  I  R
H1A1.UG.85.U455          ------------------C---G---A-------------------GGC..............-----------------G----C--C---G--------------------------------------------GTT-----T---------- 1102
H1B.US.90.WEAU160        --------------------------------------C-------------..............--------C-----------------C-------------------A-----------------------------GT------------------ 1646
H1C.ET.86.ETH2220        ------------A-------------------------C---------GGG..............--C-----G-T----G----C--C--------------------GC-----------------------------GT-------T--------A- 1021
H1D.CD.84.84ZR085        -T--T----------------------------------------------GC..............------------------------C---------------------------------------------------GT------T------C--- 1164
H1F1.BE.93.VI850         -T-----T---------------------------CA--------GGGC..............--C-----G------G----C--C------------C----A-------------------------T---GT------T--------A- 990
H1G.SE.93.SE6165         -T-----------------T----G-----------ACT-------CGGC..............--C--------------G----------------------------------G----------------------------GT-------T-A--------- 1050
H1H.CF.90.056            -T-----------------------G------------C---------GGC..............------G---------G----C--C------------------------------------G----------------T---GT-----------------A- 994
H1J.SE.93.SE7887         -T-----------------------C----------------------GGC..............----------------------G-----------------T------C-----------------------------GT---T---T----T------ 970
H1K.CM.96.MP535          ---------------C--C-----------------C----------GC..............--C-----G------G--G----C------------------T-----C-----------------------------GT------T--------- 846
H101_AE.TH.90.CM240      --------------------------A-----------------------C..............--------C------G----C--C------G----------------------------------------------GTT-----T---------- 1212
H102_AG.NG.x.IBNG        --------------------------A----------------------GC..............----------------G-------C-------------G--------------------G------------GT-------T---------- 1178
H1N.CM.95.YBF30          -T-----------A-----C-----GGCA-----GG-G-CT-------TGC-..............-----T--G-T----------T------G----------G------G-----C-G---T--G----------T---GT----------A--G--C-A- 1211
H1O.BE.87.ANT70          ----T--T--G--A-----C----AG--A--G-----TCAC---CT--C-GGCCC............--C-A-----------------C--C----G--A------G-GT-A---C----C-----G----A------C-----AGTG-----CT-A--T--T-AG 1696
H1O.CM.91.MVP5180        ----T--T-----A-----C--A-AG--A--G-----TAT-----CT--T-GAGGG............GC-AACT----------------C--C----G--A------G-G--A---C----C-----G----A------C--T--AGTG-----CT-A--T--T--G 1671
CPZ.CD.90.ANT            -T--------G--A--A--C--AG-G-----G--G--GCA-------T---CACCT.....CAACAG---GG-GGAG---------G--C--C-----G-------C---A-------------GG-G--C--G-N------T--AGT----------A--G----A- 1068
CPZ.CM.01.SIVcpzCAM13    -T--------G-----C--C--AT-G-----G-----T------GCA---GCA..............--C-----C--------------C--C----G------G----T--A--GC----C-G-G------A-------T-T--AGTA-----C--T--T----AG 1224
CPZ.CM.05.SIVcpzEK505    -T-----------A-----------GGC---G---G---C---------TGCC..............-----T--G-T----------T------G----------G--T--------------G--T----------------T---G-----------A-----C-A- 1195
CPZ.CM.05.SIVcpzMB66     -------------C--C--C---A-C--A--G--GG----G--------T-GC..............--C--T----T----------T--C--------------------TT----GC-------G-------A-------T----GT------A--A-------A- 1190
CPZ.CM.98.CAM3           -------T--G--------C--T--AGTA-----G--T-CT--------GCA..............--------CG-T-------------A--C-GG-----G-GG-T--A--GC-T--C---G-------G------T----AGTT--T--CT-A-----C-AG 1037
CPZ.GA.88.GAB2           ----T-----T-----------T-G--A--G---G--------------TCC..............--------AG----T-------G----C-G------G-GG-AT-A--GC-T-----GG-G-------------T-----GTT-----CT---------A- 1025
CPZ.CM.98.CAM5           -------T-----------C--T--A--A------G-T-CT---------GCA..............--------C----------G----C----G------G-GG-T--A--GC-C--C---G----G-A-----CT----CGT---T-----A--T--C-A- 1340
CPZ.CM.05.SIVcpzMT145    -----------C--------CT----AG----G---G----G--------GCA..............--C--T----T----------T------G----------G-------G--------------T-A-----C--T---GTT------T-------T-A- 1186
CPZ.TZ.00.TAN1           -T--T-----G--C--------G--A--A------G--CTG---------CCCCACAGGCA......C-AGG-GGAG-G----------C--C--------G-------TT-A-----------T----C-------C------GTT------T---------A- 1288
CPZ.CM.05.SIVcpzLB7      -------------C-------TG-A--A-------------------T-GC..............--C--T----A----------------C--------------TA---GC-G----GG----C-------C------GTT-----C---------AG 1193
CPZ.US.85.CPZUS          ----T--------C--C--C---T-G--A------G-G--C---------GCA..............--C----C--T----------C--A---G------G-GG---A--GC-T-----G-----A------CT-T--CGTTG-T--CT-A--T--C-AG 1695
CPZ.GA.88.GAB1           -T--------G--C--C--C--A--G-----------T--G----CA---GCA..............-----T--C----------G--TG-----G------G----TT-A--G-----C---G-G--T---------T----AGTA--T--C--T-----C--- 1708
MON.CM.99.L1             -T---------GTCT-GTCC--A-CC--AC-G--G----AG----A--C-GGGCT............--CAAC--CG-T-AG--CTC---C--C--C-G--A---G-C---T----CC-GC-GCG-TGC--C-A------C-A----GT--A----T----T--C-A- 1314
MON.NG.x.NG1             .............................................................................................................................................................. 0
GSN.CM.99.CN166          -----------GT----TCA--T-CCA-T--G--G----AG-----C----GGCAG............--C-A----G--AAT--GTC---T--A----G--AG-----C--TT----CC-GC-GCG-TGT----A--------A----GT--AT---T-------C-A- 1327
GSN.CM.99.CN71           -----------GTA--ATCA--T-C-A-T--G--G----AG-----T----GGCAG............--CAAT--AG-TAAT--G-C---T--A----G--A------C---T----C--GC-GCG-TGT----A--------A----GT--AT---T---G---C-A- 1333
MUS_1.CM.01.1085         -----------GTAT-ATC-----CA--A-----G----A------C--T-GGGCA............--CAAC-----AAAT---C------C--C--G-AG---G-G---A----GC-GC-G-G-TGT--C-A------C-AT--AGT--A---C--A--T--C-AG 1298
MUS_1.CM.01.CM1239       ----------GGTA--ATCC----CCA-T----------A------A--T-GGCAA............--CGAT-----AAAT--G-C--C--C-----G-A----------A---G---C-G-GGTGT--C-A--------A----GTT--T--AT----T--C-AG 1300
MUS_2.CM.01.CM1246       -TG-G--G---GTGT-ATCA--G-CAGCA----------A---------T-GACAG............C-AGAC--AG-AAAT-----GA-T--A----G--A------C---T----CC--C-G-G-TGT--C-A------C-AT--AGTA-A---CT----T--T-A- 1381
MUS_2.CM.01.CM2500       -T--------GGTAT-CTCC--A-CAGCA---------AG--------C-GGCAA............C-AGAC--AG-GAAT-------C--A----GG-A------T-------C--GC-A-G-TGT--C-A--------A----GTA-A---CT-A-----C-A- 1343
RCM.NG.x.NG411           -T--T--------A--A--C-ATA--GCT----------CT----CC--T-GAGCC............--CAAC--A--TG----G--A-C-------GG-AT----TG-G--T--G---C-A--T-G---C-A---TGT--G-A--------A---AGTA-A---CT-A-----C-AG 1075
RCM.GA.x.GAB1            ----------G--A-----CT-AA-AGCA-----G----CT----CC--C-GAGCA............--CAAC--C--TAAT--G--CA-TC-G--C-G--AT--------AG-A--G---C-A---TGG---A------C-AT--AGTG-A---C--A--T----AG 1075
DEN.CD.x.CD1             -T--------G--C--CTCCT-TA-AG-A--T--G---AC---------T-GACAA............C-AGGC--A--AAAT--------T--A-----G--G---G-T--TA--------C-A--GG-G--C-AG-----CCAA--AGT-TC----T---------AG 1420

MAC.US.x.239             -T--T--------A------T-AG-AG-T--------CCAG------TAC-GACAA............C-GAAC--C--A--------CA-C-----C-GG--------CCAA-----G--GC-A---TGT--C----------A---A--A-A------A--TG---A- 1913
Gag                      D  I  A  G  T  T  S  S  V  D  E  Q  I  Q  W  M  Y  R  Q  .  .  .  .  Q  N  P  I  P  V  G  N  I  Y  R  R  W  I  Q  L  G  L  Q  K  C  V  R  M  Y  N  P  T  N  I  L  D  V  K
H2A.GW.x.ALI             -------------A--A--C--AG-AG-A----G--CCAG------TTC-GGCCA............CGGAAT--G-G--------GA-C--C----G---------CCAGA-A--GC--C-G--GTGT--C--G-----C-A---A---A---C--A-------A- 1959
H2A.DE.x.BEN             -------------A--A--C--AG-AG-T----G--CCAG------TAT-GGCCA............C-AAAT--CG-A--G-----GA-C--C--C-G---------CCAAA-A--GC-GC-A--GTGT--C----A---C-A---A--T-A---CT-A-------A- 1966
H2A.SN.x.ST              ----------G--A--A--C--AG-AG-A--G--G--CCAG------TTT-GGCCA............C-AAAT---G-A--------AGC--C----G---------CCAGA-A--GC--C-G--GTGC--C--G-----C-A---A---A---C--A-------A- 1411
H2B.GH.86.D205           -T-----------C--C--C--AG-AG--------G---CAG------TAC-GGGCC............C-AAAT---G-------G---A-C-------G---------TCAAT-A-----GC-G---TGT--CC-------C-AT-------A---AT-A-------AG 1927
H2B.CI.x.EHO             --------G-----C--C--C--AG-AG-A-----G---CAG------TAC-GACCC............C-AAAT---G-------G---A-C--C----G---G-----TCAGT-A--GC--C-G---TGT--CC-G-----C-AT-----T-AT--A----------AG 1928
H2G.CI.92.ABT96          -T--------G--C--------AA-AG-A--------TCAG---CACAC-GGCAG............C-AAAC-----A--------A-T--Y----G---------CCART-A---C-TC-A---TGT--C--G-----Y-AT--A--A-A---C--A--TG---A- 1329
H2U.FR.96.12034          -T--T--------A-----C--AG-GG-A--G--G---CAG------TAC-GGCCG............C-AAAT--A--A--G--G--TA-C--C---CGG--------CCAAT----CC--C-A---TGT--T---------C-AT--A--T-AT--CT-A--T----AG 1445
MAC.US.x.EMBL_3          -T--T--------A------T-AG-AG-T--------CCAG------TAC-GACAA............C-GAAC--C--A--------CA-C-----C-GG--------CCAA---A-G--GC-A---TGT--C-----------A---A-TA-A------A--TG---A- 1399
SMM.US.x.H9              ----T---------------C--AS-AG-C--------TCA-------TAY-GGCAG............C-AAAC-----A--------CA-C--C--C-GR--G----CCAAT-A--GC-GC-G---TGC-------------AT--A--A-A----T----TG-G-A- 1398
SMM.US.x.PGM53           -T--T-----------------AG-GG-A--G-----TCAG------TAT-GGCAG............C-GAAC--C--A-----G---A-C-----C-GG--------CCAA--T--G---C-A--GTGT--C---------C-AT--A--A-A----------G---A- 1842
SMM.US.x.SIVsmH635F_L3   ----T-----------------AG-AG-----------CCA-------TAC-GGCAA............C-AAAC--C--A--------CA-C-------GG--G----CCAAT-A--GC-GC-G---TGT-------------A---A--A-A----T-A--TG-G-A- 1928
SMM.SL.92.SL92B          ----------G--A-----C---CAAGT----G----A-------TAT-GGGCA............C-AAAC--AG-------G----C--C----G---G----CCAG-----C-CC-A---TGC--C--G-------A---A--A-A---C-----TG-T-A- 1356
STM.US.x.STM             -T--------G--------CT-A-CAG-----------TCA-------TACCGGCAG............C-AAAT-----A--G-----A-C--C--C-GG--------TCAA--A--GC--C-A---TGT--C---------C-AT--AGTT-A----T-A---G---AG 1575
MNE.US.x.MNE027          ----T--------A--------AG-AG-T--------CCAG------TAC-GACAA............C-GAAC--C--A--------CA-C-------GG--------CCAA-----G--GC-A---TGT--C-A--------A---A--A-A------A--T------ 1393
DRL.x.x.FAO              ----T--------C--C--CT-TA-AG-A--G-------AG----CC--C-GAGCC............C-GGATT-C---AAT--G----C---------G-CA---G-GG--T---G---C-A-G-TGT----C------C-AT--AGT--A---CT----TG-G-AG 1157
MND_2.x.x.5440           ----T--------A--A--CT--A-AG-A--G--G----A-----CA--C-GGCAG............C-AGAC-AGG-AAAC-------GT--A--C---CA-------G-A--A--GC-GC-A---TGT-----C-------A----GTA-AT---T-A-----C-A- 1160
MND_2.CM.98.CM16         -T--T--T-----A--A--CT--A-AG-A--G--G----A-----CT----GGGAG............C-AGATG-AG-AAAT--G----G--A-----CA-------G-TT-A--GC-TC-G--GTGT-----C-------AT--AGTA-A---C----------AG 1603
MND_2.GA.x.M14           ----T--------A--A--C-G-A-AG-A-----G----A-----CA--T-GACAG............C-GGA--AGG-AAAT----G-C---C-----GCAG-----G----A--GC-GC-A---TGT--G--C-----C-A---AGTG-A-------------A- 1527
OLC.CI.97.97CI12         -T----T------C-------ATT-G---C-G---C-G-CT---G-A-ATGCAGGA.....................--AGAT------C-C-C-----G-----G-T---A-A--G--GC-G--GT-G--G-AG-----TCG--AGTT-CAG-G--A-----C--- 1509
GRV.ET.x.GRI_677         -T-------------C--CT-AA---------------AG---CCTTC---GCC................-----AGA--AGAC----G-C-CAA--C-G--A---G-T--TT----C---C-A--GG----GCAG-----C-AT--CCAA-AGG-C--A-----TC-- 1763
VER.KE.x.TYO1_patent     -T--------G--C--C--CT-AG-A--A-----GT---A------CTAT-C-GCT............--C--C-GGG-AGAT-----T-CC--C--CCGG--------T--T--A---C-TC-A--GTGT--C-A------C-A---AGTATCAG-C--A-----T--G 1823
VER.KE.x.9063            -------------C--------AG-G--A-----T---A-------TTAC-CAGCC............--C--CAGAG-AGAT--T--G-CT--C----G---------T-----A--GC-GC-G--GTGT----A------C-A---AGT-TCAG-AT-A-----C--- 1826
VER.KE.x.AGM155          -T--------G--C--------AG-G--A--G--GC---AG-----CTAT-C-GCT............--C---AGGG-AGAT-----G-CC--C---CG---------C-----A--G---C-A---TGT----A------C-AT--AGTGTCTG--T-A--T--C--- 1815
VER.DE.x.AGM3            -------------C--C--C--AG-G--A-----GC-G-A------ATAC-CAGCC............------CAGAG-AGAT--G--T-CC--C----G---G----T--------G--GC-A---TGT----A------C-A---AGTGTCTG-CT-A---------- 1327
TAN.UG.x.TAN1            --------G-----A--C--CT-TG--GCT--G-------AG----CTTTT---GCT............------AGAG-AGAT------CGT--C--C-G-G-----G-T---------C-CC-A---TGT----A------C-AT----T-TCAG-C---------C--- 1786
SAB.SN.x.SAB1C           --------G--C--------C--AA-A--A--G-------A-----CC--C-GAGCA............C-GAATG--G--AAT------A-T----------G-------C--------TC-CC-G---TGT--C-A------C-A---AGTA-A----T-A-----T-A- 1916
MND_1.GA.x.MNDGB1        -T---------------TC---AG-GG-A------T-G-C---------AT-TGCAA............C-AAATG-A---AAT------AC---C-----G--T-----T--A-----CA-G----G-T-G---A--GTC--T-T--A-TA--T--AACA--TG----- 1277
DEB.CM.99.CM40           -------------A-AC--C--AG-GG-A-----G----C----------C-GACCAGCAGGTCAAGGAC-AGG---C--AGAT------C---A---G-------G-C--A--A--TC-GC-A-G-TGT--G-A------C-AT--A----A---CT----TG-T-A- 1426
DEB.CM.99.CM5            -------------A-AC--C--AG-GG-A--G--G----CT---------C-GGCCAGCAGGACAAGGAGCAGGC--A--AGAT------C-G--A--C-GG------G-C-----A--C---C-G-G-TGT----A------C-AT--A--T-A----T----TG-T-AG 1417
TAL.CM.01.8023           -------------A--C--C----CAGCA------G-G-AG----CC--C-GGCAG............C-GAAT--AG-AGAT--G--GTCC--A--C--G--G--G-C---A-A--C---C-GCGCTGT----A------C-AT--AGTG-A---C-----TG-G-A- 1033
TAL.CM.00.266            ----------G--A--A--C--T-CCGCT-----GG-G-A-----CC--C-GAGCA............C-GAAC--CG-AGAT--G----C---A--C---------G-C---A----C---C-GCGCTGT----A------C-AT--AGTT-A---A-----TG-G-A- 1535
SUN.GA.98.L14            ----------GGTA--ATCA--AG-AG-A--G--G---CAG-----CTATGGCCTT......AATGGAGCAA--AGAG-TG-----CA---C--C------G-C---G-C--AGA---CA-GG-A-G-G----T---C-CC--CAGA-----TCAG-AG-A--A--T--- 1890
WRC.CI.97.97CI14         -T----------GTA--AGCA-GTG-GG----TG--C---CC---T--G-TC-AAAT.................-----AG-AGAT-----GTC---A------------G-CC-TAGA---C--C-G---G----TCAG--C---GTT--AGTA---------------A- 1851
WRC.CI.98.98CI04         -T---------GGTA--AGC-T-AG-AG----TG--T-G-CC---T--G-TC-AAAT.................-----A--AGAT-------C-G-G--C-----G---G-GC-TAG---GC--C-A--GG-----CA---T---GTG--AGTT------T-A--T----AG 1869
WRC.GM.05.Pbt_05GM_X02   -T----------GTA--ATC---TA-GG-----G--T---C----C-AG-TC-GCAA.................-----A--AGATA------C---A----------------TAGA---C--C-G--GG----GCA---T---GTA--A--ATCT--CT-A--T--C-A-  813
LST.CD.88.524            -TG-G-----G-TC--ATCA--AG-AG-A-----GCA--CC----CA---GCAGAC...............A----C--AGAT----GA----C-----GA----G------A--CTA-GG-G---G-G--------AC--CAG--AGT---TG-AA---------A-  813
LST.CD.88.447            -TG-C------GTC--CTCA--AG-AG---G--GC---CT----CC--TGCAGAC...............A-C---G-AGAT------A-G---------GAG---G----T-AA-CTA-GG-A---G----T-----CC--CAA--AGTT--TG-AA-A-----C-AG  819
LST.KE.x.lho7            -TG-G-----GGTC--CTCA--AG-AG-A------C---CC----CT---GCAGAC...............A-T---G-AGAT------A-----------GA----G-C--T-A--CTA-GG-A---G----C-----CC--CAA--AGTA--TG-CA-------C-A- 1903
SYK.KE.x.SYK173          -------------A--A--CT-AA--GCT--G-------A------T--T-GGCAG............--CAAC--AG-A-A-----------A----G---------C--T--A--GC-GC-G--GTGT--GCAGG-C--C-A----GTTTCT--CT-A-------A- 1730
SYK.KE.x.KE51            -------------A--C------A-C--A----------AG-----A--T-GACAG............--CAAC--A--T-AG--T---C-G--C--C-GGCAG--------T--A------C-A--GTGT--GCA-G-C--C-AT---GT--AT---T-------C--G 1358
COL.CM.x.CGU1            -T--CA----C-AC--A--CT-AG-AGCA------G-G-CT---GGAGAG.........................--C--A.......-CCA-C-------GGG-C------G-A-A-A-TC-GG-A--GG--A-CCAG--TGCA--A--AT--TCTG-A--A--T--T--- 1437
```

```
H1B.FR.83.HXB2          CAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAACTCTAAGAGCCGAGCAAGCTTCACAGGAGGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATGCGAACCCAGATTGTAAGACTATTTTAAAAGCATTGGGACCAGCGGCTACACT 1817
Gag                     _Q__G__P__K__E__P__F__R__D__Y__V__D__R__F__Y__K__T__L__R__A__E__Q__A__S__Q__E__V__K__N__W__M__T__E__T__L__L__V__Q__N__A__N__P__D__C__K__T__I__L__K__A__L__G__P__A__A__T__L
H1A1.UG.85.U455         -----G-----A--------C--G--T--------TA-A----T---------C-----T---------A-------T--------C------------------C-----------------T-----C------T-C-------G-----A--G---G------T- 1272
H1B.US.90.WEAU160       -----------A------------T---------------------------------------A-------------------------------------------T--G---------------------------------------------GA-------- 1816
H1C.ET.86.ETH2220       ----------------T----------------------T------CT------T--A------A---A--T-------------------C-----------------------------------A--C------G-----A--G---G---T--T- 1191
H1D.CD.84.84ZR085       -------------------------------------------C---------------------------A--A----------------------------------------A-----------------C-----------------G------T- 1334
H1F1.BE.93.VI850        -----G-----A----------------------A-A----T----GTC--------T----------------C-----GGG-----------C--A--------------------------------------C---------------A--G-------- 1160
H1G.SE.93.SE6165        -----G-----A----------------------TA------T----TG-T-G-----T-----------------T------GGC-----------C--------------------------------------C--C----G------A----A--GA---T---- 1220
H1H.CF.90.056           -----G-----A----------------------A----T-T-------T------T--------CA-------T--G--G-----------------------------------A--T--------C--------A----G-----A----A--G----T--A- 1164
H1J.SE.93.SE7887        -----------A-----T----------------A------T----G---------T---------A-------T--------------------T------C-------------A--T--------C-----C-----------------A---T---GA-------- 1140
H1K.CM.96.MP535         -----G-----A----------------------TA------T------C-------T--A-----CA-------A-----G--------------C---C--------------C--A----------------C-----------G-----------G---T--T- 1016
H101_AE.TH.90.CM240     -----G-----A--------C-------------TA-----------------C----G--A-----A---------------C-----------------C-A--------C-----T-----C------T-C--------------A---A---GA------T- 1382
H102_AG.NG.x.IBNG       -----G-----A---------------T--------TA------T-------T-G-----T--A-----A-G--------------C------------------C-------------T-----C------T-C-------G-----A---A---G-------T- 1348
H1N.CM.95.YBF30         -----------A--------C-------------A------C--------------A-----G--AA-------A-----G------------------AC-C--A--A--------A-----------ACAGC-CC-----------A--G----GA-----CT- 1381
H1O.BE.87.ANT70         --G---------------A---------------T---------C-----AT-------T---------A-T--A--A-------------------------TC--C-T--T--------C--T--------C--ACAG-----G---T----A--G----GA-----CT- 1866
H1O.CM.91.MVP5180       --G---------A-----A--C-----T------T-----T--C-----AT-------T---------A-T--A--A-----G-----------------------C-T--T--G---T-A-------------ACAA---C-G--------A-------AA-----TT- 1841
CPZ.CD.90.ANT           -----------A--G-----C-----T-----G--TA-A------------AA-T------A--A--G---------CCT--G---GCC------------------------AA-----------C--T---------C--ACAC--CC-G--G--T-------A---GA--CT-CT- 1238
CPZ.CM.01.SIVcpzCAM13   -----------A--G-------A---T-----------A-----T--------CA-G-----T----------T-----A--G----C-------------C---C-T-----G--------T---------------C---CAA---C----G--CC----T--G-G---CT-C-- 1394
CPZ.CM.05.SIVcpzEK505   --G--------A--GA-T--------T--C-----TA-----T--C-----------------A------G--CA-T-----A-----G---------------G--AC-C--A--A--------A--T--T---------A-A---CC-G-G-------A-------GA--CT-C-- 1365
CPZ.CM.05.SIVcpzMB66    -----G-----A--GT-A--C-----T-----G---A-A--T-----------T----G--A--A-------A---A--A-------------------------C--T--A-----G--------T--T---------------T-C--A-----G------A--G----GA-----CT- 1360
CPZ.CM.98.CAM3          --G-----T--A--G--T--C--G--T---------TA-A-----C-----C-----G--A--A-----AA----A--A-----------------T--C----C-C--A--A--G-----T--T-----C-----A-AC--C----G---CC---------GA-----CT- 1207
CPZ.GA.88.GAB2          -----G--G-----G--T------------------TA-A-----C---GT-T-------A--A--G---------A--T--G-----C-----------G---C-CC----A--G-----C-----------------CAA--AC-T-GG---C-A--------GA-----C-- 1195
CPZ.CM.98.CAM5          --G-----T--A--G--A--C--G--T---------TA----T---------T-------A--A--G--AA-------A------------------T--C----C-T--A--T---------T--T-----C--C--A-GC--C--G-G---C--A-------GA-----CT- 1510
CPZ.CM.05.SIVcpzMT145   -----------A--G--T--C--G---------------A-------C--G--CT----G--T--A-----A--------T--T------------------C---C-----A-A--G-----T---------C-----A-AC--C----G------A-------GA--C--T-- 1356
CPZ.TZ.00.TAN1          --G--------A-----A--C-----T----------A-A-----C-----AA-C-----A--A-----A--T--ACCA-------C-----------------T--AC----A---------A--------------------CA---C--------C-----G-A--GA--A---T- 1458
CPZ.CM.05.SIVcpzLB7     -----G-----A-----A--C-----------G---A-----T--C-----AT-------A--A-----CA-------A--------------------------C------C-T--A--G-----T---------C--A-G---CC-G-G---------T----GA--CT-CT- 1363
CPZ.US.85.CPZUS         --G-----T--A--G---------------------A-----------C-----AT-------A--A--------T-CA--A-------CC----------G---C-T-----A-----C--T--T----------A-----------G---C--------T-GA--C--C-- 1865
CPZ.GA.88.GAB1          -----G--------G-----------T-----G---A-A-----C------T----G--T--------------------A--G-----------------T--C--TC-C--A--G--G-----A----------------CAA---C-----------------T-G---C--C-- 1878
MON.CM.99.L1            --G-----T--A-----A--C-A---T---------A-A-----C---TGC--C-----A--A---A-AGAT----CA--C---------------GC-G--GC--C--------------C-----------C--ACTC--CC-----G--CA--...----GT----GCT- 1481
MON.NG.x.NG1            .....................................................................................................................------G--C--C---CTC--CC-C------A--...----GA--C-GCT-   47
GSN.CM.99.CN166         -----C--T--A--G-------AG--A---------A-A----T---GTGCT----G-----A-----AGAT----CC--G----------------C--T-GC--C-----T---------T-----------A--C---CTG--C--G-----CA--...----GA--C-ACT- 1494
GSN.CM.99.CN71          --G--C--T-------A---AG--A---------A-A----T----TGCT----G--T--A--G--AGAT----CA-----G-----------C--TTGC--C-----T---------T--T-----A--C---CTG--C--G--G--CA--...----GA--C-GCT- 1500
MUS_1.CM.01.1085        --G--------A-----A---AG--T-----G---A-A----T----TG------G--A-----GT-AGAC--A-C---C-------------CTC----C-----A--A--G-----T-----------G-----ACTC---C-G--GAGCA--...----GA--C--T-- 1465
MUS_1.CM.01.CM1239      --G--------A--G--A---AG--T-----G---A-------C---TGCT----G--G------T-AGAT--A-CA--T--G---------CTCC--T-----A--G--G-----T-----------G------CTC--CC-G---AGCA--...----GA--C--CT- 1467
MUS_2.CM.01.CM1246      -----C--T--A-----A--C-A---T-----G--T-------T----TGCT-------T--A--G---GAC--A-CA--C--------------C-GT--C--C-A--G--------C-----C--G--C--ACTCG-C--G-----CA--...----G---C---T- 1548
MUS_2.CM.01.CM2500      --G--C--C-GA----G--C-A---T-----G---A-------T----TG---------T--A-----CGAC----CT--C--------------C--T-T---------T--G-----C--T--T--G--C--ACTCG--C-G-----CA--...----GA--A---T- 1510
RCM.NG.x.NG411          --G--------A--G--T---AG-----------TA-------C--GG-CT-G--G--A--A--G--AGAC-CA-CA--G--G-----------C--T-AC--C--A-T--------T--T-----C--------AG-AC-T--G-GTC-----ATGAACC----C-- 1245
RCM.GA.x.GAB1           -----------A--GT-A--C-AG--T-----G--TA-A--T------G-CT-G-----A--A--G--AGAC-C--CA-----------------C--T-AC--C--A-A--------T-----T-------C-----A-TGG-AC-C--G-GTC-----ATGAACC--T-TT- 1245
DEN.CD.x.CD1            -----G--C--A---T------AG--T---------TA-A-----C---G----------A-----G--AGAT--ATCA--G--GCT---------CACTT-AC-A-----A--------A--T--------A--T--------C---CAG---C-CG-------A-----TAACC-C---T- 1590

MAC.US.x.239            -----G-----A--G--A---CAGAG----------A-------C----G-T-------A--A--GA-AGATGCA-CA-----G-----------TC----AC--C--A-T--------T-------------C---CTAG-GC-G--G-GGC----TGTGAATC-C--C-- 2083
Gag                     _Q__G__P__K__E__P__F__Q__S__Y__V__D__R__F__Y__K__S__L__R__A__E__Q__T__D__A__A__V__K__N__W__M__T__Q__T__L__L__I__Q__N__A__N__P__D__C__K__L__V__L__K__G__L__G__V__N__P__T__L
H2A.GW.x.ALI            -------------G--A--CCA-AG---------TA-A-----C----GCT----G--A--A---A-AGAT-CA-CA-----G-----------TC----AC--C----A--G----C--------C--C--ATTAG-GC-G----G---A--GATGAATC----CT- 2129
H2A.DE.x.BEN            --G--------A-----G--CCA-AG----------A-------C----GCT-G--G--A--A---A-AGAC-CA-CA-----------------CC----GC--C-AA-A--G-----C--------C--C---TTAG-AC------G-C----GATGAATC-C--C-- 2136
H2A.SN.x.ST             --G-----------G--A--CCAGAGT---------TA-A-----C--G-GCT-G--G--A--A---A-AGAT-CA-CA-----------------CC----AC--C-A--G--G-----C--------C------TTAG-AC------G-C-A--GAT-AATC----CT- 1581
H2B.GH.86.D205          --G-----------G-----CCA-AG----------TA-A-----C----GCT--C-G--A--A---A-AGAC-CA-CA--G-----------------C----AC--C--A-T--G-----T-----------C---TTAG-GC-T--G-GC------ATGAATC-C--CT- 2097
H2B.CI.x.EHO            -----G-----A--------CCA-AG----------TA-A-----C----GCT--C-G--A-----A-AGAC-CA-CA------------------C----AC--C--A-T--G-----T-----------C--ATTAG-GC-T--G-GC------ATGAACC-C---T- 2098
H2G.CI.92.ABT96         -----G-----A---------CAGTCT---------TA-A-----C--G-G---R-----A--A---A--GAT-CA-CA------------------C-G--AC-AC--A-T--------C---------------ATTAG-AC----G-GG------ATGAATC----TT- 1499
H2U.FR.96.12034         -----G--------G--A---CA-AGT---------TA-A-----C----G-T-G-----A--A---A-AGAT-CA-CA--G--------------TC-G--GC--C--A-T--G-----T--------C--C--ACTCG-G-------GT-----TATGAATC----CT- 1615
MAC.US.x.EMBL_3         ---A-G-----A--G--A---CAGAG----------A-------C----GCT-------A--A---A-AGATGCA-CA-----G-----------TC----AC--C--A-T--------T-------------C---CTAG-GC-G--G-GGC----TGTGAATC-C--C-- 1569
SMM.US.x.H9             -----------A--GT-A---CA-AG----------TA-A-----C--G-G---------A--A---A-AGAC-C--CAS-G--------------TC----AC--C--A-T--------C-----G--C-----ATTGG-GC-C--G-GTC----YAT-AATC-C--TT- 1568
SMM.US.x.PGM53          -----G--G--A--G--G--CCAGAG---C-----TA-------C----GCT-------A--A---A-AGAT-CA-CA--G--G-----------TC----AC--C--A-T--------T---------------ATTGG-GC-C--G-GTC----CATGAATC-C--TT- 2012
SMM.US.x.SIVsmH635F_L3  -----------A--G--A---CA-AG---C-----TA-A-----C----G---------A-----A-AGAC-C--CA-----G-----------CC----AC--C--A-T--------T---------------ATTGG-GC-C--G-GTC----TATGAATC-C--TT- 2098
SMM.SL.92.SL92B         -----T--------G------CA-AGT---------TA-------C----G-T-G-----A-----GA-AGAT-C--CA--G--G-----------C----AC--C--A-T--G--C--C------T-----C--ACTAG-A-------GG-----CATGAATC-C--C-- 1526
STM.US.x.STM            -----G-----A---A-G--CCA-AGT---------A-------C--G-G-T-G-----A--A-----AGAT-CATCA--G---------------TCGG--AC--C-AA-T--------T--T-----C--C--ATTGG-A--G----GTC----CATGAATC-C--C-- 1745
MNE.US.x.MNE027         -----G-----A--G--A---CAGAG----------A-------C----GCT-------A--A---A-AGAT-CA-CA-----G-----------TC----AC--C--A-T--------T-------------C---CTGG-GC-G--G-GGC----TATGAATC-C--C-- 1563
DRL.x.x.FAO             -----T--T--A-----A--C-AG------------TA-A--T------G-C--C--G--A--A--CA-AGATGCA-CA-----------------C----AC--C--A-A--------A--T-----C-----AGTG---C-C----GCC-A--GATGAACC-C-GT-- 1327
MND_2.x.x.5440          --G--------A--G--T----AG--------G--TA-A-----C----G----GC-G--G----GCA-AGAT-CAC----C----C---------GC-G--A--AC-CA----G-----C---------------AG---CCC-C----GG--A--CATGAATC-C---T- 1330
MND_2.CM.98.CM16        --------------G--G--C-AG--------G---A----T--C----G----GC-G--G----GGA--GAT-CAC-A--C----CC---------C----T---C-CA----G-----T---------------AT-C--CC-T----GGC-A--CATGAATC-CT-TT- 1773
MND_2.GA.x.M14          -----G-----A--G--T--C-AG--------G---A-A--T------G----GC-G--G----G-A-AGAT-CAC-A--C----CC---------C----A---C-CA----------C-----------------GC--C--G----G------CATGAACC-CT-TT- 1697
OLC.CI.97.97CI12        ------G-T--A--G--T--CTC-------------A-------T---GG-CT-G--G--A--------GG--C--------------CA----------G-AGC----AA-A--G-----T--T-----C--C--AGGAG-C--G-----G---...GT-AAAC--T-G-- 1676
GRV.ET.x.GRI_677        --G-----T--A---------CCAG-----------A-A-----------G-C--G-----A--A-----AC------T--T---------------C----T---C-TA----G-----C--T--G---------ATTG---C-G----G-------ATGAATC-A--CT- 1933
VER.KE.x.TYO1_patent    --G-----T--A--G-----C-AG--T-----G--A-A--T--C--GG-AA-T-----A--A-----C---GG---A--G---C-A-----------T-A--AC-CA-T---------T--T------------GTC--CC-G--G-GCC-A---ATGCACC-C--C-- 1993
VER.KE.x.9063           --G-----T--A-----T--C-AG--T---------TA-------C---G-AA-C-----A--A--G--C---GGA--A------C-A-----------T-A--AC-CA-A---------T-----------C--AGTC--CC-G----G---A---ATGCACC----CT- 1996
VER.KE.x.AGM155         -----G--C--A-----A--C-A---T---------A-A---------G-AA-------A--A---------GGA--A--C---C-A-----------T-T---C-CA-T--G-----C-----------C--AGTA-----G--G-GCC-A--GATGCACC-C--T-- 1985
VER.DE.x.AGM3           --G--G--C--A---G-A--C-A------C-----TA-------C---G-AA-------T-----G--C---GGA--A------C-G-----------T-A--AC-CA-T--G-----T--T-----C-----AGTC--CC----G-GCC-----ATGCATC-C--T-- 1497
TAN.UG.x.TAN1           -----GG-C-----G--T----AG-----C------A-A--T---C--G-C--T-----A--A---A-AC-C--A--T----------------------AC--C-CA----G-----C-----------C---TTGG---G----G-C-T---AT-CATC-A--T-- 1956
SAB.SN.x.SAB1C          -----G--T--A-----T--C-AG--T---------TA-A-----C--GG----G-----A-----GA-AGAC-CT-CT-----G--C---------C-GT-TC-----A---G-----C------------------AG-GC-C----GG--A--CATGAATC-A--C-- 2086
MND_1.GA.x.MNDGB1       --G-----------G-T----A--------------TA-A-----C--TGTAA-G-----A--A---------GGA--A-----G-TG-----CAGC-GCATC--C-TA-AG-------A--------A--C---CAG-----G-G-AGC--A--GAA--GA--A--TT- 1447
DEB.CM.99.CM40          -----G--C--------A--C-A---T-----T---A-----------G--CT--------T--A-----AGAT----CA----------------CACC---C-CA--A-T--------A-----T-----C-G--TC--CC-G----G---A--T-AGAATC-C--CT- 1596
DEB.CM.99.CM5           --G--T--T-----------C-A-------------TA-A-----C-----CT-G--G--T--A-----AGAT----CA--C--------------CACC---C-AA--A-T--G-----T-------------GA-TA---C------GG--A----AGAACC-C--TT- 1587
TAL.CM.01.8023          -----C--C-GA--G-------AG--G--C------A-------C-G-TGC--T-----A--A---A-AGAT----CA--C----------------C-G--GC----A--G--G-----C---------A--C---G-C---C-G--G--CA----T--CCA---G--C-- 1203
TAL.CM.00.266           -----C--C-G------A--C-A-------------A-------C-G-TG---T--G-----A-----A-AGAT----CA-----------------C-G--GC-----------G-----T-----------G--C---G-C--C--------CA----C----GA--C--CT- 1705
SUN.GA.98.L14           --G--------A--G--T--C---------AC---TA------TC--GG--T-------A--ATC--GC---G----A------G-G-------A---G-AAA--C-A--G-----------A----------------ATTGG-AA----G---C-----GA-A-TC-AT--T- 2060
WRC.CI.97.97CI14        -----G--T-----GGAT--C-A--C----------TA-A---------GTAT-------A--A----GAG-C-CA--A------C-A-----------C-AGC-C--AA-A--G-----T--T--G-----C--A--A--C-------G---A...-A-CAAC-A--CT- 2018
WRC.CI.98.98CI04        --G-------------GAT----A-ACT------G--TA-A-------GGTC--G-----A--A----GAG-C-CA--A--G---C-G-----------C-AA--A---A-A---------T-----------------AG----C-----G-G---A...-A-CA-C-A---T- 2036
WRC.GM.05.Pbt_05GM_X02  -----G-----A---GAT----A-TCT---------TA-A--T-T---GGTC--T-----A--A----GAGTT-CA--A----GC-A-----------C-AG--AC-AA-A---------A--T---------C--ACAG-----G----GCC-A...-A-AGTC-A---T-  980
LST.CD.88.524           --G--T-----A-----A--C-AG-------CT--TA------T---GG-CT-------A--AGG--G-------T--A--G---G-G-------A---G-AAA--C-A--G--G-----A-----T--C--C--A-TGG-AA----G--C--A---GA--GT--C--TT-  983
LST.CD.88.447           --G--G-----A-----A--C-A---A----C----A------T----G-AT-------A--AGG--GC-----T--A-----G-A-------AG--G-AGA--C-A--A---------A---------------G--TGG-GA-------C--A---GA--G------TT-  989
LST.KE.x.lho7           --------C--A--G--T----A---A----C---TA------T----G-A--G-----A--AGG--GC--T--T--A-----GG-G-------AG--G-AAA--C-T--G--G-----A-----T--C--C-G-TTAG--A--------C-A---GA--G---AT-T-- 2073
SYK.KE.x.SYK173         --G--C-----A---ATT--C-A-------------A-A-----CC-TTGCT-------A-----G--GGAC-CATCA--C---GGG---C-A--TC----AC-----A-T--------C--------A--C-GACAG--CC-C--G---A---TCAA-...C----CT- 1897
SYK.KE.x.KE51           --------T--A--A-A----AG--T-----T--TA-A------C-TTG---T-----A--A-----GGAC-C--CT-------------C-C---C-GT-A--AC-AA-T--G-----A-----T--G--C-GA-GC---C-G-------A---TCAA-...C-A--CT- 1525
COL.CM.x.CGU1           --G---AGT--A--GGAT----A-AG----------A-A--T---TC-G-C-----G--A--A-CT--AG--GGA--AA-C--GGCC-------G-TA-T-AT---C-AA-A--GC----C-----T--C--C--A-GA---C-G----G-C--CAGAAG...C-AT-TT- 1604
```


PLV Complete
Genomes

```
                                                                                                                                        Gag-Pol fusion TF protein start
                                                        Gag p24 Capsid end_Gag p2 start                                      Gag p2 end_Gag p7 nucleocapsid start
H1B.FR.83.HXB2             AGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGGACCCGGCCATAAGGCAAGAGTTTTGGCTGAAGCAATGAGCCAAGTAACAAATTCAGCTACC..............ATAATGATGCAGAGAGGCAATTTTAGGAACCAAAGA..................... 1950
Gag                        __E__E__M__M__T__A__C__Q__G__V__G__G__P__G__H__K__A__R__V__L__A__E__A__M__S__Q__V__T__N__S__A__T__.__.__.__.__.__I__M__M__Q__R__G__N__F__R__N__Q__R__.__.__.__.__.__.__.__
H1A1.UG.85.U455            --------------------C--------G-----------------A----G-----------G--------T------CA-C-GA--AGC.............................---------------------------GG--CG..................... 1399
H1B.US.90.WEAU160          ------C----------------------G-----G-----------A---------------------------------------A--------..................G------------------------G--C----.................... 1949
H1C.ET.86.ETH2220          -----------------------------G-------T-C---C--A-------G--------G-----------------AC---A--A-C.................-----------A-A---------A-GG--CT-A-.................... 1321
H1D.CD.84.84ZR085          -----------------------------G--G--G---A-------A-----------------G-------------C-----G-G-CT--G-TGCC.............----------A-A-A---------A-GG-AC----.................... 1470
H1F1.BE.93.VI850           -----------------------------G--------TA---------------------C--G--------T----C--ATTCAG-C........................-----------A-A-T-------A-GG-------.................... 1287
H1G.SE.93.SE6165           -----------------------------G---------A-------A-----------A-----G-----------G-CCT--GG-G----AG-AGCC.............--------------A---------A-GG--C----.................... 1356
H1H.CF.90.056              -----------------------------G--------TA-T-----A-----------------G---------------------A--AA---AGCC.............------------A-----C----A-GG-------.................... 1300
H1J.SE.93.SE7887           -----------------------------G--------T--T-------G-------------------------G--C---A-CAAC.........................---------A-----T--C------G----T-A-.................... 1270
H1K.CM.96.MP535            ------G----------------------G-----G--TA-------A------A---------G----------G---------C---T-...............G------------A------C----A-GG---T---.................... 1146
H101_AE.TH.90.CM240        --------------------C--------G--------TA-------A----G---------------------C-C-CA-C--G--A--................---------------------C-A-GG---G...................... 1509
H102_AG.NG.x.IBNG          -----------------------------G--------T--------A----G-----------G--------T------CA-C-GG-CAAC..............G--------------------------GG---G...................... 1475
H1N.CM.95.YBF30            ------G--------G--C--C-------G--G-----A-CA-------------GC-A--A--G--T---TCA--G--GCAGC-GC--A-A--TAGT...........G-CT-TGCA--A--G--A--C----AAGG-AT---G.................... 1517
H1O.BE.87.ANT70            ------G------GT---C-------------------AACT--------C-----GC-A--A---------GCTAC--CCCAGC-AGATTTG-AAGGAGGATACACAGCAG--T-C-----A-----GC-AAA-CCA-TTAGG-A-.................... 2014
H1O.CM.91.MVP5180          -------------GT---C-----A------------G--AACT--C-------A-A-AC-A--A---------GCTTCT-CCCAGC-AGATTTA-AAGGAGGATACACAGCAG--T-C-----A-----GC-GAA-CCA--TAG--A-.................... 1989
CPZ.CD.90.ANT              ----------T-A-----T-----A-----------C--A-C-----------------G-----A-----T---GCTTCT-CT-AT---G--CAGGGAACC........GCAG-CT-TC--------------GGA-AT-GAGG-G--AAA.................. 1380
CPZ.CM.01.SIVcpzCAM13      ------------------C-----A--G-----------A-C-------------C--A--------T---CAATG--CCA--G-CA---A-GAACA........GAC--TT-CT-T----A---GCCAGGGGCA-C--C-----.................... 1533
CPZ.CM.05.SIVcpzEK505      G--G-----------------C--A--G--G---------A-CT-------------AC-A--A-----C----CA--G-C-CA--CAG--A---GT..............G-TT-T-----A-----A-------AAGG-AT----.................... 1498
CPZ.CM.05.SIVcpzMB66       -----GGGATGA-C----C-TC--------G---------A-------A--T-----CC---A-----------T----C-CA-C-CAGTAA-GATGCA..............-A--G---ATTTAAAGGCCCA-AA........................ 1484
CPZ.CM.98.CAM3             ---------------------C----------G----A--A-C---------T-----AC----------T--------C--CA---CC-TA-A..............GGTG-GT-CC----A--G--A---AA-G----A-C--CT.................... 1340
CPZ.GA.88.GAB2             ---G--------------C--C--A-----------C--G-CA-------T------C-A--A-----T--------A--CA-C-GA--AGC..............TCA---T-C-----A----CAGGGGGA--A-CT-CTCCT.................... 1328
CPZ.CM.98.CAM5             ------------------C--C--------G------A---A-C--------T------C-----------C-----------CA------CT-A..............AATG-GT-T------------A---AA-G---GAGT--AT.................... 1643
CPZ.CM.05.SIVcpzMT145      ------------------T-----A--G-----------TCA--------T-----GC----A-----C--------G----AC--AA--..............GAA--CT-T-----A-----GGGACAA-AT-GA-CTCCTAAT.................. 1489
CPZ.TZ.00.TAN1             ----------C-C-----C-----A-----G----------TCT---------AGA--C-------------GC-TC--C----GCAGGG-GA..............GTA-AT---C-----G----A--AAGACCACC-TT--A-.................... 1591
CPZ.CM.05.SIVcpzLB7        ---G-----------------C--------C--G--T--A-C-----A--C--G--C--A--A-----T----T----C--ATGCAG--..............ACAGTC--T--------A--G--A-------AAGGA-C--A-.................... 1496
CPZ.US.85.CPZUS            ----------G-------------A-----G----------CT----------G--C--A-------C---T-----A-G-A---CC-CT-A..............AGCG--T-TC-A----A---A--CGC-G----G-C-G-G.................... 1998
CPZ.GA.88.GAB1             ---G--------------T-----A----------------ATCT-----------G---C-A-------C---TCAATG--TCAG---CA--GG-GAGCA........GATG-TT-CT-C-----AG--AC-AGG-GC-GG--C--A-.................... 2017
MON.CM.99.L1               G--G--C---C-ACAG---------------G-----T--AATG--C--AT-T--GA--C----A-----T---GCAGGG-C--T-GCCAAC...........ATGCCCATG-AT---G-A---........................................ 1593
MON.NG.x.NG1               ----------C-CCA---------A--G---------C----CA------AGC--GT-A--A--A--G--T---GC-AC--C--TT---AGCAAC-TGCCA...ATG......-AC---G-A--AG-CA-AGGAGG-................................. 174
GSN.CM.99.CN166            G---------C-T---------C--A--G-----G-----AACT--C--AT-T--GC-C--A--A-----C---GCAAC--CCCTC--AGGCA---GTTCA........TAC-AC---G------A----CCCCC-G-A-GG---.................... 1630
GSN.CM.99.CN71             G---------C-T---------C--A--G-----G-----AACT----AT-T---C-C-----A--G--C---GCAAC--C-CTC--AGGC...G-TCA........TAC-AT---G----A------CCCCCAG-A-GG---.................... 1633
MUS_1.CM.01.1085           G--G------T--CAG-----C---------G---------ACT----AT-C--GC--A---A--G--TC--GC-TCG-CCTTC...AA-CAGG-AGGC...ACACTG...-AC---G----AG---C---GGGAGCC-GGGG---G.................... 1604
MUS_1.CM.01.CM1239         ----------T--CAG-----C--A--G---------G--A-CA-----T-C--GC-AA---A--G--T---GC-TCG-CCCTT...AG-CAA-A-AGT...CAGCTG...-AC---G----AG--.........GCC-GAGGG-A-.................... 1597
MUS_2.CM.01.CM1246         -A---G--T-ACAG--C--C----GA-T-AG------ATG--C---T----GC-GA---A-----C----CTA-T-CTTTG...AGGCAA-A-ACCATTAACACCATC-AC---G-A--A---......CAG--TCC-AGG.................... 1684
MUS_2.CM.01.CM2500         G---------C-CCAG--T-----A-----T-----G--TATA--C--AT-C-A-T-AA---A-----C---GCATC--CTCT-...AG-CAG-A-ACCATTGAGACCATC-AC---G-T---G-G......CAAG-TCC-AG-.................... 1646
RCM.NG.x.NG411             G--G--G---C-A--------C--A--G-C------T--ACAG--C-----T--G--GC-A-----------CAGATG--GCA-.....................AGTAAT--C---GCC--ACAG--AGGACC---A--AGG-CCC.................... 1372
RCM.GA.x.GAB1              ----------C-A-----C--------G-------------CAG-----A--TC-G--AC-A--A--G--C---CAAATGA-GCA--G-AAT.........................--C---GCT---CA-TCAGCAAAC.................... 1360
DEN.CD.x.CD1               G-----G---T--CA---C--C----------------G--A--A--A-----CC-TCCCAC---A--G---T-AGCTTC--C-TTTC-GAAGCAAGGGTTATGCCATGGTTCA-GG-AAGG-GCCCA-CC-CCAG-AGAA-GA.................... 1732
                                                      Gag p24 Capsid end_Gag p2 Spacer start                              Gag p2 Spacer end_Gag p8 Nucleocapsid start

MAC.US.x.239               ----------C----G--T-----A--------G--G--G--A--G-----T---T-AA----A-----CC---AAG-G-CCCTCGCAC---TGC-AATCCCT......TTTGC-GCAGCC--ACAGA-GGGACCA.............................. 2213
Gag                        __E__E__M__L__T__A__C__Q__G__V__G__G__P__G__Q__K__A__R__L__M__A__E__A__L__K__E__A__L__A__P__V__P__I__P__.__.__F__A__A__A__Q__Q__R__G__P__.__.__.__.__.__.__.__.__.__.__
H2A.GW.x.ALI               ------G---C-A--C--C-------------G-----A-----G--A--C---T-AA----A-----CT-A-AGG-G-CC-TG-CAC-----C-TATCCCA.....TTTGCGGCAGCC--ACA-.................................. 2250
H2A.DE.x.BEN               ------G---C-A--C--C--C-----G-----C-----A-----G--A--C--GC-AA----------CC-A-AAG-G-CT-TGGGAC--AGCC-TATCCCA.....TTTGC-GCAGCC--ACA-.................................. 2257
H2A.SN.x.ST                ----------C-A--C--C--------G-----T-----A-----G--A--C---T-AA----A-----CT-A-AGG-G-CC-TGGCAC----CC--ATCCCA.....TTTGC-GCAGCC--ACAG................................... 1702
H2B.GH.86.D205             ---G------C-A--G--C--C--A--GA-------C--A--G--G--------GC-AA----C-----CT-A-AAG-G-CCCT--CAC-T--AC--ATACCG.....TTTGCTGCCG-T--ACA-AAAGCAGGG-A-.................... 2230
H2B.CI.x.EHO               ----------C-A-----C--C-----GA-------C--A--G--G--------GC-AA-------G--TT-A-AAG---CCTTG-CAC-TT-C---AATCCG.....TTTGCCGCCGCT--ACC-A-AGCAGGG........................ 2228
H2G.CI.92.ABT96            ---G------T-------C-----A--GR----------A--G--A--A--C--GC--A-----------T-A-AAG---CCTT---CC-TA-AG--CTACCA.....TTTGC-GCTGCT--ACA-AAA-CAGGAG-A....................... 1632
H2U.FR.96.12034            ------G---C-A-----C-----A---A----------A--A--A--A--T--GC-GA----A-----CT-A-AAG-G-CCTT-C-AC-TATGC-TATACCC.....TTTGC-GCAG-C--ACAGA-AGGCGGCCCA...................... 1748
MAC.US.x.EMBL_3            ----------C----G--T-----A--------G-----A--A--G-----T---T-AA----A-----CC---AAG-G-CCCTGCGAC---TGC-AACCCCT.....TTTGC-GCAGCC--ACAGA-GGGACCA........................ 1699
SMM.US.x.H9                ----------T-------C--------------R--G--A--A--A--AS-T---C-AA----A--G---T---AAG-T-C-CTG-CGCA--GGC-ACTCCCA.....TTTGCRGCAG-C--ACAGAAAGGACAA....................... 1698
SMM.US.x.PGM53             ----------C-------C-----------------G--A--A--A--A--T--GT-AA----A--G---T---AAG-G-C-CTG-CGC---GGCAACTCCCA.....TTTGC-GCAG-C--ACAGAAAGGACCA....................... 2142
SMM.US.x.SIVsmH635F_L3     ----------C-------C---------A-------G--A--A--A--A--T---T-AA----A------T---AAG-G-C-CTG-GAC---ACCAACTCCCA.....TTTGC-GCAG-C--ACAGAAAGGACAA....................... 2228
SMM.SL.92.SL92B            ------G---C-A-----C-----A--G--------C--A--A--G-----CC-CC-CA----A--G--C----AAG-C-CCCTC-CGGGGT--TTAGTG...............GCAGCA---TTTA-GGGAGCAGC---AGGCCA-.................... 1659
STM.US.x.STM               ------G---T-A-----------A--G--T--G-----A--A--G--A--T---C-GA----A-----CT---AAG-G--CTTCC-AC---ACC-ACTGCCC.....TTCGC-GCAGCC--ACAGCAGGGA.......................... 1872
MNE.US.x.MNE027            ----------C----G--T--C--A--------------A--A--A-------A-T-AA----A-----CC---AAG-G-CCCTTGCAC---GGC-ACTCCCT.....TTTGC-GCAGCC--ACAGAAGGGACCA........................ 1693
DRL.x.x.FAO                G---------C--TT---T--C-----T-----------A--------------G--CA----A-----T-----AG--CAGCAGGCAG-T.....................-----------CAGAATCCCCCC--AGGT-C-CCG.................... 1454
MND_2.x.x.5440             G---------T--CTG--C--C--A--G-----------TAAAT-C-------CA-A-GA----A--G--C---CAG----C-CAGGCAG-T.....................G-----------ACTCGGGGGGACC-CCG-GGG-T.................... 1457
MND_2.CM.98.CM16           ------G---CC-CT---C--C--A--G--------C--TAAAT----A---CA-A-GA----A--G--T----AGG---CCCAGTCA.....................GCTG-----------ATTCGGGAGGGCCACCG-GGG-C.................... 1900
MND_2.GA.x.M14             ----------T-ACT---------A-----G--G-----AAAAT---------CA-A-GA----A--G--C---CAGA-GA-GCAGTCA.....................GAAG-C----------ACTCGGGGGGACCACCG-GGG-T.................... 1824
OLC.CI.97.97CI12           G-----G---C-------C-----A--------G-----A---T----A--------C--A--A-----GC----A-C---GCAG-TGGTGCAACAGAGACCAGCACCA.................................................... 1785
GRV.ET.x.GRI_677           G--G------C-A-T---T--C-----------G--ACAA--------T-AGC-AA---TA---ATG-----TA-T-G-CAG..........................-AT---G-C--AGTG--ACC-CAG-AA--GGGCCCC.................... 2054
VER.KE.x.TYO1_patent       T---------T-A--G--T--------G-----C--AA--T-C--A----A---AA----A---ATG---CAGACCA-GCA----CA-.....................-AC---G-----CAG--AGG-CCA-AA-GA---.................... 2117
VER.KE.x.9063              ----------C----T--C-----A--------G--C--AA--T-C--A---------A----A---ATG---CAAA-CA-GCAG-C-CA-....................-AC---G-----CA---GGCAGCAG-AGGAGT---G.................... 2123
VER.KE.x.AGM155            T---------C-------C-----A--G--G-----C--AA-TT-C--A--C-A---CA----A---ATG---CAGA-CC-GCAG-G-CAG....................-AC---G-A---CAG--AGG-GGA---GGAAG-CC-.................... 2112
VER.DE.x.AGM3              ----------T-A--T--C-----A-----G---------AA-TT-C--A----A---GA----A---ATG---CAAA-TA-GCA--GCCAG....................-AC--------ACAG---GG-CAG--AGGAAG-CC-.................... 1624
TAN.UG.x.TAN1              ------G---C-------T--------------------A--A-----A---AGC-CA---TA---------CAA---A-GCA-GGA...................GTT-AC---G----AG---CACCGCGAG-AGGTAGGG-C.................... 2083
SAB.SN.x.SAB1C             ------G---C-T--T-----C--A--CA-T------G-ACAA---------C-TC-AA----G-----T----CAGCG-CCTTTC-GCA-CAA--AGTT......GGGAAC--CT-TG-----CA---GGCAAGACCC-GGGGTCCC.................... 2228
MND_1.GA.x.MNDGB1          ---G------T--GA-----------------T--G--ACAA----A--C---T-AA----A---ATG-----AAC---GGT-GGACA-......TCA.........CA--ATT-TGT-CAGCAG-GAGGGCCTC-AAG-.................... 1574
DEB.CM.99.CM40             G-----G---C-ACAT-----C--A--G--G-----C--ACAG--A-----C---C--A----G--G--T---GC-TC--CCTT---AGAG--AGGGTCC.........CTAGGC---G-A--ACA-A-G-GAGGACCCCCTGGTTCT.................... 1735
DEB.CM.99.CM5              ----------T--CAT-----C-----------T---CAA--G-A--CC-TC-GA---G--G--T---GC-TCG-CCTT---AGA---AGG-TCC.........CTGGGG---G-----CA-A-AGGAGGACAAGGAGG-G-T.................... 1726
TAL.CM.01.8023             G---------T--CAG--------A--G--C--G------TCAG-----AGC---C-CA----A-----C---GCGG-G-CCTTG--A...CAG-A-ACCCAGCAAGTCATAGCC---G-A---...--AC-GGGAG-TCCG.................... 1339
TAL.CM.00.266              ----------C-CCAG--C-----A--G-------------CAG-----AGC--GC--A----A-----T---GCAG---CCTT---A...CAA-A-ACCCAACAAGTCATAGCC---G-A--ACAG--GC-AGGAG-CCCA.................... 1844
SUN.GA.98.L14              ---G--G-------G---C--C--A--G--T--G--G--GA-------A-G--AGA-CC----A---------GC-TCT-CTTTT-GGCAGCAAGTAGGA......AGACAAGCT---G-----CA-AA-TTACCACCC-GAA-C--CCAG.................... 2205
WRC.CI.97.97CI14           ------G---T-A-----------A-----------G--A--AT----AAGC-A-A-AA----AC----TT-ACAAG---CTTTG...CAT...T-ACAG...........GCC---G-A--ACAG-CAGGAGCA-AACCT.................... 2142
WRC.CI.98.98CI04           ----------T----T--------A-----------G--A--AT----AAGC-AGA-AA----AC----CT-ACAGG-G-CTTTG...AGC...T-TCAG...........GC----G----ACA-CAGCCAG-G-AACCT.................... 2160
WRC.GM.05.Pbt_05GM_X02     ----------T-------------------------T--A---T----AAGT-A-A-AA----AC-----T--CAGG---CCCTG...--CCAAT-ACAG...........GC----G-A--ACAGCAGGCAG-C-A-.................... 1104
LST.CD.88.524              --------------A---------------------T-C------AAG--A-A-C--A--A--G-----GCAAC--CC-TGC-GGGGCAG-TGAAA.........CATC-C---G-----CA-ACACCACC---A--TGC-CA-.................... 1122
LST.CD.88.447              ------G-------GG--C-----A-----------A-C------A-G--GA-------A---------GCAAC--CC-TGC-AAGTCAA-TGAGA.........CAG-AT---G-A--AGTGACACCACCA-----TGC-CAG.................... 1128
LST.KE.x.lho7              ------G-------A---T-----------------T-CT-----A-G--A-A-A-----A---------GCTTC--CC-T-C-GCA-CAAGTAAGA.........CAA-AC---G-----GTCACTCCGC----A--TGCGCA-.................... 2212
SYK.KE.x.SYK173            G--G------C--CA---T-----A-----C---------TTA--C-------A-C-CA---A--G--C---GTAATG-CCCAGC-G----TA...................-AT---G-A--AG-GCCATCAAAAG--.................... 2018
SYK.KE.x.KE51              G---------C-TCA---T--C--A--G------------ACAAT-C--A----A-C-A--A--A---------GTAATGAC-CA-C-CAGCTTG..................GGC-----A--AG-GCCA-GACAAG-A-G-A-TCCC.................... 1655
COL.CM.x.CGU1              G-----C---C-AG-------C--------G---------A-A---C-------AG--GC-A--A--G--T---CAG--GT-CCAGC-AGAGAGA...............ACC-AC-----TG-AGT-AAG-CAGCA-AATG-TTC-ATTGTCAGGGAATAGGGCACTTAG 1759
```

```
H1B.FR.83.HXB2              ...................................................................................AAGATTGTTAAGTGTTTCAATTGTGGCAAAGAAGGGCACACAGCCAGAAATTGCAGGGCCCCTAGGAAAAAG............... 2022
Gag                         .__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__._K__I__V__K__C__F__N__C__G__K__E__G__H__T__A__R__N__C__R__A__P__R__K__K__.__.__.__.__._
H1A1.UG.85.U455             ...................................................................................-GA-GAA-------------C--------------A---CT-----A------T-------------------............... 1471
H1B.US.90.WEAU160           ...................................................................................--A-C-A--------------------------A----T---------------------------------............... 2021
H1C.ET.86.ETH2220           ...................................................................................-GAGCAA----A--------C--------G---------CT------------------------------A............... 1393
H1D.CD.84.84ZR085           ...................................................................................--A----------------C-------------------T---A--------T------------------............... 1542
H1F1.BE.93.VI850            ...................................................................................-GAG-------A-----T----------------A----T----------------------A------............... 1359
H1G.SE.93.SE6165            ...................................................................................-GA-C-A-C--A-----------------G-----A--TCT---------------------------............... 1428
H1H.CF.90.056               ...................................................................................--AT-------A--C-----C-----------G--A----T-----------------------............... 1372
H1J.SE.93.SE7887            ...................................................................................-GA---------------------------C----A----T---A-A---C--------------A------............... 1342
H1K.CM.96.MP535             ...................................................................................--A----------C-----C--------------------T----------------------A------............... 1218
H101_AE.TH.90.CM240         ...................................................................................--A-GAA-------C-----C-----T-G------A---CT----------------------A---C--............... 1581
H102_AG.NG.x.IBNG           ...................................................................................-GA-CAA-A----------C--------------A---CT---------------A--------------G-............... 1547
H1N.CM.95.YBF30             ...................................................................................--ACCCA----A-------------------G--C--TTTG--A-----C--T-A---------A-G-GGA............... 1589
H1O.BE.87.ANT70             ..................................................................................GGA-C-A-A--A---------C-----A-----G--A--T-T---A---------TC-A--A------------............... 2086
H1O.CM.91.MVP5180           ..................................................................................GG-CCCA-A--A--C-----------A-----G--A--T-T---A-A---C--TC-A--A-----A----G-............... 2061
CPZ.CD.90.ANT               ...................................................................................-GACC-C-C--A-----T--C--C--T-----G--C--T--T--A-----------A------A--A-GG--A............... 1452
CPZ.CM.01.SIVcpzCAM13       ...................................................................................-GA-AAA-A--A--C-AT--------A--G-----C--TCTG--------C--T-A---A--A----G----............... 1605
CPZ.CM.05.SIVcpzEK505       ...................................................................................----CCA-C--A--C--T--------T-----G--C--TCT---A-----C--T-A---A-----A------............... 1570
CPZ.CM.05.SIVcpzMB66        ...................................................................................-GA-----C--A--C--T----------G-----T----T---T--G--------A------C--A-G---A............... 1556
CPZ.CM.98.CAM3              ...................................................................................-G--AAA----A--C--T--C--------G-----A--TTT---------C-----A-----C--A-GG---............... 1412
CPZ.GA.88.GAB2              ...................................................................................-G--GA--C-----C-----------A--G-----T--T-T---A---------T-AA---------A-G----............... 1400
CPZ.CM.98.CAM5              ...................................................................................----AAA----A--C-AT--------------------TGT---T-----C-----A-----A--A-G----............... 1715
CPZ.CM.05.SIVcpzMT145       ...................................................................................--A-AGA----A--C--T--------A---------A---CT---A--G--C--------A--CC-A-G----............... 1561
CPZ.TZ.00.TAN1              ................................................................................AAGGGTCAGC-GC-------T--C-----G----T---C--T-----A---------T-----T--A--A--G--A............... 1666
CPZ.CM.05.SIVcpzLB7         ...................................................................................-GAG--A---A-------------------------------T---A---------T-AA-----C----G--A............... 1568
CPZ.US.85.CPZUS             ...................................................................................-GA-AAA-C--A-----T--------------G--A--TCT-------------------T--A--A-G--A............... 2070
CPZ.GA.88.GAB1              ...................................................................................-GA-AAA-A--A--C--T--------G--------T---CT---T---------T-AA--A--A--A-G---A............... 2089
MON.CM.99.L1                ...........................................................GCCAGAGGACCACCCCAAAGGAAGGGACAACCC--A--C----------A---TTT-------TG--G-A---------AA-----CCAA-GG-G-............... 1689
MON.NG.x.NG1                ..............................................................RGTCAGCCACGCCGACAAGGCRW-CAAA-C-G-----AT--------G---TTT--A--TGTG--A-AG--C--T-CT--T--C--A--G-CA...............  267
GSN.CM.99.CN166             ................................................................................GGA--A-CCCCA--A----AT--------AC--TTC--A---CT---A---G---TCCTAAA--C-AAG---GA............... 1705
GSN.CM.99.CN71              ................................................................................GGA--A-CCCCA--A----A----------A--GTTT--A--T-TC--A---G-C--TCCCAAA--C-A-G---GA............... 1708
MUS_1.CM.01.1085            ...........................................................GGCAACGGCAGCCAGGGAAACCGAGGA-A-CCC-GA--C-A---C----GC--T-T--C---GT------GG-----CCCAAA--C-----C--A............... 1700
MUS_1.CM.01.CM1239          ..............................................................GGCAGCCAGGGAGGTCCCAGGGGA-ACCC--G-----A---C----GC-GTTT--A----TG--A----GC---CCCAAG--A-AA-CT-G-............... 1690
MUS_2.CM.01.CM1246          ...........................................................GGAGTAATGGGAAAAAAGAGGGAA--C-GCACACGT--C-A---C----AC--TTT--A--TTTG--A---G-----CCCAAG----A-TCT-CC............... 1780
MUS_2.CM.01.CM2500          ...........................................................GGAGGGGCCGGAGGAGGGCCTAGA-G--CCCCACG---C-----C----GC--TTC--A--T-T-----A-G-----CCCAAA--A--AGTT-G-............... 1742
RCM.NG.x.NG411              ..............................................................CCTAAACTAGGGGGAGGACCT-G-T-CC-C--A--C-AT-----------AC---T--T--T--T---T------AAA-AAG---A--G--A............... 1465
RCM.GA.x.GAB1               ....................................................AGGGGGCCTCCAAGAAGATCAGGAGGAAATCCA-A-T-A-GA----A---------T--GCC---A----TTT-T---T----T-AA---------A-GG---............... 1462
DEN.CD.x.CD1                ...................................................................................-G-GGAC---GA---------------C--ATT--A---CT-----AGG----T--AAAG--C-A--G-GT-ACT............ 1807

MAC.US.x.239                ................................................................................AGA---CCAA---------GG--------G-----G--A---T-T--A--GC-A-----A-----A--A-G-C--............... 2288
Gag                         .__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__R__K__P__I__K__C__W__N__C__G__K__E__G__H__S__A__R__Q__C__R__A__P__R__R__Q__.__.__.__.__._
H2A.GW.x.ALI                ................................................................................AGA-G--CAA-------C-GG--------A--G---------T-G--A---C-A---C-A--A--C--A---C--............... 2325
H2A.DE.x.BEN                ................................................................................AGA---GCAA---G--A--GG--C-----A--G--G--A---T-G--A---C-G---C-A--A-----A-G-C--............... 2332
H2A.SN.x.ST                 ................................................................................AGA-G--CAA-------C-GG-----C--A--G---------T-G--A---C-A---C-A--A-----A-G-C-A............... 1777
H2B.GH.86.D205              ................................................................................AGAGG--CA--G-CA--C-GG--C---------C-G--A-----------GC-A---------------A-G-C--............... 2305
H2B.CI.x.EHO                ................................................................................AAG-G--CA--G-CA--C-GG--C--C-----G-CG--A--T--------GC-G----A----------A-GGC--............... 2303
H2G.CI.92.ABT96             ............................................................................AAAAGG-GC-CAA-A--A----GG-----------G--G--A--T----TT---C-G-----A-----C--A-G-C--............... 1710
H2U.FR.96.12034             ................................................................................AGG-G-CCA--C-GA----GG--C--------G-----A-----T--A--GC-G--T-AA---------A-G-C--............... 1823
MAC.US.x.EMBL_3             ................................................................................AGA---CCAA---------GG--------G--G--G--A---T-T--A--GC-A-----A-----A--A-G-C--............... 1774
SMM.US.x.H9                 ................................................................................RGA-----AA--W------RG--C---R-A--G-----A---T-T--A---C-A-T---A---------A-G-C-R............... 1773
SMM.US.x.PGM53              ................................................................................AGA----CAA-------C-GG--------A--G--G--C---T-T--A-A-C-A-----A---------A-G-C--............... 2217
SMM.US.x.SIVsmH635F_L3      ................................................................................AGG-G--CAA-C-------GG--------A--G--G--A---T-T--A---C-A----------------A-G-C--............... 2303
SMM.SL.92.SL92B             .........................................................................GGGAATAAGCCT--CA-C-G------T--C-----A--GAC---A---T-T--G---C-G-----A--T--C--A-GG---............... 1740
STM.US.x.STM                ................................................................................CGA-G--CA--A--A--C-GG--------G----C---------T--A-A-C-G--T-AA-G---G--A-G-C-A............... 1947
MNE.US.x.MNE027             ................................................................................AGA---CCAA---------GG--------G-----G--A---T-T--A--GC-A-----A-----A--A-G-C--............... 1768
DRL.x.x.FAO                 ...........................................................CGGGGACCCAGGGGACCCCCTCCA--A-ACCC--GA---CCA--C--C--GC-GTTT--A--T--CCTG---C-A--T-CAAAG-----A-------............... 1550
MND_2.x.x.5440              .................................................................CCCCCGAGACAACCCCCT-GA-ACCC--GA--CCC------C--A--GTTT--A--TGT-CTG--GG-------A-----A--A--G-GA............... 1547
MND_2.CM.98.CM16            .................................................................CCCCCGAGACAACCTCCT-GA-ACA-C-GA--CCCT--C-----A---TTT--T--TGGGCTG--GG-C--T-TAAG----C-A------............... 1990
MND_2.GA.x.M14              .................................................................CCTCCGAGACAACCCCCT-GA-ACCCC-GA---CC---C-----A-GGT-T--A--TGTTCTG---G-C--T--ACT---C--A--G---............... 1914
OLC.CI.97.97CI12            .................................................................GGAGGAGGGCCTCAGCAAGG--GGC-G-G-----AT--------T-G-TTT-----T--TAAA-AT-----TTCACAAAAGG-TCC-............... 1872
GRV.ET.x.GRI_677            ................................................................................CGAGG-CCGC-A--A--C--T------------TTT--A--T-TGCAA--GG-A----A---A--A--AC-G-TC............... 2129
VER.KE.x.TYO1_patent        ................................................................................AGACCCCCAC-A-GA----AT--------A---TTT--C--T-TGCAA---C-A--TCC--AA--A-------CA............... 2192
VER.KE.x.9063               .........................................................................AGACAGAGACCCCCAC-A--A----A------C--A---TTC--A----TGCAA--GC-A----CA-AA--A--A--G-TA............... 2204
VER.KE.x.AGM155             ................................................................................AGACCCCCGCCA-------A---C-----A---TTT--C----TGCAG--GC-G--TCCT-AG--A--A----TA............... 2187
VER.DE.x.AGM3               ................................................................................AGACCCCCA--A-------A---------A---TTT--C--T-TGCAA---C-A---CCT-AA--A--A--G-T-............... 1699
TAN.UG.x.TAN1               .........................................................................AGAGGCAGAGGCCCACCCCGC--C-----A-----AC--ATT-----T-T-CAA-A-G----TCCAAGAG-AG-ACC---C............... 2164
SAB.SN.x.SAB1C              .......................................................TTAGGGGGAAGAGGTAGACCTCTGAACCCA-A-A----A----A---------A--GCCT--T---CTG--T---TT---T-A-----------G-C-A............... 2327
MND_1.GA.x.MNDGB1           ...........................................................GGACCAGTTAGACAACCTACTGGA-G--AACC--TC--C-----C---AAT-----------TGT---A--GTTC-T--A---------A-GG--A............... 1670
DEB.CM.99.CM40              ................................................................................AGA-GACAGA---GA--C-----------AC--ATT--A--TCTCCAA-AGG-C----AAAGA-----A--G-CC............... 1810
DEB.CM.99.CM5               ............................................................................CCTAGA-GACAGC---GA--C--T--------AC-GATT-----TGTGCAA--GG----T-AAAAG-----A---GTT............... 1804
TAL.CM.01.8023              .................................................AAAGGGAGAGGGGGACCAAGAAGAACTCCCCCAGGACAAA-C-GA----AT-----------GTTT-----T-TT--A-A---C--TCC---T---CA---GCCACGGGGG.......... 1450
TAL.CM.00.266               .................................................AGGGGGAGAGGGGGACCTAGAAGGAACCCCCCAGG-CAGA---G-----AT--C-----A---TTT--C--T-T---A-AG--C---CC------GCC--G-C--AAGGTG.......... 1955
SUN.GA.98.L14               ..............................................GGGAGGTTTGTAAGAATAGGAGGAGGAGGACCTAGG---CCCA-G-CT-----T--------T---CC---A--TCT---A-ATC-G-----A-AG--A-A----GGA............... 2313
WRC.CI.97.97CI14            ................................................................................AAAGG-CCAA---GA-----T--------AC--ATT-----T-TG--A-A-G-A--TCCCAAG--A-A----CCAGTCCCTATGAGAAGA 2232
WRC.CI.98.98CI04            ................................................................................AAAGGACCAA---GA--------------AC-GATT-----T-TG--A-A-G-C--TCCCAAG--A-A----CCAGTCCCCTTGAGAAGA 2250
WRC.GM.05.Pbt_05GM_X02      ................................................................................AGAGGACC-C-A-GA-----T--------AC--ATT-----T-TG--A-A-G----T-A-AAA--A-AA-GGTCTCCTCAGCAAAGG... 1191
LST.CD.88.524               ..............................................GGGAGATTTGTAAGAACGGGGGGAGGTGGCCCAAGA-GACC-C-C-CT-------------G--GCC---A--TT----A-A--TG------CAG-----AC-GG-A............... 1230
LST.CD.88.447               ..............................................GGAAGGTTTGTAAGAACAGGAGGAGGAGGCCCTAGA---CC-T-A-CT--C--T--------A--GCC---T--T-----A----TG--T--ACAG-----AC-GG--............... 1236
LST.KE.x.lho7               ..............................................GGCAGGTTTGTGAGAACAGGAGGAGGAGGTCCCAGG--ACCCC-C-CT--C-----------A--GCC---T--------A----TG-----ACAG--C--AC-GG-A............... 2320
SYK.KE.x.SYK173             ................................................................................AGATCA--GA----A--C-A---------GC--ATT-------TGCAA-A-G-C----AAAAG--ACTT---GCT............... 2093
SYK.KE.x.KE51               ............................................................................AGGAGAGG-CCAACA-GA--C--T--------GC-GTT---A--TCTTCAA-A-G-C---CCTAGG--A-AA--GCTT............... 1733
COL.CM.x.CGU1               CCAGAATGTGTCCCAAAAGACCAATAGGAGGAGCCGGTCGAGGAAGAGGCCGAGGACGGGGAGGCTTTAGAGGAGCTCCCAGA-G-CCA--G-G------T-CC--CAATC-G-----A----TGCAA---G----TCCC.............................. 1899
```

```
                                                                                   Gag-Pol -1 ribosomal slip site
                                                                        Gag p7 nucleocapsid end Gag p1 start                              Gag p1 end Gag p6 start
H1B.FR.83.HXB2             ............GGCTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGATTGTACT..........................GAGAGACAGGCTAATTT.TTTAGGGAAGATCTGGCCTTC......CTACAAGGGAAGGCC.........AGGGAATTTTCTTC 2137
Gag                        _.__.__.__.__G__C__W__K__C__G__K__E__G__H__Q__M__K__D__C__T__.__.__.__.__.__.__.__.__.__E__R__Q__A__N__F#__L__G__K__I__W__P__S__.__.__Y__K__G__R__P__.__.__.__.__G__N__F__L__
Pol                                                                                                                  F#F__R__E__D__L__A__F__.__.__L__Q__G__K__A__.__.__.__R__E__F__S__S
H1A1.UG.85.U455            ............-------------C--G--A--------------------C--C---............................---------------------.--------A--T--------.....-A-------G----............-----------C-- 1586
H1B.US.90.WEAU160          ............-----------------C--------------------------------...........................-----------------.---------------T----......-C-A-------------..........----------------C-- 2136
H1C.ET.86.ETH2220          ............-----------------------------------C----C.........................--------------------.--------GAC-T--------......-A--------------.........--------------C---- 1508
H1D.CD.84.84ZR085          ............------------------------------------C---...................--A-------------------.--------A--T--------......-C--------------.........G-----C-------- 1657
H1F1.BE.93.VI850           ............------------------GA-------------------C--C---...................--A-------------------.--------A--T--------......-A-------G-----.........C--A-----C---- 1474
H1G.SE.93.SE6165           ............----------------------------T-----------C--C--G...................---------------------.--------A--T--------......-A-------G-----.........---------------- 1543
H1H.CF.90.056              ............------------------GA--------T--G--------C--C--A...................---------------------.--------A--T--------......-AG---A--G-----.........---A--------C- 1487
H1J.SE.93.SE7887           ............--------------------------------------C--C---...................---------------------.-------------T--------......-AG---A--G-----.........--------C-----C- 1457
H1K.CM.96.MP535            ............----------------G-----------T--G--------C--C---...................---------------------.--------A-------------......-C-------G-----.........---------------- 1333
H101_AE.TH.90.CM240        ............--T-----------C--G-----------T-----------C--C---...................---------------------.--------A--T--------......-A---------------.........G----------C-- 1696
H102_AG.NG.x.IBNG          ............----------------G--A--------------------C---...................---------------------.--------A--T--------......-AG------G-----.........---A-------C-- 1662
H1N.CM.95.YBF30            ............------------G-----GC-A--------T-----------------AAAATGAA.................GA---------------.-----------G----T--C-.........--T---A--G--A--..........---A--C--C-CC- 1710
H1O.BE.87.ANT70            ............--T--C------------C--------T-----------------C-GAAAT..................GA-A------A-----.------C--ATA--------C-......GGGGGGCACG-----.........---C-----A-G-G- 2204
H1O.CM.91.MVP5180          ............--T--C------------C--------T-----------------C-AAAAT..................GA--------A-----.-------------TA--------C-......GGGGGGCACG-----.........---C-----A-G-G- 2179
CPZ.CD.90.ANT              ............-----C----G-------C-------------GC-T---A-C---C-AGCAACA.................AATACAG--A-A-TA-----.---------ACCGACC--CA-.............-T-GT--GGGTGCAGACC------C---G-G- 1579
CPZ.CM.01.SIVcpzCAM13      ............-----C----------GC----G------------------C---T-A......................G------A-TA-----.---------GG--------------......-CG--GC--G-----.........---C--C---G-A- 1720
CPZ.CM.05.SIVcpzEK505      ............-----C----------GC-A--G-----T--------G--C--C-GAAGT...................GGA--A--G----G-----.---------G---------CT.........-C---A--G-----.........---A-----C-CC- 1691
CPZ.CM.05.SIVcpzMB66       ............-----------G-----GC-A-----G--T--G-----G-A---C--AAAT...................--A------------------.---------AC-----------......-AG--------------.........--------C---- 1674
CPZ.CM.98.CAM3             ............--A--C-----G-----GC-A--G--G--T---------A-C--CC-AGCT...................GGA--A-----------------.---------AGGT---T--C-GTG......G-GC--GG-AT-AAAGAGACC------C--C--G- 1542
CPZ.GA.88.GAB2             ............--A--------G--------A--------T--G------G--C--CCAAAAT...................GAA-GA-------TG-----.-------GG-GG---------CCG......G-GC--G-----..........------C----CC- 1521
CPZ.CM.98.CAM5             ............--A--C--------C---C-------G--TA-C---------A-----AACTGGA...................GC-----------------.---------GAT---T--C-GTG......G-GC--G--AT-AAAAAGACC------C---CG- 1845
CPZ.CM.05.SIVcpzMT145      ............-----C----G------C----G-----T--G--------------A-CAAAATAAT.................GA--G...AT-------.------A--A-AT------C-.........-GGGG---G--------..........------C-A-G-G- 1682
CPZ.TZ.00.TAN1             ............--T--C---GG------C-A--G-----T--------G--C--C--CACCAGAAAC...............AACAGC-CTGG--TA-----.---------ACG-ACC--C-T...............GT--GGGTGCAGACC------C---G-G- 1793
CPZ.CM.05.SIVcpzLB7        ............-----------------GC----------T-GG---C---------ACATG.....................G--AG...-T-------.---------AC----------.........--C-C---------..........---A--C----C-- 1683
CPZ.US.85.CPZUS           ............--G--C-----G-----GC----------T-----------C--C--AGCAGGA...................A-C------------------.---------ACAT---T-CC-ATCGTG...G-GC--GG-AT-AAAGAGACC------C--C--GG 2203
CPZ.GA.88.GAB1             ............---------CGG-----GC-A--G-----------------C--C--A.....................GA------TG-----.---------AGGT--------CCG.........-GC--G------..........---C--C---G-G- 2204
MON.CM.99.L1               ............AAA----AC--T-----CC-ACCG-------TCGCC-----C--CC-GCAGCCCCCC..........AAACAAA-C-A-GG--TG-----.---------TCCT-TCGG-C-.............CAAG-A-GGGCC.......-A----C--C-C-- 1816
MON.NG.x.NG1               ............--T--C-TCCGC--C--C---------G--TRT-TCC---A-C--CC-A.....AACGGAGGACAAAACCAACCCAGA-ACAG--TG-----.---------TCCT-TTGG...............CCCCG-CAAGAG......G---CCAAGGAACT  397
GSN.CM.99.CN166            ............AAG--C-TT--------C-GA-C---G---TTT-GC---CCA----GA...............ACAGGGACACCAATC-AGG--CT-T-----.-----ACCTTGG-C--GGGCG.............CC-CC-AG-CCC......CA----C----CAG 1829
GSN.CM.99.CN71             ............AAG----TT--------C----C--------TT-GCT-GGC-A--C-AAACA.............GGAACACCA-CC-AGG--CT------.-----ACCTTGG-C--GGGCG..................-C-AG-CCC......CA----C----CAG 1829
MUS_1.CM.01.1085           ............AAG----TC--------G-GA---------TTGGCA-G-C-A----GG...............TCTGATAATACC-A-GC--TG-----.--------G-TCC--TCGGCCT.......-C-G--AA-G------..........GA----C--C-CC- 1824
MUS_1.CM.01.CM1239          ............AAA--C-TC----------G---G------TTGGCA---C-G--C-GG...............TCTGAAGGT-CC-A-TC-------------.--------G-TCCTCTTGGCCT.......-C-G-GAA-G----------.......-A----C--C-CC- 1814
MUS_2.CM.01.CM1246          ............A-A----TT--G-----T-------------TTGGCC-G-C-G--CCGA..............ACCGACACAGGCA-A--CGCT-TG-----.---------GATCTCGGC-T.......-CGGTCTC-G-A---.........GA----C--C-CAG 1907
MUS_2.CM.01.CM2500          ............AAA--C-TT--G-----C-G------C----T-GCA---C-G--C-GA..............ACCAACTCAGACA-AG-CGCT-T-------.---------AGGT-TCGG..........--T-C-CA-G-A---..........GA----C--C-C-G 1866
RCM.NG.x.NG411             ............--G--C----G------G-------T--T-T------G----CC-A..........................-A-G---T-------.---------TGCTCCCTGGAA.............-CAGG--AAACC...GAGGAACCT-CCCT-GA 1577
RCM.GA.x.GAB1              ............--A--C-----------TCCCC--AC--T-TCT-------C--C--A..........................-AG--AATA-----.--------GAC--CCCTGGGG...........TCAGG--AAGCC...GAGGAACTT-CC-T-GA 1574
DEN.CD.x.CD1               ...CAACCAGGA-CG--C-TT-----------CT---T--T---GCA--GA------GA...........AGTAACACTAATGATC--G------TA-----.---------TGGTCTTGGGAG............-AAG-AT--TCCCAGGAACTTTGTACAGTACT 1949
                                                                                         Gag-Pol -1 ribosomal slip site
                                                                                                 Pol start
                                                                         Gag p8 Nucleocapsid end Gag p1 Spacer start                      Gag p1 spacer end Gag p6 start
MAC.US.x.239               ............--A--C-------------AATG-AC--TGTT---GCCA-A--CC-A...........................--C--------GGG---.------CCTTGGTCCATGGGG............AAAG-A---..........CC-C-----C-CCA 2397
Pol                                                                                                                    F__F__R__P__W__S__M__.__.__.__.__.__G__K__E__.__.__.__.__A__P__Q__F__P
Gag                        _.__.__.__.__G__C__W__K__C__G__K__M__D__H__V__M__A__K__C__P__.__.__.__.__.__.__.__.__.__D__R__Q__A__G__F#__L__G__L__G__P__W__G__.__.__.__.__.__K__K__P__.__.__.__.__R__N__F__P__
H2A.GW.x.ALI               ............-----C-----------C---CC------T-TT---GC-A-C--CC-A.........................--A---------GG---.--------CTTGG-CCATGGGG............AAAG-A---...........TC-C--C--C-CCG 2434
H2A.DE.x.BEN               ............-----C-----G-----C---CC-------ATC---GC-A-C--CC-G.........................--A--------AGG---.--------TT-GG-CCA-GGGG............AAAG-A---...........TC-C--C--C-CCG 2441
H2A.SN.x.ST                ............-----C-----------C----C-------ATC---GC-A-A--CC-A.........................--A--------GGG---.--------TT-GG-CCATGGGG............AAAG-A---...........CC-C-----C-C-G 1886
H2B.GH.86.D205             ............--A--C--------------AAC-------ATC---TC-A-A--CC-A.........................--A--------GGG---.--------TTAGGACCCTGGGG............AAAG-A---...........TC-C--C--C-CCA 2414
H2B.CI.x.EHO               ............--A--C-------------C--C--------ATC---TC-A-A--CC-A.........................--A--------GGG---.--------TTCGG-CCCTGGGG............AAAG-A---...........TC-C-----C-CCG 2412
H2G.CI.92.ABT96            ............--G--C-----G-----G---CC---C---ATT---GCCA-A---C-A.........................-----------AGG---.--------TTTGGACCTTGGGG............AAAG-A---...........TC-C--C--C-CCA 1819
H2U.FR.96.12034            ............--A--C----------A----CC-------A-C---GCTAGC--CC-A.........................--A-CCA---TGGG---.-------CTTGG-CCTTGGGA............AAA--A---...........TC-C--C--C-CCA 1932
MAC.US.x.EMBL_3            ............A-A--C--------------AATG-AC--TGTT---GCCA-A--CC-A.........................--C--------AGG---.------CCTTGGTCCATGGGG............AAAG-A---...........CC-C-----C-CCA 1883
SMM.US.x.H9                ............-----C-----G--------A-C---C--TGTT---GCTA-A--CC-C.........................--A--------GGG---.-------CTTGG-CCATGGGG............AAAG-A---...........CC-C-----C-CCA 1882
SMM.US.x.PGM53             ............-----C--------------AAC---C--TGTC---GCCA-A--CC-C.........................--A--------GGG---.-------CTTGG-CCATGGGG............AAAG-A---...........CC-C-----C-CCA 2326
SMM.US.x.SIVsmH635F_L3     ............-----C---GGG--C-----AACG--T--TGTT---GCCA-A--CC--.........................--A--------GGG---.-------TTTGG-CCATGGGG............AAAG-A---...........CC-C-----C-CCA 2412
SMM.SL.92.SL92B            ............--A--C----------GG-A-------GTATTCA-GCCA-C--CC-A.........................A-CCAGA---GGG---.-----CCTTGG-CCATGGGG............AAAG-A---...........TC-C--C--C-CCA 1849
STM.US.x.STM               ............--T--------------C-AC-G--C--T--G---GCCA-A--CC-A.........................--A-------TGGG---.-----CTTTGG-CCATGGGG............AAAG-A---...........CCAC--C--C-CCA 2056
MNE.US.x.MNE027            ............-----C--------------C-AATG--C--TGTT---GCCA-A--CC-A.........................--C--------AGG---.-----CTTTGG-CCATGGGG............AAAG-A---...........CC-C-----C-CCA 1877
DRL.x.x.FAO                ............--A--C-TT-GG------GCTTT--AT--TATGT-A-G-A----C.........................CCA-A-----TG-----.-----------T-CTCCCTGGGG............CA-CG--AAACC...CAGGAATTT-CC--AGG 1662
MND_2.x.x.5440             ............-------TT--G------G-TCC------T-T-----GGA-C--C.........................CCA-AGAT--TG-----.-----------T-C-CCCTGGGG............CA-TG-CAAGCC......CA----C--C-C.. 1654
MND_2.CM.98.CM16           ............--A--C-TC--G------G-TCT-------AT-----G-A----CC-A.........................-AGAT--TG-----.-----------T-C-CCCTGGGG............CA-TG-CAAACC......CA----C--C-C.. 2097
MND_2.GA.x.M14             ............--A--C-TT--G-----GG-CAC------TATG----G-A-C--CC-A.........................-A-AT--TG-----.-----------TGCTCCCTGGGG............CA-TG-CAAGCC......CA----C----C.. 2021
OLC.CI.97.97CI12           ............ACA--C-AC--G-----G--ACC---C---AT-GCA---A----C-GG.........GTGGGAGGACAACAACAAC----CA---TG-----.-----------ATGGA--GGAGG.........G--A-C-GA---TCCCCTCAACAT-C-ACAGAGGA 2011
GRV.ET.x.GRI_677           ............AAA--C-TT--G--C--C--AATT--C--TATGGCA-----C--C-AG.........................A-TG-------A-----.--------T-TGG-CATTGGGG............A---GC-AA...ACCAAG-AATTT-G-G-AAT 2244
VER.KE.x.TYO1_patent       ............AAA---CTA--G---------ATTG-------T-GCA-----C--C-GG.........................G-----TG-----.--------T-TGGAC--TGGAT..............---GGCAAAACC...GAG-AATTT-CCCGCCG 2304
VER.KE.x.9063              ............AAA--CCTT--------G---CT--------TGGCA-----C--C-GG.........................G-----TG-----.--------T-TGGAC--TGGA-..............---G-CAAAACC...GAG-AACTT-C-CGC-G 2316
VER.KE.x.AGM155            ............AAA---CTT------------CC---G---TT-GCA-----C--C-GG.........................G-----TG-----.--------T-TGG-C--TGGAT..............---G-CAAAACC...AAG-AATTT-CCCGCAG 2299
VER.DE.x.AGM3              ............A-A--C-T---------G--ACC---G--TTT-GCA-------C-GA.........................G-----TA-----.--------T-TGG-C--TGGAT..............----GC-AAACC...CAG-AATTT-CCCGCCG 1811
TAN.UG.x.TAN1              ............AAA--CCTA--------G---CC-------TT-GCA-------CCGC.........................AGTG-------C-----.---------GA--GCCAA-C-G................-----CAAAGCC...GCGCAATTT---AGA.. 2277
SAB.SN.x.SAB1C             ............-----C-----------C-GCCC--AC--T--G-----------CAA.........................-A-----T-------.--------TTTGG-CCCTGGGG............CA--G--AAACC...CAGGAATTTCCC---GA 2439
MND_1.GA.x.MNDGB1          ............--G--C-----T------GCAATG-AT--T--G-AAGCTC-A--CC--..............AAGCCAGCTCAGC--CAGAG--T-------.--------T-TGG-CCTTGGGG............TCCCTCCAAACC......G-------A--C.. 1792
DEB.CM.99.CM40             ............AAA----TT--------GC-------C---ATTGCA---A-----GGA.....................CAGACTCCGAG---A-----.-----------T-CTCCCTGGG-............CTCC---AGACC......CA----C-----AG 1925
DEB.CM.99.CM5              ............AAA--C-TC--------T-----G--C---AT-GCA--GA-C---GGA.....................CAA-TACC-AG---A-----.-----------TGCTCCCTGGG-............CTCCC--AGACC......CA--------T-GG 1919
TAL.CM.01.8023             ...CCTCCGGGGA------TC--G--C--GC--ATG-------GTGCCGC-C-G--C-GA.........................C--CC------A-----.-------------GCT-TTTTAGG............-AAG-TC--CCC......G-TC--GGGA-CGA 1571
TAL.CM.00.266              ...CCGCCAGGCACT--C-AT--G-----C--ACC---C---AT-GCA---C-A--CT-C.........................TCAGC------C-----.-------------TCT-TTTTAGG............-AA--TC--ACC......G-TC-GAGGA-CGA 2076
SUN.GA.98.L14              ...CCACCCGGCTC---C---------------ATG-----TA--CAAGCTC-A--CC-A.....................CAA-GA-A-----TG-----.--------T-TGG-CCCTGGAA............CA-GG----GCC.......------C----CAG 2437
WRC.CI.97.97CI14           GGACAGGGACCCACA--C---GG-------G-A-T------TGT-CA--G-A-C---C-A..................AAAATTAGA-------TA-----.-----------TGG-A-AGGGCA............-ATGCCCT-GAACAATTATCCAGT-CAGAAG- 2368
WRC.CI.98.98CI04           GGACAGGGACCCACT--C---GG-------G-AATT-----TGTGCAA-G-A-C--CC-A..................AAAAGTAGA--G---AT-------.-----------TGGGA-AGGGCA............-ATGCCCT-AAACAATTACCCAGT-CAAAAGG 2386
WRC.GM.05.Pbt_05GM_X02     ...GGACCTCCCAC---C---GG----------ACC--C---ATCCA----A-C--CC-A..................GAAACAAGA--G----TA-----.-----------TGG-C-AGGGCA............-A-GCCCT-CAACAATTATCCAGT......... 1316
LST.CD.88.524              ............------------C-----TGCTA---AC---AG----GCCC-G--CC-AAAACCA.....................A--G-GA---TG-----.--------T-TGGTCCCTGGAG............-AATG-A--TTC......---C--C----CAA 1348
LST.CD.88.447              ............--A--C-----C-----T-GCA-G-A---TAGGT-TGCTC-A---C-AAAGCCT.....................A--G--A---TA-----.-------CT-TGGACCCTGGAG............-A-TG-T--TCC......---A--C----C-T 1354
LST.KE.x.lho7              ............--T--------C-----C-GTA---AG--TAG-T-TGC-C-G--CC-A..................AAACCAA-AG--A---TG-----.------AT-TGGTCCCTGGAG..........--GCAA-CCT--AGG......CAATT--CC---A- 2441
SYK.KE.x.SYK173            ............AAA--C-TT--T---------AAC--------T-GCA-GG-CC----GGCAACCCAAGAGAAATCAAGGCCCTCCA-TTGC-------------.----------AGGA---GGGGT..........-TCCA-G-----CCC.......--CC--C--C-CAG 2235
SYK.KE.x.KE51              ............AAG--C-TT--C------GGAACG------AT-GCA-GGC-G----GACAACCAAGAAAAGGACAAGGCAACCCTCCTCC------C-----.----------AGGG---GGGAA....................CAGGAAACCCCC--CA-------C-G 1872
COL.CM.x.CGU1              .........................................................................A-C-AG-----A-----.-------GCA-C-A--GGGGTGGA................--TACAGACAGCCATCTTCCCC-C.. 1964
```

```
H1B.FR.83.HXB2           AGAGCAGACC......AGAGCCAACAGCCCCACCAGAAGAGAGCTTCAGGTCTGGGGTAGAGACAAC........................................................................................................ 2198
Gag                      Q  S  R  P  .  .  E  P  T  A  P  P  E  E  S  F  R  S  G  V  E  T  T  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Pol                       E  Q  T  .  .  R  A  N  S  P  T  R  R  E  L  Q  V  W  G  R  D  N  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
H1A1.UG.85.U455          ----------.....-------------------C---A-T---TG--ATG----A-A---T............................................................................................................ 1644
H1B.US.90.WEAU160        ---------T.......---A-----------------------------TCA---A---------........................................................................................................ 2197
H1C.ET.86.ETH2220        ----------.....-------------------GAGTCT-AGACCAGAGCCAAC--CCC--C-ACCAGA.................................................................................................... 1575
H1D.CD.84.84ZR085        ----------.....-------------------C---AGCT-CGG-T-TG---AGGAGATA-C-......................................................................................................... 1718
H1F1.BE.93.VI850         ----------.....-------------------C----------G---TCA-A-AG----T---......................................................................................................... 1535
H1G.SE.93.SE6165         ---A---GA-.......---A-------------T-C---A---C--G---TC--A-AG----T-G-....................................................................................................... 1604
H1H.CF.90.056            -------G--.......---A--------------C----------G---TC--A-AG----TG--....................................................................................................... 1548
H1J.SE.93.SE7887         ----------.....-------------------C-------C--G--CTC--A-AG----T............................................................................................................ 1515
H1K.CM.96.MP535          ----------.....-------------------C----------G---TC----AG----T---......................................................................................................... 1394
H101_AE.TH.90.CM240      ----------.....-------------------C---A-A--GGG--ATG----A-----T---......................................................................................................... 1757
H102_AG.NG.x.IBNG        ----------.....G--A--------T-------C---------TG--ATG----A-----T........................................................................................................... 1720
H1N.CM.95.YBF30          ---CA-C-A-AAGGAA------C---------G---CT------T-ATG---T-CA--AG--............................................................................................................ 1771
H1O.BE.87.ANT70          ----ACC-G-.......-C-C---T------------..................................................................................................................................... 2232
H1O.CM.91.MVP5180        ---AACA-GT......GTCC---T------------...................................................................................................................................... 2207
CPZ.CD.90.ANT            ---AGGAGGAAGTAGTG-----------T-----CAT-----T--A.......-CA--AG---CA......................................................................................................... 1637
CPZ.CM.01.SIVcpzCAM13    ---A---GA-.......-------------------C------T-ATG---ACCAA-AGA-............................................................................................................. 1775
CPZ.CM.05.SIVcpzEK505    ---CA-C-A-AAGGAA---------------T---AT------T-ATG---A-CA................................................................................................................... 1746
CPZ.CM.05.SIVcpzMB66     ---AG----------.....---A--------T------AT----------G---TCA---AG--......................................................................................................... 1729
CPZ.CM.98.CAM3           --G----GAA......G--A--C-----T--G--CCT------T-ATG---A-CA--A--CA---GG....................................................................................................... 1603
CPZ.GA.88.GAB2           -A--G-AGGT......G--A--------T-----TCC-----C----G---A--CA-AGC-----CA....................................................................................................... 1582
CPZ.CM.98.CAM5           --G--G-GAA......G--A--C-----T-----CATG-----T-ATG---T-CA--A-ACA----T....................................................................................................... 1906
CPZ.CM.05.SIVcpzMT145    -AG----GAA......G--A--------T-----TATG--ATC----G-A-T-CA................................................................................................................... 1731
CPZ.TZ.00.TAN1           ---A--CC--..........AGAG-A--GGAA.......................................................................................................................................... 1815
CPZ.CM.05.SIVcpzLB7      ----------.....-------C--------G---AT---------TG---T-CA-CC-A-............................................................................................................. 1738
CPZ.US.85.CPZUS          -A-A---GAA......G-----C-----T------AT----GAT---G---A-CAA-A-------GT....................................................................................................... 2264
CPZ.GA.88.GAB1           ---A----A-.........---A-----G--------GAT------T-ATG---ACCA--AG--AGA....................................................................................................... 2262
MON.CM.99.L1             TC-C-TCTGTGCAGCCCAGCG--C-GA-T-TG----C....................................................................................................................................A 1854
MON.NG.x.NG1             TTCCAGT-A-CTCCTTGC--------------G-G---CCA-C.............................................................................................CCARGGGGCYTACAAGGTGAAC  462
GSN.CM.99.CN166          TA-CTTTC-TGCTGCC-ACAG-CC-G--GGA...G--G--AGAACC-CA-CA...................................................................................................................... 1878
GSN.CM.99.CN71           TA-CTTCC-TGCTGCC-TCAG-CC-G--GGA...G--G-T-GAACC-CA-CA...................................................................................................................... 1878
MUS_1.CM.01.1085         ---C-TCT-TCTTGCC-ACTG-TC-GC-AGA--AG--G-G-GAA--GGTCAA-CTCTCTCCAG--GT....................................................................................................... 1891
MUS_1.CM.01.CM1239       ---C-TCG-TTTTACC-AC-G-CC-GC-AGAG-AG--G-GAGA-C-GGTCAA-CTCTCCCC-C--GT....................................................................................................... 1881
MUS_2.CM.01.CM1246       TT-CATCG-TCATGCCTAC-G--C-CC-AGATA---C----CCGG-G-AC-TCTCCCC...ACT-GTAGGGCCCAAGGA.......................................................................................... 1983
MUS_2.CM.01.CM2500       TT-CATCG-TCATGCC-TCAG-TC-CC-AGATA---C-----TAG---AC-TCTCCCC----GTGGGTCCACGGAAGGAGAA....................................................................................... 1948
RCM.NG.x.NG411           CTTC-TTGAGACCTAC..............TG-TCC--G.............................................GAGAGAGGAAGGATGGAATCCCCAGTACGATCCAGCAGAAGAGATGCTCAGGAAATAC 1660
RCM.GA.x.GAB1            CTTC-TTGA-TCCCTCT-CT---GGGATGGAGAGCA-TT-CGA-CC.........................................................................................TGCAGAAGAGATGCTAAAGAATTAT 1645
DEN.CD.x.CD1             -CCCTCAGATGAATCAG--TTTGC--TGGAGC----TGAG-GT...........................................................CCAGCAGCAACCACAGTACCCACAAAATTTTCAAGAGGGGGGA 2035

MAC.US.x.239             T-GCTCA..................--TG-AT-AG-GGCT--TGCCA-CTG--CCCCC.................................................................................AGAGGACCCAGCTGTGGAT 2456
Pol                       H  G  .  .  .  .  .  .  S  S  A  S  G  A  D  A  N  C  S  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  P  R  G  P  S  C
Gag                      M  A  Q  .  .  .  .  .  .  V  H  Q  G  L  M  P  T  A  P  P  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  E  D  P  A  V  D
H2A.GW.x.ALI             TA-C-C-..................--TT--G-AG-G-TTA-CACCA-CAG-ACCCCC..................AGCAGAGCCAGCAGCGGACCTGCTGGAGCAGTATATGCAGCAGGGGAGAAAGCAAAGAGAGCAGAGAGAG 2544
H2A.DE.x.BEN             T--C-CA..................------T-AG-GGCT--TACCA-CAG-ACCTCC..................GGCAGATCCAGCAGCGGAACTGTTGGAGAGATATATGCAGCAAGGGAGAAAGCAGAGGGAGCAGAGGGAG 2551
H2A.SN.x.ST              T-GC-CA.................-AT---G-AG-GGCT--CACCA-CAG-ACCCCC..................GATAGACCCAGTAGAGGACCTACTAGAGAAGTACATGCAGCAAGGGAAAAGGCAGAGAGAGCAGAGAGAG 1996
H2B.GH.86.D205           T--C-CA.................--TG--T-AG-G--T--CACCATCTG-ACCCCC..............GATGAACCCAGCAGAGGGCATGACACCTCGGGGGGCGACACCATCTGCGCCCCCTGCAGATCCAGCAGTGGAG 2527
H2B.CI.x.EHO             TCCAGGC.................-C---A...G-GGATAGTGCCATCTG-GCCCCC..............GATGAACCCAGCATTCGGCATGACACCTCAGGGTGCGATTCCATCTGCACCCCCTGCAGATCCAGCAGAGGAG 2522
H2G.CI.92.ABT96          T--C-CA.................--T---T-AG-GTCT--CTCCATCAG-ACCACC..................................................................................GATGGACCCAGCAGTAGAC 1878
H2U.FR.96.12034          T-GC-CA.................--T---G-AG-G-CT--C-CCA-CAG-CCC.............................TCCAGCAGAACCAGCAGTGGATCTGCTGACCCCGACAGCCCCTCCAGCGGACCCAGCAGTGGAT 2033
MAC.US.x.EMBL_3          T-GCTCA.................--TG-AT-AG-GGCT--CGCCA-CTG--CCCCC..................................................................................AGAGGACCCAGCTGTGGAT 1942
SMM.US.x.H9              T-GC-CA................GATG--T-AG-GGCT--CACCA-CTG--CCYCC..................................................................................AGAGGATCCAGCTGTGGAT 1941
SMM.US.x.PGM53           T-GC-CA................GATG--T-AG-GGCT--CGCCA-CTG--CCCCC..................................................................................AGAGGATCCAGCTGTGGAT 2385
SMM.US.x.SIVsmH635F_L3   T-GC-CA................GATG--T-AG-GGCT--CACCA-CTG--CCCCC..................................................................................AGAGGATCCAGCTGTGGAT 2471
SMM.SL.92.SL92B          T-CAG-C....................-A--TC-TT--CACCGTCAG-ACCCCC..................................................................................AGATCCAGCAGCCCGGATA 1902
STM.US.x.STM             T-GC-CA.................-ATA--T-AG-GGCT--CGCCA-CTG-CCCTCC..................................................................................AGAAGACCCAGCTGCGGAT 2115
MNE.US.x.MNE027          T-GC-CA.................-ATG-AT-AG-GGCT--CACCA-CTG--CCCCC..................................................................................AGAGGACCCAGCTGTGGAT 1936
DRL.x.x.FAO              CCTC........................................................................GCTGACCCCATCAGCTCCTCCATTGCCAGGATACCTCCAGGAGGATCCGGCCGAGAGGATGCTC 1730
MND_2.x.x.5440           ................T-CCATGC--TTGA-C---TC--CACC-CC.......................................................GATGCCAGGGTTGGAAGACCCGGCAGAAAAGATGCTGTTAGAC 1727
MND_2.CM.98.CM16         ................G-C-ATGC--TTGA-C---AC...-GC-CC.................................................TCCAATGCCAGGGATGGAAGACCCAGCAGAGAGAATGCTGTTAGAT 2170
MND_2.GA.x.M14           ................T-CCATGC--CTGA-C---TC--CACC-CC.......................................................GATGCCAGGGATGGAGGATCCAGCAGAGAAGATGTTATTAGAC 2094
OLC.CI.97.97CI12         CAGAG-TGGACATCTTTCCCT-TG--C-TA-C-TG--..................................................................................................................... 2048
GRV.ET.x.GRI_677         -C--AG--GACACAGTT-GT-TGGA-C-AA--G-CCCCCC..........................................................AATGGAAACAGCTTACGATCCAGCAAAGAAGCTCCTC 2321
VER.KE.x.TYO1_patent     CT-CTCTTGGAGCGGA-CC-AGTG-GC-T--T---CCGAGCG--AC--CCC-...........................................................ATACGACCCAGCAAAGAAGCTCCTG 2381
VER.KE.x.9063            CC-CTCATGGAGTGGA-CC-AGCG-GC----T---CCGACACC....................................................................ATACGACCCAGCAAAGAAGCTCCTG 2384
VER.KE.x.AGM155          CC-CTCTTGGGGCGGA-CCAAGTG-GC----T---CCGA-C-A..........................................................CTCTACACCTTACGACCCAGCAAAGAAGCTCCTG 2376
VER.DE.x.AGM3            CT-CTCTTGGGGTGGAGCCAA-TG-GC----T---CCGAGTCC....................................................................ATACGACCCTGCAAAGAAGCTCCTG 1879
TAN.UG.x.TAN1            .-CAAG-CGGAGCCGT......TC-GA-GG-T--CCCGAT-CCTGCGCACGGATTCCC.......................................................CACAGGTTCGCCAGCTGCGGGA 2350
SAB.SN.x.SAB1C           CCTC--TCAGACCAAC--CT--CC--ATGGA--GG--TT-C---AG.........................GCCGGAGGAAAATTGGTATGCAGACAGACCTCCGACCAGAGGACCAGGGCCAGACGATCCAGCAACAGCCCTGTTAAAGCAGTAT 2570
MND_1.GA.x.MNDGB1        ..................................................................................GGCACAAGAGGTGACTCCAACAGCTCCACCATTAGAGGAGAAACCTCTGCAG 1844
DEB.CM.99.CM40           -AG-A-TTGGGAAGTACATC--GGAGAG-A--G----TC--GTA-A-TC-A-...................................................................................................... 1977
DEB.CM.99.CM5            -AG-G-TTGGGAAATACATC---GAGAG-A--G----TC--GTT-ATTCAA-...................................................................................................... 1971
TAL.CM.01.8023           G--A-TTC--AGTGGTGC-A--C--C--A--G--GAT...............................................................................................AGAGCCTCTCACAGCATCA 1627
TAL.CM.00.266            G--ATTTT--AGTGGTGC-A---T-G--A--G---AT...............................................................................................AGAGCCTCTCACAGCATCA 2132
SUN.GA.98.L14            TC-TGCCTT-TGCTCCGCC.........................................................TCTAGAGGATCTGACATTGGGGAACAGAATGACACCACCCCAGTCAAAGGCAGAG 2511
WRC.CI.97.97CI14         T---G--CTTGAACAATTTTGGGC--A-AA--A--ATGCCTTCTGCTCCTC-AAT-AG----GATC-AAACAT....................................................TCAGGAGCTGGAAAAATATTTATTGAAG 2469
WRC.CI.98.98CI04         T---G--CTTGAACAATTTTGGGC-G--....-A--AT-CC-TCTGCTCCTC-AAT-AGG---GATC-AAACAT....................................................ACAAGAATTAGAGAAATACCTCCTGAAA 2484
WRC.GM.05.Pbt_05GM_X02   ..............................T-A-A--AT--ATGCAG--GTGAATACCTTTGGGC-AGTGGT..................................GCCAACAGCACCAGTGGATCCTACCATGCAAGAACTAGAG 1398
LST.CD.88.524            TC-TGG--GGGGC...T-CAGG-GT...GATT--TTC--C-CCACC..........................................................CATGGAGAAAGTGCCTCCAACAAGAGCAGAGAGG 1422
LST.CD.88.447            T--TGG-GGGGAA...T-CAGG-GT...AGT----TC--CTCC-CC...............................................................AATGGAGAGGAGCCCAACACAAGCAGAGAGA 1425
LST.KE.x.lho7            T-G-AG--G-......T-CAGGG-G...AATT---TC--CTCC-CC..................................................................AATGGAGAGTGCTCCAACAAAAGCAGAG 2506
SYK.KE.x.SYK173          T---G--CGAGCCATCT-C....................TCC-CCTCT-GA........................GGACATCGAAGATGGGCCATGGCTCACATGGTCAGCACAGATGAGCCAACAAGCACAAGCGAAGGCACAGAAC 2339
SYK.KE.x.KE51            T--TG--GGAGACACC-ACAG-CC--C-AG-GGAG--.....................TTGGCCATGGCAGCAACAGACATGGGGAAATTATGCTTCACCCCAACAGAAGCACATAGCAATGGCCACCCCGCTACAGCCACAGTCC 1997
COL.CM.x.CGU1            .......................................................................................................................................................................... 1964
```

```
                                                                                                                        Gag end
                                                                          Gag-Pol TF end Pol protease start       Gag p6 end
H1B.FR.83.HXB2           .................................AACTCCCCCTCAGAAGCAGGA...GCCGATAGACAAGGA......ACTGTATCCTTTAACTTCCCTCAGGTCACTCTTTGGCAACGACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTATTA 2324
Gag                       .  .  .  .  .  .  .  .  .  .  .  T  P  P  Q  K  Q  E  .  P  I  D  K  E  .  .  L  Y  P  L  T  S  L  R  S  L  F  G  N  D  P  S  S  Q  *
Pol                      .  .  .  .  .  .  .  .  .  .  .  N  S  P  S  E  A  G  .  A  D  R  Q  G  .  .  T  V  S  F  N  F  P  Q  V  T  L  W  Q  R  P  L  V  T  I  K  I  G  G  Q  L  K  E  A  L  L
H1A1.UG.85.U455          ...................................G--CT-----GC---A------..---T--A-----G---.......--A-AC-------GT--------AA-----------------------------T------G----A-----A--A--G--G-TA----------- 1770
H1B.US.90.WEAU160        ....................................-------T------------..----A--------------.....G-------------------------AA---------------------------------------------A---------G-------------- 2323
H1C.ET.86.ETH2220        ...............GAGCTTCAGGTTCGAGGAAGC---A--TT---C----------..---T--A-----G---...........AG-C-----------------AA-------------------------A--T--T--------A--------A--G--------G-----C--- 1716
H1D.CD.84.84ZR085        ................................CCCT-C-AGAAA---G-A---A-...-GAC-A---------ACT...........---C---G---------AA------------------------G-T--------------------A--G-----A------------ 1846
H1F1.BE.93.VI850         .......................................C---T---C----------..---A--A----GG---ACTGTA...CCC---C---G---------AA------------------------T-A------------A--------A--GA-------G--------- 1661
H1G.SE.93.SE6165         .......................................C---T-C-C----------...-AT--AG--A------..............G-TA-ATC-C-------AA-------------G--------A------G----A--------A--G----TA-----C------ 1721
H1H.CF.90.056            .......................................C---T---C---------GCA--T--AG--------ACC..........---C---G---------A-------------G--------T--T---G----A----A----A--GT---G------------ 1671
H1J.SE.93.SE7887         .......................................C---T-C-C---A------.......AG--------ACT..........-------C-----------AA-------------G--------T-----------GA------------G--G-G------------ 1638
H1K.CM.96.MP535          .......................................C---T---C--G---------...-A-C-A---------ACAGAG...CCC-----------------AA-----------------------A-T------------AG--------A--GT---GA------------ 1520
H101_AE.TH.90.CM240      .......................GGGGGAAGAGAT---CT--TTA-C-----------..---A--A---------ACATCC...TCC-------GT--------AA--------------------------T------G----A-----A--A--G--G--A----------- 1898
H102_AG.NG.x.IBNG        .......................................-C-----T---CAC---------..A----GG-------GGCTATA...TCC-----------------AA------------------------T-A--T---G---GA----A---A--G----TA-----C------ 1849
H1N.CM.95.YBF30          ....................GAAGAGCACACAGGGG-AGGAGATG---G--A-CC-GGA-AG--C---G--CTCTCTGTA...CCCA------------------A------------------------G--A-----G----A------AAAG--G---GA--------T--- 1918
H1O.BE.87.ANT70          ..................GATGGAGGAGGAAGTGAAGGGA-AGGAGA-TC--G-AC-GAAAGG-GGCCCG--C--GCT......C-----G--TG-C-------AA--C--------G-CA----AA-A--T---GC--GAG-T-----C--C---TGT----T-T-GC-G 2379
H1O.CM.91.MVP5180        ..................AATGGAGGAGGCAGTGAAGGAA-AAGAGA-TC--AGTC-GAA-GG-GATC-GG-A--GCT.........---C--A--TG-C-------AA--C--------G-CA----AA-A------GC-----G-T-----T--T---TGT--G---T--C-G 2354
CPZ.CD.90.ANT            .......................................CAAGAGGACTC--A---GTCTCAA-GGG--GG----ACT......ACC---C-CGTA------G-AA--C------------A------AA-GA-GGA-G-TCTC---CA---A---AA-TGTC-G---------- 1763
CPZ.CM.01.SIVcpzCAM13    .......................................GGAGAAGA-CC-AG----ACA-GAAGAG-GG--CAGACTATA...CCC----C--------------AA-------------G-------------A--C--G---GA----AA------T--TGT-----C--G--- 1904
CPZ.CM.05.SIVcpzEK505    ..................GGAGGAGAAGACCACACAGGGGA-AGAGAG-G-AG-AA-GGA-AA--C---G-GCTCTCTGTA...CCCA--C--------------AA-------------G--------A--C-T---G----A-----A--AG--A----A--------TC-- 1896
CPZ.CM.05.SIVcpzMB66     .............................GGAACCTGCCCAGAAGGAGG-AGT-A--AGCGA-GA--C---G--CA-GCTGTA...TCC--AGC------------AA--T---------------------G-T------G---GA----A---A--GT----A-----A--G--- 1870
CPZ.CM.98.CAM3           .......................................AAAGAGA---G--G----GAG-GA--AG--A-----ACCCCTGTATCC---G-----C-------AA-------------G----------C--GG-G----A----AA-----G----TA---------GC-- 1735
CPZ.GA.88.GAB2          ..............................GACAGG--AGAGT-AGA--GG-G----GAGTA-A-C---A-----GAG......CCC--TC-ATC---------AA-------------G----------C--C--G---GAG---CA--A------TGT-----C---C-- 1717
CPZ.CM.98.CAM5           ..................................................-G--A-AAA-AA-GAG--G-----ACCCCTGTATCC--AA-----C-------AA-------------G----------C-TGG-G---AG---AA-----GT---TA---------TC-G 2026
CPZ.CM.05.SIVcpzMT145    ..................................GGAGGAGAAGAAAGTC---G----GGT-AG-GAG--G-GCA-GCCGCC...--C-----------------AA------------------------------A--C--G---A--T-AA--A---T---TT-----CT-GC-- 1866
CPZ.TZ.00.TAN1           ..................GGCTCAGGAGCAGGAGAC-G-A-AGA-A-CAGT-GT-CCAACTG-CCC-CCACT---GATGACGA--A-AGGCGGGTTC-------A----A--------G-----AA-GA-G---G----AG-CCA---A---G-CTGTC-------T--- 1968
CPZ.CM.05.SIVcpzLB7      ............................AACCACCCAG-AA-AGGAGG-C-G-G-A--GGACAA--C---G--CA-GCTGTA...CCC---C-------------AA--T------------------------A-T------------AG-C-----CT-CT-G--A-G---A--G--- 1879
CPZ.US.85.CPZUS          ..........................................GA-A---G--A-AC-GGGAAA-GA-A-GG-...GCC......C-T--AGC------------AA-------------G----------C---G------AG---CA-----GA--GTA---------T--- 2384
CPZ.GA.88.GAB1          ..........................................GAAGAGC---G--A-...GAAAGA-GG---G-GCAGTCT......A-----ACC---C-------AA-------------G----------A--C--G----AG---AA---------TGT------T-GC-- 2385
MON.CM.99.L1             GGGACTGCGGAGCCAGTCAATCTCAACCAAGAGATTGG-T----.........-A-GAC-GG--CCCCA-----GACGAGG-G-G---T--ATC-C---T-AGCA--T--------GG-A----AA-AAGAT-TG-GCT---------A---AAGGT-TCCATG-----G 2015
MON.NG.x.NG1             TTCTCACCTCCAGTGGGACCACCAGAGACCCC.....................A-CGCC-A-AGCCCCG-----GACGAGGAA-G---T--ATC-C---T-AGCA--C--------GG------AG--AGA---G--G-A--C-CT-----GTCCGT-TCCATGT----G  611
GSN.CM.99.CN166          ....................GCTCAATCTCTCGAC-C-AGTAGGGG-ACCC---AGCAA-GGACAGA-GG----GAGCAGGAACAC--TC-ATC-C-----GTCA--T--------GG-A----AA-GAAGGA-G-CTCC---------A---G-GGT-TCCATGT------ 2028
GSN.CM.99.CN71           ....................GCTCAACCTGTCTCCGC-GGTGGGGG--CCC---A-CAA-GA-CAGA-GGCA--AAAGAGGAACAC--T--ATC-C-----GTCA--T--------GG-A----AA-GAGAGA-G-CTCC--G------A---G-GGT-TCCATGT------ 2028
MUS_1.CM.01.1085         ...........................................--GTCCCGG-GAG-GG--AC-GGACAGGAA-AG--TC-ATC-------GTCA--C------CAGG------AG--AGAC--G--G-------AA-----GG-GGT-TC-ATG--GC-C 2008
MUS_1.CM.01.CM1239       ...........................................--GTCCCAA-GAT-GAGC-AGCGACAGGAA-AG--T--ATC-------GTCA--T------CAGG-------A-TAGAGA-G---T-----AA--A---G-GGT-C-CATG--GC-C 1998
MUS_2.CM.01.CM1246       ...........................................A--GCCA--C-AG----G-A-AGACAG-AA-AC--TC-ATC-------GTCA--T--------GG-------AA--AAA---G--G-C--TAAT-----G-C-GT-TCCATG--CC-G 2100
MUS_2.CM.01.CM2500       ...........................................CC-AACC--CCA--C-CCACCGGACAGGAA-A---T--ATC-------GTCA--C--------GG--------A--AGA---G---T-G-CCAT--T--G-CGGT-TCCATG--CC-G 2065
RCM.NG.x.NG411           CTAGCATTGGGGAGACAGCACAAACAGGAGCAGAGG-AAGA............-A-CAA-GA--A--TGGG-AGAGCCTACGA-G--G-G--GT-C-------AT--T--------A-G-----TG-AACATGGG--G-A----A--------AAGGT---G--CT------ 1818
RCM.GA.x.GAB1            CTGAGGAGGGCAGGGGAACAAAAGAGACAACAGAGGCAGGA............A--GAGCAA--AGAGAG---GAGCATAT-A-G-AG-C-----C-------AT--G--------A-G-----AACTACA-T-GCTC-A----AA-----GAA-GT---G-TC------- 1803
DEN.CD.x.CD1             AACAGACAATCTTGTCAGGAAGACAAGAGGCAGAAC-AG-AGT-.........CC-TCA-GA-GAC--GG--ACAACGAGGGA-AG--T--ATC----GT-G-A---T--------AG------AA-A-GGCAG--CTG-G-TCA---A---AGGGTA--G-TC--TA-- 2196
                                                                                                                Gag end
                                                                                        Pol Protease start   Gag p6 end

MAC.US.x.239             CTGCTAAAGAACTACATGCAGTTGGGCAAGCAGCAG-GAGAAAAG----GAG-AAGCAGAGA--AGCCTT-CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T--------AGGA----AG-A-----TGCTC-T--T-AA--A--G-CTGTA----TAT--C-G 2626
Pol                      G  S  A  K  E  L  H  A  V  G  Q  A  A  E  R  K  A  E  R  K  Q  R  E  A  L  Q  G  G  D  R  G  F  A  A  P  Q  F  S  L  W  R  R  P  V  V  T  A  H  I  E  G  Q  P  V  E  V  L
Gag                       L  L  K  N  Y  M  Q  L  G  K  Q  Q  R  E  K  Q  R  E  S  R  E  K  P  Y  K  E  V  T  E  D  L  L  H  L  N  S  L  F  G  G  D  Q  *
H2A.GW.x.ALI             AGGCCATACAAAGAGGTGACGGAGGACTTACTGCACCTCGA............-C-GGGAGA--C-CCAC-CA-AGAGGTGACAG-GGAC--GCTGCA-----AT--T--------A--A----AG-A------GCCT-C--T-A---C--G-C-GT-----T-T--C-- 2702
H2A.DE.x.BEN             AGACCATACAAAGAGGTGACGGAGGACTTGCTGCACCTCGA............-C-GAGAGA--C-CCTC-CAGAGAGGAGACAG-GGAC--GCTGCA-----AT--T--------A--A----AG-A------GCGT-C--C-A--AT--G-CGGTA----TCT--C-- 2709
H2A.SN.x.ST              AGGCCATACAAAGAAGTGACAGAGGACTTCCTGCAGCTCGA............-A-ACAAGA--C-CCATGCAGAGAGACGACAG-GGAC--GCTGCA-----AT--T--------A--A----AG-A------GC-C-TG-T-A---C--G-C-GTA----T-T-GC-- 2154
H2B.GH.86.D205           ATGCTGAAAAGTTACATGCAGATGGGGAGACAACAG-GAGA............-AGCCGAGA--G-CC-T-CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T--------AG-A----AG-A----A-GC-TGT--C-A---T--GTC-GTA----TAT--C-- 2685
H2B.CI.x.EHO             ATGCTAAAGAACTACATGCAACTAGGGAAGAAGCAG-AGGA............-A-CAGAGA--G-CC-T-CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T--------AG-A----AG-A----A-GC--CT--T-A---T---TC-GTA----TAT--C-- 2680
H2G.CI.92.ABT96          CTACTGAAGAATTACATGCARCTGGGGAGGAAGCA--AGGA............-C-GAGAAAC-A-CC-T-CA-GGAGGTGACAG-RGAH--GCTGCA------C--T--------AG------AG-A-----TGC-T-C----AAA-T--G-C-GTA----TA------ 2036
H2U.FR.96.12034          CTGCTAAAGAGCTACATGCAACAGGGCAAGAAACA--AGGA............-A-CAGAGA--G-CCAT-CA-GGAGGTGAC-G-GGA---GCTGCA-----AT--T--------A-------AA-A-----TGC-T-C----A---------C-GTAA---TA---C-G 2191
MAC.US.x.EMBL_3          CTGCTAAAGAACTACATGCAGTTGGGCAAGCAGCAG-GAGA............AAGCAGAGA--AGCCTT-CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T--------AGGA----AG-A-----TGCTC-T--T-AA--A--G-CTGTA----TAT----G 2100
SMM.US.x.H9              CTACTGAAGAATTACATGAAGGTGGGCAGAAGGCAG-GGGA............-A-CAGAGA--G-CCTT-CA-GGAGGT-AC-G-GGA---GCTGCA-----AT--T--------AG-A----AG-A-----TGCCT-C--T-AA-AA--G-CC-TA----TAT----- 2099
SMM.US.x.PGM53           CTACTGAGGAATTACATGAAGATGGGCAGAAAGCAG-GGGA............-A-CAGAGA--G-CCTT-CA-GGAGGTGAC-G-GGA---GCTGCA-----AT--T--------AGGA----AG-A---G-TGCCT-C--T-AA-AA--G-CCGTA----TAT----- 2543
SMM.US.x.SIVsmH635F_L3   CTACTGAAGAATTACATGAAGATGGGCAGAAAGCAG-GGGA............-A-CAGAGA--G-CCTT-CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T--------AG-A----AG-A-----TGCCT-C--T-AA-AA--G-CCGTA----TAT----- 2629
SMM.SL.92.SL92B          GTCAA...GGAGTATCTAGAGAAAGCACAGAGGGA--AGA-............AAGGAG-AGC-GGCCTT-CA-GGAGGTGAC-G-GGA---GCTGCA-----AT--T-----C--AG-A----AG-A----G-GC-CTA---A---T--G-CCCTA--G-T----C-G 2057
STM.US.x.STM             CTGCTGAGAAGTTACATGCAGCTGGGCAAGAAGCAG-GAGA............-AGCAG-AA--C-CC-T-CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T--------AG-A----AG-A------GCCC-T--T-A---T--G-CTGTA----TA-----G 2273
MNE.US.x.MNE027          CTGCTAAAAAACTACATGCAGTTGGGCAAACAGCAG-GAGA............AAGCAAAAG--AGCCTT-CA-GGAGGTGGCAG-GGA---GCTGCA-----AT--T--------AG-A----AG-A-----TGCTC-T--T-A---A--G-CTGCA----TA-----G 2094
DRL.x.x.FAO              CAAAAGTACATGGAGCAGGGAGCTCAGCAGAAGAG-CAA-AGAG.........AC-ACAA-A-CAGA-G-G--GCCCCTACGA-G-GG---AC-AC--------T--T-----C---TCA----AACTACA-T-G--G--G---AA------AAGGTA--G---T--C-- 1891
MND_2.x.x.5440           TATATGAAGAAAGGGCAACAACAGAGGGCAGCAGC-GGAG-............CC-GAAAGA-GA-A-G----GTCCCTACGA-GCAG---AC-AC--------T--T--------A-CA----AACTACAGT-G--G-A----A-------AAGGTA--G--A--TC-- 1885
MND_2.CM.98.CM16         TACATGAAGAAAGGGCAACAGCAGAAAGCGGAAAGC-A...............AC-GGAAAA--A---G-G--GTCCATACGA-GCAG---AC-AC--------T--T--------A-CA----AACTACAGT-G--G------AA------AA-GT---G--CT--C-G 2325
MND_2.GA.x.M14           TATATGAGAAAGGGGCAACAACAGAGGGCAGCAAAGGAGA-............CA-GCA-GA--AG-------GTCCCTACGA-GCAG---AC-AC--------T--T--------A-CA----AACTACAGT-G--G-A----A---A---AAGGTA-----C---C-G 2252
OLC.CI.97.97CI12         ............ACAGATTTTGGAGAAGGGAACAAG..................---AAA-AGAGAGATG-----GGAGAAGAAACC--T--ATC-C-----AGCA--T--------AG-A----AA-GA-AGAGC-GT-T----AA--A--G-A-GTCACT--A--G--- 2189
GRV.ET.x.GRI_677         CAGCAGTATGCAGAGAAGGGACAGCGCCTGAGAGAGGAGAG............A--ACA-A-A-GGA-AC--A-GGAGAA-GAAGTGGAGGATGT----T-G--C--C--------AGGA----AA-GAAACG-G-C-TC----AA--AACG-C-GT-C----CT-G--- 2479
VER.KE.x.TYO1_patent     CAGCAATATGCAGAGAAAGGGAAACAACTGAGGGAGCAAAAGAG.........-A-TCCA---GC-ATG--TCCGGATTGGACCG-GGGA-ATT---T...G-AC--C--------AG-A----AA-AAAG---G-GT-T----AA---GTC-CC-TTA-G--A--GC-- 2539
VER.KE.x.9063            CAGCAGTATGCAGAGAAGGGAAAACAGTTGAGGGA-CAAGGGAA.........-AGAACA--ACC-ACA--TCCAGACTGGGC-G-AGGA-ATT---T...G-AC--C--------AG-A----AA-AAAA---G-C-TT--T-AA--AGTC-CTGTAA---TC--G--- 2542
VER.KE.x.AGM155          CAGCAGTATGCAGAGAAAGGGAAACAAATGAGAAATCAGAA-AG.........AA-CCCC--AGCGA----TCCAGATTGGAACG-GGGA-ATT---T...G-AC--C--------AG-A----AA-AAGG--CTGT-TA-----A--AACTGCCGTTA-G--AT----- 2534
VER.DE.x.AGM3            CAGCAATATGCAGACAAGGGGAAGCAGTTGAGGGA-CAAAGGAA.........AA-ACCA--AGC--TG--TCCCGATTGGACAG-GGGA-ATT---T...G-AC--C--------AG-A----AA-AAAA---G-TT-C----AA---GTC-CC-TCAG---AT----- 2037
TAN.UG.x.TAN1            GCATACGACCCAGCCAGGAAACTGCTGGAGCAGTATG-CAAAAAGGG......A--TCAATT--G-A-GC--A-GGAGAAGGAACTGGAGGACTA----T-G--T--C--------AGGA----AA-GAGT-AG--CT-C---CA---C--G-C-GTAC---T---G--- 2514
SAB.SN.x.SAB1C           GCTGTTCAGGGGAAACGGCAGAAACAGCAGTGGCA..................AA-CCACT--CC-C-AC--AGCCCCTACGA-G-GG---AC-GC-----------T--------AG-A----AA-GAAA--TG-GT-C----A---A--GAA-GTCACT---------- 2722
MND_1.GA.x.MNDGB1        AAAACTCTGAGCACTTATCAGAAATTAGGGAGAGGGCT...............CAGGCA-AA---GA-GG----GAAGAG-GA-G-GGA---TCA------G--TA-T------CAAG-A----AA-A-AAGAGG-CTCAG-G-AT--TGTCACT-TAAG-------C-- 1999
DEB.CM.99.CM40           ............TACCAGGGTACCAGTACAGGGGCAGGAGAA-AAC-CCC--AG---GGAAA-A-CGAGG-GA-GGAAGTA...CCCGGA-----TGGGG--G-AA--T--------AG-A----AA-GAC----G--C-C----AA--AAG-AA-GTCC-G-TG---A-- 2133
DEB.CM.99.CM5            ............AACCAGGGTACCTGTGCAAGGGGA--GGAATG---CCC--G----CACAA---CCAGG-GA-GGAAGTA...CCCAGA-----TGGGG--G-AA--T--------AG-A----AA-GAC-G--G--T-TG-G-AA--AA--AGGGTCC-G-TG---A-- 2127
TAL.CM.01.8023           TGGCAGGAGGGGACTACCACATGGACCCCTCCGACC-AGGA............T-CCGCAG-ACCCA-GC-AC-GACGAGGGA-AG--T--ATC-C--GT-AGCCAGT--------AG------AA-A--A-AGG-GCTC----AA--AAGGGAGGT---G-TA--CC-- 1785
TAL.CM.00.266            TGGCAGGAGGGGACTACCACGTGGACCCCACCAGCC-AGAA...............GCCC-AAGA-ACA-CCACAACAAGGGA-AG--TC-ATC-C--TT-GGCCAGT--------AG------AA-A----A-G-GCT-----AA--A---GAGGT---G-TA--CC-- 2287
SUN.GA.98.L14            AGAGCTCTGGAGACCTACAGACTGTTAGGACAGGGCCTCAG............A-CACA--A--AGAGG-----AAGGGGGGAA-G--AGGAGC-A-G-T-G-ATATG------CCAG-A-----T-A-AG-AGG-CTTC--TAAT-------CTGT-ACT---T----G 2669
WRC.CI.97.97CI14         GGAGCACAATTGAGGGAACAACAGGAACACAAGAA..................-ACAGA-GGACAGA-G-G---AGTGGGGG--CCGGAA-GGC-C-------AA--T-----C-AACCA----AA-GA-ACA-TGC-TA----AA--A--TAGTGTCTC---AT--C-- 2621
WRC.CI.98.98CI04         GGAGCTCAATTGAAGGAGCAGGCACAGGCAAA.....................AA-GGAAGA--AGA-G-GA--AACGGGG---CCAGAG-GGC---------AA--T-----C-AA-CA----AA-GA-ACA-TGT-TA----AA--AA-TAGTGTCTC---CT-G--- 2633
WRC.GM.05.Pbt_05GM_X02   CAATATCTTCTGAGAGGAGCTCAGATGAAGAGGGA..................-A-AGA-AAAGA-ATGC--A-GGAGGGGAAAGTAGAA-ATC-C-------AA--T-----C-AACCA----AA-GA-ACA-TGT-TC--T--TC-AG-GAGTGTCTCT--GT----- 1550
LST.CD.88.524            GCAATAGAGACATACAGGAACCTAGGACAACAACTG-GGAA...............GCAA-AACA--TTCCAC-GAAGTGTG--G-CGAGCCTTGCCTGAATTTC-TCT-----CCAG------AA-A-AG-A-G--CT-G-CAAT------GA-GTAACT--C-----G 1577
LST.CD.88.447            GCCTTAGAAACATACAGGAACCTCGGACAACAGCTG-GGAG...............GCAA-A-CAGAT-CCAC-GAAATGTG--G-CGAGCC-TGCCTGAGTTTC-TCT-----CCAG------AA-A-AA-AC--TCT-G-CAAT-----G-C-GT-TCT-----TC-T 1580
LST.KE.x.lho7            AGAGCACTGGAGACTTACAGGACTTTAGGACAACAGCTGAAAAG.........-C-GCA--A-CA--T-CCAC-GAAATGTG--G-CGAGCCCTGCCTGAATTCACTCT-...-CCAG------AA-A-AA-GT---TT-G-CAAT--------C-GT-TC---A--G--G 2664
SYK.KE.x.SYK173          TCTCCCTCGAAGAAACCCCCCACCAACAGGGAAGTGCTCT-G--.........-A-GGA-AGC-GT-GG-----GGAGACCAA-AG--TC-ATC-C--T-----C--T--------AG-A----AA-GA-AGA-G--G-TG---AA--TG-C--GGT-C--ATGT--G-- 2500
SYK.KE.x.KE51            TCTCCAAAGACAGTCAAGGACCTCTTGGTCCCTGGG-AA--TT-.........A--GAA-AAAGA---G-T-A-GGAGAAGGGACC--T--ATC-C-----TCAA--T--------AG------AA-GA-GGA------TG---AA---ACCAAGGTAAC-ATGT-GC-G 2158
COL.CM.x.CGU1            ................................GAAGATG-GCAAAGA.........CCTGCCA--C-GGAGGG--A-GGGGGAGAG...--TC-ATC-C---T-G-A---T-----C--AG------AA-A--T-T-G--G-AG---AA--A---AAGTGCC----G--G--- 2093
```

```
H1B.FR.83.HXB2          GATACAGGAGCAGATGATACAGTATTAGAAGAAATGAGTTTG............CCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGTGGACATAAAGCTATA.............. 2468
Pol                      D  T  G  A  D  D  T  V  L  E  E  M  S  L  .  .  .  .  P  G  R  W  K  P  K  M  I  G  G  I  G  G  F  I  K  V  R  Q  Y  D  Q  I  L  I  E  I  C  G  H  K  A  I  .  .  .  .  .
H1A1.UG.85.U455         ----------------------C--------C--A-A---............-------A--------------A----------------------------------------------------T--------T-----A-A--GA-----.............. 1914
H1B.US.90.WEAU160       -------------------------------C----A---............-----G-A------------------------------------------------G---C-----------------------------.............. 2467
H1C.ET.86.ETH2220       --C----------------------------------A-A---............-------A-------------------A--------------T------------------A---A---------T-----A-A--G------.............. 1860
H1D.CD.84.84ZR085       -------------------------------C--A-A---............-------AG---------------------------------------------------C-------------------------G------.............. 1990
H1F1.BE.93.VI850        -------------------------------C--A-A---............-------A--------------------------------------------A---------A-C-------------T-------C--G------.............. 1805
H1G.SE.93.SE6165        --C-------------C------------------A-A--A...........--------------------------------------------------A---------AG---CT--------TA-----A-A--G------.............. 1865
H1H.CF.90.056           -------------------------------G--A-A---..........--G---A---------------------------------------------------G--AG--GC-----------------A-A--G------.............. 1815
H1J.SE.93.SE7887        -------------------------------C--AGAC---..........---C---A-------------------------------------------------A-CG--G---CG--------TGAG---A-A--G------.............. 1782
H1K.CM.96.MP535         -----------------------------------A-A---............-------A--------------------------------------------A---------AG----T--------T--------A--G------.............. 1664
H101_AE.TH.90.CM240     -------------------------------T--A-A---............-------A-------------------------------------G----AG--A-----------------T----------------A-A--G------.............. 2042
H102_AG.NG.x.IBNG       --C----------------------------C--A-A--A............-------A--------------------------------------------A---------A--------A-----T--------T-----A-A--G--C---.............. 1993
H1N.CM.95.YBF30         -----------T-----------A-------GC-ACAA--A...........GAG----A-------------------A----------A----------G-----A------A-T---ACAG----C--ACAG---AGA-----AG-T.............. 2062
H1O.BE.87.ANT70         -----------------C------C--A-CA-C--ACAA---...........GA-----A-----------------------T--A-----------A-------A-G-A------A-TG-GACAG-------AGAA---AGGG-G-TACAG.............. 2523
H1O.CM.91.MVP5180       --------G-----------------A-TA-C--ACAA--A...........GA----------C----------------T--A-----C-----A-------A-G-----A-CA-TG-GACAG-----G-ACAA---A-GG---TACAG.............. 2498
CPZ.CD.90.ANT           -----------T-----C------G---G-G---TCA----...........-A-----AC-----G-------CA--T--T--------G--------TTCC---CA--------A-CA-AG---C----C----TG-A-ACAGG-C--TAC--.............. 1907
CPZ.CM.01.SIVcpzCAM13   --C-----G--T-----------A--C----C--TCAA--A...........-AG----TT--------------------------------C-----T-----G-A---A------A-TG--TCA-----G--AGAG--GA-A----TGCA-.............. 2048
CPZ.CM.05.SIVcpzEK505   -----------T-----------A-----------ACAG---...........GA-----A--------------------------------G---------------A---------A-TG-GA-A-----G--ACAA--GA-A-----AG-T.............. 2040
CPZ.CM.05.SIVcpzMB66    -----------------C-----GC-----------AGAA---...........GG---T------------------------------------C--A-----------A---------A-TG-TACA-----C--A------A-A-G---AG--.............. 2014
CPZ.CM.98.CAM3          -----------T-----------A----TA-TG-ACAAC-A...........A------A-----G------------------------------T-A---A-----CA-C---GCA--------TGAG---AGA---A-A-C-.............. 1879
CPZ.GA.88.GAB2          --------G--T-----------------TA-T--ACAAC-A...........GA---G-A-----G-------T-A-----A--C---------A-----T-------A---A-------A-TG--GAA--------AGAG---A-A-G--TAGA-.............. 1861
CPZ.CM.98.CAM5          --------G--T-----C------A----TA-TG-ACAA---...........A------AG----G-----G-----------------------------T-----T-A---A-------A-C---G-A--------AGAG--CAGA---A-A-C-.............. 2170
CPZ.CM.05.SIVcpzMT145   --C-----G--T-----C------A-T------CT-A-A---...........-------A-----GT-------------------------------T--------A---A-T---AAGAG--AA----------AGAA---A-A--GATA-CC.............. 2010
CPZ.TZ.00.TAN1          -----T-----------C-GT--T--TTGTA-C--C-AA--A...........AAG---CAG----C---------CC-----A-----A-----A---G-ACC---T--TG----CT--A-T--T-CAG--C----TG-CAATA-AG---TC-G-.............. 2112
CPZ.CM.05.SIVcpzLB7     --------G-------------GC-------TT-AGAA--A...........GGG-----G------------G--------A------------------------------AG--A------A-T----CA--------AG-A---A-A--G--AC--.............. 2023
CPZ.US.85.CPZUS         --C--------T-----------C----CA-C--ACAAA-T...........GA---G-C-----G---------------G-----------------------------------A--------A---A--------CG-CAAT--------AGAG---AGA-----ACAG.............. 2528
CPZ.GA.88.GAB1          -----------T-----------A----GAG---ACAA--A...........-A---CTT-------------------------A---------------------------C-A---A-T----A-TG---A------G--AGAA--GAGA----TAG--.............. 2529
MON.CM.99.L1            --------T--C--C--C---A-TA--C----TCA-CACG-A.....AATTTAGGG-AG-AT----C---T---G-AG----------A--G--AA-G--AG----T-A---A--CA-AGCA----AAG-GAT-T-TGAA-ACAGAG-G-T-TGG.............. 2165
MON.NG.x.NG1            --C--G--R--G--C-----CA-CG-CC----RCAAGACA-A.....AATCTAGG--A--AT----C---C------G----------A-----C--A--AC-G--T-AG--A-TCA-AGG-G--ACGG--TCCT--CAG--TA-A-C--TG-G-.............. 761
GSN.CM.99.CN166         --C-----G--G--C--C--CA-CG-CC----CTCAGCAA-T.....GAATTAGATAC-CC-----C----------G----------A--G-----A--AC-------GG--C--A-A--TG--GAAG--ATCT-TAA--AGA-A-G-ATC-A-.............. 2178
GSN.CM.99.CN71          --C-----G--G--C-----CA-TG-CC-----TCAGCAA-T.....GAATTAGACAC-CC-----C---------C-G-------G--A--G-----A--AC------AGG-AC--AGA--TG--GAAG-GATCT-TAA--A-A-A-G-ATC-A-.............. 2178
MUS_1.CM.01.1085        --C--C-----------C--CA--A-CC-G--TGCCCA-A-A.....CACCTAGA----CC------T--C---TCAG----------A--G--AAAC--ATC---C-A-G-A---AGGGGAG---AGG-TTGTT--AA----A-A-CTAT--GG.............. 2158
MUS_1.CM.01.CM1239       --C-----G--------C--CA--G-CC-G--TTCC-ACA-T.....CAGATAGATAA-CC------T--C---T--G-G---------C--G--AAAC--ATC---C---G-A---AGAGG-G---AAG--TCTT--AA--AGA-A-C-AT--AG.............. 2148
MUS_2.CM.01.CM1246       --C--C-----G-----C--CA-CA-CC----CTCAGA-A-A.....AAAATTAGT--GGA--T--CT--C---G-AG-G--A--TT-A--G--GC-G-----T----A---A---A-GGG-G-GA-AG--A--T--AA--AGA-A-G-ATAG--.............. 2250
MUS_2.CM.01.CM2500       --C-----G--------C--CA-TA-TC-G--GTCTGAAA-A.....GAAATAGATAA-GA-----CT------G--G---------T-A--G--A--A--T--T----AG--A---A-AGCAG--AAAG----GT-TAA--ACA-A-G-ATC---.............. 2215
RCM.NG.x.NG411          ----------------G------TA--C--A-TT-AGA---A...........GG--A--AT-----G---------T-----A-----A--G--A-A---A--T--T-A---A-T-A-C--TTGTAAA---ACT--AGCA--GA-G-CTA-ACAT.............. 1962
RCM.GA.x.GAB1           --C--------------C------A-T----G---AGAA--A...........GG-AATGAT----CC---------A-----A-----A--G--A-A---T--T----A---A---A-AA-CTGTGAA--T-----AGC----A-A-GGA--CAT.............. 1947
DEN.CD.x.CD1            --------G--------C--CA-TA---C----CAAGA-A-A...AATTTGGGAA---AGG------TC-------CAG----------A-----A-A---T--T--TCAGGCA--CCCAGG----GAA--TA-GC-TATG-ACA-AGT---GG--.............. 2349

MAC.US.x.239            --------G--T------T-TA-TG--AC--G---AGAG--A...........GGTCC-CAT-AT-CC--------AG----A-----A-----------T--TACT-A-G-A--CA-AA-TG--GAA------G-T-TA--CA-A-GGAT--A-.............. 2770
Pol                     L  D  T  G  A  D  D  S  I  V  T  G  I  E  .  .  .  .  L  G  P  H  Y  T  P  K  I  V  G  G  I  G  G  F  I  N  T  K  E  Y  K  N  V  E  I  E  V  L  G  K  R  I  .  .  .  .  .
H2A.GW.x.ALI            --C-----G--T--C--CT--A--G---C--G---AGAG--A...........GGGA-C-AC-AT-CC--------AG----------A--G--A--C--A--TACC-A-G-A-----AG-TG--GAA---A--G-ACTAAATA-A-G--TA-A-.............. 2846
H2A.DE.x.BEN            --C-----G--T-----CT--A--G---C--G---AGAA--A...........GGG-AC-AT-AC-CT--------AG----------A--G--A-----A--CACC-A-G-A--CA-AA-TG--GAA---A--G-ACTAAATA-A-G--TA-G-.............. 2853
H2A.SN.x.ST             --C-----G--T--C--CT--A--G---C--GCG-AGAG--A...........GGGA-C-AT-AT-GT-----G--AG----------A--G--A--C--A--TACC-A-G-A---A-AA-TG--GAA---AG-G-A-TAAATA-A-G--TA-G-.............. 2298
H2B.GH.86.D205          --C-------TT--C--CT--A--G---C--GG--AGAA--A...........GGTA-C-AT-AC-CC--------AG----A--G--A-----G--C--A--TACC-A-G-A--CA-AG-TG--GAA------G-AGTG---A-A-G--TA-GG.............. 2829
H2B.CI.x.EHO            --C--------T-----CT--A--G---C--GG--AGAA--A...........GGCA-C-AT-AC-CC--------AG----T--G--A-----A-----A--TACC-ATG-A--CA-AA-TG--GAA------G-AGTA---A-A-G--TA-G-.............. 2824
H2G.CI.92.ABT96         --C-----G--T-----CT-TA--G---C--G---AGAA--A...........GG-A-T-AT-AT-GT--------AG----A-----A-----A-----T--TACC-A-G-A--CA-AA-TG--GAA--TA--G-GGTA---A-A-G-ATACAG.............. 2180
H2U.FR.96.12034         A----C--G--T------T--A--G---C--GG--T-AA--A...........GA-CC-TAT-AT-CT-------CTG------A---A-A---G--C--T--TAC--A-A-A-T-C-CA-TG--AA----A--G-A-TG--CA-A-C-AT--G-.............. 2335
MAC.US.x.EMBL_3         --------G--T------T-TA-TG--AC--G---AGAG--A...........GGTCC-CAT-AT-CC--------AG----A-----A-----------T--TACT-A-G-A--CA-AA-TG--GAA------G-T-TA--CA-A-GGAT--A-.............. 2244
SMM.US.x.H9             --------G--T------T--A-TG---C--GG--AGAA---...........GGTCC--AT-AT-CC--T-----AG----A-----A-----C--C--T--CACC-A-G-A---A-AG-TG--AAA---A--G-A-TA--CA-AGT-AT--AG.............. 2243
SMM.US.x.PGM53          --------G--T------T--A-TG-G-C--GG--AGAAC-A...........GGTCC--AT-AT-CC--T-----AG----A-----A-----C--C--T--CACC-A-G-----A-AG-TG--AAA---A--G---TA--CA-GGT-AT--AG.............. 2687
SMM.US.x.SIVsmH635F_L3  --------G--T------T--A-TG--AC--GG--AGAA---...........GGTCC--AT-AT-CC--T-----AG----A-----A-----C--C--T--TACC-A-G-A---A-AG-TG-GAAA---A--G---TA--CA-GGT-AT--AG.............. 2773
SMM.SL.92.SL92B         --C-----G-----C--CT-CA--G---C--GG--AGAG--A...........GGTCC-G-G-AC-C-----G---TG-------G--------A--C--A--TAC--A-G-A---AGAACAG--AAA------G-T-TA---A-GGT-ATA-G-.............. 2201
STM.US.x.STM            --C-----G--T--C--CT--A--G-G-CG-GG--AGAGC-A...........GGGCT-CA--AC-CC-----GG-AG----T-----A-----A-----A--TACT-AGG-A-T-A-GA-TG--AA------GG-A-TA--CA-G---ATA-A-.............. 2417
MNE.US.x.MNE027         --------G--T------T-TA-TG---C--G---AGAG--A...........GGTCC-CAT-AT-CC--------AG----A-----A-----------T--TACT-A-G-A--CA-AA-TG--GAA------G-T-TA--CA-A-GGAT--A-.............. 2238
DRL.x.x.FAO             --C--------------------A-TC---G-T--GAA---A...........A------A------G---C-G--T-----A------------AAC---T--T--G-----A---TT-A-TTGTGAA--TA-G--AGCA--GA-A-TTA--CAT.............. 2035
MND_2.x.x.5440          -----T--------C----------A-TA----TT-AGAG--A...........A------AT-----G---CG---T-----A------------AGCA----G-------------TT-A-TTGTAAAG-------AGCA---A-A-TTA--CAT.............. 2029
MND_2.CM.98.CM16        --C--------T------------A-CA----TC-AGA---A...........AA---T-AT---------C-G--T--T--A------------A-CA-----T----A-----T-TTCA-TTGTAAAG--AC---AGCA--CA-A-CTA-ACAC.............. 2469
MND_2.GA.x.M14          --C--T------------------A-CA----TG-AGA---A...........A------AT--------CC----T-----A------------AGCA--T-GG--G-A-------TTCA-TTGTAAA---AC-G-AGCA---A-A-GTA--CAT.............. 2396
OLC.CI.97.97CI12        --------G--------C------A-TA-G-GG-A-GA-A-C.....TACATGGA---T-ATG---TT--T-G---T-----------A--G--AGC---T-GG----TG--A---AGATGT-GGA-A-G-ATTT-T...-AGA-AG--ATACA-.............. 2336
GRV.ET.x.GRI_677        -----------------C--TA--A-TC------A-GAC---CACTTTCCCCCA-ATAA-CC----CGTT-C--GG-AG----A--T--A----AGGG--TC-T--C-A-G-A---C-GGG-G---AAG--C--T-GGAG-ATA-A-TCATC-CC.............. 2635
VER.KE.x.TYO1_patent    --C-----G--------C--CA--A-TA------ATGA---A.....CAATTAT----TCC-----G---C-----T--------C--A----AGGCC-T--T---A-G-A---A-CG-C-GGGAAG--A----AGAA-ATA-A-TTTTG-G-.............. 2689
VER.KE.x.9063           --------G--------C--CA-CA--A------CAGACA-A.....CAACTAA-----CCT----G-------G-A-----A-----------AGGAC-T--T---A-G-----AG-A-TG--GAAG--A----AGAA-ATA-A-TTTTG-G-.............. 2692
VER.KE.x.AGM155         --------G--------C--TA--A--A-G--T-CAGA---A.....CAATTAAGG---TC-----G----------AG----A---------G--AGGGT-A--C----A-G-A-------A-TG--GAAG--C--T-GGAA-ACA-G-T-TTA-G-.............. 2684
VER.DE.x.AGM3           -----G--G-----------CA-TA--A-----GCAGA---A.....CAATTAT------C------------------A--------C--------GGGAC----T----A-G-----AG-G-T-GGGAAG--AG-T-GGAA-ACA-A-TTTTG-G-.............. 2187
TAN.UG.x.TAN1           -----------T--------CA--A--A-G---GAAGA-A-C.....CATTTAGA----C-C---CGG--T---G-AG----A--T--A-----GGGA---TC-----A---A-----CA-T--TT-AG-GA-T--AGA----A-AG--A---G-.............. 2664
SAB.SN.x.SAB1C          --C--------------CT-----A--C-G-GT--AGAA--A...........GG--AC-AT----------G---A--T--A-----A-----A-G---T--T----A-GCA--CC-CA-TCAGGAGG--A----AGAA-ATA-A-CTTG--A-.............. 2866
MND_1.GA.x.MNDGB1       -----------T--------CA-C--TA-T----GA-A-A-A.....AAATTAAA-----AT---C-G--------T-----------A--T--AAACT-A-G-----A-----------A-TG--TATG-------AA-A--GA-GGG-A-AT-T.............. 2149
DEB.CM.99.CM40          --------G-----------CA-TA-TTC-----AAGA-A-A.....GATTTAGAG-C-CCT----CC-------CAG--------T-A---------C--A--T--C-A-TGT--CCCAGG----GAA--T-C---GGC-AATA-GGT---AG--.............. 2283
DEB.CM.99.CM5           --C-----T-------------A-TA-TTC---G-AAGA---A.....GATTTAGAT-C-CCT----CC--G----CAG-G------T-A--G--A--------T----A-TGC--CCCAGG----GAG--TA----GGCA-ACA-GGT----G-T.............. 2277
TAL.CM.01.8023          --------G--------C---A-CA--C-----GG-GACA-A.....GAATTAGA----GTGCCCC-C-A-----CAG--------T-G-C---A-AC---C-G---CA-AGC---AC-GGAG--GAA--C---TGGGAACACA-A-G-----GG.............. 1935
TAL.CM.00.266           --C--------------C---A-TA--C---T-GGAGAAA-A.....GAATTGGA----GT-CCCC-C-A-----CAG--------T-G-C---A-AC--AC----GCA-AGT---AG-GCC---GAA---C-GTGGGAAAACA-A-G---A-GG.............. 2437
SUN.GA.98.L14           --C--G-----------C---A--G-GTCT---GATTC-G--......TCCATAGA---TCAC-----G--C-GGG-AG-G--T--T--A--G--ACAA--T-G-----A-G-A----CAG-TG-CT-TG-------TA-G-ATA-A-C-----C-.............. 2819
WRC.CI.97.97CI14        ------------------------A----T-T--ATTTA..................AATAAG-AT---C-G--T-G---A--T--A--G--A-----CAC---G--T----A-TTAC--AGA--AG--AAA---AG---A-TA---A-A-C-ATAGA-.............. 2762
WRC.CI.98.98CI04        ------------------------A----TATCCATTTAAAT...........AATAAG-AT---C----T-G---A--T--A-----A-----CAC---G--T--T-A-TTA---AGGA-A----AA---AG---T-TA--GA-A-C--TAGA-.............. 2777
WRC.GM.05.Pbt_05GM_X02  -----------------C--T---G----T-TTCAATTA...............GA-----AT---C------G--A-----A--T--A-----AAC---GC-------A--TA---A-AG-TG-CAGA---A----T-TA---A-GG---TA-G-.............. 1691
LST.CD.88.524           --C--T--------------A-C--TTCT----ATTCAG-T.....AGAATAGA----CCT-AT-CT----G-G-AG-------T--A--G--ACAA--T-------A-G-A-TCAGGG-TG-TT-T---C----AGCA--GA-G-CCA-AG--.............. 1727
LST.CD.88.447           --C--T--G-----------TA-C--TTC-----ATTCAG-T.....AGAATAGAG---CCT-AT-C---T-G-G-AG-----------A--G--ACAA--T-G-----A-G-A--CA-GG-TG-CT-T--------AGCA---A-A-C-A-AG--.............. 1730
LST.KE.x.lho7           --C--T--G--T-----C---A-T--TTC-----ATTCAG-A.....AGAATAGA----CCC-AT-CT----G-G-AG-------C--A--G--GCAA--T-GG----A-G-A---AGAG-TG-CT-T--------AGCA---A-A-CTA-AG--.............. 2814
SYK.KE.x.SYK173         --------G-----------TA--A-TAG----GAAGA-A-T.....CAATTA-ACCA-CC----TC----------AG----A---T-A-----AAAC--T-C---T--G--A---AGAA-T--TAGAT-CACTG-TGT-AA--CCTC--GA-A-AGAAAGCAAGTAGA 2664
SYK.KE.x.KE51           --C-----G--------C--CA--A--AG----CAAGACA-C.....CAGTTA-ATCA-CCT----CC---------AG---------T-A-----AAAC--A-GG---------A---A-C-----CAGAT-TC----AGAAA---CAG-T-GG-G-ATAAAGGAAGTAGA 2322
COL.CM.x.CGU1           ----------------------T--TA-----GAAGA-G-T.....CAATTAA----CGCT----C-G--GT-CAA--TCA------A--G--AGCA----G-----A-------AGAG-T-A-A-TC--A-G--GG-A---A-AG--TAC-G-.............. 2243
```

Pol Protease end Pol p66 and p51 RT start

```
H1B.FR.83.HXB2            .GGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCCATTAGCCCTATTGAGACTGTA.........CCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCATTGA 2628
Pol                        G  T  V  L  V  G  P  T  P  V  N  I  I  G  R  N  L  L  T  Q  I  G  C  T  L  N  F  P  I  S  P  I  E  T  V  .  .  .  P  V  K  L  K  P  G  M  D  G  P  K  V  K  Q  W  P  L
H1A1.UG.85.U455           .-----------G--------------------------------G--A-------------------T-------------A-----T--------A-------.........-------------A----A-------------G----------------- 2074
H1B.US.90.WEAU160         .------------------------------------------------------------C---------------------T--------A-------.........--------------------------------------------------- 2627
H1C.ET.86.ETH2220         .---------C----------------------------C-----CA----------C----AC----A-----C-----A-----T--C-----A-------.........-------------------------------------C-------------- 2020
H1D.CD.84.84ZR085         .-------------------C-----------------------T-------C--------C-----------C-----A-----T--------A-------.........--------------------------------------------G---- 2150
H1F1.BE.93.VI850          .--------G-------------G----------------------A-------------------T-------------AG----T--------A-------.........-----------G-------------------G---------------- 1965
H1G.SE.93.SE6165          .--G--GA------------------A------------G-----A-------------------------------A-----T--------A-------.........----------------------------GG----------------- 2025
H1H.CF.90.056             .--------------------------------T-----------G---A-A--------A-----------C-----------A-----T--------A-------.........-----------------------------G------------------ 1975
H1J.SE.93.SE7887          .--------G---A--------------------------G--CA-----------C-------T--------------A-----T--------A-------.........------------------------------A----------------- 1942
H1K.CM.96.MP535           .-------------------------------------C----C---------------------T--------------A-----T--------A-------.........-----------------------------G----------------A- 1824
H101_AE.TH.90.CM240       .-------------------------------------C-----A--------------------T----------C--A-----T--------C-------.........-------C-------------------A-----G--------G----------- 2202
H102_AG.NG.x.IBNG         .-------------------------------------C-----A--------------------T--------------A-----T--------A-----G.........-----------------------------G------------------ 2153
H1N.CM.95.YBF30           .------------------A--------T--T--T--A----------T--A--C--------C--T--------------A--A--T--------A-------.........-------------A-------------------G--A-----------T---- 2222
H1O.BE.87.ANT70           .--A--------G--G--------T-----T--T--T------------A-A-----AGGAT-A-----T--AC----C--C--T--A-----C--A-CCC-A--G.........---------C----A-------------A--------A-----------CC-AT 2683
H1O.CM.91.MVP5180         .--A--------G--G--------T-----T--T--TC----G-----CA-A-----AGGAT-A--A--T--AC-------C--T--A--T--C--A-CCC-A--G.........-------GC----A-------------A--------A-----------CC-AT 2658
CPZ.CD.90.ANT             .-CA------C-GT-G-----A-AC--A--A-----T--A--T------G-T--ATG--TTT-A--A--T---------------A-----TAAAG----A--A--G.........--C--T--------AGA------------A-----G---A-----------GC-CT 2067
CPZ.CM.01.SIVcpzCAM13     .--G-----G--G--G-----A-----------T--------G------A-T--A--A-----A--G-----AC-GGTA--C--T--C--TT-A-----A-----C.........-----C-----G--A-----------C--G-----G-A-A--G----------T--AT 2208
CPZ.CM.05.SIVcpzEK505     .--C-----C--------G--A--------T--T--T------------T-TC----A--A---------T---C------------A--A--T---------A--CA--.........-----G-----G-------T------------G-G---A-----G-----T---- 2200
CPZ.CM.05.SIVcpzMB66      .--------T-----------C---------A--T--C--A-----G--A-T---------A------A---------------C-----C--T---------A-------.........-------GT--------------------G----GG--C-----G---------A- 2174
CPZ.CM.98.CAM3            .--C--G-----G--G-----A-----A--G--T--------C------A-A-----A--A-----C--T---------------T--A--T--C-----A--A--G.........---------------------------------T-G---A-----------T--A- 2039
CPZ.GA.88.GAB2            .A-C------C----G-----A--T--A--T--T--T--A--------G---A--C--AT-A--A--T---C-----------A--A--T--A---ACTC----G.........-----------------------CA-----A----GG--A--G------------T 2021
CPZ.CM.98.CAM5            .--C-----------------A-----A--A--T--------C------A-A--A--A--A--------T---C-----------T--A--T--CG-------A--G.........--------------------------------------G---A-----------T--A- 2330
CPZ.CM.05.SIVcpzMT145     .A----T--------------A--C--------T-----A--------G-CC-A--AA-AT-A--G--T---C--------C--A--A--T--A--A--C--A---.........AA---GTTTC----------G-----C--G--C--G--A-----G-----TC--T 2170
CPZ.TZ.00.TAN1            .-CC--T--CC----G---GAA--C--CA-T--T-----A--T------A-T--A-AG--AT-A--A--T--C-----------T---------A--A---GTA---.........AA----C--------AGA------------G---------A--G--G-----CC-CT 2272
CPZ.CM.05.SIVcpzLB7       .--------T-----------C-----A--A--T--C--A-----G--CA-T--A------A------A--T---C-----------T--------------A-------.........------C-------------------------G---------A-----------C--A- 2183
CPZ.US.85.CPZUS           .---T----T-----G-----A-----A--A--T--T-----C--G---A-T--A--A--A-----A--T---------------T--C---T-----A-A-------.........-------G------A-----------------T-G---A--G---------------- 2688
CPZ.GA.88.GAB1            .--G-----------------A--------A--T---------------A-T-----A--AT-G-----T---C--GTG--C--A-----TT-A-----A-----C.........-----C-----G--A---------------T-----G--A--G-----------C--AT 2689
MON.CM.99.L1              .-CA-----------T-----A--T---A-A--T--TT-A--G------G-C--AG--A----G--GGTA--CC-C--CA-GG-TGC-G-GGGAGACCTCCAACCCCCATCTTTAGAG--C-CT-----AG------AA--G--A-----GT-G-----G---------AT 2334
MON.NG.x.NG1              .-CW-----T--GC-G-----R--C--AA-T--T--TT-G--G-----CG-CC-CT-AAG---G--GATA--CC-G--CA-GY--CAGG-AATCGCCC-AT-CCC-...TCCTTGA----G-CCC-T--A-------AA-----G------T-G-----------CC-CT  927
GSN.CM.99.CN166           .-C---T---C-G--T-----C-----A--A--T---T-A--------TGC--A-G-A--T-A--GGTA--------CA-GGTGCAGCAGAAG--A--AC-CA--.........GA---GC-CC-----G------AA--G--A-----GA-C-----G-----CC--T 2338
GSN.CM.99.CN71            .-CC--T---C-T-----G--C-----A--G--T---T-G--------TGT--A-GCA--T-A--AGTA--C-----CA-GGTGCAGCAGAAAC-A---C-C---.........GA---GC-CC---AT-----AA--A--A-----GA-----------------C---T 2338
MUS_1.CM.01.1085          .-CA---------------G--A-----A-A--T-----G--G-----CTGT--A-GCA-AT----CATT---C----CA-GATACAGG-AAAAT-A--AC-AA--.........AA---G-CCC---AGA---GCAA-A--G----TG--C--G-----------T---- 2318
MUS_1.CM.01.CM1239        .-CA-----G--G--G------------CA---------G--G-----TGT--A-G-A-AT----AAT---C--G--CA-GATACAGGAAAAA--A--AC-CA--.........AA---GGCC------GA---GGCAA-A--A----TG--C-----G-------C--- 2308
MUS_2.CM.01.CM1246        .-CC-----C--G--GA-T---------CA----T--CT-A--G-----TGC--AT-AA-AT-G-AAATA--AC----CA-GGTA--GGCAGGACAGC-ATTGAA-...CCCACTAA---T-GCC-T--AGA-A---AAA-A--G----TG--A-------------C-CT 2416
MUS_2.CM.01.CM2500        .-C------TC-----------------AA----T--TT-A--------CTGC--AT-AA-AT-G---ATT--C--G--CA-GGTAGGAG---AAC-G--AC--ACG.........AA---TGG-C-C--AG----G-AA--A--G----TG--A----------------T 2375
RCM.NG.x.NG411            .-C-CAT-----G---------A-----A--A--T---------C--G---G-----GCA-AT-A--AGC-T--C-----------A--A--TAAAGCA--A--AA--.........AAG---G---------------CAA-----G--T-G---A-----------T---- 2122
RCM.GA.x.GAB1             .-C-CAT--GC----G-----A-----A--G--T--T--A--G-----G-T--A-AGA-AT-A--AGC---AC---------A--A---AA--A--A--A--.........AAG---G--------T-T---CAA-----A---CG---A----------------T 2107
DEN.CD.x.CD1              .-AC---------------------CA-A-----CT-A--------CTATC-T-G-A--T-A--AGTA--AC----CA-GG-AG-ACAG-AA--GCG-C--CA-...GTCACC---------------A-----T-T--------T-G-A-A-----------T--A- 2515
```

Pol Protease end Pol p66 and p51 RT start

```
MAC.US.x.239              .--G---A-CA-GAC---GGAC--C--GA-T-----TT----T------T--C-A--AGCTC-G--GATGT--C-------------AGCTAAAG-A---C-------.........AA---CGCC-------------A------A------T-G--G--G---------AT 2930
Pol                        K  G  T  I  M  T  G  D  T  P  I  N  I  F  G  R  N  L  L  T  A  L  G  M  S  L  N  F  P  I  A  K  V  E  P  .  .  .  V  K  V  A  L  K  P  G  K  D  G  P  K  L  K  Q  W  P
H2A.GW.x.ALI              .-CC--CA--A-GAC---TGAC--C--AA---T--TT----C-----CA-T-----AGCCT-A--CATGT-A-----CC-A--AG--GC-AAG--A---C-AA--.........GAG----G--------------AA--C--G------T-A-G----------C--A- 3006
H2A.DE.x.BEN              .-CC--CA--A-GAC----GA---C--AA-------CT----C------A-TC----AGCCT-A--CATGT-A---------A--AG--GC-AAG--A---C-AA--.........AA-----C---G--------G-AA-----A----GGC-G-----------CC-A- 3013
H2A.SN.x.ST               .-CC--CA--A-GAC---TGA---C--AA-------TT----C-----CA-TC----AGCCT-A--CATGT-A-------C-A--AG-CGC-AAG--A--AC-AA--.........AA-A---TGC-G------------A------A-------C-G-G----------C--A- 2458
H2B.GH.86.D205            .-CA--TA--A-GAC----GA---C--AA-A-----TT----C------A-T--A-A-ACCT-G--CATG-----------C--AG-GGCAAAGG-A--AC-A---.........AA---TG-G-----A--T----AA-----G-----GA-C-G---------TC-AT 2989
H2B.CI.x.EHO              .-CA------A-GAC---GGAC--C--AA-A-----TT----C------A-T--A-A-AGCT-A--CATG---C-------C--AG-AGCAAGG--A--AC-A---.........AA---CC-G--------T-A--AA-----G------A-C-G---------CC-AT 2984
H2G.CI.92.ABT96           .-C---T--TA-GAC---GGA---C---A-A--T--CT-------G--CA-T--AGT-A--C-A--CATGT--C-----------A---GCAAAGG-A---C-A---.........AA-----GGC----A--T--C--------G------A-C-GG---------CC-AT 2340
H2U.FR.96.12034           .-----CA--A-GAC----GAC--C--AA----T--TT----T--G-----T-----AG-AC-G--CATGT--C----C-----A--AG-AAAAGAA-CACTA--C.........ATT-GGGG-C----A-----C-A------A-------T-A-G---G-----T--AT 2495
MAC.US.x.EMBL_3           .A-G---A-CA-GAC---GGAC--T--GA-T-----TT----T--G---T--C-A--AGCTC-G--GATGT--C-----C-------AGCTAAGG-A---C-------.........AA---CGCC-------------A--T---A-------T-G--G--G---------AT 2404
SMM.US.x.H9               .--A---A-TA-GAC----GA---C--AA-T-----TT----C-----------C-A--AGCT--G--CATGT--C-----C-C-----AGCTAAGG-G--RC--A--.........AA-----C-C----A-------AA-----A-------T-A-G---G-----GC-AT 2403
SMM.US.x.PGM53            .--A---A-TA-GAC----GA---C--AA-T-----TT----C--------------A--AGCT--G--CATGT-C-------C-C---G-AGCTAAGG-G---C--A--.........AA-----C-C----A-------AA--C--A-------T-G-G---G-----GC-AT 2847
SMM.US.x.SIVsmH635F_L3    .--A---A-TA-GACG---GA---C--AA-T--T--TT----C------T--C-A--AGCT--G--CATGT-----------C-----AGCTAAGG-G---C--A--.........AA-----C-C----A-----G-AA--A--A---------G-G---G-----GC-AT 2933
SMM.SL.92.SL92B           .-AA---C--A-GAC----GA---C--GA-A--T--TT----C------A-TC-AG-AGCAT-A--GGT-T---------AC--AG-GGCAAAGG-A--ATA-AC-.........AA------G--G---GAG------------G----G-T-A-----------C--AT 2361
STM.US.x.STM              .--A--TA-TA-GACT---GAC--C--CA-T-----TT----C--G--C--A-----AGCTT-G--GATGT-A-----------AG-AGCTAAAG-A--AC-A---.........AAG--G-C-------A-------AA-----A-------A-A---------------AT 2577
MNE.US.x.MNE027           .--G---A-CA-GAC---GGAC--C--GA-T-----TT----T------T--C-A--AGCTC-G--AATGT--C-------C-----AGCTAAGG-A---C-------.........AA---C-CC--------------AA-----A-------T-G-GG--G---------AT 2398
DRL.x.x.FAO               .-C-T-G--G--G-----C-----C-----A--T--T--A--T------G-AC-CTG-A-AT-A--AGCT---C-T--------TG----T-AAG-A--AC-A---.........AA-----C-C----A---------------A-----GA-A-----------CC-CT 2195
MND_2.x.x.5440            .-C-G---TC-G--G--C-----C-----A--T--T--A--T------G-AC-T-AGA--T-A--A--T---------------T-----TAAG--A-----A---.........AA-----C-C----A-------CA-----A--G-G-A-A-----G-------C--T 2189
MND_2.CM.98.CM16          .-C-T----CC----G--C--C--------A--T--TG-A--T------G--C---AGA-AT-A--A--T--AC-----------AG----TAAAG-A--A--A---.........AAG----C-C----A-------CT--C--A--T---A-A---------------T 2629
MND_2.GA.x.M14            .-C-T----GC----G--C-----C-----A--T--T--A--T------G-TC-T-AGA--T-A--A--T---C-G---------T-----TAAG--A-----A---.........AA-----C-C----A---------------A-----G-A-------G-------C--T 2556
OLC.CI.97.97CI12          .-AGG----------G---GA------A--T--TG-GG-A-----G---G----AG-AGCAG-A--GATACG-C-GGTC-GCG-ACAGCTTT-AGA-A---TCCCT...ATTACTAAG-----------AGAG---CACA-GCCT---C-GA----G--------T--A- 2502
GRV.ET.x.GRI_677          .--CT--A-TC--A-----AG-----AA---T--T--A---------A-T--AG-----GCA--CATG-AA---GT-A-GGGAG--CTAT---G-C---T--AG...GAAACAAA----C--C----AGA---G-AA-----A--T---T-G-----------C--AT 2801
VER.KE.x.TYO1_patent      .--A---A----GT-----G-A--T--CA-T--T-----A--T------T--C--G-C-C-GCA---GC-CGG---GTAA-GGGACAATTAT-AGAAA-A-T-CCT...GTCACA--T--C-----G---GA---GGCTCGG--A--CTGT--A-G---------TC-CT 2855
VER.KE.x.9063             .--C-----TC-TT-----G-A--C---A----T---------------T-AC-AG-A-CAGCA--AGCT-GAC--GTGA-GGGACAATTGT-AGAAC-A-T-CCC...ATCACC-----C---C-G--AGAT---GCAAGG--A---C-TA-A-----------TC-AT 2858
VER.KE.x.AGM155           .--A-----CC-CA-----G-A--T--CA----T--C--A--------CT-T--AG-C---GCA--AGC--AA---GTGA-GGGGCAATTGT-GCAGACA-TACC-...ATCACC--G---CGC------GA---GGCCAGA--A---CG-T-G--G----------C-CT 2850
VER.DE.x.AGM3             .--G--CA----GA-----AGC--T--CA-A-------------------A-A--AG-A-CAGCA--AGC--AA---GTAA-GGGTCAACTGT-AGAAC-A-T-CCC...ATTACC--T--G--------AGA---GGCTAGA--A--TTTCT-A-----------CC-CT 2353
TAN.UG.x.TAN1             .----AT---C--A-T-----A-----AA-A--T--TG-A----------A-TC-AG-A-CTT-A--AGCA-GA----TAA-GGGACAGCTAAA-GAGA-A-TACCC...ATAACCAAG---C----G--A--T--G---TCA--A--G---T-A-----------CC-AT 2830
SAB.SN.x.SAB1C            .-C---CA-T--G------GAA-----A--T--T--T--A----------G-C--AG-A--AT-A--AGTT---C-------GA-ACAA--AGAG--A--AC--A--.........AA---CC-C-------------CAA-----G----GGA-A-GG---------T---T 3026
MND_1.GA.x.MNDGB1         .--G--------GA----------T--AA-AG-T-----A--G-----CA-AA--GAAA-AT-A--AG-A-AA----TA--GG-ACAATTGT--GA-A-A-TACC-...ATAACAAA---G---------A------G-A-----A--C-G-A-A-----------T--A- 2315
DEB.CM.99.CM40            .TCAGATC-----------GA---C--AA-T--T--TC-A--------TAT--AG-AA-A--G--GGTAT--A----------AG-ATTAGA-TA-ACCC-A---................--------AGAG--GGCA-----A-----GA---------------C-TT 2437
DEB.CM.99.CM5             .TCAGAT--G---------TGA---C--AA--------T-G-----G---TAT---G-AA----G--GGTAT-CA-----------AG-ATTGGACT--ACAC-A---................--G--G---GAG---GCT----------GGA-------G-----CC-TT 2431
TAL.CM.01.8023            .-AC---------A----G--------AA--------T-A------------T-CC-AG-CA-AT----GGTA--AC-T---A-GGT--GC-CAGAA--ACCAGTAA--.........AA------G--G---GA----A---ACC------GA-C--G---------C---- 2095
TAL.CM.00.266             .-AC-----T--GA-T-----A--G--AA-T-----TC-A--G------T-T--AG-AA-AT--A--GTT--AC-T--CA-GGT--GCTCAGAC--ACC--TAACC.........AAG--T--------AGAG--T-A---ACC---C---A-C--G---------C--A- 2597
SUN.GA.98.L14             .--A-----G--G--G--G--A-----A--AG-T--C--A--G------A-TC-TT-AGCCT-A--GG-A-GA---GT---AG---C-CTTAG-GAAA---TCCC-...ATTACAAA---T--G-----AGA---GGCC-C---A---------G-----------CC--T 2985
WRC.CI.97.97CI14          .-CAGACA-TA-GAC---GAAA------A-A--T--T--A-----G---G-A--A-A-A-A--G--AATGCA----GCC---G-ACAATTAT-AGACA---T-CCT...ATAACAAA---G-GC--G--A------TGCTCA--A--T-G---A-----G-----C---T 2928
WRC.CI.98.98CI04          .-CAGATA-TA-GAC----AAA--T--AA-T--T--T--A--T------G-A--AGAAA-A--G---ATGCA---GGCC---G-ACAATTAT-AGA-A---TACCC...ATAACAAA----GT--G---------TGCACA--A----G---G--G--G-----T--AT 2943
WRC.GM.05.Pbt_05GM_X02    .-CA-ATT--A-GAC----AAA------A-A--T-----A---------G--C-AGAAA-AT-G--AATGCA-C--GCA---G-ACAATTAT-AGA-A-A-T-CC-...ATAACAAA---CTC------A--T---TGTC-A--A----G-A-A--G---------TC-CT 1857
LST.CD.88.524             .--G-----T--GC-------A--C--A--AG-T---G-A--------CA-T--A-AC-CTC-A--GGCT-GA---GTG--AG-ACAATTAAG-GAGA-A-TACCT...ATAACAAAG--GC-T--G--------GTGT-----G----G---A-----G-------C-AT 1893
LST.CD.88.447             .--G-----------T-G------------A--AG----TG-A--------A-T--A---A-TG-G--AGCA-GA---GTAC-AT-ACAATTAAG-GACA-A-TACCC...ATCACAAAG--C------G--------CTGT-----T--C-G---A-----G-----T---T 1896
LST.KE.x.lho7             .--G-----C---T----------C--A--AG-T--TG-A--G--G---A-T--A---GCA--G--GGC--AA--G-T---AG-TCAGTTAAG-GA-A-A-T-CCT...ATTACAAAG---TCT-----A--T--GTGT-----A----G---A--G--G-----TC-AT 2980
SYK.KE.x.SYK173           G--G--TC-T-----G--T-----G--A--A--T---T-A-----G---A-AC-T--CA-AT-G--AGT--AA---GTAA-GGTTCAA-CAG-AT-A---C-A--C.........AA-----GTC-G--A----AC-AA--ATTG--CCGGC-A-----------------AT 2825
SYK.KE.x.KE51             G----GTC-TC---------------A--A--T--TT-A--------CA-A----GCA-AT-G--AGCA-AA---GTAA-GATAGCACAAAAAT-A--AC-AA-G.........AA---CC--C------------G-A---GCTA--T---T-A-----G-----TC--- 2483
COL.CM.x.CGU1             .--AGATA-C---A------AC--C--CA-A--T---G-A--T------A----A-AAG-AGGGAAA-TAGTAG--GC-CAATTATCAGATAGG--CCCAGTCACC.........AAG--G-CG-----AGAG--C--------G--------A--G--G-----C--A- 2403
```

```
H1B.FR.83.HXB2           CAGAAGAAAAAAATAAAAGCATTAGTAGAAATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGAAAATTAGTAGATTTCAGAGAACTTAATAAGAGA 2798
Pol                      T__E__E__K__I__K__A__L__V__E__I__C__T__E__M__E__K__E__G__K__I__S__K__I__G__P__E__N__P__Y__N__T__P__V__F__A__I__K__K__K__D__S__T__K__W__R__K__L__V__D__F__R__E__L__N__K__R_
H1A1.UG.85.U455          -------G---------------AC-----------AT-----------------A----------------------------------------------T----------G-----C-----------G--------------------C-----A--- 2244
H1B.US.90.WEAU160        -------G-------------------------------------A---------A-------------------------------T---------------------------------------------------------------------- 2797
H1C.ET.86.ETH2220        ----------G------------AC--C-------GA---A----GC-------A---------G-------------C-----T---------------------A--G--G-----------G--------------------G-----C-----A--- 2190
H1D.CD.84.84ZR085        -----------------------AC--------------T--------------A---------G-----------------------------A-----------------------------G--------------------------------- 2320
H1F1.BE.93.VI850         ----------G------------AC------A---CT---A--------A----A------------------------------------------------------------G--------G-----------------A-----------A--- 2135
H1G.SE.93.SE6165         -----------------------AC-----------A---A------G------A----------------------------C-----A------------------G---------------------G-------------G--C-----A--- 2195
H1H.CF.90.056            -----------------------ACG--------------------A-----A----C---G--A-------G---------GC-----A------------A--G--G--T----------------------G--------------C-----A--- 2145
H1J.SE.93.SE7887         -----------------------AC-C--------G---A------G------A---------G-G------------------T--C------------------------------------------------------T--------C-----A--- 2112
H1K.CM.96.MP535          -----------------------AC-------------------A--------A-----A---------------G--------T--C-------G-----------------------------------------------T--------C-----A--- 1994
H101_AE.TH.90.CM240      -----------------------AC-----------A---------G------A----C------------------------------------------T----------G-----C--C---------G-----------------G--C-----A--- 2372
H102_AG.NG.x.IBNG        -----------------------AC---C---------------------------A---------------------A---------------------------------G--T-----------------------------------C--------- 2323
H1N.CM.95.YBF30          --AC----------G-G------AG----------------A-----------A--------T-G---A-------G--------T--C------A-T-----T-----A--G-----T--C----------------------------G---T-A-----A--G 2392
H1O.BE.87.ANT70          -TA-----------G----C--GAC--C---A---CAG--A------C-A----A----------G---A--A-----------T--T-----A--TA-C-----T-----A--G-----TG------------------G--------T--G---T-A--------- 2853
H1O.CM.91.MVP5180        -TAG---G------G-----C--ACT-C---A---CA---A------C------A-----C----G---A--A-----------T--T-----A--TA-T-----T-----A--G-----T--C-----G----------G-----C---------T-A-----A--- 2828
CPZ.CD.90.ANT            --A----G------G-G--CC--AA----------GATA--T-A---GCA---AAT--G-----T-G---A-----A--T--C-----------A---A-T-----A-----A--G-------C-T-A-----------GC----T-------------T-A-----A--- 2237
CPZ.CM.01.SIVcpzCAM13    ---C----------T-----C--GAC-------A---CAG--A-------------G--A--T-----A--A--------C--------------TA-T-----A--T--A--G-----T---T-A------------------------------T-A-----A--- 2378
CPZ.CM.05.SIVcpzEK505    -----------------G---C--AC------------------A--------A------A--------C-G--A----------------T------------A-T-----T-----A--G-----T--C------------------------------GT-A--C--A--- 2370
CPZ.CM.05.SIVcpzMB66     ----------------G-------AC------------------A-----------------A----------G-G-A-------------------C---A-T-----T-----A--G-----T--C------------------------------GT-A-----A--- 2344
CPZ.CM.98.CAM3           ---C----------T---------AC-----------CAG--A--------A-----A--G--------G---A-----C----------T-----A---A-T-----T-----A--G-----T-----A-----------C-G----------G---T-A-----A--- 2209
CPZ.GA.88.GAB2           --A-------G---G-G-------AC------------GG--A--------A-----A-----------G---A--T---------------T-------CA-------A----G---G---------------A--G-----------------T------T-G-----A--- 2191
CPZ.CM.98.CAM5           ---C----------C--G--C---AC-----------CCAG--A-----------G--A-----------G--A--C--A--G--C--T--T--C--A---A-T-----T-----A--G----------------A-----------C-G--------T------T-A--C--A--- 2500
CPZ.CM.05.SIVcpzMT145    --A-----------G-G-------ACT------A----A---A---------------------A-----------A-----A----------T---------A--TA-T-----A--T----------G-------G-----C--C------------C-T--T------------------------G 2340
CPZ.TZ.00.TAN1           -CA-G--G------TG-G------AC------A-----AGACAT-------------A----------TGC-G----A--A-----C-----T--C--A---A-T---------T----------G--T-CCT----------------------------------G-----A--- 2442
CPZ.CM.05.SIVcpzLB7      ----G---------T-----T---AC------C---------A-------C-------A-----------G---A---------------C-----T---------T--G-----T-----A--G-----T--C--A-------------------------T------T-A-----A--- 2353
CPZ.US.85.CPZUS          ---C----------T--G------AC-----------CCA---A--------A-----A--G---A---G-G-A-----A---------T--------A---A-T-----T--C--A--G-----T--C-------------------------C--T------T-A-----A--G 2858
CPZ.GA.88.GAB1           ---C----------T-----C---AC-----------CA---A-------------------A--G-----A--A-----------------------------TA-T-----A--C--A--G-------------------------------C------------A--------- 2859
MON.CM.99.L1             -CAG------G---G-G--TC-TAC-C-----AC-CA-------CTT--ATT---AC--T-GGA-CC--CC-A---CA-T--C-----T--CT-A--CTGT-------------------G-----AG-GC----------TGC-CA-------------------C---GC- 2504
MON.NG.x.NG1             -TA-------G---G-G--TC-CACCACC---RCCCGG------AT-G-ATT---CC--T-GGA-G--GCA-A---AACC---------T---T----TTGC--C--G----G------G--T-AA--C----------G-TGC-CA---------G---T-A--C--AGC- 1097
GSN.CM.99.CN166          -TAGG-----G---G-G--C---ACCC----AACACA--------CTT-GATT---TC--T-GGA------A--A--A---------------CT-A-----C-----------A--G--G----AAT-AC----------TGC--A-------T---A--T-A---G--GCT 2508
GSN.CM.99.CN71           -TA-G-----G---G-G--C---ACCC----AACACA---A---CTC---TT---CC--T-GGA---G--A-----A----------------CT-G-----C-----------A--G-------AAT-AC----------TG---A----------GA--T-G---G--GCT 2508
MUS_1.CM.01.1085         -CC-G-----G--C-GG--C--G-AG--C--AACCCAGC-----CT---A-T---CC----AGAGG---A--A--A--G-----T--T---T-----TGC-----A--T-G------G---AAT----G-------TGC--A----------GC--T-A---G--GCT 2488
MUS_1.CM.01.CM1239       -TC-------G--T--G--C-----A--GC--AGT-CAGC-A---TT---ATT--ATC-G--AGA-G-G--A-----A--C-----T-----CT-----TGT-----A--C-G---G-----T-AAT------------TG---A-------T---C-G--A---G--GCC 2478
MUS_2.CM.01.CM1246       -TA-G--G--G---G-G---C--AA--------AC-CAGA-A---AT---ATT---TC--C-GGAGG--GCA------A-T--C-----T---T-A--------C-------GA--G------AG--A------C---TG---A-------------A-G------G--GTG 2586
MUS_2.CM.01.CM2500       -CA-G---------G----TC---A--------AC-CAGAGA---TT----CTG--AC--T-AGA-G---CA------C----------------GC----C------T----G------G---AAT-C----------TGC--A---C-------A-G--C---G--GCC 2545
RCM.NG.x.NG411           GTA----G------TG----T---AC---G-------AT-CC-------------G--A---------T-G---A--A--A--------T-------------A-----TGT--C--A--G-----TTCC--A------------GC-------------------T-G-----A--- 2292
RCM.GA.x.GAB1            --A-----------G-----T---AC---------C-AT-CA-----G---------A-------------A-----A----C--C-----C--A---A---CTGT--T--A--G------TCC--A-------------------G--------T--------A--------- 2277
DEN.CD.x.CD1             -TAC---G---T-------T---CA-C-G--AGTG---C-------------G--A-GGT-GGA---GGCGCCA--CACC--C--T---------A--T--G----TA--TCCA--G--------AA--AC---AT--G-TG---A----------GA--T-A---G--GCC 2685

MAC.US.x.239             --A-------G---GTT------AG------C---GA-A-------------T--TC-GT-GGAGG--GC-CCC--GACC-----------C--C--CAC------T-----------G--T-AG-AC-----------TGC-GA-------T--G-----A---G-GTC 3100
Pol                      L__S__K__E__K__I__V__A__L__R__E__I__C__E__K__M__E__K__D__G__Q__L__E__E__A__P__P__T__N__P__Y__N__T__P__T__F__A__I__K__K__K__D__K__N__K__W__R__M__L__I__D__F__R__E__L__N__R_
H2A.GW.x.ALI             --A-----------G-G---C--AA------C---GA-A-A-C-----GA-----CC--T-AGAGG-GGCACCT--AACT--C--T--T-----C--CAC-----A--------G--G---AA-AC----------TGC--A-------T------T-A--C---GT- 3176
H2A.DE.x.BEN             --A----G------G-----C--AA---G--C---GA-A-A---------A--G--CC-GC-AGA-G-GGCACCT--AACT-----T--T-----C--CAC-----A--T--------G---AG-AC---------G-TGC-GA-------T--------A------GTG 3183
H2A.SN.x.ST              --A-----------G-G---C--AA---G--C---GAGA-A--------GA--G--CC-GC-AGAGG-GGCACCT--AACT-----T--T-----C--CAC-----A--C--------G---AA-AC----------TGC--A-------T--------A--C---GT- 2628
H2B.GH.86.D205           -CAGG-----G---CT---CC-CAA------C---GA-A-A-------------G--AC--T-AGA-G--GCACCC---ACT-----------C--A--CACC----------------G-----T-AG-AC----------TGC--A----C--T-----GT-A--C---GT- 3159
H2B.CI.x.EHO             -CA----G------CT---CC-CAA------C---GA-A-A---------A--G--AC-GT-AGA-G-GGCGCCT---ACT------------T-G--CACC--C---------A--G-------AA-AC---------G-TGC--A------------------A--C---GT- 3154
H2G.CI.92.ABT96          -TA-G---------C-----C-CAGG---------GACA-A------C---------C--T-GGA-G--GCACCC--CACC-----G--T--C--A--CACC---A-T--C-----G--G---AA-AC----------TG---A----C-----------------C-G-GT- 2510
H2U.FR.96.12034          -TA-G------T--G---------AA---G--A---GAGA-A-------------G--AC--T-GGA-G--GC-CCC--CACT-----------C-----CACC----------------G--G---AA-AC----------TG---A----C--T-----GT-A--C-GAGT- 2665
MAC.US.x.EMBL_3          --A-------G---GTT------AG------C---GA-A-------------T--TC-GT-GGAGG--GC-CCC--GACC-----------C--C--CAC----------------------T-AG-AC-----------TGC-GA--C----T--G-----A----G-GTC 2574
SMM.US.x.H9              -GA-------G----TT------AG------C---GA-A-A-----G--A--T--CC-GT-AGAGGR-GC-CCT--AACC---------T--C--C--CAC------T------R-GR-----G-A-A---------G-TGC--A-------T------T-A--C---GTC 2573
SMM.US.x.PGM53           --A-----------GTT------AG------C---GA-A-A---------A--T--CC-GT-AGAGG--GCCCCT--AACC---------T--C--C--CAC------T-----A--G-------AA-A---------G-TGC--A-------T------T-G----G-GTT 3017
SMM.US.x.SIVsmH635F_L3   --A-------G----TT------AG------C---GA-A-A---------A--T--CC-GT-AGAGG--GCCCCT--AACC-----G--T--C--C--CACT-----T--------G-------AA-A-----------TGC--A-------T------T-A------GTC 3103
SMM.SL.92.SL92B          -TA-G-----G---C--------AC---G--C--CGAGA-A-----------G--AC--T-GGA-CG-GCCCCT---AC------T--T-----A--GACC-----T--T--------GA-T-AGGAC--G-------TGC-CA---------------T-A--C--A-TG 2531
STM.US.x.STM             --A-------G---G-G---C--AA---G--C---GA-A-A-----G-----T---C-GC-AGA-G--GC-CCT--AAC------T---------C--CACT-----T----------G-----AA-AC--G-----G-TGC-GA------------------A-----AGTG 2747
MNE.US.x.MNE027          --A-------G---GTT------AG------C---GA-A-------------T--TC-GT-GGAGG--GC-CCC--GACC-----------C--C--CAC------T----------G-----AG-AC-----------TGC-GA-------T--G-----A------GTC 2568
DRL.x.x.FAO              -CA-G--G------TG----C--GAA------A---GAGC-A-----G--------TC----AAGT-G---A--T--A--------------C--A--G--G---TG---T-----G-----TG---C---------------------------GC--T-A------GC- 2365
MND_2.x.x.5440           -TA-------G--CCT---C---CA------A--C-ATC-A--------A-----C--G--C--T-G-G-A--T--A---------------C--C---------CTGT-----A--G-----TG-A--C-GC-------------------T---C----G--C--AGTC 2359
MND_2.CM.98.CM16         -TA-G-----G--TTT---C---CA------A----GTC-A-----G--------CC-G--C--T--G--A--T--A--------T------C--A--G--G---TG---C--A--G-----TG-A--C-GC-----------G--------T---C-GT-A--C--AGT- 2799
MND_2.GA.x.M14           -TA-------G--TTT---C---CAGAC---A---GAT--A------------------C----C--T-G---A--T--A----------T-----A---------TGT-----A--G-----TG-A--C-GC--------------------T--GC--T-A---GAGTG 2726
OLC.CI.97.97CI12         GTA-G-----G---G----T----AGA----A---GACAGAT-----G-A-----T--G--ACAG---GCA-A-G--TC-----A--GG-----C---A-----TG-----A--G--G---AA-A-G-G-AT-----G--GA-------T--G-CT--A---GCAGCC 2672
GRV.ET.x.GRI_677         --AG----------TG----T---AC------A--C-A-C-A-----G---G--A---T-A--T-GG--A--AGGA--------T--T-----A----G-----------------G--T-AA--AC---------TGC-T-----------G-----A--C--AGCT 2971
VER.KE.x.TYO1_patent     -TA----G--G--TG----T---CAG-----A---T-CC-AT-A--GC-------A----CAGT-G-G-A--AGGA------G----------C---A----TG---------G--G---AAT-CC-G-----G-TGC-------C--T-----GT-A------GC- 3025
VER.KE.x.9063            --A-G--G--------G--C---AC---G--A---GA-C--T-A-----A-----------CAGC--G--A--AGGA------G-T--T-----C---A-C--TGT--C-GA--G--G---AA---C-------G-TG-------C---------T-A-----AGT- 3028
VER.KE.x.AGM155          -TA-----------T---CC-GCA---------C-A-ACAT-A--GG-A-----A---T-AAGC-GGG-A---GGA--C---G----------A--------CTGT----G----------AAT-AC-G-------TGC-G-----------G-----C--C--AGCT 3020
VER.DE.x.AGM3            -CA-----------------C---CAG-----A---GACC-AT-A--G--A-----A------AGC--G--A--AGGA--G---G-------C--------G---TG------------------AGT-AC----------TG-----------T--G-----A--C--AGC- 2523
TAN.UG.x.TAN1            -------------------CC-GAC------AAC-CAG-CTC-A-----------A--------T--GG----AGGA--G-----T-----------C--G---TGT--CCG-----------AA--AC-G-----G-TG---A-------T----CT--C-----AGC- 3000
SAB.SN.x.SAB1C           --A-----------G-G--C---AAG-CC------GA---CT-A-----AC---AC-TT-GGA--G---A-----A-----------T--C--C-----C-----A----G--G-----T-AA---C-G-------T-C-CA-G--C-----GC-GT-A-----A--T 3196
MND_1.GA.x.MNDGB1        GTA----G------GTT-GTC-TCAGA----A---GATAGAT-A--GG-------A------AGT-GGG-A-AT--A-G---AAT--------A--TA-C---------------G--G--T-AA-A-G---------------A----C--T------T-A--C---TT- 2485
DEB.CM.99.CM40           --AGG---------G-G--C---AA---T--C-TA-AGCCC---AT-G---C---AC-G--AGTCCCGGCA-C---CACC--C-----T--CT------G---TA-------G--G----AA-G---G---C---TGC--A----C--A----CCT-A--C---GCC 2607
DEB.CM.99.CM5            --AG----------G----C---AA---G--CATA-GGCCT----T-G-A-C---AC-G--AGTCCCC-CA-A---CACT--C-----T-GCT----T--G---TT-------G--G----AG-A-----------TGC-TA----CC-G----CC-----------GCT 2601
TAL.CM.01.8023           -CA-G---------G---GGC-GCAG-CC--CATAGAC--C----TG----C---AC-GT-GGAGG-G--------G-----C--------CT------G-----T--C-GA--G--G---AG--AC---------TG--GA--------C---CC--G---C--GC- 2265
TAL.CM.00.266            -CAG----------G---GGC--CA-ATC--CATAGAC--C----TG----C---AC-GT-AGAGG-GG----A--G----------T---T------G--C--T--C-GA--G--G---AA---C---------TGC-CA---C---C-C-CC--A---C-AGC- 2767
SUN.GA.98.L14            --AGG--G------G---GCC--CAGA--------GATAGAT-----GCA-----A--G---AGC-GGGCA-AC--A-GG--C--T--T--C--A---A-T-----T--------G-------AA-A-G----------G---A-------T---T--A---GGT-TG 3155
WRC.CI.97.97CI14         --A--A--------G-G-GGC--CAGA--------GATAGAC-A---GCA-----A------AGC-G-GCA-AT--A-G---------T--------TA-------T--T--A--G-------AA-A-G--------------GA-----------A--T-A---G-AGTT 3098
WRC.CI.98.98CI04         --A----G------G---GGC--CAGA----C---GATAGAT-A---GCA-----A------AGT---GC-TCC--A-G---------T--------TA-------A-----------G-----G---AA-A-G-------G--GC-CA-------T---A--T-A---G-AGT- 3113
WRC.GM.05.Pbt_05GM_X02   --A----G-----TG-G-GGC--CA-A-G--A---GATAGAT-A---GCA-----A-----AAGT-GGGCA-AAG-AA-T-----T--------A---A-C-----A---GA--G-----T-AA-A-G---------------A-------T----G---A---GCAGT- 2027
LST.CD.88.524            -CA-----------G---G----CA--C-------GATAGAT-A-----------------A--CCC-G---AT--A-G-----T-----C--C---A-C-----T-----------G---AA-A-G----------G--GA------------A--T-A---G--TT- 2063
LST.CD.88.447            -TA-------G---G---GCC--CA--C-------GATAGAT-A-----A-----------A------GCA-ATG-A-GG--C-----T--C--A---A----------------G-------AA-A-G-G--------------A-T-----T---C----A---GCTTT- 2066
LST.KE.x.lho7            --A----G------G---GCC--CA--CT------GATAG-T-A-----A-----A-----A--TCC-G-G-AT--A-GG--C---------------A------T--C-----G-------AA-A-G----------G---A-T----------A-G--C---G-ATT- 3150
SYK.KE.x.SYK173          -G-T---G---T--G-G--CC-CAAG-C---AGTAGA---C---CTC----C---AC--T-GGAG---GC-A-T---ACC-----T--T-----C----------TA--T-G------G-------AA-AG---------G-TGC-CA-T-----T---A--T-A---G--GCC 2995
SYK.KE.x.KE51            --AG-------C-TG-G---C-GCG--C---AGT-T----T---CTTG-AA-G---C--T-GGA---GGCCTCC---AC------T--T-----C-----T----TA---C-A--G--G---AATTA--------G-TGCCTA-G-----T--GA----C---G-AGC- 2653
COL.CM.x.CGU1            GCA-----------G---GT---CA-A-G--A---GA-AGAT-----GCA--G--A-----AGA----GCA-AATTA-GG--------T-----A---A-----TG---T-G---------T-AA-A-G-------G------A---------G--G--A--C------ 2573
```


PLV Complete
Genomes

```
H1B.FR.83.HXB2             ACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCA..................GTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAGTATACTGCATTTACCATACCTAG 2950
Pol                        _T__Q__D__F__W__E__V__Q__L__G__I__P__H__P__A__G__L__K__K__K__K__S__.__.__.__.__.__.__V__T__V__L__D__V__G__D__A__Y__F__S__V__P__L__D__E__D__F__R__K__Y__T__A__F__T__I__P__S
H1A1.UG.85.U455            -----G----------------------------C--G---A-A--G--TC-------------------------------------------------A--------G--C-----------------T---------AG---T--A-----------G--C----------- 2396
H1B.US.90.WEAU160          -----------------------------------------TT-----------------------..................--------------------------------------A--------------------------C-------------------- 2949
H1C.ET.86.ETH2220          -----------T-----------------G--------C--A------------------------..................--G--------A--------G-----------C--------T----------GT-----A--A-----------C----------- 2342
H1D.CD.84.84ZR085          -----------------------------G-----------T-----A-----G--------G---..................A-------------------C---------------A--------TG---------T--------C-------------------- 2472
H1F1.BE.93.VI850           --------T--------G-----------------------T--------------------G---..................--------------------G-----------------------------A-G--T----A------C--------C----------- 2287
H1G.SE.93.SE6165           -----------------G--C--------------T-----T------------------------..................-----------A--------G-----------------------------------T--A--------------C--T-------- 2347
H1H.CF.90.056              -----------------------G--------------C--A------------------------..................---T----------------G-------------------C--T------A----A-----A--------------C----------- 2297
H1J.SE.93.SE7887           -----------------------G---------------A------------------------..................--C-----------------G-------------------C--T---T-------T--------A-----------C--T-------- 2264
H1K.CM.96.MP535            -----------------------------------------A-----A------------------..................--------------------G-------------------C--T------A-------------------------C--T-------- 2146
H101_AE.TH.90.CM240        -----G-----T-------------------G-----A-----T------------------..................-----------A--------A--------------------T---------AG---T--A--------------C----------- 2524
H102_AG.NG.x.IBNG          --------T--------G--C--G-----------T--------G--A------------------..................--------G--A--------G-------------------------------A-------T--A-----------------T-------- 2475
H1N.CM.95.YBF30            --C-----T--T--------G--GC-------T--------A-----A-----GC-----------..................--G-----TT-------A--A-----T---------TG------G--CA----T--T--A--------A--T-------------- 2544
H1O.BE.87.ANT70            --A-----G--------G--A--GC----T--C--------G-GG--T--G--GC-A--GC----T..................--T-----CT-A-----A--A-----T-----C---TG---------CCC---T--T--A--A--------T--C--T--T----- 3005
H1O.CM.91.MVP5180          --A-----T--------G--G--------T--T--------A-GG--T-----GC-A-GGC----T..................--T-----CT-A-----A--A-----T-----C---TGC--T------CC------T--A--A--C-----C--C--T--T----- 2980
CPZ.CD.90.ANT              --A-----T--T-----GA-A-----------T--T-----A--C--A-----GC-A--G------..................--G------T----------A-----C-----C--CA-A---------C-G-----T--A--A-----A--T--C--A--T--A-- 2389
CPZ.CM.01.SIVcpzCAM13      --C-----T-----------A-----G--C-----C--C--A------------G----G------..................-----------A--------------------C--TTG---TC-----A-G-----------------A-----C--A--T----- 2530
CPZ.CM.05.SIVcpzEK505      --------T--T--------A-----------C--------A------C----G------------..................--------G--A-----A--A-----T--C------TG-------------A-T--T--A--------A--T-------------- 2522
CPZ.CM.05.SIVcpzMB66       -----------T--------C--GC----------C-----T------------C-----------..................--------TT-A-----A--G--C--T--C------TG----C-G------A----T--------C-----T-----------C-- 2496
CPZ.CM.98.CAM3             --A--G-----T--------C--G-----C--------C--G-----T--------------G---..................--------G--A--------A-----T--------CTGC--A---------A-T--T-----A-----A--------T-----C-- 2361
CPZ.GA.88.GAB2             --A-----T--T--------A--G-----C--T-----C--T-----T-----G--A--G------..................--------G-C---------A-----T--------CTGC------------A----T--A--------A--C--C-----C----- 2343
CPZ.CM.98.CAM5             --A--G--A--------G--C--------C--------C--A-----AC----G--A--G--G---..................-----------A--------A-----T-----C--CTGC--A---------A-T--T--A--A--C--A--T-----T-----C-- 2652
CPZ.CM.05.SIVcpzMT145      --------A--T--------A---C-G--C--------C--A--T--------G------------..................--------CT----------A-----C---------TG---TC-T--CCC------T--A--A-T---------C-----T----- 2492
CPZ.TZ.00.TAN1             --------T--T-----GT-A--GC--------C-----G----------G---A-G---TATG..................--G--------------A--G-----C--C-----CA-----C-G---CC---------A--------A--T-----------C-- 2594
CPZ.CM.05.SIVcpzLB7        -----------T-----------------------T--C--T--G--A-----G-------GG---..................--C------T-------A--A-----C--C------TGC---C--------A----T--A--A-----------C----------- 2505
CPZ.US.85.CPZUS            --A--------------------G-----C--------C--A-----A-----G-------G----..................--------C--A-----A--G-----C--------TTG---AC-G---A-G--A--T--A--A--C--A-----------C----- 3010
CPZ.GA.88.GAB1             --A--------T--------G--G-----C-----T--C--A------------------------..................--G------T-A-----A--A-----C--C--C--TTG----C-G---A----T--------A--------------A--T--C-- 3011
MON.CM.99.L1               -----G-------TT--G--G--G--G-----C-----C--A--G--AC-TG-G----TGG-CCAT..................--G---A--G-A--CA--AAG--C--C-TC-ACAGCA-------GTGG---CCG-----A--A--C--G--G--CT-A-----CTC 2656
MON.NG.x.NG1               --A-----T----TT--G--A-----------T-----C--G--T--AC-CC-G--C-TGGC-AAT..................A----TA--G----CA--AAA-----C-T--ACAGCA-C--TC-G------A-T--T-----A------------T-A------TC 1249
GSN.CM.99.CN166            --G--------GGCT--G--C--G--G-----T--G--C--A-----AA-TCCGC-CCCTGC-GG-CTAGAGCAAAAGGAGCAT-----TA--A-A---A--AAA-----T----A-AGTA-A--A--GT----G--T-----------C--C--T---T-AG-----TC 2678
GSN.CM.99.CN71             --G--------GGCT--G--C--G--G--G--T--G--C--A-----AC--..................GAACAAAAAGAGCAC-----TA--A-A---A--AAA--C--T----ACAGT--A--A--GT-------A-----A-----C---------T-AG----ATC 2660
MUS_1.CM.01.1085           --C-----AC--GCA---------C-------C--------A-----A---..................ATAGAGCATGAGCAC--T---A--G-----A-CAAG--C--C-T--ACAGT--C---C-----CC------T--------C--A--C---T-GC-G---CA 2640
MUS_1.CM.01.CM1239         --C-----G---ACA--G--A--GC----G--C--------A-----A---..................GCGGAACATGAGCAT--C---A--G-A---A-CAAG--C--C-T--ACAGC--C--A------CC--CA-----A--A--C--A--C--CT-GC-G--CCA 2630
MUS_2.CM.01.CM1246         --A-----A--TGCA-----C--GAC---G--C---------T--------GCCC-A--G.........GCA.........CAT--C--TA--G-A---A--CAG-----C-T--A-AGT--C--A-----CA----A---C-ACCC--C--G--C--CT-GG-C---GC 2738
MUS_2.CM.01.CM2500         --A----------CA-----C--GAC---G--C--G--C--AT-----C-TTCT--A--G.........GCT.........CAC--C--CA--G-A--CA--CAA-----C-T--ACAGTA-C--AC-CT----G--G--T---CCT--C--A--C--CT-AG-T----- 2697
RCM.NG.x.NG411             --A-----T----TT--G--A--GC-G-----C--G-----A-G---A------C-ATGC-GGCA-..................A-T-----CT-A--A--T--G-----G---------TGC---C--T----GCCA--T--A-----C-----T--C--G----TAGC 2444
RCM.GA.x.GAB1              --A--G-----T-TT--G--G--GC-G--------------A-G---A-----GC-ATGTG-GAG-..................A----G---T-A--CA-C--G---------------TG--TTC-GT----GCCT--T-----A--------------A-----AGC 2429
DEN.CD.x.CD1               --C--G--------------A--------T-----------T-----A---C--C----GG--CAC..................A-C--T--GT-A--CT--AAA----------A-AGT--G--T--G---CC---T--TGCAC-A--C--A--------AG-G--GTC 2837

MAC.US.x.239               -----G-----TAC------C-----------------C--T-----AC--GC---A-GG---AG-..................A-T------------A-A--------------C--CA-A--TC----------A--T---C----C-----C-----TT----ATC 3252
Pol                        _V__T__Q__D__F__T__E__V__Q__L__G__I__P__H__P__A__G__L__A__K__R__K__.__.__.__.__.__.__R__I__T__V__L__D__I__G__D__A__Y__F__S__I__P__L__D__E__E__F__R__Q__Y__T__A__F__T__L__P
H2A.GW.x.ALI               --------T---ACA--GA----G-----G--T--------A-----A---GCC-------G-AG-..................A-C--T--G--------A--G-----T--C-----CA-A--AC-GC----GAG---T--AC----------------TC----ATC 3328
H2A.DE.x.BEN               --------T---ACA---A----GC-------T-----C--G-----AC--GCC--A--G---AGG..................A-CT-TA--T-A-----A--G-----C--------CA-A--AC--C-------T--T---C-----------------C----AGC 3335
H2A.SN.x.ST                ------------ACA---A-C--G--------T-----C--A-----AC--GCC-----G---CG-..................A-T--T--C--A-----A--G-----T--C-----CA-A--AC--C----G--T--T--AC----------------TC----ATC 2780
H2B.GH.86.D205             --C--G------ACA-----CA-C-.G--T--T--C--C--G-----TG..GC-G-A----GGAG-..................A--------A-A--------A-----C--C--CAGTA-C--AC-----CC-A-------AC-------A-----C---T-G--ATC 3308
H2B.CI.x.EHO               --C-----A--TACA--G--C--GC-G--T--T--T--C--A-----AC-GGC-TCA--G---AG-..................A----------A-----A--A-----C--C--CAGT--C--AC-----CC---------AC-A-----A--------TT-G--AGC 3306
H2G.CI.92.ABT96            --C--G-----TACT-----C--G--G--C-----T--C--T-----T--GGC-G-A----GGAG-..................A--------------A-A--G-----C--C--CAGCA----TC------T---T------C----C--A--T--C--TT----ATC 2662
H2U.FR.96.12034            --------T--TA-T-----C--G-----T-----C--C--A-----AC--GCTG-------GAGG..................A-------T--A-----A--G-----C--C--CAGCA-C--GC-GTG-------------C-A--------T-----TT----ATC 2817
MAC.US.x.EMBL_3            -----G--AC-T-ACAG-AG--C-A-TA----------C--T-----AC--GC---A-GG---AGG..................A-T------------A-A--------------C--CA-A--TC----------A--T---C----C-----C-----TT----ATC 2726
SMM.US.x.H9                --------T--TACA-----A--G-----G--------C--TR----AC--GC-----G--G-AGG..................A-C-----GT-------A--------------C--CA-A--T-----------A------C-A--C-----C------T----ATC 2725
SMM.US.x.PGM53             --------T--TACA-----A--G-----G--------C--T-----AC--GC-----GG-G-AG-..................A-C------T-------A--------------C--CA-A--TC----------A------C-A--C-----T-----TT----ATC 3169
SMM.US.x.SIVsmH635F_L3     -----------TACA-----A--G--------------C--T-----AC--GC-----G--GGAGG..................A-C------T-------A--------------C--CA----TC----------A------C----C-----C-----TT----ATC 3255
SMM.SL.92.SL92B            --A-----A---ACA-----A--GC-G--T--C--------G-----CC----GG-A-TGG--AGG..................A-------G--A---A-A--G-----C-----------A---C----CCCG--A--T--AC-A--C--A-----C--T-----ATC 2683
STM.US.x.STM               --A-----T--TACA---A----GC----T--------C--T-----AC--GC-----GG-G-AG-..................A-C------T-------A--G----------------CA-A---C-------G-G---T---C----------T--C---T----ATC 2899
MNE.US.x.MNE027            -----G-----TACA-----C-----------------C--T-----AC--GC---A-GG--GAGG..................A-C--------------A-----C---------C--CA-A--TC----------A--T---C----C-----T-----TT----ATC 2720
DRL.x.x.FAO                --A--G-------TT--G--G--G--G--G--------C--G-G---C------GCATGCG--CAG..................A-C--T-----------A--------C-----C---TGC--TC-G---CCG-----T--A-----C--------C-----T--ATC 2517
MND_2.x.x.5440             -----G--T----TT--G--G--G--------CA-------A-G---CC----GC-ATGTG-GCA-..................A-T------T----CA-A--A-----C---------TG------G----TG-----T--A-----------G--C-----T--ATC 2511
MND_2.CM.98.CM16           -----G-------TT--G--G--G--G-----C-----C--T-G---TC-C---C-ATGTG-GCA-..................A-T-----------CA-A--G-----C-----C---TG---T--G-----G-----T--A-----C--------C-----T--ATC 2951
MND_2.GA.x.M14             -----G--T----TT--G--G--GC-------C--------G-G---C------C-ATGTG--CA-..................A-C---A--T-----A-A--A-----C---------TG---T--G-----GAG---T--A-----C--------------T--ATC 2878
OLC.CI.97.97CI12           --A-----T--------G--A---------------T-----A------------G-A-GGG--CAG..................A-T------T-A---A-A--A-----T-----C--TA-A--GC-----CC---A-AT--AG-A-----A--------A--T--ATC 2824
GRV.ET.x.GRI_677           -----------T-TT--G-----GC-G-----T--T--C--A--G--CC-TC-G--A--G--GCA-..................A-C------A-A--CA-A--G-----C----A----A-A--A---TGCA-G--A-----A--A-----A-----------C--CTC 3123
VER.KE.x.TYO1_patent       --C-----T----TT-----G--------G-----C--C--A-----A----G-----TG-G-CAG..................A-------TT-A-----A--A--C--C----A---CA-A--A--G---CC-A-T--T-----A--------T-----T--T--C-C 3177
VER.KE.x.9063              --A----------TT-----------G--------T--C--G--------GC-G--A-TG---CAG..................A--------T-A---A-A--G-----T----A---CA-A--A------CC---G--T--A--A--C--A--------A--C---TC 3180
VER.KE.x.AGM155            --A----------TT-----C--------T-----C-----A-----------G--A-TG--GCA-..................A----CA-TA-A--------G----------A-AGCA-A--AC-G---CCT--G--T--A--A--C--A--T--C-----C----C 3172
VER.DE.x.AGM3              --A-----T--T-TC-----A--G-----C-----T-----AT-------CG------TG-CGGA-..................A--------T-A--CA-A--G-----C----A----A-A--A-----CCC---G--T--A--------C--T--------T--ATC 2675
TAN.UG.x.TAN1              --A--G-------TT--G--A--------T--T--C-----A-----A--G----GA--G---CAG..................A-C--------A---A-A--A-----C----A-AGTA-A--A--G---CC---A------CCC-----A--T--------C--ACA 3152
SAB.SN.x.SAB1C             -----------TCA------C--G-----G-----C--C--A-----C---C-GC-A-GGG-GCAG..................A-T-----GT----CA-A--A-----C------AGCTGC--T------CC------TCA---A-----A-----C-----C--ATC 3348
MND_1.GA.x.MNDGB1          --A--G--T--TCAT---T-A--G-----T-----T--C--A-----AA--------TGT---AG-..................A-------C--A---A-A--G-----C------AGTA-A--TC-G---CC---T-AT--ACCC--------C-----GG----ATC 2637
DEB.CM.99.CM40             --A--G--GG-G-----G--G--GAC---T--------C--AT-T-CAC--CC-C-A-TGG--CAG..................A-C--T--CT-A---C-T-CA-----C----A----A-AG-AC----CCC-ACA--TGCA--A--C--A-----C--AG----ATC 2759
DEB.CM.99.CM5              --C--G--TG----------G---AC---G--C--G--C---T-G-----GCC-C-A-TGG--CAG..................A-------C--A--CT---CA-----C----A----A-A-----G---CC-C-G--TGCACGT-----A-----C--A------TC 2753
TAL.CM.01.8023             -----G-----TGCA--G--C---CAG--G--------C---T-T--AC--CC-C-------GCAG..................A-------CT-A---A--AAG-----C----A-AGCA-C--A------CC-------GCCCGT--C-----T--C-----C--GTC 2417
TAL.CM.00.266              --C---------GC------A---CAG--G-----T--C---T----CC-GCCTC---TG--GCA-..................A----G--C--A--CA--AAG-----C----ACAGCA-C--------CCC------TCAC-----C--A--C--C-----C--GTC 2919
SUN.GA.98.L14              -----------TCAT--GT-A--G-----G--T--T-----A-----A-----G---TGT---CAG..................A-T------G-A--------G--------C--C---A-A--A--G---CC---T-AT--A--------A-----C--A-----GTC 3307
WRC.CI.97.97CI14           --A--G-----T-TT-----A--G-----------T--C--T-G---AC-G-G------T--GAAC..................A-TT-----G-A---A-A--G------------C--TA-A--A------CC---A-AT---C-A--------T-----AG----ATC 3250
WRC.CI.98.98CI04           --A--G-------TT-----A--G--G--C--T--T--C--A-G---AC---G------T---AAT..................A-TT-----G-T---A-A--G-----C---------CA-C--AC----CCC-----AT---C-------A--T-----TG----ATC 3265
WRC.GM.05.Pbt_05GM_X02     --A----------TT--G--A--------C--T--T-----A-G-------C-T-------G-AAT..................A--------T-A---A-A--A-------------TTA-C--A------CC---A-AT--AC----C--A--------AG-----TC 2179
LST.CD.88.524              --A--------TCAT--GC-A--G--------------C--A-----AA----G--ATGC---CA-..................A-T------G-A---A-A--A-----C-----CAGCA-A--T------CCT--T-A---AC-A--------G--C--T-----GTC 2215
LST.CD.88.447              --A--------TCAT-----A--------C--T-----C--T-----AA----G--ATGC--GCAG..................--------GG-A---A-A--A-----C------AGCA-----C-G---CC-----A---AC----C--A-----C--AG----ATC 2218
LST.KE.x.lho7              -----------TCAT--GC-A--------T--T-----C--A-----TA-C--G--ATGC---CAG..................A-T------G-A--CA-A--A-----C-----CAGTA-C--T------CC-A-T-A---A--------A--------T------TC 3302
SYK.KE.x.SYK173            --A-----T--T-TC--G--A-----G--------C--C--A-G---AC-----CG-C-----.....................C-----A-TA-A---T-AAAA----------AC-----C---C-G--CA----G--T---CCA--C--G--------AG-G--A-- 3144
SYK.KE.x.KE51              --A--G-----T---------A--G----------T-----A-G----C-C-GT-CA----G-.....................A-G------T-A--CT--AAA--C--C----ACA----A--TC----------G-----AC-------A--------AG-G--C-- 2802
COL.CM.x.CGU1              --A---------AT------C--------G--------C--A-G---TC-T-T-G-A--G-CCCAT..................A--------T-A---A-A--A-----C-----C---A-------G---CCG--T-ATCAA-----C--A--T-----A--C--ATC 2725
```

```
H1B.FR.83.HXB2         TATAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTATGTAGGAT 3120
Pol                     I  N  N  E  T  P  G  I  R  Y  Q  Y  N  V  L  P  Q  G  W  K  G  S  P  A  I  F  Q  S  S  M  T  K  I  L  E  P  F  R  K  Q  N  P  D  I  V  I  Y  Q  Y  M  D  D  L  Y  V  G
H1A1.UG.85.U455        --------------------AG-C--G-------------------G-------------------T---------G-------------------------C-----TC----C-------------------------------C--------G--- 2566
H1B.US.90.WEAU160      ------------A------------------------------------------------------------------------------A----------------------------------------------------------------------- 3119
H1C.ET.86.ETH2220      --C---------A-------------------A--T-----C--C-------------------C---------G---------C-CC-------------C-----GGCC-CC--C-----A---------------T--------C------------- 2512
H1D.CD.84.84ZR085      -------------------------------------------A---------------------------------------T--------------C-----------------AG-------------------------------------- 2642
H1F1.BE.93.VI850       -G-C--------------T--A-----G--C---------------------A-------------------------T------------------------C------TGA----C-----------C-----C-----------------------G- 2457
H1G.SE.93.SE6165       ------T--------------G----------------------------------------------T--G-----------G-----------C------GC-A-T--------A--G--G-----C-----------------A---------- 2517
H1H.CF.90.056          ------------------------------------T--------T---------------------------------------------G--------------C---C-----G----------T--A--G-----T--C------------------------ 2467
H1J.SE.93.SE7887       ------------------------C-------------C-----A-----------------C-----------T--GT-------------------A-A---------G--AG---C-----A-----------C--G----------C--------G--- 2434
H1K.CM.96.MP535        ---T--------A---------G----------------A-----A-------------------------CA---------------------------C-------T-A----------A--G-----A--C-----------------A--------G- 2316
H101_AE.TH.90.CM240    ---------------------A--C--------------G--------------------G----------G---------------------------C-------T-A----------A--G----------------A------C------------- 2694
H102_AG.NG.x.IBNG      -G----T-------------------C----------------------------G------G---------T--GGCA--------------------C-------C-A----------G-----G-----C-----------------A---------- 2645
H1N.CM.95.YBF30        ------------------T--T--------C-----T--------G-----A--C---------G---------T--T--G----CA-----------TC----A--A--C---G-GA--C--------G---A-C--T--C--G-----------CC-C-----G---- 2714
H1O.BE.87.ANT70        -G-G--T--------C-----A--A-----C-----------C--C--G--A----------T--------T------------TCA-----------TC----T--A--C--G-G-G-C--C-----AT---AA--T-G---G-----------CC-A---------- 3175
H1O.CM.91.MVP5180      -G-G-----------C-----AG-A-----C-----------C--C--G--A--G--------T--------C-----T--G---TCA--------G--TC----T--A----------AGC--C-----AG---AA--T-----G-----A-----C--A---------- 3150
CPZ.CD.90.ANT          CG-G-----------C-----A--A------TGT--T-----T--A-----A--C------------------T--T--T---GCA--------C--G--------CA--A-----GG-TA-GT------CAG---AA--T-----------------C-C----------- 2559
CPZ.CM.01.SIVcpzCAM13  -G----T-----A--T--------C-------------T-----TT-A-----A----------G--T-----T--C--T--------T-----------TC-G--A--C--C---GC-A----C--T--T--CAC-G---C--------------C--A--------G- 2700
CPZ.CM.05.SIVcpzEK505  -G----T-----A-----T--T--------C-----------------G-----A--C--------------------C--T--G--C-CT------------A-----A--A-----G--GA-C---------AC----C--A-----G-----------C-C--------C- 2692
CPZ.CM.05.SIVcpzMB66   -G---------------------A---------C-----T--------A--G--A--G-------G------------T---------------------T-----------C--------G--------A---A--------------T-----------A--------C- 2666
CPZ.CM.98.CAM3         -G----T-----A---------A-----------A--T------T-A-----A----------G---------T--T-----C-------T-----TC--ACC--A--C---C-G--GC--------T--A-A--T--C--G--T----------C-T-----G---- 2531
CPZ.GA.88.GAB2         -G----T-----A--T-----T--A--G--C-----T------T--A-----A----------G---------C---------C----------GG-T-----C-----C---C-GA----C-----AT-G--A--T--C---------------C--A---------- 2513
CPZ.CM.98.CAM5         -G----T---------------AG--------------T------T-A--R--A----------G-----G-----C--T-----C---------------A-----T--A--C--------GC-------TC--A-A-----C-----T---------CC-T---------- 2822
CPZ.CM.05.SIVcpzMT145  C-----T-----A--T-----A-------------A--T-----A-----T--A---------G-----T--TT-C--C--T--G---------------G---------A--C--------C-A-GC-------A---A----T--C--G-----------C--------G--C- 2662
CPZ.TZ.00.TAN1         -C-C--T---A-C-----------AA----T------T--C---T-A--T--A--T-----G-----T---------T--T--G--C--T------------C-----T-----C-------G--C-C-----TG-G-AC--T--C-----T----------C-T--CA----T- 2764
CPZ.CM.05.SIVcpzLB7    ------------A---------A--C----------A--------T--A--G---------------------T-----T------------------G--T---------A--C----G--G--C-----T-----A--T-----G---------------A--------G- 2675
CPZ.US.85.CPZUS        -G-----------C-----A---------C-----T-----T-A-----A----------G-----------T-----G--C--T-----------TC---T-----------------C-----TG-TA-A------------T----------C-C--------G- 3180
CPZ.GA.88.GAB1         ------T--------C-----AG---------A--T-----TT-G-----A----------G--T---T-T--C----------C--T-----------TC----A--C--C---G--A-G-----T-----TAC---T--C--G-----------CC-A-----G--G- 3181
MON.CM.99.L1           -T---------CTGA---T-CT-AA--G-------T---G--C--C--G-----C--------G--C---A-TG-G--T---CA--CAGC-G--G-GC-ACC-C--GAGA-C--G---A--T-C-----GG--AC-T-AATA--G-----------CC-ACTGA------ 2826
MON.NG.x.NG1           AG----T----CAGA-----CCG-G----------T--GA---T-G--C-----R-----------C---A---------T--GTACTCAGCAT--G-GC-GC-RA-AGRAA-C--G-G-A--T-------GG-CACCC--ATC--G-------------C-ACTCA----T- 1419
GSN.CM.99.CN166        AG----T---C--G-C----CTGAA--G-------T---A--CT-A--G-----T-------CG--G---A-C--C--T----AC-TTG-AG-TTCCT-GC--CGA-AGA-----G-T--GG-G-----TG-G--AT-AATA--G---------------ACTGA------ 2848
GSN.CM.99.CN71         AG---------C----C----CTGAG--G-------T---A--TT-A--T-----T-----G-CG--T---A-C--T--T--GGCC-CTG-AG-TTCCT-G---T-T-AGA-----G-C--GG-G-----TG-G--A--AATA--G-------------C----TGA------ 2830
MUS_1.CM.01.1085       GG-------CTT-GG-----CA-GA---------A-T----A-----A---------C--------G--T--TA-C--C-----GTTC-CAGCAG-C---C-------AGAGA---AG---GC---C-----TG-CAC-C--ATC---------------CC-CCTGA------ 2810
MUS_1.CM.01.CM1239     GG----T--CC---GG-----CA-GA--G-------T----A-----G----------G-----------C--CA-C--C-----TAC-CGGCAG-----T-A---C--GAAA-C--GG--GC---------T--CACCC--ATC--G-------------C-C-TAA------ 2800
MUS_2.CM.01.CM1246     AG-G-----CAT-GG-----CT-AA----TC--A-TT--G---T-G--------G-----G----------CA-C--T--T---GCC-CA-CAG-C--GC-TC------AAG-----G--GG-C-----AG---GCCT-A--AGTC----------------CC--CT-A----G- 2908
MUS_2.CM.01.CM2500     -G----T--CCTAGGG--C-C--AG--G-----A-T---G---C--G--C--A--G-----G--G--C--CA----T--T---GCA-CA-CTG-----T-A-------AAA-C--G---G-----AG--TGC-T-A--AGTA-----T----------CC-ACTCA----G- 2867
RCM.NG.x.NG411         ---T--T---C-AGG---T--C-----------A--T------T-G--C------------G-----T-----T-----T----CA-CAGCT-AC-----T-----A--A-----G---G-G--------TG----C-------G--T--------------T------G- 2614
RCM.GA.x.GAB1          AG----T---C-AGG------AG-G--G-----A--T----------G--T-----------G-----T--C--C--C--T--GGCATCAGCC-AT--G-------C----A-----GG--G-G--------TG-CA-C--T--C--G-------------C-C-T------C- 2599
DEN.CD.x.CD1           C--T---------AGG------ACAA--G--CATT-T---A--CT-A-----A------GT------G--TA-C--T--T--GGCC-CAGCAGAT-----GC-TAGC-AA-----G--GGCCC-----T--TG-GA--T-AATC--G--T----------C-C-TG----C-- 3007

MAC.US.x.239           AG----T----CAGAG-----A-AAC----CATT--T--G--T--G--T-----------G--G--------C--C------TAC-CT----G-C-TG-GC----A--C--C--G--GGC----------TG-GACCT-AGTC--G--T---------CA-C-TAA---CTA 3422
Pol                      S  V  N  N  A  E  P  G  K  R  Y  I  Y  K  V  L  P  Q  G  W  K  G  S  P  A  I  F  Q  Y  T  M  R  H  V  L  E  P  F  R  K  A  N  P  D  V  T  L  V  Q  Y  M  D  D  I  L  I
H2A.GW.x.ALI           AG---------CAGA------A-AA-----ATA--T--A--CT-A--G-----------G---------------T--T---CAC-CA----G-C-G------------A--C-----GGC---C-AG--TG-CA--C--AT--------------A-C-TAA---CTA 3498
H2A.DE.x.BEN           AG--------AT-GA------A-AA-----ATA--T--A--CT-G-----A---------G---------------T--T---TAC-CA----GGC--G-------A-----C------GC---C-----TG-CA--C--ATC--G-------------A-C-TAA---CTA 3505
H2A.SN.x.ST            A----------CTGA------A-AA-----CATA--T--A--CTCA---------------G---------------T--T--GTAC-CA----GGC-GG-------A--A--C------GC---C--G--T--CA--C--AT---G-------------A-C-TGA---CCA 2950
H2B.GH.86.D205         AG---------CAGAG-----A-AG-----CATT--T--A--T--A-----A--G-----G--G--C-AGT----T-GT---TACTC-----G---GG-A-----T-----------GGCC--CAGC--TG-CA----AAT---G------------CA-CCT-A---C-A 3478
H2B.CI.x.EHO           AG----T----CAGA------A-AG-------TT-----A--C--A-----------------G-----C---------T-----GTAC-C----G---GG-AC----C-----C------GCC--CAAT--TG-CAC---AATC--G------------CA-TCTC--G-C-A 3476
H2G.CI.92.ABT96        AG---------CAGA------A-AA-----CATC-----G--C--A--G--A---------G--C--------T--C------CA--CA----GG--GG-AC----C--C--C------GCC---GAG--TR--A----YATC-----T---------CA-TCT---G-CCA 2832
H2U.FR.96.12034        C--T--T----CAGAG-----A-AA-----CATC-----G---T-G--T--A--C-------------------A----T--T--GTA--CA----GG--CC-AC-G-----A--------GGC----GA---TGCT--C--AAT---G--T---------CA-TCT-A---CCA 2987
MAC.US.x.EMBL_3        AG----T----CAGAG-----A-AAC----CATT--T--G--T--G--T-----------G--G--------C--C------TAC-CT----G-C-TG-GC----A--C--C--G--GGC----------TG-GACCT-AGTC--G--T---------CA-C-TAA---CTA 2896
SMM.US.x.H9            GG---------CAGA------A-AA-----CATC--T--G--AT-A--T--A--G------------------T--T--T--RCA--CTR---G---TG-A-----A-------------GC----------TG-GACCC-GATC---------------CA-C-TAA---CTA 2895
SMM.US.x.PGM53         AG---------CAGA------A-AA-----CATC--T--G--AT-A--T--A--G--------G---------T--T--T--GCA--CT----G---TG------GA-----C-------GC----------TG-GACCC-GATC---------------CA-CCTAA---CTA 3339
SMM.US.x.SIVsmH635F_L3 AG----T----CAGA------A-AA-----CATC--T--G--AT-A--T--A--G-----G--G---------T--T--T--GTA--CT----G---TG-A-----A-----C-------GC----------TG-GACCC-GATC---------------CA-C-TAA---CTA 3425
SMM.SL.92.SL92B        AG-G-----CC-AGA------A-AG--G---ATT-----G--C--G-----A---------G-----T--T-----T-----GGCA-C-----GGC-GG-A-----A--C--C------GCT--C------TG-GC--C--GTC--G--T-----------ACTAA-T--CA 2853
STM.US.x.STM           AG---------CAGA------A-AA-----CATC-----G--T--G---------------G---------------T--T---TA--CA----G---T--A---------A--C--G-G-GCC--C------TG--AC-C-GATC-----------------A-C-TGA---CCA 3069
MNE.US.x.MNE027        AG----T----CAGA------A-AAC----CATT--T--G--C--G--T-----G-----G--G--------C--C------CAC-CT----G---TG-GC-G--A--C--C--G--GGC----------TG-GACCC-AGTC--G--T---------CA-C-TA----CTA 2890
DRL.x.x.FAO            GG-G--T---C--GGC-----A--C--G--------T-----C--C-----------------C-----T-----C------GC--CAGCTGAC--G------C----A--C---G-GA------------TG-GA----A-----G-------------C---T-------A 2687
MND_2.x.x.5440         GG-G-----CC--GGC-----A--C-----C-----T-----C--A-----------------C--T---------T-----GGCA-CAGCAGAC-----TC--A--G------AGG--A-GC-------AG-CT-A---------G--T----------C----T---G--GA 2681
MND_2.CM.98.CM16       GG-G--T---C--GGC-----A--C--G--C-----T-----C--A--------------G--C--C-----C--T-----GGCA-CAGCAGAT---------C----C--C---G-GAG-C-------TG----G------------T----------C-C-T---T--GA 3121
MND_2.GA.x.M14         GG----T---C-AGGC-----A--C--G--C-----T-----T--------------------C--T--G---------T---GCA-CGGCAGAC-----TC--A-AA-A--C-A-G-GG--T-C-----AG--T-A--T-------------------C---T---G---A 3048
OLC.CI.97.97CI12       AG----T-----T-G------A-GA--G---ATA--T------T-A--T--A---------G--G------TGT--C-----G----CAG--G-TC-TT-GA-G-------GG--G--GG-------T--G----A-T-A--C--G--T----------C---------G---- 2994
GRV.ET.x.GRI_677       AG----T---ACAGGG-------A--G-------T----TGT--G--T-----------------T--TA----T-----G-A--CGGCAG----C--T------GAGA-C-A--GG--C-C---T-GGT---AA--TGTC------------C-------GGT-G-CG- 3293
VER.KE.x.TYO1_patent   AG-G--T---C--GG---C--------G-----A-T---CTGT--C--G--A--G-----------T--TA----C-------A--CAGCAG--TCC--T--G---GAGA-A-A--G-A-CTTG----CAC--ACC--TGTA-----------------A-GG-----T- 3347
VER.KE.x.9063          AG----T---C-AGG------TG----G-----A-T----TGC--------A--C--------G--T---A-T--T--T--G-A--CAGCAG-------T------AAAA-A-A-G-GG--TTAAA-CCAT-GACA--AGTG--G-----------------GG-----G- 3350
VER.KE.x.AGM155        GG-----------GG------C--A---------A-TT---TGC--A--G-----C-----G-----C--GA----T-------AC-CAGCAT-------TC----AGAAA-A-AG---G--TTAAA-C-GC-GACG--TGTC--G------------CC-C-GG------- 3342
VER.DE.x.AGM3          AG----T---C-AGGG-----T-C-------------T---CTGT--------A---------G-----C---A-T--T--T--G-AC-CAGCAG-TTCC--TC----AGAAA-A-A---GG-GTTAAA-CC-C--ACC--TGTG--------------CC-A-GG-----G- 2845
TAN.UG.x.TAN1          AG----T----TAGGG--T--A-----G-----A-T----TGT--A--T--A-------------------CA-C-----T--GCAC-CAGCACAGCG-G--C--A-AGAAA-AG-----A-GGT--A---TT---AA--AAT---G------------CA---GGA------ 3322
SAB.SN.x.SAB1C         AG-C--T---AG-GAG-----C--A-----------T--G-----C--------G--------C-----CA----T--T--G-CA-CAGCC-AC-----TC-GC--GAA-----GC--A-G--C-----TG---A---A-----------------A---TAA---CCA 3518
MND_1.GA.x.MNDGB1      AG-T--T---C-AG-------A-AA-----CAT---T-----T-----T--A--G-----G---AGT---TGT--C--T---G-G-CAG-AG--TC-C-GC-G---GTA--------GA-CC-----ACAG--CAGT-A------------------------T------G- 2807
DEB.CM.99.CM40         -G-C--T---ATACA------AGAA------G---T--GA--A--G--T----------GC-CT--------C--T--T---GCCTCAG-AGGG-G-CAG---C--ATC-----GG--A-G---------GC--A-A--AGTC--G--T------------ACTA-----G- 2929
DEB.CM.99.CM5          GG-G------ACACA------AGAG------G---T--GA--A--A--T--A--C----GT-C----------C--------GGCA-CAG--GGC-GGCAGC-GA-A-T-----AGG--A--T-C-------T--A----AGTC-----T---------CC-A-TG------- 2923
TAL.CM.01.8023         -C-G--T-----C-G---------GGC-G--C----T---G--A--C--C--A--G--------------CA-C-----T---TAC-CTTCAGATC-GC-G---AG-AAG-----GG--A--T-C---C-GG---CGC---TC--G-----------CC-ACTAA-----A 2587
TAL.CM.00.266          -C-G--------T-GG-----A-GA--------A-T---G-----A--C--A---------G--T--G--CA-C--C-------CA--C-TC-GATC--T-A--GAG-AAA--C--GG-CA-GT-C--TC-GG---CAC---T---------------CC-A-TG------A 3089
SUN.GA.98.L14          AG----T---C--G-------A-AA-------GT---T-----C--A--------C----------AGT--TTGC--C--T--GG-A-CGG-AG-G-GCC--------AAA-------G---GT-------AG-TCAAT-G--C-----------------ACTCA------ 3477
WRC.CI.97.97CI14       AG----T------T-T-----AG------GTA--T-----A--G--T------------------T--TTGT--T-------TCA-CTG-AG--TCTC-A-----AATA------GG-A-G------A-T---CAAG-TCT------T---------C-A---A-T---- 3420
WRC.CI.98.98CI04       AG----T-----AT-T--T----AG--G---GTC--T-----A-----T--A--G-----G--T--C--TTGC--T------GCA-CAG-AG---GTT-G-----TGTA--C-AG-----G-----TACT--GCAG--TCTC--G--------------A---A-----A 3435
WRC.GM.05.Pbt_05GM_X02 ------T--------A------T-CC------GTA--T-----A--C--T--A-----------G--T--TTGC-----T---GCA-CAG--G-TTCTT-GC-G---ATC--------GA--C----TA-AG-GCAG--AGTC-----T----------C------------ 2349
LST.CD.88.524          AC--------C--GA------A-AA-----------T------T-G-----A---------G--GAGT--TTGT--T--T---GAA-CAG-T-GTGG-C--C-T-CAGAG--C--G---TT-------------C-CC----C--G-----------------TCA----T- 2385
LST.CD.88.447          ---T--T--CC--GA--------GA-----------T------A--A--C--A--------G--GAGT---TGC--T--T---GC--CAG-AG--GG-C--C--TCAGAA----------TT---C--T------A-CT-G----------------------A-TCA------ 2388
LST.KE.x.lho7          -C-C--T---C-AGAG-----A-AA--------------------G-----A--G-----G--GAGC--TTGT-----T--GG-A-CAG-AG--GG-C-TC-CTCAGAG-----G---TT------------GA-C--T--C-----T-----------A-TCA------ 3472
SYK.KE.x.SYK173        C-----T----CCT-C--G--AGAA---------A-TT-CA---T-A--------------------T--CA-T--C---------C-CA--A-ATC-G--AC--C----A--------GA-GT--T----TC-CAC-C-AATA--G-----------C--A-TGA-T--CA 3314
SYK.KE.x.KE51          ----------TCTT-T-----AGAG-----C----T---G---T-G-----A---------G-----G--CA-T--C--T--GGCA-CAG-A-ATC-G--AC-CC-A---A-A--GG--A-G---T----T---A-A--AGTA--G-----------CC-T-TG-----G- 2972
COL.CM.x.CGU1          AG-G------TT-G---------C-------GT------A-----A--G-----T---------G--T---A----C------G-A-CAG-AG---G----C----A--G--C--G--GAG-GGA-A-TTG...CAA--AGC---G-------------A---C------ 2892
```

```
H1B.FR.83.HXB2          CTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAGCTGAGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTGCTGCCA 3290
Pol                     S  D  L  E  I  G  Q  H  R  T  K  I  E  E  L  R  Q  H  L  L  R  W  G  L  T  T  P  D  K  K  H  Q  K  E  P  P  F  L  W  M  G  Y  E  L  H  P  D  K  W  T  V  Q  P  I  V  L  P
H1A1.UG.85.U455         ----T--------------A------G------------AT-A---GCT-----A-----C------T-C-TT--C-----------G-------------------T----------G-------T-------C----------T---------CA------- 2736
H1B.US.90.WEAU160       ---------------------------------------------------------------------T-----------------------A-----C--------------------------------------------------------------AA------- 3289
H1C.ET.86.ETH2220       -------------------A------G-CCC------A---T-A---G------A--A-A------T----------------G-------------------T----------G-------T-------C--------------------CA------- 2682
H1D.CD.84.84ZR085       ----T-----------A---------G-----------A-AT-A---G------------------G-----------------------------------T---------------------------------G-----------T-----ACA------ 2812
H1F1.BE.93.VI850        ---------------------------------------T-A---G------AC----A-----T----T-------------------------------------------------GC------------------------G---------CAAT----- 2627
H1G.SE.93.SE6165        --------------------------G-------------T-A---G------A----AA----GT------------T-----------------------------------------A-----G----------C--------G-----A-----CA------- 2687
H1H.CF.90.056           -------------------A------G-------------T-A---GCT---T------AA------T----------------------------------C-----T----------A-----------------C---------------A--G--AAA------ 2637
H1J.SE.93.SE7887        -------G-------AA----------G-------A----------GG------A----A-------T------------T----------------------------T----------A-----G--T-------C----------C--A------CAA------ 2604
H1K.CM.96.MP535         -------------------A-C-----------------A--A--A---G------A----AA-----T------------------G-------------C-----T----------A-----------------C-----------G---------CAA------ 2486
H101_AE.TH.90.CM240     ----T----------------C------------------------A---GCT----A----C-----T---T-----------G-------G-------------------A-----------------C-G--------C----------AA------ 2864
H102_AG.NG.x.IBNG       ----T---------------------G-------------T----GG-----AC--AA-----T-----------------G--------------------T------------------------------C------G--AA--A--- 2815
H1N.CM.95.YBF30         -----------C---CA--A------GAGGC-G----A--C--T---G-T-----T----A-----CT----G--C--T----------------G-----C--G-----C--------A--------------A--C-----------C-----A---AA-T-A--- 2884
H1O.BE.87.ANT70         -A--T---CCCC-GACAG-A-------A--GG--T--ATT---T---G----C--A-ATCA------T-C--T--C--T-------G-----A--G--------C--T--A--------G-------G-----G-------A--C--------------T-C--CCAAT----T 3345
H1O.CM.91.MVP5180       -A--T---CC-T-G-CAG-A-------AG-GGG-C--ATT---T--GG-----T-A-ATCA------T-----T--C--T--T-----G--------G--------C--TT-A--------A-----G-----C--A--C--G---------------C--CCAAT----T 3320
CPZ.CD.90.ANT           ----TA-G-----TACTGCA------GA--TG-----AA----T---------T-ACA-GTC--------AGAG--T--T--------G-------------------T-AA--------A-----GT-A----A--C--------T------AAA---AA---A--- 2729
CPZ.CM.01.SIVcpzCAM13   ----TC-T--C--TACC--A--C---CAG--GG----A---T-A--G------T--C-TCAC-----GT----A--C-----T-----G-----A---------A--C---T-A--------G-----G--T--C--A---------------G-----A---AA---A--- 2870
CPZ.CM.05.SIVcpzEK505   ----TC--------ACA--A------GAGGC-G-G--AAGA--T--GAGT-----T----CT------T----A--C--T--------G----------------A-----TT-G--------A-----------------A--------------C---A-A---CAAT-A--- 2862
CPZ.CM.05.SIVcpzMB66    --------A------A--TA------GA----G-G--A--AT----GGCT--C--A----AA-----TT-C-----T--T-------G-----A--G--G--A-----T--C--------A----------C--A--C------------T---------CA-T-A--- 2836
CPZ.CM.98.CAM3          ----T--G---C--A-AG-A------GA---GG-----CTA--C-----G-----AC-T-CT-----GT----T--C--T--T-----G--------G--G--A--T--TT-G--------C-----G----------------G-------------G-----A---CA-T-A--- 2701
CPZ.GA.88.GAB2          -A--T---TCTT-G-AAG-A--C--G-A--G-G----AC-A--C---G-----T-AC-C--A-----TT--------C--T-----G--G--C--A--G--------T--T--C--------G-----G--T-----A--C--------------G-----A---CA------T 2683
CPZ.CM.98.CAM5          -----C-G--G-----AA-A------GA---GG----ATTA--C--------------C-T-CT-----TT----T--C--T--------G-----A--G-----A--C--TT-G--------A-----G---------A--C--------------G-----A---CAAT-A--- 2992
CPZ.CM.05.SIVcpzMT145   -A--T---CCT----A-G-A-------AG-G-G----A-----C---GC----------A-AA------T----A-----------G--G-----------------C-----T--C--------A-----GT-A-----A-----G----------C----AA---AAAT----- 2832
CPZ.TZ.00.TAN1          -A--TC-TA-TGA--A-G-A-----G-A-CTG---A--A-----------G--------A-CA------T-AGAG--C--T--------GT-----G--A------------A-G--------C-----G--A-----AA---------------T--AAA---CACAT-A--- 2934
CPZ.CM.05.SIVcpzLB7     -A--T--GA------A---A--C--G-AC--GG----A---------GCT--CT--C-A-A------GT-------C--T----------------G-----A-----T--C--------A--------------A--C-----------------C---CA-T-A--- 2845
CPZ.US.85.CPZUS         -A--TC--A-CT---AAA-------GGA---GG----ACT---C------T--T--C-T-CT------T-C--T--C--T--T--G--G------G-G-----A------T-A--------C--------------A-----G--------G-----C---CA-T-A--- 3350
CPZ.GA.88.GAB1          ----TC-T-----T-AT--A-------A---GG-G--A--A--A-----------T--C-T-AA-----GT-C--A--C-----------------A--G--G--A--C--TT-A--------A-----G-----C--A--C-------------C--------TCAAT-A--- 3351
MON.CM.99.L1            -AA-TCACTCTC-GCCAG-A--C--G-A--T-G---G--TCA-A---GCGACC--AC-A---AAA--GA--CAA-----CCCAG-G---TT----G-T--CTAC--GG-G-AA---T-A--G-----G--A--C--ATCCC-T----A-A-TGCCAAA----A---A--- 2996
MON.NG.x.NG1            -A---CATTCCC--CCTG-A--C--G-AG-T-G--ACA-CCA-T---ACG-TR--A-CAGAAAA---GA--CAA------ATAG-G---T-----G-CA-GTAC-----G-AG---T----G-----------C--CACCC-T--------C--AAAG----AA------ 1589
GSN.CM.99.CN166         -A--TAGGA--C-G--AG-A--C--GCA-GT-G-GC-CA-AA-A---A-TTTA--AACATCC-ATAACA-CCAA--C--C--GGCC-----C--ACC---TTA---GG--AAG---C-C--G-----GT-G------A-GGG-----GGC-C-----GG-G-AA--C--- 3018
GSN.CM.99.CN71          -A--TAGGA--T-G-CAG-------GCA-GT-G-GC-CA-AA-A---A-TTTGT--AC-TCC-ATAACA-CCAG--C--C--GGCC-----C--AC----TTA----G--AGA---C-C--G-----GT-G--C---A-GGG-----GGC-C--A--AG-G-A---C--- 3000
MUS_1.CM.01.1085        -CA--AG----T--TCAAGC--C--G-A-GT-G---CAC--A-T---A-CATG--C--A-A-CTT--GA--CAA--C--TC-AG-----TT----G----CTAC---G---AA---T-A--G-----G--A--C--CCGGGGC----A-A-------CG---CAT----- 2980
MUS_1.CM.01.CM1239      -CA--AG---G---TC-GGA-----G-G-GT-G---CAC--A-T--GA-CATG--C----AACT---GA-ACAA--C--G--AG-T--GT-C---------CTAC---G---AG---C-A--G-----GT-A--C--CAGAGGG----A-A-------AG--ACA--C--- 2970
MUS_2.CM.01.CM1246      -A---AG----------AG-A-------GGCT-G-GA-AA--A-C---G-T-TC--C-CA--TAA---GA-ACAG--C-----AG-C-----C---CC---CTAC--GG-AGAA---C-A--G--C--GT-A-----GA-AGG-----G-A-CA-A--AG-G-A-------- 3078
MUS_2.CM.01.CM2500      -A---AG----------G-A-----G-GGCTGG-CA-AAGAG-T---G-C-TC--AACA--CAAA--GA-ACAA--C-----AG----G-----C----CTAC---G-AGAG---C-A--G-----GT-A--C--A-GGGGT----G-A-C-G---AG---AA--A--G 3037
RCM.NG.x.NG411          -A--TAGGAC-GC-CATG-A-----TCAG-TG--CA-AAG---C---G-G------C---CC-----TT--GAG--T-----T--G---TT----G-CA-G-----C--TGAA--------C----TG-------A--C--------------AGAG--CAA-T-A-T- 2784
RCM.GA.x.GAB1           -A--TAG-AC--AGTT-G-A---GAC-A--TG--CA-AC-AT-A---G-T-----AC--TTC-----CT--GAG--C-------------TT----G-TA---------ATT-G--------G-----G--G--C--AA-A-GT-------------GAG--CAA-T-A--- 2769
DEN.CD.x.CD1            -A--TAG-TCCC---AA-----C--G-A--TGG-TA----A--A---G-CTTCT--GCAGTA-----CT--GAG--C--T--TGCC---TT---AG--AC---A--CA-A-AA--------G------T-A--C--CA-AC-C---CA-A-----AAA----AC------ 3177

MAC.US.x.239            G----AGGAC-GACCT-G-A---GAC-GGGT-G-TTTAC--TCA-AGG---TCT--AAT--CATA--GT--T-T--C-----AG-G---TTC--A-----T--C-----T-AA--------G--C---T-GTGG--AACA-------AGT-G--AAAG----A-T----- 3592
Pol                     A  S  D  R  T  D  L  E  H  D  R  V  V  L  Q  S  K  E  L  L  N  S  I  G  F  S  T  P  E  E  K  F  Q  K  D  P  P  F  Q  W  M  G  Y  E  L  W  P  T  K  W  K  L  Q  K  I  E  L
H2A.GW.x.ALI            G----AGGAC-GATTTAG-A---GAC-AGGTGG-CCT-C--T-A-AGG---TC--AAATG-CCTA---T--T----C-----TG-G--GTTC--A-----C-------A-AAA--------C----G---GTGG--AAC--------AGC-G--AAAA---CAAT----C 3668
H2A.DE.x.BEN            G----AGGAC-GGTTTAG-----GAC-A-GTGG-CCT-C----A-A-G---T---AAATG-CCTA-GT--T-T--T-----TG-G--GTTC--A-----C--------T-AA--------C-G------ATGG--AAC--------AGC-G---AAAC--CAA-----C 3675
H2A.SN.x.ST             GC---AGGAC-GATTTAG-A---GAC-G-GTGG-TCT-C----A-AGG---T---AAATG-CCT----T--T----C-----TG-G--GTTC--A-----C-------A--AA--------C---------GTGG--AAC--------AGC-G--AAGA---CAAT----C 3120
H2B.GH.86.D205          G----AG-AGTGATCT-G----CGAC-GGGT-G-GTCCC-A--A-A-G-GTTAT-AAATGACAT----T-CT-T--C-----AG----GTTC--A-----C-----G---AAA--------------G---TGG--AA-A--G----A-C-G--AAAA---CAA------ 3648
H2B.CI.x.EHO            G----AGGAGCGATCT-G-----GAC-GGGT-G-GTCTC-A--A-A-G-G-TAT-AAAT-ACAT----T-CT-T--T-----AG----GTTC--A-----C---------AAA--------G-----G---TGG--AA-G-------A-C-G--AAAA---CA---A--- 3646
H2G.CI.92.ABT96         G----AG-ACTGATTTAG-A---GATCA-GT-G-TTT-C----A-AGG---TCT-AAAT-AAACR--GT-CT-T--C--T--TG----GTTC--A-----C-------AT-AA--------A------T-RTGG--AA-A--G----A-C----AAAG--TAAC-----M 3002
H2U.FR.96.12034         G---TAG-A-TGATTTAG-----GATCAGTT-G-TTCAC----C-AGG-G-TA---AAT-A-ACA---T---G---C------G-----TT---------C--A---CA-AAA--------C--------ATGG---A-A--G----A-T-G---AAA--T-A---C--- 3157
MAC.US.x.EMBL_3         G----AGGAC-GACCT-G-A---GAC-GGGT-G-TTTAC----A-AGG---TCT-AAAT--CATA--GT-CT-T--C-----AG-G---TTC--A-----T--C-----T-AA--------G--C---T-GTGG--GACA-------AGT-G--AAAG----A-T----- 3066
SMM.US.x.H9             G---TAG-AC-GATTTAG-A---GAC-GGGT-G-TTTAC-AT-A-AGG---T----AAC--TATA---T--T----C-----AG-G--GTTC--A-----T--C-------AG--------A------T-GTGG--CACC-------A-C-G--AAAA----A-T----G 3065
SMM.US.x.PGM53          G----AG-AC-GATTTAG-----GAC-GGGT-G-TTTAC--T-A-AGG---T----AAC--CATA---T--T----C-----AG-G--GTTC--------T--C-------AG--------A------T-GTGG--AACC-------A-C-G---AAA----A-T----- 3509
SMM.US.x.SIVsmH635F_L3  G---TAG-AC-GATTTAG-----GAC-GGGT-G-TTTAC--T-A-AGG---T----AACG-CATA---T-CT----C-----AG-G--GTTC--------T--C-------AG--------A------T-GTGG--AACC-------A-C-G---AAA----A-T----- 3595
SMM.SL.92.SL92B         G-A--AG--GCC--ACAG-----GAC-A--TGG--ACAC--T-A---G-CATG--CAAC-ATCT---GT-C-GT--C-----AG-C--GTT---A--GA-T--A-----A-AG--------A---TTGT-AT-C--CA-G--G----A-T-----AAA----AA--C--T 3023
STM.US.x.STM            G----AG-AC-GATCTAG-----GAC-GGGT-G-TTTAC--T---AGG---TC--AAAT-AT-TA---T--T----C--T--AG-G---TT------G--T--------T-AA--------A-----GT-ATGG--AACA--G----A-C-G--AAAA----AAT-A--T 3239
MNE.US.x.MNE027         G----AGGAC-GACCT-G-A---GAC-GGGT-G-TTTAC--T-A-AGG---TCT-AAAT--CATA--GT--T-T--C-----AG-G--GTTC--A-----T--C-----T-AA--------G------T-ATGG--AACA-------AGT-G--AAAG----A-T----- 3060
DRL.x.x.FAO             G---TAG-CC--A-CATTTA---GAGCAG-TG--TA-A--AT-A---AC----T-AAA-TTC-----G--CGAG--C--G---------TT----G-CA-T-----T-ATGAA--------G--CC-G--G-----AA-G--G--------G--AGAA--TCGCT----T 2857
MND_2.x.x.5440          G---TC--AGTGCCTCTG-A--C--T-A--TG----TA--AT-A--GG-----T--CA-TTC-----G--CGAG--C-----------GTTC--A--G---------C--TGAA--------A----TG--G--C--AA-G------------T---AAA---CA---A--- 2851
MND_2.CM.98.CM16        G---TAG--TTGCCCCAG-A--C--C-G--TG--TC-A---T-A---G-C--C--C---TTT-----G--CGAG--C---------G--GTT---A--G-----A--C--TGAG--------A--CAT---G--C---A-G------------G---AAAG--CA---T--- 3291
MND_2.GA.x.M14          G---TC-GAGTGCGTCAG----C-AC-A--TGG--A-TAG-T----GG-----T-AAG-TTC-----G--CGAG--C------------GTTC--A------------T--TGAG--------G----TG--G--C--AC-G------------C--AAAG---CA------- 3218
OLC.CI.97.97CI12        -A-----GCC--G--CAG-----TT-GAG---G--A-AC-------GG-G-TCT--AG---T-ATAAC--AGAA--T--TCCA-----GT-----G-G--T--A--T-ATGAG--------G---AC-T-AT-C---A-G-G-----AGA--GCAAGA---AA-T----- 3164
GRV.ET.x.GRI_677        -A---CAT--TGAGACTAGA---ATCA-C-GG----CATAG-A---A-GATG---C-AGAAAAA--T--AGAA--C---------G---GTC--A-G------G---GGGAA--------G---A--T-G-----GA-----------CA-TA-CAAA----AAT-A--C 3463
VER.KE.x.TYO1_patent    --C-AGA-A-TGA-CACACC---GAC-A-TT-G----AC--T-A---AC-A-AT-ACAAGCC-----CT-AGAA--C-----A--G--GGTG--A--------A--T-ATGAG--------A--CA----TTGG---C-C------GA-C--AGCAGA---CAA---GAG 3517
VER.KE.x.9063           --C-AGA---TGA-TACACC---GAC-AGTT-G----AC----T---GTCA-G--AAAA-CA------T-AGAA--C-------------GGTG--A--GA-G--A---ATGAA--------A--CA--T-ATGG--AC-A-------CA-C--A-TAG----AACT-AGA- 3520
VER.KE.x.AGM155         -AC-AGA---GGGTCCAA-----GATCAGCT-G--C-AACA--T--GA-TAGAT--CAAGAA------T-AGAA---------G-----GGTG--A-G----------C--TGAG--------A---A--T-ATGG---C----------AGT----AAG-----AAT-AGAG 3512
VER.DE.x.AGM3           --C-GGA---TGA-TACAC----GATCGGTTGG----AC-A--A---ATGA-AT-AAGTGCC------T-AGAA-----------G---GTA------A---A--T-ATGAG--------A--CA--T-GTGG--AC-C--G---CAGA--AGCAGC----AAT-AGA- 3015
TAN.UG.x.TAN1           -A--T--GCC-GAGAAAG-A---GATGA---GG----A---G-A---A-G-T----C-ATTAA-A--CT-CGAG--G-----T--G--GGTG------AG---------ATGA---------A--CA-GT-GT-C---TC-C-C----GTC-CA-CAAA----A-A-A--- 3492
SAB.SN.x.SAB1C          G---TAGGCC--AG-CAG-A---CT-GT--TGG--C--C--T-A---G-CT----AGAA-CT-----GT---AA--C--T--A--G--GTT---A--G--T--A----A---C--------G------T-GT----AA-G------CA-C-G---GAG--CACC--A--- 3688
MND_1.GA.x.MNDGB1       -A----ATAC-GC--AAG-----GAG-A-GCT----TA--AT-A--GGCTTTAT-AA---CA---AACT-AGAA-----T--A--G---T---------------C--T-A--------G-----GT-A--C--A-----G---AGA--G-AAAGG-TCAA--A--- 2977
DEB.CM.99.CM40          -A--T---AGG-A--AT--A--CCT--AG--GG---CAC-AT-A-----GTTC--CC-AGA-A-----T-G-GG------CCTG-G--GT-----G---C--A--C--T-AC--------G-----GT-A--C---A-A------CGGT-GAT---AG-TACC--A--- 3099
DEB.CM.99.CM5           -AA-T-----T-AG-TCTCA---TT-GA-C-GG--A-AC----T--GA--TT---CC-AGAAAA--GT-A-GG-----CCCGG-G---T-C---G-G--T--C----A--A--------G-----GT-A-----AA-G------CACT-GTT---A----A---A--- 3093
TAL.CM.01.8023          G---TGA-AGCT---AA--------G-G--T-G-G-GA-----T---GCC-TC--TAAC-A-GTA--GA-CCAG-----G--GGCC---TTC---G-CA-G-----TA-TAAG---C-T--G--C--G--GT-C--A--G-------A-------AAA-----A---A--- 2757
TAL.CM.00.266           GCA-TGAGAGCT-G-A-G-A-------GG-T-G---GA--A---GCT-TCT-AAAC-A-GTA--GA-CCAG--C-----GGCT---T-C--AG-CA----G--CA-TAAA---C-A--G-----T-GT----A-----G----A-------------T-AAT----- 3259
SUN.GA.98.L14           ----T-AT--G-A-AA-A-----GA-GA--T-G-GA-AC-AT----G--G-TA--CA--GAA---AATT-AGAA--------G------T-----GG---G--A--C-ATAAG--------G--CAT-T-A--C--A--C-G---GAGA--G--AAG---AAAT-A--C 3647
WRC.CI.97.97CI14        -A-----G--C-AGACAG-A---CT--AC--GG----A--AT-A---A-G-TA-------CA---AATT-GGAG--T-----G-----G-TA---TT---------C-GTAAC--------A---AT-T-G-----A-----G---CA-A-----GAG---AA-T----- 3590
WRC.CI.98.98CI04        G------T----A-CCAAG----TT-GAG---G----A---T-A-------TG-------CT---AATT-GGAA--T-----A--G---TTG--ATT---G--G--C-G-AAG-----------ATT--A----------G---CA-A-----GAA---AAA--A--- 3605
WRC.GM.05.Pbt_05GM_X02  -A--T-AT----AGAAAG-A---TTGCA--TG--TC------T-----G-T-T---T--A-CT---AAC--GGAG--T-----A--G--G-TA--ATT-----------GTAAA--------G---ATCT-G-----A---------GA-A-----GAG---AAAT-A--- 2519
LST.CD.88.524           -A---AGG--G-A--AAGGA-----TCAGGC-G--A-A--AT-A---G---T-T-AA-A-CT---AACT-AGAA--T-----G--G---TT---AG-CA---A--C-A--A--------G-----TC--G--C--A----G---GA-A--G--AGAG--AAA------ 2555
LST.CD.88.447           -A--TAGG--C-AGAAAGG----GATCAGGC-G--A-A--AT-----G-T-TCT-AG---CA---AATT-A-AA--------G--G---TT---AGC---------C-AT-A--------G-----T---T-----A---GG---GA-A--G--AAA-CTAAA------ 2558
LST.KE.x.lho7           -A--TAG---G-G-AAAGGA---GATCAGGC-G--A-A--A--C---G---T---TA---CA---AACT-AGAG--------A--G--GTTC--AGC---G--A--C-AT-A------------------T---G----------GG---GA-A--G-AAAG--TAAAT-A--- 3642
SYK.KE.x.SYK173         ----TAG-TC-GAGAA-GCA---CAGGAG-T-G--C-AC-AA-AGTGAC-GCA--A--A-AAGT----T---AGGTC---A-AG-----TGG---G-CC--TA----A-G-AG---T-A-----CACT--T-----A--C--G---CAGT-G---AAA----AA--C--T 3484
SYK.KE.x.KE51           --A--AG-TC-GA-AA-G-A---G-TCA--TTG--AGT--AA-TGT-A--AG---A--AGCAGTA---T-CT-T-TT---CCAG-----TGG---G-CA--TTC--C--G-AA---T-A--G--CACG--A--C--------G----GTC----AAAGG-TCAA--A--- 3142
COL.CM.x.CGU1           -A--TC-GCCTT---A-G---------GGGC-G--C-------C---G-TGCA--C---TAC-AT--GT-AGAG--------A--G---T-C---GC---T-------A-AAG--------C------T-A------A-GC------A-T----AAAGG-CAAAA-T--- 3062
```

```
H1B.FR.83.HXB2           GAAAAAGACAGC......TGGACTGTCAATGACATACAGAAGTTAGTGGGGAAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAAACTCCTTAGAGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAGCTAGA 3454
Pol                       E  K  D  S  .  .  W  T  V  N  D  I  Q  K  L  V  G  K  L  N  W  A  S  Q  I  Y  P  G  I  K  V  R  Q  L  C  K  L  L  R  G  T  K  A  L  T  E  V  I  P  L  T  E  E  A  E  L  E
H1A1.UG.85.U455          ------------......--------------T--------A-----A--A---C-A-------------------TG---------------A----C-G--------T--C--G---G---------------TA--G--A-C--G--T-----------AT---- 2900
H1B.US.90.WEAU160        --------A--T......--------------------------A------------------------------G---------------A----C-----------------G--G-----------------A--------A---------------------- 3453
H1C.ET.86.ETH2220        -----G--T---......--------------T-----A-----------A----A--C------------------------------------C-G-----------------G---G---------------CA--G--A-------T-----------AT---- 2846
H1D.CD.84.84ZR085        --G-----A---......--------------T------------------A-----A-----------C-------T----A----------A------------------G-------G----------G----------------------------AT---- 2976
H1F1.BE.93.VI850         A-C--G------......-------------------------A--A--C-A--------------------T-------------------C----------------G---G----G----------CA--G-G----G--T-C---G-----T---- 2791
H1G.SE.93.SE6165         --C--G--A---......--------------T-----A-----------A--C-A--------------------T---------------C---C-----------------G--GG---------------CA--G--T----G-----C---------AA-G-- 2851
H1H.CF.90.056            ------------......--------------T------------------A--C-A--------------------T---AAT---------A----C-----------------G--GG-------T------CA-----------G---A----G-----AT-G-- 2801
H1J.SE.93.SE7887         --------AGA-......--------------T-----A-----------A--C-A--------------------T-----A----------A----C---------------A---GG-T---------------CA--G-----T-G--TAG---------AT-G-- 2768
H1K.CM.96.MP535          --C--G--T---......--------------T------------------A-----A-------------A--A--T------------------AA-----G---------------G---GT---------------CA--G-----T----T-C-----------T---- 2650
H101_AE.TH.90.CM240      ------------......--------------T--------A--------A--C-A--------------A-----TG---------G---A---C-G-----------C--G---G-T--------------CA--G-------G--T-----------AT---- 3028
H102_AG.NG.x.IBNG        ------------......--------------T--------A--------A--C-A--------------------TG---A------A--A---C-G---G------C--G---G----------------TA--G--G----G--T--G-------AT---- 2979
H1N.CM.95.YBF30          -----G--TGTA......--------------T--------A-----A--A--G--A-------------------C--T-----A--C-G-----AA--GC-C------T-AA-C------------------TT-G----------G-CAACT-T---------------AT---- 3048
H1O.BE.87.ANT70          A-C--G--TGTG......-----A--A-----T-----A--AC--A-A--A--GC-A---------------A--C--T-A---A----G---G--AG---G-----GT-AA-------C------GT--T-----------G----TT---GTAG---G--------G-- 3509
H1O.CM.91.MVP5180        --C-----AGTG......-----A--A-----T-----A--A-----A--A-----A---------------A--C--T-A---A----G-----AAG---G--C--GT-AA-C-----------------T--T-G-----G---G----TT---GTA----G-----A----- 3484
CPZ.CD.90.ANT            --GCC---TGAT......-----A--T---------C-----AC----A--A-----A--------------------C---------A--C---ACT-A---G--G----G----A-C------GT---GAGT--------TAG-G---A-A-G--TAGG---------AT---- 2893
CPZ.CM.01.SIVcpzCAM13    -----G--GGTA......-----A---------TC-T-----AC-CA-A--A-----A--C-------------------T-----A--A---A---AA--G--G--C---T-AA-A---------A---AGGT-------T---G-T--T--C---CC---------AT---- 3034
CPZ.CM.05.SIVcpzEK505    --G-------CA......---------------T-----AC-------A--A-----A--C-----T--C-----C---------------------AA---C-T--C---T-AA---------G-A-----T------------G-CA-----T---AG---------AT---- 3026
CPZ.CM.05.SIVcpzMB66     --------A---......---------------T--C-----A---A-A--A-----A---------T-------A--T-----A--C-G-----AA---C-G--C--G---A-----------A-----TT----------TG--A-CT-T---AC----------AT---- 3000
CPZ.CM.98.CAM3           C-G-----G-C-......---------------T--T-----AC-------A--C-A---------------A--A--T-----A--A-------AA--G-----C--G--GA-A-A---G--A-----T---------CA--G-GA-CA-G---C-G---------AA-G-- 2865
CPZ.GA.88.GAB2           C-G-----G-AT......----------A--------T-----AC-G--A--T-----A------------------C--T-----A-----GA---AA-----G------T-AA-C------G-T--G--TT---------A--G-CGAGA-G--TAG---------AT---- 2847
CPZ.CM.98.CAM5           C-G-----G-T-......---------------T--T-----A-----A--A-----A---------------A-----TT--------A-------AA--G-----C-------A-A-A---G--A-----CT-G--T------G--G-CA-G---C-G---------AA-G-- 3156
CPZ.CM.05.SIVcpzMT145    ---C-G--AGTG......---------T-----T--T------------A--C-----A---------------A--A--T-----A------A-T-A---GC-T------T-AA----G---G-G--G-GT--G----------G--G-T--C--CCC------T--A----- 2996
CPZ.TZ.00.TAN1           ---CC---GCAG......-----A--G---C-T--C---------G--A--C---C-T---------C-----A-----T-AT--A--A---ACT-AAG--C----C---T-GA--------GTA----G-T----T--GCC-G--GA-A-G--CAGG---------AT-G-- 3098
CPZ.CM.05.SIVcpzLB7      -----G--A--T......-----------------------A--G-----A-----A---------------A--A---------C--C--GA---AA--GA----C-GG--T------G---GTA--GAATTC-----------G-CA-CT-T--TAG---------AT---- 3009
CPZ.US.85.CPZUS          C-------A-TT......-----A---------T--T-----A-----A-----C----C--------------A--T-----A--A---A---A---G-----C---T-AA-A-A----G-T-----T-----T-----TG--AATT-T---C-T--G-----AA-G-- 3514
CPZ.GA.88.GAB1           --------GGTA......---------------T-----A--AC-GA-A--A--G--A-------------------T-----A--A---A---A-----------------G--GA-A---------A--GAA---G-----T---G-T--T--C---CC---------AT---- 3515
MON.CM.99.L1             --CC-G--ACAA......---------------G--C--A-----GA-A--A--GC-T---------GCA--A--C--TT----A--A---ACC-AAA-CC-G--C---AGTA-C------GG-ACGGAGC--G-----TAC-G-GGTC--CT-----TTT---TC-AGCT-- 3160
MON.NG.x.NG1             ---C-----CAG......-----A-------GG--C--A--A-----A---------A--C-----GCG-----C--------A--C--GAC--A---TC----C-R-TGTA-CC-R--GRTA-----C-----G--CA--G-GGAG---T-----CTT---CC-AGC--- 1753
GSN.CM.99.CN166          --CC-G--T-C-......-----A--G---------C-----A--G--A---------A-------T-TCA--A--AC---T--T--A-GGAC--AA--G--G--C---TGTA----G---G-T----GG--------CCAGG--GAG--C--T--G--G---C--A---- 3182
GSN.CM.99.CN71           --TC------T-......-----A--A---------C--A--A--G--A-----GC-C-------T-TCG--A--AC-T--T--C--A-GGAC--AA--G-----C-G-TGTA---------G------GG---------CAGG-GGAG--C--T--G------C-AA---- 3164
MUS_1.CM.01.1085         C--C----AGTT......-----A--T-----TG-G--A---C-G--A---------A---------GCA--AGCA--------T--C--GAC--A---------C---TG-A-C--G---GTA---T-T--G--T----AGG-C-AGT-G-----G------C--GC--- 3144
MUS_1.CM.01.CM1239       C-GC----A-A-......-----A--A-----T------------G--A--C-------------GCG--AGCC--------AG-----ACC-AA--GC-----C---TGTA-A--G---GTT---T-C--G-------A-G-T-A----------G-----TC-AT---- 3134
MUS_2.CM.01.CM1246       --CC-G----C-......-----A--G-----A--T------C-G--A-----G--A--C------GCC--AG-C--TT--------C---ACT-AA--T-----C-G-TGTA-C--G----TA--G-GG---------GC--G-GGAG--C-GT--GC----CC-AT---- 3242
MUS_2.CM.01.CM2500       --CC----A-C-......-----A--A-----G--T-----AC----A--A------------------GCA-----A---T----A--C--GACC-AA--TC-G--C---TGTA-C------GTA------C--T------AC-G--GA---T--C--GC-G--TC--T---- 3201
RCM.NG.x.NG411           --------AGAA......-----A--A-----T--------A--G--A---------A-------------------A--TT----A--A--GAC--A-G-G-----C-----AA-A--G---GTT--G---T-GGAT--CAAGG--GA-T-T---AG---------AT---- 2948
RCM.GA.x.GAB1            --G-----AGAA......-----A--T--C--T--T-----A-----A--A--G--A--C------------A------AGT---C-A-GGACT-A-G-----G--C--GT-GA-A--G---G-A------T--GAT---AA-G--GA-A-G---A----------AA---- 2933
DEN.CD.x.CD1             TTGCC---GGATGGGCAG-----G--G----CA--T------C----A--AGT-------------------T-G---T----A--A--GAC--AAG-----------GAG-A---------GTA----ACT--GG----ACC--TT-----C--G--G-G--------GC--- 3347

MAC.US.x.239             C---G---G-C-......-----A--G-----T------------------A--AGT---A----------GC---A-----T-----T--A---ACC-AA--TC-C----GGT-AA---------AA-TGA-T--------G-A-G-T-AGTGG--T--GATG-----AGC--- 3756
Pol                       P  Q  R  E  .  .  T  W  T  V  N  D  I  Q  K  L  V  G  V  L  N  W  A  A  Q  I  Y  P  G  I  K  T  K  H  L  C  R  L  I  R  G  K  M  T  L  T  E  E  V  Q  W  T  E  M  A  E  A
H2A.GW.x.ALI             C-G-----AGTA......-----A-----------C--A--ACAT-----TGTCC-A---------GCA--A--C--------A--A--GACC-AA--C-------G---AA---------AA-TGA---C-------G-G-G-AGTGG------CT------AGC--- 3832
H2A.DE.x.BEN             C-G-------TA......-----A-----------C--A---C-------AGTC--A--------GGCA--A--C--TT----A--A---ACC-AA--C-------G---AA---------AA-TGA---C-------A-G-G-AGTGG------CT------AGC--- 3839
H2A.SN.x.ST              C----G--AGTA......-----A-----------C--A--AC-G-----TGTCC-A---------GCA--A--C-----------A--GACC--AA-C-------GGT-AA-C-------AA-TGA---C-------AGG---AGTGG------TT------AGCG-- 3284
H2B.GH.86.D205           --------AGTT......-----A--G----CA--T--A--AC-G--A--AGT---A--C------GC---AC-C-TT--T--A-----GAC------CA----C-----AA----G----AG-TGA-C----------A-G---AGTGG------CT------AGC--- 3812
H2B.CI.x.EHO             --------GGTT......-----A--A---------T---------------AGT---A---------GC---AC---T---G---------GACC-----TA----------AA-A--G----AG-TGA-C----------AGG---A-TGG--T---TTG------GC--- 3810
H2G.CI.92.ABT96          --------G-CT......-----A--A-----T--T--A--AC----A---GT-C----Y------GC----C-A--T-----A--C---AC--AAA-CC-G--Y---T-AA-------T-AR-TGA-CT-G--G----A-G---AGTGG------TTG-----AGC--- 3166
H2U.FR.96.12034          ----G---TGTT......-----A--A---------T-------------A--AGT-C-A---------GCA--AG----T-----A--A--GAC--AA--TC-T-------T-AA---AG--G-AA-TGA----C-------AGG---AGTGG------ATG-----AGC--- 3321
MAC.US.x.EMBL_3          C---G---G-C-......-----A--G-----T-----------------A--AGT---A---------GC---A-----T-----T--A---ACC-AA--TC-C----GGT-AA---------AA-TGA-T--------G-C-G-T-AGTGG--T--GATG-----AGC--- 3230
SMM.US.x.H9              C---G---A-C-......-----A--A---A-T-----A--A-----A--AGTGC-A---------GCA--A-----T-----A--A--GACT-AA--TC-T--C-G-T-AA-C--------AA-TGA-TT---------AGG-T-AGTGG--T--GATG------GC--- 3229
SMM.US.x.PGM53           C---G---A-C-......-----A--A-----------A--A-----A--AGT-C-A---------GCA--A-----T-----A--A--GACT-AA--TC-T--C---T-AA-C--------AA-TGA-TT---------AGG-T-AGTGG--T--GATG-----AGC--- 3673
SMM.US.x.SIVsmH635F_L3   C---G---G-C-......-----A--A-----------A--A-----A--AGTGC-A---------GCA--A-----T-----A--A--GACT-AA--TC-T--C-G---AA-C--------AA-TGA-TT---------AGG-T-AGTGG--T--GATG------GC--- 3759
SMM.SL.92.SL92B          -----G--GC-A......-----A--A-----------A--AC-------AGTTC-A---------GCA----A--T-------------AC--AAA-CC-G--C---A-GA-C-----G-AA-TGA-T--G-------A-G-T-AGTGG------TT------AGC--- 3187
STM.US.x.STM             C-G-G---GTT-......-----A--G-----T--C--A--AC-G--A---GTC--A--C------GCA--A-----T-----A------ACT-AA--TC-G--C---T-AA-C-----T-AG-TG-----G--T--G-AGG---AGTGG------ATG-------GC--- 3403
MNE.US.x.MNE027          C-------G-C-......-----A--G-----T-----------------A--AGT---A---------GC---A-----T-----T--A---ACC-AA--TC-C----GGT-AA---------AA-TGA-T--------G-A-G-T-AGTGG--T--GATG------GC--- 3224
DRL.x.x.FAO              --T-----A-TT......---------T-----T--T-----A-----A--AGTGC-T--------T-----A-----TT-C--A------AC--A-G-GC-C------T-AA----G---G-----C-C---GAT----A-G--GA-TGG--TAGG---------A----- 3021
MND_2.x.x.5440           --------A---......-----A--A-----T--T-----A-----A------C-T-----------C--A--C--TT-C--A------AC--AAG-GC-C------T-GA-A--G--GG-A-G-C-C---GATA--A--G-GGA-TGG--TAG----------A----- 3015
MND_2.CM.98.CM16         --G-----AGAA......-----A--A---------C--A-----G--A------C-T-----------------TT-C--A------AC--AAG-GC-C----G---AA-------GG-A---C-C---GAT---AA-G--GA-TGG--TAG----------AT---- 3455
MND_2.GA.x.M14           --------G-CA......-----A--A-----T--T-----AC-G--A--A---C-C--------------A--C--TT-C--A------AC--AAG-GC-C--C---T-GA-C-----GG-A---C-T---GAT------G--GA-TGG--TAGG---------AT---- 3382
OLC.CI.97.97CI12         --------AGAA......-----A--------T--------A-----A--C----A--------TTCA--A--C--TT----A--A---ACG-AAGCT-----------GA---AG---C-T----GGT-GCT---TAA-G-GGAGTGG---------G---A--GA--- 3328
GRV.ET.x.GRI_677         CCCTT---AG-A...GAA-----A--A--CA-A---------G---A--AGTTC-A---------------A----T-----A------ACC-AA--TACC--GCCA-GT-G-----G-AA--GAAC--CCT-----A----GT-TGG---------G------GC--- 3630
VER.KE.x.TYO1_patent     --------TGAA......------------------C----------T------C-A---------GCA--AT-G--T-----TC---GGACC-A-A-TA----C--GT-AA---------AG---AAT--GTT---GC--G-GA-TTGG---CCT--G-----AGCT-- 3681
VER.KE.x.9063            -----G--ACAA......---------T-----T--C--A--AC---A------------C------GC----C-------C---T-A-G-AC--AA--GA----C-GGT-AA-C--G----AG---AAT---TT---G-A-G-G-A-TGG--T------G--G--AGC--- 3684
VER.KE.x.AGM155          A-G-----ACAA......-----A--G-----TC-T-----A--G--A---------A---------GCA--AT-A--T-----AT-G-G-AC--AAA-TA-C------G--A-----------AG---AATT--TT---C--GG--GA-TGG--CCC----G-----AGC--- 3676
VER.DE.x.AGM3            --C-----AGAA......---------A-----T-----A-GAC------------C-A---------GCA---C-----------AC-C-G-ACT-AAA-C--G-------T-AA-C--------AA--GAACT--CT----AC-G--A-CTGG-----G---------AGC--- 3179
TAN.UG.x.TAN1            --GC----ACA-......-----A--G-----T--T--A--A-----A--AGTG--A--------T-----AT-A--T-----A-----GAC--AAGCCC----C---T-A--AC-----G-TC---ATT------CT---G-G-A-TGG--G--GC-G-------G---- 3656
SAB.SN.x.SAB1C           ----GG--AGAA......-----A--G-----T--T--A--A-----A--A--G--A---------------A-------A--------A--GACT-AA--T--G----G---AA--------G-T-G-C-TT-G-----GA--G-C-A-TGG---------G-----A----- 3852
MND_1.GA.x.MNDGB1        ---TT--CAGAACAG...CCA--A--A-----A--------A--G--A--T-----A--------TGCA---T-A--T--T-----C---AC--AA---C-G--C--G--AA-A------GGACT-AACA-------GAA-G-CA--A-G-------------AGA--G-- 3144
DEB.CM.99.CM40           --TG----AGAA......-----A--GC--A-A--T-------------A--AC-G--A-------------C--A-----TT-------A---AC--AA--C--------GGGCAA-C-----GGT-CCG-G-T--------CCCGG-TA-C--CT-T---------TC-AGC--- 3263
DEB.CM.99.CM5            ---G-G--GGAG......-----A---C-CA-A--C-----AC-G--A--AC-G--A-------------------A---T----A--A--GACC-AA--C--G--C--GGCAA-C-----GGTTCC--GG--------TAA-G--AGC--TT----G------C-AGC--- 3257
TAL.CM.01.8023           ------C-G-CA......-----A--A-----TT---------C-G--A--A---C-A-----------------A--A--TT-------C---AC--AAGCCC----C---TG------G--TGTA----GCT-GTT-----AGG-TA--T-G----------------AGC--- 2921
TAL.CM.00.266            --T---CT--CA......-----A--------GT-----------A-A--A---C-T----------T-C--A--A--TT-T--T------AC--A-GCCC-C--C-GGTA---C--G---GTA----G---TCT-----A-G---AGT--T----G---------GC--- 3423
SUN.GA.98.L14            CCTCT---AGAAGAG...CCA--A--G-------------AA-T--A--AGTTA-C--------TTCA--AT-A--TGA----C-C-G-AC--AAG-------------AA-------G-AT---C--T--G----GAA-G--AAGA-G-----------TAG-GA--- 3814
WRC.CI.97.97CI14         ---G---TA-AG.........--A--A-----TG-------A--G--A--A-----A--C-----TGC----C-A--TGA---A--AC-GAC--AA--------C-G-T-GA-------G-AT---C--T-G--------A-G--GA-TGG-----------TTTAGA--- 3751
WRC.CI.98.98CI04         --GG---TA--A.........--A--A-----TG----A--A--G----A-----------------GCA--------TGAG--A---C--AC---A--------C---T-AA---------A----C-C--C--------A-G--A--TGG-----G-------TTAGA--- 3766
WRC.GM.05.Pbt_05GM_X02   ---G--ATA--G.........--A--A-----TG-------AC-GA-A--A-----A---------GCA--A--A--TGAG--T--A---AC--AAG----G------T-AA---G----AT---C-T-------C--A-G--GA-TGG--CATG---------AGC--- 2680
LST.CD.88.524            --G-T-----CAAGGAAAACT--A--A-----A--T-----AC-G--A--AGTT--A---------GCA--AC-G--TGAT--A--C-G-AC--AAG-GC-C-----GT-GA--------GTA--GC-CT-GGA-C--C--G-GA-TTGG-----G------CTAGA--- 2725
LST.CD.88.447            --C-T----CTTCAGAAGACA--A--T--CA-------------C----A--AGTC-------------GCA--AC----TGAG--A----G-ACT-AAG--C-T-------T-AA---G---GTA-----T---GAT-----TG--A-CTGG-----T-C-----TTAGA--- 2728
LST.KE.x.lho7            --G-TG--TCTAACAAAAACT--A--A---C-A---------AC-G-----AGT-C-T---------GC---AT-G--TGAT--T----GGAC--AAG--C-C--C---T-AA-A--G---GTA---C-CT-GGA----A-C---AACTGG-----G-----CTT-GA--- 3812
SYK.KE.x.SYK173          A-T-T---TGA-GAA...ATC--A--G---C-GT-------A--GA-A--AGT-C-T-----------C---------TT----A-----GAC--AAG-GC-C-----GTGTA-C----------T---CA-T-----------C-CA----T--------C---T---GC--- 3651
SYK.KE.x.KE51            ----TTACAGAATCA...CCA--A--T-----GT-------AA--A-A--AGTCC-A---------------A--A--T--T--T--C---TGT-AAGC---GA----GTG-A-A-----GG-A---------T--C--CA-TG--GA---G--------G-----AGC--- 3309
COL.CM.x.CGU1            ---C-G--T-AG......-----A--A-----T-----A--A--G--A--AGCC--A------TT---C--AT-A---------A--C-GGAC--A-G-GA--A-----TG-AC--A---T-AG--GCATT-GTTG--G-A-G-G-AG--T-GT-C---------AGAG-- 3226
```

```
H1B.FR.83.HXB2             ACTGGCAGAAAACAGAGAGATTCTAAAAGAACCAGTACATGGAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAGAGCCA...TTTAAAAATCTGAAAACAGGAAAATATGCAAGAATGA 3621
Pol                        _L__A__E__N__R__E__I__L__K__E__P__V__H__G__V__Y__Y__D__P__S__K__D__L__I__A__E__I__Q__K__Q__G__Q__G__Q__W__T__Y__Q__I__Y__Q__E__P__.__F__K__N__L__K__T__G__K__Y__A__R__M__
H1A1.UG.85.U455            -T-------G----G-------------C--T--G--------A-------------------C--G-----------------A--A------A-----------------------...-------------------------------------AA- 3067
H1B.US.90.WEAU160          G-------------G--A--------------G-------------------------------------GC-------------------------------G-----------...---------------------G---------G--- 3620
H1C.ET.86.ETH2220          -T-------G----G--A-------G--------------A-T---------------------T----------A-----A-T-A----------T----T----C-----A---...--C---------------G--G-T-----A--GAG 3013
H1D.CD.84.84ZR085          --------------G---------G-----A-G-------------------------------T------A--A-------T-----------------------A---...---------------------G----------- 3143
H1F1.BE.93.VI850           -T-----A---T--G---------G-------------G--A-------------------------------A--A--AG-C--G--------------------GA-C---...------------A----------G-------A-G--- 2958
H1G.SE.93.SE6165           GT-------G----G--A-------G------T----------C-----------------A----------G---A--A--A---T--A----------------------...-AC--------------------G-------A--G-G 3018
H1H.CF.90.056              -T-------------G-------G-G-------A----------A--------T------------------------G-----A----------------------G-------...-------------G----------------A----- 2968
H1J.SE.93.SE7887           -------------A-----------------------G--A---------T--G-------A----------G-G----A--A----TG-A---------------------G-------...--------C----------G---------A--G-- 2935
H1K.CM.96.MP535            -T-A----G------------------G----G--A--------------------------------A--A---A-T-A------------------A------------...CA-----------------G--G-------------- 2817
H101_AE.TH.90.CM240        GT-------G----G-----------ACC--T--G--------A--------------------G-------G------A--A------G-A-----------------------...-----------A------------------C----GAG 3195
H102_AG.NG.x.IBNG          -T-------G----G--A-------------T-------G--A----------A----------G-------T------A--A------A-----------------------...----------T-A-------------------A--A-- 3146
H1N.CM.95.YBF30            ---A-----------G-----AT----------CC-G--------C----------GG-----A---G-----------T--A-----A--A-----T--G---------G---------G---TT-...CA-------T-A------------G--------A----- 3215
H1O.BE.87.ANT70            -T-A-AG----------A-GGT-----C-------G-----G--A--C---C-A--TGAT--G--TC--TGG-TTA-T--T--------A--AGGG-AG-------T--C--G--A-----G--AGA-...CA---G--C--C--------G------A-T--GCAA- 3676
H1O.CM.91.MVP5180          -T-A-A-----------A-AG-------G--------------A-----CC-G--TGAC-------GTGG-TTAGT--T--------T--AG----G-------T--C--GG-A-----G--TGA-...CA---G--C--T----------------T--GCAA- 3651
CPZ.CD.90.ANT              -T-A-A-----T-A-C-A------C-GC-GAA-A--G-G---TAC--C---C-G--TGGTCT-CCAC---A----ACC---------A--A---TC---A--------C---------A------A-TGA-...GGA---CTGT-A---G------------T--GCCT- 3060
CPZ.CM.01.SIVcpzCAM13      -T-A--------------A--AG-T-GTAC-----------G--T-----A-T---GAT--G--GC-C---------G----A--A-----AG---G--G-----------G--C-T---G--A-A-...CA--G----T-A--G-------------C---CA-- 3201
CPZ.CM.05.SIVcpzEK505      GT-A--T--G--T--G-----AT----------T-------G-CC------A----GAC-----A-----------G--C-----------A-----T-----------G-------------CTT-...CA-------T-A--G-----------------A----- 3193
CPZ.CM.05.SIVcpzMB66       -T----------T-----A--AT--------------------G-CC----------C--------A-----------------------A--T-----A--G-----T----------TC--G--A-AG...-A---G--CT------------G--------T--G---- 3167
CPZ.CM.98.CAM3             -T-A-A---G-----R------T---RG--T-----G-----G-----C---------GA---G--A------------N-------A-----AA-TA-T--G-----T----------TC--G--A-AG...CA---G--C--A---------C--------T--GCAA- 3032
CPZ.GA.88.GAB2             GT-A-A------T--G--A---T--------------G-----CACC--C----------GA-------AC-----------G-T--A------AATA-T--G--------T-----G----T--G--AA--...GGA-GG-----A-------------------C---CAA- 3014
CPZ.CM.98.CAM5             -T-A-A------T-----A---T----G--T--T--G---------T--C-----T---GA---G--G------------GG----A--A-----AA-TA-T--G-----T--C--G----TC--G--A-AG...CA---G--CT-A---------G--------T--GCA-- 3323
CPZ.CM.05.SIVcpzMT145      ------------TCA------A-----------------------CC--C---------TGA---G--A-----------------------A--A--G--T--G-------------A-T---G-------...CA------C--A--G------------GC-A-CAA- 3163
CPZ.TZ.00.TAN1             GT-A-A------T-AGC---------------AAG--T--A----CA--C-----T--TAA-TT-CCTC-GCA-----C--------------------A--G-------------G--A-----G--AGA-...GGG------T-A--------------------A-TCAC 3265
CPZ.CM.05.SIVcpzLB7        -T-A--------T--G--A--AT-----A-C------------G--C-----------GG-T------C-G---------G----A--A--A--A--GA-T--------C--------C-----G--A-A-...---------------------------------C---CAAG 3176
CPZ.US.85.CPZUS            GT-A-A--------------A-------G-------------G--C-----------GA------A---G---------G----A--A--A--AAGGA-T--G------C------T-----A-GG...CA---G-----A--G--------------C---CAA- 3681
CPZ.GA.88.GAB1             -T-A--------T--G-----AG---GCAC------------G--A--C-----T---GAC-----GC-T--------------------A-----CA-CT----G-----T-----G--A-T--G--A----...CA---G---T----G-----G--------C---CAA- 3682
MON.CM.99.L1               G--A--G-----T--G--------C--G--GGA-TCCACC--G-GA-----C-----CAACG-GCCA----AG-TG--GT--ACCTCA-TA-CAG--AA-------GGC--CAG-T-C-TC-----AAA-.....TC-GTAT----G-----------T---C-A---T- 3324
MON.NG.x.NG1               -T-A--------T--G--A-------GTC---AGTCGAC---G-GA-----C-----CCAGCTC------G-AG-T---G--TACCTCC-T--YAG----A--G---GGG--CRGGTA--TC--G--ARAT......-GGGT---A--G-----C----TC----AG--T- 1917
GSN.CM.99.CN166            -T-A--------T--G-----CT------C--T-TTC-G-G--T-GC----------G--GA---GCCGC-CG-G-TG-----TACCTCC-TA--AG--CAG------GGG--CATGT-CAGC-----AGGC......--G-TGT-A---T---CC----TC----AG-CC- 3346
GSN.CM.99.CN71             GT-A-------------------------C--T-CTC-G-G--CAGC--C--C---G--GA---GCCAC-TG-G-T------TACCTCC-TA--AG--CAG------GG---CATGT-CAG---G--CAGC......--G-TG--A-G-T----C----T-----A--CC- 3328
MUS_1.CM.01.1085           -T----------TCAG-----C--TC--C-GA----C--G--G-CA--C----------GA-C-GCCA---G-G--------TGTCTCA-TA--AG-CTCA--G---GG----A-CT-CACC--GAGCAAG......GGGGTAT-A-GG-----------T---C-A-G-A- 3308
MUS_1.CM.01.CM1239         -T-----------CAG--A-----T--GC-GG----C--G--G-CA--C-T---T---GA-C--CC-C-GG-G--T-----TGTCTC-TTA--AG-CTCA------GGG---A-CT-CCTC---A-TAAG......GG--TAT-A---T-T--------T---C-A-G-C- 3298
MUS_2.CM.01.CM1246         ---A--------T--G--A--C--G--GC--GAG-CCGG---G-CC--------------GA----CC-C--G-CTT---G---GT-TCCTTA---G--CAA------GGG--CACCT--ACC-----CAG-......--C-TG--A-G----------G-T---C-AG--C- 3406
MUS_2.CM.01.CM2500         GT--------------G-----CA-G---C--GAG--CA-A----CA--C-----T---GA----CCA------TCTG-------GT-TCC-TC---G--CAA------GGG--CAC-T-CACC------AAG......--C-TGT-A--G-----------T---C-A-G-T- 3365
RCM.NG.x.NG411             -TAT-A--------A--CTG-C--G--G--GAA-A-------G--T------A-----GA----CCA----------TA--G-T-----ATTA-AAGG---G--------GC--C--G--AG-G--G-GAGAC...-CA---CC---A--G-----G--G-------A-CAA- 3115
RCM.GA.x.GAB1              -TAT-A--------AGAT-------------AAGT-G-----G------------TGA-AAG---CC----G-G---A-C--T------TTA-AAGGT--A--G---T-C---------AG-A--G---T--...GGA---CCA-----------C--------T-AGCA-- 3100
DEN.CD.x.CD1               G--TAA----GCACA---A--AT---------AGTA-TTCA---AGT------AGG---GAG---CCT--G---------G--CTCT--ATTA--AG--AAT--G---GG----AC---CA-A--G--AAA-......C-TCTA--C---------C-----CCAT-AGGAC- 3511

MAC.US.x.239               -TAT-AG-----T-A-ATA-----C-GTC-GGA-CA-G-A---TGT-----CC-AGA-GGC--GCCA---GA---CACGG--ATA---AGTCA-G-CAAT--G---T-T---A-----C-C-----AGAC..........---TA------GT--------T-----AG--A- 3920
Pol                        _E__Y__E__E__N__K__I__I__L__S__Q__E__Q__E__G__C__Y__Y__Q__E__G__K__P__L__E__A__T__V__I__K__S__Q__D__N__Q__W__S__Y__K__I__H__Q__E__.__.__D__K__I__L__K__V__G__K__F__A__K__
H2A.GW.x.ALI               ---A-AG--G------ATT--CT---GTC-GGA-CA-G-G--GCAC-----CC-AGA-GA---G--G---GA----AC-G-C--A--AG-TCAAG-CAAT-----------A----AC-C--G-GAGA-.........---T---A---GTG-A---G-------A----- 3996
H2A.DE.x.BEN               G--A-A---------A-ATT--CT-G-GCC-GGA-CA-G-A---TAT-----CC-AGA-GA------A---GAG---AC---C--A--AAGCCAAGG-CAT-----------CA----AC-C--G--AGAG.........---TC--A---GT------G-------AG--A- 4003
H2A.SN.x.ST                ---A-A---------A-ATC--CT---GCC-GGA-CA-G-A---TGC-----CC-AGAGGA---G--GC--GA----AC-G-C--A--AG-TCAAG-CAAT--G---------A-G--AC-C--G-GAGG-.........---T---A---GT---------------AGG-A- 3448
H2B.GH.86.D205             G--ACAG--G--T-A-ATC--CT--G--C-GGA-CA-G-A---TCC--C--CA-GGA-AGGGT-CCGC--GA----AC-G--------AA-CCTAGC-AAT--G---------CA-----C----G-GAAAT.........---GTC--A---GT---------------AGG-T- 3976
H2B.CI.x.EHO               -T-CCAG--------A-ATC--C---G--C--GAGCAGG-A---TCC-----CA-AGA-GGGGT-CCT---GA----AC-G-G-----AA-TCTAGC-AAT--G---------CA-G---C----G-GAGAT.........----TC--A---GT---------------AGG-T- 3974
H2G.CI.92.ABT96            -T--CA---------AGATT--C---R--C-GGARCAGG-A--C-CC-----CA-AGA-GR-G-GCCT--RGA---CACTG-------AA-TTTAG-CAAT--G---------CA--G--C------GAAAC......-G--TC--A---GT---T-----------AG--A- 3330
H2U.FR.96.12034            -T-T-----G--T-AGATT---T-G-GTC-GGA-CA-G-G--G-CT-----CAGGGA-GA-G--A-TC--GA----AC-G-G-TT--AA-TCAAG-TAAT---------C--CA-G--CC----G-GTGAC......-GG-CAT-A---GT---------T---C-AGG-A- 3485
MAC.US.x.EMBL_3            -TAT-AG-----T-ATATA-----C-GTC-GGA-CA-G-A---TGT-----CC-AGA-GGC--GCCA---GA---CACGG--ATA---AGTCA-G-CAAT--G-----T---A-----C-C-----AGAC..........---TA------GT-A------T-----AG--A- 3394
SMM.US.x.H9                -TAT-A---------AGATA----CC-GTC--GA-CC-G-G---TGT--C---C-GGAGGG---GCCAC--GAG---CC-G--ATA---AGTCA-G-TAAT------T-----A-----C-C------GAC..........---TA--A---GT---C----T-----AG--T- 3393
SMM.US.x.PGM53             -TAT-A---------AGATA-----C-GTC--GA-CAGG-G---TGT--C--CC-AGAGGG---GCCA---GAG---AC-G--ATA---AGTCA-G-TAAT------T-----A----CC-C-----AGAC........--G-TA------GT---C----T-----AG--T- 3837
SMM.US.x.SIVsmH635F_L3     -TAT-A---------AGATA-----C-GTC--GA-CA-G-A---TGT--C--CC-AGAGGG----CCAA--GAG---AC-G--ATA---AGTCA-G-TAAT--G---T-----A-----C-C-----AGAC..........--GTA--A---GTG--C----T-----AGG-T- 3923
SMM.SL.92.SL92B            -T-------G--T---ATT---------TC--GAGCAGG-A---AGA-----CAGAGA-GATG--CCGC--GA----ACGG-T-TA---A-TCAAG-CAAT--G----GT---A-G---C------GTGAC......-GG-T---C--GGT---G--G-T-----A---A- 3351
STM.US.x.STM               -TAT---------T-A-ATT-----C-GTC-GGAGCAGG-A--GCAA--C--CCGAGA-GAC--GCCTC--GAG---AC-G--GT---AG-TCA-G-CAAT------T-C---A-G--AC-C-----AGAC..........---TAT-----GT---C--G-T-----A---A- 3567
MNE.US.x.MNE027            -TAT-AG-----T-A-ATA-----C-GTC-GGA-CA-G-A---TGT-----CC-AGA-GGC--GCCA---GAG--CACGG--ATA---A-TCA-G-CAAT--G---T-T---A-G---C-C-----AGAC..........---TA------GT--------T-----AG--A- 3388
DRL.x.x.FAO                -TAT-A------T-A-CTA--AA--------G-A-A-G-----G--A------C-G---GA----CC-C---A---CA-G-----A--A-T---TA-T--A------T-------G--AG-A-----TGAT...AAC--GCCA--A----------------T-AGG-A- 3188
MND_2.x.x.5440             -TAT-A---------AGCTA--AG-GC-G--GGA---------------C--CC-G---GA----CCT-----G---A-GG-G--A---TTAACA-----A--G----GT--C-----AG-A--G--AGAC...AAC---CC---C--GGT------G-----C--G-CA- 3182
MND_2.CM.98.CM16           -TAT-A---------A-CTC--AG-GC-G--GGA---G--------A--C---C-A---GAG--GCCT-----G---A-GG-T--A--ATT-ACA-----G-------GT--C-----AG-G--G---GAT...AA---GCCG-----GGT---G--------C----CA- 3622
MND_2.GA.x.M14             -TAT-A------T-AGTTA--AG-GC-G--GGA------------T------C-G---GG----CCTC-G--G--CA-GG-------A-T-ACA--G--A-------GC-----G--AG-A--G--AGAG...AA---GCC---T--GGT---G--G-----T--G-CA- 3549
OLC.CI.97.97CI12           -T-A-A---G-----------T----G---GA-CA-G-A------ATG---A-AGAGGAGG-GCCTC-CCAGTGTACT...TTATGCT-TACAA-G--AACGGTAG-----AG---A-GG--GA---AT......GG--CCT-A--G-------GG-T-CA---G-A-- 3489
GRV.ET.x.GRI_677           -TATAAGA-C--TCA--G---AG-GC-G---A--CA-G-A---ACA--C--------TCTC-----A---------AC-G-T--A-----A--AG---G-----------C---T-CACC-----AGGG......GC-GTAT-A--GGTG----G------C-AGCAA- 3794
VER.KE.x.TYO1_patent       -TAT--------TGC---------T---AC-GA-CAGG-A---ACC-----CA-A---GG--T-CCTA-T-GG----C-G-------ATT--AAGG---A--G----GT--C---T-CA-A-----AGG-......C--GTCT-----GT---------CA-C-AGCAA- 3845
VER.KE.x.9063              GTAT-----G----A---A---T-G---AC-GA-CA-G-A---ACT-------CA---GG---GCCTC-T-G-----C-G-C-----ATTA-AAG-T--G-------GT-----GT--A-A--G--AGG-..........---TCT-----GT-------G-------A-CAA- 3848
VER.KE.x.AGM155            GTAC-A---------AG-----C-------AC-GAGCA-G-A--TACT-------CA---GA----CC-C-T-GG----C-G-------ATTA--AG-T--G------T-----C---T-CA-G--G--AGG-..........---TCT-A--GGT---G--G-TC--C-A-CA-- 3840
VER.DE.x.AGM3              -TAT-----------A------CT-----ACGGA-CAGG-A--GACC--C--CA-A---GG--G-CC-A-C-G-----C-G-G--A--A-TA-AAGG---T-------GT--C---T-CA-G-------GG-......C--GTAT-A---GT---T-----CA---AGCA-- 3343
TAN.UG.x.TAN1              GT-CAG-C----TCA------CT---------GAGCA---G----CT--C----------AGT---TCT--G-G---CACT--TACC----T--ATG-T--G------GG-------T-CA-A--G---GG-.........---GTA--A------------------A---AG-GC- 3820
SAB.SN.x.SAB1C             -T-A-A------T---C----CT---G-C--AAGCAG--G---CAA------------TG-TCTTCCT----GG--TA-GG-T-TT--A-T---AG-T---------GG------G--A--C--GCCAGA-...AA-----TA-----GGTT--C-----------A---C- 4019
MND_1.GA.x.MNDGB1          -TAT-A-C----T-A------CT-GGCT---GA-CA-G-A--GTCT--------T--TAAT--G--A---TAT-T-AG-T-T-----AACAACAGG--AG-TATAT---T----TGGA-G----GA.......AA---GGT-T-A-G-G----G-------GG-A-CA-- 3308
DEB.CM.99.CM40             ---TAG---G----AG--A--C--G--GC--GA---GTCA-----T--C---A-AGAGCAGG--CCA------------T--TCC--ATTA---A----A------GG----GTG--CAGA---CCTAAG......GG--TAT-A--G----------TCT-T---GAT- 3427
DEB.CM.99.CM5              ----AGG--G--T-A---A-C--CC---C-GGA---CTCA--G--A--C--CA-AGA-GATG-GCC-C-----T----G--TTCC--AATA--AAG---A--G---GG---GT----AGA---ACAAAG......GG-TTG--A--------G----TCT---AGGAC- 3421
TAL.CM.01.8023             -T---------------CA------GC---GGAGACC-----TTCA-----CA-GGA-GG-G--CCA---GA------GC--ACC--ATT--CAG-----------GGC--CATGG-AA-A--G-GAAA-......CGGCTG--A--G-----G--G-T-----AGCA-- 3085
TAL.CM.00.266              -T-A-----------G--A--CT--GC----GA-ACT--C--TACA-----CC-GGA-GG-G-GCC-C-GGAG------C-GACC--ATT--CAG----T------GGC---ATG--AA-A----GCAA-......CGGCTCT-A--------C--G-T---C-A-CA-- 3587
SUN.GA.98.L14              -TACCAGAGT--TCAG---G-AT--C-G--GT-----TCA--CAGT--C-----A---GAT-----A-----TTGCAG-G-G--A---GTAAA------AAT--TA--C-T----TGG-TG--G-GAAA-......C--GTAT-A-GGGT-----G----CA--AG-GAG 3978
WRC.CI.97.97CI14           GT-A-A-ATC--T-AGC----AT--------GA-CA-TCA----CC--C---A-AGG-AA-G-CTTAA--GC--A-TT----A-A-CAGG-CCAGG-CAATGGGCCTATACAAT--GAG-AAT--G-GA-...AAG---CCA------GT---C-------GG---CAA- 3918
WRC.CI.98.98CI04           GT-A-A-AT---T-A-AGC-----------GGAGCA-ATG--G-CT------ACTGGGAA-G--TTAG-GGC--A-TT----A---CAGGATCAGG-CAATGGGCCTATAG-ATT-GAG-AGC--GAGA-...AAG---CCAT-A--GGT---G-------GG--GCAA- 3933
WRC.GM.05.Pbt_05GM_X02     ---A-A-ATC--T----CT--------G---T--CA---A---AGT-------GAGG-GG-G-CTTAA--GCT-AG-T----A-A-CAGGAACAGGGCAATGGGCCTACACAAT--AAG-AGC--GAGAG...ACA---CCAT-A---GT--------T-A-------C- 2847
LST.CD.88.524              -TAT-AGACC--T-AGACA------------AA------G--G-CT--C-----T---AGT-----AC-TG---T-AGGG----AC--A-CAA-A----GGTTATA--T-TC---TGGAGG----GAAAT......GCC-T---A-G-G------G----CA----CAA- 2889
LST.CD.88.447              -TAT-A-C-----CA-C----A---------AA-C-G--G----CT--------T--TA-T-----A--GG---T-AG-G-----C--AGTAA-A-G--AATTATT--C-TC--GTGGAGA----GAAAT......TCT-T-----G-G------GG---CA----CAA- 2892
LST.KE.x.lho7              -TAT-GGC-G----A----G-A--T------AAGA-G--G----CC------------GA---G--AC-T--T-TCAGGG-----C-AA-CAAAA-G--GAT-ATT--T-TC--GTGGAGA----GAAAT.......--C-TCT-A-G-G-T--G-GG---CA----CA-- 3976
SYK.KE.x.SYK173            -T-A-A---G--T---C-A--CT----G---GA-CAGGCA--GTCT-----------TAA----CCT---GA----C-C--CACT--A-TA---TCCCAG------GG---CATG---A-G-------A-AAAGGCCC-CCC--A-T---------GACA----A--CCT 3821
SYK.KE.x.KE51              -A----T-----T--------A--C----T-GA-CA-ACA--GTCT------C-G---GA----CCA---GA----C-C--TAGC---TTA--A---CAA--G---GGG--CAT----A-A----G----CAAGAA---CC-----TC---------ACA-G--A-GCAT 3479
COL.CM.x.CGU1              G--ACA---T----A-C----AA-C------GA---T--G--------C---ATA---CATG-G---A-TTGG-T---C--TTTC-G-TTA---A----A------GGC---GC-G-C-GC---A-A-AT......GG--CG--A-G--G-------C--AAC-C-GGC- 3390
```

```
H1B.FR.83.HXB2          GGGGTGCCCACACTAATGATGTAAAACAATTAACAGAGGCAGTGCAAAAAATAACCACAGAAAGCATAGTAATATGGGGAAAGACTCCT...AAATTTAAACTGCCCATACAAAAGGAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCCTGAG 3788
Pol                     R  G  A  H  T  N  D  V  K  Q  L  T  E  A  V  Q  K  I  T  T  E  S  I  V  I  W  G  K  T  P  .  K  F  K  L  P  I  Q  K  E  T  W  E  T  W  W  T  E  Y  W  Q  A  T  W  I  P  E
H1A1.UG.85.U455         --TC---T---------------------------A-TG---------G-GT--------------------------------TC---...--------G---A-------------------------------G---------TG-----------G--T-------------A 3234
H1B.US.90.WEAU160       -------------------------------------------------G-----------------------------------------...-------------A-----------A----------------------------------------------------------------- 3787
H1C.ET.86.ETH2220       --AC---------------------G--G-------C--T---------G---G--CTG---------------------------------...-------G-T-A-----C--G--A--------------G------------C-------------------------A 3180
H1D.CD.84.84ZR085       -------------------------G----------------------G---T-----------G--------------------...-------G---A-------------------------------T---C------------------------------ 3310
H1F1.BE.93.VI850        --TCG----------------------------A-----A----G---G-TTT-------------------G----GAT--CCT--G-------A-------TG--A--G---------T-------------T-------------------------A 3128
H1G.SE.93.SE6165        --TC----------------------G---------A-T------------G-------G---------------------------C...-------------A-------G--A--------------T-----------C---------G-------------------- 3185
H1H.CF.90.056           -AAC----------------A----------------A-----------G---T-T-------------------------A-T----...-------G---A--T-------A------------G--C---------------------------A----------A 3135
H1J.SE.93.SE7887        --A-------------------------G----A-TG-----------G--TTG---GC--------------------A------...-------G---A---------GA-----------G-------------C--C-----G--------------------- 3102
H1K.CM.96.MP535         -ATC--------------------G----------A-------------G--------G--------------------A------...-------G---A----------A-----------G--------------A---------------------C-----A 2984
H101_AE.TH.90.CM240     --TC---T---------------G----------A-TG---------G--G-------------------------------C---...-------G---A---------GA--------------------TG------------G--T-------------A 3362
H102_AG.NG.x.IBNG       --TC---------------------G--------------T---------G--G-T-TG---------------------------C---...-------G---A---------GA--G------------------TG------------G--T-------------A 3313
H1N.CM.95.YBF30         -ATC------T---------A-------G---GTT--A-TG--AAGG---G-GG-A--------T--------T------------------...-------G-T-A--AG------------GTG-----GG----------C--TC-----------A--T------------ 3382
H1O.BE.87.ANT70         A--CCT-------A------A---G-------G----A-T-A-C--G--GG-GT-TCA----TCT---A-T--C---------ATTG---...--------G-----AG-CACT-GA-----T-----------------G-G--C--------------------------A--A 3843
H1O.CM.91.MVP5180       A--CCT-------A------A---G------GG----A-T---C--G--GG-GT-TCA----GCT-----T---------G--ATTA---...------C-GG-----AG-TACT-GA-----T--------T------G-----A---------G------------------A 3818
CPZ.CD.90.ANT           CA--AA-T-----------G--T-GG------G-T-GA-T-------------GGATT-------T-----T--C------G--GTA--A...------C--T-A------ACT-G-------T------TG-C-------T------C----------G------------A--A--A 3227
CPZ.CM.01.SIVcpzCAM13   -ATCAA-A--T--A------A-T-G------GG----A-TCA-A---------TG-TGTG-----T-----C--T---------A--A---...-G-----G-T-A--AG--------A----GT-----G------------G---------C--------G------------C--C--- 3368
CPZ.CM.05.SIVcpzEK505   --TC-A-T--T--------CA---G-----------T--A-TG--A-----GG-GG-ACT------T-----T--C---------A---------...-----C-G-T-A--AG-G-----A---GT-----G--C--------------------------G--A--T------------T 3360
CPZ.CM.05.SIVcpzMB66    -ATC---A--T-------------------G------T----T--A-----G--G--CT------------C--C--------AGTA---...-G---C-GG--A--T--------------------------GG-G--------------C---------------A--A------C-----A 3334
CPZ.CM.98.CAM3          -ATCA--A--T--------CA---G-------G------T--A---------G-T--T-----T--------C----------A--A--A...------C-G-T-A--TG-T---------GGTC---------C------T---------------------------------C--A--C 3199
CPZ.GA.88.GAB2          -A-CA--T-CT--------CA---G----C--G-T--A-TG----C-----G-TT--C-T-----T---A-CC-T--------CTGT----...---------GT-A--AG-G----G----GT------G-C----------------T-----------------------C-----A 3181
CPZ.CM.98.CAM5          -ATCA--A--T--------CA---G-------G------TG--------------G----T-----T--------C----------A--A--A...------C-G-T-A--TG-T---------GTC---------C------T-----A-------------------------A--A--C 3490
CPZ.CM.05.SIVcpzMT145   -A-C---A--T-----------A-T-GG--GC-TG-T--AA----A----------G-T--------------T-----------------A--A--A...--------G---C--TG-T-----A----TG------G-C------G---T--C-----G--T--------------- 3330
CPZ.TZ.00.TAN1          CA---A-------C-----GA---G-------G---GACTGA-A--G------GG--AT--G------A----T-----T-TTGTG---...------TT-T-A--TG--TCC--A--G-----AGCCAG---------T--T--C-------TT------G-A------ 3432
CPZ.CM.05.SIVcpzLB7     --TC---T--T--C--------------G---G------T-A-------G---G---T-------------C-----------A-TA--A...--G----G-T-------T-----A---G-----------G---------------T------------------T------C-----A 3343
CPZ.US.85.CPZUS         -ATCA--A--T----------A-C-G---GC--GTT--A-TG-------------G-T--T--G-----T--C--T---------GTA--A...----------T--TAG-G--G-----GTC-------------T----A----------G------------------A--T 3848
CPZ.GA.88.GAB1          --TCA--A-----A-----CA-C-GG---C-GG-G--A-------A--------TG-T--T---------------T------------A--A---...------GGT-A--AG----G--A----GT-----G------G----A--C-----G--A----------C-----A 3849
MON.CM.99.L1            -AA--A------T-C---ACCTATC-G--G---G----T--GC-AGCC---T-GGGA-A---GGC------C-C-------CGACT---C...-TT---CGC--A--AG-GGTG--A--GCA-----CG----------G-----TA---------AT--AT---G-A--A--C 3491
MON.NG.x.NG1            AAAA-------T-C--CACATACC-G-----GG----T--T--AACC---C--GGG-AG--GGC-T-G--C--T-----------CTA--A...-T-----G-T-A--AG--A-C--A---CAG-----C----------G-C--CA-----------T---T------A--A--- 2084
GSN.CM.99.CN166         -AAA---A----TA--CTCCTATC-G-----TG----T--TT--ACC---G--GGT-GG--G--TCAG--T-CC----------GTA--AGAC-------GGA-C--AG--GTC-------CA------TG----------TGA-T--------G--A--A------A--CACC 3516
GSN.CM.99.CN71          -AAC---A--T-TA--CTC-TATC-G-----GG----T--TA--ACC---G--GGT-GG------TCAG--T-CT---------AGTA---GAC-----C-GGA-C--AG--GTC------GCA--------T--G---------TGA-C--------G--A--A----------CACC 3498
MUS_1.CM.01.1085        --TCA-TA----------AGCTATC-G-----GG----C---A--GCC-------GGA-G--------C-----C--T-----G---C-C---GAC--GG-C-G-A-C--AG-GATC--A---CAG-----CCAG---------T--CAC-C-------G-T-T-G------A---ACC 3478
MUS_1.CM.01.CM1239      -ATCA-T---T--C--CAG-TATC-G--G---T-T--T--CA-AGCC-G-----GGA-GG---TCAT-----C--T-----G---C-C--CGAA--GG-C-GGA-T--AG-GATC------GCAG------TCA------------AC----------G-TAT-A------A------C 3468
MUS_2.CM.01.CM1246      -AACA-----T-G---CCCCTACC----GC-GG----A--CC-ATCT-G-GCT-GT-AG--GGC-C-----CTGT------G-----C--AGAC----GC-GGA-A--GG--GTC------CAG-----C-AT--------G-----C-CA------A-A--T-----------C--C 3576
MUS_2.CM.01.CM2500      -AACA-----T-TC---AC-TACC-G---C--G-----------T-AGC---GGCTG-A-G----TC-T-----CTGT--------------C---GAC----G--G-A-A--GG-GGTC------GCAG-----CCAC-------G-----A-CA-----GA-T---A---G------A--- 3535
RCM.NG.x.NG411          A-ACA--A--T--A-----AA-T-G-ATGC-GG---GCTT---A--G--------TG--A----GCT--------T----------CT---A...-------G-G-----A---G---GA---GT-------CAG-------T-G--T--C------G-TTT--------------A 3282
RCM.GA.x.GAB1           AAACA--A------C-----AA-T-G-ATGC-GG---G-TT---A----------TG-A-A----GC-------T--T-------G-CTG--A...-C---C-G-------A---G-G-GA--GGT-------CTA-----C--T-CC-A--C------G-TT--------A--A--T 3267
DEN.CD.x.CD1            AA--AAG---TTA---CCCATATC---------G-CA-A-TGA--GC--------GGA-G----GC-T--------C-----C-GACT-----...-CT--C-G---C--AG--AGC-----T-GAT------TG--------T-----A-----------TT-AT------A-----C 3678

MAC.US.x.239            A-AA-A-A--T--C----GA--G-G--T----G--C-T-T-A-A--G------GGA-AG---GCA-----G--C------C--GTC--A...-----CC-CT-A--AG-TG-G-----TGT-------CAG-----------C--------G-TA--------A--G--A 4087
Pol                     I  K  N  T  H  T  N  G  V  R  L  L  A  H  V  I  Q  K  I  G  K  E  A  I  V  I  W  G  Q  V  .  P  K  F  H  L  P  V  E  K  D  V  W  E  Q  W  W  T  D  Y  W  Q  V  T  W  I  P
H2A.GW.x.ALI            AAAA-A----T--C--C-GG--C-G-TTG---G--C---T---T---------GGA-A----GCAC-G--C--T------CGA-TA--A...-G----C-C--A--AG--G---GA-----C------CAG-----GAT--C--C-------TG--A------C--A--C 4163
H2A.DE.x.BEN            AAAA-A----T--C----GG--C-G-TT-C--G--C---T---T-G--------GGA-A---GGCAC----C--T------CG--TA--A...-------C-C----AG-GG-G-GA--G--C-----GCAG------GATA-C--C--------TG--A------C--A--- 4170
H2A.SN.x.ST             AAAA-A-------C--C-GA--C-G--TCC--G--C-A-T---T---------GGA-A----GCAC----C--T------CGA-TA--A...-------C-C--A--AG--G---GA--T--C------CAG------GATA-C--C--------TG--A------C--A--C 3615
H2B.GH.86.D205          AAAACA-G-----C--C-GG----G--T-C-GG--C-T-T---T-G-------GG--A----GC-C-----C--C------G---TA--A...GTG--CC-T-----AG--G---GA--G---------CCAG-----------T--C--------TA---------C--A--- 4143
H2B.CI.x.EHO            AAAACA-T-----C----GA----G--T---GG-TC-T-T---C----------GGA-AG---GCAT-G--C--C------G---TA--A...-TG--CC-T--A--AG--G---GA--G---------TCAG-----------T--C--------TA---------C--A--A 4141
H2G.CI.92.ABT96         AAAA-A----T--A----GA----G-TT---GG-TC-T-TG--C-----R---GGA-AG---GCAC------------------ATTA--C...TTC---C-T-----AG--G---G---C--C------CA-------------C-----------TA--T------C--A--T 3497
H2U.FR.96.12034         A-AA-A-T--T--A----GA--TCGC-TGC--G--C-T-TG--A--G--G---GGA-A----GC-----------------A-T---A...-TG--CC-T-----AG--G---GA--GGTC-------CA---------T--C--------G-T-------G-C-TA--C 3652
MAC.US.x.EMBL_3         A-AA-A-A--T--C----GA--T-G--T----G--C-T-T-A-A--G------GGA-AG---GCA-----G--CGTA---C--GTC--A...-----CC-CT-A--AG-TG-G-G---TGT--------CAG-----------C--------G-TA---------A--G--- 3561
SMM.US.x.H9             AAAA-A-A--T--A----GA--C-GCTT----G--C-T-T-------G------GGA-A----GCA---A----T-------C--GTG--A...-G----C-TT----AG--G-G-RA----TT-------CA------------T-----------TA---------A--A--A 3560
SMM.US.x.PGM53          AAAA-A-A--T--A----GA--C-G-TT----G--C-T-T---CG-G------GGG-A----GCA--------T-------C--GTG--A...-------C-TT----AG--G-G-G------TT------CA------------T-----------TA---------A--A--A 4004
SMM.US.x.SIVsmH635F_L3  AAAA-A-A--T--A----GA--C-G-TT-C--G--C-T-T-------G------GGA-AG--GGCAC--------T-------G--GTG--A...-----CC-TT----AG--G---GA-----TT------CA------------T-----------TA---------A--A--T 4090
SMM.SL.92.SL92B         AAAA-A-A--T--A----GAA---GGTT-C-GG-TA-T-T---T-----G---GGA-A------TC----G----------------A--A...TTT---C-C--T--AG--G---G---GGTT-----TCA------------T-----------T--T----------A--- 3518
STM.US.x.STM            AAAACA-A--T--A----GA--T-G-TT---GG--C-T-TGA-T--G--G---GGG-A----GCA-----G--C-------C---TC--A...-----CC-CT----AG--G-G-GA--GGTC-------CA-------G----T----------G-TA--A------A--A--A 3734
MNE.US.x.MNE027         AAAA-A-A--T--C----GA--T-G--T----G--C-T-T-A-A--G------GGA-AG---GCA-----G--C-------C--GTC--A...-----CC-CT-A--AG-CG-G------TGT-------CAG-----------C--------G-TA---------A--A--A 3555
DRL.x.x.FAO             A-AA-AGT-----C------A-G-G-ATG---G---GACT---A-----G---G-A-A---GGCAC-G-----C--------G-TTG--C...-TT--CT-CT-A--AG--G---G----GTT------CA-------CAG-----C-----G-TA--A------C-----T 3355
MND_2.x.x.5440          A-AA---------A-----AT-G-GGGT-C-TG---G-TTG--A-----GG--G-T-AG--GT-TC-----------C---------CTA--G...-----CT-T--C--GT--G-G-G---GGT------CCA-------CAT--C-------------TA--A------C--A--- 3349
MND_2.CM.98.CM16        A-AA---T--T--A-----GC-T-G-GTGC-TG---GACT---A--G--G--TG-A-A----GCTT----G--C--------ACTC---...--G---T-T-----A---G---GA---GTC-------CCAG------C----A------------------A--------C--A 3789
MND_2.GA.x.M14          AACA---------A-----GT---GGGTGC-TG---GATT---A---------G---A---GT-------G--T-------G-ATTG--A...-----CT-T--C---C-GG-G-G----GT------TCAG-------CAT--T--------G--A--A------C--A--A 3716
OLC.CI.97.97CI12        T-AC---T--T--A---A-GT-T-G---GA-G----GA--CA-AG----G---GGA-A-------T-----G--T----------TTA--G...-TT---C-GG-A--T-CTAC--GAA--GAT------G------GCATTT--TG----------TACAT------A--A--C 3656
GRV.ET.x.GRI_677        -A-AAA-T--T---------C---GGACTC--G--C-CCTT--C-------CTGT-AG---GCAC-TACC--T-------GACT---A...CG-G-AC-----C--AG--G-C---A-----------T-TG------CAG--C----------TAT-------A--A--A 3961
VER.KE.x.TYO1_patent    A-AACA----T--A-----AC-TCGCAC----G-T-GTTT------G--G--TTG--A---GCTC----T--T------G-TATTA--A...GTTC-AG----C--G---G---GA--GGT--------CA-------G-G--T--C------G-TA-G-----------C--A 4012
VER.KE.x.9063           A-AACA----T--A-----AC-TCGCAC---GG-C-GCCT---A---------CTGT-A---GGCAC-G--C--T------G-ACTA--A...-TT---G----C--AG--G---G----GT-------CA-------G----C--C------TG-G-------------A 4015
VER.KE.x.AGM155         AA-C-A-T-----C-----GT-GCGTGT-C--G---GA-T---A--G------GGG-A---GGC-C--------T------C-ATTA--C...-CT---G----C--AG-GG-G-G---C-----------CA-------G----C-----------T--GT------A--C--A 4007
VER.DE.x.AGM3           AAAACA-T--T--C-----GT-CCGTGT---GG---GATT---A-------C-TTGT-A----TCTT-----T---------G--TTG--A...GTCC--G----C--A---G-G-G-----GT-------CA-------G-T--T--C-----G-TA-GT---------A--C 3510
TAN.UG.x.TAN1           AAAA-A-T--T--A-----GT-T-G-GT---GG---GAATG--C----GGG-GTG--AG---GCG--CACCT-T---------ATTA--C...-TCA-GG--G-A--AG--G---GA---GT------CA-------GT--C----------T---A---C-A--A--- 3987
SAB.SN.x.SAB1C          AAACA--A-----C-----AT---G-ATG---G---GCCT---A--G--G---GGA-AG-----------C--C------C---TA---...-TTA-GG-------AG--G---GA--GCTC------CA-------T-T--C---------G-T---------C--G--A 4186
MND_1.GA.x.MNDGB1       AAACA--A--T-G------CC-C-TGA----GG---GT--TAC---G--GG--GGA-G------T--------C-----TTTTGTA--A...---A-GC-GA-A----CTAC--G---G-T---------GAT------CAT------CAT-----------GTGT--A------A--A--A 3475
DEB.CM.99.CM40          AA--A--T---TA-------T-CC-T------G-TA----TA--T-T------GGA-------------TT-T--------GA-TA--A...C-G--C-GG--C--AG-GGT---A---GA------C-AT-------CATA-TC-C---------G---------C--A--C 3594
DEB.CM.99.CM5           A---AA-T--TTA------CT-TC-T---C--G-TA-----A--T-T------GG------G--T-----CTAT--------AGT---C...--------GGT----AG--GT-------GA-------T-AC-------CATA-CC-C-----G------T------C-----C 3588
TAL.CM.01.8023          -AACA--A----G---CCCCTATC-----C--GTG-GA--CA-------------GGA-A---GGC------CT-T------C--GTG--C...-T-----GGG-A--AG-GGTC------GGAG-----GCA--------T---------------TG-T-A----C-G--CACC 3252
TAL.CM.00.266           -AACA--A----G---CCCCTATC----G----T--GA--TA----G------GG--AG--GTC------CT-T-----GC--GTA--A...GT---C-GGA-C--AG-GGTC------GGA-------CAG-------------AC-C------TG-T-A-----A--CACC 3754
SUN.GA.98.L14           ---CA--A--TGAG---CC-TGTC----G---G---CT--CC-A-----G---GGA-G-------------T----------TTTGTG--A...--GA-AC-GG-A-----------GA--G-T----AGTCAG-------G-----C--C-----GTGT--A------C------ 4145
WRC.CI.97.97CI14        -ACA-A-T--T-G------GT---G-TC-A-GG-TC-T--CA----G--G---GGG-------T---------C--------G-GTA--C...-----C-G-A-A------G--------T-T------TT-T--------T----C-----------TGT-A------C--A--C 4085
WRC.CI.98.98CI04        -ACA-A-A-----C-----GC-T-G-AGC--GG-TC-T--CA----G------GGG-------T---------C-----------GTA--C...---------GA-A------G--------T-TC------CT---------T----TC-C-----G-T-T-A----------A--T 4100
WRC.GM.05.Pbt_05GM_X02  --CA-A-T--T--A-----GC---G-AC----G-TC-TTGTA-A--G------GGA-----GGCT--------T-----G--ACT---A...-------G-A-C--AG--G--------T-T------C-TG------T-----A--C------G-----A------C--A--- 3014
LST.CD.88.524           A-A----T----TG---CCCC-TC-G------G----A--T--A--G------GGA-A-------------C--T-----TTTTGTA--A...--GA-AC--G-A--TG-CACT-GA--------------AT---------T-A-C-T---------TA--T------A--A--C 3056
LST.CD.88.447           AATCA--T-----A---CCCC--C--A----GG----------A----------GGA-A---G--T-----C--T-----TC-TGTG--C...--GA--C-GG-A--GG--AC--GA---GTG-----TCA---------GT--TC-------G-T----------A-----A 3059
LST.KE.x.lho7           A--CA--A-----A---CCCC--C-GA-----GT---A--TA-T--G--G---GGA-A-------------C--C-----CTTTGTG--A...---A--C--G-C--AG-GACT-GA---GTC-----GCAC-------GC--CC-C-----G-TT--A----------A--C 4143
SYK.KE.x.SYK173         TT-CA--A--T-G------CTATC--TCTC-TG-TC-ATTGT-AA-T------GGG-TCC--TCAC--TGGTAT----------GTA--A...-C---CC-T--C--AG--A---GA--GGA-------A-----------------C--------------A---G-A--A--A 3988
SYK.KE.x.KE51           AT-CGA----TCAC------TACC-GGC---GG-TC--CTGA--A-T------GGA-T-C--GC-C--TGG--T---------A-TC---...G----CC-C--C--AG--A---G----GA--------A-------------TC-C----------A--A---T-G--A--A 3646
COL.CM.x.CGU1           AAAA-CAG---TT----AG-A-GC-GG-G---G-TTCA-TTA-A-----G---GGT-GG---GC-------C-CC----------GTG--C...---A-G---G-C--AGC-A-G-GA--GGAT------CAG--------GT--C---------AGTG-T------C--A--- 3557
```

p51 end and p66 RT continue Pol p15 RNAse H start

```
H1B.FR.83.HXB2          TGGGAGTTTGTTAATACCCCTCCCTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCCTAAC 3958
Pol                      W  E  F  V  N  T  P  P  L  V  K  L  W  Y  Q  L  E  K  E  P  I  V  G  A  E  T  F  Y  V  D  G  A  A  N  R  E  T  K  L  G  K  A  G  Y  V  T  N  R  G  R  Q  K  V  V  T  L  T
H1A1.UG.85.U455         -----------C------------C---A--------------------------C-------C---------G--A-----------------------T----A----GC------------G-----C---G-C-------------G-----TT------- 3404
H1B.US.90.WEAU160       -----------C--------------------------T---------------------------------T-----------------------A---------------------------------C-------------G------T------- 3957
H1C.ET.86.ETH2220       ------------------------C----A---------------C-----------------C-----T-----T-----------A--------T----A-----A-----------G---------G---------G--G---A----TT-T----- 3350
H1D.CD.84.84ZR085       -----A-----C----------T-----A-------------------G----------A------------T-------------------T-AA--------------------------------G-C------------G---------C--T-T-- 3480
H1F1.BE.93.VI850        -----------C--------C--TC--------C-------------A-C---G------C---------C--------------------T----T--A-----C---AA--------------------G---A-----A-----------T------- 3298
H1G.SE.93.SE6165        -----------C----------TC----A--G--------T-G------A-C---------CC-----------T-A------------------T-------A-----------G---------------G-C-A-----AG------A--A-T------- 3355
H1H.CF.90.056           --------------C-------ATC----A----------T-------A-C---G------C-----------T-A---A-----------------T-----A-------------------------C---G----------AG------------T------- 3305
H1J.SE.93.SE7887        -----------C--------C--TC----A--G------------A----A--G--------A-G----------T----------------T--------------------AC------G-----G---------G-C-A---------G-----A--T------- 3272
H1K.CM.96.MP535         -----------C-----T------C----A---C----------------A-C----------------G-------T-----C--------------C-T------A-----AAG---G---------------G-C--------------G------T--A---- 3154
H101_AE.TH.90.CM240     ----------------------TC----A-------------A-----A-----C-----------------G--T------------------------GT-----------GC-----------G-----C---G-C-------------G--A---TT------- 3532
H102_AG.NG.x.IBNG       -----------C----------TC----A-------------------A-----C-----------------------T-----------A--------T--A---------GA-----------G-----C---G-C-------------G-----TT------- 3483
H1N.CM.95.YBF30         -----A-----C--C--T------C-T--A----------T--------A-C---G--A--CAGT--G---------T-------------A--------T----A--A-----G-----------T-T---G--AG-----G--------G-----G---T-TA-TG- 3552
H1O.BE.87.ANT70         -----A-----C-GC--A--C--A--GA-C------------AG-C----A-GT-----T--TA-G--G-----------AT--------------A--------T--A-----A---C------G-------------AG-ACA---G-A---G--GA-AA-A-AAT--GA 4013
H1O.CM.91.MVP5180       -----A-----C-GC--A--C--A--GA-C--------------------A-C-----T--T-----G-----------T--------------A--------T---A-T--A---C------G--G-------------AG-ACA-----A---G--CA-AA-A-AGT--GA 3988
CPZ.CD.90.ANT           -----------------A--C--A---A-T-G-C-C-----TA-TC--TT-GC---C--T--TCC--AG--T--------T-----T--------------T--AA-CT--C----G--G--G-----C-----G--AG-C---------AGT-GG--AAAACA----CA 3397
CPZ.CM.01.SIVcpzCAM13   --------CA----C-----C--A-----A-----------TAGT-------C---C--T---CC-ACG-----C--A-A--------------A-----A--T-----A--A--GAA--------------CA-A--AG-C-AG---A-------A-AA---GT---GA 3538
CPZ.CM.05.SIVcpzEK505   -----------C--C--------AC----A-----------T-----------C------A--TAGC--G-----------AT-------------A--------T-----A-------------G---------T---G--AG----G--T--G-----G--GAC-T-TA-CT- 3530
CPZ.CM.05.SIVcpzMB66    -------A------C-------------A---C-T--------A------C----C------CC-----------A--T---------------------------A--A--C---------G--------------A--AG-C-A---------------GA-CA--T-TT---- 3504
CPZ.CM.98.CAM3          -----N--C--G------------C-T--A-----------------A--G--A-C---------CC-----------G--------G-----------------A-----A--GAA-------------C--C--C---G--------C-----G---A-A--A-A-T--GA 3369
CPZ.GA.88.GAB2          -----A--C--C-GA-----A---C-T--A--G--------T--A--G--A--------T---CA--------------AT-----T----------C--T--------A-G-AG--------------------T-----AT-----A---G--G--CA-TT--T-GGA 3351
CPZ.CM.98.CAM5          -----A-----A------------C-T--A-----------------A--G--ACC---------CC---G--G----A-------------------------A--A--A--GAA-------------T--CA----AG----G--C-----G---A-AA-A-ATT--GA 3660
CPZ.CM.05.SIVcpzMT145   -----A--CA-A--C-----A---C----A--G-------------------C---G--T---A-------------A-A----------------C-----A--A--A---AG------------C-----G--A--C--G--G--------GA--CAT-GT--TGA 3500
CPZ.TZ.00.TAN1          -----A---A----C-----A--AC--A-C-GGC--------A-TC-GTT-TCT--C-----CCC--A-------------T----------------A-----A--C-G----AAG----G-----------------A--A--C-----C---T-C-GGTCAAAGGA-T--GA 3602
CPZ.CM.05.SIVcpzLB7     -----A-A-------------T--CAT--C--G--T--------A-------C------T---CC--G---------A--T-----------------------A--A----G-------------------------A--AG----G--G-AG--G---A-AA--C--T---- 3513
CPZ.US.85.CPZUS         -----A-----C------------C-C--A---C-T-----TA---------C---GG-----AG--G---------A--T----G----A----CC-AC-A--A--C--GAA-----G-----------------G---G--T---------A--A-A--TT--GA 4018
CPZ.GA.88.GAB1          -----A---A-C--------A---------C-----------AGT-------C------T---CC-ACCA---T--T-AT------------A-----A--T-----A--A---ACT--G---------------A--AG-C-A---G-A-------A-CA-T-G-T--GA 4019
MON.CM.99.L1            ATA---GCG--GT--T----A-A-C-GC-C-G-CAG-----TACTC-G-TAC----C-------A---G--TCC---A-A---C--G-------C--C--T---ACAT-A--GGA---G--------T-----A--C--CTTT--GGC-AG---G-CC-----AT--GA 3661
MON.NG.x.NG1            ATA---CCA--AC-CT-T--A-A-C-GC-T-G-CA-------ACC----T-C--R-T-----T-A-CAC--TCCC----A-----------------G--------A-CAT-A---GA---G--G------------C--C-G--C---CCA-AG---G--AA-TC-AT--GA 2254
GSN.CM.99.CN166         ATA--AGCAA-AC-C--A--A-A-C--C---G-CA-------ACCC---TCTCG------C---A--AT----TT--T-AT-------------A-----A--T---ACCT-C---C-----------T--T-----G--C----C-----A---T-GG-CCA-AGAAT--GA 3686
GSN.CM.99.CN71          ATA--AGCA--AC-C--A--A-AT--GT---G-CAG------ACCC---TCTCG------C---A-AAT----TC----AT--------C--------A--T---ACCT----------------T--C-----G--C----C-----A-T-T-G--CCA-AGAA---GA 3668
MUS_1.CM.01.1085        ATA--AGCA--GC-----A-C-A-C-GT-A---CAG----T-ACA--G-TACC------AT--AA--A---TCCC--A-AT-----G-----A-----A--T----TAT-C-----G---------------C--C--A--C-C---G-A-G----G--AA-T-G----GA 3648
MUS_1.CM.01.CM1239      ATA---GCA--AC-----A-C-A-C-GC--CG-CAG----T-ACA--G-T-CC---G--AT-GTC--A---TCCC--A-AT-------------A-----AC-T-AA-TAT----G-----G--G------------C-C-----A-G----G--G--T-G-T--GA 3638
MUS_2.CM.01.CM1246      ATA--AGCA--AC-C-----ATA-C-TC---G-CAG----TTACC--G-TCCC---G-------A-AC-----CCC----A------G-----A-----A--T--A-CT-----C-----G---T--------A--CG----G---CAGG-G-G--C-A-A-ATT-GGA 3746
MUS_2.CM.01.CM2500      ATA---GCG--GC-C--A----A-C-GT---G-CA-----T-ACTC-G-T-CC---G--TC---AG-CG---CCC---AT-----T---------C----T--AA-CT-C--------T-GG-----G-----C--AG--------CAGG----G-CCA-AGA----GA 3705
RCM.NG.x.NG411          -----A-------G---T--A--AC--A-A-GGC--------A-TC---TA--G--T--T---CC-----A---GT--AT---------------------A--T---A-T-G---G--G-----G------------A--AG----G-----A-G----G--AAAAGAAT--GA 3452
RCM.GA.x.GAB1           -----A-------GC--A--A--TC-TA-T-G-C-CG-----A-CC---TT--G--T-------CC-----AG---GT--A-------G-------------------AA-T-G----A----G--G-----------A----------A-G----G--AAAAGAAT--GA 3437
DEN.CD.x.CD1            -----AA-G--ATCA--T--A---C-TC-A-G-AAT-TT--TA-T--G-TATC---G-------ACAA---GCCC----AT---AC-----A-----A--T--AC---------CA-----G------------GG-C-CC-AT--C-AG-----G--A--A---T-GGA 3848
```

Pol RT end Pol p15 RNAse H start

```
MAC.US.x.239            -----T---A-CTCA--A--A--GC----A-G----GTC-T-A-TC---T---G--C--T----AG----A--------AT---AC------AT--TG---T-AAC--T-A---GA---G------------A-C--AG----G--C-A-G-C-----AAAAGTGT--GA 4257
Pol                      E  W  D  F  I  S  T  P  P  L  V  R  L  V  F  N  L  V  K  D  P  I  E  G  E  E  T  Y  Y  T  D  G  S  C  N  K  Q  S  K  E  G  K  A  G  Y  I  T  D  R  G  K  D  K  V  K  V  L
H2A.GW.x.ALI            -----C-----ATC------A--AC-G--C-GGC--GCA-TTA-CC---TA-----T--T---C----C-----G--------CAC---C---C-CTG---T---C-AT-A---GA----------------A-A--AG-------G---G-C--G--GAAGGTA---GA 4333
H2A.DE.x.BEN            -----C-----ATC------A--AC-G--C-GG---ACA-TTA-CC---TAGG---T--T---CC---C-----G--------CAC------AT--TGC--T--AC--T-A---GAG------------------A--AG---------A-G-C-----AAAAGTAT--GA 4340
H2A.SN.x.ST             -----C--CA-ATC------G--AC-G--C-G----GTA-TTA-CC-G-T------T------C----C----------------CAC------AT-CTGC--T-A-C-AT-A-G-GA----------------CA-A--AG-----------G-C--G--GAGGCTAT--GA 3785
H2B.GH.86.D205          -----C-----CTCG-----A--A---A-A-G-C--GCC---A-CC---TC-----C---C---A---GAG--------A---CAC-------T-CTGC--T--AACCT-A--GGA-------------------C---G-C--G---A-G-T--G---AAAGTGT--GA 4313
H2B.CI.x.EHO            -----T-----CTCA-----A--A---A-A-GG---GCC--TA-CC-G-TC-----C---C---A-----T------T-A---CAC------AT-CTG-----AA-CCT-A---GA---G--------------C--AG-C--G---AGG-T------AAAC-AT--GA 4311
H2G.CI.92.ABT96         -----C-----GTC---A--G-ATC----A-GG---GTC---A-CC---TC--------TC---AGCAGA----G----AT---AC---C---T-CTGC--T-AAACCT-C--GGAR-----------C--C--C--AG-C--------A-G-C--G--AAARGT-T--GA 3667
H2U.FR.96.12034         -----T-----ATCC--A--A---C-G----G-C-GGCA--TA-TC---TA--------AT--C--...-A---G--A-AT---AC------CT-TTG---------CTT-A---CAG--G---------C-----G--AG----------A-G-C-G---AAAAC-TT--GA 3819
MAC.US.x.EMBL_3         -----T---A-CTCA--G--A--AC----A-G----GTC-T-A-TC---T---G--C--T----AG-T--A---------AT---AC------AT--TG---T-AAC--T-A---GA---G------------A-C--AG----G--C-A-G-C-T---AAAAGTGT-G.. 3729
SMM.US.x.H9             -----C-----GTCA--A-----------C-G-C--GTT-T-A-CC---TA-----G--T---CAG----------A--T----------A-AT-CTG---T---C--T-A-G-GA-------------C-----G--AG-------C---GGC--GACAAAACTG--GGA 3730
SMM.US.x.PGM53          -----C-----GTCA--A-----------C-G----GTC-T-A-CC---TA--------T---CAG---------G--A--T------------AT-CTG---T---C-AT-A-G-GA------G------------C-----A--AG-------C---G-C--G-CAAAAC-TT-GGA 4174
SMM.US.x.SIVsmH635F_L3  -----C-----GTCA--A-----------C-G----GTC-T-A-CC---TA-----G--T---CAG------------A--T--C---------AT-CTG---T---C--T-A-G-GA------------------C-----G--GG----G--C---G-C----CAAAACTTT--GA 4260
SMM.SL.92.SL92B         -----C--CA-CTCC--A--A--G---A-C-GG---GT--TTA-T----TA-----T--A----A-AAG-AG--GTA-AT--CA-T------CT-CTGC--T--AA-C-G----GA---G---------------C--AG-C--G----A-G----G--CC-GC-AT--GA 3688
STM.US.x.STM            -----T-----ATCC--A-----T-----A-G----GTC-TTA-T----TA--------T-G-A----A-------A--T-----G-----CT--TG-------C-AT-----GA---G---------C--CA-A---G---------A-A-C-----AAAAG-TT--GA 3904
MNE.US.x.MNE027         -----T---A-CTCA--G--G--AC----A-G----GTC-T-A-TC---T---G--C--T----AG----A---------AT--------------T--TG---T-AAC--T-A---GA---G-------------A-C--AG----G--C-A-G-C-----GAGAGT-T--GA 3725
DRL.x.x.FAO             -----A---A-ATCA--A--C-ATC-CA-C--G---------TA-C--GCT--GT--C--A---CC---G-A---GTT-AT-----G-----A------------A-TT-G----GA----------T--C--------CTCC--G----A-G-------AA-AG--A--GA 3525
MND_2.x.x.5440          -----C--CA--TCA--A--A--A---A-A-GG---------A-CC-GCTA-----C--A---CC---G-A---TGTA-AT-----------------A-----AA-CT-C--GGA---C--G-----C---TA---CTC------T-A-AGC--G--A--AG--T-GGA 3519
MND_2.CM.98.CM16        -----A-----ATCC--T--A-A-C-TA-CGGG--G-----TA-CC-GTT--G------TG-GCC-----A---TGTG-AT-----G-----A------------AA-CT-C---GA---G------------A--AGC----AAT-A-TCC-G---AA-AG-----GA 3959
MND_2.GA.x.M14          -----A--CA-CTCA--A--A--AC--A-A-GG--G-----TA-CC-GTTA--------A---CC---G-AG--TGT--AT-----G-----------A--T--AA-CT----GGA---G-----------CTA---AGC---G----A-AGT-----AA-AG--T--GA 3886
OLC.CI.97.97CI12        ATA--------ATC--ATT--GAAAATCAA-GGAC--TC-GGG-------T-CC--C-------G...-AG-T---------------G-----A--TA---T---A-A-GC--GGA------G------------C-G-TCAG-G----AG--T---A-A--ACA-A--CA 3823
GRV.ET.x.GRI_677        -------------GC--A--A-T-C----A---C-G-----TTCC----TA---------A--CAA-----A---TGTT-AT-----G-----------ATC--AA-T---C--------T--G---------C-GT-AG-G-------A-AGT-G-A--AGGGAAT--GA 4131
VER.KE.x.TYO1_patent    -----T-----C-GC-----A--T--GC-C---C---------ACA---ACA-----------CCCAAG-AG--CGTT-A-----------A--TGC-----AA-TT-A---GA---------------CA-CT-AC-ATAC----A---G-G---A-AA--AT--GA 4182
VER.KE.x.9063           -----A-----G-G-GTA--A--TC-----------G------ACTC-GACC--G-----T---CC-----AG--TGTA-AT--------------CTG---T--AA-CT-A---GA---------------CA----AC-ATAT----A---G-G---A-AG-AAT--GA 4185
VER.KE.x.AGM155         -----C-----C-G-GTT--G---------A-CT--G-----TACAC-GACT--G--------CCCG---AG--TGT--A-----------A--CTG---T--AC--T-G---GAG-----------C--CA-A--CC-ACA---C-A-----G---ACAACAG---GA 4177
VER.DE.x.AGM3           -----A-----C-G------A---C----A------------TACCC-GACA-------------CC-AAG-A---TGT--A------G-----A--TTG---T--AA-TT-A-GGGA-----------------A-C--AC-ATAT--G-A-----GG--G-AA-AAT--GA 3680
TAN.UG.x.TAN1           ATA--A--CA-CTCC--A--G------A-C------------TA-C--G-TA-----T------CC-AA----A-GGTG-AT---TGC---------TGC--T--AA-TT-C----------G-----C-TCC-A-----C-C----GA-TC-C-G--CA-AGAAT--GA 4157
SAB.SN.x.SAB1C          -----AA-G--C-G---T----AA--GA-T-G---G------A-A----TA-----T-----CCC-----A--C-GT--AT--------------------T--AA-TT-C---GA---------------CT-A--AG-------GGAT-----G--G--AG-AT--GA 4356
MND_1.GA.x.MNDGB1       GTA--A---A-C-GC--A---ATG----AA-GGGA------TAGC--GTCCCC------TC---AG--G-T------A-AT-----T-----A-------------C-G----A-G--------------A----AG-------TTTT----GG--A-AAGAATATCT 3645
DEB.CM.99.CM40          ------GCAA-AC-------A-ATC-G--C--GC-C------TG--C-G-T-TC----------CC--AT-----C----A-----------------------T-C-------G-----T--C-----C--AG-GT-G----A-----GT--AAAATG-T--GA 3764
DEB.CM.99.CM5           ------GCG--GC-------A-AAC----C-GGC-------TG--C-T-TATC---G--A---CC--AT--T--C----AT-----------A--------T-----AT-A---A-G----------C--CA-AT-AG-AT-G----A---GGG---AAAA-GT--GGA 3758
TAL.CM.01.8023          ATT-TCCCCA-AC---------------C-GGC-------TA-CC---TATC---G-------A------------A-A-----G-----T--T--CC----A--AT-A--GGA---C-----G--------C-----G--C-CGTC-----A-G--C-T--A--AGGA--T-G 3422
TAL.CM.00.266           ATC-T-CCCA--C-C--A--A--A-----T-GGC-------TA-CC---T-C-G--G-------A---------C--G-AT--C-----C--T-----CC--C-A---T-A---GA---C--G--------C--G---GCA-C-----A-G--C-T--CA---AGT--GA 3924
SUN.GA.98.L14           -TA--A-----C-GC--T---AAG--G-A-C--GAG-----TACTC---CC-C------AG-GCC----...--C----A------G-----A------G-A-A-CTAGAA--GAG---G-----------A-A--CC-GG-------TCC-G----AAG-AAT-GGA 4312
WRC.CI.97.97CI14        ACA--ACA---GTC---G---TT-C-T-CT-GGGA------TA------T-CT------TC--CC--A--T------T-AT--A-------AT-T--A-----A--C-G----GA---G-------------...TGG-CC-A-TC--A-G----------A-GTG-GCA 4252
WRC.CI.98.98CI04        ACA--ACA---ATCA------TA---G-CA-GGGA-------A-AA---TACTG--G--T-------AG-TG----A-AT--CA-------AT-T--A--T--A--C-G---GGA---G--------C-TC...TGG-CA-ACTCT-A-G-------AA-T-ATT--TA 4267
WRC.GM.05.Pbt_05GM_X02  GTA--ACA---GTC---T---TTTC--AAA-GGGA------ACA----TATC------T----A---G-T----GAAA-G---A--------T-T-----T-AA----GC---GA---G--------------...TGG-C--AGACT--GA-C--GA-ACAG-AAA-C-- 3181
LST.CD.88.524           -TA--A--CA-CTCC--A----AA----AAC-GGA------TAGC--G--AGC-----------A---TA----T----A------G-----A--T--AG-A--A-TAGG----AC-----------G---A-A---C-ATC-----A-G-G---A-AAAAGAAT--CA 3226
LST.CD.88.447           -T---A---A-CTCA-----A-AG----AAC-GGA------TAGC-----AGC------T----A---G-T---T----AT-----------A--T--AG-A-AA-TAGGG---AC------G-----T-----C--AC-ATC-----AGG-G------AAAGAAT---- 3229
LST.KE.x.lho7           -TA--A--CA--TC------G-AA----AAC--GAG-----ATT-GG---GC----------A----G-T---C----A-----------A-----AG-A-A--TAGGA---AC----------------A-A--AC-ATC---G-A-G-G-----AAAGGAGT---- 4313
SYK.KE.x.SYK173         GTCA-A--CA-A-G------G--A--G--T-GG-GG-AC--TA-C--G-TACC----------CC--A-----T---------C---------A-------T--A--C-GC--GAC------T----------AG-A-G-GAT--G-C---G-G---ACAGTAT--GGA 4158
SYK.KE.x.KE51           GTAA-A-G---GC-C-----CATA-----T-G--GG-AT---A-T--G-TTTC---G--AG--...-A---------A--T-----C--------C-----T---A-T-GC---GA----C-------T--CA-A--AG----G--T-A---G-C-T-AAAG-GGT-GGA 3813
COL.CM.x.CGU1           ATA--------CTCAT--T--TATG----A------GTA-GGA-C--G-TTTT---G--TC-GAAG-AG---CC---T-AT-GGAC------A-G-TGCTC----A-ATTAGG-GCG-----G-------GGA-A-A-TCA-A----GA-G--G----A----AT--TCA 3727
```

```
H1B.FR.83.HXB2           TGACACAACAAATCAGAAGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGATTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAACCAGATCAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAA 4128
Pol                       D  T  T  N  Q  K  T  E  L  Q  A  I  Y  L  A  L  Q  D  S  G  L  E  V  N  I  V  T  D  S  Q  Y  A  L  G  I  I  Q  A  Q  P  D  Q  S  E  S  E  L  V  N  Q  I  I  E  Q  L  I
H1A1.UG.85.U455          ---G-----------A--------A-----T-----CC-------C-----------A----C---------T-----------------G-----------G--------G-----------CAGG-----------AA-------------------A--C-----G 3574
H1B.US.90.WEAU160        -----------------------------------T---C-----------------------------------------G--------C-------------------------------------------G-----------------G------------A---------- 4127
H1C.ET.86.ETH2220        ---A--------------------A--------G--CC-G-----G--A--------A----C-----------------------G----------G----C-TG-----------A-G-----------A-----------------A---------- 3520
H1D.CD.84.84ZR085        ----------------------------------------A---------------------------------T-----------C---------------------------------A-G----------A------G------------A---------- 3650
H1F1.BE.93.VI850         G--G--C------------G-------------------T--------------A----C-----G------------------------------------------G--G----------A-G-----------AA---------------------A------C 3468
H1G.SE.93.SE6165         ---A--------C--A---G----A-----G------C-G-----------A--C--AA---C-----------------G--------G-----C-------------------------AGG-----G----A-------------------A---C------ 3525
H1H.CF.90.056            G--A--------------------A--------------------------A-----A--G--------G--------G-----T-----G-----C---------------------C---A-G---------------T-------------G-A-------- 3475
H1J.SE.93.SE7887         ----------------------AC----C--C------T-------A-G------A--GC---------------------T---------------G--G--T--A---------------A-G-----------A-------------------G--------- 3442
H1K.CM.96.MP535          ---A--------------AG---A-----------C-G-T-------A--------A----C-------------------------G------------------------G---------A------------T-----T----------------A------- 3324
H101_AE.TH.90.CM240      ---G-----------A--------A-----T--G--CC--T----C-----------A----C---------T----------------------------G-----------CAGG-----------AG-------C------------G--C------ 3702
H102_AG.NG.x.IBNG        ---G-----------A--------A-----T------C--T----C--A--------A----C---------T--------------------G------------G-----------CAGG-------------------------------------A--C------ 3653
H1N.CM.95.YBF30          A-----C--C-----A---G-------------T--CCT-A-G--C--A--A--G--A---CGG--T-----------C--T-----T--G-----TA-G-----A-----TT----G-------A---------------A--G--G-GC-----------AG--C-C---- 3722
H1O.BE.87.ANT70          ---G--C--C-----A--AG----A---ATG--G--A-TA-----C--A--------CAA-GA-AC------T-----------T----------------G--CG----CTCCT-C-----TACA--G--------CCCTA----TC-G------------G-AC---C-- 4183
H1O.CM.91.MVP5180        A--G-----C-----A---G----A---ATG--TG-A-TAA----C-----------CAAGGAGC-------------------------------T---G--C-----ATCCT-C------ACA--G-----C--CCCTA----TC-G--G---------G-AC---C-- 4158
CPZ.CD.90.ANT            AA-G--C--C-----AC-AG-A--A-----G--G---CT-A-G---C-AG-------AAC-GGCCC---C--T-----C-----T--------------G---G--T-G----GTACC-------------------CCCC----GG-AG-------CC--A-A-------- 3567
CPZ.CM.01.SIVcpzCAM13    GA----C--T-----AC--G-A------GGG--CT-G-----------A--A--C--T-A-CA-C----G----------T-----T-------TG-----G--T-----GT-----------------------C-----------A-----A----------------AG-A-----T- 3708
CPZ.CM.05.SIVcpzEK505    A--A--T--C--C--AC-AG----A-----G--TG-CCTCA-G--C--A--A---G-A---CA--------T-----C--T-----T--G---TCC-------A-----CT--------------------A-------------A-----A-----G----------AG-AC-T---- 3700
CPZ.CM.05.SIVcpzMB66     A--A--C--T--------AG-A--A--------T--AC-AT-G----A--A-----A-A-G----------T-----C--T--T-----------C-G-----A-------GT-----T---AC--------------AA-----A-----G----------G-AC-G--C- 3674
CPZ.CM.98.CAM3           AA-T--C--T--------AG-A--AC-GAC---TG----------A--A--A---G-A-A-CATAT---T--T-----C--T--T--C-------TG-----G--A-----CT-C--G------------------G-----A--G--T----T----------AG--------- 3539
CPZ.GA.88.GAB2           A--G--C--T--C--A---G-A--A---G-G---C-CAGAA---CC-A-----C--A--GCA---------TG-G-----------C-------TG--G---------AA-TT-C--------C---------G-----AA----G---------------AG--C------- 3521
CPZ.CM.98.CAM5           AA-T--T--T--------AG-A--AC--AC---TG-------G--A---------G-A-A-CATAT---C--T-----T--T--T--T-------TGC-------A-----CT-T--G----------------------A-----AGGGTCT---------AG--------- 3830
CPZ.CM.05.SIVcpzMT145    A--T--C--T-----A--AG-A--A-----------G---TAA-G--C--A-----C--A---CAG-----C--TG-G--C--T-----C-------TGC-G------T-A---T-----C-----C--C-----------TA-----A-----G-------AG--------- 3670
CPZ.TZ.00.TAN1           GA----C--T-----AC-AG-A--A---TGG---G-AG-------C--AA-A--C--A---GC-C-G-----T-----C-----T--C-------TTA-G---G-TT-A--G-G-TT-------------C--CCCCA----AG-G-----T--TC-AA------C-C 3772
CPZ.CM.05.SIVcpzLB7      A--A-----C---------G-------------C---C-AT----C--A--A--C--A---CCT--G-----T-----T--T--T---------------------C-------C-----G-----G-----T-----A--G--G--------C-----AA-TC------ 3683
CPZ.US.85.CPZUS          AA-T--T--T-----A---G-A---C-CAC---GG-A--CT----A--AA-A-----A-AGAATAC---T--TG----C--T-G---C-------TC------------C--CT-T--G----C--------------A--G-------G--------AG-A------- 4188
CPZ.GA.88.GAB1           AA-T--C--T------C-AG-A--A---A-G--TT-GCT---G--C-----------A-ATCA-C-G--T--------G--T--------------TG-----G--T-----GT----G-------C-------------A--------------G--------AG------T- 4189
MON.CM.99.L1             AC-G--C--C--C--A---G-A---C--G-G--CG-G-TA-----CC-CAGA---GGACC-CCTC-GA-G------A-T-----T--T--G---T-C-G--C--TC-AGCCT-CTGC--G--GATCTCA---AGTCCCA----GG-GGCC--C---C--G-AC-CT--- 3831
MON.NG.x.NG1             GC-G--T--C--C------G----AC--G-----G-GCT-T----A-----AA--GG-CCCGCTAGTA-G--T--------------------G----TCA-G------C-AGC-T--A-G-----GGTCTCG--G--CCCAA----AG-GA----C--CC-----C-CC--G 2424
GSN.CM.99.CN166          A--A--T---------C-AG-A--AC-C--T---G--CT-T-G--A--AA-A--AGG-CC-CC-A--A-G--TC-G--C-----T--C--G---T-A-G--G---T-G---AGT--G-----GGT-TCCACC--T-C-C----GG-A-----C---C-AG-AC-GC--- 3856
GSN.CM.99.CN71           A-GA--T--------AC-AG-A--AC-C--G---G--CT-T-G--A--AA-A--AGGACC-CC-AG-A-G---T-G--C-----T--C-----C-T-A-G--G---T-A---AGC--------AGT-TCCACC--TCCC-----GG-A-----C---C-AG-AC-GT--G 3838
MUS_1.CM.01.1085         AAGT--T--C--C-----AG-A--A--GG-G---G-C-TAT-G--C--AA-A--AGG-CCTCCTAGGA-G--TT----C-----T--T--G---TC-----G--AG-AGCTT-T--G---C-AG--TCCACT---CCA-----AG-G----------C--G--C--C-GT 3818
MUS_1.CM.01.CM1239       AA-T--T--C-----A---G-A--AC--G----TG-C-TGT-G--C---A-A--AGGCCCTCCCAGCA-G--T-----C-----T--C-------TGC-G--G--AG-GGCTT-T-----CC-AG--TCCACC--GCCC--G--AG-AG-------TC-A--A---C-G- 3808
MUS_2.CM.01.CM1246       AA-------C--C---C-AG-A---C-C-----C---CTCT-G---C-AAT----GG-CC-CC----A-G---C-----------T--T--------------G-----A--CAC----GC------GGT-TCA--G--CCCTC-G--GG-G--G-------C-----AA-GC-G- 3916
MUS_2.CM.01.CM2500       AA-T-----C--C--A---G-A--A--G--G--T--A-TAA-G--CC--ACA---GGACC-CC-A--A-G---------G-----T--C--G-----CC-G--T-----CAC----GC------GAC-TCA--G--TCCCC-G--AG-GA-G--C---C-A-----GC-C- 3875
RCM.NG.x.NG411           AA-T--T--C-----A--AG-A-------GG----G-A-TA--G--C--AA-A--C--A---AA-AGG-------------------T--T-G-C---TG--T--C--AT-AGC--G-AGT-----CAC------GAGCCCCC----AC-A--G--T-----A--AA-G---G 3622
RCM.GA.x.GAB1            A--A--T--T-----A--AG-A--A---G-----G-A-TA--G--A--AA-A-----A--GCCTA-------------C-----T-----G---TT-AT-----TC-AG--------------------ACT-----C-----G------G-CAG----T---A-T---A-G---G 3607
DEN.CD.x.CD1             A--A-----C-----A---G-A--A---AT--G-G-C-TAA-G--C--A--A---AGTCCCAACA-G--C--------C--T--T-----------CC----C-----AGC--GTAC---CAC-AC-TC----CATC--A----GG-A--G--C---C-AGCTA-GCA-G 4018

MAC.US.x.239             AC-G--T--T-----AC-AG-A--A--GG-----T--CTCA-G--A---ACA--C--A--GCC-A-G-C---T--TA--GT---T----------TTA-G-----A--AAC--G-TGC--TACAG--TCA--GAGCAG-C----T-------------AG-AA-G--T- 4427
Pol                       E  Q  T  T  N  Q  Q  A  E  L  E  A  F  L  M  A  L  T  D  S  G  P  K  A  N  I  I  V  D  S  Q  Y  V  M  G  I  I  T  G  C  P  T  E  S  E  S  R  L  V  N  Q  I  I  E  E  M
H2A.GW.x.ALI             AC-A--T--C------C-AG-A--A---G---CT-CGCG--G--AG-AACA--C--A--TCC-A---CC--T--TA--GT---T-----G----T-A-G-----AG-AGC--GC--G---ACAG--TCA---AATAGAA----A---------C-----AG-AA-G---- 4503
H2A.DE.x.BEN             AC-A--T--C------C--G-A--A---G---TCT--CGGA-G--AC--GCA--C--A--CCC-A-G--T--T--CA--GT---T-----G----T-A-G--G-AG-AGC--GC--G---ACAG-GTCA---AATAGAA----G--C--G--C-----AG-AA-G---- 4510
H2A.SN.x.ST              GC-A--C--C------C-AG-A--A---G----CT--GCGA-G--AG-AACA--C--A--TCC-A-G-CC-----TA--GT---------------T-A-G-----AG-AGC--GC------ACAG-GTCA--GAGTA-AA----A---------C-----AG-AA-G---- 3955
H2B.GH.86.D205           AC-G---------C--AC-AG-A--AC-TG-----T--GCAT----A--AACA--C--A-A-CC-C-----T-----CA--GT---T-----------TCA-G-----A--AGCT-----G---ACAG---CA------CCAA----AGCAA-------T--AG-AA-G--C- 4483
H2B.CI.x.EHO             AC-A-------------C-AG-A---C-TG-----T--GCA-----AC-A-----C--A---CC-C-G--C--T--CA--GT---T-----------TCA-G-----AG-AGCT---------ACAG---CA------CC-A----A-GAG-------T--AG-AA-G--C- 4481
H2G.CI.92.ABT96          GC-G--------C---C--G-R--AC-TG-G---T--GCAA-G--C--------C--A--TCC-A----C--T--CA--GTR--T--C--G---TCA-G--R-----AGC--GR------ACGG---CA-----CCC------A---A-------T--AG-AA-R--T- 3837
H2U.FR.96.12034          AC-G--C--C--C--AC-AG-A--A---G----CT--GC-A-G--C--AG-A-----A---CC-A-----------T--G-----T---------AT-A-G-----A--AGC--G---G---ACAG---CA--G--CCCAC----A-G-A----C-----AG-AC-T---- 3989
MAC.US.x.EMBL_3          .ACA--T--T-----AC-AG-A--A--GG-----------AT-GCA-TG-A--C--A--GCC-A-GAG---T--TA--GT---A-T----GTATGT-AT--G-AT-A-A-CAGGTTC--TACAG--TCA--GAGCAG-C----T--C------------AG-AA-G--T- 3898
SMM.US.x.H9              AC-G--T--C-----AC-AG-A-----GG----CT----------C--AGCA-----A---CC-A---C---T--TA--GT---T--C-------TTA-G--C--AG-AGC--GT-----CAC-G--TCA---AGTAGA-----A--C--G----------A-AA-G--T- 3900
SMM.US.x.PGM53           AC-G--T--C-----AC-AG-A-----GG----CT-C---------C--AGCA-----A---CC-A-G-CC--T--CA--GT---T--C-------TTA-G--C--A--AGC--GT-----CAC-G--TCA---AGTAGA-----A--C--G---------G-AA-G--T- 4344
SMM.US.x.SIVsmH635F_L3   AC-G--T--C--C--AC-AG-A-----GG----CT-C---------C--AGCA------------CC-A---C---T--TA--GT---T--C-------TTA-G--C--AG-AGC--GT--G--CAC-G--TCA---AGTAG------A--C--G--------AG--A-G--T- 4430
SMM.SL.92.SL92B          AC-GG-T--C------C-AG-A---C-G--G--CC-G-TAT-G--C---A-A-----TCC-AGTA-G--C--TG-------G--------G---TC--GAAC--A--AAC--G---G---TCAG--TCG--T-----TA----GGCA-----T--T--A--AC--G-TC 3858
STM.US.x.STM             AC-G--T--T-----AC-AG-A--A---G-----T--GCCA-G--C---GC---C--A--TCC-A---C---T--T--GGT---------------TGA-G---------AAC--G---G--CACAG--TCG--GAGCA-------A-----G---------AG-AA-G--T- 4074
MNE.US.x.MNE027          AC-G--T--T-----AC-AG-A--A--GG-----T--CTCA-G--A---ACA--C--A--GCC-A-G-C---T--TA--GT---T----------TTA-G-----A--AAC--G-TGC--TACAG--TCA--GAGCAG-C----T--C------------AG-AA-G--T- 3895
DRL.x.x.FAO              A--A--C--C---------G----A-----G--C---CTGT----C---A----C--T--TCC-A----C--T-----G--G--T-----G--C--CA-G--T--A--ATCCT-CTCC-----AATC-----CAATCCTA----G--------T-----A--AA-G--C- 3695
MND_2.x.x.5440           A--A--G--C-----A---G-A--A--GAC---T--AA-GA-G---------A------A---CCCAGG--G--T-----C--T--T--C--G------C----C--AC--TC----GCC---------GTCA--CAACCCCA----T-GGG----------A-TCA-G---G 3689
MND_2.CM.98.CM16         GA-T--C-----C--A--AG-A--A--GG-G-----AA-AA-G--A--A--A--C--A---CC-A-------T-----G-----T-----G-----TA-G--T---C-ATCG---GC---C--C---TCA--CAACCCCA----A-GAG----T-----A-TCA-G--CC 4129
MND_2.GA.x.M14           AA-------T---------AG-A--A---G-G-----ACTGT----C---A-A-----A--TCCTAGG-C---T-----C---------------------------C--AC--TC---CACT-------G-TCA--CAACCCTA----T-GAG-------CA-C--AA-G---G 4056
OLC.CI.97.97CI12         G--G-GT-GT---------AG-A--AA--AG---T---ATAGA---CC-TAG---C--AAA-AGT--G--C--T--T--------T---TTG-----CA-G-----AG-GTT--AGA-------AGTGTC----GA-CCCC-CA-AC-AG--T-----A--G----GGAG- 3993
GRV.ET.x.GRI_677         AA----C--T--C--AC-AG-A--A---AC----G--A-GA-G--A---G----CAGT---GA-A-T-----T-----C-----T--T-------T-A-GAAC---T-GAC----TGT---C-GG-----A-C---CCC-----GG-A--G------C-AGCCC----G- 4301
VER.KE.x.TYO1_patent     AA----T--C------C-AG-A--A---AC---T--AA-AA-G------G-A--CAGT--GCCTA-T--G----------------T---------A-G-----TT-GAC---------CACA--------T---CCA-----AG-G-----T----CCTTAA-G---C 4352
VER.KE.x.9063            GC-G--T-----C---C-AG-A--A--GGT---C--AA-AA-G--C---G-----AGT---GA-A-------------------T-----------CA-G-----AT-GAC------G--CACA--------C--CCCA-----AG-A--G--C----CATTAA-GG--- 4355
VER.KE.x.AGM155          AA------------AC-AG----AC-GAC---C--AA-AA-G--C---G-----AGC--CCCTA----C--T-----------T-----------GA-G--C--AT-GAC------G--CACA--G-----C--CCCAC----AG-A-----------CA---A-GG--C 4347
VER.DE.x.AGM3            AA-T--------C---C-AG-A--A---ATG--C--AA-AA-G--AC-AG-A---AGT--GCCTA-T-----------------T-------------A-G-----AT-AACT--C-----CACA--G-----C---CCC---A-AG-A-----T----CA-TAA-GG--C 3850
TAN.UG.x.TAN1            G--T--C--G--C--AC-AG-A--A---------G-CCT------A--AA-AC--AGTAAGAGCA----T--T--------G--T-----G----TCA-G------C-GAC---------TACA--------C--C-T-C----A-----------T--AG-AA-G--T- 4327
SAB.SN.x.SAB1C           GA-T--T--C--C-----AG-A------G-G--C---CTGT----C--AAGA--C--T---AGCA-------T---A-------------------TA-G--T--T--AGC--GGG---TACAG--TCA--TAATA-CA----AC-A-----T------G-AC------ 4526
MND_1.GA.x.MNDGB1        AA-T--C--C------C----A--A-----T---G-AA-A------C-AG-A---AGT---AGTT-T--T-----------------T---------T-G----T--AC-AGC-AGCAG---TAC-G---CA--TCACCCCA----A--GG----------ATTAA-G-A-G 3815
DEB.CM.99.CM40           GA-T--C--T---------AG-A--A---G----C--A-TGT-G--C---G-A--AGGCCC--C-A--A-G--T--------------C--G------C-G--G--A---TT--A---C---TC-G-G-CA---CATA--A----AG-AA-GG-----C--GCC--GCA-G 3934
DEB.CM.99.CM5            AA-T--T--C-----A---G-A--A---G----C--CCTA--G--C---G---GGGACCC-CCA-GA-G-----------------C--G-----T--------AT-GTT--A---C--CACGG---CA---CAT---A----AG-AA-G--C---C-AGCC---CA-G 3928
TAL.CM.01.8023           CA-T--T--T---------G-A--A---G-----G-GCTAT-G--A--AA-AC-C--T--TCC-A-G------------C--------T-------TC-AT--C---T-AGT--G-A-T--CACAG-G-CA--CAACA--A----GG-AG-C--C---C-----C--T--G 3592
TAL.CM.00.266            AA-A--T--C--C--A---G-A--A--GG-----G-ACTAT-G--AC-CA-A-----A--GCCCA-G-----T-----C--------T--G---TC-AT--C--AT-GGC--G-AGC--CACAG---CA--TAACA--A--A-AG-AG-C--T--CC---TAC--C--- 4094
SUN.GA.98.L14            AA----C--C------C-AG-A--AC--G-G--T--AA-AA-G--C---G-----AGTA-GAGCTC---T--T-----------------G-----C---A--C-GC-CAGTAAGAG---CACAG---CA--T------C----A--GG-G--CG------TCA---GGC 4482
WRC.CI.97.97CI14         AA-T-----T------C-A--A--A--GAT----T-AA-A-----AC-AG-A---GG---TCA-A--A-G--T---A-C------AGT-------TCA-G--G--AG-AA---AT--T--CACAATG-----G--TCCT-----AC-GG----T--T--ATT-----A-G 4422
WRC.CI.98.98CI04         -A----C--T------C-AG-A--A---GCG--TT-GA-GT----CC-AG-A---GG---T-C-A-GA-C--T--TA-T-----TAGT-------TTA-G-----AG-AA---AC--C--TAC-ATT-----G----CTC----AC-AG----T-----T-TTC---A-C 4437
WRC.GM.05.Pbt_05GM_X02   AA-T--T---------C--------C-GATT---T-AAGAA-G--A--AG-A--CAGTTC-GATA-G-----T---A-C------AGT-------TT-----------AA---AT--C---AC------------TCCC---A-GC-AG----------C-TTC---A-G 3351
LST.CD.88.524            A--T--T--T------C--G-A--A---G-G---G-A-TGT-------A--A---AGCAA--C-AC---T-----------T-----C-------T-A-GAAGG--T-AAGTCAGAG---TACAG---C---GCATCCTA----G--GA-C--C---A--G--C-GCAG- 3396
LST.CD.88.447            A-----C--C-----AC-AG-A--A---G----TG-C-TAT-G--A--AAGA--CAGTAAGAGCAGG--C-------------------------TGA--AA-G-TT-ATCTCA--GG--TACAG---CA---CATCCTG----A--A--G--T------G-A-GT-A-- 3399
LST.KE.x.lho7            ------C---------C-AG-A--AC--G-G---G--CTAA-G--A--A--A---AGCAATAGTA-------T-----------T-----------T-A-GAA---AT-GTC-CA-AG----ACAG---CA---CACCCCA----A--AG-C--C-----A--A-GC-AGC 4483
SYK.KE.x.SYK173          GC-A--------C--AC--G-A--A---G-G-GCT-GCTCA-G--A--------C--AAA-GACA-------T------GT---T--------CT-T-AT--C--TC--ATGA--TGC--CAC-A-C-CA---CACCCAA----GG-G--G-------C-AG--GCC---- 4328
SYK.KE.x.KE51            AA-------C-----A---G--------G-----G-A-TAA-G--A--A--A--C--A---CC-------T--------------T--T--G---------------AT-AATGAACTGT--TACATCCTCG---CACCCAG-----G-A--G--T--GC-AA--GCT--GG 3983
COL.CM.x.CGU1            ---AGGGT--------C-AG-A---C-GAC--G-G---TG------C-AA-AC--GG-CC-AA-AG-A-G--TT-G---------AGT-TT-----------G--A---ACT-GG--G---T--G-TCAGCTTC--TCTAGTAGAGGGG-G--T--TC--GCAGGT--GG 3897
```


PLV Complete Genomes

Pol p66 RT, Pol p15 Rnase H end Pol p31 Integrase start

```
H1B.FR.83.HXB2          AAAAGGAAAAGGTCTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTCAGTGCTGGA...ATCAGG......AAAGTACTATTTTTAGATGGAATAGAT...AAGGCCCAAGATGAACATGAGAAA...TATCACAGTAAT 4283
Pol                     K  K  E  K  V  Y  L  A  W  V  P  A  H  K  G  I  G  G  N  E  Q  V  D  K  L  V  S  A  G  .  I  R  .  .  K  V  L  F  L  D  G  I  D  .  K  A  Q  D  E  H  E  K  .  Y  H  S  N
H1A1.UG.85.U455         ----------A-----C---T------------G--------G-----------------G------------------T-----..-------......--G--G------------G-------..-----T-----G--C-----A---..------T-C--C 3729
H1B.US.90.WEAU160       ---------------C--A--------------------------------------------------------T-A---..-------......-----------------G-------..-----------A-----------..------------- 4282
H1C.ET.86.ETH2220       GC-------G------C---T---------------T-----------------------------------A---AG----..-------......----G--G---C-----------------..-----T-----A--G-----A---..--------C--- 3675
H1D.CD.84.84ZR085       ---------------C-----------------T------T----------------------------------AG----..-------......----------------G------------..-----T-----G-----------..--------C--- 3805
H1F1.BE.93.VI850        ---------G------C---T------------------------G------------------------------------..G------......---A----G----------G-------..-----A-----A-------A---..-------AC--- 3623
H1G.SE.93.SE6165        ---------------C---T-------------------G-------------------C-----------AG----..-------......---------------C------C-------..--A-------A--G-------G-..--------AC--C 3680
H1H.CF.90.056           -G-------------C---T--------------------------------------------------T--T-----..G---A......----G-------C------G-------..--A--T-----A-------A-GG..-----T-AC--- 3630
H1J.SE.93.SE7887        -G-------------C---T-G-------------------------------G-----------------T-----..-------......----G--G---C------G-------..--A--T-----A--T-----A---..-----T--C--- 3597
H1K.CM.96.MP535         ---------G-A----C--AT---------------------------------------------------------..-------......----G------C-------------..-----T-----A-------A---..-----T-AC--C 3479
H101_AE.TH.90.CM240     -------G--A-----C---T------------------G--G----------------------------T-A---..-------......--G--G------------G-------..-----T-----A-------A-G-..--------C--- 3857
H102_AG.NG.x.IBNG       -------C--A-----C---T------------------G-------------------------------AA----..-------......--G----------------C-------..--A-------A--G-----A-G-..--------C--- 3808
H1N.CM.95.YBF30         ---------GA--T-----CT-T---------T-----T-----T-----------G--G-----C--------T--CT-A---..--T--A......---A--T----CC--------T-----A...--A-------A--T-----C-G-..--------C--- 3877
H1O.BE.87.ANT70         --------C----G-----TA-------T--T--C--T-----C--A-----------A--A-----------A--CAAG-AT..--T--A.....-G---C--G--CC----A--------C...C----A-----A--T-----A---..-----T------ 4338
H1O.CM.91.MVP5180       --------CGA--G-----TA-------T--T--T-------C--A-----------A--A-----------A--CAAA-AC..--T--A.....-G---C--G--CC-G--A-----------...C----A-----A--T-----A---..-----T------ 4313
CPZ.CD.90.ANT           ---GA---C--A-T--C--CT-C---------------T-----C--A--------------G-----C------------CAA--G..-----A.....C----C--C---C-G--A--C---------..--A--T-----G--C-----T---..-----T--C--- 3722
CPZ.CM.01.SIVcpzCAM13   -------C-----T--C--TT-C--------T-----T--G--C--A-------------------G--------T-----..-------......--------------CC-G-------------..--A--A--G--A-------A-G-..-----T--C--- 3863
CPZ.CM.05.SIVcpzEK505   ---------GAA-A-----TT-C--------T-----T-----T---------------GA----C----------CA-A---..-------......-----CT----CC-------T-------..-----T-----A-------A-G-..-----T------ 3855
CPZ.CM.05.SIVcpzMB66    ----A--G--A--G-----CT----------T--------------------------A------------------CT-A---..--T---......-----C--C--CC-G-----G-------..------G--A--G-----A---..--C--T-A---- 3829
CPZ.CM.98.CAM3          -G--------AAGT--C--TT-------G--------T------G---------------------------T---T----G..A-----A......-----T--C-------------T-------..T---CA-A-G--A---------G--A..-----T-AC--- 3697
CPZ.GA.88.GAB2          -G--A--G---------T--A--------------------------A-------------G--------G--------AG---C..-----A.....--G--G--------------T-------..-----A-----A--C-------G-..-----T--C--C 3676
CPZ.CM.98.CAM5          ----------AAGT--C--CT-------G------------------------------------------C----T---T-A--C..--T--A......-----C--C--C---------T--C---..-----A--G--A-------A--G..-----T-AC--C 3985
CPZ.CM.05.SIVcpzMT145   ----------AA-T-----AT-----------------G---T--A-----------G--------C----T---T-A--G..-----A......-----T--T--C--------------A...--A--T-----A-------A-G-..--C---TC---- 3825
CPZ.TZ.00.TAN1          -----AC-GCAA-T-----A-------------C--T-----T--A--G--T------G-------C-----G--T---AAAAAT..--T--A......---A--T---CC-G--------TA--...G-A--A--G--A--C-----T---..------------- 3927
CPZ.CM.05.SIVcpzLB7     ----------A--T--CT-AA----------T-----T--G-----A---------------------C-----------T-A---..--T---......-----C--T---C-G--------------...CGA-----G--A--G-----A---..--C--T-A---- 3838
CPZ.US.85.CPZUS         -G--A-----AAGT---A-CT----------------T-----------------------A--------C--------T-A--G..-----A......-----T--T--CC-------C-------..--A--A--G--A--C-----C---..-----T------ 4343
CPZ.GA.88.GAB1          ----A-----AA----C--CT-C---------------------T--A----------G-----G---------------A---..-------......-----G-----CC-------------C...-G---------A-------A-GG..-----T------ 4344
MON.CM.99.L1            -------CC-----A-TCT--T-------------T------------A-----------G--------CGCC----GG-CCAG---..--T---......C----C--G---A-G--AAATT-G--A...CCC--AAGG--A--T-----A--G..-----T------ 3986
MON.NG.x.NG1            G---AA-GC-----A---T--T-------C--G--C---------C--A--G--------G----R--C---------A--YCAA--G..-----A.....C-G--C--C--CA-G--GAAC--T--A...CCR----G---G--T-----A---..--------C--- 2579
GSN.CM.99.CN166         --------GCCA----CT-AAGT-----T--G---------G-----A--------CAC-G-G--G--C-------------CAAG----..--C-C......C----G--T--CC-G--AAAC-----G...CCT--AGT---A--C-----A--G..-----T------ 4011
GSN.CM.99.CN71          GG--A--GGCC------CT--AGT-----T----------------G--A--------CAC-G-G-----C-------------CAGA--G..--C-C......C-------C--CA-G--GAAC-----G...CCC--AGT---G--C-----A---..-----T------ 3993
MUS_1.CM.01.1085        C---A---GCCA-A--CT-AT--------G-----C--T--G--T--A------------GG----G--C---------A-A-AG----..-------......C----C-----CA-G--ACAG-----A...CCT--AA-G--A--T-----A---..------------- 3973
MUS_1.CM.01.CM1239      CC-----GGC---A--C--AT-G-----G-----C--T-----T--A------------GG-T--G--C---C-G--A--CCA-----..--T---......C-G--G-----CA-G--GCA------G...CCT--AA----G--C-----A--G..--C--T--C--- 3963
MUS_2.CM.01.CM1246      GC--AA-TGCCA---T-A-AT-CC-A---------C-----G--C--A-----G-----GG-------CC-CC----T--CAGG----..--C-C......C-G---------A-G--AAAT-----C...CC---AGTT--A--C--------G..--C--TTC---- 4071
MUS_2.CM.01.CM2500      CC--A---GCAA-A--C---T-C-----------C--T--------A--G-----CC-GG-------CC-CC----T---AAA--G..--C-C......C-G--T--G---A-G--AAAC--T--G...CCA---AT---A--C-----T---..--------TCC--- 4030
RCM.NG.x.NG411          GG-------G-A--G--C--CT-C---------G-----T-----C-----G---------G--G----------------G--CAG----..--T--A......C-G-------CC----C--G----GA...--A--A--G--A--------A---..-----T--C--- 3777
RCM.GA.x.GAB1           GG--A---GCA--G--C--TT-C-----G--C-----T--G--C--C----------GG-------------------T--AAA----..-----A......C-G-----G--CC-------G-----A...--A--A-----A--------A--G..-----T-A---- 3762
DEN.CD.x.CD1            GC------G-AA-A--CA-CA-C-----T-----C-----------A-----C---AC-G--A----C--GC----T--AAA--G..-----A......C--A-C-----C-----ACA---TCCA...C-A--T-----A--------T---..--C--T--C--- 4173
```

Pol RNAse H end Pol p31 Integrase start

```
MAC.US.x.239            -----TC-G-AA-T---G-A----------------------T--A--------CC--G--A----CC-CC----T---CAA--G..--T--A......C----T--C--C--G--AAAG-----G...CCA--A-----A--------T---..--C--T------ 4582
Pol                     I  K  K  S  E  I  Y  V  A  W  V  P  A  H  K  G  I  G  G  N  Q  E  I  D  H  L  V  S  Q  .  G  I  .  .  R  Q  V  L  F  L  E  K  I  .  E  P  A  Q  E  E  H  D  .  K  Y  H  S
H2A.GW.x.ALI            --------GCCA-----G-T--G-----C-----C--------C--A--G-----C-GG-------CC-T-----A---CAG--C..-----A......C-----T-G---C---AAA------G...CCA--T--G--A--------A--G..-----T--C--- 4658
H2A.DE.x.BEN            -G------GCA------G-T--------C-----C--T-----C--A--------CC-GG-------CC-T-----A---CAA--C..-----A......C-----T----CC-G--AAAG-----G...CCC--T-----G--------A---..-----T--C-T- 4665
H2A.SN.x.ST             --------GCAA-----G-T--------C-----C--T-----C--A---------C-GG-G------C-C-----A---CAG--C..-----A......C-----T----CC----GAA------A...CCC--T--G--G--------A---..-----T--C--- 4110
H2B.GH.86.D205          ----A--GGCA--A---G-A-G------------T-----G---C-G--T--T---C-GG-------CC-CC----A---CAA---..-----A......C-G--CT-G--CC----AAA------A...CCA----G--A--G-----A---..-----TG-C--- 4638
H2B.CI.x.EHO            ----------AA-A---G-A-G------------T-----G---C-G--T--T---C-GG-------CC-CC----G--CCAA---..--T--A......C--A-C------C----AAA------A...CCA--T-----A--------A---..-----T-A---- 4636
H2G.CI.92.ABT96         ----A---RCTC-A---G-A-G------------T--T--------A-----T---C--G-G-----CCRTC----A--CCAG---..--T--A......C-----------C--GR-AAA------A...CCT--T-----A--------A--R...-T---T------ 3992
H2U.FR.96.12034         GG--A---GCA--G---G-T-GC--------T--T--T--G--T-----T--C---C--G---------C-T--G--G---CAG---..--T---......C-------G---C----AAA------A...CCA--T--G--A-------------..--C--T--C--- 4144
MAC.US.x.EMBL_3         --...-TC-G---T---G-A----------------TTAG----T--A--------CC--G--A---G-CC-C----T---CAG--G...T-T--A......C----T--C--C--G--AAAG-----G...CCA--A-----A--------T---..--C--T------ 4050
SMM.US.x.H9             --------GCAA-T---G-A-----------T-----T---RR---A----A----C--G--------C-TC-----T---CAA-A-..--T--A......C----C-----C-----AAA------A...CCA--G--G--A--G-----A---..--C--T--C--- 4055
SMM.US.x.PGM53          --------GCAA-T---G------------T--T--T--------A---------C--G--------C-CC-G--T---CAA---..--T--A......C----C-----C-----AAA------A...CCA--A-----A--G-----A--G..--C--T--C--- 4499
SMM.US.x.SIVsmH635F_L3  --------GCAA-T---G-A-----------T-----T--------A---------C--G--------C-CC-G--T--CCAG---..--T--A......C----C-----C-----AAA------A...CCA--A-----A--G-----A--G..--C--T--C--- 4585
SMM.SL.92.SL92B         -G--A---GCA--A--CA-A-G------T-----T--T-----C--A-----G--CA-TG-G-----CCGT-----T---CAA---..--A--A......C-G--C--G--CC----AA-T-----G...CCA--A-----A--C-----C---..-----T--C--- 4013
STM.US.x.STM            --------GCAA-T---G-A-----------------T--------A--------C-GG-G------C-T--G--A---CAG---..--T--A......C-G--C-----CC----AAA------A...CCA--T-----A--G-----A---..-----T--C--- 4229
MNE.US.x.MNE027         -----AC-G-AA-T---G-A--------G---------------T--A--------CC--G--A----CC-CC----T---CAA--G..--T--A......C----C--C--C--G--AAAG-----G...CCA--A-----A--------T---..--C--T------ 4050
DRL.x.x.FAO             GT--A---GCA--T--CT-AAAT-------------T--------A-----T-----GG-G-AG--C---C-----T---AGA---..--A--A......C----GT----CC-----AAC--G--A...--A--A--G--A--------TTT-..--C--T-A---- 3850
MND_2.x.x.5440          GC------G--A-T-----CT-------C-----C-----G--T--A-----T---------GA-------------A---AAG---..--A--A......C-G---------------G----------..--A-----G--A--------C---..--------A---- 3844
MND_2.CM.98.CM16        -C------GCA--A--C--A--T--------------------C--A-----G-----G---------------C-----G--CAGA---..G-A---......C--------G--------A--C-------..--A--A-----A--G-----T---..--------A---- 4284
MND_2.GA.x.M14          CC--A---GCT--A--CT-A--------------T--T-----C--A-----C---------A-----C-----------A--CAAA---..-----A......C-----------CC-G--A----------..--A--T--G--G--------C---..--C--T-A---- 4211
OLC.CI.97.97CI12        -G--A----GAA-A---A-T-G------T---------T--G--CT-A--G--C---C--G--------C-------------AAA---..--A---ACTAAGC----GA-G-----G-CA-AC-----G...CCA--A-----GAGC--CA-T--G..-----TCAA--- 4154
GRV.ET.x.GRI_677        -----AGGC-------CT-ACA-----------T--T--G--G--A-----C---AC-G--A--------------A--CAAA---..--A--A......C-GA-C--C--C------A-------A...G-A--A--------C----CA--G..--C--T-AC--- 4456
VER.KE.x.TYO1_patent    -----C--C-AA-A---T--CAG-----------------T--------A-----------GG-GA-------------G---AAA--C..--T--A......-G---TT-----C-----AAA------A...G-A--T-----A--G-----A-G-..-----T-A---- 4507
VER.KE.x.9063           -G--AAGGGCAA-A--C--ACA-------G--T-----------T--A--G---------G--A----C-----------A---CAA---..--A--A......-G---T------A---GAA-------A...G-A--T-----A--------C-G-..-----T------ 4510
VER.KE.x.AGM155         -G--A---GCCA---------CA---------T--T--T-----T--A--G--C------G--A----C-----------A--CAAG---..G-T--A......-G-A--T-G--CA-T-GCA-G-----A...G-A--A-----A--------T-GG..------------C 4502
VER.DE.x.AGM3           ----AC-TC--A-A--CT--CA-------------G-------G--A-----C------G-GA--------C----A---CAA--G..--G---......---A-TT-----------AAA------A...G-A-----G--G--------A-GG..--C--T-A---- 4005
TAN.UG.x.TAN1           -T--A--TGCTA-----T-AA-C------------T---------G--A--G-------C--G--A---------GC----A--CCAG--G..--T---......C----G-----C-----CA----G--A...G-A-----G--ATC------T---..------T-C---- 4482
SAB.SN.x.SAB1C          -G-----GGCA--G---A-A----------T--C--T------G-A--T--C-----GG--A-T------C-----T---CAA---..--A--A......C----------CC-G--CA----T--A...G-A--A-----A--G-----T---..--C--TGCC--- 4681
MND_1.GA.x.MNDGB1       G---A-----AA-T---T-AAGT---C-------------------G--A-----G-----G--A--------GC----A--T-A---..-----A......-----CT----CC--C-AAAT-----A...CCA--A--G--A----------------..-----T--C--- 3970
DEB.CM.99.CM40          GG--A---C-AA-T-----A--------------C-----G-----A-----------C--G--------------------AAA---..--A---......-GGA-C--C--C--G--AA-GG----G...G-A-------G--C-----A---..--C--TG-G--- 4089
DEB.CM.99.CM5           --------C----A-----TAAT------------C-----------T--A-----G-----T--G-----C-----G--A--AGA---..--A---......-G-A-C-----C--G--AA--G----A...G---T-----G--T---------------..-----TG----- 4083
TAL.CM.01.8023          GC------GCT--G--C--TA-----C-------C--T--------A--------------------C--GC-G--G---CAA---..-----------......-G-A-T--C---G----ACA---CCCA...G-A--A--G--A--G-----A-GG..--C--T-AC--- 3747
TAL.CM.00.266           GC-----GGCA-----CT----T---C-G--G--C-----------C--A--------------G--G-----G-----A--CCAA---..------------......--GA-C--T---G----ACAG--CCCA...G-A--A-----A----------G-..--C--T-A---- 4249
SUN.GA.98.L14           -------CC----A--C--A-GT--------G-----------------A---------C--G--A-----C--C----T--CCAG---..-----AAAGAGAC-G---A-G---A----AAAG-----A...CCA--TGT---A-------G---G...-TC--T-AC--- 4643
WRC.CI.97.97CI14        GG--A---G-AA-A--CA-A-AT-----------T--T-----G--A--G--C---AC-G--A------G-----G-----A--G..--T--ATCA...---A-TA-G---C----AAAT---ACA...CCA--AAT---AAGT---ACT---...--C--TCAAGG- 4580
WRC.CI.98.98CI04        -G--A---G-AA-T---A-T-AG-----G--------T-----G--A-----T---AC-G--A-T-----------A-----A--..--T--ATCA...-------A-G--C----AAAT---AC-...CCA--GAT---AAGC---ACT---...------TCAGGGC 4595
WRC.GM.05.Pbt_05GM_X02  GG--A---G-AA-A---A-CA-------G--------T-----G--A--------AC-G--A-----GC-------A-----A--C..--T--ATCT...G---A-G---C----AAAT---ACC...CCA--AAT---GAGT---ACA---...------TCAAGGG 3509
LST.CD.88.524           -G--A--TGCA--T--C--A-G------------T--T-----T--A--G-----CC-GG-GA-----C--------T-T-AAA---..--T---ACAAAAC-G--CA-G---C-G--GAAG-----A...CCA--TGT---A--G---GA--G...-T-----A---- 3557
LST.CD.88.447           -G--A--CC-A--G-----A-G------------T-----------T--A--G------C--G--------CC-CC---------AAA---..--T--ACAAAAAC-------G---C---GAAAG-----G...CCC--A--G--A--GT--AGC---...-T---T-A---- 3560
LST.KE.x.lho7           -G--A--TC-A--T-----A-G-------G--T--T--T--------A-----T---C--G-G----CC-CC----A---AAA--T...--A--ACAGAAAC-G--CA-G---C----AAAG-----A...CCT--AGT---A------AGC---...-TC--T-AC--- 4644
SYK.KE.x.SYK173         --------GCCA-A---G--A-------C-----C--T--------A----------GCT--T------C--------AAA---..--A--A......---A-C-----CC----AA----TCCC...C-A--A--G--A--T-----A-GG..--C---TCA--- 4483
SYK.KE.x.KE51           -G--A------A----CA-AA-C-----T-----C--T--------A--------C--G--------C-------A---AAG---..--T---......----A----G---C-G---AAC--CCCA...--A--A--G--A-------------..--------C---- 4138
COL.CM.x.CGU1           CT--A---GCCA-TC--G-A--C---TGT--------T-----G--A------------TGA-----C--AAG--A......---...G-A--A......C-G---A-G-GGA-----AA------A...GCA--AG---A--C---C-A--G...-T---T------ 4046
```

```
H1B.FR.83.HXB2              TGGAGAGCAATGGCTAGTGATTTTAACCTGCCACCTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAA...GGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAGGAAAAGTTATCCT 4450
Pol                          W  R  A  M  A  S  D  F  N  L  P  P  V  V  A  K  E  I  V  A  S  C  D  K  C  Q  L  K  .  G  E  A  M  H  G  Q  V  D  C  S  P  G  I  W  Q  L  D  C  T  H  L  E  G  K  V  I  L
H1A1.UG.85.U455             ---------------------------T-----------G-----G--G-------------------------A-----------A------...--G------------------------------------G---------T-------C-----C-------------A----- 3896
H1B.US.90.WEAU160           -------------------------------------------------------------------------------------------...------------------------------------------------------------------------G-----A------- 4449
H1C.ET.86.ETH2220           ----------------A---A-----TA-C-----C------C-C-------------TT-------------------------...--G-------A-----------A-T-----------G---------T-------------C--------G---A-C----- 3842
H1D.CD.84.84ZR085           -----------------------------------G-------G-------------------T---------------------------...---------A--------------------------------------------T---C------------------------------- 3972
H1F1.BE.93.VI850            --------G------------------T--------A-----A-------------------T----------G--------------...--G---------C-----------------------G---------T-------C-------------C--G--------- 3790
H1G.SE.93.SE6165            ---------------C-----------TT--------A-----A----------------G---------------------------...--G-----------------------------------------------T-------------------------A-----A- 3847
H1H.CF.90.056               ---------G----------------T--A------A-----A-------------------T---------------------------...--G----------------------------C-----------------T-------C--------G------C-------T-- 3797
H1J.SE.93.SE7887            ---------------------------T--------------------G---------T-------------------------G...--G---------------------------------G---------T-------------C----------------------- 3764
H1K.CM.96.MP535             ---------------------------T--------A-----A-------C--------------T-------C--------------...--G-------A-------------------------------------T------------------------A------- 3646
H101_AE.TH.90.CM240         ------A-------------------TT--------A-----A---------G--------A---A--------------A------...--G----T-------------------------G---------T-------C-----C-------------C----- 4024
H102_AG.NG.x.IBNG           ----AG--------------------T---------A-----A-------------------G-------------------A-G---...--G--------------------------G-------------------T-----------------------G---A-----T- 3975
H1N.CM.95.YBF30             ----A---------C-------------T-A--C--CA----C--CA---G--------------------C-----C-----------...--G-----------------G--CA-T-----------G-G------T-------------C-----G-----A-C----- 4044
H1O.BE.87.ANT70             ----A----C-A--------A---GGA--A-----A--G----A--G--G--C--G-----CA-T--T------CC---------TA-----...--G-----A--T-----T--G-------C-------A-G-------A-------C------A-G-----C---A-C--AA- 4505
H1O.CM.91.MVP5180           ---------T-A--------C---GGAT-A-----AA-------AA-------C--G-----CA-T--T--T---CC------C--TA-----...--G-----A-C------T---------AC--C----AG---------A-G-------------------C---A-C--AA- 4480
CPZ.CD.90.ANT               ------T--T-A--AGA---A-AC--T--T--C---A-T--G--T--------TA----ACAG--------------CG----G...----------G-------------------C-------------------G-------C--C----------T-----A---A- 3889
CPZ.CM.01.SIVcpzCAM13       ----A---TT----GG----A------T-A-----CA----CA--------------G-----G--T-A----------------G------...------G--TC-------G--G--T--T-----C-------C------GA-G------C--------------C---A-C---A- 4030
CPZ.CM.05.SIVcpzEK505       ----AG--------C-----------T--A--C--CA----C--CA------T-----------------T--------G----AT-G---...-------------A-----G---A-CA-------------GG-----------GG--------T--------C--------G--G------A------- 4022
CPZ.CM.05.SIVcpzMB66        -----G--C-----C-----------T--------CA----CA--------------G--------A-AT--------G----AT-------...------G-----A-------G---------------------------T---C--C-----C--G---------GA-C----- 3996
CPZ.CM.98.CAM3              --------T-----A--------C--TA-A---A-A---A-A---G--C------C--------C-T-----------C--A-----------...------G-----------G-----------C--C---------C--C-----T---C--C-----CC----------------------- 3864
CPZ.GA.88.GAB2              --------C------GA---A--CC-A--C-----G-------G--------C-------------------CA-----------------T---...------------A--A-------------G--T--C-------G--C-----G--G--C--C--C--C----------G---A-A----- 3843
CPZ.CM.98.CAM5              ----A---T-----C--------C--T--A---------------------G--C------------------T-----C--G--C---T-----...--G---------------------------------------------C--T------T----C---------C-----------A-C----- 4152
CPZ.CM.05.SIVcpzMT145       ----AG---------CAG--C------T-A-----AA----AA---G--T---------------CAA----------------------------...-----------A--------------T--------C-----------GA-T-------C--C-------------A-A--TA- 3992
CPZ.TZ.00.TAN1              ----A---TT-A---GA---A-A---T-----C--A--T--C--A--T--G--T---------TA-T--TCAG---CCA-----C--TA-----...------G--T--A---------G--G----AC-------A---C-----A---C-----C--CC------------G--C---A- 4094
CPZ.CM.05.SIVcpzLB7         --------T-----C---------C--T--------CA-T-----CA-T-------------G-----T--T--------------------G...------G--A-----------G-----------C-------------GT-------------CC-----------------C---T- 4005
CPZ.US.85.CPZUS             ----C---T-----C---------------------A--G---A--G--C--T--G--G----G------------------C-----------...------G-----A----G--G-----T--C-------T--C-------G----------------------------G---A-C--TA- 4510
CPZ.GA.88.GAB1              ----A---T------------------TT-A-----CA----CA------------------G---CAT---------G--C---G--------...-----------------G--G-------------C--------G--T-------G--------C--CC-------C-----G---A- 4511
MON.CM.99.L1                ----A-TACT-AAGAGA-ACC-ACC--A-T--TA--C-GT-G--C-----------CAA-CAT---C-C--------AAACC-T...------C---AATCA-----G---A--GCAGAAGT----G-------GA-G--C--C--C---C----G----GCA-A--TT- 4153
MON.NG.x.NG1                ----A-TATC-CAGGGA-AC--A-GG-A-T---A-AC-CC----C-----G-----CAA-CAT--CCCC--G-----CAC-C--...------C---AA---------G-----Y-CAGA-AT---G---------A-G-----C--T----AT-------GTA-C----- 2746
GSN.CM.99.CN166             ----AGTACC-AAGAGA-C-A-AC--AA-C--CG-AC--T-G--T-----------GAA--AA---TCC--G------G--C-T...--G---C-A-AA-----G--G---A-TGCAGAATT---C--G-----GA-G--C--C--G-----G--------G--C---T- 4178
GSN.CM.99.CN71              ----AGTATT-AAGGGA-C-A-AC--AA-C--GG-AC--T-G--T------------AA--AG--CTCC----------A--C-T...--G---C-A-AA-----G------A-TGCAGAATT---C---------A-G--C--C------C-G-----C---------T- 4160
MUS_1.CM.01.1085            ----AGTACC-CAGGGA-A-A-ACC-AA-C---G--T-GT----C------------AA-CAT---CCA-----C---AC-C-T...------C-A-AAACA--G-----CA-TGCAGAGAT---GG---------A-G-----C--C--CC-G--G------A-C--T-- 4140
MUS_1.CM.01.CM1239          ----AGTATC-CAGGGA-A-G-AC---A-C---G--T--C-G--T--G---------AA-CTT---CCC---------AC-C-T...--G--GC-A-AAACA----G--CA-TGCAGAATTG---G-G------A-G---------T--C--------------A-C---T- 4130
MUS_2.CM.01.CM1246          ----AGTATC-CAGGGA-A-A-AC---A-A--CA-CA-CT----C--G---------AA--AA--CTCAGC------AAC-C-T...--G---C-A-AA-----G-----CA-TGCAGACTTG--GG---------A-G-----C-------C-G-----C--G--C---T- 4238
MUS_2.CM.01.CM2500          ----A-TATT--AGGGACA-A-A---TA-A--CA-AC-TC-T--C--G---------AA---A--CTCAGC---C--AAC-C-T...------C-A-AA-----T-----CA-TGCAGACAT---GG---------A-G--C--C--C--------------A-C--T-- 4197
RCM.NG.x.NG411              --------------AGAA--A---C-AA-A--T-AAA-------------------------CAA--CCCC--G------G-----...----------AG-A---------G-----TGC---------T-CT------A-G-----C-------C-----------A-C---A- 3944
RCM.GA.x.GAB1               ---------T----AGAA------C-AA-T----AAA----G------------------ACAG---CCA---------AG-----...--G-----A--C-----G-----G--TGCA---------G-CT------A-G--C--C--------G---------A-A---A- 3929
DEN.CD.x.CD1                ----A--AG--CAGAGA-A-G-A-GG---A---A--A-G-----T--G-----CA-TAATGT----CCA-----------AC-C-T...--G---C-A--CAGA-----G--G--TGCCTC--T-----C-------GA-G-----C--C--C-----GAAT-----A--AA- 4340

MAC.US.x.239                GTA-A--A-T---TATTCA-A---GGAT-A--CAGAA----G--C-G-C-G-------A--C----------------T-AG---...------G--T--A-----G--G-C-A-T-CAGA--T---G-CT------A-G--------C---C----G------A-A---A- 4749
Pol                          N  V  K  E  L  V  F  K  F  G  L  P  R  I  V  A  R  Q  I  V  D  T  C  D  K  C  H  Q  .  K  G  E  A  I  H  G  Q  A  N  S  D  L  G  T  W  Q  M  D  C  T  H  L  E  G  K  I  I
H2A.GW.x.ALI                GT--A--A-C-AT-CCA-A-A---GGAT-A--CAA-C-G--G----G-C-G-----AA--CA----CCC----------AG--G...------G--T--A-----G-----GA-TGCAGAA-T---C-CT------A-G--C--C-----C--------------C---A- 4825
H2A.DE.x.BEN                ATA-A--A-C-AA-CCA-A-A---GGAA-A--C-T-C---------G-C-G-----AA-TCA----CCC----C--A-AG---...-------------A-----G------A-TGCAGAAAT---CG-T------A-G--C-AC-----C-------------A-C--TA- 4832
H2A.SN.x.ST                 GTA-A--A-C-AT-CCA-A-A---GGA-----CAAAT----G----G-C-------AA--CA--CACCC----------AG---...--G--G--T--A-----G------A-TGCAGAATT---C-CT------A-G--C--C-----C-------------A-C--TA- 4277
H2B.GH.86.D205              GTA-A--A-C---TCCA-A-A--CGGAA-T----AAT----G------C-G-----AATTC-------------C--A-A----...--G-----T--T--------G---A-TGCAGAC-T---G-C------GA-G--C-------------------------A----AA- 4805
H2B.CI.x.EHO                GTA-A--A-C-A-TCCA-A-A---GGGA-T----AAT----G----G-C-------AATTC-------------C--A-A----...--G-----T--T--------G---A-T-CAGAA-T---G-C-------A-G--C----------------G-----G-----AA- 4803
H2G.CI.92.ABT96             GTA-A--AGT-AA--CA-A-G---GG-A-T--T-AAT-----------C-G-----AATTCA-ACC-C-GT--C--A-AG---...-----------T--------G---A-TGCAGAATTG--T-CT------A-G--C--C--C-------------------A---A- 4159
H2U.FR.96.12034             GTT-A--AGC-ACAGCA-A-A---GGTA-A----AAT-------C---C-------AAT-CT---C--GTT--C----AG---...-----------------------G---A-TACAGAAGT---G-C-------A-G--C-----C--C-----G---------G--A- 4311
MAC.US.x.EMBL_3             GTA-A--A-T---TATTCA-A---GGAT-A--CAGAA----G--C-G-C-G-------A--C----------------T-AG---...---------T--A-----G--G---A-T-CAGA--T---G-CT------A-G--C-----C---C------------A-AG--A- 4217
SMM.US.x.H9                 GTA-A-AA-T---TATTCA-A---RGTT-A--TAGGC-----------C---------A--CA-----------C--C--G---...---------Y--A-----G------A-TGCRGRA-T---G-CT------A-G--C---------CC----G--C---A-A--TA- 4222
SMM.US.x.PGM53              GTA-A--A-T---TATTCA-A---GGTA-A--TAGGC-----------C-G-------A--CA---C-------C--C-AG---...-------------A-----G------A-TGCAGAA-T---G-CT------A-G--C---------CC----G--C---A-A---A- 4666
SMM.US.x.SIVsmH635F_L3      GTA-A--A-T---TATTCA-A---GGTA-A--TAGGC-----------C-G-------A--CA-------G---C--C-AG---...-------------A-----G------A-TGCAGAA-T---G-CT------A-G--C-----G--CC-------C---A-A--TA- 4752
SMM.SL.92.SL92B             GTA-AG-AGT-A---CAGA-A-A----A-C----AAT-------T---C-G-----AATGCT---A-C-----C--A-AG--G...---------G--A-----T--GAC-A-TGCAGAGGT---C-C------GA-G--C--C--C--C-----------G-------A- 4180
STM.US.x.STM                GT--A--AGC-A-TCTTCA-G---GG-A-A---AGGC----------GC-------G-AT-CA-----C--G--C--C-AG---...---------T--A-----G------A-TGCAGAGTT---G-CT-----GA-G--C----------C-------C--GA-A--TA- 4396
MNE.US.x.MNE027             GTA-A--A-T---TATTCA-A---GGAT-A--CAGAC----G--C---C-G-------A--CA----------------T-AG---...---------T--A-----G--G---A-T-CAGA--T---G-CT------A-G--------C---C----G------A-A---A- 4217
DRL.x.x.FAO                 ------T-CT-A--ACAA--A---GGTT-A---GGGA----G---------------G--ACAG---CC---------AA--C-T...------GC-A--A-----C---------TGCA--C-------CT------A-G--C--C------------------GA-A---A- 4017
MND_2.x.x.5440              ---------GT--T--CAG--C---------A--C--CA----G---------G-----T--ACAA---CCA---------AC---G...------C-TG-T-----G--G---A-TGCAGA------G-C------GA-G--C--C--C----------------A-C---A- 4011
MND_2.CM.98.CM16            -----G---T-A--ACAA--C--CTGTA-A---AACA----G-------------------CAG---CCA--G-----AAC----...------GC-A--A-----C--G-----TACATCC-------CC------A-G--C--C------A-G-----------A---A- 4451
MND_2.GA.x.M14              ------T--T-AAG-CAG--A----G-A-A--C---A----C---A----G--------T--ACAG--CCCA--G------G------...------C-T--C-----------G--G--TGCGGCC-------C-------A-G-----C--C---C---------G------A- 4378
OLC.CI.97.97CI12            GCACA-AGTT--CAGCAA--G--CC-AA-A----TAA----G------AGG------ATGAA----AG-----A-AGGG---...--G-TTC-T-G-------AT---G-CAGTGG-GGTTTAT-GG-----GA----C--C--T-----------G--G--G---TG 4321
GRV.ET.x.GRI_677            ------AGT----TACAG--A---GGAT-A--TAA-A-------------G--------GGCA---CCC-----C--AA---G-...------C-T-A---------G------GCCTCCATT-A--CT-----GA-G--C--C--C------------------------AA- 4623
VER.KE.x.TYO1_patent        ----A-AACC-A--AGA-ACA-A-GGG--T----AAA-------------G-----G---TG---CCA---------A----G...------C-AG-------------G--TGCCTCA--T----C------GA-G-------T---C------------AG--A- 4674
VER.KE.x.9063               ----A-AATT-A--GGACACA---GGAT-A----AGA----T--C-----G-------A-TG---CCA--TGCCA-AA------...--G---C-A--A-----G---------GCCTCC-----GG--------A-G--C--C--G---G-----------A-AG--A- 4677
VER.KE.x.AGM155             ------AATC-A--AGACACA---GGAT------AAA-------T--------T-----A-TG--CCCA---------AG------...--G---C-A--A-----------------TGC-TCA------G-G-----GA-G--C--C------A-----------A-AG-GA- 4669
VER.DE.x.AGM3               -----GAACT-A--AGACAC--A-GGG--A----AAA-T--G------------------TG---CCA---------A-------...--G---C-AG-C-----G--------TGCCTCG-----GG--------A-G--C------------------C--G--A---A- 4172
TAN.UG.x.TAN1               ---CA-TTT--AAGAGA--CA--CGGTA-A---G-AC-G-------------------TGCT--CCCA---------AA-TCG-...------C-G--A-----G--------TGCC---GT---GG-C------A-G--C------------G-------A-C--TA- 4649
SAB.SN.x.SAB1C              -----GAGT---CAACAA--A---GG-T-A--TG--A--------------------GGCA---CC------C---A-------...--G--GT-TG-A-----G--------TGC----------G-G------A-G-----------CC-------------A-C---A- 4848
MND_1.GA.x.MNDGB1           GAAGC-CA-T-AAGAGAGA-A--CC--T-A---G--C------C---C-G--T--GCAA--T--CAG---G--CTGT-ATC-T...------GC---AA-G-----GAC---TGC-TCA-TT---G-C-----GA-------C------C-G---AATC--A----TA- 4137
DEB.CM.99.CM40              ----A-CA-T-AAGGGA---A-ACC-A--A--TA-AC-CA----T--GC----TA--A-ACAA--CCCC----------AC--G-...--G--GC---CA-----T--G--G--TGCCTCAATG--C--C-----GA-G--C--C--C---C-G---AAC---A-A--AA- 4256
DEB.CM.99.CM5               ----A-CA-T--AGAGA---A---C-GT-A--TA-CT--A----C--GC-G--TA-T--GCAG---CCA--G--C---A----G...------GC-T---------T---------TGCCTCCATG--GG--------A--------C--C--CC-----AAC-----C---A- 4250
TAL.CM.01.8023              ----AG-ACC-AAAGGCCAGG-----A--A--CA-CA-------C----CC--CA---AGGCA--CCCC---------G-CC--...------C-A-A-ACA-----GACCA--GCAGCAGT---G-C------GA-G-----C--C--CC-G------C-G--A--ATG 3914
TAL.CM.00.266               -----G-ATC-TAAGGCAAGA-----GT-A--CA-CA-------C--G-CT--TA---AAGCA--CCCC---------AG--C--...------C-T-A-ACA-----GAACA-TGCAGC-GT---G-C-------A-G-----C--C-----G--G---C--A-C---TG 4416
SUN.GA.98.L14               GCAGCCT-TT--CAGGAGATG---G--A-A--C-TA--G--------GC-G-----CAATGAA----CAC-------A-AG---...--G--------AACA-----------TGCCTCAGTG--C--C-----GA-------C--C--CA-G----A------C--AA- 4810
WRC.CI.97.97CI14            ----AG--CT-ACAAGAA--A---C-T--T--T-AAT-GA--------AG----A---AGGAA---AG--CC--C--AGGG---GGA--G-C-C-A--------TATG-----T-A-GAGGTT--G-C-------A-G--C-----T-----------G---A----TT- 4750
WRC.CI.98.98CI04            ----A----T-ACAAGAA--A---C-TA-T----AAT--A-T------AG--------AAGAA---AG--CC----AGG---GGA---CTC-AC-A-----TATG-------A-GAAGTG----C-------A-G--------T---------------A-C--T-- 4765
WRC.GM.05.Pbt_05GM_X02      ---G-G---T-A-AGGAA--A---C-T-----CA-CA-TC-T------AG--------AAGAA----C--CC------GG---GGG---C-CC-T--T-----CATG-------A-GAGAT---G-C------GA-G--C--C--C--C--G---------A-C--AT- 3679
LST.CD.88.524               GCA-A--ACT-A-AAGAAA-G-A-GGGA-A--C--CA-G-----T--GC-------AATGAG----C---G-------AG--G...---------------AACA--T--------TGTATCAGTG--T--T--------G--------T--CC-------GC--A-C---A- 3724
LST.CD.88.447               GCT-A--ATT-A-AAGAAA-A---C--A-A--C---A-G-----T--GC------GAAT--T---A-----------AC---G...---------------AACA-----------TAC-TCA-----T--T------A----C--C--T--CA-G--G---C-G--A---A- 3727
LST.KE.x.lho7               GCA-AG-ATC-A-AAGAAA-A--------A--C--CA-G-----C---C----T--CAATGA-----CA--C-----AAAG---...---------------AACA--------G--TGTCTCAGTG--T--T-----G-----C-----T--C---------C-G--C--TA- 4811
SYK.KE.x.SYK173             AT-GAGTATC-TAGGCAG--G---C-T--A--C-GACAG----------C---TA--CAACAA---CCA---------AAAT-G-...------C-A-AA-----G--------TGTAGA-AT-TAT-AT-----GA-G--C-----T---GAG-----C-----A---TG 4650
SYK.KE.x.KE51               GTAGA-TAC--AAGGCAA--A---C-T--A--CAGACAG-----C----CT-------AA-TG---CCC-----C--AA-C-G-...------C-A-AA-----G---------ACAGA--T---T-C-------A-G-------C--C---------------A--TTG 4305
COL.CM.x.CGU1               GT-CAGTATC-AAAAGAAC-A---GGGT-A--CA-------------------TTGG-AA--A---TCAG-G------AAC---...----C----TG-T-----C--GC----T-A---CTAT--CT--------T---------T---GAG--------G------T- 4213
```

```
H1B.FR.83.HXB2            GGTAGCAGTTCATGTAGCCAGTGGATATATAGAAGCAGAAGTTATTCCAGCAGAAACAGGGCAGGAAACAGCATATTTTCTTTTAAAATTAGCAGGAAGATGGCCAGTAAAAACAATA...CATACTGACAATGGCAGCAATTTCACCGGTGCTACGGTTAGGGCCGCCT 4617
Pol                        V  A  V  H  V  A  S  G  Y  I  E  A  E  V  I  P  A  E  T  G  Q  E  T  A  Y  F  L  L  K  L  A  G  R  W  P  V  K  T  I  .  H  T  D  N  G  S  N  F  T  G  A  T  V  R  A  A
H1A1.UG.85.U455           A--------C-------------C--C-----------------C-------------A-----G---------C---A-AC-----------------------------GT----...--C--A-------------------A-C---G-A---AA--A-T-- 4063
H1B.US.90.WEAU160         ------------------------------------------------------G--------A-----------C---A-C------C---------------------------------...-----A-------------------A--A---------A-------- 4616
H1C.ET.86.ETH2220         ---------C-------------C--C-----G-----G-----------------A--A-----------C----AC------------G----------C-GGGT----...-----A--T--------T--C------A--AA-G-A---AA--A---- 4009
H1D.CD.84.84ZR085         -----------------------C-------------------------------------G--------C----C-----------------------------GT-G-G...-----A-------------------A-------A---A-------- 4139
H1F1.BE.93.VI850          ---------C--------T----AG--C--------------------------A----------C--C--CA-A----G----------------------------T----...-----A-------------------A----CG------A---AT--- 3957
H1G.SE.93.SE6165          A----------------------C------------------------------A----------------C---A-A-----------------G---------C-GT----...-----A--T-----------------A-----G-A--A-A---A--A- 4014
H1H.CF.90.056             ---------C-------------C-----------------C--C-------------AA----------------C--C--G--G---C-----A-C---------------GT----...-----A------------------GA-----G------A---A---- 3964
H1J.SE.93.SE7887          -----------------------C-----------------C-------------A-----G------T----A-A-------------C-----------------GT----...-----A--------------C------A---G-G-T--G-A---A---- 3931
H1K.CM.96.MP535           ---------C--------T--C--C-----------------C----------G--A--A-------C--C----A-AC-----------------------------GT----...-----A---------C----------A-CA--GT-----A---A---- 3813
H101_AE.TH.90.CM240       ---------C--C--G----------------------------C-------------A----G---------C----GC-----C-----------------------GT----...--C--A-----C--T------------A-C---G-A---AA--A---- 4191
H102_AG.NG.x.IBNG         A--------C-------------C-----------------C-------------A----G---------C---A-A-----------------------------G--GT----...--C--A-------------------A-----G-A--A-A---G--A- 4142
H1N.CM.95.YBF30           T--G--G--C-----G--------C--CT-------------------T----G----A------------------A------G-----T----------------GTT---...--C-----T-----ATC---------TA----C--T--A-AA--A---- 4211
H1O.BE.87.ANT70           A--T--T--C-----G--A-----G-TC-------------G--A-------------A--A-----T--C--C--C--G------C-G--T-C---------T--T--GT----...-----A-----C--GCCT-----T--AA--A-A--TA-G-A---T--A- 4672
H1O.CM.91.MVP5180         A--T--T--C--------A----AC-T-----------G--G--A-------------A-----------T--C-----C--G------------C---------T--C--GT----...-----A--------ACCT-----T--AA----AG-CA-G-AA--T--A- 4647
CPZ.CD.90.ANT             A--------G------T-T-----C-TC--------T-----A--GG---AT--G------A-AAGT---------C--C--G----------CA-C--------------------...--C-----------AGCT--------AA----AG-A--A-AA--G--A- 4056
CPZ.CM.01.SIVcpzCAM13     A-----------A-----------C-----------------C-------------A--A-----T-----C--CA-A------------A-C-----------G---GTCT--...--C--A--T-----G-CA-----T--AA--AA-G-A----AA--A--A- 4197
CPZ.CM.05.SIVcpzEK505     T--------C-----G--------C--CT-------------C-----T------------A---------T--------A-CC---G----T-----------------------GTC---...---------T-----G-----C-------A----C--T--A-AA--A---- 4189
CPZ.CM.05.SIVcpzMB66      A--G-----A-------------G-----G--------------A--A--T--T---------A--A------------C---A-C---------G----C-------------------GTC---...---------T-----A--T--C--T---A--AG---T--C-A----A--T- 4163
CPZ.CM.98.CAM3            ---G-----A-------------G--CC----G-----G--C-----T----G--T--A-----------T--------A--C-----------------------T------GT---C...--C-----T-----GCCT--C--T---A--AA---A-----A---A---- 4031
CPZ.GA.88.GAB2            ---G-----C-----G--A--------CT------------------A--------T--G--T--A-----------T---------A-C--------------G--------------G-AG---...--C--A--T-----A-CA----------A-----AG-A--A-AA--A---- 4010
CPZ.CM.98.CAM5            A--G-----A-----G-----------CC----G-----G--C-----T-----------A--A--------T-----A---------------------------T------GT----...---------T-----GCCT--C--T--TA--AAC--A--C-AA--A---- 4319
CPZ.CM.05.SIVcpzMT145     A--G-----C-----G-----------C---------------C--C--T--T---------A------------------G-CC-T--G---------------------G--T---...--C--A--------GCCT--------TA--AA-G-A--A-A---A--A- 4159
CPZ.TZ.00.TAN1            A--------G--------T-----T-TC--------------C--A----A---------AAGA----C--T--C--CA-CC-------G---------------T-----G-A----...-----A--T-----ACCA-----T--TA--A-AG-A--G-A---A---- 4261
CPZ.CM.05.SIVcpzLB7       A--------G--------T-----C-----G----------C--A--T-----------A----------T-----A-C-----------TCG----------------TT----...---------T-----T--------T--TA-CA--G-A--C-A---T--A- 4172
CPZ.US.85.CPZUS           A--G-----A-----G-----------CC-------------G--C-----T-----------A--A--G----C-----A-C-----------------------T-----GTG--T...---------T-----ATCT--C--T--AA--AG---A-----A--A--T- 4677
CPZ.GA.88.GAB1            A--G--------C-----------C-----------------C-----T--G-----A--A-------T-----C--G--------------T------------------T--T...--C--A--T-----GCCA-----T--AA------G-A--C-A---T---- 4678
MON.CM.99.L1              ---------A-------------T---G--TGG--CA-GA-CC-A--AGG-------AA-ATGT---GGATAGCC--G--GG-GC-G--C-CT-TG-----G--C-CTCAG---...-----A--T-----GCCA----------A-CCAAGAAT-CGAA--A--TG 4320
MON.NG.x.NG1              A-----C-----------T-----C--CG-TTGG--CA--A--C----CAG-----------A-A-------GCAT-GCC--G--GG--C-T----CG-C-------A--C-TCATG---...--C-----------ACCA----------A--CAGGAAT--GCA--Y--TG 2913
GSN.CM.99.CN166           A--G--G--C--C-----A-----G---G-CTGG--TAGGA-A--A--CCAG--------CAGAC----T----TGAAG--C---G----------C--C------T----CCCACC--...-----A--T------CC-------GT-TCCAAGGAAT-GGAA--A---- 4345
GSN.CM.99.CN71            A--G-----C--C-----------G---G-TTGG--TAGGA-A--A--CCA---------CAGAC----T----TAAAG--A---G----------C--C-------------CTCACC--...-----A--T------CC--------T-TC-AAGGAAC-GGA---A--T- 4327
MUS_1.CM.01.1085          A--------C-----G--T-----G--C--GTGG---A-GA-AT-G---AG----------AG-C-G-----T-TAGGC--AC-GG-G------TC-CT------T-CG-C-CA---C...-----A--T-----GGCT--C-----T-A----AGAAT-CGAA--A--A- 4307
MUS_1.CM.01.CM1239        A-----C--C--C--G------------C-CGTGG---A--A-AT-A--CAGG-------AAG-C-------CCTAGGC--G---G---------CTCTT-----T-----C-CA------...--C--A--T-----AGC----C----TTA---AAGAAT-CG-A--T--A- 4297
MUS_2.CM.01.CM1246        ---G-----A-------------A-T-TC---TGG--CAGGA-A---T-CCA------T--AAG-CT-------CCTAGAG--G-C---C--G----CC-C--------A-T-GTCAG---...--CC-A---------T-C---C--T--TA-CAAGGAAT-CCAA--T-TTG 4405
MUS_2.CM.01.CM2500        A--------A---------------T-T----TGG--TAG-A-----T-CCA-----AT--CAG-CT--------TAGAG--G-----C-------C--C-------CA---C-CAG--C...--C--A---------T-C---C-------A--AAGGAAT-CCAA--A-TTG 4364
RCM.NG.x.NG411            A--------C-----------------C---------------A--A--------G-----AA-A--G-----TC----C--G--------------CC-----------G---CATT--...--C--A--------A-CA--C--T--AA----AG-A--GCAA--A-A-- 4111
RCM.GA.x.GAB1             A--------C-----G--A-----------------G-----------A--A-------------AA-A--G------C----C--G--G---C------C---------------GG-AGC--...--C--A--T-----AGCA--------AA----AG-A--ACA---G-T-- 4096
DEN.CD.x.CD1              A--------AA---C---------G--CC----GA--A--A--T-G-----G--G-----CA-A----------CTA---T-AC-GC-GA----T-CT---------A-C-CC-A-T--...--C--A--------ACCA-----T-T-A--CAAG-CA-GCA---A--A- 4507

MAC.US.x.239              A--T-----A--------T-------TC-----------G--A------CA---G----AAGAC-G------CTA-----G--------G-----C--------TA-T-C-CATC--...--C--A--T-----TGCT--C--TG-TTCGCAAGAA--A-A-ATG-TTG 4916
Pol                        I  V  A  V  H  V  A  S  G  F  I  E  A  E  V  I  P  Q  E  T  G  R  Q  T  A  L  F  L  L  K  L  A  G  R  W  P  I  T  H  .  L  H  T  D  N  G  A  N  F  A  S  Q  E  V  K  M
H2A.GW.x.ALI              AA-------G-----T----------TC-----------G--C--C--GCAG---T----AAGAC--------CTC--C--A------C-G--TA-T---------A---C-CACT-G...--C--A--T-G---TGT---C------TCACAGGAA--A-A-ATG-TAG 4992
H2A.DE.x.BEN              A--------A-----T--A-------TC--------------C--C---CAG---T----AAG-C-G------CTC--C--A------C-G--CA-T--G------A---CGCACT-G...--C--A----------CC---C-----TTCACAGGAA--G-A-ATG-TGG 4999
H2A.SN.x.ST               A--------A-----T--A-------T---------------C--C---CAG---T----AAG-C----G---CTC--C--AC-----C-G--CA-T--G------A---C-CATT-G...--C--A---------TGC---C-----TTCACAGGAA--A-A-ATG-TGG 4444
H2B.GH.86.D205            A--G-----C-------------G-T------------G--A--A--CCA---G-----AAGAC-G-----TCTC--C--AC----G--G--CA-C---------TA-C-C-CACC--...--C--A-----C--TGC---C------TCAC-A-GT--A-A-ATG-TAG 4972
H2B.CI.x.EHO              A--G----------------------TC--------------A--A--CCA---------AAGAC-G-----TCTC--C--G-----GC-G--CA-C---------TA-C-C-CACC-G...--C--A-----C--TGC---C-----TTCACAAGAT--G-AAATG--AG 4970
H2G.CI.92.ABT96           A-------------G--------C-T-G--------------A--C---CA---------AAG-C---------TA--C--G----------------G---------TA-C-C-CACC--...-----A---------TGC---C------TCGCAAGA---G-A-ATG-TAG 4326
H2U.FR.96.12034           C--G-----A-----G--------C-T---------------G--C--GCAG-----T--TAG-C-------CCTG--C--AC----GC----------G-----TA-C-CTCATT-G...--C--A--T-----TGC---C-----TTCACAGGAA--A-A-ATG-TAG 4478
MAC.US.x.EMBL_3           A--T-----A--------T-------TC--------------A------CA---G----AAGAC-GCACTAT...--C--G---------G-----C--------TTATTT-CAC--CTACACACACAGT-----TGCT--C--TG-ATCGCAAGAA--A-A-ATG-TTA 4384
SMM.US.x.H9               A--T-----A-----G-GT-------TC-----G--------A--C--GCRG--------AAGAC--------CTG-----GC----------CA-C--------CA-C-C-CATC-G...---------T-----TGC----------ATCACAAGAA--G-AAATG-TTG 4389
SMM.US.x.PGM53            A--------A-----G--T-------TC-----G--------A-----GCAG--------AAG-C--------CTG-----GC----------TA-C--------CA-C-C-CATC-G...---------T-----TGC----------ATCACAAGAA--G-AAATG-TTG 4833
SMM.US.x.SIVsmH635F_L3    A--T-----A-----G--T-------T------G--------A--C--GCAG--------AAGAC--------CTG-----G-----GC----T--C--------CA-C-C-CATC-G...---------T-----TGC----------ATCACAAGAA--G-AAATG-TTG 4919
SMM.SL.92.SL92B           A--------A-----G--T-----C-T-------------G--A--A--TAGG--------AAGAC-G-------TG--C--C------A-----TC---------TA---GCATC-G...-----A-----C--TGC-----------TCCCAGGA---G-A-ATG-TAG 4347
STM.US.x.STM              A--------A--------T-----G-T------G--------A--C---CA-----------AGAC-G-------TG-----G--G--G------A-C--------T--C-C-CACC-G...--C--G--T------GCT--C-------TCGCAGGA---A-A-ATG-TAG 4563
MNE.US.x.MNE027           A--T-----A--------T-------TC--------------A------CA---G----AAGAC-G------CTA-----G--------G--A-C--------TA-T-CGCATC--...--C--A--T-----TGC---C--T--TTCGCAAGAA--A-A-ATG-TTG 4384
DRL.x.x.FAO               A--G-----C--C-----A------C-C--------------A--A-------------AG--A-------T--------GC----------T--C-------------TC-CACT-G...--C--A--T-----ACCA----------A---AA-A---AGCAA-A-TA- 4184
MND_2.x.x.5440            A--------C------------------T----------G--A--A--------G-----AA-A--------GC-C--C--G-----GC-------T--G-----------CATC--...--C-----------TCCT--C--TGT-A---AA-A---GGCCA-A-T-- 4178
MND_2.CM.98.CM16          A-C------C--------A---A-G---C-------------A--A--CA-----------AA-A---------C-C--C--G---------------T---------------CATC--...--C--A--T------CCT----------A---AA-AA--AGCCA-A-T-- 4618
MND_2.GA.x.M14            A--------C-----G------------C-------------G--A--------------AA-A---------C----C--GC----C--------------------G---CATT--...----------------TCCA--C--TGT-A---AG-A---AGCCA-A-TG- 4545
OLC.CI.97.97CI12          TT-G----CA---AC--GGTCA--G-TCC--T-T-G----A-CC--AG-------GAT--AG-AA-T----TGA-A---G-GC-T-G-C----TA-TCA------T-------T---T...--C-----------G-CA-----TGTGAA-CAG-A-A-GCA---A-TA- 4488
GRV.ET.x.GRI_677          A--------A-------------------TC-----------G--G--C---AG------T---A---G------C-C-----GC-G---C-GTT--C------------G---CATC--...--C-----T------CCA--C--T---TC-CAG-AT--GGCA--G-TG- 4790
VER.KE.x.TYO1_patent      A--T--G--C----------------TC--------------C--A--TAGG--------AA-A----G--A-G-----A-----A--CTGA-T--------TA---C-CAGT--...--C--A--------GCCT--C--T---TCCCAAGAA--GGCA--AATA- 4841
VER.KE.x.9063             A-------------G-----C---TC--------------C--A--TAGG--------AAGA--G-----CA-A--C---C-----A--TTGA-T---------A--GTCCAGT--...--C--A--T-----ACCA-----T--TTCCCAAGA---AGC---AAT-- 4844
VER.KE.x.AGM155           A-----G--C-------------G-T-----------G-----C--TAGG--------AA-A--G------A-G--CT-G------A--AT-------------CA-C-CTCACC-C...-----A--T-----ACCA---------TTC-CAGGAA--AGCT--TATG- 4836
VER.DE.x.AGM3             A--------C----------------TC-----------------A--TAG-----------A-A----------A-A---T-A-----GA--CT-A-T--------CA---CCCA-C-G...-----A---------ACC------T--GTC-CAAGAA--AGCA--AATG- 4339
TAN.UG.x.TAN1             A--------C-----G--------C-TCT-------------C--AG-TAGG--------AA-A----------A-A--C--G------A-CATTA-T--G-----TA-T-C--A-T--...----------------GCCA-----TGTTA--CAGGA---ACAAA-TATT- 4816
SAB.SN.x.SAB1C            A--G--T--------T--T-------T---------------A--C---CAG--------AA-A-CT------C-C--C--GC-G--G-----CA-C--------CA-C-C-CAGC-G...-----A--T-----T-C---C--T--TA-CCAGCAA--AGCA--TATT- 5015
MND_1.GA.x.MNDGB1         A--------G--------TTCA--C-TC--GA-G-----------AA----------T--AA-AA-G-----G-G----G-----G------C-CA------TA-T-GT-A-C--...--C--A--T-----GCCT--C--T--TA--CAGGAA--AGAAA--ATG- 4304
DEB.CM.99.CM40            A--G-----C-----G--T---------------GA-CA-GA-CC-A--------------AA-A--G----TCTG--C--G---C-G---G--C--------------CTCAG---...--C--A--------ACC---C--T-T-A--CAAC-A--C-A---A--T- 4423
DEB.CM.99.CM5             A--------G--------T-------T---G--GA-TA-G--CT-AT----------T--AA----------C-TA--CT-A---C-G--G-GG-C---------CA-T---CAG---...--C--A--------ACCA-------TAA--CAGC-A--C-A---A--A- 4417
TAL.CM.01.8023            T--------A-----G--A-----T--C--T--GA--A-GA-CT-G--CAG------C--A-G------C--CCTG-----C--GC-GG-G--CA-C----------A--G--CATC--...--C--A--------GCCG--C--TGT-TCG--AGA-A-GCA----A-AG 4081
TAL.CM.00.266             T--------A--C--G--------G--C-----GA-TA-GA--C-G---AGG--G--G---G---G--C--CCTC-----C---C--G-G--CA-C--------TA---GCCACC-G...-----A--------GCCA--C--TGT-TCA--AGAAA-GCA----ATGG 4583
SUN.GA.98.L14             A-----G--G-----G--TTCA--G-----G-T---------AC-A--CAAT--GCA----A--AC------CAC--GG--G-----GC-CTGT-CC-TG-----------CAG---...-----A--T-----GCCA-------T-A--AAAGAT--GGA---A-TA- 4977
WRC.CI.97.97CI14          AAAT-----C---------TCA-----CC--------A--A-C--AT--------GAT-CCT-TC-G------CTA---ACC--GC-GC-G---A-T--------CA-C--G-A----...----------------G-CA-------AGTC-CAG-AAA-GCA---TAT-- 4917
WRC.CI.98.98CI04          AAAT-----C--------ATCA--G---C--------A--A-C-------------GAT-CTT-TC-------TCTA--C--CC--C--A-C-----T--------TA-G---AG--T...-----G--T-----A-CA-----T-AGTCACAG-AAA-GCA-A-TAT-- 4932
WRC.GM.05.Pbt_05GM_X02    AAAT-----C-----G--ATCA--C--CT-------GA--A-A-----A----GAT-CCT-TC-------CT---CACCA--C-------------------------AG--T...---------T-----A-CA-------AATCACAG-AAA-GGA----ATA- 3846
LST.CD.88.524             -AAT-----C---A-G--ATCA--C-TC--G-TT--------G-------AT---------A-AACT------A-----T-G-----G---TGTA-C---------T------CA---T...--C--A--T-----TCCA-----TGTGA--AAAGAT--ACA---A-TGA 3891
LST.CD.88.447             TAAT------------G---TC-----T---G-T---------A-------AT---------A-AAC-------C-C--C--G------C-GTGCA-C-------------C-GCA-G-C...-----A--------ACCA--C--TGT-A-CAAGGA---ACAA----TAA 3894
LST.KE.x.lho7             AAAT-----C-----G---TCA--C-TC--G-TT-----------A----AT--G-----AA-AAC---GT--C----C--G---------TGCA-T---------T------CA-------...-----A--------TCCA-----TGTAA--AAGGA---ACA---A-TAA 4978
SYK.KE.x.SYK173           T--------AA-CAC---T-----------------A-CA-GA-CT-AAA-AGG-----G---G-T-----------TG-----CA-GC--A----CA-T----------A-----CAG---...--C-----T-----ACCA--C--TGT-A---AC-A-T-C-AA--A--A- 4817
SYK.KE.x.KE51             T--------AA-CAC----------G-----C-----A-TA--A-AT-GAA-AG-------------A-------G-CTA--C-----GC--A------C-----------A-T---CAT---...-----A--------TCCA--C----T-A---A-G-CT--GC---A---- 4472
COL.CM.x.CGU1             A--------A--C--CTGT-C-CTG-TCTGCTGG---ACCA-AT-AAAGAG----------G-A-------GTAGGGCC--CA-------G--TA-TCAG---GAG--T-G-CA-G-T...--C--A--T-----GCCA-----TGTAA--CAGCACT---A---A--AG 4380
```

```
H1B.FR.83.HXB2              GTTGGTGGGCGGGAATCAAGCAGGAATTTGGAATTCCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAAAGAATTAAAGAAAATTATAGGACAGGTAAGAGATCAGGCTGAACATCTTAAGACAGCAGTACAAATGGCAGTATTCATCCACAATTTTAAA 4787
Pol                         C  W  W  A  G  I  K  Q  E  F  G  I  P  Y  N  P  Q  S  Q  G  V  V  E  S  M  N  K  E  L  K  K  I  I  G  Q  V  R  D  Q  A  E  H  L  K  T  A  V  Q  M  A  V  F  I  H  N  F  K
H1A1.UG.85.U455             ----------AAAT---C-A----------G--C----------------------------G-----C-----C--G-------------C-----G-----------G--A--------C----------------------------------T----------- 4233
H1B.US.90.WEAU160           -------------G----------------C----------------------------GA-------------------------------------------------A---------------------------------------------------- 4786
H1C.ET.86.ETH2220           ----------A--T--TC-A--------------------------------------------------------------C-----G----------A--A----G--C----------------------------------T----------- 4179
H1D.CD.84.84ZR085           -C--------A--T------------------------------------------------------------------------------------------------A----------------------------------------------T----------G 4309
H1F1.BE.93.VI850            ----------A--T---C-A-------------C------------------------------A-------G-----A--G--C----------------A---------------------------------------------------- 4127
H1G.SE.93.SE6165            ----------AAAT----CA-----------------------C------------------------G-------------C--C--G-----C--G-----A-------------------------G-----------T----------- 4184
H1H.CF.90.056               ----------A-AT---C-A----------G---------------------G-------------------------G--C-----G----------C--A--A-----C-----------------------------------T----------- 4134
H1J.SE.93.SE7887            ----------A-AT------------------------------------------------------------------C----------------A--A--------C-----------------------------------A----------- 4101
H1K.CM.96.MP535             ----------A--TG---------------G-------------------C---------------------------------C--------------G---------------C--A--------------------------------------- 3983
H101_AE.TH.90.CM240         ----------CAATG--C-A----------G--C----------------------------------G------------C-----G-----------G--A--------C----A--------------------------------T----------- 4361
H102_AG.NG.x.IBNG           ----------AAATG---CA--A-------------------------C---------G------------------------------------------C--------A--------C-----------------G--------------T----------- 4312
H1N.CM.95.YBF30             ----------AAAT-----A----------G--A----------T--------G----C-----G--C-----------------------------AA-C--------A--A---------A----------G---------G--T-----T----------- 4381
H1O.BE.87.ANT70             ----------CAAC--AC-A--T--G---------A--A--T-----A------------------G-C--------G--------ATC-------CAG-----G--G--C--A--A-----CT-A-GA----------------------TG-T----------- 4842
H1O.CM.91.MVP5180           --------A-A--C--AC-A--T--G-----G--A--A--T-----A------------------G-C-----------------ATCT------CAG-----G--G--C--A--A--G---T-A--A--------------------C--TG-T----------- 4817
CPZ.CD.90.ANT               ----------TAAT---C-A--------------A--A---------A-----------------G-----C---------C----G---C--------------AA-T-----C--A--A-----AT-A-----------GT----------GCA---T----------- 4226
CPZ.CM.01.SIVcpzCAM13       ----------CAAC-----A--AA-G------------------C--A------------G-------C-----C-------------------C--T-A---A--T----------A---------A--A---------------------G----GT----------G 4367
CPZ.CM.05.SIVcpzEK505       ----------AAAT---C-A-----G--------A------------T--------G--T-C--------C------------------------------AA-T--------A--A------T-A--A---------------------C-----T----------G 4359
CPZ.CM.05.SIVcpzMB66        -------------T---C-A--A--------------------------T----------G------------------------------------A-----T---A---------A--A------------A--------C---------------T----------- 4333
CPZ.CM.98.CAM3              -C--------A--C---C-A------------------T-----A-------------G-------A-------G-------A--G--C-------AA-T-----A--A--A-----CT-G-G--------C--G-------G-----T----------- 4201
CPZ.GA.88.GAB2              ----------A--C--A--A--A---------A--T--T-----A-----------G---------C-----C---------A--G--------G--A--T-----C--A--A-----CT-A-------------------T--------T----------- 4180
CPZ.CM.98.CAM5              ----------T--C---C-A-----G--------C-----------A--------------G-----C----------------------C---------A-T-----C--A--A--G--G--A--------T--C-----------C-----T----------- 4489
CPZ.CM.05.SIVcpzMT145       ----------A--T--T-GA-------------------A-----C--A-------------G-----------------------A--G--C----AG--A--C-------------------T-A--A-------------------C-----T----------- 4329
CPZ.TZ.00.TAN1              -C--------ACA---TC-A--T--------G-----A--T-----T--------------------------C--------C---------AG--A--C--G--C--A--A--G--A--G-G----------ATC--------G-AT------------- 4431
CPZ.CM.05.SIVcpzLB7         -----------CAG---C-------------------------------------G----------G-------A-----C------C-T-------A-----G---------------------A--G--C--A--A----------------------T----------- 4342
CPZ.US.85.CPZUS             -C--------A--C--AC-A--A--G-----------A---------A-----------G---------C------------------------------C-------AA-C--G-----A--A--G---T-A-------T-----G----------------T----------- 4847
CPZ.GA.88.GAB1              ----------A-AC------------------A------T----A--G----------G--------CT-A--------GC-----------A---------T--G-----A--A-----T-A--A-----------------------G-----T----------- 4848
MON.CM.99.L1                CA--------AAAT--TC-T--CACCACA--GG-A---------C--A-----C-----G--T----GAA--------G-C--C----AG-G-CC--TAAGA-AA-C--G---G---T----T-C--CCCA------T--GC-CAA---C-G------TTA---------- 4490
MON.NG.x.NG1                CA--------AAAC--AG-A--CAGTAC----G-G--------C--A-----------T--------AAC--------GC--A-C--A-CC-CAT----GA-AA-C--G---G-A-T-C-GTTC--CTCT------C-GGC-CAAA-GC-C-A---T-TA---------- 3083
GSN.CM.99.CN166             -C--------A-AC--AC-A--CACCACA--GG-G--A-----C--G-----------G--------AAC-----C---C--C-C---G---CA--CACGA-AA-C-----AG---TCAC-T-CT-AG-A---------GC-CA----TGC-A---T--T--------- 4515
GSN.CM.99.CN71              ----------AAAC--AC-A--CAGTACA--GG-G--A-----C--G-----------G--------AAC-----C---C--C-C---G---CA--CCA-A--A-T------G---TCAC-T-CT-AG-A---------GC-CAA---TGC-A---T--T--------- 4497
MUS_1.CM.01.1085            -C--------CA-C--AG-A--TACCACA--GG-A--A--T-C--T--G------------G-----AAC-AA--C---C-G-----AG--GCC---CACA-AA-C--G--AG---TCACTT--T-GG-A---------GC-CAA-----C-----TTTA---------- 4477
MUS_1.CM.01.CM1239          ----------TA-C--AG-T--CACCACA--GG-A--A-----C--T-------G------------AAC-AA--C---C-G-----AG---C----CAGA-AA-C-----AG-A-TCACCT--T-AG-A---------GC-CA------C------ATG---------- 4467
MUS_2.CM.01.CM1246          CA--------TAAC--AC----CACCACA--TG-C-----------A--G----------------AA-GCA--C---C-G------G---CC--TCACA-A-----G--AG-A-TAAC-T--T-AG-A---------GC-CAA-----C-----T-TA---AA---- 4575
MUS_2.CM.01.CM2500          CA--------A--C--AG-A--CAGTACA--TG----T--T--C--A--G--C-----G--------AA-GCT-----GC--------G-G-CC--TAACA-A--C-----AG-A-TCC-GT-C--AG-----------GC-CAA-----C-----TTTA---AA---- 4534
RCM.NG.x.NG411              ----------TCA---AG----CACC------G-A-----------T-----C--------G--------A------C-TC----G----T--A-----G--AA-T-----------A---A-GT-AG---------T-----------GC----T----------- 4281
RCM.GA.x.GAB1               -C--------TCAG--AG----C-CC------G-A--T--------T-----------------G---AGC----------C-------A-T---C----A---A--------A--A--A---A-AT-AG-A--------C-----------T--GG-T----------- 4266
DEN.CD.x.CD1                ----------AAAT-----A--CACCACA---G----A--T-----A--------G--TA-T-----AA--A---C--GC--------G--G------C--TA-----G--AG-T--A---AG-T--A---AGAT-AG-A-----C--GC-------ACCC----T-TT--------- 4677

MAC.US.x.239                CA--------A--G--AG----CACC-----GG-A--A---------A--G-----G--------G---G-A-----C-CC-CC-G--A--TCAA---ATAGAA-C--G--A--A--AA-TTCAG-AG-A--CAT----TT---------TCATTG-ATG--------- 5086
Pol                         V  A  W  W  A  G  I  E  H  T  F  G  V  P  Y  N  P  Q  S  Q  G  V  V  E  A  M  N  H  H  L  K  N  Q  I  D  R  I  R  E  Q  A  N  S  V  E  T  I  V  L  M  A  V  H  C  M  N  F
H2A.GW.x.ALI                CA-------TT--T--AG---ATCC------G-A--T---------A-----C----------------G-A-----C-CC-CC---A--TCAG--A-CAGAA-T----G-----AA-TACAG-GG-A---AT-----T---------TCATTG-ATG--------- 5162
H2A.DE.x.BEN                CA-------TA--T--AG-A--ATCC------G-A--T-----C--A-----C--G-----------G-A-----C-CC-CC-------TCAG--A-TAGAA-T-----A-----AA-TACAA-AG-A---AT-----TG--------TCATTG-ATG--------- 5169
H2A.SN.x.ST                 CA------ATA--T--AG-A--ATCC--C---G-A--T--------A-----C----------G---G-A-----C-CC-CC----A--TCAG--A-CAGAA-T-----G-----AA-CACAG-AG-A---AT-----T---------TCATTG-ATG--------- 4614
H2B.GH.86.D205              CC-------TA-----AG-A--AACT------G-A-----T--C--A----------------G---G-A-----CC-TC-CC-G--A--TCAA---ACAGAC-C-----C--A--A-T-TCAA-AG------TT----T-------ACTCA-TG-ATG--------- 5142
H2B.CI.x.EHO                CC------ATA--G--AG-A--AAC---C---G-G-----T-----AG-------G------------G-A-----CC-TC-TC-G--A--TCAG---ACAGAA-T-----------A-T-TCAA-AG------TT--GTT-------ACTCA-TG-ATG--------- 5140
H2G.CI.92.ABT96             CC-------TA--C---G----A-C-------G-A--A-----C-----G--------G-----------A-----C-TC-TC---------CAA---ATA-AA-T--------A--AA-TTCAA-AG-A--CAT---GTT----------CATTG-ATG--------- 4496
H2U.FR.96.12034             CA-------TA--T--AG----TACC------G-A-----T--C--G-----C---------G-----G-C------C-CC-TC-------CCAA---AGAGAA-C-----A-----AA-TACAA-GG----CAT---C-T--------CTCA-TG-ATG--------- 4648
MAC.US.x.EMBL_3             CA--------A--G--AG--GCAC-CC--T-GG-A--A---------A--G-----G--------G---G-A-----C-CC-CC-G--A--TCAA---ATAGAA-C--G--A--A--AA-TTCAG-AG-A--CAT----TT---------TCATTG-ATG--------- 4554
SMM.US.x.H9                 C---------A--R--TG-A---ACC---R-GG-A--T--T-----A--G--C---R-G-----G---G-G-----CC-TC-TC---A-CCCAA---ATAGAA-T-----A-----AA-TTCAA-AG-A--TAT-----T-------A-TCATTG-AT---------- 4559
SMM.US.x.PGM53              CC--------A-----TG-----ACC--C--GG-A--T--T-----A--G--C------------G---G-A-----CC-TC-TC---A-CCCAG---ATAGAA-T-----A-----AA-TTCAA-AG---TAT-----T---------TCATTG-ATG--------- 5003
SMM.US.x.SIVsmH635F_L3      CC--------A--G--TG-A---ACC------G-G--T--T-----A--G--C------------G---G-A-----CC-TC-T----A-CCCAG---ATAGAA-T-----A--A--AA-CTCAG-AG---TAT-----T---------TCATTG-ATG--------- 5089
SMM.SL.92.SL92B             CA------TTA---G-GG-A---TCC------G-A--G-----C--A--------G--------------G-C---G-CTT-C-TC-T-------AC---ATA-AA---G--A--A--A--GTCAG-AG---CCT-----T--------C-CATTG-ATG--------- 4517
STM.US.x.STM                CC--------------AG-A--AACC------G-A--T---------A--G--C-----------G---G-C------C-TC-T-------CCCAG---ACAGAA-T---------A--AA-CACAG-AG-A--T-T-----TC--------TCA-TG-ATG--------G 4733
MNE.US.x.MNE027             CA--------A--G--AG----CACC-----GG-A--A---------A--G-----G--------G---G-A-----C-CC-TC---A--TCAA---ATAGAA-C--G--A--A--AA-TTCAA-GG-A--CAT----TT---------TCATTG-ATG--------- 4554
DRL.x.x.FAO                 -C--------TAAG--AG----TACCACA---G-------T--C-----G----------G-------A-----C--TC-GC-------------------AA-T-----------A---A-A--AG----------T------------C-G--T----------- 4354
MND_2.x.x.5440              -C--------TCA---AG----CACCACA---G-A--T-----C--A--G----------------G-A-AA---C-TC-TC-T-----G--C-----A---A--T--G-----A------AGG--AG-A--------G-----------GC-A--T----------- 4348
MND_2.CM.98.CM16            ----------TCA---AG----CACTACA------------------T--G-----------A---GG-A-AA--CC-TC-TC-C--AC----C-----G--A--T-----C--A--A--GA-A--AG-A--------T--------------A--T----------- 4788
MND_2.GA.x.M14              ----------TCAG--AG----CACCACA-----A-----T--C------------------T--G---G-A-AA---C-TC-TC-T---GT--------AG--A--T--G-----A--A--GAGA--GG-----------------------C----T----------- 4715
OLC.CI.97.97CI12            -C------TTA--G--TG----T---ACG--C------G-GG----T-----------A-CA-----G-A-AA---G--TG------G-G--AT---GAGAA-C-----A-----AAC----T-GG-A------T-GGCT-----TA-C-----T--T--------- 4658
GRV.ET.x.GRI_677            -C-------GTAAT--AG----CACCAC------A--T--T-C--A--GTCA--G--TAGT-----AGC-----C-G-C-GC-C---G---C--CTCT--AA--------G-TTG---GAGAT-GG---------G--------TACGCAT------------- 4960
VER.KE.x.TYO1_patent        ---------GAAA---TG-A--TAC-ACA--T--A--A--T--C------TC-------TC-A----AGC-----C---C-------AG-G--A--T--GA-AA--------G-TTGCC--T--ACAG-----------TG-----TTGCCAT--T----------- 5011
VER.KE.x.9063               ---------GAAAG--AG----CACCACA---G-G--A--------A--GTCA--G---TC-------AGC-----C---C--C-C--AG---A--T---A-AA--------G-TTG-C-GT--ACAG-A----------TT------TGCCA--------------- 5014
VER.KE.x.AGM155             -C-------GAAAGG-AG-A--CAC-ACG--GG-A--A--T-----A--GTCC--G---TCTA-----AG------C---C----G-AG-G--A--T---A-AA--------G-CTG-C--T--ACAG-A----------TT-----CTGCCA---T----------- 5006
VER.DE.x.AGM3               ---------GAAA---AG-A--CACCACA--TG-A-----T--C--T---TCA-----CTCTA---GAG---------C-G-----AG---A--T--GA-AA--------G-CTG-C--T-CACAG-A----------TT------TGCCA--------------- 4509
TAN.UG.x.TAN1               ---------GACA-G-AG----CAC-ACA--G--C-----T--C--A---TCA-----AGT------AG--------G-C-------AG---C-----A-A-AA-T-----AG-CTG--C-TT-T-GG-A--G-----C-TC-----CTGCCA---T----------- 4986
SAB.SN.x.SAB1C              -C-------GAAA---AG----CAC------GG-A--------C----------G------------------GC-G-----AG---C--------AA-T------G-T--A---AGAT-GG-A--------CAT---------GCA---T----------- 5185
MND_1.GA.x.MNDGB1           --------TTA--G--AG-A--CAC----------A-----T--C--A-------G--G-----G--AA--AA----GT-TC---AG--T-G--T-AGA-AA-------AG-TTGCA--G-AT-A--A---------GCC-----CAC------T--T--------- 4474
DEB.CM.99.CM40              -C------CTA--T--TG----CACCACA--GG-A--A--T--C--A-----C--G-----G-----AG--A---C-GGATTC---AG-T-CC--CAA---AA-C-----G-T--A--GAGAT-GG-A---------GC------TACC-A---AATT--------- 4593
DEB.CM.99.CM5               --------TT------AG-T--TACTACA--GG-A--A--T--C--A-----------G--------AGC-AA---GG-TCC---AG---C---CAA---AA-T-----AG-T--A---AGA--AG-A---------GC------CACC-----TATT--------- 4587
TAL.CM.01.8023              CG------CT-AA---AG----CACCACA--GG-A--A--T-----A---TCC-----GTC------GAA--A---C---C-G--G-----G-CA--TCA---AA-T--G---G-A-TGC-GT-C--GTCC-----G--GGC-CA---TACC-----T-TT--------- 4251
TAL.CM.00.266               TA------CTCAA---AG----CTCCGCA---G----A--T-----T---TCC--G----C-------AAC-AA-----C-GC-C-----G-CA--CAC---AA-T--G---G-A-TAC--T-C--GTCC--------GGC-CA---GACT-----TTTG----A---- 4753
SUN.GA.98.L14               --------TTA--C--AC----TAC-ACA--G--A--T-----C--T--G--------------G---G-A-A---C--G-T------AC------CA-CAGAA-T-----AG-T--AC-GG-AT-G--A--T--------TC-----TC-GCA---T--T--------- 5147
WRC.CI.97.97CI14            -C------TTA-----AG-A--A--------G--C-----T-----T---------------A-----AG--A---C-GGAGG----AG-C------CACAGTA-T-----AG-T--A---A-AT-AG-A------T--GCG------T-------T----------- 5087
WRC.CI.98.98CI04            -C------TTA-----AG-A--A--------G--C--T--T-----T---------------A-----AGC-AA---G-AGGC---AG---A---CACAGAA-TC----AG-C--A---A-GT-AG-A------T--GC-------C-------T----------- 5102
WRC.GM.05.Pbt_05GM_X02      -C------TTA--C--AG-A-----------G-------------T--G-----------A-T--GAG--AA---G--T------AC---A---CA-AGAA-------AG-T--A---A-AT-GG-A------T--GCT------A-C-A---T----------- 4016
LST.CD.88.524               CC------TTAA-T--AG----CAC-ACA--G--A--A--T--C--T------------------GG-A-AA-------TT------C----C----A-AGA--------AG-T--AC-G--G--A-----------GTT----------CCA---T--T--------- 4061
LST.CD.88.447               T-------CTA--T--TG-A--TAC-ACA--------T--T--C--T--G-----G--------T---G-C-AA-------C-C---AC---------A-AG---C----AG-T---C----AT-A--------T--C-T-------AC-CA---T--T--------- 4064
LST.KE.x.lho7               CC------ATA--G--AG-A--TAC-AC---G--A--T--T--C--T--G--------G--T------G-A-AA--C----TT-----A-GT--------A-AGA--------AG-T--AC-G--A--A--A-----------T----------CCA---T--T--------- 5148
SYK.KE.x.SYK173             --------TGT--C--AG-A--CACTACA--T-----A-----C--A-----C--G---A-------G--C-AA----GGT-T-----AG--GC---CTC---AA-T-----CG-C-TCAC----T-GC-A--------GGC-------ACC------TTA---------- 4987
SYK.KE.x.KE51               ----------A---G--G-A--CACCACA-----A--A--------A-----C--G---A--------AAC-AA----G-C-------AG---CC--CAAT--AA-T--G--AG-A-TA---AGAT-AG-------G--GGC------CACG-----TTTA---------- 4642
COL.CM.x.CGU1               TG------CTA-----AGCA--CACCACA---CAC-----T-----T---TCA--------T-----GCAG-GA--C--G--TG-T--A-G--AG--TAAGA-AA-G-A------A--A---ACAT-AG-A-GCAA----GC---------C-ATGCG-T---------- 4550
```


PLV Complete
Genomes

```
H1B.FR.83.HXB2            AGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATT...ACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAAATCCACTTTGGAAAGGACCAGCAAAGCTCCT 4954
Pol                       _R__K__G__G__I__G__G__Y__S__A__G__E__R__I__V__D__I__I__A__T__D__I__Q__T__K__E__L__Q__K__Q__I__.__T__K__I__Q__N__F__R__V__Y__Y__R__D__S__R__N__P__L__W__K__G__P__A__K__L__L
H1A1.UG.85.U455           --------------------A----------------------A---------------------------------------------------...T-------------------------------------------------G-C--CA-------------------A--A-- 4400
H1B.US.90.WEAU160         -----------------------------------------A-----------------------G--C--C-------------------...----------------------C------------------G-----------------------------T-- 4953
H1C.ET.86.ETH2220         ----G------------------------------------A---T---------T---------G--------C-C-----C--------...TT--------------------------A--------G-C--TA--------------C--A--A-- 4346
H1D.CD.84.84ZR085         -----------------------------------------A-------------T----T---------G-----------------C...------------------------------------------G-----A-----------------------T-- 4476
H1F1.BE.93.VI850          -----------------------------------------A-----------T-----------------G--------------------...--G-------------C------------------------------C---G-------------------------A-- 4294
H1G.SE.93.SE6165          --------------------------------A--------A-------------T----T-------------C------------------...------------------------------------------G-C---G---------------C--A--A-- 4351
H1H.CF.90.056             -----------------------------------------A--------------------------------------------------...T---C-------A---------------------------G-C---A-------------------A---- 4301
H1J.SE.93.SE7887          --------------------A--------------------A----------------------------G---------------------...----------------------------------------G-C---A-------------------A--A-C 4268
H1K.CM.96.MP535           ----------------------------------------------T-----------------------------------------------...TT--C-------A---------------------------G-A---A----------------------A-- 4150
H101_AE.TH.90.CM240       -----------------------------------------A--------------------------------------------------...------------------------------------------G-C---A-------------------A--A-- 4528
H102_AG.NG.x.IBNG         -----------------------------------------A---T---------T----T-------------C---------G-----...-T----------------------------------------G-C--CA-------------------A--A-- 4479
H1N.CM.95.YBF30           -------------------------C---------------A--------------------------G--A-C-A-T-------C-------...TT---G-----------------A--------G---CA---------------C--A--T-- 4548
H1O.BE.87.ANT70           -------------------------C------A-----G--A--------T-----T--C-A--------A-C---------------------...TT---N---C--A-------C-----C-------A--------G-C--TA-C----------G---C------- 5009
H1O.CM.91.MVP5180         -------------------------C--------G---C--A--------C-----T--C-A--------A-C----C----------------...TT------CA-C---------C-----C-------A--T-----G-C--TA-----------G---C-A---- 4984
CPZ.CD.90.ANT             -------------------------CAC-T--AC-G------T--------C------------------G--A-CTC----------T------...TT------------C-A---------C------C---------T-----G-C--TG-G-----------C-A--T-- 4393
CPZ.CM.01.SIVcpzCAM13     -------------------------C---------------A------------------------A-----C------------------...TT------------A------------------A--------G-C--GA-------------G--C-CC--G-- 4534
CPZ.CM.05.SIVcpzEK505     --------------------A----C---------------A-------------------T--------A-C-A-GC------C--------...TT---G--------------------------A--------G-G--CAC---------------C--A--A-- 4526
CPZ.CM.05.SIVcpzMB66      --------------------------------A--------A--------C------------T----------C-A-------C---------...T---T--------A---------------------------G-C--TA----------------T--A--A-- 4500
CPZ.CM.98.CAM3            -------------------------C------A-------TA-------C-----------------------A-CTA-TC-------------...TT---G---------------C-----C------------G---CA--------------T--C-GA--T-- 4368
CPZ.GA.88.GAB2            -------------------------C------A------------------GC-----------T-------AC-----C--------------...CT-------------A------------------A--------G-C--CA---------------CTCT----- 4347
CPZ.CM.98.CAM5            -----------------------T-C------A-----G--TA-------C--------------------A-CTA-T----------------...TT---G---------------------------------G-----A--------------T--C-GA--T-- 4656
CPZ.CM.05.SIVcpzMT145     --------------------------------A--G-----A--------C--------------------A-----C----------------...TT-----------C-A---------C------A------A-G-----A------------G--A---- 4496
CPZ.TZ.00.TAN1            -------------------A-----C-------------C--T--------C--A-T---A-T-----G--A---C------------------...TT---G--------------------------T------GC---G-----A-------G--------GCGA--A-- 4598
CPZ.CM.05.SIVcpzLB7       --------------------A-----------A--------A---------------------T-----------C------------------...T----G--------A--------------------C-----G---TA------------------C--A--A-- 4509
CPZ.US.85.CPZUS           -----------------------T-C------A-------CA-------C-----T---AC------AG-CTT------------------...TT---G---------------C------------------G-----A------------------C--A--T-- 5014
CPZ.GA.88.GAB1            -------------------------C---------------A-------------------------A-GC---------------------...TT---G--------A------------------A--------G-C---A--------------G----CCT-A-- 5015
MON.CM.99.L1              -----G--A--A--A---ATATGGCCC-T-TA--C---T-TA--A---G--CAC----AT--G--CTAC--ACC-C-A-C--C-------...-------T--TC--A-----------------T----ACAG-TGCCG-C--T-AG--------------GCGC--A-- 4657
MON.NG.x.NG1              ----RG--A--A--A---ATAT----C--ATA--G--GT----CA---G--TCAT----TT--G--CTAC--C-C-C-A-C--C-------...-------T--TCG--A----------------T---AACAG--GCTG---TTCG-----G---------C-T--T-- 3250
GSN.CM.99.CN166           -----G--A--A--A---ATATGT-CC-TACA-------------A-T--G--TCAC----AT--G---AC--C-C---A-C-C-----A-...T-C---T------A-------------------------C-AG-AGCT--C--T------C-----------GTA----- 4682
GSN.CM.99.CN71            -----G--A--A--A---ATATGT-CC-TACA--G--GT----GA---G--CCAC----AT--G---AC--C-C---A-C-C-----G-...T-C---T------A-------------------------C-AG-AGCT-----T-A----C-----------GTA--T-- 4664
MUS_1.CM.01.1085          -AG--G--A--AG-G----ACCTATCAAAT-CT--G--G---A-CA---GC--CAC----AC--G--CTAC--C-C------C-----AA...TAC---T---------------------------T---ACAG-TTCTG---GTC------G-----C--CGTCT-A-- 4644
MUS_1.CM.01.CM1239        -AG--G--A--A--A----ACCTAGCA--T-CT--G------A-TA-T--GT--CAT----AC--G--CTAC--C-C------TC----AA...T-C---T---------------------------T---ACAG--TCTG-C---TC------G-----T--TGCCT-GT- 4634
MUS_2.CM.01.CM1246        -----G--A-----A----ATATG-CCC-TACA-----GT----TA-T--GC--TAC----AC--G---TAC--C--C-------T--C-A-...------T--TC--A------------------------AAGAG-AGC---C--TTA----CTGA-------TCGA--T-- 4742
MUS_2.CM.01.CM2500        -----G--A--A--A----ATATG-C-C-CACA-----GT----CA---GC--TAT----A---G---AC--C--A-------CAC--AA.........T--TC--A------------------------AAGAG-AGCG-------AG----C-------T--TC-A----- 4695
RCM.NG.x.NG411            --------------------------------A----G--A---------------T-----T--------A-CCA-----------C-------...T----------------------------------T---A--AG-A--GG---AG-AG-----G--------T--A---A- 4448
RCM.GA.x.GAB1             --------------------------------A----------A----T---------------T--GC---C--TA-----------T-------...T-------------------------------------A--AG-A--GG---A----G-----G---T-----T-----GA- 4433
DEN.CD.x.CD1              ----GG--A--AC-A----ATATG-CA----CA------TACA-TA----GC--AAT---C-TT--G---TAC--C-T...ACC-------......C-----T--TCG------AA---C---------A--GG-AGCT------G-AA---C-----------TG----G-- 4838

MAC.US.x.239              ----GG--A--A--A----ATATG-C-C---CA------T--A-TA----G--CA-T-----ACA-G-G-TAC--TTTCA----TC-A--AC...T-----T--A-------------C---------A--AG-----G---A---G-----G-----C-GTG----AT- 5253
Pol                       _K__R__R__G__G__I__G__D__M__T__P__A__E__R__L__I__N__M__I__T__T__E__Q__E__I__Q__F__Q__Q__S__K__.__N__S__K__F__K__N__F__R__V__Y__Y__R__E__G__R__D__Q__L__W__K__G__P__G__E__L
H2A.GW.x.ALI              ----GG--A--A--A----ATATG-CCC---CA------C-CA-CA-T--G--CT-C-----ACA-G--TAC--TTCC-C----C-A--A-...TTG---T--A--------C-C---C----------AG-----G---AG--G----------T-GGG----A-- 5329
H2A.DE.x.BEN              ----GG--A--A--A----ATATG-CCC---CA------C--A-CA----G--TA-C-----ACA-G--TAC--TTCC-C----G-A--A-...T----TT--A-------C-A---C--------A--AG-----G---AG--G----------T-GTG-A--A-- 5336
H2A.SN.x.ST               --G-GG--A--A--A----ATATG-CCC---CA------C--A-CA-T--GG-CA-TG----ACAGG--TAC--TTCC-C---GC-A--A-...T-----T-A--------------C-----T---A--AG-----G---AG--G----------T-GGG-A--A-- 4781
H2B.GH.86.D205            ----GG--A--A--A----ATATG-CCC-T-CA------C----TA----G---A-C-----GCA-G--TAC-GTTC--C---GC-A--A-...TT----T----------C-A---C--------A--AG-----G---A---C-----G-----T-GTG-A--AT- 5309
H2B.CI.x.EHO              ----GG--A--A--A----ATATG-CCC-T-CA------------CA---G---A-T-----ACA-G--TAC--TTCC-C----C-A--A-...TT----T-C--------C-----C--------A--AG-----G---A---C-----G-----T-GTG-T--AT- 5307
H2G.CI.92.ABT96           ----GG--A--A--A----ATATG-CCC-T-CA----------A-TA---G--TA-T-----ACA-G--TAC--TTTCA-----C-A--A-...T-----T--A--------------C--------A--AG-----G-C-AG--C-----------T-GTG----TT- 4663
H2U.FR.96.12034           ----GG--A--A--A----ATATG-C-C-T-CA--------CA-TA----G---T-C-----ACA-G--TAC--TTCC-C---TC-A--A-...T-----T--A--------------C--------A--AG-----G-C-A-T-G--------C------G-CT-A-- 4815
MAC.US.x.EMBL_3           ----GG--A--A--A----ATATG-C-C---CA------T--A-TA----G--CA-T-----ACA-G--TAC--TTTCA----TC-A--AC...T-----T--A--------------C--------A--AG-----G---A---G-----G-----C-GTG----AT- 4721
SMM.US.x.H9               ----GG--A--A--A----ATATG-C-C---CA------T----CA-T--G--TA-C-----ACA-G--TAC--TTCCA----TC-A--A-...T-----T--A--------------C--------A--AG-----G-C-AG--G-----------C-GTG----AT- 4726
SMM.US.x.PGM53            ----GG--A--A--A----ATATG-C-C---CA------T----CA-T--G--CA-C-----ACA-G--TAC--TTCCA----TC-A--A-...T-----T--A--------------C--------A--AG-----G-C-AG--G-----------C-GTG----AT- 5170
SMM.US.x.SIVsmH635F_L3    ----GG--A--A--A----ATATG-C-C---CA------T----CA-T--G--CA-C-----ACA-G--TAC--TTCCA----TC-A--A-...T-----T--A--------------C--------A--AG-----G-C-AG--G----G------C-GTG----AT- 5256
SMM.SL.92.SL92B           ----GG--A--A--A----ATATG-C-C-T-CA-------------A-T--G--CA-C-----AC--G----CC--T-C---A-TTC-----A-...T-----T-------------------------A--AG-A---G-C-A---G-----G---------G------- 4684
STM.US.x.STM              -A--GG--A--AC-A----ATATG-CCC---CA------T----TA-T--G--CA-C-----ACA-G--TAC--TTTCA----TC-A--A-...T-----T-CA--------------C--------------AG-----G---A---A-----G-----C-GTG----AT- 4900
MNE.US.x.MNE027           ----GG--A--A--A----ATATG-C-C---CA-----GT--C-TA----G--CA-T-----ACA-G--TAC--TTCCA----TC-A--AC...T-----T--A--------------C--------A--AG-----G---A---G-----G-----C-GTG----AT- 4721
DRL.x.x.FAO               --------------A----A-----------CA-------------------T-----CT-T---C-CTT---A-GCA--C-----C--A-T---...CTC----------------------------------A--GG-A---G---AGT-G----------------G-A--TG- 4521
MND_2.x.x.5440            --------------A----A-------C-T--A----------------------A-C---------C-T---A-CCA-------G-C--A-T---...T-------------------A------------------A--AG-A--GG-C-A--AG----------------G-----A- 4515
MND_2.CM.98.CM16          --------------A----AA------C----A-----------------T--------------C-C-T---A-CTA--------C-CA-T---...CA-----------------------------------A--AG-A--GG-C-A--AG----------------G-A---A- 4955
MND_2.GA.x.M14            --------------A----A-------C----A------------G--T------A-------T--T-T-----CCA--------C--A-T---...T--------------------------------------A--AG-A--GG-G-AG--C----------------G-A---A- 4882
OLC.CI.97.97CI12          -AC--G--A--ACAA-----AGGT-CA-TG-CA--GGA-T-----A--CA----TAC-----T-C-...CAA-T-----ACAC-C--A-T-AA...TTT----AATTTTCAGGATATCGAGTG--TTATA-A-A--A-G-AGGC-AG---TGT--C------G-A----- 4822
GRV.ET.x.GRI_677          -----G--A--A--A-----TAT-TC-AGT-CA------T-G--TA-T--GC--A-----C-AC--G--CTA--TACTC-------C-----C...CA-------TTG------AA---C--C-----A--AG-T---G-----G-G-------------GCGA---A- 5127
VER.KE.x.TYO1_patent      -----G--A--A--A-----AC-G-C-T---CA--G---C--A-TA-T-----A-----C-AT--G---TAC--C-T-------CCA-----...CA-------TT-------A-A--C--C-----A--AG-G---G-C--TG-G----------------C-AT-AA- 5178
VER.KE.x.9063             -----G--A--A--A-------TA-CAC---CA------C--A-TA-T--G---A-----C-AT--G--CTAC--AC---------CCA-----...CA-------TT-------A-A--C--C-----A--AG-G---G-----G-G-----G---------C-AT-AA- 5181
VER.KE.x.AGM155           -----G--A--A--A------CTA-CA--T-CA--G---C--A--A-T--G---A-----C-AT--G---TC--CACTC------CCA----C...CA-------TTG------A-A--C--C-----A--AG-----G-----G-G-----G-----T--TCGC--GA- 5173
VER.DE.x.AGM3             -----G--A--A--A-------TA-CAC-G-CA--G---T--A-CA-T--G--TA-T---C-AT--G--TTAC--C-CC------CCA-----...CA-------TT-------A-A--C--C-----A--AG-G---G---TG-C--------------G-C--T-AA- 4676
TAN.UG.x.TAN1             -----G--A--A--A----TATG-CCC---CA--G--GC--A--A-----G---A-C---C-AC--G---TAC--C-CA------C----CAA...CA-------TC-------AA---C--C-----A--AG-A---G-G---G-G--------------C-CT----- 5153
SAB.SN.x.SAB1C            ----------------------TC-----CA------C--A-CA-T------CAT-----AC--G----C---AC-C-T---C--A-----...T---------------------------C-----C------------AG-----G-C--TG-G--------T--G--T------A- 5352
MND_1.GA.x.MNDGB1         CA--GG--A--AC-A------ATG-CA-----A--G------------A-T--G--CAAT-----AT--G--TA-C--T-TCA-----T-------...T-----A-TT-------AA------C-T---A--AG-A---G---AG--G-----------T-GT-TC--TT- 4641
DEB.CM.99.CM40            ----GG--A--A--A----ATATGTC-C-GATA------C-----A----G--A-------A---G---A---ACT------C--A-------...T-----C--TT-GGC--CAA---C--------A--AG-TGCTG-----AC---------G-------CC--G-- 4760
DEB.CM.99.CM5             ----GG--A--A--A----ATATGGCA---ATA--G---C-----A----G--A-T------T-GG----AC--ACC------C--A-------...T-----C--TT-GG--ACAA---A--------A--AG-AGCTG-C---AC------G-----C--T-CCT-A-- 4754
TAL.CM.01.8023            ----G------AT-A----ATATGT--C---CA---GC-C-CA--A----G--CTAC-----AC--------A-C-ACCC-------C-------...CAC--TT--TC-G-----AA---G--C--T--AA-GG-TGCT-----T------C-G--------TC-C---G- 4418
TAL.CM.00.266             ----G------AT-A----ATATGT-CC---CA---GC----A-TA----G--CTAC-----AC---------C-C--C-------C-------...CA---TT--TCGG-----AA---C--------AA-GG-TGCT-----T------C-G--------TC-C---A- 4920
SUN.GA.98.L14             CA--GG--A--AC-A-----AATG-CA----CA--G--GT-TA-CA----G--TAATG----TC--G----AC--T-C-----G---AT--A-...T--------TT---A---AA---C--C-----A--AG-A---G-----T-AG------------T--T--A--T-- 5314
WRC.CI.97.97CI14          GA---G--A--AT-A-------...TCCC-T-CA------T-----A-T--G---A-CT----TC--G----AC--C--ACT---C------AA...TT----T--A---------A---G--------AACAG-AGCC----AG-AG---C----------GT-------C 5251
WRC.CI.98.98CI04          GA---G--A--AT-G-----A...TCCC---CA------T-----A-T--G--CA--G----AC--G----C--TC--AC----C------AA...TT----T--A---------A-------------AACAG-AGCC----AG-AG---C----T----GT-------- 5266
WRC.GM.05.Pbt_05GM_X02    GA---G--A--AT-G-----A...TCCC---CA-----------A-T--G---CA-T----T---G------TC--AC----C------AA...TTC---T--A---------A------C------ACAG-AGCCTCG-AG-AG---C----------GT--AT-GG- 4180
LST.CD.88.524             CA--GG--A--AC-G-------TATCAC---CA--G---TACA-CA-T--G--CAATG----TT--G-----C--T-CA---------TT--A-...T--------TT---A---AA------C-----AC-AG-A---G------AG----G------T--TC-A--T-- 4228
LST.CD.88.447             CA--GG--A--AC-A------CT-TCAC---CA-----GT-TA-TA-T--G--TAATG-----T--G----AC--T-T------------T--A-...T--------TT---A---AA------C-----AC-AG-AC--G----C-AG-----G-----T---C----T-- 4231
LST.KE.x.lho7             CA--GG--A--AC-A------CTT-CAC---CA--G--GT-TA-TA-T--G--TAATG----AC--G----AC--T-TC---------TT--A-...T--------TT---A---AA------------AC-AG-A---G----T-AG--------------GC-A--T-- 5315
SYK.KE.x.SYK173           -----G--A--A--A------ATATCAC----A-----GTACA-TA-T--GC--TAT-----AC-----TTAC--C--AATAC--C-TC-......C-----T--TCG------A-------C---C-CC-GG-G-AG...AACGAG-----G---------TCGT--A-- 5148
SYK.KE.x.KE51             -----G--A--A--A------ACT-CAC---CA--T--GTATA-TA----GC--TAC-----AC-----CTAC--A-TAC---C-C-.........CA----T--TCG------AAA--A-----T--AC-TG-T-CT...T-TGA----C-G---------TTCT--A-- 4800
COL.CM.x.CGU1             -----G--A--AT-A------A-G---C--T--------CA------A-G-GC-AT--T---AT--G-T--AC--A-CC--AC----TT-CAAAATCA----T--A--------AA--CC--C-GG-AA--GCA-....-CAGG-GAG---C---------G-G-A--GG- 4717
```

```
                                                                                                            Vif start                              Pol p31 integrase end
H1B.FR.83.HXB2          CTGGAAAGGTGAAGGGGCAGTAGTAATACAAGATAATAGT...GACATAAAAGTAGTGCCAAGAAGAAAAGCAAAGATCATTAGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAG......................... 5096
Vif                                                                                                   M  E  N  R  W  Q  V  M  I  V  W  Q  V  D  R  M  R  I  R##
Pol                       W  K  G  E  G  A  V  V  I  Q  D  N  S  .  D  I  K  V  V  P  R  R  K  A  K  I  I  R  D  Y  G  K  Q  M  A  G  D  D  C  V  A  S  R  Q  D  E  D  *
H1A1.UG.85.U455         ----------------------------------C------...--T----G----A---------------------------------------------------------------A-----G-----------A------......................... 4542
H1B.US.90.WEAU160       ----------------------------------------...------------------------------------------------------------------------------------------------------......................... 5095
H1C.ET.86.ETH2220       ----------------------------------------...--------G-----A-----G--G--------A------------------------------C------------G------------A------......................... 4488
H1D.CD.84.84ZR085       ----------------------------------C------...--------G-----A-------------------------------------------------------------------------------------A......................... 4618
H1F1.BE.93.VI850        -----------------------C--------C------...--A----GA---A-----G--G-------------------------------T--------------G-----------------A......................... 4436
H1G.SE.93.SE6165        ----------------------------------C----AC...--A----G-----A---------------------------------------------------------G------------------------......................... 4493
H1H.CF.90.056           ----------------------------------C------...--A-----------A--------G-G-----A--------------------------------------------------------------A......................... 4443
H1J.SE.93.SE7887        -----------------------------G--C------...--A----G-----A---------------------A---------------------------------G-----------------A......................... 4410
H1K.CM.96.MP535         ----------------------------------C--C----...--A----G-----A-------------------T-------------------------------G-----------------A......................... 4292
H101_AE.TH.90.CM240     ----------------------------------C------...--T-----------A---------------------------------------------------G-------------------------......................... 4670
H102_AG.NG.x.IBNG       -----------------------------G--C------...--T----G-----A-------------A-------------------------------G-------------------------......................... 4621
H1N.CM.95.YBF30         G--------A-----------G-----T---------CG-G...--T-----------C---C-T--G--------A--A-----------------------G-------------G-----------AA--C--GAAATGGAATAG............. 4702
H1O.BE.87.ANT70         G-----------G--A--------C-----------GG-A...-----T--G-----A--------G--G----A--A--C--A--G--------------------AC----A--A--------G-----ACA--AAG-G-AAGCGTGGAACAGCCTAGTGAAATAC 5176
H1O.CM.91.MVP5180       G-----------G-----------C-----------AG-A...-----T-----G--A-----------G----A--A--C--A--------------------AC----A--A------A-------ACA--AAG-G-AAGCATGGAACAGCCTGGTGAAATAC 5151
CPZ.CD.90.ANT           G-----------------------C--CA----CC-AGAG...--A--T--G-----C--TC----G-----------T--A-A---G---A-----A----A-A--A-AGA---GA-T-----G--------A----------ATATCTGGAAATC............. 4547
CPZ.CM.01.SIVcpzCAM13   G---------------------------T--G---C-AG-A...--A-----------A--T----G--------A--A----A---------------A-------------------------------------AAGCATGGAATAG............. 4688
CPZ.CM.05.SIVcpzEK505   G--------A-----------G-----T---------CG-G...--T-----------C---C-T----G--C--A--A----A---------------------------G-------------G-----------AG-C--GACATGGAATAG............. 4680
CPZ.CM.05.SIVcpzMB66    -------------G--A--------GC-C-----CC-AGAA...--A-----G-----C--TC----G--------A--------------------------------------------------------A--C-AAACATGGAATAG............. 4654
CPZ.CM.98.CAM3          G---------------------G-----TA----G-GAGAG...--AG-T------A-A--C-----G--------A--A-----------------------T---------A--A-----G--G-----------AG-C-AGGCCTGGAATAG............. 4522
CPZ.GA.88.GAB2          G-----------G--T--------T--T--G--C--CG-A...--T-----------A--C-----------T--A--T----AAC----------G---G-----------------C----TG--G----------A--C--AAC......................... 4492
CPZ.CM.98.CAM5          G-----------------------G--TA----G--CGAG...--AG-T---------A--C--------G--T----A--C--A------------------------G-A--A-----G-------A--------C-AAACCTGGAATAG............. 4810
CPZ.CM.05.SIVcpzMT145   G---------------------G-----T------C-GGAC...--A-------G-----T-----G-----T----A--------------------------CA---A--------------------------AG---AAACCTGGAATAG............. 4650
CPZ.TZ.00.TAN1          G---------------------------------A----AGGAGAA...-----T--------A--C--G-----------A-----A-AA--G--------------------CA-G-G--A---ATGA-------A-----AC----AACATGGACAAG............. 4752
CPZ.CM.05.SIVcpzLB7     G----G--------A--------GC-G------C-GGAA...--A-----G-----C---C-T--------T--A--T--------------------A-----C-A-----C-----GA--------------A----AAGCATGGAACAG............. 4663
CPZ.US.85.CPZUS         G------------------------CA-G--A--CGAG...--GG-T---------A--C---------------A--T--A-AA--C--------------------C----A--A-----G-------------AG---AGGCTTGGAATAG............. 5168
CPZ.GA.88.GAB1          G-----------------------G--C--G---C-AG-G...--AC----G-----A---------G-----------------A-------------------------------------A----------AGACTTGGAATAG............. 5169
MON.CM.99.L1            ---------------T--C-----C---A--AC-G-GGAGGGC--T--CCTGAC---C--T--G----G-------A--C-AACCC-----C-CCAC-GAAAAT-TG-G--G-GACACCAACCAATATA-TCTCAGAA-GC-AGATGGATTGGCTAATTAG....... 4820
MON.NG.x.NG1            ------G--G-----A--TT-G--G--TA-GACAG--CAAGGT--GG---T-AC---A---C-C--G--G-----------C-CCC------CTCCA--GAA-AT-TG-G-AG-AAGCCCAGTGCACAT---ATCAGA--GGT-GATGGAATGGCTGATTAG....... 3413
GSN.CM.99.CN166         ----------------A---------G-C---AC-C-AGCAGGT--G--T-TTAC---C--T-----G---------------C-A-CCA------G-GGCA-A-A---AA----TGG---GT----CCCATACAA--A-T--C-GAGAGGAAGGTAGACTGGTTGCTTAG. 4851
GSN.CM.99.CN71          ------G--------A--------GG-C---ACCC-AGCAGGT--G--C-TCAC---C--T-----G---------------C-A-CCA------G-GGCA-A-----AA----TGG--AGT---GCCCATACAAG-A-T--CAGAGAGGAAGGTAGAATGGCTGATTAG. 4833
MUS_1.CM.01.1085        G----G--------A--------C--CA--ACAG-CTCAGGA--GG---TCAC---T--T--------G-----A--A--C-AACCC-----CGCCA--GAA-AT-TG-G-AG-AAGTCC-AT-CAG--G-CCTCAGAA-GG-AGATAGATTGGATAATTAG....... 4807
MUS_1.CM.01.CM1239      G-----G--------A--------CG-CA--ACAG-ACAAGGAC-GG---TCACT--A--T--------G-----A--A--C-A-CCA-----CGCCA--GAAAAT-TG-G-AG-AAGTCCAAT-CAG--G-CC-CAGAA-GG-AGATGGATTGGATAATTAG....... 4797
MUS_2.CM.01.CM1246      T--------A-----A--TC----G--CA--ACA--GGAAGGT--G--CGT-AC---C-----------------A--A--C-A----------C-CCAG-CAAAAT-TG---AG-GAGCCC-ATTCAGT---T-TCAGAA-GG-AGATGGATTGGCTGATTAG....... 4905
MUS_2.CM.01.CM2500      ------G--------A---A-------CA--ACAG-CCAAGGT--G--T-T-AC---A-----G--G--G--------T--C-AA--C-----C-C-A--GAAAAT-TG---AG-GACCCCAATCCAGTTT-T--CAGAA-GG--GATGGACTGGCTGATTAG....... 4858
RCM.NG.x.NG411          ----G---A--------T--C--G--C--G---C-AGAG...--TC----G-----C--T--------GTGT--------A-AA--------C-G-A-AGA--TG-A-AGA--GACCAATATGGAAG-TAGAC-G-----AAGTTGA..................... 4593
RCM.GA.x.GAB1           ---------A-----A--------C--C--G--GG-G-C-GGA---T-G--G-----C--C--G-----------A-----A-AA--A-----C-G-A-AGAT-TG-A-AG---GGCCAATTTGGCGG-TAGAC-G-----AA-TTGA..................... 4581
DEN.CD.x.CD1            T--------------A-----T--C--CA-GACAC-GGCAGGA-----CTTCCCG--A--T---------G-----A-----C-A-CCA--AATGC-G-A...T-TA-GA-G-GAGAACA-AAC-AAG--AT----AGT--GACTGACATATAA.............. 4989
                                                                                                            Vif start                                Pol, Gag-Pol, and p31 Integrase end
MAC.US.x.239            G--------G-----A-----CA-CT--A-G-TAGGG-CA...-----T--G-----A--C--------G--T--A--T--C-AA---------GG-GGA-AA-AG-TG---AGCA--TCCCAC-TGGAGG-TACC-GA--GGCTAGAGAGGTGGCATAG.......... 5410
Pol                       L  W  K  G  E  G  A  V  I  L  K  V  G  .  T  D  I  K  V  V  P  R  R  K  A  K  I  I  K  D  Y  G  G  G  K  E  V  D  S  S  S  H  M  E  D  T  G  E  A  R  E  V  A  *
Vif                                                                                                   M  E  E  E  K  R  W  I  A  V  P  T  W  R  I  P  E  R  L  E  R  W  H  S##
H2A.GW.x.ALI            G--------G--C--A-----CA--G-TA-G-TAGGG-CA...---------------A--------G--G--C--------C--A--C------GG-AG-CAA-AGCTG---AG-G--CCCCACTTGGAGGGT-CCAG---GG-TGGAGAAGTGGCATAG.......... 5486
H2A.DE.x.BEN            G-----G--A-----A-----CA--G-CA-G-TAGGG-CA...---------------A--------G--G--C-----T--C-----C------GG-AGACA--A-CTG---AG-A--CCCCACCTGGAGGGT-CCAG---GG-TGGAGAAATGGCATGC.......... 5493
H2A.SN.x.ST             G-----G--G--C--A-----CA--G-CA-G-TAGGGGC-...----------A--A-A--------G-----T--------C-AA--C------GG-AG-CAA-AGATG---AGCG--TCCAACTTGGAGGGT-CCAG---GG-TGGAGAGGTGGCATAG.......... 4938
H2B.GH.86.D205          G--------G-----A-----CA-C---A-G-TAGGG-CA...--A--C--------A--C--G------------A--T--A---C-C------GG-GGA-AA-G-TTG---TG-A---CC-AC-TGGAGG-TACCAG-C-GGCTAGAGAGATGGCACAG.......... 5466
H2B.CI.x.EHO            G--------G-----A-----CA-C---A-G-TAGGG-CA...--A--C------A-A--C--------------------A---A-C------GG-GGA-AA-A-TTG---TGCA---CC-ACGTGGAGG-TACCAT-C-GGCTAGAGAGGTGGCACAG.......... 5464
H2G.CI.92.ABT96         G--------G-----A--T--CA-C---A-G-TAGGG-CA...--A--C-----T--A--C--------------A-----A-AA---------GG-GGA-AA-AG-TG---AGCA--ACCAAC-TGGAGG-TACCAG-C-GACTGGAGAAGTGGCATAG.......... 4820
H2U.FR.96.12034         G-----G--A-----A--G--CC-T--TA-G-TAGGG-CA...-----T--G---A-A--C--G-----------------C--A---------GG-GGA-AA-A-TTG---AG-A--TCCAACTTGGAGGGT-CC---A-GGTTAGAGAGATGGCACTC.......... 4972
MAC.US.x.EMBL_3         G--------G-----A-----CA-CT--A-G-TAGGG-CA...-----T--G-----A--C--------G--T--A--T--C-AA---------GG-GGA-AA-AG-TG---AGCA--TCCCAC-TGGAGG-TACC-GA--GGCTAGAGAGGTGGCATAG.......... 4878
SMM.US.x.H9             G--------G-----A-----CA-CC--A-G-TAGGG-CA...--G--C--G-----A-----G--G-----T--A--T--C-AA--C------GG-GGA-AA-A-TTG---AG-G--TCCCACTTGGAGG-TACC-G---GGCTAGAGAAGTGGCATAG.......... 4883
SMM.US.x.PGM53          G--------G-----A-----CA-CT--A-G-TAGGG-CA...--G--C--G-----A--------G-----T--A--T--C-AA---------GG-GGA-AA-A-TTG---AG-G--CCCCACTTGGA-G-TACC-G---GGCTAGAGAAGTGGCATAG.......... 5327
SMM.US.x.SIVsmH635F_L3  G--------G-----A-----CA-CC--A-G-TAGGG-CA...--G--T--G-----A-----G--G-----T--A--T--C-AA--C------GG-GGA-AA-A-TTG---AG-G--TCCCACTTGGAGG-TACC-GA--GGCTAGAGAGGTGGCATAG.......... 5413
SMM.SL.92.SL92B         ---------A--------T--T-----TA-G-TAGGG-CA...--A--------G--A---------------A-----C-AA---------GG-AG-CAA-AGATG-G-AGCA---CCAGT-TGGAGG-TC-GT--A-GA-TAGTGGAGCGCTGGCATAG....... 4844
STM.US.x.STM            G--------G-----A-----CA-CG--A-G-TAGGG-CA...-----T--------A-----G-----G--T--A-----C-AA---------GG-GGA-AA-AG-TG---AG-G--TCCCACCTGGAGG-TACCAG---GGCTGGAGAGGTGGCATAG.......... 5057
MNE.US.x.MNE027         G--------G-----A-----CA-CT--A-G-TAGGG-CA...-----T--G-----A--C--------G--T-----T--C-AA---------GG-GGA-AA-AG-TG---AGCA--TCCCAC-TGGAGG-TACC-GA--GGCTAGAGAGGTGGCATAG.......... 4878
DRL.x.x.FAO             ---------G-----A--G--G--C--CA-G--AGG--C-...---T-G------A-A--T------------------------AA-------------A-TG---ATA--C--ACCAACA---AG---T--......................................... 4648
MND_2.x.x.5440          T--------A-----C--T--G-----CA----AGGAGT-...--TT----G--G--T--T--------------A--A----A---C----------ACAG---ATA-CA--CCCCACA---G---ATAG........................................ 4642
MND_2.CM.98.CM16        T--------A-----A--G--G-----CA----GGGG-CA...---T-G--G-----A-----G-----------A-----C--A------------ACAG---ATA-----CCCAAC----A-GC-T--T--................................. 5085
MND_2.GA.x.M14          T--------A-----A-----G-----CA----AGGA-CA...---T-------T-----C-----------T--A--A--C-----------------ACAG---ATA--A--CCCCACA---AGGA---------GACA-C--GA.......................... 5024
OLC.CI.97.97CI12        G--------A-----A-----CC-GC-TA----AGGAGA-...A-GTATTTCTCT-----TCTG--G--------A--AG-G-A-CC-------GTGG-T-GCAGT-CCA---TA...............T--..................................... 4937
GRV.ET.x.GRI_677        ---------A-----C--G--G-----TA----GGGGGAA...-----C--G-----C--C--G-----G--T-----T--C-AA---------G-GAGA-AAA--ATG---AG-GAG-GTAGT-TGGAGGGT-TCAGA--GGCAAATAAGCAGATGGAGGGGGATAGTG 5294
VER.KE.x.TYO1_patent    ---------G-----A-----G--CC-CA-G...G-CG-AAGT---C----G--T--A--------G-----T--A--T----A--------A-CCCA-ACAAAG--TG-G-A--GAG-GT-ACGTGGA-GGTACCAG--GA-CTGATAACTAAATGGCAGGG...AATA 5342
VER.KE.x.9063           ---------G-----T-----G--C--CA--...G-AG-AGAA--GC-G---A-T--C-------------------A-----C-AA-------A-CC-AGA-AAAG-ATG-G----GAG-GTAAC-TGGA-GGTCTCAG--GACCTGATAACCAAATGGCAAGG...AATA 5345
VER.KE.x.AGM155         ---------A-----C--G-----TC-CA-G...G-AG--GAA--AC-G--G-----T--G-----G--------A-----A-AA--C----AGCC-AGA-AAA--TTG-G----GAGACTCACCTGGA-GGT-CA-GA-GAAGTGATCACCAAATGGCAGGG...GATA 5337
VER.DE.x.AGM3           T--------G-----T-----G--C--CA--...GGAG--GTG--AT-------A-TAC---------G--------A--T--A-A---------A-CC-AGA-AAAG-ATG-G----GAGAGTAACTTGGA-GGT-CC-GA-GAGCTGATAACTAAATGGCAAGG...GATA 4840
TAN.UG.x.TAN1           ------G--A-----A--G-----C---A----AGG-GAG...--GT-G--------A-----G-----G-----A--T--C-AA--A------G-GAGA-AAA-T-TG-G---CAAGACTAAC-TGGAGGGTT-G.................................... 5286
SAB.SN.x.SAB1C          ---------A-----A--------T--------AC-GG--...--AT-G---AC-A-C---------------C--A--T--A-AA----------------.......---TTG---AG-CAG-CCC-TCTGGA-GGTA-C-G-CGGACAGCAGGAGAGGTGGACTAG....... 5506
MND_1.GA.x.MNDGB1       G--------A---------------TT--A--T--C-AGAA...--G-----GA----A--T---------GTGT--A--A--A-AA---------G-G...-GT-G-AAGA--AG-CAG--TAACTTGGA-AGT-TCT--.............................. 4774
DEB.CM.99.CM40          G--------A-----T-----G--GTGCA-GACAG-GGTAGGA--T--C------G--C--T-----G---------AG-A--C-AA--A----------G-C----ATA--A-G-T-AACACACA--A-TCCA-T--AT-A................................. 4899
DEB.CM.99.CM5           G--------A-----A--G--G--GTGTA--ACAG-GGTAGGA---------G------A---------------C---G-G--A-AA-------------GGAG---ATA--A-GAC-GACACACA--A-CCCAGT--AT-A................................. 4893
TAL.CM.01.8023          G--------A-----A-----C--CC--AGGACAG-CGAAGGA--GG---TTAC---T--C-----G--G--T--A--A--C-AACCC------C--GCT----GTAACA-GAC-GACT---A-G--TCCA----AC-----GCAGAGATGGGTAGGGACAATTAG.... 4584
TAL.CM.00.266           T--------A-----A--G--T--CT-G-G-ACAG-CGAAGGA--GG---TCAC---C--T-----G--G-----A--A----A-CCC------C--GCTT---GTAACA-GAT-GACC---A----TCCA----AC-----GCAGAGATGGGTAGGGACAATTAG.... 5086
SUN.GA.98.L14           G--------G-----A--------C--CA--C-G...G--GAGA----CCTG-----A--T--G-----------AC-AG-A-A---C------GG-G-ATCCT-CTCA-TG--GATG----G-T-A.......................AAGAGAATGGGAAGATAGAA 5458
WRC.CI.97.97CI14        T--------A-----A---T---------G--ACACCGGAAGGT...----T-AC---T--G---------------C-T--A-AA-T--GGAATGG-G-AGGAATG-A-AGGAG--CCTCT-AT-AG-ATT--............................ 5381
WRC.CI.98.98CI04        T--------A-----A--CT-----G-GG--ACCCCAGAAGGT...----T-AC---T--T-A---G--------A--A--A-A--TC-GGAATGG-G-AGGAATG-A-AGGAG--CCTCTAAT-AAGATT--............................ 5396
WRC.GM.05.Pbt_05GM_X02  ------------G--A--TT----TG-G--GAC-CC-GAAGGA--TT--CTTAC---T--C-----------------A-----A-A-AC--GG-ATGG-A-AGGAATG-A-AGCAG-ACCACTAAT......T-A.................................. 4307
LST.CD.88.524           G--------A-----T--T-----GG--A-G--GGGAGAG...A----TTTCTCTA-C--C--G--G--------AT-AG-G-AA---------G-GGGACCAAAG-A-AG-C-GA--AGCATGGA--AC--AT-G.............................. 4361
LST.CD.88.447           G--------A-----T--T-----GG--A----AGGAGAC...AC---CTTCTC---C--C-----------T---C-AG-A-AA---------G-GGGACCAAAG-A-AG---GG---GCCT-CA--ACA-CT-G.............................. 4364
LST.KE.x.lho7           G--------A-----T--T-----GG--A----AGGGGAG...A----CTTCTC---C--C--------------AC-AG-A-AA---------G-G...GGAC--AAG---AG-GAGAGTAGCCTGGATA-CA-CT--.............................. 5448
SYK.KE.x.SYK173         T--------------T--------GG-G---ACAG-GGAGGGA-----TTTT-C---A-----------G-----A--A----CA--CC-----G--AGA---ATA---G-TCCCAC----AG-A-GAC-CCA-GACA----GATAG....................... 5296
SYK.KE.x.KE51           G-----G--A-----A--------G--C--GACACCAGA-CAAC-GG---T--C---A--T-----G--------A-----C--CAG-G-----G--AGAG--ATA---G-AC-CACC---A--T--C-TCCA-ATCAA---GATAA....................... 4948
COL.CM.x.CGU1           G--------------A--G-----C--CAG-A--TC-CAAGGTAC-T-GTTT--GAAA--T----------T--------CC--ACTCC-ATATGG-G-AGAT-TG--CAG---GAA-T--TT-TCC-ATGG-C-AA----AGCCGAAACCGTTAAGGGAATGGATTAA. 4886
```

```
H1B.FR.83.HXB2              .........................................AACATGGAAAAGTTTAGTAAAACACCATATGTATGTTTCAGGGAAAGCTAGGGGATGGTTTTATAGACATCACTATGAAAGCCCTCATCCAAGAATA...AGTTCAGAAGTACACATCCCACTA 5217
Vif                         .__.__.__.__.__.__.__.__.__.__.__.__.__.__T__W__K__S__L__V__K__H__H__M__Y__V__S__G__K__A__R__G__W__F__Y__R__H__H__Y__E__S__P__H__P__R__I__.__S__S__E__V__H__I__P__L_
Pol
H1A1.UG.85.U455             .........................................---------C-------------T--C--------C--CAA-------CAA--T---------------------------TAGA---T-----G--...------------------------ 4663
H1B.US.90.WEAU160           .........................................----------------------------------A-----AA--------A--------C---------------C------A-------------...--------------T--------- 5216
H1C.ET.86.ETH2220           .........................................---------T-----------G--------C--A-----A---G-----AT------G------------T-----C---AGA-------AGG--...----------------------T-- 4609
H1D.CD.84.84ZR085           .........................................C---------------------T--------C--A-----AA--------AA------------------C--------C--------C---A----...----------------------T-- 4739
H1F1.BE.93.VI850            .........................................C---------------------T-T---------------AA------C-AA-------C--------------T-C-----AGG---------G--...------------------------ 4557
H1G.SE.93.SE6165            .........................................-------C-C-----G--------T-----------C---AA---------A--C-------------C--------C----AGG---------G--...------------------------ 4614
H1H.CF.90.056               .........................................C-----------C--------GT--------C--A-----A---------A-----------------T-T---------A----------G---...----------------------T-- 4564
H1J.SE.93.SE7887            .........................................C--------------------GT--------A--------AA---------AAA----C-G-----G-----------C--AA-------A----...------------------------ 4531
H1K.CM.96.MP535             .........................................T--T------------------T--------C--A-----AA--------ACC------A-------------------AGG-------A----...------------------------ 4413
H101_AE.TH.90.CM240         .........................................---------C-------------T--------A-C---AA--------AAAAG-----------------T-----------AG------AGG--...--C----------T--------- 4791
H102_AG.NG.x.IBNG           .........................................---------T-------------T-----------C---AA--------A---T------------------T------TAGG-------A-G--...T------------------------ 4742
H1N.CM.95.YBF30             ...................................................C-----------T-----------G---AAA--G--A-AA-------A-------------T-------CA-A---C----A----...--------------T------G-- 4811
H1O.BE.87.ANT70             CATAA.............................................................---CAGG--TA-A--GA-CGA-AAC---GG--C---------T------TC-AGAA--------G-T...-------AGT---T-T--T---G-- 5272
H1O.CM.91.MVP5180           CATAA..............................................................A--CA-G--TAA---G--CGC-AAC---CG------G-----T------TC-AGGA------A-G-C...-----G-CG--GT-T--T---G-- 5247
CPZ.CD.90.ANT               ...................................................C-----G--------C--A-GG-AGA-TAAAGT-CT--AACC----AAG------------T-------ATGA---C----AG-A-...G-AGA----------T--A------ 4656
CPZ.CM.01.SIVcpzCAM13       ...................................................CC----T---T-T----------AAAA--AA------A-AA-AT-----C--C-----C----------CAT---A---------G-T...TC------G---------T--TT-T 4797
CPZ.CM.05.SIVcpzEK505       ...................................................---G-----G--T--C--------G---AA---G--AC-A--T-----C------------T-------CAGA-A------A----...----------A-----------T-- 4789
CPZ.CM.05.SIVcpzMB66        ...................................................CC-----------T--C--------G--TA-------C--AAAT-----C---------C-------------AA-A------A----...-------------------------G 4763
CPZ.CM.98.CAM3              ...................................................CC--------TGT---------AAAA--AA------A-A--AC-----------G--C--T---------TT-------T-A-G--...----------A----------TA-- 4631
CPZ.GA.88.GAB2              ...........................................------T------------GTTT---------CAAG---AAA--G--C--A---------------G--C--T-----T--T--CA----T-A-G--...-------------C-----A--CA-- 4610
CPZ.CM.98.CAM5              ...................................................CC---------AC-----------AGAA--AA---G--A-AAAAG-----------G-----T-----G--TGA-A----T-A-G-G...-------------T------TT-- 4919
CPZ.CM.05.SIVcpzMT145       ...................................................------G---T-T----------CAGA---AA------A-AAAAT-----C-------------------TG--A-CA-C--T--G--G...GC----C-GA--T-T--T---A-- 4759
CPZ.TZ.00.TAN1              ...................................................CC----T-----T-----C-T-ACAA-CAAATGCTG--AA-AT---AAG------------T--------CTGA-ACA----A--G-...GCAGGG---A-------A--T--- 4861
CPZ.CM.05.SIVcpzLB7         ...................................................---G------T-T-----C------A--CA-------A--A--C------------------T-----G--TGAAA-------A-G--...----------A----T--------- 4772
CPZ.US.85.CPZUS             ...................................................-C-G------T-T--C------AGA---AA---------AA---------C--C--G-----T-----G---A--A----C---G--...-A------------T-----TA-- 5277
CPZ.GA.88.GAB1              ...................................................------T---T-T-----T---AGG---AAA--G--A--A--T---------C----------T-----TCA----A-------A-G-T...GC---T---A-------A--TT-C 5278
MON.CM.99.L1                ...................................................A-GCACT---TGG--C--ACTCACAGG-A-T...--ACCCTTTGTC...---GT---------CC-GCTG-ACA---AG---T-CTCACAGAACA--A-TA-AC----CT-- 4926
MON.NG.x.NG1                ...................................................AGC-ACC---T-TA-C--A---ACAGGGAAA...-TGCCATTTGT-...---GT----------C-GCTTA----C-AGC-CT-TTCGCAA-GCA-GA-TA-GC-T--TT-- 3519
GSN.CM.99.CN166             ...................................................AGCTACT--------C--A-GGTCAGGCAAA...A-ACCATTTGTA...---GTC---------CC--TTG-AA--C-A----T-TACACAAAACA-GA-T-GAT-AG--A-G 4957
GSN.CM.99.CN71              ...................................................AGC-ACT-----T-----A-GGTCAGG-AAA...A--CCATTTGT-...----TC--C-------C--CTA-AA----A---GT-TACACAAAACA--A-T-GAT-AG--T-- 4939
MUS_1.CM.01.1085            ...................................................GGC-AC----T--T-C--CCT-TCAGGGA-A...--AGCCTTTACA...---GTG----------C--TTGTAC----AG--GT-CTCCCAGAGCA--A-CAGA--A---T-G 4913
MUS_1.CM.01.CM1239          ...................................................GGCTAC------T-----C-T-TCAGG-AAA...A-ACCATTTGT-...---GT---C-----CC-GGTA-AA--C-A----T-TTCCCAAAACA-GA-CA-GC-AG--T-G 4903
MUS_2.CM.01.CM1246          ...................................................A-GCACT-----TT----C-TCTCAGGGAAA...---CCATTTACA...---GTC--C-------C--TTG-AG----AG---T-CACCCAGAACA--A--A-A--A--TT-G 5011
MUS_2.CM.01.CM2500          ...................................................A-GTACC--------C--C-TATCAGGGAAA...--ACCATTC-CA...--C-TC-----T--CC--GTG-AG----AG---T-CACTCAAAATC-GA-TA-G--A--CT-G 4964
RCM.NG.x.NG411              ...................................GCAGTGGCAT-CCTT-GTAAAGTATC-T--GT---A-GGAAAGAAG-AAGCTAAAGATT-GGA--A-GTACCC---TTTA-A-T-CCTTGGGGGTGGT-GTCT...C----------C--------CT-G 4720
RCM.GA.x.GAB1               .................................ACA..ATGGCAC-CTTTGGTAAA-TATC-T--GT-C-AAGGAAAGAA--CAGC-AAAGA-T-GGA--A-GTGCCT--CTTTA-A-T-CCATGGGGATGGT-GTCG...CA---------T-----A--TT-- 4708
DEN.CD.x.CD1                ...................GATTGGGGAAACCAAATTAGAACATTGGAT--GC-C---GAA-TCTATAATATACA-AACAAAT-A-TT-AAAG--AC-GT-A--ATACC----TTTAGAATC-CATGGG-ATGGTAT-CC...-TGA--A--A-------------G 5134

MAC.US.x.239                ................................................CCTCA--AA-T-T-TGA-ATATA-AAC-AA--ATCT-CAA-A--TT--C-A-GTGCCC---TTTA-G-TCG-ATGGGCATGGT-G-CC...T-CAGCAG----AT-T-------- 5522
Pol
Vif                         .__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__L__I__K__Y__L__K__Y__K__T__K__D__L__Q__K__V__C__Y__V__P__H__F__K__V__G__W__A__W__W__T__.__C__S__R__V__I__F__P__L_
H2A.GW.x.ALI                ................................................CCT-G-CAAGT-C-TAA-GTACAGAACAAA--ATCT--AA-A--TG--C-A-GT-CCC--C--TA-G-TGG-ATGGGCATGGT-G-CT...T-CAGCAGG---ATAT-----T-- 5598
H2A.DE.x.BEN                ................................................CCT-G-CAAGT-C-TGA-ATACAGAACAAA--ACCT--AGGA--TGC-C-A-GT-CCC--C---A-G-TGG-ATGGGCATGGT-G-CT...T-CAGCAGG---ATAT-------- 5605
H2A.SN.x.ST                 ................................................CCT-A-CAAGT-T-TAA-ATACAGAACAGG--ATCT--AG-A--TG--C-A-GT-CCC--C--TA-G-TGG-ATGGGCGTGGT-G-CT...T-CAGCAGG---ATAT-----T-- 5050
H2B.GH.86.D205              ................................................TCTGA-TAAGT-T-TTA-GTATAGAACAGG--A-TTGCAACA--TC-CT-A-GTCCCT--C---A-G-T-G-ATGGGC-TGGT-G-CT...T-CAGTAG-A--ATAT-T--C--- 5578
H2B.CI.x.EHO                ................................................TCTAA-TAA-T-C-TGA-ATATAGGACAAA--ACTTGCAACA--TC-CT-A-GT-CCT--C--TA-G-T-G-ATGGGC-TGGT-G-CT...T-CAGTAG----ATAT-T--C--G 5576
H2G.CI.92.ABT96             ................................................TCTGA--AA-TTCTTAA-GT-TA-AACAAA--ACCT-CAG-A--CGGTT-A-GT-CCT------A-A-TTG-ATGGGC-TGGT-G-CT...T-CAGCAGG---ATAT-----T-- 4932
H2U.FR.96.12034             ................................................CCTGA-CAA-C-C-TGA.GTATA-CACAAA--A-CT-CAG-AA-TG--T-A-GTGCCC------A-G-T-G-ATGGGC-TGGT-G-CT...T-CAGCAGG---ATAT-T--T--- 5083
MAC.US.x.EMBL_3             ................................................CCTCA--AA-T-T-TGA-ATATA-AAC-AA--ATCT-CAA-A--TT--C-A-GTGCCC---TTTA-G-TCG-ATGGGCATGGT-G-CC...T-CAGCAG----AT-T----C--- 4990
SMM.US.x.H9                 ................................................CCTCA-CAA-C-C-TGA-ATATA-CAC-AA--ACCT-CAG-A--CT--T-A-GTACCC-----TA-A-TTG-ATGGGCATGGT-G-CT...T-CAGCAG----AT-T----GT-- 4995
SMM.US.x.PGM53              ................................................CCTCA-CAA-T-CTTGA-AT-TAGAAC-AAG-ACCT-CAG-A--CT--T-A-GTACCC-----TA-A-TTG-ATGGGCATGGT-G-CT...T-CAGCAG----AT-T-------- 5439
SMM.US.x.SIVsmH635F_L3      ................................................CCTCA-CAA-C-C-TGA-ATATA-CAC-AAG-ACCT-CAG-T--CT--C-A-GTGCCC-----TA-A-TTG-ATGGGCATGGT-G-CT...T-CAGCAG----ATTT-----T-- 5525
SMM.SL.92.SL92B             ................................................CT-CA-TAA-TTC--TA-GTATA-AACAAG--A-CT--AG-AA-CC--C-A-GTGCC-------A-A-T-G-ATGGGCATGGTAC-C-...GCC--TAGG--GATTT----CT-G 4956
STM.US.x.STM                ................................................CCTCA--AA-C-TTTAA-GTATA-CAC-AA--A-CT-AGC-A--C---C-A-GTGCCC--C--TA-G-TTG-GTGGGCATGGT-G-CT...T-CAGTAG---GATTT-T--CT-G 5169
MNE.US.x.MNE027             ................................................CCTCA--AA-T-T-TGA-ATATA-AAC-AA--ATCT-CAA-A--TT--C-A-GTGCCC-----TA-G-TCG-ATGGGCATGGT-G-CC...T-CAGCAG----AT-T-------- 4990
DRL.x.x.FAO                 ......................GAGGTACATGGTAGAAAGGTGG-AC-CCCTGT-CAA-TATC-TA-G----A-GGA-AAAAGTATTT--AA--AT-GCAT-A-GCACCT--CTT-C-ATGCTCA...GG-TGGT-G-CT...CA-AGTC-GTGGAC------CT-C 4787
MND_2.x.x.5440              ...................ACAGGACGTCAAGTTGAGAGATTGG-AT-CA-T-GTAAA-TATC-TA-GT---AAGGA-AGAA-CATTT--AA-ATT-GCAA--A--CCCC---TTTC-ATGCTCA...GGATGGT-G-C-...CAC--TC--AAGA-G-----CT-T 4784
MND_2.CM.98.CM16            .........................ACATGCAGTAGAAAGATGG-AT-CACTGGTAAA-TACC-CA-AT---AAGGA-AAAA-CATTTG-AC-AAT-GGAT--AA-GGC----TTTC-ATGCTCA...GGATGGT-G-C-...CAC---C--AA-T-T--T--C--- 5221
MND_2.GA.x.M14              ...................................AAAATGG-AT-CATT-GTCAA-TACC-CA-AT---A-G-A-AAAA-CACTT--A--A-T-GGAAC-A-TCCAC--CTTTC-ATGCTCA...GGATGGT-G-CG...CAC--TC--AAGATT-----CT-C 5148
OLC.CI.97.97CI12            .........................GCTAAATAATGGGCAGGT-GA-CATGTG-AA--CA-GGCTA-GTTAGCC--CAAA-TTT--GG-TACCATT-GATA-GGC-A-ATTG-T--A-GT-CCCAAGGG-ATGGT-G-C-...CA-GG-A--A-TT-T--T---G-- 5076
GRV.ET.x.GRI_677            ACTTACAAGATCAGGAATAA.........................................................AC--TTGCCTT-GGAA-ACAGACAT---TGGC-G-TGCAATGG--GTTTT-G-CC...TACAGCC-GT-CATT-----CT-- 5393
VER.KE.x.TYO1_patent        GTCAGATATTGGATGAGACA.......................................................-A-A--TTTG-AAT-GAAC-A-----TG---T--C-AATT-CATGGGC-TGGTAC-CC...-TGAGTAG-TATGTA--A---A-- 5444
VER.KE.x.9063               GTCAAATATTGGATGACTAA.......................................................-A----CCT--AAT-GCAA-ACATG-TG--CT--C-AATT-ATTGGGC-TGGTAT-CT...-TG-GCAG-TACAT-------A-- 5447
VER.KE.x.AGM155             GTTAGATATTGGATGAATAA.......................................................-A----TCTG-AAT-GGAA-ACA-A-TG---T-TC-AATC-CTTGGGCATGGTAC-CT...-TGAGCAG-TATGTA--A--C--C 5439
VER.DE.x.AGM3               GTGAGGTACTGGATGAGGAC.......................................................TA-A---TTAGACT-GAAA-A-CGA-TG--CT--C-AATT-CATGGGCATGGTAC-C-...-TGAGTAG-TATG-G--A--C--- 4942
TAN.UG.x.TAN1               ..............................TGGAGAGCATAT-GACAAATGG-AGGGCA--GTGA-AT-CCACATGAGGAACC--TT-CAGGATT-GACA-A-CT--TG--CT--C-ATGTG-ATGGGC-TGGTAC-C-...T--AGTAG-T-C-T------CT-- 5419
SAB.SN.x.SAB1C              ................................................CCTAG-GAA-T-C---ATGCATGTGTCCAA-CA-TGT-T-CATT-GC---A-ACCCCT---ACTA-AAT---ATGGA--TGGTACTCC...TA-CA----TGGGT---A--C--G 5618
MND_1.GA.x.MNDGB1           ............TCAGAGAATAGAAAAGTGGCACTGGTTAGTAAG--G-CA--TGGCA-GG-CC-CTGCAA---ATG-G-AAGG-T-TTGGTGGCT-TATCCTCA--T--TGGC-T-TA-----T-GTA-AC-TGC--T-A-...GTAGTGATTA--ATA-ATAGGGAC 4928
DEB.CM.99.CM40              ................................ATTAGATAGGTGG--C-GTGC-A--AA--G-G-AA-ATACA-GACCAA--AATT--A--CT-T-CAA-GGATACAC---TTTG-ATT---GAACTCATATTAT-C-...CAAA--A--A--GTAT----CA-C 5029
DEB.CM.99.CM5               ................................ATTAGATAAGTGG--T-GTGCCA--AA--G-GGAA-GTATA-GACCAA--AATT-CAAGCAAT-AA--GGATACAC---T-TG-ATT---AAGGGCATTCTAC-CT...CAGA--A--A--GT-T----CG-C 5023
TAL.CM.01.8023              .....................................................GTATC-CATGT-C-AAACCAAG...TACTT-CAA-A-TT-ATA-GGGTCCAT---TT-C-ATT---AATGGCCT-TTTC-CG...CAGG-CA--TGGAT-C----C--- 4687
TAL.CM.00.266               .....................................................GT-TC-CATA----AAACCAAG...TACTT-CAA-A-TTGATT-GGGT-CAT--CTTTC-ATTG--A-ACGCCT-CTAT-C-...CAAG--A-GTGGATAC----GT-- 5189
SUN.GA.98.L14               GACAGTATAAGATAGTTAGAATAGTATGGCTTATAGATAGAATTGCTGT-GA----TT-----ATTT-AGAAGA---C--AGGGA-ACA----A-GATT-GGTA-C-ATGTATGGAACAGG-ACTG-ATGGG-GTGGTAC-C-...TACAATA--A-TATT--T---G-- 5625
WRC.CI.97.97CI14            ................................TAAAGGTAA-TTCA---G-T-GACCTATT-TATAGCA-ATCTAA--AA--CAGGGCAGGCAAAT-TTC-A-GGG--TC-TGGC---C-GCAG-AG-ATGCTTT-CT...--GCATAT----A-A--A---A-- 5511
WRC.CI.98.98CI04            ................................TAAAGGAAGGTTTAAA-G-T-GGC-TATTGGATAGC-CATCTAA-GCA-AAT---AAGGCAA-T-TT--A-GGG--AC-TGGC--ACTCAGG-AG-ATGCTTT-CT...--A-ACCT-T-GAGA--A--CT-- 5526
WRC.GM.05.Pbt_05GM_X02      ...........................AATCTCAAAAGGAAG-TTTA-A-G-T-GAGCTATTGGATAGCAGA-CTAA-ACA--AA---CAAGCAA---TCC-G-T-G-CTCATGGC---T-CATG-AG-ATGTTTC-CT...--ACATAT----AGA--T--T--T 4443
LST.CD.88.524               ...................................GAGCTT-GT-GA----TT-C---AT-TCA-AAGG-ATC--TCAGA-A-T---AA-GATT-GGTTAG--TACATGG-ACAGGA-T-G-ATGGG-GTTTTAT-CT...TACAATA--T---TA--T---T-- 4488
LST.CD.88.447               ...................................GAGCCT-GT-GA---GTTCC---AT-TCAGGAGAGCAC--AAAGAGACTC--AA-GATT-GGTAAGC-TACATGGAACAGGG-T-G--TGGG-ATTTTAT-CT...TA-AATA-TA--GTG--A--TT-- 4491
LST.KE.x.lho7               ...................................ATCCTT--T-GAA--GTT-----AC-T-GT-AGGT-AA-CACAGAG-CA----AA-ATT-GGTAGGC-TACATGGAACAGG--T-G-TTGGG-ATTCTAT-CC...TA-CATA--T--GTG--A--T--- 5575
SYK.KE.x.SYK173             ........................................-TTGCA-C-T-TAA--AA--C---TA-GTATA-AAGCAA--AATTG-AG-AA-CCACA-A-A-ACAT--CT--C-AAT-GAATGG--ATGGTAC-CT...TA--GCC--TGGACA--A---G-- 5417
SYK.KE.x.KE51               ........................................--TGCA-C-T-TA---AAGGT----A-GTATA-AAGCAA-CAATT--AG-AA-CCACA-A-A-ACAT---T-TC-GATT-CATGGG-GTGGT-G-CT...C-AG--C--TGGG-A--A---G-- 5069
COL.CM.x.CGU1               ..................................................................-A-GCAG-TTCCAG--TGCCAATAT-AT-A--TAG--A-AACAG-GTC-CC-GAATGGTGG-C-GA-...G--AA-AT--AGTTTTA-TT-GGG 4977
```

```
H1B.FR.83.HXB2           GGG......GATGCTAGATTGGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGAGACTGGCAT........................... 5280
Vif                       G  .  .  D  A  R  L  V  I  T  T  Y  W  G  L  H  T  G  E  R  D  W  H
H1A1.UG.85.U455          ---......---A--------A---G---G----------------------------A---------C........................... 4726
H1B.US.90.WEAU160        ---......---A-G--A---A-----------------------------------------------........................... 5279
H1C.ET.86.ETH2220        ---......---G--------AA------A-------------T----A-------------T------........................... 4672
H1D.CD.84.84ZR085        --A.........-------C-----G-------------------------------------A------........................... 4802
H1F1.BE.93.VI850         -A-......---A-T--A---A------------------G-----CC------------A------........................... 4620
H1G.SE.93.SE6165         --A......-------C-C-A---G-----------------G-----------------A---------A........................... 4677
H1H.CF.90.056            --A......--A-----G--A--C-----C-----C---------A---------------A------........................... 4627
H1J.SE.93.SE7887         --A......--G----T---A--------------------------G-------------T-----C........................... 4594
H1K.CM.96.MP535          ---...........AGAG--A---G---------------------T-----------G---------........................... 4476
H101_AE.TH.90.CM240      --A......--G--------A-------G------------------A----------AG--------C........................... 4854
H102_AG.NG.x.IBNG        ---...........--------G---A---G---G-----------------------C---------C........................... 4805
H1N.CM.95.YBF30          --T.....C-G--A-----A--G-C-GTC--T--------G--AACA---------CAGTCT------........................... 4874
H1O.BE.87.ANT70          ---......-TG--CCATG-A---G----C----------AT--ATGC----G--------TGAA---........................... 5335
H1O.CM.91.MVP5180        -CA......--A--GATA-A--GG-C--C-----------AT-AATGC----G--------GGAA--C........................... 5310
CPZ.CD.90.ANT            CCAACATTG---AAA-A-------G-G---GTG-TC-----AT-A--ATGT------C-CCCA-----C........................... 4725
CPZ.CM.01.SIVcpzCAM13    AAA...GATA--T-A-A---AA----C-----C-------C-T-AACC-------------CG------........................... 4863
CPZ.CM.05.SIVcpzEK505    --A......--A--A-A---A------GTT--C--------AT-AATGC----------GCCA------........................... 4852
CPZ.CM.05.SIVcpzMB66     --A......--A---A---A--------------------A--CACA-------------C-------........................... 4826
CPZ.CM.98.CAM3           -AA.....A-G--A-A---A--GG-------C--------AT-AATGC-----C-GCAT-C-------........................... 4694
CPZ.GA.88.GAB2           A-A.........-----A-A-A-A--GG-CTGT-----C---AA-T-AATGC--------CA-CCA------........................... 4673
CPZ.CM.98.CAM5           -AT.....ACA--A-A---A--GG-C--------------A---ACA-----GC--CAT-CT-----C........................... 4982
CPZ.CM.05.SIVcpzMT145    AAT.....C-G----AGC-AA-TG-C--------------A--AATGC-------------T------........................... 4822
CPZ.TZ.00.TAN1           ACA...GAAAGAT-A-A---A--GG-TTT-CAT-----------AGCCTGT---------CCA------CTAGGTCATGGC........................... 4939
CPZ.CM.05.SIVcpzLB7      ---.........-G--C-A---A-----GT---T--------A--TACAC--------C-G--------C........................... 4835
CPZ.US.85.CPZUS          -AT.....A-A--A-A-A-A--G--T-T---G-------GT--A-A---------CA--C------C........................... 5340
CPZ.GA.88.GAB1           A-A...GATT--T---A---AA--G-G-----C-------CCT-AAGCC----------G-C------........................... 5344
MON.CM.99.L1             -ACCTAGGAATCA--CA-GAA-G-GAGT--TGGGCAACATA-T--G-G...AT-AC--TTT-T---G-CTTAACAAACGTGGGACCTAGTGCACTCTCACCAACAGTGTACAACAAACAGGCCTACACAATAGAATGGGTTTATTGGATGCAAAAGGTGGAACGG..... 5088
MON.NG.x.NG1             -A-...ATG-G-ATC-C-CAARAGGRGGA-CA--GGGCAACATACATRGAGATCACG-TCCTT---G-CAYTACCAATGTAGGGCCCGCGTCACTATCTCCATCAGTCAAATGGAAGCAAGCATACATCATAGAATGGGTATATTGGAGGCAGCCCAGGGTGCCACAGCA 3686
GSN.CM.99.CN166          -ATACTAGGA-G-TAG-GGAA-A-...GT-GA-GCAACATA-A-AG-G-T-ACTATCCT-TGGGACACAACCAGTCATGGACCAGCTAGCTTATCCCGGTCCACCTATTACCAACAAGCAGTCATAATAGAATGGGTATACAAC........................ 5098
GSN.CM.99.CN71           -ATACTAGGC-G-TAG--GA--A-...GT-GA-GCCACATA-A-TG-G-T-ACCATCCT-TGGGACACAACTAGTCATGGACCAGCCAGTCTATCCCGATCCACCTACTACCAACAGGCAGTCACAATAGAGTGGGTTTACAAC........................ 5080
MUS_1.CM.01.1085         ACACAAAATACCCGGGA--CA-GTGC-GT-GA-AG-ACTTACA-AG-ACTGAC--TCCTGTGGGATGTGACCAATGTGGGACCAGCAAGTTGCTCTAAAAGCACATGGTGGAAGCAATCCTTTATCCTAGAGTGGGTATATCTA........................ 5057
MUS_1.CM.01.CM1239       ACACAAAACACCTTAGA--CA-G-GAG---GA-GTCACCTTCA-AG-G-T-AC--TCCT-TGGGATGTGACTAATGTAGGACCAGCAAGCCTATCCAAAAGTACCTATTGGAAGCAGTCATACATCCTGGAATGGGTATACATG........................ 5047
MUS_2.CM.01.CM1246       A-CATAAACAGAATAGA-GGA-GCCC----GAGAC-ACATACA-AG-G-T-ACCAT-TTGTTAGATGTCACAAACGTAGGGCCAGCCAGTTTATCAAGGTCCACCTATTGGCAGCAATCCTATATCCTGAAGTGGAGATACATC........................ 5155
MUS_2.CM.01.CM2500       ACAATAAAGA-GAGAGAGGAA-AC-A-GT-GA--C-ACATACA-A--A-T-AC-ATTCT-TGGGATGT-ACCAATGTAGGGCCTGCCAGCCTCTCAAAGTCTACTTACTGGCAGCAGTCATATATATTGGAGTGGGTTTATGTT........................ 5108
RCM.NG.x.NG411           --A............-ATAACACT-AG-T--AGGT-ACAACATATTGG-GTCTCAC--T---AAA-GGATGGTTAGGAACCTATGGA........................... 4795
RCM.GA.x.GAB1            -A-............GA-GGATC--AGTT--A-ATAACCACCTATTGG-ATTTGAC-GT---AAA-GGATGGCTAGGGACATATGGG........................... 4783
DEN.CD.x.CD1             A-A...............GC--CTTAC-TGGA-GTAAA-CTCTATTGGCATTTGACCCC---AAAAGGATATCTGTCACAAGTAGCA........................... 5206

MAC.US.x.239             CA-............GA-GGAAGCCATTT-GA-GTACAA--GTATTGGCATTTGAC-CC---AAAAGGGTGGCTCAGTACTTATGCA........................... 5597
Vif                       Q  .  .  .  .  E  G  S  H  L  E  V  Q  G  Y  W  H  L  T  P  E  K  G  W  L  S  T  Y  A
H2A.GW.x.ALI             CAA............G--AGAAGTCATCT-GAGATACA--CATATTGG-ACCT-AC-CC---AAAAGGATGGCTCTCCTCTTATGCA........................... 5673
H2A.DE.x.BEN             CAA............G--AAAAGTCATCT-GA-ATACA--CATATTGG-ACCT-AC-CC---AAAAGGATGGCTCTCCTCTCATGCA........................... 5680
H2A.SN.x.ST              AAA............G--GAAAGTCATCTGGAGATACA--CATACTGG-ACCT-AC-CC---AAAAGGATGGCTCTCCTCCTATTCA........................... 5125
H2B.GH.86.D205           AAC............-A-GGA-C-TGGCT-GA-GTCCAA--ATATTGG-ACCT-ACCCC---AA--GGATTCTTGAGCTCCTATGCT........................... 5653
H2B.CI.x.EHO             AAA............GA-GGA-C-CATCT-GA-GTCCAA--ATACTGG-ACCTGACCCC---AA--GGATTCTTGAGTTCCTATGCT........................... 5651
H2G.CI.92.ABT96          ACA............-A-GAA-C-CATTTGGA-ATACAA--GTATTGG-ACTTGACCCCT--AAA-GGATGGTTAAGCCAGTATGCA........................... 5007
H2U.FR.96.12034          AAA............-A-GGATC-CATTTGGAGGTGCAA-CCTATTGG-ATTTGAC-CC---GAA-GGATGGCTATCCTCTTATGCA........................... 5158
MAC.US.x.EMBL_3          CA-............GA-GGAAGCCATTT-GA-GTACAA--GTATTGGCATTTGAC-CC---AA-AGGGTGGCCCAGTACTTATGCA........................... 5065
SMM.US.x.H9              A--............GATGA--CTCATTTGGA-GTACAA--ATATTGG-ATTTGAC-CC---GRAAGGGTGGCTCAGTACTTATGCA........................... 5070
SMM.US.x.PGM53           A-A............GATGAATCTCATTTGGA-GTACAA--ATATTGG-ATTTG-C-CC---AAAAGGATGGCTCAGTACTTATGCA........................... 5514
SMM.US.x.SIVsmH635F_L3   A--............GATAA-ACTCATTTGGA-GTACAA--ATATTGG-ATTTGAC-CC---AAAAGGATGGCTCAGTACTCATGCA........................... 5600
SMM.SL.92.SL92B          -A-............GAGGGAAGTCACTTGGAGGTACAA-TGTACTGG-ATCTGACCCC---AAAAGGATGGTTATCAAGCTATGCA........................... 5031
STM.US.x.STM             CAA............G--GAA-C-CACCT-GA-GT-CAA--GTATTGG-ATTT-AC-CC---AAAAGGATGGCTCAGTGAATATGCT........................... 5244
MNE.US.x.MNE027          CAA............GA-GAAAGCCAGTT-GA-GTACAA--GTATTGG-ATTT-AC-CC---AA-AGGGTGGCTCAGTACTTATGCA........................... 5065
DRL.x.x.FAO              AAA............GACAAAAGT-A--T-GT-GTAACA-TC---TGG-ACCTCACCCC----AAAGGGTGGTTAAGTAAATATGCCATA........................... 4865
MND_2.x.x.5440           AA-............GATGGTTC--AG-TT-TCATCACC-TCT--TGG-ACTTGACCCC---AAAAGGGTGGCTGTCTCAATACGCC........................... 4859
MND_2.CM.98.CM16         -AA............GA-GAT-A-CAG-TC-T-ATCACCATC---TGG-ACTT-ACCCC---AAAAGGATGGTTATCAACTTATGCC........................... 5296
MND_2.GA.x.M14           AAA............GATGGCTC--AG-TT-TTAT-ACT-CC--ATGG-ACTTGACCCC---AAAAGGATGGTTATCTCAATATGCA........................... 5223
OLC.CI.97.97CI12         CCA...............GG--GTGAG-T-GA-AT-ACACTCTACTGGCATATTAG-CAT-GGAAAGGATTACAAAATTCAGGTGGGGTG........................... 5151
GRV.ET.x.GRI_677         TCA............-A-GAT-ATTAC-T-GA-GTGAATAT-TAT--C-ACCTCACCCC---AA-AGGATGGCTCTCAAGTCATGGA........................... 5468
VER.KE.x.TYO1_patent     --A............-A-CAT-GGGA--T-TGTGTAGACCTATATTGGCATTT-AC-CC---GCAAGGATGGCTATCCACATATGCA........................... 5519
VER.KE.x.9063            --A............TT-GAA-G-GA--T-CAGATAGACT-CTATTGGCATTT-AC-CC---AA-AGGCTGGCTCTCTACTTACGCC........................... 5522
VER.KE.x.AGM155          CCA............G--AGT-G-GA--TCCATGTGGATATCTATTGGCATTT--CTCC-A-ACAAGGATGGCTCTCAACTTATGCA........................... 5514
VER.DE.x.AGM3            ---............CA-CAT-G--GT-T-CATGTAGATCTATATTGGCATCTGAC-CC---AAA-GGATGGCTATCAACATATGCT........................... 5017
TAN.UG.x.TAN1            --A............G--GAA-G--A--T-GT-GTAGATT--TACTGGCACCT-ACGCC---ACA-GGATGGCTCTCTACCTATGCA........................... 5494
SAB.SN.x.SAB1C           AA-............GATGGA-CTC---T--A-GT-ACTAA-TATTGGCATTT-ACCCC---GAAAGGATGGTTAGAAACCTATGCA........................... 5693
MND_1.GA.x.MNDGB1        ........................A---G-TT--T-GT-A-AA-CTATTGGCATTTGC---T---GGTAGGATGCTTAAGTACTTATGCAGTA........................... 4997
DEB.CM.99.CM40           ACA............-A-GAA-GGGG-GT-GT-GTAGAATTA--ATGGTGCTTGACCCC---GAAAGGATGGCTACCCACTATGGCA........................... 5104
DEB.CM.99.CM5            ACA............GA-GAA-G--A-GT-GT-GTGGAAATAT-CTGG-GCCTG-CTCC---AAAAGGGTGGATGCCCTCTATGGCG........................... 5098
TAL.CM.01.8023           -AAGAAAAA--GAA----G-A-AGTAC-TT--CATCACTCAGTATTGGTGCTT--C-CC----AA-GGATACACATCAGACTGGGAT........................... 4774
TAL.CM.00.266            -AAGAAAAG--CA-C---G-A-A-TATTT---CATAACACAGTATTGGTGTTT--C-CC----CATGGATATACGTCCGATTGGGAC........................... 5276
SUN.GA.98.L14            ACT...............-AT-GG-C-GT-GT-GTAA--ATCTATGGACACCTCAC-CC--CAAAAGGGTTGATTAACCAGTGGGGA........................... 5697
WRC.CI.97.97CI14         CA-............G--CACTGGGAG-TGCA-AT-A-AT-CTATTGG-ATTTGTC-CC---AAA-GGATGGATGGACACCATGGGATTT........................... 5589
WRC.CI.98.98CI04         CAA............G-GCATTGGGA-TT-CA-ATCA--T-CTATTGG-ATCT--C-CCT--AAA-GGATGGTTAGACACCATGGGGTTT........................... 5604
WRC.GM.05.Pbt_05GM_X02   CAA............G-GCATTGGGA--TGCAGATAA--T-CTATTGG-ATTT--CTCC---AATAGGATGGATAGATACTTTGGGATTC........................... 4521
LST.CD.88.524            -AT...............ACA-G--CTTT-GT-GTAA-AAT-TATGC-CATCTG-C-GC--GGA-AGGATGGATTTCACAGTGGGCA........................... 4560
LST.CD.88.447            -A-...............ACA-G--CCTT-GT-GTGA-AT-CTATGCACACTTG-C-GC--GAA-AGGATGGATCTCTCAGTGGGCA........................... 4563
LST.KE.x.lho7            -A-...............-CA-G--CTTT-GT-GTCA-AATATATGCACATTTG-C-GC--GAA--GGATGGATCTCACAATGGGCA........................... 5647
SYK.KE.x.SYK173          --A...............GAT-G--CT-T-TGGATCAC-TTCTAT--C-ATTTG-CCCC---AA-AGGATGGTTACATATGCAGGGA........................... 5489
SYK.KE.x.KE51            --A...............CAA-GTG--TT--TTATAAAATTCTAT----ATTTGAC-CC---GAAAGGCCTGTTAAGAACAGAGGGA........................... 5141
COL.CM.x.CGU1            TCA...............GCTA-T-AGTT--GCATAATATACT--GGC--TTT-ACCCC---TA-AT--CAT........................... 5034
```

```
H1B.FR.83.HXB2              ..............................TTGGGTCAGGGAGTCTCCATAGAATGGAGGAAAAAGAGATATAGCACACAAGTAGACCCTGAACTAGCAGACCAACTAATTCATCTGTATTACTTTGACTGTTTTTCAGACTCTGCTATAAGAAAGGCCTTATTAGGAC 5419
Vif                         __.__.__.__.__.__.__.__.__.__._L__G__Q__G__V__S__I__E__W__R__K__K__R__Y__S__T__Q__V__D__P__E__L__A__D__Q__L__I__H__L__Y__Y__F__D__C__F__S__D__S__A__I__R__K__A__L__L__G__
H1A1.UG.85.U455             ..............................--------T--G------------------CTG--A-------------------T-----C-----------T--------C---C----T-----------------A-----C-----G-GA---A--------- 4865
H1B.US.90.WEAU160           ..............................--------A---------------------C-------------------------C--------------------------C-T-----T-----------G--C-----------T---A--------- 5418
H1C.ET.86.ETH2220           ..............................--------T-------------------ATTG-GA--C----A---------------GC--G--------C--------A--C----T-----T-----G---A-----C-----G--A---A--------T 4811
H1D.CD.84.84ZR085           ..............................C-------------------------------G------------------------GC--G-------------------A-------T-----T-----G------------C-----A---A--------- 4941
H1F1.BE.93.VI850            ..............................C-------------------------C-GGG--AG----G-----A---------GC--G--G-------G--C---A-A-----T-----T----------A-----C-----G--A---A--------- 4759
H1G.SE.93.SE6165            ..............................--------T--G-----------G------C-G-GA--------A-----------G-----C-----------T--G----------C----T-----T-----------------C-----G--A---A--------- 4816
H1H.CF.90.056               ..............................--A--C----------------------ACTG--A--G-----------------G---GC--G------------------A--C----T-----T-----------A-----C-----G--A---A--------- 4766
H1J.SE.93.SE7887            ..............................-----------------------------C-G-GA--G--C--A---------------GC--G------------------A---G---------T-----------------C-----G-----A--------- 4733
H1K.CM.96.MP535             ..............................-----------------------------CTG--A--G--C--A-----------A-----C---------------------A-A-----T-----T----------A-----AG----G-----A--------- 4615
H101_AE.TH.90.CM240         ..............................--------T-------------------C-G-GA-A--------------A---T-----C---------A---G--------AC-A--T----G-----------------C-----G--A---A--------- 4993
H102_AG.NG.x.IBNG           ..............................--------T--G---------------A-C-G--A---------------A---------CT--------------G-----C--------T---A-T-----------------C---------A---A--------G 4944
H1N.CM.95.YBF30             ..............................C-A--A--T-----A-------------ACT--GA-A---C-AG--------T--T------A-G------A-G-----A-----TC----T-----T------A---C------C-----GC-A--GG-C-----GA 5013
H1O.BE.87.ANT70             ..............................-----A--T--G--TAGT---------C-TT-T----AG----AA-----GA-T---------AC-------AGGA-G--A------C----T---AC-------A---CA--A--AG-C--G--A---A-CC------ 5474
H1O.CM.91.MVP5180           ..............................-----A--T--G--TAGT---------CAAT-C---GAG----AA-----GA-T--T------AC-------AGGA-G--A------C----T--CACA------A----A--A--A--C--G------A-TC----G- 5449
CPZ.CD.90.ANT               ..............................--A--A--T--T--GAG------------ATGTGGA-AG----T----------------A--AC---T--T---A-G--A--C-AA-----T--CCC------------TCAA--AG-C---C-A--AA-CC-----G 4864
CPZ.CM.01.SIVcpzCAM13       ..............................--A-----T------------C-G-------TGGGATC----GTT-----G-----T---T-TG-------TA--T-G--C--CTCTC-A--T-----T-----G-----------C--C----GA--TA-------G- 5002
CPZ.CM.05.SIVcpzEK505       ..............................--A--A--T-----A---------------AC--GGA-T---C--G------A----T------T-G-------A-G-----A--C--T-----T-----T-----A---C------C--C--GC-A--GG-C-----GA 4991
CPZ.CM.05.SIVcpzMB66        ..............................C-A--C--T--G--AAGT------------AC-G-GA--T-----G------A----T------T--G---------T----C---T----C--T-----T-----G----A-----C-----G-----A--------- 4965
CPZ.CM.98.CAM3              ..............................C----G--T-----G---------------ATGTGG-TC------T-----------------GA-G-----TA-----------C--TC-A--T-----T-----G---CT-----C--C--G--A--TG-T-----GA 4833
CPZ.GA.88.GAB2              ..............................C-A--C--T-----G--------G------C-G-GA-C------A--T--GA----T---C--AC------T-GG--G--A--CTGT-----T-----T-----G---C--------G-C----GA---A-CC----GG 4812
CPZ.CM.98.CAM5              ..............................C----A--T-----G---------------ATGTGG-TC---C-AA------------T------ACG-----TAG-----------C--TC----T-----T-----G---CT-----C--C--G--A--TG-T-----GA 5121
CPZ.CM.05.SIVcpzMT145       ..............................C----A--T-----AG-A------------AG---GAG-T--C--T--C--G------------AC------TAGG-----A--C--TC----T-----T-----A---CAA----C--C--G--C--TA-C----A-A 4961
CPZ.TZ.00.TAN1              ..............................A-A--ATTA.............---------AC--GGA-A---C--T-------A-----------AC-----------T-G-----CACTAGG--T---AC-------G-T-CAGGA--AG-TC-GC-A--AA--------G 5066
CPZ.CM.05.SIVcpzLB7         ..............................--A-----T--G--A--T--C--G----AAC-G----A------T-------A----------GT-G-----T--TA-G------T-A-----T---A-T------G----A-----C-----G-----A--------- 4974
CPZ.US.85.CPZUS             ..............................--A--G--T-----A--A------------AC--GGA-CT--C-AG------A----T-----GA-G-------A-G-------------------T-----T-----A---C-------C--C-----A--TG-T-----GA 5479
CPZ.GA.88.GAB1              ..............................--A--C--T-----------TC--------CT-GGATCT---GT-------G--G--T--CTTCAC------TAG-T-G--C--CTCTC-A--T-----T--C---G-----AA-----C--C----GA--TA-------G- 5483
MON.CM.99.L1                ....GTCCCTTTAGATCCCGTGTGG......AATCCAA-CCC-AGAGAT-G---GGTC...-G--CAT-G---GCT--C--TC-GACG--G--C--G--GA-T--GA--G-A---ACAC----T----CA--C--CCA-C-G-TG-ACG-GC-C-GA--AA-CAG----G 5245
MON.NG.x.NG1                AGAGGGAGTAGTAGATCGAGTGTGGTTTCCTAGTCCCAGAA--CCGGA-CGG---ATTGAT-G-GTCT-G---GCT--CT-CC--ACA--A--TG----T------A----C--CACCCGC------TCA-----CCAG-C-----ACG--C-T-GA--AA-TAG---GG 3856
GSN.CM.99.CN166             ..............................AGAA-CA-ACT--GAG-A-G---GG-AGAC-T--TTT-G--C-----TA-TC-GAC---A-G-G-----ATG---A-T--A---GGAA-A------CCA--C---CAGA--GAA-AC---C-G-G---AA--AGG--TG 5237
GSN.CM.99.CN71              ..............................AGAA--A--CT--GAG-A-GG---ACAGAC-T--CAT-G--C--T--TA-TT-GAC---A--CG-----ATG---A-C--A--CACAAGG------CCA-----CAGA--GAA-ACG--C-G-G---AA--AGG--TG 5219
MUS_1.CM.01.1085            ..............................AG-A-A--TCC-CAGGA--GG---CA-GATT-C-TCT-G---TCT---T-CT-GAG---CAC---G--CATGAG-A-C--A---ACCC-G--T---TCT--C--C-ATAG-CAG-AC--CCC--G--T-A-CAG----- 5196
MUS_1.CM.01.CM1239          ..............................AG-AAAGGTCC-AGAGAT-G----C--GACT-TGTTT-G---CA----T-CC-CACA-----GA----CATGA-GA-C--A--CACCC-CC----CTCT-----C-ATAG-CAG-A---CCCG-G--T-A-CAG----- 5186
MUS_2.CM.01.CM1246          ..............................AG--CAGCCCC-CAAAAT-G---GATAGAC-TGGT-CAT--CGAG--CTTCC-------A---G-GA-CATG--GA-T--A--CACAC-C-----CTCT-----CCA-C-GAGA-AC--CCA--G---AA-CAG----G 5294
MUS_2.CM.01.CM2500          ..............................AG--C-GCCCC-CAAGAT-GG--GGT-GAC-TGGTAT-G--C--T--CT-CC-T-----A---T-G--CATG---A-T------ACT---C----CTC-------CAG--ACAG-AC--C-T--GA---C-CAG---GG 5247
RCM.NG.x.NG411              ..............................G-A--AATA-C-TA-ATAGATA-CA-AT-TG-TCCCT-T--CTT----G-TA-T-----AAC-G-------TA--T--G-A---AATC-G--T---CCT------A----TAAA--A--T---C-A---A-C------G 4934
RCM.GA.x.GAB1               ..............................G-A--AATCTT-................----TA---GGAGAT---GTA---G-T---TTT---TGGACT-----TAGTT----A---AAAAT---T---CCA------A----TAGA--A--C---C-AAGTC--C----GG 4910
DEN.CD.x.CD1                ..............................A-CA-CATA-CTTAT...G--C-CCTC--C-C-GG-T-G...GCT---G-----------T-AC------TG-GA----A---AAAG-CC-T--CTCT--C---A-CAGACAA--AG-C---C-A------AGG---G 5339

MAC.US.x.239                ..............................G--A-GATAACCTGG-A-...........TC-----AC-T-T-G---G-T---ACA--AA-CTAT------ATTT--C-G---AGCAC---T--CCCT--C---A---CGGGA-AAG-G----G----A-CAGG---G 5724
Vif                         __.__.__.__.__.__.__.__.__.__._V__R__I__T__W__Y__.__.__.__.__S__K__N__F__W__T__D__V__T__P__N__Y__A__D__I__L__L__H__S__T__Y__F__P__C__F__T__A__G__E__V__R__R__A__I__R__G__
H2A.GW.x.ALI                ..............................G-AA-AATAACTTGG-AT...........-C-G-A-AG-TCT-G---G-T--TAC---A--CTGT------TCC-----A---GGCAC---T--CTCT--C---A-G-CAGG--AAG------GA---A-CAG---GG 5800
H2A.DE.x.BEN                ..............................G-AA-ATTAACCTGG-AT...........-C-G-A-AG-TCT-G---G-T--TAC---A--CTGT------ATC-----A---AGCAC---T--CTCT--C---A-G-CAGG--AAG------GA---A-CAG---GG 5807
H2A.SN.x.ST                 ..............................G-AA-A-TAACTTGG-AT...........-C-G-A-A--TCT-G---G-T--TAC---A--CTGT--G---TCC-----A---AGCAC---T--CTCT--C---A-G-CAGGC-AAG------GA---A-CAG---GG 5252
H2B.GH.86.D205              ..............................G-AA-A-TAAC-TGG-AT...........G-G-G--AC-T-TAT---G-T---ACA-----TG-G---------G---C-G---GG--C---T--CTCT--C--------C-AA--AAG----G-GA---A-CAGG---G 5780
H2B.CI.x.EHO                ..............................G-AA-A-TAAC-TGG-AT...........G-G-G---C-T-TAT---G-T---ACT-----TG---------G-T--C-G---GG--C---T--CTCT-CG---A---CTAA--AAG----G-GA---A-CAGG---G 5778
H2G.CI.92.ABT96             ..............................G-AA-ATTAACCTGG-A-...........-C--GA-A--T-TAT---G-T--GACA--A--GAC------T---T--T-G---GGA-C---T-----T--C---A---CTGG--AAG------GA---A-CAGG---G 5134
H2U.FR.96.12034             ..............................G-AA-GTTAACTTGG-A-...........-C--GAG-T-T-T-G---G-T---ACA-----TTGC---------T--C-G---GG-AG---T--CTCT--C---A---CGGG--AAG-G--G-GA---A-CAG----G 5285
MAC.US.x.EMBL_3             ..............................G--A-GATAACCTGG-A-...........TC--G-GACCT-TTGGACAG-TGTA--A-CAG----TGCAGG--TT-TTCCTGCA-AGCACTTA--TC-CTC---A---CGGGA-AAG-G----G----A-CAGG---G 5192
SMM.US.x.H9                 ..............................G--A-AATAACCTGG-A-...........TCC-GA-AT-TCT-G---G-T-C-ACA--A--TTAT------ACTT--C-G---GGCAC---T--CCCT--C--------GGGA-AAG--C---G----A-CAGG---G 5197
SMM.US.x.PGM53              ..............................G-AA-AATAACCTGG-A-...........TCC-GA-A--TCT-G---G-T---ACA--A--TTAT------ACTT--C-G---GGCAC---T--CCCT--C--------GGGA-AAG--C---G----A-CAGG---G 5641
SMM.US.x.SIVsmH635F_L3      ..............................G-AA-AATAACCTGG-A-...........TCC-GA-AT-TCT-G---G-T---ACA--A--TTGT------ACTT--C-G---AGCAC---T--CCCT--C--------GGGA-AAG--CA--GA---A-CAGG---G 5727
SMM.SL.92.SL92B             ..............................G-TA-GATTACTTGG-A-...........TC-G-A-A--T-T-G---G-T---AC---A--TG-------T--GT--GCA---AGTAC---T--CCCT--C---G---CGCAC--GG-G---C-A---A-CAG----G 5158
STM.US.x.STM                ..............................G--A-GATAACTTGG-AT...........-CC-GA-AT-TCT-GT--G-T---ACA--A--CTGT-----T--G---C-G---GGCAC---T--CCCT--C---A---CGGGA-AAG-G----GA---A-CAGG---G 5371
MNE.US.x.MNE027             ..............................G--A-AATAACCTGG-A-...........TC--G--AC-T-T-G---G-T---ACA--A--CTAT------ATTT--C-G---AGCAC---T--CCCT--C---A---CGGGA-AAG-G----G----A-CAGG---G 5192
DRL.x.x.FAO                 ..............................ACAATAG-ACATA-T..............C-GG---AG-T-T-T--CTT-A-------AAC-AC-------AG-A----C--CT--G-A--T--CCCT------A----TCAA--A------GGT-T----C-G---T 4989
MND_2.x.x.5440              ..............................--TACAGTA-A-TAT...........CAA--GG-A-AT---TT---TT-CA-------A-TGACG-----TAG-A----C--CGG-G-A--T--CCCA------A----TCAA--A---------A-----G--T---G 4986
MND_2.CM.98.CM16            ..............................A--ACAATA-A-TAT............-A-TTC-CA-AC--CTA---T--CA-------CC--AC-------AGGA-G--A---TG-A-A--TC-CCC-------A----ACAA--A------C-A---C-GC-----A 5423
MND_2.GA.x.M14              ..............................A-CACAGCA-A-TA-............-AC---GG-GAC---TAT--T--CA-------A-T-AC-------AG-A-G------TG-G-A--T--CCCA------A---CTG-C---G-G-------TGC-G-AC---G 5350
OLC.CI.97.97CI12            ..............................AATTT-A-ATATA-AGTA...........GGG-TTTC--GG-TG...T-TT-GACT--A--GTAT-T---TTTCA-GG-A--C--AGTG-------C-----GGA-TC-A-----------G--A--TA-T--G---A 5275
GRV.ET.x.GRI_677            ..............................G-A--GTTATCCTAT-A-......CATCAA--GGGA...----AG---G-------T--A-G-AC-------AG-A-G--A--C--A-----T---A--------A----TAGA--C--CCA-C----TA-CAG---GG 5598
VER.KE.x.TYO1_patent        ..............................G-A---ATACA-TATGTA...AGCAATTTAG--TCT-A------A---G--T---------CTAC------TAGTA----A---GGTC-C--T---A-T------AA---AAGA--C--CCA-C-A--TC-GAGG--C- 5655
VER.KE.x.9063               ..............................A-A--AATACA-TA--TG...-G-GCAGTAGG-....-AT-TC--A---G--T-------A-C-AC------TGG-A-G--A--C--AC-C--T---TCA------A----AAGG--C--TCA---A--TA--AGG---G 5655
VER.KE.x.AGM155             ..............................G-A--AATACA-TATGTT...AGCCTAGTA--TG-T-A------A---G--T----T--CA-TAC-------TCCA-G--A---TGTC----T---AC-------A----TAGA--C--CCA-C----AC--AGG---A 5650
VER.DE.x.AGM3               ..............................GA---GATACAGTATCTA...AGCAAT---G-TCCTT-G-----G---G--T-G--T----C-AC------TAGC--G--A---ACCC----T---ACT------A----AAGG--C--C--G--A---C----G---- 5153
TAN.UG.x.TAN1               ..............................G-A-CAATAA-TT-T......--GAATT--C----C-CG--C-AA---G-----ACT-----TG-G--G-----TA-G------TGCC-C-----CCCT--C---A----TAGA--C--CCA-C-----A-CAG---GG 5627
SAB.SN.x.SAB1C              ..............................ACA--AATA---TA-............--T---GG-GA--GGTTT---G-GC-T--T---TGGAC---G-----TA-T------TG--CC--T--CCC-------A----TAGA---G-CCA-C-A---A-CAGG---G 5820
MND_1.GA.x.MNDGB1           ........................AGCATAGAA-CAGTA-TTAGA...CCGCC-CCCTTTG-G--AGAG-GGT-T---G-GA--ACT--A--GG-------T--T-----A---T-AC---TT-A------C--CATG---AG---AG-T-TG--A---A-CAGG---G 5139
DEB.CM.99.CM40              ..............................G-A-CAATA-C-TGG...........GAA-C---A-C--GG-AG---G--C--ACA--A--TT-------T--C--T--A--C--AAGA------CCA-----C--CC-ACAG--AG-G---C-A---A--AGG---G 5231
DEB.CM.99.CM5               ..............................G-A--AATCTCTTGG...........GA-C-G-CAG-G-GGCAG---G--T--ACA-----C-----T--T--TT-G--A--CT-AAGA-----CCCT-----C--CC-ACAG--AG-C-C-C-A---C--AAG---G 5225
TAL.CM.01.8023              ..................ACAATATCCATTGA-T-GTCA-TCTCGGATG-GACC-T-GATGC----GTC---TC----G-GT--ACA--A--GG----------C--G-----CACCA-G--T--CCCA------A---C---A--AG-C---C-A---A-CAG---GG 4925
TAL.CM.00.266               ..................ACAATATCCATTGA-T-GTCA-TCTCGGAT----TCC-AGACG-----GTG--CGC---TG-----AC---AA-TG-G-----T--CT-G------ACCA-A--T--CCCA------A---CG--A--AG-C---C-A---A-CAG---GG 5427
SUN.GA.98.L14               ..............................-GTA--AT--A-TGG............-TTT-T--C--T---CAA---G--A----T---CTTG-G-----T--GA-G--C--CT-TA-G--T-----T----GGA-CTCTAGGTG---C----GA--TA-G-----GG 5824
WRC.CI.97.97CI14            ..............................AACAT-AGAATG--G-AT...........G----T----GGGAA---TTTA--ATG--ATC-GAG---A-TA--A-C--A--CAGTC-AC-T-----A----GGA--C-GAA--TGG-G-----TA-AA-TAGGAC-G 5716
WRC.CI.98.98CI04            ..............................AACATAAGAATG--G-AT...........G----T----GGGAA---TTTA-TATG--CTCTGA------TGT-A-G-----CAGTC-AC-TG---------GGA--C-AAA--TAG-T---G-TA--A-TAG-TC-G 5731
WRC.GM.05.Pbt_05GM_X02      ..............................AACATAAGAATG--A-GG...........G----T-A--GGGAA--TTTTA--ACTGTAA--T-G-----TGTCT-G--A---TCTC-AC-------G----GGA-CC-AAG---AG---A---CA-AA--AG-CAGG 4648
LST.CD.88.524               ..............................G-ATCAATA-A-TGG............-TCT-TGGA-AC---CAG---G----G--T--CATT-----T--T---A-GC-----G--A-G--T-----------GGGTGC-TAGG-AC---GTT-G----A--AGG---G 4687
LST.CD.88.447               ..............................G-ATCAATA-A-TGG............-TTT-TGG--AC---CAA---G--A----T--CATT------A-T---A-G--A--CTGTA-A--T---A----C-----TACAAGG-AC--T--G-----T----GG---G 4690
LST.KE.x.lho7               ..............................G-TTCCATA-A-TGG...........GTCT-T-GAGAC---CAA---G--A-------CAT-------------A-G--C--CTATA-A--T---A----C--CAT-C-AAGA-AC--T----G---AC-TC-G---G 5774
SYK.KE.x.SYK173             ..............................A-TA-GATTCA-TA-CAA...TGGAACCA-TGG--C.........--TG-CT--ACA--A-C-G---------GC--T--A---AAT-T---T--CCCT--C---A---CAAGA--GG-C-ACC----AG-CAG---GG 5616
SYK.KE.x.KE51               ..............................G-A--ATTAATTTGGG--...C-TCC-TCTGG-TG-.......--T---G-----ACA--AC--AC------TGC-T--G-A---AGTC-A-----CCCT-----------TAGG--AG-CCA-C-A--AC-TAG---GG 5271
COL.CM.x.CGU1               ..............................AGA--A-CACC------TTAT----TACA--T-GGT-AG-GGGAA---G-TT--AG-TGGC---A-----TAGGGTAC--G--CA-CATAAGGGGGA-G---G-ACGCATAGAG-AAGA-GC--GAAAAGCTA--A-GG 5173
```

```
H1B.FR.83.HXB2             ACATA......GTTAGCCCTAGGTGTGAATATCAAGCAGGACATAACAAG..................GTAGGATCTCTACAATACTTGGCACTAGCAGCATTAATA...............ACACCAAAAAAGATA...AAGCCACCTTTGCCTAGTGTTACGAAACTG 5547
Vif                        H  I  .  .  V  S  P  R  C  E  Y  Q  A  G  H  N  K  .  .  .  .  .  .  V  G  S  L  Q  Y  L  A  L  A  A  L  I  .  .  .  .  .  T  P  K  K  I  .  K  P  P  L  P  S  V  T  K  L
H1A1.UG.85.U455            -A--.......-----A---------------------------------..................-----------------T-----T--GAA-------G--...............--C----C--G-GC-...----------------------A----T-A 4993
H1B.US.90.WEAU160          -TT--......----TT-----------------------G---------..................---------------------------A-----------...............---------------...----------------------A------- 5546
H1C.ET.86.ETH2220          -T-G-......-----------------C-----------------T---..................--------------------------GA-------G---...............-AG---------GC-...---------C------------GT---T-A 4939
H1D.CD.84.84ZR085          -T--.......-----T---------------------------------..................--------CT----G--T--A------A-----------...............G------------G-...--------------C-------A---G--A 5069
H1F1.BE.93.VI850           -T-G-......A--------------A-C-----------------T---..................--------C--------T---------A-----------...............G-T---G-G----C-...--------GC-----------CCA-----A 4887
H1G.SE.93.SE6165           -A--.......-----------------------------------TC--..................-----------------T------T-GAA--T----G--...............---T------G--GT...-G------------C---------G--T-A 4944
H1H.CF.90.056              GTG--......-----A---------A-C----C------------AC--..................------A-----------------T--A--------G-G...............G--------------...----------------------GA--G--A 4894
H1J.SE.93.SE7887           -A--.......-----------------C--C------------------..................-----------------TC--------A-----------...............-G--------G--G-...---------------------CA------A 4861
H1K.CM.96.MP535            -T-G-......-----T-------------G-------------------..................--------CT-------T--------CA--------G--...............G-T----G--G-CC-...---------G------------A------A 4743
H101_AE.TH.90.CM240        -AG--......-----A-G--------------C-T--------------..................--------C--------T--------GAA--------C-...............----------G----...-G-------C----------A...G--A-A 5118
H102_AG.NG.x.IBNG          -A--.......-----A-----------------------------T---..................--------G--------T--------GAAT-T----G--...............G------C-----C-...----------------------G---GT-A 5072
H1N.CM.95.YBF30            GACC-......--ATTA----------------C------G--C--AC--..................-----CA-C--------TC-A------A----C-GGG-G...............GG-G----G----G-...--------C--A--------G--T--G--A 5141
H1O.BE.87.ANT70            -G-G-......--ACTGA-C-A---------C-CTA---------GTC--..................-----GA-A------CTGC-A--T---AG----G--G--...............-A-G---G--GC-GG...-----T--CC-A--C-----CCA----T-A 5602
H1O.CM.91.MVP5180          -G-G-......--GCTGA-C-A---------C-TG----------GTC--..................-----GA-A-------T---A--CT-GAA----G--G-G...............-A-GT-----GA-AT...-----T--CC-A--C-----CCA--G-T-A 5577
CPZ.CD.90.ANT              -A-G-......A-ATTGA-ATAC--CC-C--CA-GAAG-------GTC--..................-----GA-A--G--G--T--A--TT-TTGTAAGA-CC--......GAATTCAGAGG-TACCCC---GGTCCA-GAAG--AG--C------C-GT-T-TCT-A 5004
CPZ.CM.01.SIVcpzCAM13      -T--T......--AGAA--CC------CT--CAG--A------C-GGC--..................-----CA-C-------T-C-A--TT--AA---T---G--...............-A-GA-----G-C-T...-GA-----C--A--C-----AG-T---T-A 5130
CPZ.CM.05.SIVcpzEK505      GACC-......--A-TA----A---------C-CT-----G-----AC--..................-----C--C--------TC-----T--AT---C-GGG-G...............GG-GT-C-G----G-...-GA-----C--A--------G--C--GT-A 5119
CPZ.CM.05.SIVcpzMB66       -A--.......----------------------C-A--------------..................-----CAG-T-------T------T-GT--A--------...............-A-AGC---CCA-G-...-GA-----C-----C-------GA---T-A 5093
CPZ.CM.98.CAM3             -ACA-......--ATAT----A---------C-C------G---C-AC--..................-----G--AT-------T--A--C---AG---C-GGG--...............-G-GT-GG-----A-...--A-----C-----------C--T--GT-A 4961
CPZ.GA.88.GAB2             -G--T......--G-CAAAGGT--------TC-CTAGG--G---CCTC-A..................--TCAG--A--------TC----CA---T-CA-G--G--...............-AGA----G--A-AG...---------CC-A-------C--T-TCT-A 4940
CPZ.CM.98.CAM5             -ACA-......--ATAT----A---------C-C-A----G--CC-AC--..................-K---G--A--------T--A--C---AG---T-GGG-G...............GG-GT-C------A-...-CA-----C-----C---A--GSC--GT-A 5249
CPZ.CM.05.SIVcpzMT145      -GGC-......--CCT-----A--------TC-C--A---G--C-G----..................-----G---T---------G---T---AAGCAG--GG-T...............-A-A-C--G--A-AT...-GA------C----CTCA--AG-A--GT-A 5089
CPZ.TZ.00.TAN1             -A-G-......A-ATTGA-ATTC--CC-C-T----T-------C-GAC--..................-----GA----G----T---A--TT-CAG-AAGG--G-TGAGAGCCAA...GAT-A--AGCC----GG-CCA-G-AGG--C-----ATC------A-----A 5209
CPZ.CM.05.SIVcpzLB7        GAG--......-----------A-------T--C---------C--G---..................------A-C--------T------A-CT-T--C---T--...............CA-AG-C-GCCA-C-...-G---G--------------C-GC---T-A 5102
CPZ.US.85.CPZUS            GACC-......--A-TT--C-AA--------C-CTA----G--C--T---..................-----C---T-G----T------TT--AAG--C---G--...............GG--AG-GC--A-G-...-GA-----C--A--------G--T---T-- 5607
CPZ.GA.88.GAB1             -GC-T......--AGCG---C-T--------CA---A------C-GGC--..................---------CT---G-T------C---AA---T------...............T--GAG-G--G-CAT...-GA--------A--------CG-C---T-A 5611
MON.CM.99.L1               -G-AG......C-CTTGGGACA----C--C---T---AC---CT--................CCGAAG-----GA-C-CGTTGAC---A-AGAA-TT---C--CC-TGCCCTAGTGCGATATGG-G---CCCCC-C-AAGTCAT-CG--CCATTGGACC-CC--G..... 5389
MON.NG.x.NG1               -AGCG......M-CCTAGGG-AC---C-CC-CAT-CAGAC---CT--...............AATAAG------A---CT-C-AC-C-A-AGAA-TT---C--CC-TGCTTGCTTCAAGTATGG-AA-TGC-CC-CCGAGCCAT-CT-CACATTGGACC-CC--...... 3999
GSN.CM.99.CN166            -GCAG......A-CCTAGGA-AT------C-C-CG-A--CT--CCTGC--..................------A-A-CT--GACGC-T-AG-AGTT---C--CT-TGCATATGTCAAACATT............................................... 5336
GSN.CM.99.CN71             -ACAG......C-CTTAGGATAT-----CC-C-CG-A--CT--CCCAC--..................------A-A-CT--CACGC-T-A-AGGCT---C--CT-TGCGTACGTTAGATACT............................................... 5318
MUS_1.CM.01.1085           -GCAC......AC-CTAGGGGAT------C---C--AC-C---C-CA---..................------C---CT-CT-CT--ACA-G-CCT---T---T--GCATTAAAGAAGC.................................................. 5292
MUS_1.CM.01.CM1239         -ACC-......CCCTTAGGGGAT------C-C-C--G--C-----CA...AAG...............------C-C-CT-CCACTC-ACAGG--TT----C--T--ACATTAAAGAAACATGG-AAG-GT-CC-CCCAGCCA-AGA-C-CC.................. 5314
MUS_2.CM.01.CM1246         -ACAG......C-CCTAAGA-A------GC-CATCAAGAC---CT--CC-AAG...............------A---CT-T--CTC-A-A-AAGCT------CT-TGCATACATCAAACAT---AATGG-GCA-GCGCCTCCAAGT-ACCCT--CTC............ 5431
MUS_2.CM.01.CM2500         -A-GG......A-CCTAGGG-TT---C--C-C--GA---C---CCT-CCCAAG...............------A-C-CT-CT-C-C-A-AGAAGTTG--C--CT-CGCATATATCAAACAGGG-GATGG-GCA-C-CCCTCCAAGT-ACCCT--CCC............ 5384
RCM.NG.x.NG411             -GCA-......--GCTG-T-...----GG-T-G-GAA--------GA...............GATCAG-----GAGC--------TC-----A--AA---T-GGGCTAGGCAGCAGCTAAAGG............................................... 5033
RCM.GA.x.GAB1              -A-A-......---CTAGTG...----CC-TC----GG-------GA...............GATCAG-----GA----G----TT------A--CA---G-GGGCTAAA...AGTCAGTTAGATAGGT......................................... 5012
DEN.CD.x.CD1               -AG-G......--CCTGAGC-CT--C-CCAGGA-GAAG-C----T--C-TCCA......AGGACAGGTCC---TAGGGA--CGCCA-CATT--AGTATAT-G--T--......CAAGCAGCC-TTAG-T-TGGA-AGGTT-CCA-CATCACA--C-CCAGC-TT-GTT-C 5491

MAC.US.x.239               -ACA-......C-GCTGT--T-C--CAGG-TC-CGAG--CT-----GT-CCAG...............---CC-AGC-----G-----A-----GAA--T-G---GCG.............................................................. 5811
Vif                        E  Q  .  .  L  L  S  C  C  R  F  P  R  A  H  K  Y  Q  .  .  .  .  .  V  P  S  L  Q  Y  L  A  L  K  V  V  S  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
H2A.GW.x.ALI               -A-AG......T-ATTGT-CT-C--CA-T----CCCAG-CC-----GTCACAG...............---CCG--A--C----TTC----CT---TG-T-G-GCA-CAAA........................................................... 5890
H2A.DE.x.BEN               -A-AG......T-ATTGT-CT-C--CA-C----CCCA--CT-----AGCACAG...............---CC---A--T------C-A--C----T--T-G--CA-CAAA........................................................... 5897
H2A.SN.x.ST                -A-AG......C-ATTAT-CT-C--CA-C--C-CCCA--CC-----GT-CCAG...............---CCG--A--C--G-TTC----CT---TG-T-G-GCA-CAAA........................................................... 5342
H2B.GH.86.D205             -A-AG......A-ATTGT-CTAC--CA-C----C-T---CT--CG-AGG-CAG...............---CC-AGCT----G-TTC-A--C---AGG-TCG--CAGGAAGGAAAAA..................................................... 5876
H2B.CI.x.EHO               -A-AG......A-ATTGT-CCAC--CA-C--C-C-T---CT----CAGGCCAG...............---CC-AG-T----G-TTC-A--C---AG--TTG--CA-GAAGGAAAAG..................................................... 5874
H2G.CI.92.ABT96            -GCAG......A-ATTGT-CT-C--CA-C----C-A---CT-----A-G-CAG...............--GCCTAGCT-G----TTC-A--CT--CA--T-G-GCA-AAAGGAC........................................................ 5227
H2U.FR.96.12034            -ACAG......T-ATTGT-CT-C--CA-C-T--C-T-G-CT----GAC--CAG...............---CC----T------TTC-A-----CAG------GCA-G.............................................................. 5372
MAC.US.x.EMBL_3            -ACA-......C-GCTGT--T-C--CA-G-TC-CGAG--CT----GGT-CCAG...............---CC-AGC-----G-----A------AA--T-G---GCG.............................................................. 5279
SMM.US.x.H9                -A-A-......T-GCTGT--T-C--CA-G-TC-CGAA--CT-----G--TCAG...............---CC-AGC-----G--TC-A------A---T-G---GTC.............................................................. 5284
SMM.US.x.PGM53             -G-A-......T-GCTGT--T-C--CA-G-TC-CGAA--CT-----G--CCAG...............---CC-AGC-----G--TC-A------A---T-G-G-GTC.............................................................. 5728
SMM.US.x.SIVsmH635F_L3     -G-A-......T-GCTGT--T-C--CA-G-TC-CGAA--CT-----A--TCAG...............---CC-AGC-----G--TC-A------A---T-G---GTC.............................................................. 5814
SMM.SL.92.SL92B            -GCAG......--GTTAT-CTAC----GC--CGCG-T--CC---C--TCTAGT...............--CCA-AGC-----GCTG-----C---AAG-T-G--T--CAAA........................................................... 5248
STM.US.x.STM               -A-AG......T-ACTGT--T-C--CAGG-TCACGAA--CT-----G--CCAG...............---CC-AGCT----G-----A--TT--AA--T-G-GGA-C.............................................................. 5458
MNE.US.x.MNE027            -ACA-......C-GCTGT--T-C--CAGG-TC-CGAG--CT-----G--CCAG...............---CC-AG------G-----A-----GAG--T-G---GTT.............................................................. 5279
DRL.x.x.FAO                -T-G-......A--CAAG--...--CT-T-TC-CTAG---G--C--G...............GGTCAG-----GAG---T--G--TC-T--CT-GCA---CCAC---AAGTACCTTAAGC.................................................. 5085
MND_2.x.x.5440             -A-G-......C-GGTAG--...---T-C-TC-CCTGG-------GA...............GGTCAG-----GA-------G-TT-----CT--CA---C-ATC-CAGAGGTAGAAAGC.................................................. 5082
MND_2.CM.98.CM16           -G-G-......C-A-CAGTG...--CT-C-TC--CTGG--------A...............GGGAAG-----GAGC--G--G--------TT-GCT-AGT-AT-C-GCCTACTGTAACA.................................................. 5519
MND_2.GA.x.M14             -A-G-......A-AGTAG-G...---T-CAGC-CCTGG--------A...............GGTCAG-----GA--------CT------TT--AG---C-AC--CAAGTTTTGCAGAA.................................................. 5446
OLC.CI.97.97CI12           TGCC-......-A-CAT-GATAC--CC-GCTGG-GA----T---CTT...............GGG...--CCTG--GT----G-TGC----TT-GAG----G-C-G-A.............................................................. 5359
GRV.ET.x.GRI_677           -G-AG......TA--CGTGG...--CAC--TCA-G-A---------AGGTCAG...............---CA---A--G---CTT----------TT----AT-C-A.............................................................. 5682
VER.KE.x.TYO1_patent       ---G-......T--GT-TTC...---C-G-T--C--A---G-----A-GCACA.........GGACAG---CC----T-G--G---C-A--T--GCTC---CATCA-A.............................................................. 5745
VER.KE.x.9063              -G-G-......T-CGT-TTC...---C---TC-CC-AG--------A-CCACA.........GGAAAG---CCC--A--G--G-----A--C--TCTG--CCATCAGA.............................................................. 5745
VER.KE.x.AGM155            ---GG......T-C-T-TTC...---C---T--C--G---------ACTAACA.........GGTCAG---CCC--CT-G-----T--A---T--CT---CCATCA-A.............................................................. 5740
VER.DE.x.AGM3              -G-GG......T-C-C-TTC...---C-G-TC-CC-AG-----C--G--AACA.........GGACAG---CCC---T-G--------A--T--CCTT---CACCA-A.............................................................. 5243
TAN.UG.x.TAN1              -A-GC......T--TTGTGG...--CAC---CA-G-A------CGTAGCAGAAAACCATTGGGGTCAG---C-G--C--G--G-TT--A--C--GA---T--AT-C-GACTTCCTAAGAA.................................................. 5738
SAB.SN.x.SAB1C             -A-AG......TA-CTTTGG...---A--C-C--G-TG------C-GCCAACA.........GGGCAA---CC---AT----G--T--A-----TAGG-T--AT-C-A.............................................................. 5910
MND_1.GA.x.MNDGB1          -AGA-......--GTTAAAAGTT---AG--T--C---T--C-----AGCA............CAAGGT--TCTC---T-G--G-TTC-CTGCT-GAG--TCA-CT................................................................. 5226
DEB.CM.99.CM40             -G--C......C-ACTG-AACA----C-GC-GT-CCATA-GGGACCAG-AGGT.........AAAGGACC-CCT--A-----G---A-A-----TAG--T-GC-G--AAAC........................................................... 5327
DEB.CM.99.CM5              -G-A-......C-ACTA-AACA----C-GC-AT-CCATA-GGGACCAG-AGGT.........AAAGGACC-CCC--AT----G-----A------AGGA--GC-G--AAAT........................................................... 5321
TAL.CM.01.8023             -GG--......C--CTTG-GGT-------GTCAG-CATAA-GG--GG...............GACTCGCC-CC---CT----G-T-C-A--CT--AAG-T-GC-G-TAGGC........................................................... 5015
TAL.CM.00.266              -GG-G......C-ACTAG-AGTT----G-GTCAG-CACAA-GG--GG...............GACACGCC-CCG--------G-T-C-A--CT-GA-----GC-G-TAGAT........................................................... 5517
SUN.GA.98.L14              -A-A-......A-ACTA-A-GA---CAG-A------T--CT--C--A...............GGCCTG--TCTC---T-G--G-TTC-CTGCT-GAG--TTC-CC................................................................. 5908
WRC.CI.97.97CI14           GT-A-ATAAAC--G-TAGA-GCA---C-G-----CAA-------GCA...............GGA...--GCTTA-CT-G--GAC----TGCT--CAG-TC--GCA-CATGAGCAACTCCAG-G-G--.......................................... 5826
WRC.CI.98.98CI04           GA-GGTTAGATA-A--AAACGCA---C-------CAA-------GCA...............GGA...--GCTTA-CT-G--GAC----TGCT--CAG-TC--GCA-CATGAGCAACTCCAGGG-A--.......................................... 5841
WRC.GM.05.Pbt_05GM_X02     GTTATTTACATCC--TAGA---C---A-------TCA-------CTA...............GGA...--GCTTA-CT-G--GGC---ATGTT-GCAG--C--GCA-CATGAGCAACTCAGG-AGAG-.......................................... 4758
LST.CD.88.524              -A-AG......A--CTG-AC--C---A--ATG-CTA-T--------A...............GGTTGT--TCTC--AT-G--G---C-CTGCT-GAG-CA-C--CA-C.............................................................. 4774
LST.CD.88.447              -A-AG......A-A-TGG-ATTC----CTATG-CCA-T--------A...............GGTTGT--TCTC--AT-G--G-T-C-CTGCT-GAG-CA-C--CA-C.............................................................. 4777
LST.KE.x.lho7              -A-G-......A-A-CTATGT-T--CC-G...-CTTT---------A...............GGTTGT--TCTC---T-G--G-TTA-CTGCT-GAG-CA-C--CA-C.............................................................. 5858
SYK.KE.x.SYK173            -AC--......T-G-CAT-ACAC--CTGGACC-C-CATACTG-CC-G.....................---CCC--A--T--G--T--A---T--CA--TG-ACC--AAAG........................................................... 5700
SYK.KE.x.KE51              -G-AG......T-G-CTT-GCAC--CTGGA--TTCCATAA-G-AC-G.....................---CTT--AT-G-----T-----TT--CA-AAG-ATT--AGCAAG......................................................... 5357
COL.CM.x.CGU1              G-T--......CCCT-GAA-CCT-----T-TC--G-T------CCTA...............GGA...--GTTGAAGT-G---G-AC-ATGT--TCTG--CC---G-GAGAAA...GAGAGG--C-ACC-G---.................................... 5280
```

```
H1B.FR.83.HXB2               ACAGAGGATAG............................................................................................................................................................... 5558
Vif                          _T__E__D__R#__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__
H1A1.UG.85.U455              -----------............................................................................................................................................................... 5004
H1B.US.90.WEAU160            -----------............................................................................................................................................................... 5557
H1C.ET.86.ETH2220            GT--------A............................................................................................................................................................... 4950
H1D.CD.84.84ZR085            -----A-----............................................................................................................................................................... 5080
H1F1.BE.93.VI850             GT------C--............................................................................................................................................................... 4898
H1G.SE.93.SE6165             G----A-----............................................................................................................................................................... 4955
H1H.CF.90.056                GT---------............................................................................................................................................................... 4905
H1J.SE.93.SE7887             GT---------............................................................................................................................................................... 4872
H1K.CM.96.MP535              GT---------............................................................................................................................................................... 4754
H101_AE.TH.90.CM240          -----A-----............................................................................................................................................................... 5129
H102_AG.NG.x.IBNG            G----A--C--............................................................................................................................................................... 5083
H1N.CM.95.YBF30              -----A-----............................................................................................................................................................... 5152
H1O.BE.87.ANT70              -----A-----............................................................................................................................................................... 5613
H1O.CM.91.MVP5180            -----A-----............................................................................................................................................................... 5588
CPZ.CD.90.ANT                -----A-----............................................................................................................................................................... 5015
CPZ.CM.01.SIVcpzCAM13        -----------............................................................................................................................................................... 5141
CPZ.CM.05.SIVcpzEK505        -----------............................................................................................................................................................... 5130
CPZ.CM.05.SIVcpzMB66         G----------............................................................................................................................................................... 5104
CPZ.CM.98.CAM3               -----A-----............................................................................................................................................................... 4972
CPZ.GA.88.GAB2               G----A-----............................................................................................................................................................... 4951
CPZ.CM.98.CAM5               G----A-----............................................................................................................................................................... 5260
CPZ.CM.05.SIVcpzMT145        G----A-----............................................................................................................................................................... 5100
CPZ.TZ.00.TAN1               --------C--............................................................................................................................................................... 5220
CPZ.CM.05.SIVcpzLB7          G----A-----............................................................................................................................................................... 5113
CPZ.US.85.CPZUS              -----A-----............................................................................................................................................................... 5618
CPZ.GA.88.GAB1               -----------............................................................................................................................................................... 5622
MON.CM.99.L1                 .......................................................................................................................................................................... 5389
MON.NG.x.NG1                 .......................................................................................................................................................................... 3999
GSN.CM.99.CN166              .......................................................................................................................................................................... 5336
GSN.CM.99.CN71               .......................................................................................................................................................................... 5318
MUS_1.CM.01.1085             .......................................................................................................................................................................... 5292
MUS_1.CM.01.CM1239           .......................................................................................................................................................................... 5314
MUS_2.CM.01.CM1246           .......................................................................................................................................................................... 5431
MUS_2.CM.01.CM2500           .......................................................................................................................................................................... 5384
RCM.NG.x.NG411               ...AT--CAGAG......GGCAGAGAAAGAGTGCCA......GAGGCCCCCACTGGGGCTGGAGATGTAGAGTTTGCCCCCTGGCTTCACAGAATGCTAACAGAAGTCAACTTAGAAGCC...AGGTTGCACTTTCATCCAGAGTTCATTTTCCGTCTGTGGCGTACTTG 5185
RCM.GA.x.GAB1                ...AT--CAGAA..............AGAGCCCCC......GAGGTCCCCACTGGGGCTGGAGAAGCAGAGTTCCAGCCTTGGCTACGGGACATGCTAGAAAAGGTCAACTTGGAAGCC...AGGTTACACTTTCATCCAGAGTTCATTTTCCGTCTGTGGCGCACTTG 5155
DEN.CD.x.CD1                 -AGA...................................................................................................................................................................... 5495
                                Vpx start

MAC.US.x.239                 ...AT-TCAGAT......CCCAGGGAGAGAATCCCA......CCTGGAAACAGTGGAGAAGAGACAATAGGAGAGGCCTTCGAATGGCTAAACAGAACAGTAGAGGAGATAAACAGAGAGGCGGTAAACCACCTACCAAGGGAGCTAATTTTCCAGGTTTGGCAAAGGTC 5966
Vpx                             M  S  D  .  .  P  R  E  R  I  P  .  .  P  G  N  S  G  E  E  T  I  G  E  A  F  E  W  L  N  R  T  V  E  E  I  N  R  E  A  V  N  H  L  P  R  E  L  I  F  Q  V  W  Q  R  S
Vif                          _##D__V__R__.__.__S__Q__G__E__N__P__.__.__T__W__K__Q__W__R__R__D__N__R__R__G__L__R__M__A__K__Q__N__S__R__G__D__K__Q__R__G__G__K__P__P__T__K__G__A__N__F__P__G__L__A__K__V
H2A.GW.x.ALI                 ...AT--CA-AC......CCCAGAGAAACAGTACCA......CCAGGAAACAGTGGCGAAGAGACTATCGGAGAGGCCTTCGAGTGGCTAGACAGGACAGTAGAGGCCTTAAACAGAGAGGCGGTGAATCACCTGCCCCGGGAGCTCATTTTCCAGGTGTGGCAAAGGTC 6045
H2A.DE.x.BEN                 ...AT-ACAGAC......CCCAGAGAAAGGGTACCG......CCAGGAAACAGTGGAGAAGAGACCATTGGAGAGGCCTTCGAGTGGCTAGAGAGGACCATAGAAGCCTTAAACAGGGAGGCAGTGAACCATCTGCCCCGAGAGCTCATTTTCCAGGTGTGGCAAAGGTC 6052
H2A.SN.x.ST                  ...AT--CAG-C......CCCAGAGAGACAATACCA......CCAGGAAACAGTGGCGAAGAAACTATCGGAGAGGCCTTCGAGTGGCTAGACAGGACGGTAGAAGCCATAAACAGAGAGGCAGTGAACCACCTGCCCCGAGAGCTTATTTTCCAGGTGTGGCAAAGGTC 5497
H2B.GH.86.D205               ...AT-...GAT......CCCAGGGAGAGAGTGCCA......CCAGGAAACAGCGACGAAGAAACAGTAGGAGAAGCATTCGCTTGGCTAGAAAGAACAATAACAGAGCTCAACAGGGTAGCGGTCAACCATTTGCCCCGAGAACTTATTTTCCAGGTCTGGCAGAGGTC 6028
H2B.CI.x.EHO                 ...AT-...GAT......CCCAGGGAGAGAGTACCA......CCAGGAAACAGCGGCGAAGAAACAGTAGGAGAGGCATTCGAATGGCTAGAGACAACATTAGAACATCTCAACAGAGTAGCAGTCAATCACTTGCCCAGGGAACTTATTTTCCAGGTCTGGCAGAAGTC 6026
H2G.CI.92.ABT96              ...AT-...GAT......CCAAGGGAGAGAGTMCCA......CCAGAAAACAGCGACGAGGAGACAATCGGAGAGGCATTCGAATGGCTAGAAAGAACAGTMACAGAGATCAACAGGATAGCAGTCAATCATTTACCCAGGGAGCTTATTTTCCAGGTTTGGCGAAGGTC 5379
H2U.FR.96.12034              ...AT--GAGAT......CCAAGAGAGAGAATCCCA......CCAGGAAATTCAGAAGAAGAAACTGTAGGAGAGGCTTTCAGTTGGCTAGAAAGAACTGTGAGAGATATCAACAGGGCAGCAGTGCAACATCTTCCCAGGGAACTTATTTTCCAGGTCTGGCAAAGGTC 5527
MAC.US.x.EMBL_3              ...AT-TCAGAT......CCCAGGGAGAGAATCCCA......CCTGGAAACAGTGGAGAAGAGACAATAGGAGAGGCCTTCGAATGGCTAAACAGAACAGTAGAGGAGATAAACAGAGAGGCGGTAAACCACCTACCAAGGGAGCTGATTTTCCAGGTTTGGCAAAGGTC 5434
SMM.US.x.H9                  ...AT-TCRSRT......CCCAGGGAGAGRATCCCA......CCTGGAAACAGTGGAGAAGAAACRRTAGGCGAGGCYTTCGACTGGCTACACAGAACASTAGAAGAAATAAACAGGGCAGCAGTGAATCATTTGCCRAGGGAGCTAATTTTCCAGGTCTGGCGAAGGKC 5439
SMM.US.x.PGM53               ...AT-TCAGAT......CCCAGAGAGAGGATCCCA......CCTGGAAACAGTGGAGAAGAAACAATAGGAGAGGCATTCGAATGGCTAAACAGAACAGTAGAAGAAATAAACAGGGCAGCAGTGAATCACTTGCCGAGGGAGCTAATTTTCCAGGTTTGGCGAAGGTC 5883
SMM.US.x.SIVsmH635F_L3       ...AT-TCAGAT......CCCAGAGAGAGGATCCCA......CCTGGAAACAGTGGAGAGGAAACAATAGAAGAGGCTTTCGAATGGCTAAACAGAACAGTAGAAGGAATAAACAGGGCAGCAGTAAATCACCTGCCGAGGGAGCTAATTTTCCAGGTTTGGCAAAGGTC 5969
SMM.SL.92.SL92B              ...AT-ACAGAC......CCAAGGGAAAGAATCCCA......CCAGGAAACAGTGGAGAAGAAACAATAGGAGAGGCCTTCGAATGGTTACACAATACAGTGGAGGCCCTGAATCAAACAGCAGTACAACACCTACCAAGAGAGTTAATTTTCCAGGTCTGGAGAAGGTG 5403
STM.US.x.STM                 ...AT-TCAGAT......CCCAGAGAGAGAATACCG......CCCGGAAACAGTGGAGAAGAGGCAATAGAGGAAGCATTCGAGTGGCTACACAGAACGGTAGAGGACATAAACCGAGAGGCAGTAAACCATCTACCGAGGGAACTGATTTTCCAGGTCTGGCGAAGGTC 5613
MNE.US.x.MNE027              ...AT-TCAGAT......CCCAGAGAGAGAATCCCA......CCTGGAAACAGTGGAGAAGAGACAATAGGAGAAGCCTTCGAATGGCTAAACAGAACAGTAGAGGAGATAAACAGAGAGGCAGTAAACCACCTACCAAGGGAGCTGATTTTCCAGGTTTGGCAAAGGTC 5434
DRL.x.x.FAO                  ...AT--CAGAAAGACAGTCAGTGGAGAGAGCTCCAGCGGAGCCAATGGGAGCAGGAGAG.........GTAGAGTTAGAAGAATGGCTACAGAGGAGTCTCTTAAGAATCAACCAGGAGGCTCGATTACACTTCCACCCAGAGTTCCTCTTCCGTCTTTGGAACACCTG 5243
MND_2.x.x.5440               ...AT--CAGAG..............AGGGCACCAGAGGCGCCAGAAGGAGCAGGAGAG.........GTAGGACTGGAGCAATGGCTGGAAACGTCACTGGAGAGAATCAACCGGGAGGCCCGGTTACACTTCCACCCAGAGTTCCTTTTCCGTCTCTGGAACACATG 5225
MND_2.CM.98.CM16             ...AT--CAGAA..............GGGGCCCCAGAGATCCCAGAAGGAGCAGGAGAG.........GTGGATCTAAATACCTGGTTGGAAAGATCTCTGGAGAAAATCAATCAGGAGGCAAGATTACACTTCCACCCAGAGTTCCTTTTCCGTCTCTGGAACGCATG 5662
MND_2.GA.x.M14               ...AT--CAGAA..............AGAGCACCAGAGGCACCACAGGGGGCAGGAGAA.........GTGGGACTAGAACAGTGGCTGGAAAGATCATTGGAACAAGTGAACAGAGAGGCTCAATTACACTTCCACCCAGAGTTCCTTTTCCGTCTTTGGAACACCTG 5589
OLC.CI.97.97CI12             .......................................................................................................................................................................... 5359
GRV.ET.x.GRI_677             .......................................................................................................................................................................... 5682
VER.KE.x.TYO1_patent         .......................................................................................................................................................................... 5745
VER.KE.x.9063                .......................................................................................................................................................................... 5745
VER.KE.x.AGM155              .......................................................................................................................................................................... 5740
VER.DE.x.AGM3                .......................................................................................................................................................................... 5243
TAN.UG.x.TAN1                .......................................................................................................................................................................... 5738
SAB.SN.x.SAB1C               .......................................................................................................................................................................... 5910
MND_1.GA.x.MNDGB1            .......................................................................................................................................................................... 5226
DEB.CM.99.CM40               .......................................................................................................................................................................... 5327
DEB.CM.99.CM5                .......................................................................................................................................................................... 5321
TAL.CM.01.8023               .......................................................................................................................................................................... 5015
TAL.CM.00.266                .......................................................................................................................................................................... 5517
SUN.GA.98.L14                .......................................................................................................................................................................... 5908
WRC.CI.97.97CI14             .......................................................................................................................................................................... 5826
WRC.CI.98.98CI04             .......................................................................................................................................................................... 5841
WRC.GM.05.Pbt_05GM_X02       .......................................................................................................................................................................... 4758
LST.CD.88.524                .......................................................................................................................................................................... 4774
LST.CD.88.447                .......................................................................................................................................................................... 4777
LST.KE.x.lho7                .......................................................................................................................................................................... 5858
SYK.KE.x.SYK173              .......................................................................................................................................................................... 5700
SYK.KE.x.KE51                .......................................................................................................................................................................... 5357
COL.CM.x.CGU1                .......................................................................................................................................................................... 5280
```

```
H1B.FR.83.HXB2              ............................................................................................................................................................................... 5558
Vif                         .__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__._
H1A1.UG.85.U455             ............................................................................................................................................................................... 5004
H1B.US.90.WEAU160           ............................................................................................................................................................................... 5557
H1C.ET.86.ETH2220           ............................................................................................................................................................................... 4950
H1D.CD.84.84ZR085           ............................................................................................................................................................................... 5080
H1F1.BE.93.VI850            ............................................................................................................................................................................... 4898
H1G.SE.93.SE6165            ............................................................................................................................................................................... 4955
H1H.CF.90.056               ............................................................................................................................................................................... 4905
H1J.SE.93.SE7887            ............................................................................................................................................................................... 4872
H1K.CM.96.MP535             ............................................................................................................................................................................... 4754
H101_AE.TH.90.CM240         ............................................................................................................................................................................... 5129
H102_AG.NG.x.IBNG           ............................................................................................................................................................................... 5083
H1N.CM.95.YBF30             ............................................................................................................................................................................... 5152
H1O.BE.87.ANT70             ............................................................................................................................................................................... 5613
H1O.CM.91.MVP5180           ............................................................................................................................................................................... 5588
CPZ.CD.90.ANT               ............................................................................................................................................................................... 5015
CPZ.CM.01.SIVcpzCAM13       ............................................................................................................................................................................... 5141
CPZ.CM.05.SIVcpzEK505       ............................................................................................................................................................................... 5130
CPZ.CM.05.SIVcpzMB66        ............................................................................................................................................................................... 5104
CPZ.CM.98.CAM3              ............................................................................................................................................................................... 4972
CPZ.GA.88.GAB2              ............................................................................................................................................................................... 4951
CPZ.CM.98.CAM5              ............................................................................................................................................................................... 5260
CPZ.CM.05.SIVcpzMT145       ............................................................................................................................................................................... 5100
CPZ.TZ.00.TAN1              ............................................................................................................................................................................... 5220
CPZ.CM.05.SIVcpzLB7         ............................................................................................................................................................................... 5113
CPZ.US.85.CPZUS             ............................................................................................................................................................................... 5618
CPZ.GA.88.GAB1              ............................................................................................................................................................................... 5622
MON.CM.99.L1                ............................................................................................................................................................................... 5389
MON.NG.x.NG1                ............................................................................................................................................................................... 3999
GSN.CM.99.CN166             ............................................................................................................................................................................... 5336
GSN.CM.99.CN71              ............................................................................................................................................................................... 5318
MUS_1.CM.01.1085            ............................................................................................................................................................................... 5292
MUS_1.CM.01.CM1239          ............................................................................................................................................................................... 5314
MUS_2.CM.01.CM1246          ............................................................................................................................................................................... 5431
MUS_2.CM.01.CM2500          ............................................................................................................................................................................... 5384
RCM.NG.x.NG411              TGTGGAACACTGGCATGATAGGCTTGGAAGAAGCCTTGAGTATGCAGGCTATAGATATCTGCTTCTGATGCAAAAAGCTCTGTTTATTCATTGCCAATCAGGGTGTTCTCAGAGA.........CATGGACAGGGACAAGCAAGGGAAGCAGGAGAAAGAATCCAGATTC 5346
RCM.GA.x.GAB1               TGTGGAACACTGGCATGACGTACATCAGAGAAGTTTAGAATATGCTGCTTATAGGTACTTACTGCTGATGCAAAAGGCTTTGTTTATCCACTGTCAGACAGGGTGTAGTCAGAGA.........CATGGACCCAATCCTAGG...GCAGTAGGAGAAAGGATAACCATCC 5313
DEN.CD.x.CD1                ............................................................................................................................................................................... 5495

                                           Vif end
MAC.US.x.239                TTGGGAATACTGGCATGATGAACAAGGGATGTCACCAAGCTATGTAAAATACAGATACTTGTGTTTAATACAAAAGGCTTTATTTATGCATTGCAAGAAAGGCTGTAGATGTCTAGGGGAAGGACATGGGGCAGGGGGATGG...AGACCAGGACCTCCTCCTCCTCCCC 6133
Vpx                         _W__E__Y__W__H__D__E__Q__G__M__S__P__S__Y__V__K__Y__R__Y__L__C__L__I__Q__K__A__L__F__M__H__C__K__K__G__C__R__C__L__G__E__G__H__G__A__G__G__W__.__R__P__G__P__P__P__P__P__
Vif                         _L__G__I__L__A__*
H2A.GW.x.ALI                CTGGAGATACTGGCATGATGAACAAGGAATGTCACAAAGTTACACAAAATATAGATACTTGTGCTTAATGCAGAAAGCTATGTACACACATTTCATGAAAGGGTGCACCTGCCTGGGGGGAGGGCATGGGCCAGGAGGATGG...AGATCAGGACCTCCCCCTCCTCCCC 6212
H2A.DE.x.BEN                CTGGAGATATTGGCATGATGAACAAGGGATGTCAGCAAGCTACACAAAGTATAGATATTTGTGCCTAATGCAAAAAGCTATATTTACACATTTCAAGAGAGGGTGCACTTGCTGGGGGGAGGACATGGGCCGGGAAGGATTGGAAGACCAAGGACCTCCCCCTCCTCCCC 6222
H2A.SN.x.ST                 CTGGAGATACTGGCATGATGAACAAGGAATGTCAATAAGTTACACAAAGTATAGATATTTGTGCCTAATGCAGAAAGCTATGTTCATACATTCTAAGAGAGGGTGCACTTGCCTGGGGGGAGGACATGGGCCGGGAGGATGG...AGATCAGGACCTCCCCCTCCTCCCC 5664
H2B.GH.86.D205              TTGGGCATACTGGCGTGAGGAACAGGGCATGTCAATTAGCTATACCAAATATAGATACTTGTTGCTAATGCAGAAAGCAATGTTTGTGCACTATACAAAGGGCTGTAGGTGCCTGCAGGAGGGCCATGGGCCAGGGGGATGG...AGATCAGGACCTCCTCCTCCTCCTC 6195
H2B.CI.x.EHO                TTGGGCATATTGGCGTGAGGAACAGGGCATGTCAATTAGTTATACCAAATATAGGTATTTGTGTCTAATGCAAAAAGCAATGTTTATACACTTTGCAAAGGGTTGTGGATGTCTGCGGGAAGGCCATGGGCCAGGGGGATGG...AGATCAGGACCTCCTCCTCCTCCTC 6193
H2G.CI.92.ABT96             CTGGGAATATTGGCATGATGAACAGGGCATGTCAGCTAGTTACACTAAGTATAGATATTTGTTGCTGATGCAAAAGGCTATGTTTGTACATTGTAAGAAAGGATGCACATGCTTGCAGAAGGGCCATGGGCCTGGAGGATGG...AGACAAGGACCTCCTCCTCCTCCTC 5546
H2U.FR.96.12034             CTGGCAATACTGGCGTGATGAACAGGGCATGTCAGAAAGTTATACTAAGTATAGATATCTCTGTCTGATACAGAAAGCTTTATTTATGCATTGCAAAAAGGGATGTCAATGCTTGGGGGAAGGCCATGGACCAGGAGGATGG...AGATCAGGGCCTCCTCCTCCTCCCC 5694
MAC.US.x.EMBL_3             TTGGGAATACTGGCATGATGAACAAGGGATGTC.............................................................................................................................................. 5467
SMM.US.x.H9                 CTGGGAATACTGGCATGATGAAATGGGGATGTCAGTCAGCTACACTAAATRCAGATACTTGTGCTTGATACAGAAAGCTCTGTTTATGCATTGCAAGAAAGGGTGTAGGTGCTTAGGAGGAGAGCATGGGGCAGGGGGATGG...AGACCAGGGCCTCCTCCTCCTCCCC 5606
SMM.US.x.PGM53              TTGGGAATACTGGCGTGATGAAATGGGGATGTCAGAGAGCTACACAAAATACAGATACTTGTGCTTAATACAGAAAGCTCTGTTTGTGCATTGCAAGAGAGGGTGTAGGTGCTTAGGAGAAGAGCATGGGGCAGGGGGATGG...AGATCAGGGCCTCCTCCTCCTCCCC 6050
SMM.US.x.SIVsmH635F_L3      CTGGGAATACTGGCATGATGAAATGGGAATGTCAGAAAGCTACACAAAATATAGATACTTGTGCTTGATACAGAAAGCTCTGTTTATGCATTGCAAGAAAGGGTGTAGGTGCTTAGGAGAAGGGCATGGGGCAGGGGGATGG...AGAACAGGGCCTCCTCCTCCTCCCC 6136
SMM.SL.92.SL92B             TTGGGAATACTGGGTAGATGAGCAAGGATACTCACCAAGCTATGCAAAATATAGATATGTACAGTTAATGCAAAAAGCAATGTTTCAGCACTTCAGAAAGGGATGCACATGCAGAGGAGAAGGGCATTCACAGGGGGGCTGG...AGAACTGGACCACCACCTCCTCCCC 5570
STM.US.x.STM                TTGGGAATACTGGCATGATGAGCAAGGGATGCCCGGGAGCTATGTTAAATATAGATACTTGTGCTTGATACAGAAAGCTTTGTTTATGCATAGCAAGAGGGGATGCAGATGCTTGGGGGAAGGACATGGGGCAGGAGGATGG...AGACCAGGCCCTCCTCCTCCTCCCC 5780
MNE.US.x.MNE027             TTGGGAATACTGGCATGATGAACAAGGGATGTCACCAAGCTATGTAAAGTACAGATACTTGTGTTTAATACAAAAGGCTTTATTTATGCATTGCAAAAAAGGCTGTAGATGTCTAGGGGAAGGACATGGGGCAGGGGGATGG...AGACCAGGACCTCCTCCTCCTCCCC 5601
DRL.x.x.FAO                 CATGGAGCACTACCATGATGCTCTTCAGTTATCTTTTACTTACAGCAAGTATAGATACCTACTTTTGTTACAGAAGGCCATGTTCATGCACTTTCAGCAAGGATGCTCATGTCTGCAGGGAAGGCATCCACCTCCCCTCAGACCAGCAGGAGATAGACTTCCTCCTCCTC 5413
MND_2.x.x.5440              TGTAGAACACTGGCATGATAGACATCAGAGATCTCTTGATTATGCCAAGTATAGATACCTGCTGTTGATGCATAAGGCCATGTATACTCACATGCAACAGGGATGCCCATGTAGAAATGGG...CGCCCAAGG..................GGA.........CCTCCTC 5365
MND_2.CM.98.CM16            CATAGAACACTGGCATGACAGACATCAGAGAAGCTTAAGCTATGCAAAGTATAGATACCTGCTGTTGATGAATAAAGCAATGTTTACTCACATGCAACAAGAATGTCCTTGTAGAAGTGGG...CATCCG...............AGA...GGA.........CCTCCTC 5802
MND_2.GA.x.M14              TGTAGAACACTGGCATGATAGACTTAGAAGAACCTTAAATTATGCCAAATATAGGTATCTTCTGCTGATGCAGAAAGCCATGTTTGTTCATATGCAACAAGGATGCCCATGTAGAAGTGGA...CATTCA...............AGA...GGA.........CCTCCTC 5729
OLC.CI.97.97CI12            ............................................................................................................................................................................... 5359
GRV.ET.x.GRI_677            ............................................................................................................................................................................... 5682
VER.KE.x.TYO1_patent        ............................................................................................................................................................................... 5745
VER.KE.x.9063               ............................................................................................................................................................................... 5745
VER.KE.x.AGM155             ............................................................................................................................................................................... 5740
VER.DE.x.AGM3               ............................................................................................................................................................................... 5243
TAN.UG.x.TAN1               ............................................................................................................................................................................... 5738
SAB.SN.x.SAB1C              ............................................................................................................................................................................... 5910
MND_1.GA.x.MNDGB1           ............................................................................................................................................................................... 5226
DEB.CM.99.CM40              ............................................................................................................................................................................... 5327
DEB.CM.99.CM5               ............................................................................................................................................................................... 5321
TAL.CM.01.8023              ............................................................................................................................................................................... 5015
TAL.CM.00.266               ............................................................................................................................................................................... 5517
SUN.GA.98.L14               ............................................................................................................................................................................... 5908
WRC.CI.97.97CI14            ............................................................................................................................................................................... 5826
WRC.CI.98.98CI04            ............................................................................................................................................................................... 5841
WRC.GM.05.Pbt_05GM_X02      ............................................................................................................................................................................... 4758
LST.CD.88.524               ............................................................................................................................................................................... 4774
LST.CD.88.447               ............................................................................................................................................................................... 4777
LST.KE.x.lho7               ............................................................................................................................................................................... 5858
SYK.KE.x.SYK173             ............................................................................................................................................................................... 5700
SYK.KE.x.KE51               ............................................................................................................................................................................... 5357
COL.CM.x.CGU1               ............................................................................................................................................................................... 5280
```

```
                                                Vpr start                                                                                                          Vif end
H1B.FR.83.HXB2          ....................ATGGAACAAGCC...........................................................................CCAGAAGACCAAGGGCCACAGAGGGAGCCACACAATGAATGGACACTAGAGCTTTTAGAGGAG 5633
Vif                     _.__.__.__.__.__.__. W  N  K  P__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__._ Q  K  T  K  G  H  R  G  S  H  T  M  N  G  H  *
Vpr                     ____________________M  E  Q  A__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__P  E  D  Q  G  P  Q  R  E  P  H  N  E  W  T  L  E  L  L  E  E
H1A1.UG.85.U455         ....................------------...........................................................................----------G-----------------GT--GC-----A-G--T-----T-G-----A--- 5079
H1B.US.90.WEAU160       ....................------------...........................................................................-----------------------------T--------------------------------A 5632
H1C.ET.86.ETH2220       ....................------------...........................................................................----------G-----G-------A---T-----------G--------------------- 5025
H1D.CD.84.84ZR085       ....................------------...........................................................................--G-C-----------G-----------T-------------T----------------- 5155
H1F1.BE.93.VI850        ....................------------...........................................................................----G-----G-----------------T-----------G-------AA----------- 4973
H1G.SE.93.SE6165        ....................------------...........................................................................----------G-----------A--A--CT-----------G--T----A--G-----A--- 5030
H1H.CF.90.056           ....................------------...........................................................................----------G---------------------------------------------------- 4980
H1J.SE.93.SE7887        ....................------------...........................................................................----------C-----------A-----T--C---------------------G------- 4947
H1K.CM.96.MP535         ....................------------...........................................................................----------G------------------A--------------T-----A----G------- 4829
H101_AE.TH.90.CM240     ....................------------...........................................................................--------..G-----------A--A--CT--------------T----A--G--------- 5201
H102_AG.NG.x.IBNG       ....................------AG----...........................................................................----------G-----------------GTC-------------------A--G-----A--- 5158
H1N.CM.95.YBF30         ....................-------G---A...........................................................................--------TGC--------------A--CT-T---------G-------AT-A-----A--A 5227
H1O.BE.87.ANT70         ....................-----------A...........................................................................--T--GA-T--G--A---GCT-AA-----CTT----------G----------C-----A--- 5688
H1O.CM.91.MVP5180       ....................------------...........................................................................-TG---A-T--G--A---GCT---------TT---------------------CC-G--A--- 5663
CPZ.CD.90.ANT           ....................------------...........................................................................--GC-G--TG-G-----------A--A---ATG-----G---TTGT----AAC------A--A 5090
CPZ.CM.01.SIVcpzCAM13   ....................--------C---...........................................................................-----G-------------CC--A--A---T-------G---G--T----AA-A-----A--C 5216
CPZ.CM.05.SIVcpzEK505   ....................--------G--A...........................................................................--------T-----A----------------T----C------G-TT----AT-G--G--A--C 5205
CPZ.CM.05.SIVcpzMB66    ....................---------CT-...........................................................................----------G--A-----A-----A---T-T-------------------A-A-----A--A 5179
CPZ.CM.98.CAM3          ....................--------G--A...........................................................................--------T--G--T----CA--A------T-----------G--T----A--GC-G-----A 5047
CPZ.GA.88.GAB2          ....................------------...........................................................................----------G--A---GCAG-A--A--TT-------G---G-------AA-A--G-----C 5026
CPZ.CM.98.CAM5          ....................---------ATA...........................................................................----------G-------CA--A--A---T-----------R-------AA-A-----A--- 5335
CPZ.CM.05.SIVcpzMT145   ....................------------...........................................................................-----G-----G--A---TCA--A-----CT-TG-GC-G---CTGT----CACA--G--A--A 5175
CPZ.TZ.00.TAN1          ....................-----------A...........................................................................--GA-C---A-C-----G-----A--A---T---C---G---CTGT----CA-CC----A--A 5295
CPZ.CM.05.SIVcpzLB7     ....................----------T-...........................................................................-----------------------------T-T-----G---G--T-----G-A-----A--A 5188
CPZ.US.85.CPZUS         ....................----------TG...........................................................................------------------GCC--A--A---T-------G---G--T----TG-C-----A--A 5693
CPZ.GA.88.GAB1          ....................-----------A...........................................................................-----G--------T----CA--A--A--TT--C-G------G--T-----ACC--G--A--A 5697
MON.CM.99.L1            ....................----TT.................................................................................................CCCAT-G-A-GAC-GGCCCT--AG-CCGCCA-TTG---- 5434
MON.NG.x.NG1            ...............................................................................................................................CGT-GTTCCCAT-G-GCGC-CCGCTCTG-AG-CCGCCA--TG---C 4045
GSN.CM.99.CN166         ....................------AG-AT-CCACCCAGTCATCCT............................................................ATGCCTTGG-TCA-C-G-AG-GT-CCCA--AC--TG-CTATCG-T-A-A-CGC-C-GTG---A 5426
GSN.CM.99.CN71          ....................------AG-A--CCACCAAGTCACCCT............................................................-T-CCTTGGATCA-C-G-AG-GT-CCCA--AC--TG-C-GTTG-T-A-AGCGCCA-GTG---A 5408
MUS_1.CM.01.1085        ....................------AGCTTGCCCCCCAGCCACCCA............................................................G-GCCCTGG--CTCCAG-GT-GTCCCAA-CAC--TGCGCCA-G---A---AGCCA-GTG---A 5382
MUS_1.CM.01.CM1239      ................................................................................................................-TGG--CTCCAG-GT-GTCCCAA-TAC--TGC-GCA-G-C-A-C--GCCA-GTG---C 5372
MUS_2.CM.01.CM1246      ................................................................................................................-TGG-T-A-T-GT--A-C-CCAA--ACT-TG-TTCA-G-G-A--CCGCCC-GTG---A 5489
MUS_2.CM.01.CM2500      ................................................................................................................-TGG-T-A-C-GC---GTACCAA--AC--TG-CGCA-G---A-A-CGCCA-GTG---T 5442
RCM.NG.x.NG411          TTCCGGGAATG...TAA.........................................................TCAAGCACTCCTCTCATTGCAGAGATGGAGATGCTC--G-----TG---AA--------A--A--CT-TG--------CT-A-G--TACCC-TATA--- 5459
RCM.GA.x.GAB1           TACCGGGGATG...TAA...........................................................TGATGGCCCTCTCTTTACAGAGATGGAGCTGCCC--T--G--TG----T--G--AC-A--A--TT-TG--------TT-A-G--TACCC--ATA--- 5424
DEN.CD.x.CD1            ....................----C-...........................................................................................................--AA-CACTC-AAGGCA-CT--AG--AGCAG--TG---T 5540
                                  Vpx end   Vpr start
MAC.US.x.239            CTCCAGGACTAGCATAA..---................................................................GAAGAAAGACCT------A-TG----A-----A-----A---TGGG--------GT-G-G---G--C-G--A--A 6228
Vpx                     P__P__G__L__A__*_
Vpr                                         _M__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__._.  .  .  .  .   E  E  R  P  P  E  N  E  G  P  Q  R  E  P  W  D  E  W  V  V  E  V  L  E  E
H2A.GW.x.ALI            CTCCAGGTCTAGTCTA...---..........................................................ACTGAAGCACCAACAGAGTTTCCC--G-CC-GAATG--A--C--CCA--GAG-CAGGG----G---GT-A----AG-CC-GAGA--A 6318
H2A.DE.x.BEN            CTCCAGGTCTAGTCTA...---..........................................................ACTGAAGCACCAACAGAGTTTCCC--------TGGGACC----G------CTT-GGG-G---C---GT-A----AAC-C-GAG---A 6328
H2A.SN.x.ST             CTCCAGGTCTAGTCTA...---..........................................................ACTGAAGCACCAACAGAGTCTCCC--G--G--TAGGACC----C----------GGGG----G---GT-A----AACCC-GAGA--- 5770
H2B.GH.86.D205          CCCCAGGCCTGGCCTA...---..........................................................GCAGAAGCAGCCCCAGAGATCCCT-----GA--G-GAAC-----A--A--A--GTGGG-A--G----T-GGG---A-CC-G-----A 6301
H2B.CI.x.EHO            CCCCAGGTCTAGCCTA...---..........................................................GCAGAAGCAGTCCCAGAGATTCCT-----G--TA--AAC-----A--A--A--GTGGG-AC-G---GT-G-G--CG-CC-G-----A 6299
H2G.CI.92.ABT96         CTCCAGGTCTAGCATAG..---...................................................................GCAGAAGAAATCCCC--------TG-G-CC--C-----A------TGGG--------GT-G----AG--C-G--A--A 5644
H2U.FR.96.12034         CACCAGGATTGATCTA...---......................................................................GCAGAAATACCT-----G---GGG-CC-----A--A-----TGGGG-C--G---GTGAG---AA-AC-G-----A 5788
MAC.US.x.EMBL_3         .................................................................................................................................................................... 5467
SMM.US.x.H9             CTCCAGGACTAGCATAA..A---.....................................................................ACAGAAAGACCT--------TG---CC-----------A--GTGGG--------GTRG----AG--C-G-----A 5702
SMM.US.x.PGM53          CTCCAGGACTAGCATAAA.A---.....................................................................ACAGAAAGACCT--------TG---CC-----------A--GTGGG--------GT-G--A-AG--C-G-----A 6147
SMM.US.x.SIVsmH635F_L3  CTCCAGGACTAGCATAAA..---.....................................................................GCAGAAAGACCT--------TG---CC-----------A---TGGG--------GT-G-G--AG--C-G-----A 6232
SMM.SL.92.SL92B         CACCAGGTTTGGCCTA....---.......................................................................GAACATGCT-----G--TG-GACCAAC-CA--A--A---TGGG----G----T-AG---TG-AC-G--A--A 5661
STM.US.x.STM            CTCCAGGCTTAGCATAG..A---.....................................................................ACACACAGACCT-----G--TG----C--------A--A---TGGG----G---GT-G-G--AG--C-G-----A 5876
MNE.US.x.MNE027         CTCCAGGACTAGCATAA..---......................................................................GAAGAAAGACCT--------TG----C-----A-----A---TGGG--------GT-G-G---G--C-G-----A 5696
DRL.x.x.FAO             CTCCTCCTCCA...TGA.....................................................TTGGTCTAATAATTGTGTCTTACAGATGGAAAGGCCA--G--G--TG----A--G-CA-----A--CT-TG----G---CT-G----TAC-----TA--A 5527
MND_2.x.x.5440          CTCCAGGGATGGCTTAA.............................................................TTTCTCTGTTTTACAGATGGAACAGCCA--T--G--TG-G-CT--C-----A-GA---T-T---A-----CTGG---GTAC---G-CA--A 5473
MND_2.CM.98.CM16        CTCCAGGGATGGTCTAA...............................................................CCTCTGTGTTACAGATGGAGCAACCA--C--G---G-G-CC--T--A--A--A---T-----------CT-A-T--CACC----CA--A 5908
MND_2.GA.x.M14          CTCCAGGATTGGCCTAA...............................................................CTTCTGTCTTGCAGATGGAACAGCCG-----G--TG-G-CT-----A--A--A---T-T-----G---CTGA----TACA----CA--A 5835
OLC.CI.97.97CI12        ....................----C-.....................................................................GAAAGAGTT-----G--T--G-------CC----------TGG-G--T---GTGAG---CACAA-G--A--A 5437
GRV.ET.x.GRI_677        ....................----C-TC--GAAGAGATCCAAGAGAA....................................CCTTTACCAGGATGG-TG---AT-TGG-AT-T-G-C---------TGGG-C------CT--A---CA-GC-CAG---T 5787
VER.KE.x.TYO1_patent    ....................----CCTC--GGAGA.....................................GATCCAAGAGAGGCAAGACCAGGAGAAGT----ATTTGG-AT-T-AGC---------GTGGG-C------CT-AG---TA-G---C-A--T 5850
VER.KE.x.9063           ....................----C-TC--GGAGA.....................................GGTCCCAGAGAAAACAGGCCAGGAGAAGT---GAT-TGG-AT-T-A-C--A--A---TGGG-C------CT-AGG--CA-G-----A--T 5850
VER.KE.x.AGM155         ....................----CCTC--GAAGA.....................................GATCCCAGAGAGGAGAGACCAGGAGGA-T----ATTTGG-AT-TCAGC---------GTGGG-C------CTCAG---TA-GG----A--A 5845
VER.DE.x.AGM3           ....................----CCTC--G-AGA.....................................GATCCCAGAGAAGCAAGACCGGGGGAA-T----ATATGG-ATTTGAGC---------GTGGG-C------CT-AG---CA-GC----A--T 5348
TAN.UG.x.TAN1           ....................----C-G---GAAGA.....................................GATTCCAGGGAAAGAAGACCCGGATGGTT----ATTTGG-AT-TCAGC---------GTGGG-C-----ATT-AG---CA-GG---CA--A 5843
SAB.SN.x.SAB1C          ....................----CCTC--GAGGGTGGCTCCCACCAGTAGGAGGGGATCCTCCCAAGGATCCCCCCAAGAATCCCAGAGAAGAGATACCAGGATGG-T----ACATGG-ATTTG-CC--A-----GTT-G-C------CTC-G---CA-G---C----C 6060
MND_1.GA.x.MNDGB1       ....................---...................................................................GGCCAGAAGAGAGATGAGCAAGTAT--------T-----A--T-CC--A------T-----C-G---CT-GC---TAC-A-G-----A 5316
DEB.CM.99.CM40          ....................------AGGTA-CCACCCTCACACCCA...........................................................---C-TTT-ACCTCAAG-ACA-TACCCATGACA-GGCT-G--TTG-AGC--GCCA-GC-A--T 5417
DEB.CM.99.CM5           ....................------AGGTATCCACCCTCACACCCA...........................................................---C-TTT-AC-TCAAG-ACAGTACCAATGACA-GGCT-GCACT--A-C-AGCCA-GC----T 5411
TAL.CM.01.8023          ....................------AGGTTG.................................................CCACCCAGTCATCCT-TGCC-ATAAC-TCAAGG-TC-TACCAA---CA-GA-CTCAATT-GA-AGA--CA-GAGA--A 5105
TAL.CM.00.266           ....................------AGGTTG.................................................CCACCCAGTCATCCT-T-CCGCTGACCTCCAG--TC-TTCCAA---CA-GGCCTCA-CT-GA-AG---CA-GAGA--A 5607
SUN.GA.98.L14           ....................----CCAGC..........................................................AGGAGAGAGCCAGTAGAACAGCCA---------G-G-------C---A------TTTGCAC-----CTGGC---TACCA-G--A--A 6004
WRC.CI.97.97CI14        .............................................................................................................................GGCGCCTT--GGAG-ACCT----CA-T-CC-GA-GAC--TCTAGG-A-GC-A--- 5881
WRC.CI.98.98CI04        .............................................................................................................................AGC-CCTA--GGAG-ACCT----CA-C-CC-GA-GAC--TCTAGG-A-GC-A--A 5896
WRC.GM.05.Pbt_05GM_X02  ............................................................................................................................-AGG-CCCCAA--GGAG-ACCC-AGGCA-TACC-GA-GAT--TCTAGG-A-GC-A--- 4816
LST.CD.88.524           ....................---TCCAG-CAAAGA..........................................CAGCAGCGCAGCCCAGAAAGGCCC-----G---TTT--T--T-CA--A------TGGGC-C-----CT-GC---CACAA-G-----A 4876
LST.CD.88.447           ....................---TCCAG-A-AAGG..........................................CAGCAAAGGAGTCCAGAAAAACCT-----G---TTT-------CA---------TGGGG-C-----CT-GC---CACCA-G-----A 4879
LST.KE.x.lho7           ....................---TCCAGGCAGAGG..........................................CAGCAAAGAAGTCCAGAGAGACCA-----G---TTT--A----CA---------TGGGG-C-G---TT-GC---CACCA-G-----A 5960
SYK.KE.x.SYK173         ....................----CGG-G--TTTC......................TTCAATCCCTCCCAGCATGTGCAAGGAACACCATGGTTCTTCATTCCAAG-AATGT-G--TT-ACC---A--GTGATCAATGTCACT-TAAAGGC---ACT- 5817
SYK.KE.x.KE51           ...................G----C---G--TTTC...................TTCAGAGTTTACCAGCGGCAGCCAGGGGAACCACTATTTTACATTCC-AGAA-T-TCAATTG-GAC---G--GGAGTAAT-A-TG-CACT-CCAAAGC---AC-- 5478
COL.CM.x.CGU1           ....................---AG-GTCCTG.............................................................................................A-AGGGACTTCAC-.---TTG-AG--C-AGG-----A-C 5330
```

frameshift insert in HXB2

```
H1B.FR.83.HXB2          CTTAAGAATGAAGCTGTTAGACATTTTCCTAGGATTTGGCTCCATGGCTTAGGGCAACATATCTATGAAACTTATGGGGAT.......................ACTTGGGCAGGAGTGGAAGCCATAATAAGAATTCTGCAACAACTGCTGTTTATCCAT..TTTCAG 5777
Vpr (frameshifted)                                                                                                                                                                           F R
Vpr                     _L__K__N__E__A__V__R__H__F__P__R__I__W__L__H__G__L__G__Q__H__I__Y__E__T__Y__G__D__.__.__.__.__.__.__.__.__T__W__A__G__V__E__A__I__I__R__I__L__Q__Q__L__L__F__I__H_##_F__Q_
H1A1.UG.85.U455         ------C-----A-------------------CAG----------A-----A-----------A-C--C---------.......................--------AG----T-----T-----------T--------T-A-------T---..------ 5222
H1B.US.90.WEAU160       -------G----------------------G-A---------A------A------------------------------------.......................-------A--------------------C------------------T---..------ 5775
H1C.ET.86.ETH2220       --C---C-G--------C-----C--------ACCA---------AA------A---T------------C------------.......................------T-G-----A-----TT--------C-----------AA-----------..------ 5168
H1D.CD.84.84ZR085       -------G---------C-----------------A-----T---A------A---T-----------------------------.......................--C----------T-----T----------------------A--------T---..------ 5298
H1F1.BE.93.VI850        -------------------G------------CCA-----T--C--------C-----------A-C--C--------C.......................--------AG----T-----T--------G--AT---------A-----------..------ 5116
H1G.SE.93.SE6165        -----A------------------------C-------T-----------A--G---------A-C-----------------.......................--------A------T----------------A--A--------A-----------..------ 5173
H1H.CF.90.056           A----------G--------G----------AG-A---------CAA-----A--G---------A-C--C-----A---.......................--------T------T-----TT---------CG----------A--------T---..------ 5123
H1J.SE.93.SE7887        -----A-----------------C---------CCA-----T---A-T-----A--GT--------AGC--------A---.......................--A----A-----C-------------------A----------AT-------T---..------ 5090
H1K.CM.96.MP535         -------GA----------------------CCA---------AATC-G--A--G---------AC-------------------.......................--------A----C-T-----T-----------AT----------A--------T---..------ 4972
H101_AE.TH.90.CM240     -----A-------------------------CCC---------A---------A--GT-C------A-C-A-----------.......................--------A---G--T-----T-----------GT----------A------G-T---..------ 5344
H102_AG.NG.x.IBNG       -----AC---------------C---------CCG------------A-----A-------------------A----------.......................--------A---G--TA----T-----------T----------AT-----G-T---..------ 5301
H1N.CM.95.YBF30         T-A--A-----------GC-C-------A------------A-----G-----A-----C-------A-C--A-----A--C.......................--C----AG--G--A--G--A--T--C--G--A--A-------T-A-------------...-AT-- 5370
H1O.BE.87.ANT70         --A--AGCA-----A--A---------C-------CC------A--C-C---G--A---T-C--T-----G-----------C.......................--------T------TATG--A--T---------CT-A------T---------C-----...-AT-- 5831
H1O.CM.91.MVP5180       --G--AG-A-----A--A---------C-------CC----T-A--A-C--GT-----GT-C--T-----G---------A--C.......................--------A------TATG--A--T---------CT-A---------A-------C-----...-AT-- 5806
CPZ.CD.90.ANT           A-A--A-----G--A--A---------C--ACA-CC-ACAT-A--ACA-C-T--AA-TTGGG-A----C--A---C--A---.......................T-A---AG-------C-G-TTT-G----TTC-C--T---A--GCT--------CA----...-AT-- 5233
CPZ.CM.01.SIVcpzCAM13   T-A---------------A-A---C-A---C--A-G------T--CCAA-----C--G-----T---ACC--C---------C.......................--G---A---G----G--A--T--------C-A--------A--------G-T-----..---T-- 5359
CPZ.CM.05.SIVcpzEK505   --A--------G---C-GC-C--C------C--CC------A-----AC---------T-CT-----A-T--A------A---.......................--C----AG-----A--G-----C--T--G-CA--A----------T--------A----...-AT-- 5348
CPZ.CM.05.SIVcpzMB66    -----A-GG-----A--A-----C--C---C-AGA-----------CAGC-----A--G--------ACC--------A---.......................----------G----G--T---------C--------------T--------T-----...-A--- 5322
CPZ.CM.98.CAM3          --A---------------AC-C---------C-AGCA-----A-----A-----A--GTT---T-----C--------A---.......................--A--------T--A-----A--C--T-----A-------G--T-------------..---T-- 5190
CPZ.GA.88.GAB2          T-A--A--C--G-----C--------------ACCA-----A---CAG-----A---TTA-----CA-T-A------A---.......................--A----A------A--G--------C----A--T----CT--------G-G-------..---T-- 5169
CPZ.CM.98.CAM5          --A--A-C---G---C-A-------------C-CCC------A---A-AC----A--GT-C-------------------A---.......................--A---AT---G--A-----A--T--T-----AT-------T--T-------G-------..---T-- 5478
CPZ.CM.05.SIVcpzMT145   A----AC-A-----A---------C------C--ACCC---T-A----CAC----AG-------T---------A-----A--C.......................T-C-----G-----AC----T-----------T-A-------G---------------..------ 5318
CPZ.TZ.00.TAN1          A-A--AC-A-----A--G-A---C------A--ACCAATAT-A--G--GG----AA-TTGGG---TCACC-T-------A--C.......................T-C----AG-----AC-G-AAT----C-AG--CT----GAG-GCTT-------C---C...-ATC- 5438
CPZ.CM.05.SIVcpzLB7     --C--AC-G-----A--A-----C--------ACCA---------CCAA--G--A---T----A-----C--C------A--C.......................--C----AG-----TC----T--C-------C-----------C---------T-----...-A--- 5331
CPZ.US.85.CPZUS         --A--AG-A--G---T-GC-C------------ACCC-----A--A---------A--GT-C------CAGT---------A--C.......................-------AG-----A-----A--T-----G--CT-A-------TT------------...-AT-- 5836
CPZ.GA.88.GAB1          T-A--A----------A--A---------C-----ACCA---T-G---CAA-----A---TTC--T-----C---------A---.......................--A----T-----A--G-----C---------CT-A------CT--------------..---T-- 5840
MON.CM.99.L1            --A--CG-A-----CC-A-A---C--CT-A---GAAGA-----G----A-TT--G----GG-AACG--GCTGCC--CT--CCCGGCCTGGAATGCAGACCAGGCATGGGCA---T-T-CCATT-ATTAT-CT---TGGG-C---ACCA-C--T-A--GG--C...-A--- 5601
MON.NG.x.NG1            --A--CG-G--G--CC--TYG--C--CT-A-A-GAAGA-----G----A-CT--G-T--GG--AC----TTGCC--CC--CCCAGGGTGGAACGCAGACCAGGCATGGGCA---T-Y-CTATT-ATTAC-CC--GCGGG-----AC-A-RA-A----GG--C...A-G-- 4212
GSN.CM.99.CN166         A-A--CG-G--G--A-AG-A---C--CT-A---GACGA-----G---AA-TT----TG--G-GACG--GCTCCC--CC---CCAGATTGGACAGTGGATCAGGCAG-AATA---T-T-CAAT--ATTAC--C...-GGG-----AC---T--C----GA--C...--T-A 5590
GSN.CM.99.CN71          A-A--CG-A--G--A-AA-A---C--CT-A---GAAGA-----G---AA-TT----TG--G-AACG--GCTCCC--CC---CCGGATTGGACAGTGGATCAGGCAG-AATA---T-T-CAAT--ATTATG-C--G-GAG----GAC-T-A--C----GA--C...--T-- 5575
MUS_1.CM.01.1085        G-A--CG-A--G--A-AA-A---C---T-C--AGAAGA-----G---AA-TT--A-TG-GG-AACG--GCTGCCA-CC---CCTGATTGGACAGTGGATCAGGCGG-AATA---T-T-CCAT--A-TATG-T--G-GAG-----AC---TG-C----GG----...------ 5549
MUS_1.CM.01.CM1239      --A--CG-A--G--A-AA-A---C---T-C--AGAAGA-----G---AA-CT--A-TG--G-AACG--GCTACCA-CC---CCTAATTGGACAGTGGATCAGGCTG-AATA---T-T-CCAT--ATTAC--T--G-GAACA---AC---C--T--C-GG----...-ATC- 5539
MUS_2.CM.01.CM1246      --A--CG-A--G--A-AA-A---C--CT-C--AGACAA-----G---AA-TT--G-TAGAG--ACA--GCTCCCA-CA---CCACAATGGTCTGTAGATCAGGCAG-CATAA--T-T-CCATT-ATTAC--T-A-TCAG-C---ACC--CT----C-GA--C...-AT-- 5656
MUS_2.CM.01.CM2500      --A--CG-A--G--A-AA-AG--C--CT-C---GACGA-----G---GA-TT--G-C-GAG--ACG--GCTCCCA-CA--CCCTAATTGGACTGTTGATCAGGCAG--ATA---T-T-CCAT--ATTAC--T---TGGG-A--GGGT--C--C--C-GG---...----C- 5609
RCM.NG.x.NG411          A-AC-AG-A------AAG-A-------A-ACCTGAA-T---TACACAAGA---AA-TT----A-----GCAGC----A---.......................-G--TA-AG-----AA----T--G----CCT-G--TA-TA--GCA--T--CT-G--C..------ 5602
RCM.GA.x.GAB1           T-GC-AG-A------AAG-A-------A-ATATGC--T---AACGCAAA----AG-TT--G-A-----GCAAC---------.......................-GCATA-AG------C-G--A--G--TC-GC-G--T---AG-GCCT-----C-T---..---T-- 5567
DEN.CD.x.CD1            A-C-CAG-G--G--AAGA-A---C--CAGC-AAGAAGA-A--AC---AA--T--G-T--CTG--GGTC-TTACCA-C-CTCCCACATTGGACAGAAGGACAAGTC-TGGCA--C-CC--CAT--ATT-C---------CA----GA--GAAA-A-GG-AG---...-AT-- 5707
                                                                           Tat exon 1 start
MAC.US.x.239            --G--AG-A-----T-A-A-------GA-CCTCGC-T---AAC--CAC-T--TA-T---------A-T-GAC----A--C.......................--CCTT-AG----CA-G--AAC-C--T-----C--C----G-GC---C--C--G---..------ 6371
Vpr                     _L__K__E__E__A__L__K__H__F__D__P__R__L__L__T__A__L__G__N__H__I__Y__N__R__H__G__D__.__.__.__.__.__.__.__.__T__L__E__G__A__G__E__L__I__R__I__L__Q__R__A__L__F__M__H__.__F__R
Tat exon 1                                                                                 M__E__T__.__.__.__.__.__.__.__.__P__L__R__E__Q__E__N__S__L__E__S__S__N__E__R__S__S__C__I__.__S_
H2A.GW.x.ALI            A-A--AG-A------T-A--G------GACCCTCGCAT---AAT--CTC-T--CGGCT---------ACT-GAC----A--C.......................--CCTT-A-A---CCAGG-AGC-C--T-ATGCC--A----G-GCC--C--C--G--C..------ 6461
H2A.DE.x.BEN            A-A---G-A-----CT-A----------GA-CCCCGC-T---AAT--CTC-T--CT-CT-----C--A-T-GAC----A--C.......................--CCTT-A---C-CCAG--AGC-C--T-A--CC--A----G-GCC--C--CG-G--C..------ 6471
H2A.SN.x.ST             A-A--AT-A------T-A-AG--C---GACCCTCGC-T---AAT-ACTC-T--CA-CT---------CT-GAC----A--C.......................--CCTT-A---C-CCAG--GGC-C--T--G--C--A----G-GCC--CC-CT-G--C..------ 5913
H2B.GH.86.D205          A-A---C-A-----CT-A-AG------GA-CCTCGC-T---AAC--CGC-T--TA-CTT------CAGT-GGC----A---.......................--CCTT-------CA-G--AGC-C--T-A---C--C----G-GCC--C--CC----C..------ 6444
H2B.CI.x.EHO            A-A--AC-A-----CT-A-AG------GACCCCCGC-TAT-AAC--CAC-T--CA-TTT------CA-T-GGC----AA--.......................--CCTT-A-----CA-G--AGC-C--T-A-C-C--C--G-GGGCA--C--CC----C..------ 6442
H2G.CI.92.ABT96         A-C--AG-A-----CT-A-A-------GACCCTCGC-T---AAC--CGC-T--TA-CT----------T-GGC----A--C.......................--CCTC-AG----CA-G--AAC-C--T-A---C-----GAG-GCT--C--CC-G--C..--TC- 5787
H2U.FR.96.12034         --A--AG-A--------A--G------GA-CCTCGC-T---AAC--CGC-T--TA-CT---------A-T-GGC----A---.......................--CCTT-A-----CA-G--AGC-C--T-A---C------AG-GCAA-C--CC----C..------ 5931
MAC.US.x.EMBL_3         ..........................................................................................................................--C----G-GC---C--C--G---..---T-- 5496
SMM.US.x.H9             -KA--AG-A-----CC-G-AC--K---GA-CCTCGC-T---AATY-CGC-T--TA-CT-----------T-GAC----A--C.......................--CCTT-A-----CA-G--AGC-C--T-----C--C--GGG-GCA--C--C--------..------ 5845
SMM.US.x.PGM53          G-A--AG-A------C-G-A---C---GA-CCTCGC-T---AAC--CGC-T--TA-CT-----------T-GAC----A--C.......................--CCTT-A-----CA-G--AAC-C--T--G--C--C----G-GC---C--C--A--C..------ 6290
SMM.US.x.SIVsmH635F_L3  A-A--AG-A-----CC-G-A-------GA-CCTCGC-T---AAC--CGC-T--TA-CT----T-----T-GAC----A--C.......................--CCTT-A-----CA-G--AAC-C--T-----C--A--GAG-GCA--C--C--------..------ 6375
SMM.SL.92.SL92B         T-G--AG-A--G--AC-G-A-------GACCCTCGC-T---TAC--CT-----AA-TT-CG-A-----T-----C-------.......................--AATA-A---G-CA-GG-AA-----T-AG--AT-A---A-GGCAT----CT-A--C..---T-- 5804
STM.US.x.STM            A-A--AC-A-----CC-GC-G------GA-CCTCGC-T---AAGC-CGC-T--TA-CT-----------T-GAC----A--C.......................--CCTT-A-----CA-G--AGC-C--T-AG--C--C--G-G-GCC--C--C--G--C..------ 6019
MNE.US.x.MNE027         --G--AG-A------T-A-A-------GA-CCTCGC-T---AAC--CGC-T--TA-TT-----------T-GAC----A--C.......................--CCTT-AG----CA-G--AAC-C--T-A---C--C----GGGC---C--C--G--C..------ 5839
DRL.x.x.FAO             A-AC-AG-G------T-G-AG------GA----CGC-TAT-G---AATG-----TC-TGGG---------GA-----A---.......................T-A-TA-A-------C-CCAGC-T----T-C-CT-----AG-GCA--T--CT-G-----..---T-- 5670
MND_2.x.x.5440          A-AC-AG-G------T-A-AG-----CGA----CGC-T---A----CAG----CTC-TGGG-G-----GCAAC-A--A--C.......................----TA-A---T--CC--CAGC-------C---T-----AG-GCA-----CT-A--C..---T-- 5616
MND_2.CM.98.CM16        G--C-AG-A------T-A-AG------GA---ACGC-TAT-G----CAG----AAGCTGG-----CA----C---------.......................--A-TA-A---G---CG-CAGC-T--CTC-------T---A--GCT----C---G--C...-AT-- 6051
MND_2.GA.x.M14          A--C-AG-A------T-G-AG------GA----CGC-TAT-G----CAG----AAGCTGG--T--------A--------C.......................--C-T--AG--G---C-GAAGC-T--C-CCC-CT-A--GAG-GCT--A---C-G-----..---T-- 5978
OLC.CI.97.97CI12        A---TTGTA--G--AAAA--G--C-----C---GAG-TAT-GGTCC-AG-TTTT-T-TTCTGTGC-C---GA.....................CATATCAATGATGAGAT-AGTACCTCTTTGCTGT-C--T--GC-A--A---AG-GCCT-----CAA----...-AT-- 5583
GRV.ET.x.GRI_677        --A--CG-A-----CAGA--G--C---GGA-T--ACAT---AATCC-AG--T--A-TT-CTGTGTA--GGAGGGAA--AGA............CATAATACCCCATGGAAT-AGAT--GCT-CAAGTACTAT------G-T---A-GTCTA-----G-A-----..------ 5942
VER.KE.x.TYO1_patent    --C--CC-G--G--GAGGCTG------GGG--AGAACTC--GTTCCAAG--T--A-CT-CTGTC-G--GGAAGGA-AAAGA............CATGGTACTCCC-TGAT--A-A-G-CCT-CAAATATTAT--GC-AG-A---A-GGCT--C---G-G-----..---TC- 6005
VER.KE.x.9063           --C--TC-G-----CAAATT-------GGC--AGAGCTC---TTCCAAG-TT--A-TT--TGTC-G--GGAAGGA-A-AGA............CGGGGAGCCCCA-TGAC--AGA-G-CCT-TAGATACTAT--GT-GG-T--GA--GCA--C---G-T-----..----C- 6005
VER.KE.x.AGM155         A-C--C--C-----CAAACTG------GGCC-AGAACTC--AT-CCAAG--T--A-TT--TGTC-G--GGAAGGG-A-AGA............CAGGGAAGACCC-TAGC--A-A-G-CAT-TAAGTATTATC-CT-AG-T--GA--GCT--C---G-G-----..----C- 6000
VER.DE.x.AGM3           A-C--TC-G-----CAAG-TG------GGGC-CGAGCTC--GTTCCAAG--T--A-CT--TGTC-G--GGAGGGA-AAAGG............AATCGCACTCCC-TGCTA-A-A-G-CTT-TAAATATTAT-A-T-GG-----A--GCT--C---G-G-----..----C- 5503
TAN.UG.x.TAN1           --C--TC-G-----ACAAC-C------GGG--AGAACTC---TTCCAAG--T--A-CTTCTGTC-G---GAAGGG-AAAGG............AATGGGGCACCA-TGATA-A-A-G-CTT-TAGATATTAT--GT-GG-A--GA-GGCTT-A---G-G-----..---TC- 5998
SAB.SN.x.SAB1C          --C--TTCA--G--CCAGT-C--C--C--A----A-CTC--TTTCC-TC-TT--TGGA----TGTG---GAACCA-CCAT-..........GATCATGGGCAAACC-GGCT--A---CTG-T-TAAATATTGT------------A-GGCTT-A---G-G---...A-G-A 6218
MND_1.GA.x.MNDGB1       A-A---G-A-----AAGA-AG--C--C---CTC---ATC--AA---CAG--TCAG--T--TGTGTGC---ACACA---AG-GAG.....................GAAGA---CT-T-A-A--TTT--T-CC...T-AA--A-TAG-GCCA-T-GGG----C...C-AGC 5459
DEB.CM.99.CM40          --A--CG-G-----CC-A-A------CA-C--AGAAGAA--ATT---AG-TT--A-T---TGTGTA--CTTACCC-CCC-ACCAGATTGGACGGGAGAACAAGCATGGGCT---TCG--CATT-A-TAC----A---AG-T---A--A--T-A-GGT-G--C...C-G-- 5584
DEB.CM.99.CM5           --A--CG-G------T-A-A---C--CA-A--AGAAGA-T-ATGG--AG-CT--A-T--CTGTGTA--CCT-CC--CCC-ACCAGATTGGACAGGAGAACAAGCATGGGC---GTC---AATT-A-TAT----A---AG-A---AG-A----T-GGC-T----...C-G-- 5578
TAL.CM.01.8023          G-A-G-G-C--GATGC-C-T---C--CA-A--AGAAGA-G--AT---AG-TT--A-T--CTGTGTA--GTTACCA-CA--GCCTGATTGGACAGGCGAACAAGCATGGGCT--TTC---CAT--ATT-TGCTC-G-AGGGTA-CG-GA--T-A--CC-A---..----CA 5272
TAL.CM.00.266           G--CGAG-C---ATGC---T---C--CA-A--AGACGA-G--AT---AG-CT--A-T---TGTGTA---TTGCCA-CA--CCCTGATTGGACAGGAGAACAGGCATGGGCT--CTC---CAT--ATT-TGCT--G-AGGGTA-T---A--T-AC-CC-A---..----CA 5774
SUN.GA.98.L14           A-C-G-G-C-----G-AA--G-----C--AGCAGCAATAG-TGCCCAAG-CTT----T-CTGTGCG.......................GACTCTACAGGGTCAGAAAA--A---GTGTATGAGAGC---C-CTT-G--CA-TAG-GCCT----C--T---...C-GCC 6147
WRC.CI.97.97CI14        --CTTTGCAATG--AAAACAG-----C--A---GAAGTC--TGTAA-AG--TTT-TTT--TG--C-C--.....................CAGAGAGAGTATGAGGAAGTCC-TACTA--C-TTATT-CCAC--GG-A------AG-GG---C--C--G---...C---C 6027
WRC.CI.98.98CI04        --CTTTGCAATG--AAAACTG-----C--A---GAAGTC--TGTAA-AG--TTT-TCT-CTG--CGC--.....................CAGAGAAATTATGAAGAAG-TC-CACTA--C-TTATT-TCAC--GG-A--A---AG-GG---C--C--G--C...C-T-C 6042
WRC.GM.05.Pbt_05GM_X02  --GTTTGCCATG--AAAAC-G-----C--CC--GAAGTC--TGTAA-AG--TTT-TCT--TG--CGC--.....................CAGAGAAATTATGAAGAACA---CACTA--C-CTATT-TCAC--GG-C--T---AG-GGAT-A-A-C-A--C...C---T 4962
LST.CD.88.524           A-C---G-A--G--AAGAC-C-----C--AGAAGCAATTT-GGCCCAAG-CTC-G--T-CTGTGTC.......................AGCAGTACCGGGTCAGAAGA--AG-CTTGCATGAGAT-C--C-CTT-AG-TA-TAG-GCT--C-A-T-A---.....CT 5016
LST.CD.88.447           A-A--AG-G--G--CAGAC-C-----C--AGAA-CAATTT-GGCCCAAG-CTCAG--T-CTGTGTC.......................AGCAGTACCGGGTCAGAGGA-AA--CCTGCATGAAAT-C--T-CC--AG-AA-TAG-GCA--C-----T---.....TT 5019
LST.KE.x.lho7           A-A--AG-G--G--AAGAC-C-----C--AGAAG-AAT-T-GGCCCAAG-CTC----T--TGTGTA--C-G-ACA--........................TCAGAAGA--AG-CCTGCATGAAAT-C--C-CCC-AG-CA-TAG-GCTT-----C-T--C...C--CC 6103
SYK.KE.x.SYK173         G-AGT--CA-----ATCC-AG------A--CCTCAAGAAA-AT----AG-TT--A-T--ATC--TAA-CGAAG-G-CA-GGACTGAT...TCCCCGACAATGGCTTGGGAAAG-ACCA--TTG-A---GG-----GCA--CA-T-TTA----C---GAA-----..--TGC 5981
SYK.KE.x.KE51           G-AGCCC-A-----ATGC-AG------A-AGC-GAAGAAA-ATT---AG-TT--A-T--ATG--TA--GGTGG-A-CA-G-CCAGAAGAAACACCCACCATGGCTTGGATTAG-T-CA--CT--AT--GCAT---GCA--CA-CTTTA----T---GAA-----..--TGC 5645
COL.CM.x.CGU1           A-A-CCCGG--G--CAGA--G--C--C---TT--C-CACA--TCA-TGA--AT-TGGGG-TGTAC-AG--GGC-CT---GGAGA.....................C-G---A-G-CT-GTCTG--TTATG-G-----CA--G-GTGTGCC--C-ACT-G--C...G-TCA 5476
```

```
                              Vpr premature end (HXB2 only)                             Tat exon 1 start  Vpr end
H1B.FR.83.HXB2              AATTGGGTGTCGACATAGCAGAATAGGCGTTACTCGA...............CAGAGGAGAGCAAGAAATGGAGCCAGTAGATCCTAG...................................... 5850
Vpr (frameshifted)          __I__G__C__R__H__S__R__I__G__V__T__R__.__.__.__.__.__Q__R__R__A__R__N__G__A__S__R__S__*
Tat exon 1                                                                                         _M__E__P__V__D__P__R##__.__.__.__.__.__.__.__.__.__.__.__
Vpr                         _N__W__V__S__T__*_
H1A1.UG.85.U455             ---------C-A-----------------A---T--C-...............GG---A----GC--G-------------------A...................................... 5295
H1B.US.90.WEAU160           -----------------------------A------C-...............-----A---------------------------A---.................................... 5848
H1C.ET.86.ETH2220           ------A--C-AG---------------A--TTA---...............-----A---------------------------A...................................... 5241
H1D.CD.84.84ZR085           -----------A-----------------A------C-...............-----A-------------T-----------A...................................... 5371
H1F1.BE.93.VI850            ---------C--T---------------A--GT--C-...............-----A---T---------------------...................................... 5189
H1G.SE.93.SE6165            ---------C-A-----------------A----A-C-...............-G---A--G-T---GG-----C--G---------A...................................... 5246
H1H.CF.90.056               ---------C-A---------------AA----------...............-----A---T---------C-----------A...................................... 5196
H1J.SE.93.SE7887            ---C-----C-AT---------------A---TA-C-...............--A--A----GG--G-------------------A...................................... 5163
H1K.CM.96.MP535             ---------C-AC----------------A---TA-C-...............-----A----GG--------T-----------A...................................... 5045
H101_AE.TH.90.CM240         -----------A-----------------A---TG-C-...............GG---A----GC--G------A-TG---------A...................................... 5417
H102_AG.NG.x.IBNG           -----------A-----------------A---T--A-...............GG---A----GC--G--------TG---------...................................... 5374
H1N.CM.95.YBF30             G-----C--C-AG--C-----------GA-C----CT...............--A...--GAG---G-----A-----------...................................... 5440
H1O.BE.87.ANT70             ------A--C-A------T--------AA---AC-C-............AGAGGA--AG--AG---------T-----------GA...................................... 5907
H1O.CM.91.MVP5180           ------A--C-A------T--------AA--CTC-C-...TCTAACACAAGAGGA--AG--AG---------T-----------GA...................................... 5891
CPZ.CD.90.ANT               -CAC--A---GCT--------G-----ACA-GGA-CC.................................C-----CG--GA...................................... 5282
CPZ.CM.01.SIVcpzCAM13       -G-------C-AC--------------TA-A-TA-C-...............--A------AG-GC-----------A----C---GA...................................... 5432
CPZ.CM.05.SIVcpzEK505       G-----C----A---------G-----AA-C----CT...............--A...--GAG---G-------------------GA...................................... 5418
CPZ.CM.05.SIVcpzMB66        -----------AT---------------A-C-A-ATT.........AATCAG-----A---AG-------------A---------A...................................... 5401
CPZ.CM.98.CAM3              GC----C----A---------------GA-C-TC-CT...............GG-...--GAGG---------------------GA...................................... 5260
CPZ.GA.88.GAB2              G-----C----A---------G-----GA--GAA---...............----------AG---G-GA-...............................ATGGATCCTATAGATC 5241
CPZ.CM.98.CAM5              GC----C----A---------------GA--CT--YT...............G--...---AGG---------A---------Y--A...................................... 5548
CPZ.CM.05.SIVcpzMT145       ------C--C-AT---------------A-A-AC-CT...............-----A---AGG---G-----A---A----G---GA...................................... 5391
CPZ.TZ.00.TAN1              CCA---T---ATC--C-----------ATCA.......................................................................TGAATCCCATAGATC 5484
CPZ.CM.05.SIVcpzLB7         ------C----A-----------------A----A-C-...............--A--A---AGG---G---------A---------...................................... 5404
CPZ.US.85.CPZUS             G-----C---AAT--------G-----GA--T-CTTGGCAAGGAGAACTCCT--A...G-GAG---G------T-----------A...................................... 5921
CPZ.GA.88.GAB1              -C-------C-AG---------------A-CCTC-C-...............--A--A---AG-TCC------T---A--------GA...................................... 5913
MON.CM.99.L1                GGAA--C---TAC-----ATATGC--AGCAA-TAA--AGATACCCTGTTCTTAG-CCA-TGAG-G-G-C--CTC--G-CCC-A--AGTTCTGTTCCACAGGCCGATCCAGACAATCCCAGAAGACCTTCCAGATATAGAATGGATGAATGA....TGGAACCTGTTGACC 5767
MON.NG.x.NG1                GGA-------TAT--C--GCATGCCC-A--GRT---CAGGTACCCTYGRCTG-GACCACT-AG-G-C-C--CCC-AG-ACCCA--AGTTCTGTCCCTCATGCAGACCCAGAGAGGCCTTTGCACCCGTCCAGATACCACCACGATGAATAG..TGCAASGATAGTGAACC 4380
GSN.CM.99.CN166             -GA---A--CTTT--C--GGTC-ATTTATCA...G-GAGATACCCAGCAATTAG-CCCTC-AGGG-C-CC-CCC-ACCAGACAGTA-CTCTGTTTCTCATGCCGATCCTGAGCAGCCTAGGAGACCCTCAAGATATCGCATGGATGAATAG.....AGACCCGGAAGAGT 5752
GSN.CM.99.CN71              -GA-------TTC--C--ATAC-ATA-AA--GTCA--AGGTACCCAGTAATCAG-CCCTT-AGGG-T-CC-CCC-ACCAGATAGTA-CTCTGTTCCTCATGCTGATCCTGAGCAGCCCAGAAGACCTTCAAGATACCGCATGGATGAATAG.....AGACCCGGCAGAGT 5740
MUS_1.CM.01.1085            -GAC--A--CTAT-----GTAT-ACA-A--A-TGA--AGATATCCAACCATAAG-CCTTTGAG-G-C-CACA-C--CC-CCT--TA-TTCAACCCCAAATGCTGACCCTGTACCTCCTTTGAGTCCTTCTAGGTACAGGATGGATGAATAA....TGGATCCATCAGTAG 5715
MUS_1.CM.01.CM1239           -GAG--A---TAT-----GTACTCTAA-ACA-TCA-GAGATACCCCAACATAAGACCCTTGAG-G-G-CACA---TCC-CCT--TA-TTCAATGCCTAATGCTGATCCTACTCCTCCTTTGAGACCCTCTAGATATAGGATGGATGAGTAA....TGGATCCTTCAGTAG 5705
MUS_2.CM.01.CM1246           GGAC--A---TAC------TATGCTCAAACA-TCA-GAGATACCCACCGTTG-GCCCATTGAG-G-G-CGCAGC--CCAGAC--TA-TTCCATGCCAAATGCTGACCCCACACCATCTTTGAGACCCTCCAGATATAGAATGGATGAATAA....TGGACCCTTCAGTTG 5822
MUS_2.CM.01.CM2500           -GAG--A--CTAC-----GTAT-GTCAG----T-A--AGATACCCACAGGTG-GCCCCTT-AG-G-G-CGCA-C-TCCAGAC--AA-TTCAACACCAAATGTTGACCCCTTTCCTCCACGAAGACCCTCCAGGTACCGCATGGATGAGTAA....TGGATCCAGAAGTAG 5775
RCM.NG.x.NG411              GCA---A---GA-GGG-----G--T--A-GAG-GA--GGAGGA.........A-TCCTCTTAG-TCC-T-CTTCA-TCG---AA-ATCTTGTAA...................................... 5687
RCM.GA.x.GAB1               --A---A---GC-GGG--T--G--T--AACAT-CA--GGAAGT.........A-TCCTCT-CG-TCC-T-CCGCAA-CG---AA-ATCATGTAA...................................... 5652
DEN.CD.x.CD1                -G-A--A---TTC-----GGA-GC--AAAGAGT-A--CACTAT.........-CA-AC-T-AG-CCCTTAC-CC-TC-C---GAAG--CCATGA...................................... 5792
                                                                                                    Vpr end
MAC.US.x.239                -GGC--A--CATC--CTC------C---CAAC--G-GGGAGGA.........A-TCCTCTCT--GCT-TACCGC--TC----AG-AT-CTATAA...................................... 6456
Vpr                         __G__G__C__I__H__S__R__I__G__Q__P__G__G__G__.__.__.__N__P__L__S__A__I__P__P__S__R__S__M__L__*
Tat exon 1                  E__A__D__A__S__T__P__E__S__A__N__L__G__E__E__.__.__.__I__L__S__Q__L__Y__R__P__L__E__A__C__Y__N##__.__.__.__.__.__.__.__.__
H2A.GW.x.ALI                -GCA--A---G-C-G-TCA---G-T---CAA--AA-GGGAAGA.........A-TCCTCTCT--GCT-TACCGA--CC----AA-AT-CAATAA...................................... 6546
H2A.DE.x.BEN                -GCG--A---AAC-GCTCA-----T---TAA--AA-GAGAAGA.........ACTCCTT-CC--GCTGCACCGA--CC----G--AT-CACTAA...................................... 6556
H2A.SN.x.ST                 -GCA--A--CG-C-GCTCA--G--T--TCAGC-CA-GGGACGA.........A-TCCTTT-T--GCT-TACC-A--CC----GG-AT-CGATAA...................................... 5998
H2B.GH.86.D205              -GCC--T----A---CTCA--G--T--ACAAT-AG-GGGAGGA.........A-TCCTCTCT---CT-TACCGC--CC-TA-GG-AT-CGATAA...................................... 6529
H2B.CI.x.EHO                -GGG--T--C-A---CTCA--G--T--ACAAC-GG-GGGAGGA.........A-TCCTCTCT--GCT-TACCGC--TC-TA-GG-AT-CAGTAA...................................... 6527
H2G.CI.92.ABT96             -GGA-AC---A--TGCTC---G--T--A-GG-ACGC-GGCGGA.........A-TCCTCT-T---CT-TACCGC--TC----GG-GT-TTCTAA...................................... 5872
H2U.FR.96.12034             -GGG--A-----G--CTCA--G--C---AGC-GCA-GGGAGGA.........A-TCCTCTCA---CTGTACCGC--TC-T--GG-GT-CTCTAA...................................... 6016
MAC.US.x.EMBL_3             -GGC--A--CAAC--CTC------C---CAAC--G-GGGAGGA.........A-TCCTCTCT---CT-TACCGC--TC-T--GG-GT-CTATAA...................................... 5581
SMM.US.x.H9                 -GG---A--CA----CTC------C---CAGT--G--GGARGA.........A-TCCTCTCT---CT-TACCGC--K-----GG-GT-TTATAA...................................... 5930
SMM.US.x.PGM53              -GGA--A---GCT--CTC------C---CAA-G-A--GGAGGA.........A-TCCTCTGT---CT-TACCGC--TC----GG-GT-CTCTAA...................................... 6375
SMM.US.x.SIVsmH635F_L3      --G---A---GCC--CTC------C---CAAT--A--GGAGGA.........A-TCCTCTCT---CT-TACCGC--TC----G--AT-CTATAA...................................... 6460
SMM.SL.92.SL92B             GCA---A--CACC---TCA--G--T--ACAGC-AA-GGGACCA.........A-TCCCCTG-G-TCT-TACCGT--GC----GA-GT-TTATAA...................................... 5889
STM.US.x.STM                -GGC--T--C--C--CTC---G--T---CAAC-AG--GGAGGA.........A-TCCTTTG----CT-TACCGC---C----GGAGT-TTGTAA...................................... 6104
MNE.US.x.MNE027             -GGC--C--CAAC--CTCT-----C---CAAT--G--GGAGGA.........A-TCCTCTCT---CT-TACCGC--TC----AG-AT-CTATAA...................................... 5924
DRL.x.x.FAO                 -TA---A--C-AGG-A--T---C-T--A...C-AGC-GGACGCTGG......A-CCCATTTCG--CCTT-CC-AGGGCG-ACCAAC-ATTGTAA...................................... 5755
MND_2.x.x.5440              -TA---A---A-GGGA--TC-C--C--ACAAGAAG--GGAAGATAT......A-CCCACTTCG-TCCTT-CCGAG-CCG-ACAA-CCCTTGTAA...................................... 5704
MND_2.CM.98.CM16            -TA---A--CA-GG-A---C-T--C--ACAAGAAG--GGGGGATGT......T-CCCCTT-CG-TCCTT-CCGAGATCGGACAA-CCCCTGTGA...................................... 6139
MND_2.GA.x.M14              -CA---A--CA-GG-A---C----T--ACAAGAAG--GGCAAA......TATA-CCCCCTTAG-TCCTT-CC-AGGCCG-ACAA-CCCTTGTAA...................................... 6066
OLC.CI.97.97CI12            --A---A--C-C-GGAGTGGAGC-TA-AAGA-AC-CTCACAGA..............................AGATC-GA-GGAAT-CAAGAA...................................... 5647
GRV.ET.x.GRI_677            -TG---T---A--AGG--AG--CCTTTTTCCC--TACGAAGAGAGGAGAAATGGACAAG---G-G--...-CCC-ACCCCCT.................CCTCCAGGACTTGCATAG.................... 6039
VER.KE.x.TYO1_patent        -TG------CA-GAGA--GCAGCCCTTT-AAC-ATACGAGGAGAGGAGAGATGGACAAG-G-G-G-C-GA-C-AATC--GTC.................CCACCAGGACTTGAGTGA.................... 6105
VER.KE.x.9063               TTG---T-----TAGA--ACAGCCCTT--AAC--TACGAGGAAAGAAGAGATGGACAAG-G-G-G---GA-CGAA-C--GCT.................CCACCAGGACTTGATTAG.................... 6105
VER.KE.x.AGM155             GTG---A-----CAGG--ACA-CCCTTT-AGC-ATACGAGGAGAGGAGAAATGGACAAG-G-G-G---GACC--GAC--GTC.................CCACCAGGACTTGATTAA.................... 6100
VER.DE.x.AGM3               GTG---A--C--CAGA--ACA-CCCTTT-AAC-ATACGAAGAAAGGAGGGATGGACAAG-G-G-G--CGA-C--GGC-CGT-.................CCACCAGGACTTGATTGA.................... 5603
TAN.UG.x.TAN1               -TG---C--CA-GAGA--A-CTCCTTTT-AGC-CTACGAAGAAAGGAGAAATGGAGTCG---GGG--CGG-AT-G---AGAG..............CCTCCTCCAGGACTTGCATAG.................... 6101
SAB.SN.x.SAB1C              -GG-A-A---T-TA-GCCA-AG-CTCA-CC-G-GTATGGACCAGGAGCAGGAGGCCC-CCCC--G-TCTG----GAGC-TC-GGAGGAGCTGCATCGGCCGCTCCAGGCCTG...TGA.................... 6333
MND_1.GA.x.MNDGB1           TCAA------GATGGA-C-TTC............A-GGAAAGAAGACCACAA-T-CCCCCCT--G--TTCA-GC-A--AG--GATAGATTATAA...................................... 5541
DEB.CM.99.CM40              -GAA-CT--CTTT--C--GGAG-G--AG-C---CA--GGG...TATCCTAACATT---CC-TTG-C-GGAA--AAT--AGA-GTTAG-GACGGAGAATAA.....................CCTTTCACCTTTGCTT 5697
DEB.CM.99.CM5               -GAA-CA--CTTC-----AGAG-GG-AA-C----A-GAGA...TACCCCAATATT--ACCCCT--C-GGAA--AA---GGAGGT-AG-GATGGAGAATAA.....................CCTCTCTTTTACATCA 5691
TAL.CM.01.8023              GCAA-----CTTT---GAGCAGCA-A-AAGA---GCCCACTACCCCAACATA-GACCCCTTT-CG-C-GGA-G...-AC-CCATGC-ATAA...................................... 5360
TAL.CM.00.266               GCAA-----CTAT---GAGCAGCA-A-AAGA----CGCACTACCCCAATATC-GACCCCTTT--G---GA--G-AG--G-CCATGC-ATAA...................................... 5865
SUN.GA.98.L14               -...AT-----CT.....................TCTT-C-TC---ATC----G----T-AG-A-ATGT-A-CCCAAGGCCACCAACAAGACCAAGACCAGGGCAAGGGGACATTAGAAGAGGCCTACAAGACTAA.................... 6259
WRC.CI.97.97CI14            CCAA-------CTA-ACAACA-GAGAAAAGGGAA-A-CAATTTCAAAGAGCAG-ACAATTGTAT-----CTTG-AG--A-C-G-AGGAGATCAGGGAGAAGTTGAGCAAAGA...................................... 6139
WRC.CI.98.98CI04            CCAA-------C-A-ACAACA-GA-AAGAGAGAA-A-CAATTTCAAAGAGCAG-ACAATTCTTT-----CTTG-AA--AGC-G-AGGAGATCAGGGAGAGGTTGAGCAAAGA...................................... 6154
WRC.GM.05.Pbt_05GM_X02      -CAA--A----CCA-AGAA-A-GAGAA-AGAGAAA--GATTTTCAAAGAGCAG-ACAATTTTAT------TTG-AA--GGCT--AGGAGATCAGGGAGGAGTTGAGCAGAGA...................................... 5074
LST.CD.88.524               GCC-CAA----CCAG-TATGC-GC--C-AGA.........GCAGGAACAGCA--CCCA-C--A-GC-GCACCTAGACCA-CTGGAAGAGATATACAAAGAGGCCATTACTGA...................................... 5119
LST.CD.88.447               GCC-ACC----CCAGCTATGC-GC--CAAGA.........GCAGGAACAGCT--CCC--C--A-GC-GCACC-AGACCC-CTAAAAGAGACATACAAAGAGGCCGTTACTAA...................................... 5122
LST.KE.x.lho7               -...ACT----CCAGCTATGC-...-CA-CC-GA......GCAGGAACAGCC--TCCA-C--A-GC-GCACC-AGACC--ATGAAGGAGACATACAACGAGGCAGTGACTAA...................................... 6203
SYK.KE.x.SYK173             TGCA--A----CT--AC-G-CC-G-TAT-CC-GA-ACAGAGGGTACCCCCAT-CTTCTTAG...................................... 6042
SYK.KE.x.KE51               TGCA--A----C---G--G-C--AGTA--CG-GA-ACAGGGGATACCCACAT-CTTCTTAG...................................... 5706
COL.CM.x.CGU1               TGCG---CAAT-T-GCTTTCCCTCTACGCAA-AC-CTTAG...................................... 5516
```

```
                                                
```

```
                                                        Tat Rev exon 1 end    intron start
H1B.FR.83.HXB2          AAC...............AGTCAGACTCATCAAGCTTCTCTATCAAAGCA...GTAA........................................................................GTA.....................GTACATGTA........ 6061
Tat exon 1              _N__.__.__.__.__.__S__Q__T__H__Q__A__S__L__S__K__Q
Rev exon 1              __T__.__.__.__.__.__V__R__L__I__K__L__L__Y__Q__S__
H1A1.UG.85.U455         ...............GGC---A--GA-------A--CT-A--C-------...----........................................................................---A.................TTA-CTT-----........ 5510
H1B.US.90.WEAU160       ...............AAC----------------A---------------...----........................................................................---.....................---------........ 6059
H1C.ET.86.ETH2220       ...............AGC---A--GA-------AA-CT-A----------...----........................................................................---CCAAATAATAG..............-----...ATGGT 5463
H1D.CD.84.84ZR085       ...............AGC----------------A-C--A--C-------...----........................................................................---.....................---------........ 5582
H1F1.BE.93.VI850        ...............AGC-----AGTG------AA-------C-------...----........................................................................---GTGTTA...........AATA---T-----........ 5410
H1G.SE.93.SE6165        ...............AGC---A-AGGC-------A-C--G--C-------...--G-........................................................................---GAAAAA...........ATTA---T-----........ 5467
H1H.CF.90.056           ...............AGTTTG--AGA-------AA----A----------...----........................................................................---.....................T--TCATA-........ 5407
H1J.SE.93.SE7887        ...............GGC---A------------A-CT-A--C-------...----........................................................................---AATCAA............GTA-C-T-----........ 5383
H1K.CM.96.MP535         ...............AAC---G--AC--------A-C-----CG------...--G-........................................................................---GTGCTA...........AGTA---------........ 5266
H101_AE.TH.90.CM240     ...............AGC---A--GA-------AA-C--A--C-------...----........................................................................---.................ATAA---T-----........ 5632
H102_AG.NG.x.IBNG       ...............AGCC-----GA-------AA-C--G--C-------...--G-........................................................................---GTAATA...........ATTA-C-T----G........ 5594
H1N.CM.95.YBF30         ...............AGC---A---G--------A-CT-A--C--G----...----........................................................................---AAACCTGTATAT.......................... 5654
H1O.BE.87.ANT70         ..................CA--CAGA----A---A-C--G--C-------...----........................................................................---ACGCTA................................ 6109
H1O.CM.91.MVP5180       ..................TA--CAGA-A--A---A-C--G--C--G----...----........................................................................---ACGCTG................................ 6096
CPZ.CD.90.ANT           GCTGAA.........AGC---G-A-A-A------A-C--G--------T-...----........................................................................---CTCTTA................................ 5496
CPZ.CM.01.SIVcpzCAM13   -G-...............---A-A-A-A------A-C--G--CG------...----........................................................................---TATTATCTGACCCT........................ 5648
CPZ.CM.05.SIVcpzEK505   -G-...............-A-A---G--------A-C-----C-------...---C........................................................................---AAGCTATATAT........................... 5631
CPZ.CM.05.SIVcpzMB66    -G-...............-------A--------A-C--G--C-------...--G-........................................................................---ATTCATGCAGATATATGCTTA-GCTT-A.......... 5631
CPZ.CM.98.CAM3          GCA............AGC---AC-CA-A------G-------CAGC----...----........................................................................---ACTTCACATGCT.......................... 5477
CPZ.GA.88.GAB2          GG-...............---A-A-GCA------A-C-----CG-C----...--G-........................................................................---TCCCCTCTTT............................ 5457
CPZ.CM.98.CAM5          TCA............AGC---AC-CA-A------A----A--CG-C----...----........................................................................---ATTTTGCATGCT.......................... 5765
CPZ.CM.05.SIVcpzMT145   --T...............CAGG-T...A--A-G-A-C-----CG------...----........................................................................---AAAGGCTCAC............................ 5600
CPZ.TZ.00.TAN1          -GT..................AC--A-A------A-C--G--CG-C----...----........................................................................---CCC................................... 5687
CPZ.CM.05.SIVcpzLB7     -G-...............----C-CA-T------A-C--A----------...----........................................................................---CTCCTATAAGTTGTAGTCAATA-............... 5629
CPZ.US.85.CPZUS         --GAGC............---GC-GA-A------A-C-----CGC-----...----........................................................................--GCCCTTTTGTGCT.......................... 6138
CPZ.GA.88.GAB1          GCAGGC............---A-A-A-A------A----A--C-------...----........................................................................---T..................................... 6119
MON.CM.99.L1            GT-CCT.........GGCCT-AGC-GCAG-A-GAA-C--GC-AG------...--G-........................................................................---...................................... 5976
MON.NG.x.NG1            -CG.........ATTGGCGA--G--GCT----G-A-AA--CGGG-GTCA-...----........................................................................--G...................................... 4592
GSN.CM.99.CN166         G-ACCT...GATCCTAGCGTGTC-T--A------AC---GC--AGGTC--...----........................................................................---AGGC.................................. 5974
GSN.CM.99.CN71          G-ACCT...GCTCCTAGCGTG-C----A------ACC--A--CA--TC--...----........................................................................---AGGC.................................. 5962
MUS_1.CM.01.1085        ..................CTG-C--GC---TC--A-C--G--CA--TCA-...----........................................................................---GC.................................... 5920
MUS_1.CM.01.CM1239      ..................GTG----A-----CGAA-C--G--C---TCA-...----........................................................................--GAT.................................... 5910
MUS_2.CM.01.CM1246      ..................CTG-C-T-AG----GAA-C--G--CG-TTC--...----........................................................................---AT.................................... 6027
MUS_2.CM.01.CM2500      ..................CTG-CA-ACA------A-C--A--CA-CTCA-...----........................................................................--GAT.................................... 5980
RCM.NG.x.NG411          --G............AAG-A-----T-A---C-TT-GA-AC---TG----...----................................................................................................................. 5841
RCM.GA.x.GAB1           --G.....................G--G--A--TT-C--G--C-T--T--...----................................................................................................................. 5797
DEN.CD.x.CD1            ............CAAAGCG--AC-C--G------AAG--G----------...----.GTTTGATGGATCTCCTCTCATCACTGATTTTAAGCTTTTCAATAATCATTATTTGCTTAGTTTGTTATATT--TATTATTACAGTGATAGTTCTT-AC-CATAGG....... 6091
                                                        Tat Rev exon 1 end    Tat Rev intron start
MAC.US.x.239            ..................GC-A--G--A--AC-T-----GC------CA-...----........................................................................--....................................... 6603
Tat exon 1              _.__.__.__.__.__.__A__K__A__N__T__S__S__A__S__N__K
Rev exon 1              __.__.__.__.__.__.__L__R__L__I__H__L__L__H__Q__T__
H2A.GW.x.ALI            ..................-T-A--G-----TCGT-----GC----G-CA-...--G-........................................................................--....................................... 6693
H2A.DE.x.BEN            ..................-C-A------C-TCGC-----GC-C--G-CA-...--G-........................................................................--....................................... 6703
H2A.SN.x.ST             ..................-C-A--G-----TCGT-----GC----G-CA-...--G-........................................................................--....................................... 6145
H2B.GH.86.D205          ........................---AC-GC-C-----GC-C--G-CA-...--G-........................................................................--....................................... 6670
H2B.CI.x.EHO            ........................---AC-AC-T-----GC-C----CG-...--G-........................................................................--....................................... 6668
H2G.CI.92.ABT96         ........................G--A--TC-TT----GC------CG-...----........................................................................--....................................... 6016
H2U.FR.96.12034         ........................G--T--TC-T--G---C------CA-...----........................................................................--....................................... 6157
MAC.US.x.EMBL_3         ..................GC-A--G--A--AC-T-----GC------CA-...----........................................................................--....................................... 5728
SMM.US.x.H9             ........................G--A--AC-TT----GC----G-CA-...----........................................................................--....................................... 6071
SMM.US.x.PGM53          ..................GC-A--G--A--AC-TT----GC-C----CG-...----........................................................................--....................................... 6522
SMM.US.x.SIVsmH635F_L3  ........................---A---C-CT-C--GC------CA-...----........................................................................--....................................... 6601
SMM.SL.92.SL92B         ........................G--A---C-TT-GA-AC------C--...----........................................................................--....................................... 6030
STM.US.x.STM            ........................---T---C-AT----GC------CA-...----........................................................................--....................................... 6245
MNE.US.x.MNE027         ..................GT-A--G--A--AC-T-----GC------CA-...----........................................................................--....................................... 6071
DRL.x.x.FAO             ............TCAAGAGTATCTGAGAC-GGT-ACA-G-CTGTGGCAAG...----........................................................................--GAAAAAG................................ 5918
MND_2.x.x.5440          ............GAGAAAGTA-CA-GCA--AGTCAAG-AGCT-TGGGAAG...----................................................................................................................. 5858
MND_2.CM.98.CM16        ............GAAAAA-TATCAGAAG--AGT--AG-G-CTGTGGGAAG...----................................................................................................................. 6293
MND_2.GA.x.M14          ............GGAAAA-TATCACAGG--AGT-AAG-AGCTGTGGGAAG...----................................................................................................................. 6220
OLC.CI.97.97CI12        ........................T-G------CAAGA-TCTAGGG-CAC...----........................................................................--....................................... 5791
GRV.ET.x.GRI_677        ........................GA-A------ACC--A--CGGC----...----........................................................................--....................................... 6201
VER.KE.x.TYO1_patent    ........................CT-G--TTG---C--T--CAGC-C--...----................................................................................................................. 6277
VER.KE.x.9063           ........................C--A---GCCTA----A-CAGG-C--...----........................................................................--....................................... 6279
VER.KE.x.AGM155         ........................G--G---GCAT-C--G--C-GC----...----........................................................................--....................................... 6274
VER.DE.x.AGM3           ........................CT-G---GCTT-C--GA-CA---A--...--G-........................................................................--....................................... 5777
TAN.UG.x.TAN1           .....................TCA-A------GAA-CT-G-----C----...----........................................................................--....................................... 6266
SAB.SN.x.SAB1C          .....................ATTT-G---A-TCAAGT-TCTCT-C-CA-...C--G........................................................................TA....................................... 6477
MND_1.GA.x.MNDGB1       ..................-TA-CTGG-AG-GGT-AAGAAGCTAT-CGAAG...----........................................................................--....................................... 5736
DEB.CM.99.CM40          -CGCCAGCTGCTGAAAGT-A----C--G------A--T-G--C--G----...----........................................................................-C-AATCTGCATGTAAATGCTAAA-GGA-CA-T........ 5946
DEB.CM.99.CM5           GCT.........GCGAATGA-A--GA--------AG-T-A--C--G----...----........................................................................---AAGCTGCATATGTATATAGTATC---............ 5927
TAL.CM.01.8023          -CTTCT......AGCAGC-A-----AGG--A-G-A-C--G----------...----................................................................................................................. 5514
TAL.CM.00.266           -CTCCT......AGCAGC-A----GAGG--A-G-A-C--G----------...----................................................................................................................. 6019
SUN.GA.98.L14           ........................-AGG--A-GCAAGAGGCTGT-CGAAG...----................................................................................................................. 6416
WRC.CI.97.97CI14        .......................................GCTG-TTCAAG...----........................................................................--....................................... 6289
WRC.CI.98.98CI04        .......................................GCTA---CAAG...----........................................................................-A....................................... 6304
WRC.GM.05.Pbt_05GM_X02  .......................................GCT----CAAG...----........................................................................-C....................................... 5224
LST.CD.88.524           ........................GAAAC-G-GCAGGA--CTGTGG-CAG...----........................................................................AA....................................... 5284
LST.CD.88.447           ........................GAAG--A-GCAAGA--CTGTGG-CGG...----........................................................................-A....................................... 5287
LST.KE.x.lho7           ........................GAGAC---GCAGGA--CTGTGG-CAG...----........................................................................-G....................................... 6368
SYK.KE.x.SYK173         ........................T--G-GG-T-A-AGCGC-C-T-CAGG...----................................................................................................................. 6218
SYK.KE.x.KE51           ..................GAGG-CTT-GC-GCGCA-AGCTC-A---CT-G...----................................................................................................................. 5894
COL.CM.x.CGU1           .............................................-CA--...----........................................................................--....................................... 5637
```

```
                              Vpu start (ACG in HXB2)
H1B.FR.83.HXB2                .ACGCAACCT.........ATACCAATAGTAGCAATAGTAGCATTAGTAGTAGCAATAATAATAGCAATAGTTGTGTGGTCCATAGTAATCATAGAATAT.............................AGGAAAATATTAAGACAAAGAAAAATAGACAGGTTAATTG 6191
Vpu                            T  Q  P  .  .  .  I  P  I  V  A  I  V  A  L  V  V  A  I  I  I  A  I  V  V  W  S  I  V  I  I  E  Y  .  .  .  .  .  .  .  .  .  R  K  I  L  R  Q  R  K  I  D  R  L  I
H1A1.UG.85.U455               .-T-AC----.........T-GGA---CTGG------AC--GGC-GA--------C----CT----------A------A-T------GGT--------AAA.........................-AA---T-GC---AG----A---------------T-AA 5642
H1B.US.90.WEAU160             .-T----T-..........T-G-A----T-------------C-----------GG-----------------------------C-----------...........................------------------------------------- 6189
H1C.ET.86.ETH2220             TGATTT--TAGCAAAAGTAGATTAT-G-A---T------------CA--------C----C-----------------A------C-TAT---------..........................------T-G-------------G------T---------A 5603
H1D.CD.84.84ZR085             .-T----T--.........T---A----T-----------------T----GC----------------------A--------T-----------..........................C-T-G----AA---G---------------T--------- 5712
H1F1.BE.93.VI850              .-T-TC-TA-.........T-GTT-GC-A---GG----C-------A--------C---------C------------A-T------TAT---------..........................-A----C-GG----G------------A-T-AA---TA-A 5540
H1G.SE.93.SE6165              .-T---GT-A.........T--GT----T-----------G----A--------T-C---GC---C------------A--------T-T---------..........................---G-----AG--A---G-A--G-----GA-A----C--- 5597
H1H.CF.90.056                 .-T-T-TATA.........T--GG-T---G-AT-GG--CGCT-G-----AC-TTT--C---GCC...G--A--------A--------TAT---------..........................-AA---T-GG----G----AG------------C------ 5534
H1J.SE.93.SE7887              .-T-AT----.........T-G-A-----C---------------TA----------TT-CC-T-------G-A-----A-T-----GTAT---------..........................-A----C-G-----------G--------T-A-------- 5513
H1K.CM.96.MP535               .-T-GTGT-C.........T--G----TTC-AT-G---C-TT-G------C-CT----T--GC-..----A----A---A--------TAT---------..........................------T-GG---A-----A--G----A--T-------A- 5393
H101_AE.TH.90.CM240           .-T-AC----.........T-GGA---TAGT----------G-C-GA-------GC----CT-----------A-----AA-T------GCT------GT-..........................-A-------C----G---------------------G--A 5762
H102_AG.NG.x.IBNG             .-T-----------.....T--A-----ACT----------G-C-----------T-C---GC---C------------AG-------TAT---------..........................---------AGG-A---G-AG--------------C--- 5724
H1N.CM.95.YBF30               .-T--TG.................TC-T-G-G-T-CA----G----G--C----G-T-GC------G--A-A--C---G-AT--C--TAT-G-------..........................-A-------AA-TTG--GGA-------A-ACACA---GAC 5775
H1O.BE.87.ANT70               .-T---T-A-.........-GGGACC-GC--------A--ATTA-TAGT-CTTTGC-GT-T---AATG--A--T-A---GGGT-TA-T......CTTAGAAAATAT............TTAGAACAAAAGGAAC--GACAG--AGG-G---G-----CTTGA-AGGT-AA 6251
H1O.CM.91.MVP5180             .-T---T-AA.........GAGAACC-GC----CT--A----T---AGT-CTTTGTGTC-T---AATG--C--A-A----TGT-TAAC......CTTAGAATTTAT............TTAGTGCAAAGA-AAC--GATAG---GG-GCAGG-----CTTGAAAG-T-AA 6238
CPZ.CD.90.ANT                 .-T-ACTAA-ATATTTGAGTATG-TT-TC-T--CT-TAGTAT-G--C--TGGATT--CTGT---C-T......A-ACTC-AT-AGC-T......T-TA-AATATAT.........AAGCAGCAGCAAATAGAT--T-AGAG--AT-------TC-----AGT-----G-A 5644
CPZ.CM.01.SIVcpzCAM13         .-T-AT-TTA...................----CTTT-GGTTGCTT-GCCA-T-CCTT---TT--C-TAAC--TT--A-A---AGG-ATT--......TGG............AGATTATGTAAA............C---ATAAGCTTG--GT-G----T--G-ACC-TGC-T 5772
CPZ.CM.05.SIVcpzEK505         .-T-TTGTTG.........C-TAT--AGT---G-T-CA---G-----C-A---AG-CTC------T------A--T---G-A-----......T-TAGAATATAC............AGA............G--G--AA-GT-G--GA-------TCTCA-C---GAC 5761
CPZ.CM.05.SIVcpzMB66          .-T-G-TATA.....GTGCAG-A-G---GCTT-T------TTC--A-TA---A-T--G-T----TC-----AA-A---GTA-A---T......T-T............AAACTCTGTAAA............G-GGATAGG--G----AG--------T--A------- 5764
CPZ.CM.98.CAM3                .-T--T-A--......TGG...GA-CA-A---GCT--A-T-GTA---G-A---A---T--T-----T--T--A-CC---GGG----CT......TTTA-AGAGTGG............AGAAAAGGAAAAGAAG--GA-AAC---GC-CA------TAGG-CT-----A- 5619
CPZ.GA.88.GAB2                .-T--TTT-C.....ATGTGGGT-GCTA---G-C--A-T-G-A---GTACTCTTT--G-T---AAT-----G------GGT--T---......-G-ATCTCTGTCTAT......AAAAGGTGGAAAAGACAC---GA-GAGCA-AG--TC...--T--TTT-A-T...A 5602
CPZ.CM.98.CAM5                .-T--T-AT-......TGG...GAGCA-A---GTT--A----T----G-A---A-C-T--T--T-TT--T--A--T---GGA----CT......T-TA-AAAGTGG............AGAGAAATAAAAGAAG---GGAAC---ACCCA--G---TT---AT---GCA- 5907
CPZ.CM.05.SIVcpzMT145         .-T---G-TA............GA---T--GCT----A--TT---CA---CTTT---GC-TG----T---T---CT---ATAGC--CT......T-T........................AAAGAGTAT-AA--GT--CA-CA-GT---GCG------A--AC-TCAG- 5727
CPZ.TZ.00.TAN1                .-T-AT-AAA.....ATAG--GTGGG-AGT-TGTC-ACTAATG-CA---GCATTC-TTGT---TT-C-GA--T-AATAGGGGG-..........-GC-TGCTAATAGGT......ATAGGTATAAGAAGAGA-TT-GA-AGGGA-AGGCA-C-TCA-AGAGTA---GAAA 5835
CPZ.CM.05.SIVcpzLB7           .-T-G-C-TA............AT-GA-T---G-T--A---GC-----TA---A-C----C----T------A--A----TA-A--C-......T-TC-GCTTTAT..........................-AAG-G-ACA---------AGGC----A-T-AA------- 5756
CPZ.US.85.CPZUS               .-T--TTAA-......TGGT-TGA-----GTTT-----C-TT-GG-A--...-A-GGT--TT---T-G-GA-AA-C---GGAC----T......-CTAGATTGTGG............AGACAAATAAAA-TA---GA-AAT-CT---CA-G-G--TC-A-ATC-TT-A- 6280
CPZ.GA.88.GAB1                .-T-ACT-TG........TTAGT-.GGCT---TTC-CA--TT-G---GCT--...--TGC-TGGAAC--TTG-A-A---GGATACA-T......ATTA-ATGGGGATAT......AGAAGGTATAAAAGACAT-GGC-TGAG-C-G--.........--T--G--AC-T-ACC 6258
MON.CM.99.L1                  .-T-A-TTA-...TGGTGGTC--TGG---C-AT--CTTACAGCC-GA-CC-CAT-GCCC-CCCC-T-GC--C-TG-GCT-GG...TGG....................CGATATTATAAAATCACTAAG---TTT-A-AG--T-G-CCA-G-----C-G---C-G---C 6118
MON.NG.x.NG1                  .-T-A-TA-C................TATTGGAT-GC--CCAT--CT--TTGGTGTG-T--TG-GCTTT---CGCCT--C................ACCCTATACCTTCTCTATCAGTCTTATAAGGAGCAC--GCGGG-GTCC-GCTTCG--CAG---TTTCA--GGC 4728
GSN.CM.99.CN166               .-T---T---..............G--GC---TTGGTGGTGG-GTGCT-CTA-TAT--CCT-C---TATC-TTG-...GTAG-GC-TC-T......-CTCTCTAC....................TTAGCTT--G-T-AG-GGGT-A--G-----CC-A-ACCAAC-CAA- 6101
GSN.CM.99.CN71                .-T-AGTG--..............G--GCTC-CTGGTGGTGG-GGGCT-CT---AT--CCT-T---TATT-CTGC...-TAG-A--CT-T......-CTCTCTAC....................TTGGCGT--G-C-AG-GG-T-A-GG----GCC-A-G...A--CCA- 6086
MUS_1.CM.01.1085              .-T-A-CTA-...TGGTATT--G--GCT--CATTG--AC--GCA-CTACT-T-TG--TGCT---TTTGCCT------TAG-TTACCA-....................AGGTGGTGTAAGCCCAAGAAG.....................GT-GAG-T---TG-C--CC 6044
MUS_1.CM.01.CM1239            .-T-A-TTA-...TGGTATC--G--GC--CTTT-G--AC--GCA-CTACT-TCTG---GCCT--TTTGC-T-C---CTAG-ATATCAG....................AGGTGGTGCAGACCAAAG.......................GT-GA--TG--CG-C--CA 6031
MUS_2.CM.01.CM1246            .-T---TTAC...TGGTATC--GG-GCT-CCAT-G--AC---TA-TTATC-T-TT--TGCTT---T-GCCT--A-CCTAG--TATCA-....................AGGTGGTGTCAACCTAAG..................-C-CC--T-GA--TG-ACG-CC-CC 6154
MUS_2.CM.01.CM2500            .-T-A-TTGG...TGGTGGT-TG--GC--C--T-G--ACT--CA--TATT-T-TG---GCTT---TTGC-T-----CTTG-ATATCAG....................CGCTGGTGCCAGCCTCAG................-A-GGGCAGGT-GAG-T--ACG-T--CC 6110
RCM.NG.x.NG411                ..................................GTAT-G-A-G...TCCTGGTCTG--CCT--TT-TAGATGC-TAG-TTATAAG............GTGATAGGAGATATAATTAGAGAAGTTTTT-TTTGT-G-G-TTATAG--A-T-G-AG-TA-ATC-TTAGA 5960
RCM.GA.x.GAB1                 ..................................GTAT-G-G-G...CCCTGG-CT-GT-CT--TT-TAGA-CC-CAG-T-A-AAG............GTTTTTATAGAT....................CTTTT-G-TAC--T-ATT-A-GG--A-T-G-ATAAG--AA 5898
DEN.CD.x.CD1                  .-T--TTAAG.....................ATC-AGT---G-CC-A---A-TACG-TTGCT-G-T-T-T-CA--TGTTATT-CCTGG--AGC--CCATAGGT............GTAGGGTATCTTGCTTAT-G-GCC-AT-A-AGTTAT-G-GA---GCTTAGGTA-A 6227

MAC.US.x.239                  .................................................................................................................................................................... 6603
H2A.GW.x.ALI                  .................................................................................................................................................................... 6693
H2A.DE.x.BEN                  .................................................................................................................................................................... 6703
H2A.SN.x.ST                   .................................................................................................................................................................... 6145
H2B.GH.86.D205                .................................................................................................................................................................... 6670
H2B.CI.x.EHO                  .................................................................................................................................................................... 6668
H2G.CI.92.ABT96               .................................................................................................................................................................... 6016
H2U.FR.96.12034               .................................................................................................................................................................... 6157
MAC.US.x.EMBL_3               .................................................................................................................................................................... 5728
SMM.US.x.H9                   .................................................................................................................................................................... 6071
SMM.US.x.PGM53                .................................................................................................................................................................... 6522
SMM.US.x.SIVsmH635F_L3        .................................................................................................................................................................... 6601
SMM.SL.92.SL92B               .................................................................................................................................................................... 6030
STM.US.x.STM                  .................................................................................................................................................................... 6245
MNE.US.x.MNE027               .................................................................................................................................................................... 6071
DRL.x.x.FAO                   .................................................................................................................................................................... 5918
MND_2.x.x.5440                .................................................................................................................................................................... 5858
MND_2.CM.98.CM16              .................................................................................................................................................................... 6293
MND_2.GA.x.M14                .................................................................................................................................................................... 6220
OLC.CI.97.97CI12              .................................................................................................................................................................... 5791
GRV.ET.x.GRI_677              .................................................................................................................................................................... 6201
VER.KE.x.TYO1_patent          .................................................................................................................................................................... 6277
VER.KE.x.9063                 .................................................................................................................................................................... 6279
VER.KE.x.AGM155              .................................................................................................................................................................... 6274
VER.DE.x.AGM3                 .................................................................................................................................................................... 5777
TAN.UG.x.TAN1                 .................................................................................................................................................................... 6266
SAB.SN.x.SAB1C                .................................................................................................................................................................... 6477
MND_1.GA.x.MNDGB1             .................................................................................................................................................................... 5736
DEB.CM.99.CM40                .................................................................................................................................................................... 5946
DEB.CM.99.CM5                 .................................................................................................................................................................... 5927
TAL.CM.01.8023                .................................................................................................................................................................... 5514
TAL.CM.00.266                 .................................................................................................................................................................... 6019
SUN.GA.98.L14                 .................................................................................................................................................................... 6416
WRC.CI.97.97CI14              .................................................................................................................................................................... 6289
WRC.CI.98.98CI04              .................................................................................................................................................................... 6304
WRC.GM.05.Pbt_05GM_X02        .................................................................................................................................................................... 5224
LST.CD.88.524                 .................................................................................................................................................................... 5284
LST.CD.88.447                 .................................................................................................................................................................... 5287
LST.KE.x.lho7                 .................................................................................................................................................................... 6368
SYK.KE.x.SYK173               .................................................................................................................................................................... 6218
SYK.KE.x.KE51                 .................................................................................................................................................................... 5894
COL.CM.x.CGU1                 .................................................................................................................................................................... 5637
```

```
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
```

```
                              Env signal peptide end Env gp120 start
H1B.FR.83.HXB2          ......GCTACAGAAAAATTG...............TGGGTCACAGTCTATTATGGGGTACCTGTGTGGAAGGAAGCAACC.........ACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACC 6451
Env                      .  .  A  T  E  K  L  .  .  .  .  .  W  V  T  V  Y  Y  G  V  P  V  W  K  E  A  T  .  .  .  T  T  L  F  C  A  S  D  A  K  A  Y  D  T  E  V  H  N  V  W  A  T  H  A  C  V  P
H1A1.UG.85.U455         ......AA-G--C--C-----...............--------G----C-------------------A--T--GTT.........-----C--------------------------------G---A--G--------C-----T----------------- 5896
H1B.US.90.WEAU160       ......---G-------C---...............-------------------------------A---------.........--------------------------------------------A----------------------------------- 6446
H1C.ET.86.ETH2220       ......-GA-TG-GG--C---...............-----------------------A-----A--T--T-G-.........C-T-------------------------------------A-----------C----GT---TT---------C-- 5860
H1D.CD.84.84ZR085       ......---G----T--C---...............-----------T--------------A-----A---------.........----G-------------------------T-----A-A-------C-------A-C-----T----------------- 5972
H1F1.BE.93.VI850        ......---G----C--C---...............-------------------------------A---------.........--T-------C----------------------------A-G-----C------C-------------------------- 5797
H1G.SE.93.SE6165        ......--CT--A-T--T---...............-----T-----------------------G-A--T--GAT.........-------------------C---------T-----AG-G-T--AAGC--------C-----T----------------- 5854
H1H.CF.90.056           ......---G--C----C---...............-----T----A------------------A--G---AA.........-----------C----------------G--------G-----AAAG--------C-----T--------A-------- 5791
H1J.SE.93.SE7887        ......--A-A----G-C---...............-------------------------------A--T---AG.........------T-------C------T-------------AG---T--AAA---------C-----T--------T-------- 5770
H1K.CM.96.MP535         ......AA-G----C------...............-------------------------------A-------A.........C----------C--------------------------G-AG---------------C-----T--C------------- 5650
H101_AE.TH.90.CM240     ......--CT----C--C---...............-----T-----T---------G---------GA--T--GAT.........-----C-------------------C--G---C----G-----A--G--C-----C-------------------------- 6019
H102_AG.NG.x.IBNG       ......AA-G-T---C-----...............--------G----C--------------------ACG--GAG.........-----C-------------------------------A-----------C-----T----------------- 5978
H1N.CM.95.YBF30         .........T-T---C--CAT...............-----A-----G--C----------A--A----GA------GAG.........--A-----T--C-----T-------------CC--AG--------CT--C--CA-C------------A--A-----T-- 6038
H1O.BE.87.ANT70         AGC.....CTTAG-C-GC-A...............-AT-CA--------GC------G----A--G-A--T-----A.........C-AGTA-----C-----T----------CCT-ACAAGC--T--AAAG------A-------AT----A--------T-- 6499
H1O.CM.91.MVP5180       AGT...TA--GTA--C-----...............-AT-C------T----C------------A--G-A--G--G-A.........C-AGTA-----C-----T----------CCT-ACAAGC--T--ACAG------A-------AT----A-----C--T-- 6489
CPZ.CD.90.ANT           AAG...-GG--GA-TG--GAC...............-AT--A-----A-TC-----A--C-----C----GAA-T--G--A.........C-T--------------CA--A---CTCCATGACAAG---------G--C-----A-----A--TACCAGT-----G-- 5877
CPZ.CM.01.SIVcpzCAM13   ACA........TCTG-TGAA...............-----A-----G--------A-------------GA----TC-A-.........--AGTA--C--C-----T--------------C--AGC-------CT-----CA-A-------------C--A-------- 6029
CPZ.CM.05.SIVcpzEK505   ............TCTG-----...............-----A-----G--C--------------C----GA--T---GAG.........--AGT------C-----T----------G--CC--AG--------CT--C---A----------G--A--A--C--C-- 6000
CPZ.CM.05.SIVcpzMB66    ............--T--CA--...............-----G-----T--C-----A--------C----GA--G---GA-.........--A-----G------------------A------C-GA-CC----A--G--C-----A-----TT-T--A--T------T- 6033
CPZ.CM.98.CAM3          ACT...-G--GT--TC-G---...............----CA-----G--C-----------------------A--G---GA-.........--A--A--C-------C-----G--GTC---TCT-A-C-A-----CC------A---------T----A--------T-- 5870
CPZ.GA.88.GAB2          AAGACA-AAT-T--TC---G-...............-----A-----A--C--------T--A--C---------G--T--T.........--AC-AT-G--C------G--A-----A-C---T-TAA-AGAG--ACCT------A-------------A--------G-- 5856
CPZ.CM.98.CAM5          ACT...-GCCAT--G...-G-...............----CA-----G--C-----A--------A------A--G---GAA.........--A--A--T--C-----C----G--CTC----CT-A-C-A---A-CC--C---A--------T----A--------C-- 6155
CPZ.CM.05.SIVcpzMT145   GGT......GA------T-G-...............--------T--T-----------C--A--A-----G---G----AA.........-----C--T--C-----C--C-------GAGT---AG------A-CT-----CA-C---------C--A--T--------- 5996
CPZ.TZ.00.TAN1          ATT...-GATTTAGC-CT-AT...............-ACAC------G-T------A--------T-----A--G--CCAA.........C-A--CT-G-------C--T------G-TATTAC-AG--G---TAA---C--CA-A-----A------AA------G-- 6071
CPZ.CM.05.SIVcpzLB7     TTG...ATA-TTAGT-GCACTAAAAATGAAAATTTA-----G-----G--C-----A--G-----C-----A-----GAA.........G-A----C-------G---------A------C-AA-GG------C---C---A-A------T-T--A--T--------- 6016
CPZ.US.85.CPZUS         CCA...-GATTGTCT-GC--A...............----CT-----A-----------------A----GA--T-T-GAG.........--A--CT---C-----C--T-----A--G-----CA-GCAG----CC--C--CA-------------G--A-----C-- 6522
CPZ.GA.88.GAB1          ACC.....T-T--G.....A...............-----A-----A--------A--------T---C-T--T--TGA-.........C-GGTA--C-------C-----C-----G---C--AG--------CT------A-------------G--A-------- 6515
MON.CM.99.L1            TTG...-TG-AT-GGTCTAAC...............---AC---C-----C--C--T--T--A--A----ACCG-----T.........C-TC---C-------C--T---C-A--TTACGGGTC--A-----C-GGG--CAA----CTGG-GTC-T----CC-C-- 6366
MON.NG.x.NG1            RAT...T-AGG-A-C----CA...............-----------A-TC-----A--G--G--C-AT--ACCT--G--G.........C-TC-AT-G-C-----T-----CC--G-YTATGGCTCA-A-----C-GGG---AA----CTAG--TC-T----CC-C-- 4979
GSN.CM.99.CN166         GGT...TTG-AGTGTG--AAT...............---ACG-----G--C-----------G--A----G---C-----T.........C-AC-AT-G-----C--C------C-AG-TAT-GC-AG--AT--ACC-GGA--CA-----ATTT-TACA--A---T-G-- 6340
GSN.CM.99.CN71          TCG...AGGT-TA--GGGAAC...............---ACG-----G-TC-----------G--A----GA--T----AA.........C-AC-AT-G-C----C--------AG-CAT-AC-AG--AT--ACC-GGA--CA-----AT-T-TACG--A--CT-G-- 6325
MUS_1.CM.01.1085        ATA...AT-TGG-G--G-CAA...............---------------T-----------G---------A--T--T--.........C-AC-A--C--C--------------C--CAT-GCAA---G---ACC-GGG-----A---ATA--CACA--A--CC-C-- 6295
MUS_1.CM.01.CM1239      ATA...ATATGG-GC--TCA-...............------------G-TC-----------G--A------A--T-----T.........C-AC-A--C--C-----------------CAT-GCAA-C-GG--ACC-GGG-----G---AT---CACA--A---C-T-- 6282
MUS_2.CM.01.CM1246      AGA...AAACAT-CC--GCAA...............---------G--G-T----A-------G--A-CA---A-T--C--A.........C-AC----C--C---------A------TAT-GCAA-C-G----CC-GGG--C--G-T-ATA--GACA--A--CC---- 6414
MUS_2.CM.01.CM2500      CTG...ATATGG-G----CA-...............------------G-T-----A--C--G--C---------T-----A.........C-AC-AT----C-----C-----------CAT-GCGA-C-GG---CC-GGG-----G---ATA---ACA--A--CT---- 6358
RCM.NG.x.NG411          ATA...-TAGG---G---CAA...............-AT--G-----A-T------T--C--A--A------A-CAGT--A.........GTTC--A-G-------T........................--T--TAACGCAC-ATCA----GA--GTTAAAT---A---- 6134
RCM.GA.x.GAB1           GGA...T-ACA-A--CCGCAA...............-ATA-A-------TC-----T--C--A--A----GAA-CAGC--A.........GTGC--A-G------TG........................--T--TAATAC-C-ATCA----GA--T-TAAAT---A---- 6078
DEN.CD.x.CD1            TGT...A-A-------GCA-...............-AC--------A-TC-----CA-------CA----G-A-CAGTG-A.........G-TC-G--G--C--CA--......................G-GAT--ACAGAGGA--A-----AG--GCAT-A-----T-- 6445
                                     Env signal peptide end Env gp120 start                                   V1 loop start                                   V1 loop end
MAC.US.x.239            GGG...ATCTATTGT-CTC-A..............-AT---------T-----T-----A-CT---G-A-T--G--A.........-TTC-C--C---------..............................--CA--AATAGGG--AC-----GA---AC-CAG--CC---- 6779
Env                      G  .  I  Y  C  T  L  .  .  .  .  .  Y  V  T  V  F  Y  G  V  P  A  W  R  N  A  T  .  .  .  I  P  L  F  C  A  .  .  .  .  .  .  .  .  .  .  T  K  N  R  D  T  W  G  T  T  Q  C  L  P
H2A.GW.x.ALI            TTA...-TATATTGT---CAA..............-AT--G--T--T-T------C--G--A-CA-----AA-T---T--.........-TTC-C--C---------..............................--CA-AAATAGAG--AC-----GA--CATACAG--CT---- 6869
H2A.DE.x.BEN            TTA...-TATATTGT-GCCA-..............-AT--G--T--T-TC-----CA----C-C------AA-T---T-T.........-TTC-CT----------..............................--TA-AAATAGAG-CAC-----GG--CATACAG--CT-G-- 6879
H2A.SN.x.ST             TTA...ATATATTGCGTCCAA..............-AT--G--T--T-TC-----C--G--C-------GAA-T---T--.........-TTC-C--C---------..............................--TA-AAATAGAG--AC-----GA--CATACAG--CT-G-- 6321
H2B.GH.86.D205          GGG...TT-GT-TGT---CAA..............-AT--T--T----TC-----CA----C-CA----G-A-C-----A.........GTTC-C--CA--------..............................--CACAAACAGAG-CACC----GA--TGTACAG---C-C-- 6846
H2B.CI.x.EHO            GGG...TA--TG-GC--GAAC..............-TT-----T----TC-----TA----C-CA-----AA-T---T-A.........-TTC-C--C--------T..............................--CAGAAACAGAG--ACC----GA--TGTACAA--CC-C-- 6844
H2G.CI.92.ABT96         GGG...TTCTATTGT-CCAAA..............-AT-----T--T-T------CA----C-CA----G-A-C--G-G-.........GTAC----C--C------..............................--CA-AAATAGGG-CACA----G---TACACAG--CC-T-- 6192
H2U.FR.96.12034         GTG...ATATACTGT-CTCA-..............-AT-----T----T------CA------CA----G-A-C--G--A.........-TTC-C--C---------..............................--CA-CAATAGAG--ACA----GA--CACACAG--CC-T-- 6333
MAC.US.x.EMBL_3         GGG...ATCTATTGT-CTCAA..............-AT----------T------T-----A-CT----G-A-T--G--A.........-TTC-C--C--C------..............................--CA--AATAGGG--AC-----GA---AC-CAG--CC---- 5904
SMM.US.x.H9             GGG...AT-TATTGTGTTCAA..............-AT--A-----A-TC-----T-----A-CA----G-A-T--G--A.........GTTC-C--C--C------..............................--CAGAAATAGGG--AC-----GA--CACACAA--CT-G-- 6247
SMM.US.x.PGM53          GGG...ATCTATTGC-CTCAA..............-AT--------T-T------T-----A-CA------A-T--G--A.........-TTC-C--C--C------..............................--CA--AATAGGG-CAC-----GA---AC-CAA--CT-G-- 6698
SMM.US.x.SIVsmH635F_L3  GAG...AT-TGTTGTGTTCAA..............-AT--A-----A-TC-----T-----A-CA------A-T--G--A.........-TTC-C--C--C------..............................--CA--AATAGGG-CAC-----GA---ACACAA--CT-G-- 6777
SMM.SL.92.SL92B         GGG...A-CTGGTGTGC-CA-..............-AT--A---A---T------TA-T----CA----G-A-C--T--G.........-T-C-C--C--C----G..............................--CC--AATAGAG-CACA----GG--CGT-CAG--CT-G-- 6206
STM.US.x.STM            TTG...A-CTATTGC-CTCA-..............-AT-----G--A-T------T-----A-CA----G-A-T--G--G.........-TTC-C--C--C------..............................--CA--AATAGGG--AC-----GA---AC-CAA---C---- 6421
MNE.US.x.MNE027         GGG...ATCTATTGC-TTCAA..............-AT----------T------T-----A-CT----G-A-T--G--A.........-TTC-C--C--C----T-..............................--CAG-AATAGGG--AC-----GA---AC-CAG--CC---- 6247
DRL.x.x.FAO             AGT...ATAGG-..----CAT..............-----A--G--G-TC-----AACC--CAAA----GACC----T--.........--ACA---GA-C-----T..............................--G--TAAT---TCAT-C----TG--TACAAGT--CA---- 6088
MND_2.x.x.5440          GTA...ATAGT--G-G--CAA..............-----G-----G--C-----CAC---CAAA---C-TA-G----GG.........--ACAC--G---------..............................-----CAATA--TCAT-C----TG--CACAAG------C-- 6039
MND_2.CM.98.CM16        CTA...-GGTT--G---CCAA..............-----G-----A--------CAC----AAA---C-TA-G---GAG.........--ACAC--G--C------..............................-----AAATG--TCCT-A----TA---AC-AGT-----T-- 6466
MND_2.GA.x.M14          ATA...-G--T--G----CAA..............-----A-----A--------CAC----AAA---C-CCC-----GA.........--GCA---T---------..............................--T--CAACA-CTCCC------TA--CACAAGT-------- 6395
OLC.CI.97.97CI12        GGAATCAG--T--GG...CAA..............-AT--G-----G-TC-----AAC--AAA-----GA--TAGT-AT.........GAT--GT--A-G----TG..............................----CTAACA-A-G-C-G----T----ACA--T--CC---- 5958
GRV.ET.x.GRI_677        GGA......-T--GT--CC--..............-AT--G-----G-T------AA-C--A--A-----AA-TT----A.........GTTCAGGC------CATG..............................--GCCCAATACC---A-G-----A--CACCAA---CA---- 6371
VER.KE.x.TYO1_patent    .......-TGTT-AGT--GCAA..............---A-A-------T------AA----A--A-----AA-CAGCT-G.........GTGCAGGCT--C--CATG..............................--TCCCACCACA-GCT-A-----T--TAC-AAT--CA---- 6449
VER.KE.x.9063           ......TTA-GT----G-CAA..............------------G-TC------A-------A------A-CAGCT-A.........GTACAAGCC--C---ATG..............................--TCCTACTACC-GGC-A-----T---AC-AAT--CA---- 6452
VER.KE.x.AGM155         ATA...AG----A--C-GCA-..............---A-A-----G-TC-----A-----A--A-----AA-CAGCT-A.........GT-CAAGCT-----CATG..............................---CCTACTAC--GGT-G-----A--TAC-AAT--CA---- 6450
VER.DE.x.AGM3           CTT...AA----AGGC--CAA..............------------A-T------A-----A--A-----AA-CAGCT-A.........GTACAGGCT--C--CATG..............................---CCCACCACC-GAC-A-----A--TAC-AA--CGA---- 5950
TAN.UG.x.TAN1           CTG...TTATAT-GG-C-CA-..............-ATA-T-----G-T------TA-C--A--A-----AA-CAGCT-A.........GTGCAAGCT--C---ATG..............................---CCTAATACC--CC-------A---AC-AAT---A---- 6448
SAB.SN.x.SAB1C          CAG................................-AT--T-----G-T------TA-C--A--A-----AA-TAGTT-A.........GT-CAAGCC------AAG..............................--CCCCAATACG--CC-A------T-TAC-AAT--CA-C-- 6647
MND_1.GA.x.MNDGB1       GGT...AT--T--G----CAA..............-AT--G------TC-----A-----A--A----------T-AA.........--ACA-T-GA-------T..............................-----TAATTCA-G-C-C----TA--CAC-AAT--CA---- 5905
DEB.CM.99.CM40          AAA...CAG-A-A-GC--CAA..............-ACT-G-----A--------A-----A--C---GTA--T--C-AA.........GTTGA-T-G--C---A--G-CA-CT-ATC--AG...........................TCA-GA-----AGTGACA--A--CC-G-- 6131
DEB.CM.99.CM5           GCA...C-A---A-GC--CAA..............-----G-----A--------A--G--C--A---C-CA-T--C--T.........GTTGGCT----C---A--G-TA--AG-TC--A-...........................TCA-GA-----TGTGACA--A--CT-G-- 6100
TAL.CM.01.8023          ACC...T-CG-----C-GCA-..............-TCC-------G--C--C--A--G--A--A-----AA-T-----G.........C-TC-AT-G--C--C--C..............................--TAT--C-AGGG-CCAG---T-T---TCCCAA-------- 5723
TAL.CM.00.266           TTT...ATA-T--GG--CCA-..............-AT-----CA-----------A-----A--------AA-T-----A.........C--C-AT-G--------C..............................--CAT--C-AGGG-CCAA---T-A---TCCCAA--C--C-- 6203
SUN.GA.98.L14           GGT...AG--T--C--T-CA-..............-AT--A-----A-TC------ACT---AA----G-ACC----GTA.........GTGC-AT--A-A-----T.............................T---CTAATA---GCT-G----TG--TAC-T-A--CT-G-- 6585
WRC.CI.97.97CI14        ATC...AG-TT--CTGT-CAA..............-AT--A-----G--C-----A-----ACA----G----CAGT--A.........GA-TTC---C-G------..............................-----AAATAGATCAT-A----TA-GCACACG----A-T-- 6450
WRC.CI.98.98CI04        TTA...-GG-T--GG---CAA..............-ACA-A-----G--C-------------CAA---G----GAGT-AA.........GATTTCT-GC-C------..............................-----AAACAGATC-C-G----T--GCAC-CA---CA---- 6465
WRC.GM.05.Pbt_05GM_X02  TTA...-GATT--G---GCAA..............-ACA-A-----G-----------G--ACA----G---------AG.........GA-TT-A-G--G------..............................------AATAGATC-C------TA---ACACAT---C---- 5388
LST.CD.88.524           GGA...-AA-AG--G---CAA..............-AT--A-----A-TC-----A-----AAAA---G-A--T--TGTA.........GTAC--T--A-C-----T.............................T---CTAATA---GCT-G----T---TACAT-T---C-C-- 5460
LST.CD.88.447           GAG...AAA---A-G--GCA-..............-AT--A-----G-TC--------G---AAA---G-A------GTA.........GTGC-AT--A-------T.............................T---CTAATA---G-T-A----T----ACAT-A--CA---- 5460
LST.KE.x.lho7           ATA......TGG-G----CAA..............-AT--A-----A-T---------G--AAAT---G-T--TAATGTTTCA.....GTGC-CT-GA-------T.............................T---CCAATAC--GCC-C----T---GACAT-T--CT-G-- 6538
SYK.KE.x.SYK173         TTT...ATGGA-A--C--CAA..............-AT--A-----A-T-------A---ACAT---G-A--T---TAT.........G-AC----C------A-C..............................--CTCCCATA-AGGA-GA-----T--TA-GAAT-----G-- 6400
SYK.KE.x.KE51           GTA...----GGTT-G--CAA..............-AT--------A-T------A--T--CAAT---G-A--T--TCAA.........GTAC--A-G---------..............................--CCCCCATACGGGA-G------A---A-GAA-------- 6078
COL.CM.x.CGU1           GTG...TTA...AGT--TAAA..............-----A-----A---C-A--A-------CA---G-A------GATGTAAATGATCAGCAGT-C--C--CTTT--TTCCT-CCC--AG.....................A--CAACAGGTGTTG-GG---T---- 5807
```

```
H1B.FR.83.HXB2            C...ACAGACCCCAACCCACAAGAA.........GTAGTATTGGTAAATGTGACAGAAAATTTTAACATGTGG......AAAAATGACATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACCCCACTCTGTG 6603
Env                         .  T  D  P  N  P  Q  E  .  .  .  V  V  L  V  N  V  T  E  N  F  N  M  W  .  .  K  N  D  M  V  E  Q  M  H  E  D  I  I  S  L  W  D  Q  S  L  K  P  C  V  K  L  T  P  L  C
H1A1.UG.85.U455           -...---------------------.........A---AT-----C------------G-A------------......-----A---------T----------A--------------C--------C-----------------------G--------T-----C- 6048
H1B.US.90.WEAU160         -...---------------------.........----------A----------------------------......-----A-----------------------A-------T------------------------A---------------------------- 6598
H1C.ET.86.ETH2220         -...----------G----------.........T-G-GT---A-----------------------------......--------------G--G---------C-------------------------G--------------------G--G------------- 6012
H1D.CD.84.84ZR085         -...---------------------.........A--A-C--AA-----C-----------------------......-----CA-------G--C-------------------T------------------T-----A----------------------G----- 6124
H1F1.BE.93.VI850          -...---------------------.........---T-TC--AA-----A-----------G-T--------......-----A--------------A-----ACA---------------------------------------------G---------------- 5949
H1G.SE.93.SE6165          -...---------------------.........A--A-TA---A------A------T--------------......-----A--------G-----A-----------------------------G-----------------------GC-------T------- 6006
H1H.CF.90.056             -...--------------------G.........A-G--CA---AG----A-----G-GC-----T-------......G----A--------G--G--------ACA--C-------------------------T-G--A---------------------------- 5943
H1J.SE.93.SE7887          -...-----T--A-G---------G.........A-GAATC--CC-----A--------C-------------......-----------------C--------G--A-----------G--------G------T----------------GA-------T--A---- 5922
H1K.CM.96.MP535           -...--G------------------.........----AGA--A-----A--------C--------------......------A----------G--------ACA---------------------G-----------A--T------G-G--------T------- 5802
H101_AE.TH.90.CM240       -...---------------------.........A--CACC---A-----A----------------------......------A----------G--------G------G----------------------T-----------------G-----T--T-----C- 6171
H102_AG.NG.x.IBNG         -...--------T------------.........A--CAT---A-----A-------G---------------......------A----------------------A-------T---C--------C-----------------------G--------T-----C- 6130
H1N.CM.95.YBF30           T...--T--T-----T---------.........--GC---ACCC-----A--T----A-----T--------......G-----A-A----C---C--A-----A----------T---C-G-----A--G---T--------C-----T-----------T-A----- 6190
H1O.BE.87.ANT70           T...-----------CT---T-T--.........TATCC----CAC----------TG-C-----T--A----......------T-------------A-----G--A--C-T--T------------C--G--T--T--A--T-----TC--A-G--TTTC--G---- 6651
H1O.CM.91.MVP5180         T...-----T----T-----T----.........T-TCC-C-A-GC-----------T--C---G-T--A---......------T-------G--C--A--------A--C-C--T-----G-----A--G--TT--------T----AG--A--G--TTTCT-A---- 6641
CPZ.CD.90.ANT             A...-T---T--AG-T--TATT.............-----TAG-C-C---...--CTC-GTC-GGTTT-ATGCTTAT..-------T-T---------AGT---ACA--A----GN-ACAA----TCC-A-------AT----------------C----A--TA-G---A 6026
CPZ.CM.01.SIVcpzCAM13     T...--------T--T--T------.........-----TC-TCAT----A-----GG-C-----T----AAT......------C-G-----------------A--A--C-C-GT---C-C------C--G--T--T--A-----------GC----T--TT-A---- 6181
CPZ.CM.05.SIVcpzEK505     -...--T------------------.........---C--A-ACC-------------CG---G-T----AAA......-----CA-T--------T--A-----A--A--C-T--T--C--G------A--G--TT----A--C-----T--G--------T-A----- 6152
CPZ.CM.05.SIVcpzMB66      -...---A----T------G-G--G.........A---A-C-AAC----A--T---T-C--C--TGCA--GAA......-----CA-T-----G--A-----------------G---A-C-T------C--G--TG----A-----------------TT--TT-G--- 6185
CPZ.CM.98.CAM3            T...-------T-------------.........--GCAGA-CCC-----A--T----C-----T--------......--C---AC-----C------A--A--ACT---C-T--T---C-G------------TT-G--A--------T--GC-G--G----A-A--- 6022
CPZ.GA.88.GAB2            T...---------TCT---G----G.........--CA-CC-AA-T-----------G-A----TG-A-----......G-C--CT-------G--A--------G-------T-CT--CC--------C--G-----T-GA--T-----------G--A--TT-A---- 6008
CPZ.CM.98.CAM5            T...GT---T--T-----C-----G.........--GTATA--CC-----AT-T----GC-------------......------A-T----C------A-----G--A----T-GT---C-G------------T-----A--------T---C-G--A----A-A--- 6307
CPZ.CM.05.SIVcpzMT145     T...--T--T----CT--T-----G.........---C-C--ACCT-----C--T---G-G------------GAA...------T----------T--A-----G--A--C----A---C-C------A------T--T-A-----------------A--CT-A---- 6148
CPZ.TZ.00.TAN1            T...TT---T-----T--TT-T---.........---ACCC-A-CC------T--AT--GG----T----GAA......G-----T-------GC-G----A-A--A-------T-ATCAC-T-TTC-A--G--TT-T------T--------------A---T-T--CA 6223
CPZ.CM.05.SIVcpzLB7       -...-----T--TG-T--TA---C-.........A---A-C-AAGC----A---------------T------ACA.........--CA-----A-T-------AC-----G----T--C---------------------A-----A--GC-------T--T------- 6168
CPZ.US.85.CPZUS           T...--T-------T----------.........---CAT---CC--------T----G---G----------......G-----A-T---C-------A-----G--A----T--T--CC--------C--G---T----A--T--A-----G--------T-G----- 6674
CPZ.GA.88.GAB1            T...-----T---GT--T--G----.........---T-TC-TCC-----A-T---TCA--------------......------A-T-----G-C--A---------A-C--T------C-C---------G--T--C-----T--------G--G-----TT-A---- 6667
MON.CM.99.L1              G...--------T-C---G---TCT............C---ATT-G--CA-C--T-GG-G--C--TGCT-ATCAA.......--CT-------G---G-A---GTA-------G-AGTC-C-G-TCTCG--GGCTT--------T-----G--------A--CA-G---- 6515
MON.NG.x.NG1              T...T----T--TGTT----GC-CT............C-CCCTC-C---A-A------GC-----TGCA-ATAAC.......--CT-------G---G----GTG--A--C----AATCAC-C-TCTCA--GGC-T-----------------G--------TA-G---- 5128
GSN.CM.99.CN166           -...T-G------TCA--GGC---G..............--GCC-T----CA-A-----G--A--C-----C-AT......------T---------TG--G---GA--T--C--GG-ATCAC---TTA----GGC---G---A--------C-----G--T---A-G---- 6489
GSN.CM.99.CN71            -...T-------TCA--GTCT--G-..............--GCCAC----CA-A-----GG-A-----T--C-AT......------T-T-------TG--G---AGA--T--C--GG-ATC-C---TTA----GGC------GA--------C---C-G--T--GA-G---- 6474
MUS_1.CM.01.1085          -...--G--T--TGCA--C--G---..............--GCA-T-G--A-CT-T-GGG-AAAGTTTGACATCTATAAG..-----T-T--------C-----AC---T---G-G-GGTC-A-C-TTA----GGC-T-----A--------C-----------TA-G---- 6447
MUS_1.CM.01.CM1239        -...-----T---GCA--C--G---...............---CAAT----A-CT-T-GGG-A-A-TTT-ATGTCTATAAA..-----T-------GACC------GT---T---G--GGTCCA---TCA-C--GGC-------A--C-----C-----G-----TA-G---- 6434
MUS_2.CM.01.CM1246        -...--G------GCG-----G--G...............---CAAC----CA-CT--A-TG-GCACTTT-ATATATATAAG..---CT-T--------T-----AC----T--G-G-AGTCCA-C-TTAC---GGC---T------G-----G---C----A--TA-G---- 6566
MUS_2.CM.01.CM2500        -...--------AGC---T-----G...............---CAAT----CA-C----GGG-A-A-TTTGATATCTATAAA..---CT-T-----G--T--A---AC---T---G-GCGATCCA---TCA----GGCT--T--A--C----------C----T--CA-G---- 6510
RCM.NG.x.NG411            A...GA--GAGG--TAT---C----...............---CCA---------T----G--A---G-TGCT---GAA.........---AG-T--TAT----AGCAA-GTCTA--G-GTGGCA-C-T-AT---TCC-CTT-G--A------------GG----A---A-G---A 6283
RCM.GA.x.GAB1             A...GAG-GAGGA-TAT-T-C---G.............--T-CAA--------T-----CGA---G-TGCT---AAT.........---AG-T-ATAT----AGCAA-A--TA--G-GTGG-A-C-T-AT---TCC-CT------A---------C-G-----G----T-A--A 6227
DEN.CD.x.CD1              T...GA-------T-T-A-AGG--GGTCTGGATG-A-AATGCCAC-G--TACTTT--C-TC-A-C---ATCCA.........-TA-TAGACCAG-T-GT-CA-A--AT-A-AGAT......C-G-ATA-A--G-C-T-C--A-T------C-------A--TG-A---- 6597
                                                                                                                                                 |_V2 loop start

MAC.US.x.239              A...GATA-TGGTG-TTATTC----...........--GGCCC-T----T------GC---G-TGCC---AAT.........---ACAG-CAC------GCAATA-----G--TGGCAAC-C-TT--GACCTCAA-------T-----------T-----T-A--CA 6928
Env                         .  D  N  G  D  Y  S  E  .  .  .  .  .  V  A  L  N  V  T  E  S  F  D  A  W  N  .  .  N  T  V  T  E  Q  A  I  E  D  V  W  Q  L  F  E  T  S  I  K  P  C  V  K  L  S  P  L  C
H2A.GW.x.ALI              A...GACA-TGATG-TTAT--G---..........A--GCTT-G-----------GGC---CG-TGCA---GAT.........---ACAG-AAC------AGCAGTA--A---G-CTGG--AC---TT--GAC-TCAA---A--------C--G-----A--TT-A---A 7018
H2A.DE.x.BEN              A...GACA-TGATG-TTAT--G---..........A--A-TT-------------GGC----G-TGCA---AAT.........---ACAG--AC------AGCAGTA--A---G-CTGGCA-C---TT--GAC-TCAA---A--------C--GC----A--T--A---- 7028
H2A.SN.x.ST               A...GACA-TGATG--TAT--G---..........A--GCTT-------------GGCC--CG--GCA---AAT.........---ACAG-AAC------AGCAGTA--A---G-CTGG---C---TT--GAC-TCAA---A-----C--C---C----A--CT-A---- 6470
H2B.GH.86.D205            A...GACA-TGGTG--TACACT--G.........A-CAG-C----CA-A-----GGC----G-TGCA---GAT.........---ACAG--AC-C-----GCAGTA--T---G-GTGG--AC-C-TT--AACCTC-A---A--------C---C----------G---- 6995
H2B.CI.x.EHO              A...GATA-TGATG--TATACC---..........A-CCAAT-----A-A-----GGC----G-TGCA---GAT.........---ACAG--AC---T---GCAACAA-----G-GTGG---C-C-TT--GACCTCAA---A--------T---------------G---- 6993
H2G.CI.92.ABT96           T...GATA-TGATG-TTACAGT---..........C--GCAT-G-----A------GC----G-TGCC---GAT........--CACAG-AAC-------GCAGTA--A---G-TTGG-A-C-G-TT--GAC-TCTAC---A--T-----T--------A--T--G--CA 6341
H2U.FR.96.12034           T...GATA-TGGTG-TTATTC---G.........T--GCAT-------C-----GGCC---G-TGCC---AAT........--CACAG-AAC---G--AGCAGTG--A---G-CTGG---C-T-TT--GAC-TCTA---A-----------G--G--A--C--T---A 6482
MAC.US.x.EMBL_3           A...GATA-TGGTG-TTATTC----..........T-GGCCC-T-----T-------GC---G-TGCT---GAG.........---ACAG-CAC-------GCAATA-----CG--TGGCAAC-C-TT--GACCTCAA-------T------------T-----T-A--CA 6053
SMM.US.x.H9               A...GATA-TGGTG-TTACTC----..........T-GGCAA-T-----C-----GGC----G-TGCT---GAT.........---ACAG-CAC------AGCAATA------G--TGG-ACC-C-TT--AAC-TC-A-T-----C---------C-C---------A---A 6396
SMM.US.x.PGM53            A...GATA-TGATG-TTATTC----..........T-GGCCA-------T-------GA---G-TGCC---AAC.........---ACAG-CAC-------GCAATA------G--TGG-A-C-C-TT--GACCTCTA---A--C-----------------CT-A--CA 6847
SMM.US.x.SIVsmH635F_L3    T...GATA-TGATG-TTACTC----..........T-GGCA--C---A-C-----GGC----G-TGCT---AAT.........---ACAG-CAC------AGCAATA------G-GTGG-ACC-C-TT--AAC-TC-A-T-----C---------C-T---------A---A 6926
SMM.SL.92.SL92B           A...GATA-TGGAG--TATTC----..........T-GGCCC-T-----T-----GGCC--CG-TGCT---GAT.........---ACAG-AACT-----AGCAATA--A---G-TTGG-A-C-C-TT--AAC-TCTA-T--A--T--------G-----T---T-A---A 6355
STM.US.x.STM              A...GACA-TGGTG-TTACTC----..........T-GGCAA-------C-----GGC----G-TGCC---GAC.........---ACAG-CAC------AGCAATA------G-CTGG-ACC-T-TT--AAC-TC-A-T--A--T-----------G--T-----A---A 6570
MNE.US.x.MNE027           A...GATA-TGATG-TTATTC----..........T-GGCCC-T---A-T-------GC---G-TGCT---GAG.........---ACAG-CAC-------GCAATA------G--TGGCA-C-C-TT--GACCTCAA------T------------------T-A--CA 6396
DRL.x.x.FAO               A...-GTTTATTGC-TTATG----.........AC--A-A-TAAC---A-AGAG----------CTG-CCCTATGACAG-------GG--A--A-----GCATGG-GT-CGC-GTCATCCA-GATA---GC-GTG--T--A--------T---A-C-AT--TTAT---- 6246
MND_2.x.x.5440            -...-GTCTGTTGC--TATG----.........CAGCATA-CCCC---A-A-----G--------CAGGCCCTATAGAAG-G-----A--A--GAC---AGCTTGG-GA-CA--TTCTTCCA-GATA---GC-GTAT----A----------GG-----A--TTAT---- 6197
MND_2.CM.98.CM16          -AGCTTGTTG-A-T-TGA-G-GC--.........CC-A-TCC-AAC-TAACATGGA-TTT...C-CAGGTCCCATGGAAG-------GG-A---ATG--AGCATGG-GA-C---CTC-TCCA-GATA--CGC-GTAT---A--T-----G---C-G------TA----- 6624
MND_2.GA.x.M14            A...-GCTTATTGC-TTATG-G--G.........CA-C-CA-CCCC---A-A-------------CAG-CCCTATACAAG-G-----AG-----AGG--AGCATGG-GA-C---TTC-TC-A-GATA---GC-GTT-----A--G-----------G--A---TA----- 6553
OLC.CI.97.97CI12          -...T-TTTTGATG-AGATG-GAT-.........-A-CA-CCCA-CCC--GAGT----G-GG-ATTTTATAATGTGTCAG-C-----AG-A---AG---AGCATGGTTT-GC--TGCTTCC--G-TA---GC-GTT--G--A--------G--G-----T--TT-T---- 6116
GRV.ET.x.GRI_677          A...GAT--T-ATG-TAATAC---G.........--GCCTC-A...--CA-T------GC---CG-GGCT---GAT.........---CCGC-----A----AGCAG-GAGTA-----CATCTAC-C-TT--A----CGA-G-G---T-----T--GC-CT----CA-A---A 6520
VER.KE.x.TYO1_patent      A...GAT--T-ATG--TATAC----.........---CCTC-A...---A-C--T---CCA---G-GGCA---GGTGAT-G---CCCAT-AA---C---AGCAGCCAGTA-C--CCATTTAC-C-TT--G----CTA-G--A--T-----G--G---T-A-----A--CA 6604
VER.KE.x.9063             A...GAT--T-ATG-TTATAC---G.........--GCC---A...--CA-T--T--GCCC---G-AGCC---GCAGAT-GG---CCAT-A--G-C----GCAGGAAGTA----TCATTTG--G-TT--A---CAT-G--A-----C---------T-T--T--A---A 6607
VER.KE.x.AGM155           A...GAT--T-ATG--TATAC----.........---CC-C--...---A-A----GCCA---G-AGCA---GCAGAC-G----CC-T-A----C---AGCAGGAAGTA-C--TCA-CTGC-G-TT--A--G-CAT-------C---------GC--T-A--T--A---A 6605
VER.DE.x.AGM3             A...GAT--T-ATG--TACAC---G.........-----CCAT----CA-C--T---CCA---G-AGCA---GCTGAC-G---CCC-T-A----C---AGCAGGAAGTA-----CA-CTGC---TT--G--G-CT--G------------------T-A--TT-G--CA 6105
TAN.UG.x.TAN1             A...GAT--T-ATG-TTATACT---.........---CAG--A...-----CT-T--G--A---G-AGCA---AAGGAC-GG---CCAT-A--G-C----GC-G-GAGTA-C---CATTTGC-C-TT--AAGC-CT--G--------C------C-G--A--TA-G--CA 6603
SAB.SN.x.SAB1C            -...GAT--TGAACCAGA-GG-AC-ATAGCAGAA---CCCA-ACCT---A-C---------A---G-TGCT---AAAAATCGG--CCCTC-A----GC--AGCAG-GAGTA-C---CA-CT---G-TT--AAGT-CTT--------------T--GC-CT-A---A-G--CA 6814
MND_1.GA.x.MNDGB1         T......TCATTGCCAGATT-T--TGAG..........----A-A-TCCTG--A-A-AG----------CAGGACTTATAAGGG-----C-G--A--TT-T--AGCATGGC-T-C---GGGA---A-G-TA---ACC-TA--T--------------GA-T-A----TAT---- 6063
DEB.CM.99.CM40            ACATG-CTTAGTG-GAGA-G---TCCCAATGCCAAAT--GACCCA---C...TTTA-TGC----G-T-ACCCC...............--A-AG--G---C-ATGGC-------G-CT-----G-ACA-A--GTCAT-C--A-----------GC----A---TAT---- 6283
DEB.CM.99.CM5             ACATG--ATGGTT-GAGA-----TCCCTATGCCTAAT--G-CCCAG...TATTTT--CTC----TCA-ATGTT...............----A-C----AT-ATGG--C--C--G-CT--C----ACAGA---TCTT-T--A--C---------G----------TAT---- 6252
TAL.CM.01.8023            A...T-------TC-GAAT-T---G..........--GCCA--C---T----T--GT-C---G-TG---AT..........-----CT-------G--C--C--AGCA--T-----GGA-GCA--G-TCTTAAC-GCAACC------T--CACC---C-----T--TA-G---- 5872
TAL.CM.00.266             A...T----T--TC-GAAT-TC--G..........---CCA--G---C-C--T--GT-C--CG-TG---AT..........------T-------G--C--C--AGCA--A--C--GGAT--CC-C-TTTTGAC-GC-AC--GA------ACT--G-----A---A-A---- 6352
SUN.GA.98.L14             A...GATTTA-AA-CTTATGC----..........--GCCCA-AAC-GGGT-AGA---G-------CAGA-G-AATAAGC--T---C-A--A--GC-G--AGCATGGC-A-C---G-C-TC-A-GGTA---GCC-TTA-G-----T--------GA-C-AT--CTAT---- 6743
WRC.CI.97.97CI14          A...CATTTGGAGC-TTTGG----G.........CAGTACA--CCTG-GA-A-A----------C-CA-GAAAAGTAGCAG-----C-A--A--GC-G---GCATGGACA-GAG---CTTCAA-GATT--CACCTTAT----------------A-C--A--CTAT---- 6608
WRC.CI.98.98CI04          A...CATATAGAGC-TTGGG-----.........CA-TATC-TCCT-----A-A-----------CT-G-AA-GTAGCAG-C---C-G--AA-CC----AGCATGGACA-GGG-T-CATC-A-GATT---ACT-TAT----A--T-----T--GA----A---TA----- 6623
WRC.GM.05.Pbt_05GM_X02    T...CATATAGAAG-AAT-AC----.........CA-TATA---A-GG-A-A-A-----------CA-AAAA-GTGGAAG-------GG-A---C-G--AGCATGGACA-GAG-G-CT---A-GATA---ACTCTGT-G--A--C----------A---AT---TAT---- 5546
LST.CD.88.524             A...GATCTG-AATCTTATGCTC-G.........A-CCCTA-TCAT---A-T--TCT---------CTCAAGAAATTAAGG-C---C-A--AA--C-G---GCATGG-GA-C---G-CG--CA-GGT----GCT-TTA-G----------AC----A---AT--TTAT---- 5618
LST.CD.88.447             A...GACTTG-AGTCTTATGC-C-G.........A-TCCTA-TCAC--CA-A-GTTT------C-CT-AAGA-ATAAAAG-T---C-A--TA--C-G--AGCATGGAGT-C---G-C----A-GGT----ACT-T-A-G-----T---------G---AT---TA----- 5618
LST.KE.x.lho7             A...GATTTG-AATCTTATGC---G.........---CCTA-CTAT---A-TT-------C----CA--ACCAGTAAAAG-C---C-GG-TA--C-G--AGCCTGGTCT-C---G-ATGC-A-GGTA--CAGT-TTA-G--A------------A-C-AT---TAT---- 6696
SYK.KE.x.SYK173           G...T---CAGATC-AAT-G---T-............AGGG-CAAC-T-ACAGG---T-C---CCAGCA---AATAGC-GTC-CATG--AAGGC-G--A--ATTA-----C--GTCAGC-----TCTTG---GC-AAT-GA--C-----G--GC--G-A--CA-G---A 6555
SYK.KE.x.KE51             A...T-TTCTGAAC-GAT-G---TC............AG-G-TAAT-T-AGTGGG---T----C-CTGC----AATTCTTCCC---G---AAG-C-----C-TTTGC----C--GTCTGA-C-T-TCCTG--GGCTGAT-G---T------------T-T--TA-G---A 6233
COL.CM.x.CGU1             ACCTC-GCCTGGA--G---GT----.........CA-AACA--CC------A------GC----G-TC-T-TT......--G---AG-T-CAGT-GGG--G-ATGGATAATA-C-CAG-C-ACC-TA--A-----G----GA-------CT--G--G--AG--TA----- 5962
```

```
                                 V1 loop start
H1B.FR.83.HXB2              TTAGTTTAAAGTGCACTGATTTGAAGAATGATACTAATACCAATAGTAGTAGCGGGAGA.............................................................................................. 6662
Env                         V  S  L  K  C  T  D  L  K  N  D  T  N  T  N  S  S  S  G  R  .__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.
H1A1.UG.85.U455             ---CG---G-T---CA-A-CA-C-CC-T-A-C-AC-CC-AT--C-AC-CC-AT..................................................................................................... 6101
H1B.US.90.WEAU160           ---C------T------A--G----TGTGAC--A-TTG-AG---GAG-C--ATACC-ATAGTAGTAGTGGAGGG................................................................................ 6672
H1C.ET.86.ETH2220           ---C------C--T-A--C-A-C--A--CA-------GGTT-CA-A--A---T..................................................................................................... 6065
H1D.CD.84.84ZR085           -C-C------T---------GCATC--GAA-C-GC-C-GAT----AC--C-CA..................................................................................................... 6177
H1F1.BE.93.VI850            -G-C------C--T---A--GCC-CC---A-C-G-C-GGAA--GCCAG-GGCA..................................................................................................... 6002
H1G.SE.93.SE6165            ---CC-----C--T------G-A-CC--CA-AGGC--C-AA-GA-A--A----ACTGATAATAGCACTGAAACCAACAAT.......................................................................... 6086
H1H.CF.90.056               ---C-C----C--T---A--G-C-GA--CA----CTC--A--GC-C---C--T..................................................................................................... 5996
H1J.SE.93.SE7887            -C-C------T--T-G-A-CA-C-CT-G-A-C-G------A-C-----AC--TA-T.................................................................................................. 5978
H1K.CM.96.MP535             -C-C------T--T-------AC--AGGAACC-A--GC-------A-GCA-CTA-C-CT............................................................................................... 5861
H101_AE.TH.90.CM240         ---C------T--T--CA--GCT--TTTGACC-A-GGC-GT-GC-AA-CC-AT-TCTCTAAC............................................................................................ 6233
H102_AG.NG.x.IBNG           ---C----G-T--TCA-A----C--C--CAGCTAC-GC-A--G---C-ACCTAACT--T............................................................................................... 6189
H1N.CM.95.YBF30             -A-C-A-GCTT--T-AC---AGCT-TGGG--GGAA-GG-A-----CA-A-.....................................................................................ATGACAACAAGA 6252
H1O.BE.87.ANT70             -ACAAA-GG----T--AA-CA-AGCTGGAACA--A---GAA................................................................................................................. 6692
H1O.CM.91.MVP5180           -ACAAA-G--C--TGTA---C--C-A-CAA---AA-CAGG-................................................................................................................. 6682
CPZ.CD.90.ANT               -A-AAA-G--T--T--A-GA-AC--TGGAACAC---CA--ACCA----CA-CAACA--TACA.......................................................................GTAACACCAAAGACAACA 6106
CPZ.CM.01.SIVcpzCAM13       -A-CAA-G--A--T-G-A-CG-T-CT-GAC-A-A--GC--T-GC-ACTC--A-............................................AATAACCCTACCCTCAAGAACAATACTTCAGGT 6267
CPZ.CM.05.SIVcpzEK505       -A-CCC--TCA--T-GCTCA-G--G--G--TG-AC----GT................................................................................................................. 6193
CPZ.CM.05.SIVcpzMB66        ---C------C--TT-ACTG--T--ATGCAT--AAG-A-ATGGC-AC-CC-C-AACT-C........................................ACTGTTCAAATTAGCACTGGTAATGATAGCACTGCTAAC 6283
CPZ.CM.98.CAM3              -C-CCA-GG-A--T-GAA-GG-T-CTTTCA---GC-C-T-T-----G-A--AAACTTCTACTATGACTACT...........................................................................AACAGT 6099
CPZ.GA.88.GAB2              -GCAA----C---T---TC-A-A--TGC-AC-GGA--CGAA-C--A-GC--CT--A.........................................................................................AACGGAATT 6073
CPZ.CM.98.CAM5              -C-CCA-GC-A--T--GACCC-T-----CAT--GC-C---T----C--AC--T.............................................................................................ACA 6363
CPZ.CM.05.SIVcpzMT145       -G-CCC-T-CT--T-A-A--CCC-CT---AC-T-CTGC--T.........................................................................................................AAC 6192
CPZ.TZ.00.TAN1              -A-AGA-G-CA--T...ACAA---CT---ACC--A----AA-CC.............................................CTGAATTCGGCAACAACAACCTTAACACCAACAGTAAAT 6303
CPZ.CM.05.SIVcpzLB7         ---C----C-T----G-ATCCCT--ATT-----ACTC--G-................................................................................................................. 6209
CPZ.US.85.CPZUS             ---C-A-G-CT--TCT-A-CCCCG-T-G-A--..................................................................................................................AGT 6709
CPZ.GA.88.GAB1              -A-CC--GC------G-A-GGCT--CTT-AGCCAGGCA-AA--CCTA-CA................................................................................................AAC 6720
MON.CM.99.L1                ----GA-GCTC--TGT---AG-T--C-CC-T-T-G---G---G--CC-C-CCT-CCCCTAGTACCCCCACACCC...........................................................TGGGGAAATTGG 6601
MON.NG.x.NG1                -A--GA-GCG---TCAAARGA-C--TG-AAGC--A-CAGTG-CRCCA-C--C-ACC-CYWCC...GCCACCACACCA.........................................................TGGGGAAGTTGG 5214
GSN.CM.99.CN166             -GG-GA-GT-T--T-ACCTCACA--C-CATCC--A-GCGAACCA-CA-CA-CTCCA--GCCACCAAATGTCTCCACAACCACG.............................................CAGTGGGGAAGTTGGGGG 6590
GSN.CM.99.CN71              -AC-AA-G--A--T-AGCTCCCA--T-CGAGC--A-CA--A-CACCGTCC-CAACA-C-GCCACAACATCACCCTCATCCCCAACGCCAACACCA...........................................TGGGGAAATTGG 6581
MUS_1.CM.01.1085            -A-AAA-G--A--TG-GA-AG-A-C----T-C--C-GAT-T..................................ACGGGAGGGAATAGCACCGGAACCACAGCACCCCCAACAACACCATGGGGATCATGGGGCAAT 6551
MUS_1.CM.01.CM1239           -C-AAA-G--A----G-AGAA-A--TGTAACA-GGCCAT-A......................................ACAACAACAACAGCATCAACACCAACAAAAACAACACCAGTGACCCCTTGGGGGAAATGG 6535
MUS_2.CM.01.CM1246           -A-AAA-G-----TGACCT-C-C--T-T-AG--TGTC--GT..................................ACCACCGCAGCAAGCGCCACAACAAGCGGGCATACAACCGTCACCCCTTGGGGCAGATGGGCA 6670
MUS_2.CM.01.CM2500          -G--AA-G--A--T--CA-GC-A--TTCAACA-GGCCG---.............................................ACAAAGCCGACCACGCCCAGTACCACTCCATGGGGAAGGTGGTCT 6596
RCM.NG.x.NG411              -A-CAA-GT-T--T----CA-----TGG-AG-TGGG--GGT-C-GCC-CCCAATCACCTACAACAACAACAAGAGCAACAACAAGAACAACA...AAAGAAATAGGGAAAGATTGTTATGCGAAT..................CAAACAGAGTGCAGT............ 6420
RCM.GA.x.GAB1              ---CCA-G--T--TT-A-CAA-A--TGG-AGCTGGG--GGA-TCCC--CCTCA-CACC-CCAACAACAACAAAAACAACAACACAAAGAACT...ATAGGTGTAGAAAAGGAATGTACTGCTGGC.................AACGAAACATGTGAG............ 6364
DEN.CD.x.CD1                -C--AA-G--T--T--AAGAAAAG--G--A-C-GC-CA-A-GCCGTC-CCGTGTCCTCTACCACAACCACAACCACAACCATAAAGACCCCAACCACA....................................................GAGTGG 6701

MAC.US.x.239                ---C-A-G-GA----A-A-AAGTG---CA----GATGGGGATTG-CA-AATCAATA-C-ACAACAGCATCAACAACATCAACGACAGCA..........................TCAGCAAAAGTAGACATGGTCAATGAGACTAGTTCTTGTATAGCCCAGGATAAT 7071
Env                         I  T  M  R  C  N  K  S  E  T  D  R  W  G  L  T  K  S  I  T  T  T  A  S  T  T  S  T  T  A  .__.__.__.__.__.__.__.__.__.__.S  A  K  V  D  M  V  N  E  T  S  S  C  I  A  Q  D  N
H2A.GW.x.ALI                -AGCAA-G-----T-GCA-CA-A-GC-CA--G-GC-CA---CATCCCCG---CCAG-GAGCACACTC..........................................AAACCCCTGATAAATGAGAGCGATCCATGCATAAAGGCAGACAAC 7131
H2A.DE.x.BEN                -GGCAA-G--T--T-GCAGGG-TC-AGGGA----C-CG---CCG-A-CCC--GACCTCGAGTTCCACAACCTCGAGACCACCCACATCC.....................GCAGCCTCCATAATAAATGAAACTTCTAACTGCATAGAAAACAACACA 7165
H2A.SN.x.ST                 -AGCAA-GCGT--T-ACAGCACA-CTGCAA-A-AC-CA---TCC-CACCA-CAACC-CCACAACAGCA..........................................AACACAACAATAGGAGAGAATTCTTCATGCATACGCACAGACAAC 6583
H2B.GH.86.D205              -GGCAA-G--C--T-G-A-AACCG-A-CAA-CC-AGGG-ATGCC-------CTACC-CCACTAAGCCTACTACCACCTCT.............................CGTGGGCTGAAAACGATTAACGAAACAGACCCATGCATAAAAAATGACAGC 7126
H2B.CI.x.EHO                -G-CAA-G-----T-A-A-GAC-TG--GCTCAG---GC-AAG-G-CC-C--CGTCCTCCGCGTCCCTCAGATCTTCTACTCAG.............................ACCCTACTCAACGAAGATAGCAAATGTATTCAAAATGACAGC 7118
H2G.CI.92.ABT96             --GCCA-G-GA----A-A-AACTG---CAA-C-AATGGGGTCTA-CAG-G-C-ACC-CTGTAACCCCTGCAACAGTTCCCACAACTAAAATGGTGACAGCT.......................GAACTAGTGAATAGCACTAGCCAGTGCTTAATGTACGATAAT 6484
H2U.FR.96.12034             -AGCAA-G-GA--T-G-A-AACTG-A-CAA---GATGGGGATTA-CAG-A-CA-CAGC-ACAACAACTAGCCCAGCAATGCAAGAATCAAGT................................AAAGTAATAAATGAAACAGAGCCTTGTATACGAAATAATAGC 6616
MAC.US.x.EMBL_3             ---C-A-G-GA----A-A-AAGTG---CA----GATGGGGATTG-CA-AATCATCA-C-ACAATAACAACAGCAGCACCAACATCAGCA.....................CCAGTATCAGAAAAAATAGACATGGTCAATGAGACTAGTTCTTGTATAGCTCAGAATAAT 6202
SMM.US.x.H9                 -A-CAA-G-G---T-A-A-AAGTG--MCA-MC-GATGGGGTTTA-CAG-AGSACCAGC-CCAACAACAACACAAACATCAACAACACCACCATCACCAATAATAGCA..................AAGGTTGTAAATGACAGTGATCCTTGTATAAGAAGTAATAAT 6545
SMM.US.x.PGM53              ---CCA-G-GA----A-A-GACAG-A-CA--C-GATGGGGGTTA-CA-AG-AT-AA-T-CCAACAACAACAACAACAACAACAACAAAAAGCCCAAAAGCA............GAAGCAATAACAGCAAAAGTTATAAATGAAAGTGATCCTTGTATAAGCAACAATAAT 7005
SMM.US.x.SIVsmH635F_L3      -AGCAA-G-GA--T-A-A-AACTG---CA----GGTGGGGTTTG-CAG-A--A-CAGAGACAACAACAACAGCGAAATCAACAACATCAACAACAACAACAACA..............GTAACACCAAAGGTTATAAATGAAGGTGATCCTTGCATAAAAAATAATAGT 7081
SMM.SL.92.SL92B             --GC-A-G--A--T-A-A-AAATG---CA--C-GATGGGGGTTG-CA--AGCA-CT-CTACTACTAGCTCACCAACTACTACTAGCCCCTTAACTGCTGCTAGCCCATCAGGAGAG............GAAATCGTTAACGACACTATGTCTTGTACAAAGAACAACAAT 6513
STM.US.x.STM                --GCAA-G-G---T-A-A-GAATG-A-CA--C-AATGGGGATTG-CAG-G-AAACAGT-ACAACAGTGACACCAACAGCAGCAGCAGCAGCAACA.........................AAGCCAGAGTTAGTAAATGAAACAAGCTCCTGTGTAAGTAACAACAAC 6713
MNE.US.x.MNE027             ---C-A-G--A----ACA-AAGTG-A-CA----AATGGGGATTG-CA-AATCATCA-C-ACAACAGCAGCACCAACAACAAAAACAACAACAACA.........................AAGGAAATAGAAGTGGTCAATGAAAATAGTACTTGTGTAAATCGTGATAAT 6542
DRL.x.x.FAO                 -G-AAA-GG-A--T-G----C-C--A-TA-TA-A-GCAG---GC-CA-CACCAACACC-ACTACA...........................GTAGCTACTACTGTTAATACTAGTGTTGCCTCAACAACAACAGCAAATGACACTTTGGAAATAGCTGTAAATAATACA 6389
MND_2.x.x.5440              -G-AGA-G--A--T--A--AGGTC-A-----A--AG-GCAAGCA-CAGC--AGACC-C-ACACCTGTACCCACAACAACAACCCCCTCCACTACCACT.....................AGCTCCAGCACAAATAAGACAACAACTCCTGTGTTGGTTGTTGAGAAACAA 6346
MND_2.CM.98.CM16            -C-AGA-GC----T--AA-AGGA--AG-CCT-G---CCC-T-TACCC-C--CATCC-CCACAACTACCACAACTACCACC...............................AAGACGGTGGCAAATACTACTGAGTTGGACATAGACACCAAT 6746
MND_2.GA.x.M14              -C-AGA-GC-A--T--AA-GGGAG-A--CAG--G--CC-G--C--CC-C-CC-AT-CCCACTACCACCACTCCCACCACTCCA.......................................GCAGTGACGGGGCTGGAGGTGACGGTACAG 6666
OLC.CI.97.97CI12            ---AAA-GC-A--TGAA---CCATCAGCAA--TT--CAGAACC--T--CACCTATC--TACAGCTGCTCCA.......................ACAACTGTCACTAGCAGTACAGGAGGTCAGACCACTACAATAGGGTTAGGAGTTGATAATAGT 6250
GRV.ET.x.GRI_677            ---AAA-GTCC--TGTA--GC----TGG-ACAG-C-CG--A--GGCC-CC-CTACTGC-ACTACAACAATGACT......................ACCCCCTGTCAGAAT...............................TGCAGTACAGAGCAGATA 6627
VER.KE.x.TYO1_patent        -C-AGA-G--T--TGTA--G--A--TTCCACA-GAG-A-GGGCG-CA-CACCTACA-CGACGCCGAAATCTACCGGC.......................CTACCCTGTGTAGGCCCGACGTCAGGT...GAAAATCTACAGTCCTGTAATGCAAGCATTATA 6741
VER.KE.x.9063               -A-AAA-G-GC--TGTA--G-----CTCAAG--GA-GA-A-CCA-C-CCA-CTTCAGCTTCAACAACAAAAGCTCCAAAAACAGGT...................GATCCCTGCATTAAATCTACCAATAATAATGTTAATTTACAGCCCTGTAATGCCTCCCTTATA 6756
VER.KE.x.AGM155             -C-AAA-G--T--TGTA--G--A--AGGCTCCG-A-CCT-T-CCCCAGCA-C-TCT-CTACGGCAGGAACCAAA......................CTACCCTGTGTTAGAAATAAAACAGAC...TCCAACCTACAGTCATGCAACGACACCATCATA 6739
VER.DE.x.AGM3               ---AAA-GTCC--TGTA--A-----CTCCTC-GAGCC-----CC-C-CC--AAA-T-CCACGGCCTCAACAACCAATATCACAGCCTCAACAACC.........ACTTTGCCGTGTGTCCAGAACAAG...ACAAGTACTGTGTTAGAATCATGTAATGAAACAATCATA 6263
TAN.UG.x.TAN1               -A-AGA-G--T--T---A-A--A-CA-G-ACCG-CCC---ATCA----CCCCTACATC-AGCAGCACTACG.......................GATCCCTGTCCAAATACC...............GACGAAAGTAGCTGTAACGCCACCTTAGTT 6722
SAB.SN.x.SAB1C              -A-AAA-G--C---TA-CGC--AG-AGGA-GAG-AGCA--A-CA-CATCACCATCA-C-TCAACAGCACGCCCCGAGGTAGTGAGT.....................................GTGGGCTTTAATGACTCAGTGATA 6924
MND_1.GA.x.MNDGB1            ---AGA-GC-A--TCAG--AACAG-A----TAT-AGCA--A-CAGC--AGCCTATA-CTACACCTACTACTACATCTACAGTTGCAAGTAGTACA..........................GAGATTTACTTAGATGTAGATAAAAAT 6185
DEB.CM.99.CM40              ----CA-GC------TAA-AACATCA-CAA--C-A-CAC-A-CA-AC-CATCAACA-C-ACAACAACAATAGCAACAACAACA............................................AAAACAACTACAGATTGGTCA 6387
DEB.CM.99.CM5               -GC-GA-G--C--T-G-AG-ACA-CA-CAACA--A-CAG-AGCA-CA-CA-CAATA...........................................................AAAACAACAACTGACTGGTCA 6329
TAL.CM.01.8023              -G--GA-GGC---T--A--CA-C-CA--AA--CTC-C-GG--CAGTAGCACC-CCA-C-ACAAATTCAACAACAAACTCAGTAACATCC..................................AAGAGAACCAATGACACCAATGTGTGG 5988
TAL.CM.00.266               -A--GA-G-----TGAG--CA-A-CA--GA-CCT--CA-T--G-GCC-A--C-ACA-C-CCAGCCCCGACG..........................................................GCCATGAGCAATAGCAATATGCCGTGG 6450
SUN.GA.98.L14               -C-AAA-G-----TC-ACCGAAACC--CAAC-C-CTCCT----CTCA-CAGT-AAA--TAGCTGTGATTATTGGACGACCACCACGGCTAAGACTACAACTCAGACTACAAGTAGTACCAGCTCCACTGCTAGTACAACTACTCCAATGCCTCTAGATTGGAATTGTACT 6913
WRC.CI.97.97CI14            -A-AAA-GC-A--T---CCAGGAG-AGGGAG-GAG-CATTT.............-ATCC-TGTGAATCTAATACTACAAGAGCT...................ACTACAAGTGCTACTACAACAGTTGGTACAGTTAGTACTACTCCAGTTATGCAGGTAGAAAACAATTGC 6748
WRC.CI.98.98CI04            -G-AAA-G-CA--T--ACCAGGTGCACCA---T----C-AT-CG-AGGACTTTAACCCCTGTGAAATTAGCACAGGTGCACCTACTATTAGTACCATAAGTGCTGCACCTACTACTAGTACCATAAGTGCTGGTACCAGGACACCAGCTATGCCAGTAGATAATAATTGT 6793
WRC.GM.05.Pbt_05GM_X02      -G-AAA-GTCT--T--ACTAGG-G-TG-GA-A-GC---TAT-CAGAAGAC-TA-AACTGACAACAGTAAAA...............................ACAACAACACCAGTAACTCCGCCACAGAATACTACCCCTAGTATGGATGTAGAGAATAACTCT 5680
LST.CD.88.524               --C-CA-G--A--T--A-GGGAAGGACTA-G--AG--G-GA-CA-CA-CA-CAACA-C-CCAACAACAACACCC...ATGCCTTGCTTTATAACAGAATCTACTACTACTACTACAAAGAGTTCAGGGACTGAACAACCATCTTTAATAGAAGATACAAATTGTACACGA 5785
LST.CD.88.447               -G--GA-G--A--TGAA-GGGAAGGA-TA-GA--A-CA-AAGCA-CA-AA-CAAAA-CG...TCAACAACACCC...ATGCCTTGCTTTATAACAGAAAAAACCACAACCACCACCAAAAGATCTACAACTGGGGGACCAGAATTAATAGAGGACATAAATTGTATTAAA 5782
LST.KE.x.lho7               -G--AA-GC----TGGG--AG-T...............--A--A-CACCA-CAACA-C-CCAAAAACTACTACACAGATGCCTTGTTTTATCAACGAACAGGTA.........ACAGTTAAGAACCCGGGGAATGAAACAAGATTAGAGGAAGATCTTAATTGCACAAGA 6842
SYK.KE.x.SYK173             ----GA-GTT---T--ATTGGAC--C-GCCC-G---CA-GT.............................................ACCCCTACTACAAGTCCCCCTACTACACCTCCAAATGAAACGTGG 6641
SYK.KE.x.KE51               -C--AACGTTA--TGTA--GGAC---G-CAG--G--G-GGT---CT--CC-CAACA-TC...........................AGTCCAACAACAACAACAAACAAAGCAGCAACACCACCTGCATGG 6337
COL.CM.x.CGU1               C-CCAA-G-TT-----CA-AG-A--T-GGAC-GAA---GG--C-----CAGTA-CCCC-ACAACA..............................................ACAAATAATAGTGCTTCAGATTGGGAT 6054
```

```
                                                   V1 loop end   V2 loop start
H1B.FR.83.HXB2             ...ATG...ATAATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGC...ACAAGCATAAGAGGTAAGGTGCAGAAAGAATATGCATTTTTTTATAAACTTGATATAATACCAATAGATAATGATACT.................................AC 6790
Env                         .  M  .  I  M  E  K  G  E  I  K  N  C  S  F  N  I  S  .  T  S  I  R  G  K  V  Q  K  E  Y  A  F  F  Y  K  L  D  I  I  P  I  D  N  D  T  .  .  .  .  .  .  .  .  .  .  .  T
H1A1.UG.85.U455            ATC-CA...GATGGA-T--GG-A---A--G-----T--------------G-C-...---GAAC------A----AAA-------T----T--C---------G----------G---A---CA----AAC-GA-AAC..............................-A 6235
H1B.US.90.WEAU160          ...GAA...-AG------G-G-----A--G------------------G---C-...---CT-------AA----AGAA---CT---------C-----------------G----G--------C------A-.................................-- 6800
H1C.ET.86.ETH2220          ...--C...-AT-GT-CC--T-AT--A--G-----T--T-----------A-C-...---GAAC------A----AAAAG---G-C-------C--------------------G-----C-TA----C-G-T--..........................-- 6193
H1D.CD.84.84ZR085          CTGCCA...-CGG--A--CC------A--G--------------------A-C-...---GTAG---CG-A----AG-A--C---T-C-----C---------G-------G--G---A------------GGG-AAAAATGAAATCAAT.........GATACTTATGG 6332
H1F1.BE.93.VI850           .........................--GC-------T-----------G-C-...---GAAG------A----AA--T----TT--C---GC---------GG---------G---------GC---A-C-A-AGC.............................-G 6112
H1G.SE.93.SE6165           AGC-CT...G-GGATA-CCC------A-----------------G-A-C-...---GAA-------A----AAAA----------C--G-----C----G-------G--G-------TA------CA.......................................-- 6214
H1H.CF.90.056              ...........--GGA--CAGG---G--AC---C---T-----------G-A-CT...---GTAC------A----CA--------T-C-----C-C------G-------G--G-------T-----CA---G-..........................-- 6121
H1J.SE.93.SE7887           .........G---GTAGTCCT-ATATC--G-C-----------------A-CT...---GAA------AA---AAG-A--C----G--C--GC----C----G--AA---G--G-------T----G-A-C-A-..........................-A 6103
H1K.CM.96.MP535            .........G-GG-AAGCCC--C---A--------T-------------A-C-...---GAA---A--A----AA-A-A--G----C----C----C----G-------G--C-G---C-TA--GGC--GGG-AATAACAGTAGT.....................-- 5998
H101_AE.TH.90.CM240        ATA--A...GG--ATATA-C--AT--AG---G------TA----T-----G-C-...---GAAC---C--A----AA------G-TCC-----C-C--------G---------G---A---T--AG--A-G-AGACTAGT.......................-G 6370
H102_AG.NG.x.IBNG          ............GACAT---T--G--A---------------------A-C-...---GAAG------A----AAAA----GATGC-----C---------G-------G--G---A---TA--G-AA---A-GGT............................-G 6314
H1N.CM.95.YBF30            GAACCA...GAC--A-GAT-CAA-C-A--G-----T-----A------GCA-C-...--TGAGC---C--A---AAA-A--C---TT--CT-TC-G------GT-GAA---G--G-------CA--GCC............................TATAAT-A 6383
H1O.BE.87.ANT70            ......................AACCTT--G--G--G--TGAG--T---G-A-C-...--TGTT--C-A--AC--AAA-G-----A--C-G--TC-A--C---GT-TCA---T-G--GGA-C-GA--G-GACA-GCAGCACAAAT...........AAGACAAACAGC-A 6826
H1O.CM.91.MVP5180          CTAT-A...-ATGA-ACA-T-AAT-----G-G---T--TAG---T---G-A-CT...---GT-C-C-C--AC--AAA-G--C--A--C-G--TC-A--C---GT-TCA---C-G-GTAAGG-TA--G-CTCA-A-GCAGTAAAT.................GGA-- 6825
CPZ.CD.90.ANT              ACACCA...---G-A--TGGCATGA--C--C--G-A--TAAC--T---CAG---...---G-AT-T-A--A----AAA--A---ATGA-A--CA-A---------GGA---C-T--GAAGTGTC-GG-CA-C-A-GAGACTAAC........................ 6244
CPZ.CM.01.SIVcpzCAM13      AAG-AT...GA--CT-GAGTTTT-C-A--G--------TA-A--T--C-CA-CA...---GAAT-G----A----AAA-A-AC-G-TC---T----A------GT-GA----C-GCA-AGCC---GA-G-........................GGAACTGGTGA 6401
CPZ.CM.05.SIVcpzEK505      GTT-AC...CA--CAA-CC-C-T-C-A--GC----T-----A--T---G-G-C-...--TGAGT------A----AAA-A--C----TC--CT-TC-G-------TGGGA-----------CT----C-C-A---A-AGCAGTGGT.................AATAAC-G 6339
CPZ.CM.05.SIVcpzMB66       AAC--T...-CGG-T-GC-C-AT---C--GT-C-----T---------GCA-C-...---GAAT------AC-GAAA-A-AC-G-TG---T----G-------GG-AA---C-TGAG---C----AGGAA---AGCCACCAGAA............GGAGATAAGAACG- 6435
CPZ.CM.98.CAM3             CCT-AT...GA--AAAYAG-CA-CACAG----G--T--TA-C--T-----G-CT...--TGAGG------A----A-A-A-A----T----T-CC--------GT-GA----C--G--TTG--T-----C-----CGAT.............................. 6231
CPZ.GA.88.GAB2             GAA-A-...GATGG-TT-GC-CA---C--G-G---T--TA---------CA-C-...---GAAT------A----AAAA---C--ATT---T-C--A--C-GG---AA----C-TG--GGT-CTA-C---............................... 6196
CPZ.CM.98.CAM5             GAGT-A...-A--ATATA----A-ATA--G-----T--TA-A--T-----G-CT...--TGAAG------A----AAA-A--G---T----T-TC-A------GT-GA-------G----T--T-GAGG-A---AGAGTAATGAT........................ 6501
CPZ.CM.05.SIVcpzMT145      TCT-CT...GACGATCGCTTG-----C--G-G---T--TAGC--T---G---CA...--TGAAC------A----AAAAGAC---TG---T-C--A--C---GTGGAG-------C-G------GA---.............................AAT-G 6320
CPZ.TZ.00.TAN1             TTG-GT...TCT--ACCT--CTAT---G-GT-T--T--T--A--T---CAG-CA...--TGAGT-T----A----AAAA-AC--AT----T-C--G-------G-GAA-----TG--AA-GAG---GG-A-C..............................AAT-A 6434
CPZ.CM.05.SIVcpzLB7        ...--T...-AT-GCTCT--T-AGACT--G-----T--TAGC--------CA-CA...---GAAC------A----AA-A-AC-GACT--CT----G------GT-GA----T----CT-AG---A-C--G..........................ACAACCGA 6337
CPZ.US.85.CPZUS            AGTGCT...G---ATACT-CT-ATATA--G-G------T-----T-----A-CT...--TGAAT------AC--AAA-A-AC---TG---T-C--A------GT-GA----C--GCT-AT--CA-----A--.............................. 6835
CPZ.GA.88.GAB1             CAG-CA...TCTTCTCC-CCTCTC--A--G--------T--A--T---G-G-CA...---GAAT------A----AA-A-AC-G-TC---T--C-A------GTGGAG---G-GG--AACC---GG-----G-ACAAC........................... 6852
MON.CM.99.L1               GGAGG-...-ATGGAACAGG-CAGCCAG--T-T--T-----A--T--CCAG-CT...---GAAT-C--G-A---AAA-A--C--ATG---AGC--A----GG--GGAA--C------GAAGGAG--AGGAAGC-A-GGC.........................-G 6735
MON.NG.x.NG1               GGAGGC...-ATGG-AC-GG-CAGCCTC-TT-C--T--T--A--T--CCAA-C-...---GAAT-T----AC--AAR-A-AC--ATG---AGC--G---KGG--GGAG--C--T--G-A-GACA-G..........................GATAACACAACTGATGA 5351
GSN.CM.99.CN166            GGGGAA...-ATGGAACAGG-CA-CCCT--T-T-----T--C--T---CAG-C-...---GAAT-T----ACC-AAA-A-AC-GATG---AGCC-C--C-GG-G-GAA--C--T--GGA-GAGACCC---G-.........................AACCAG-G 6724
GSN.CM.99.CN71             GGAGGT...-ATGGAACAGG-CAGCCA---T-T-----T--C--T--CCAG-CT...---GAAT-T--G-AC--AAA-AGAC-GATG---AGCC-C----GG---GA------T--GAGGGC-C-GG---G-.........................AATGAC-- 6715
MUS_1.CM.01.1085           GGC-AT...GAT-GT-CTGG-CA-CCT--GTTC--T-----C------CAG-C-...---GAGT-C--G-A---AAA-AGAC-GATG---T----G--C-T-GTTGA----C-G--GAG-GACAGCTCA--G-G-GAGAATGAT........................ 6689
MUS_1.CM.01.CM1239         GACGCC...-ATGGTACAG-GCA-AGC--GTTC--T--T--A------CAG-CT...---GAAT-C----AC--AAA-A--C-GATG--CT-C--A-----C-CTGA----C-G--G-GTGAG--------................................ 6661
MUS_2.CM.01.CM1246         GAC-AT...-ATGA-ACAC--TTGACC--GT-C--T--T--C------CAG-C-...--CGAGT-C----A-GTAAAAA-AC-GATG--CT-TC-C--C-T-GT-AAA--CC-C--GAA-GGG--G-----G................................. 6796
MUS_2.CM.01.CM2500         GATGGC...-ATGATACA-TGCA--TC--GT-C--T--T--C------CAG-C-...---GAGT-C--G-AC--AAA-A--C--ATG---T-TC-G----T-GTGGAC---T----GAGGGAT-GG-GC---............................GA 6724
RCM.NG.x.NG411             .........-GTG-CCCTG-C-C---TG-C-TG--T--TGAA--T-CGG-GGCA...GGGTTA-A-----A-G-AAAAAT-----T-A---ACACA-GG--CTC-AGA---C-CTGGTGT-AG-CA........................GGTAATAATAATACT-G 6551
RCM.GA.x.GAB1              .........GA-G-TC--G-T-C---TG-G-TGTCT--TGAA--TGC-G-AGCA...GG-CTA-AG----A-G--AAA--C--GT-TA---ATACC-GG---TCTAGA--CC-TTGGTGTGA-A-G...............GAAACAAATTCTACAAATTCTACAAAA-A 6507
DEN.CD.x.CD1               TATGGA...GG--ATA---CTACCATCT-CCTT--G--TAA---------G-C-...---G-A-A-G-CTT--G--ATA---G--T-ATGTATGAAGC-C--TGTAGAATAGA-GATGT---G--A......................GATAGGGAAACTAACCAA-- 6841
                   V2 loop end                                                                                                V3 loop start
MAC.US.x.239               TGC-CA...GGCT----AC---AGC-A--G-T--G---TAAA-----C--G-CA...GGGTTA-A-----AC---AAAA-AG-GT-CA---A-AC--GG--CTCTGCA---T-GG--TGTGA-C-A.....................GGGAATAACACTGGTAATGAA-G 7214
Env                         C  T  .  G  L  E  Q  E  Q  M  I  S  C  K  F  N  M  T  .  G  L  K  R  D  K  K  K  E  Y  N  E  T  W  Y  S  A  D  L  V  C  E  Q  .  .  .  .  .  .  .  .  G  N  N  T  G  N  E  S
H2A.GW.x.ALI               TGCCCCAGGGG-C-A-G-G-T-A------GGTC--T--TCGG-----C--G-CA...GG-TTACAG----A----CCAA-AC-GT-TA---A-ACA-GG--CTC-AAA---G-GG-TTGTGA-CCA.....................TTTAACACCACCACAAACCAG-- 7277
H2A.DE.x.BEN               TGCGCA...GG-T-A-G-T-T-AG-----G-TGC-A--TGAG---------G-AG...GGGTTAGA-CA--A----AAAAG--GGT-TA-G-ACACA-GG---TT-GAA---G-GG-TTGTGACA-C..........................ACAACAGCTGG 7296
H2A.SN.x.ST                TGC-CA...GGGT---GAG---A------GGTCG----TCAG---------G-CA...GG-TTAGAG--G-A----AAAA-ACT-T-TA---A-ACA-GG--CTC-AAA---G--G-CTGTGA-TCA.....................AATGACACCAAGAAAGAG-A 6723
H2B.GH.86.D205             TGC-CA...GG-C-A-GAG---AG--A----TGC-A--TAA---T-G---G-CG...GG-CTA-G-----A-G--C-AA-AC--T-TA-A-ACACC-GG--CTC-GAA---T--GAGTGT-ATA--.......................ACCAGGAAGTATACC-G 7263
H2B.CI.x.EHO               TGTGCA...GGG--A-GACT--A---A--G-T-G----TCAA-----A--G-CA...GG-TTA-A-----A-G--TCAA-AC--T-TA-G-ACACC-GG-------AA---T--G-GTGTGA-A-GGGGACA-GGAGTAAT......................GAA-G 7258
H2G.CI.92.ABT96            TGT-CA...GGT--AC--TC--A-TCC--GGTTGG----AAA--T--C--G-CA...GGGCTA-A---G-A-C--AGAA--G--T-CA---AGAC--GG---TC--AA---T-GCA-TGYGA-C-G.....................TCCAATAAGAGTGAGAATGAG-G 6627
H2U.FR.96.12034            TGCTCA...GG-T-A--AC---A-CCAT-GGTG-GT--TAGA--T-----G-CA...GGCTTA-A---G-A----AAA--AG-GT-TA-A-A-ACA-GG---TCT-A---C---G-TTGTGA-C-G.....................AATACTAGTGAAAGTGAA-G 6756
MAC.US.x.EMBL_3            TGC-CA...GGCT----AC---AGC-A--G-T--G---TAAA----CC--G-CA...GGGTTA-A-----AC---ACAA--G-GT-CA---A-AC--GG--CTCTACA---T-GG-TTGTGA-C-A.....................GGGAATAGCACTGATAATGAA-G 6345
SMM.US.x.H9                TGT-CA...GGCT----AC-G-AGCCC--GGT--GT--TAAA--T--C--G-CA...GGGTTA-A-----AC--AAA-AGAG--T-TA---A-ACA-GG---TC-AGA---T--G-GTGTGA-C-A.....................AATAGCAATGGAAATGAAACTGAC-G 6691
SMM.US.x.PGM53             TGT-CA...GGCT-A---C-G-A-CCA--GGT--G---TAGA--T--C--G-CA...GGGTTA-A--AG-AC--AAA-A-AG--T-TA---A-ACA-GG---TC-AGA------G--TGTGA-C-A.....................AGTACTACTAGTAGAAATGAA-G 7148
SMM.US.x.SIVsmH635F_L3     TGTGCA...GGCT----AC-G-AGCCC--G-T-GGT--TAAA--T--C--G-CA...GGGTTA-AG--G-AC--AAA-ATAG--T-TA---A-ACA-GG---TC-AGA---T----CTGTGAGC-G.....................CCAGCAAATGGAAGTGAG-G 7221
SMM.SL.92.SL92B            TGCTCT...GGC--A---C-G-A-CCA--G-T-GGT--TCAA--T--C--G-CA...GG-CTA-A-----ACC--AAAAG-C-GT-CA---A-ACC-GG--CTCCAGA---C--G-CTGTGAGC-A.....................GGAGGAAATGAAAGC-G 6650
STM.US.x.STM               TGC-CA...GGCT----AG---A-TCAC--GT-GGT--TAAA--------G-CA...GG-CTA-A-----A---AAA-AGAG-GT-CA---A-ACG-GG--CTCTAGC---C----CTGTGA-C-A.....................AATGTCACTGGCGAAGAG-G 6853
MNE.US.x.MNE027            TGC-CA...GGCT----AC---AGCCA--G-T--G---TAAA-----C--G-CA...GGGTTA-A-----AC---AAAAGAG-GT-CA---A-AC--GG--CTCTGCA---T-GG-TTGTGA-C-A.....................GGTAATAGCACTGAAGATGAA-G 6685
DRL.x.x.FAO                GAACCA...G-------A-C-AACAGAG--TGT--A-A-AA-G-G-CAGGATTA...TGC--GGACT-TAAGG-A-A-ATC-C----A-CTTTAGG-A-G--G-TG-A-T-TGT-A-TGTGGCA--G-AAC-GAATCACCAGAAAGCAATAACTGTACTAATAAAAAA-G 6553
MND_2.x.x.5440             AAT-AT...GAG-CAACA-C-CAGC-A-AT-G-GTA--TAAA-----C-CG-CA...GG-TTATGT--G-ACTGCAAATT-G--AT-G-G-A-AAC---AGAT-TGAA---G---CTTGC-C-A-GCT-A---AAACTGGTAGTGCTACTAATAGTACTGAACCTGAATA 6510
MND_2.CM.98.CM16           AAC-CT...GA--CTACA-C-CA-C---AC-GGGT---TAAA-------CT-CA...GG-TTATGT--G-A-TGCAAATT-G-GAT-G-A-A-AA----AGAT-TGAA---G---CTTGT.........................ACAAATGGCACAAATAACACATA 6883
MND_2.GA.x.M14             AAT-AT...GAG-GTACC-C-CA---A-AT-GGGTA--TAAA--T--C-CA-CA...GG-TTATGT--G-A-TGTAAAATAG-GATTA-A-AGAGC---AGAT-TGAG---G-G-CCTGT...................ACTAAAAAGGAAAATGACACTGAAACTGA 6809
OLC.CI.97.97CI12           ACAG-A...-AGCCA--A-C--A-A--C-T---C-A--TAGG-AT-----T-CA...GGGTTATGT---ACTGTAAAATAG--ATGA--CA-ACA---G--TTGGAA---G--GC-TGT-C-A----C..................................... 6373
GRV.ET.x.GRI_677           GAAGGA...GA-----CAG-G-A-CCAGC-TCC------A----TGCA--TGCA...GG-TATCA-----A-GTAAAAA----TT-TAGCATGACC-GG---G-T-AG--GT--G-CTGC-ATA----AACAGGAAGTGAAAAGGGAAGTAAG................. 6774
VER.KE.x.TYO1_patent       GAA-G-...GAG------G-T-AGCCCGCCTCT--T--TA-A---GCA--GGCT...GGCTATG------A-C--AA-A-A--TT-T-C--TGG-G-GGA--G-TGCA--A--CTATTGC-A-A----GAC--A-AGCACTAGCAAA....................... 6882
VER.KE.x.9063              GAAGA-...GA-C-----G---A--CAGCTTC---T--TA-A--TGC---GGCA...GGGTA---T----A-C--AA-A-A--TT-T-CA-T-G-A-GGA--G-TGCA--G--CTATTGT-AGA-C-GCTC--GCTCTAACTCTAGCACAAAA................. 6903
VER.KE.x.AGM155            GAA-A-...GAG---A-TG-C-AG-CAGCGTC-------A-C--TGC---GGCT...GGGTA---T--G-ACC-AAA-A----TT-C-CA-T-G-A-GGA--G-TGCA--A--CT-TTGT-AGCG--G-ACAT-GCATAATGGGACAAAA.................... 6883
VER.DE.x.AGM3              GAA-A-...GA-T-AA-TG---AGCCTGCTTCT--T--TA-A--TGCA--GGCA...GGGTATG------A-C--AAAA----GT-T-CA-TGG-G-GGA--G-TGCA--A--C--GTGT.........................AAGAAGGGTAACAATTCTAAC-G 6400
TAN.UG.x.TAN1              ACA-AT...-GC-----TT-T-AGA-T-GTTCT-TA-----C--TGC---GGCA...GG-TAT-GG----A-GTAAAAA-----T-TA--AGTAC--GG---G-T-AG--GT-GG--TGTGAGA-GG-AA-C-AC...ACCACAGGCACGAGA................. 6866
SAB.SN.x.SAB1C             GAACAA...GAG-----A----A-C--GCC-TG-----T--C--TGC---GGCA...GG-TAT-G---G-A-GTAAA-A-A--TT-C-CAA--G-A-GGG--G-T-AA--GG-GG-TTGTGA---GGGAAGGGAAAAATCAAACGCAACA............CATACTGT 7076
MND_1.GA.x.MNDGB1          AAT-CA...GA-GAAA--GT--AGAG--ATC-TGTA--TAGG-AT--C--A-CA...GG-CTATGC--G-A-TC-AA-G-AG--AT-GTAA--AA----AGAGGGGA----G-G-A-TGTGA-A--......................................-A 6307
DEB.CM.99.CM40             GGGGAA...-AC--AACC-TGAC-C-ATATTGG--T--T--A--T---G--TCA...GGTCCTTAT----A----AAAG-A--GAGC-CA---G-A-GGCTAG-GGA------CCAGTGGGCG------A-AGA-GGTAGTGGGAACAGA....................... 6531
DEB.CM.99.CM5              GGAGA-...-AT-A-ACAGTGAC-C--TATTGG--T--TA-C------G-T-CA...GGGCCATAT--G-AC---AA-G-A-TGC--AG----G-G-GG-TAG-GGA---C---CAGTGGGAGACAG-AA-CCAAACAAATACTAACAAA............AGTATA-G 6481
TAL.CM.01.8023             TGGGGA...-AC-AT-CC-C--A-CCAG-GT-C-----TA-A-GG-----G-CA...GG-G-AT-T--G-A----AA-T-C-CCT-T-T-T-TCA--GG---CT-GCA--CT----GAAGGA-CCAGGCA--GGCAAT..........................GACT- 6125
TAL.CM.00.266              TGGGGA...GAC-ATACA-C--A-CCAG--T-T--T--TA-A-GG--C--G-CA...GG-GCAT-T--G-A----AG-T-T-CCT-C-T-T-TCAC-GG--CCT-GCA--CC----GAAGGA-CCAGGCA--GGATCC.............................. 6582
SUN.GA.98.L14              GAC-CA...GAG-ACATAGCG-AGAGC-AT---GTA--TAAA-AT---G-G-CA...GGGCTTTGT----A-TGTAAAAC-G-G-TGG-ACA-AAC---AGGG-TACA--AG-T-CTTGC-AT-G------C.................................. 7039
WRC.CI.97.97CI14           TCACA-...G-T-ATACTTC--AGATT-ATC-TGTG--TAGA--T--C-C--CA...GG--TGTGT-AG-ACTGTAGAATAG--AT-A-A-A--C----AGAGG-GA----G-GG-TTGT-ATA-----AG-................................ 6874
WRC.CI.98.98CI04           ACA-CA...G-G-ATACCTCT-AGATA-ACC-TGTA--TAGA--T----C--CA...GG--TGTGC-A--A-TGCAG-ATAG--ATTA-A-A--C----AGAGG-GA----G-GG-TTGC-ACA-A----GG---.............................. 6922
WRC.GM.05.Pbt_05GM_X02     ACTTCC...T-G-ATCTC-C--ACACA-ATC-TGT---TAGA--T---G-A-C-...GGGTTATGC----A-TGTAAA--AG--AT-GTA-AG-CC---AGAGG-GA----G---G-TGC-C-C-AC--.........................GGTAATGGTACACA 5817
LST.CD.88.524              TGGT-A...-ATGA-ACT-C--A-ATT-AT-GGG-A--T--A-AT---G-G-CA...GGGCTTTGT----A-TGCAA-G-AG-GAT-A-GCA-AA----AGGT-TGA---G-C--C-TGT-AT--G----G-.............................. 5911
LST.CD.88.447              TGGG-A...-ATGA-ACT-CT-AGATT-AT-G-G-G--TAGA-AT---G-T-CA...GG-CTATGT--G-ACTGTAG-ACAG--AT-A-GCAGAG----AGGT-TGA---G-C--C-TGTGATA-G----GA.............................. 5908
LST.KE.x.lho7              GGGC-T...-ATGA-ACC-C--AGAG--ATGC-G-A---CAG-AT---G-A-CA...GGGTTATGC----ACTGCAG-AC-G--AT-A-ACA-AGC---AGAT-TGA----G---CGTGCTC--GAG-GAGG............................GAGAAC-G 6976
SYK.KE.x.SYK173            TGGGGA...GAC-ACAGT-C--AGCCA-G-TTC--T--------T--CT-G-CA...GGGG-GT-C-AG-A---AAAA---C--T-TAGA--T--C--C-----GGA---CT-G--GAA-GAG--A..........................GGCAATAGT-G 6772
SYK.KE.x.KE51              TGGGG-...GAC-ACTCT-C--A-ACT-GGT-C--T--T--A--T--C--G-CA...GGGG-TT-T-AG-A----AAA--TC-GT-TA-------C-----C--GTCA--------GAG-GAG--A..........................GGAAATGAA-G 6468
COL.CM.x.CGU1              GAGTCA...-ATTG-A-AG--TATCC-TGGT-T--T---AGAA-G---TCA-CAGCTTTCTTAC-T-A--A--GAAA-G--TT-----TA-G-----CAGTAG--GA-TTA-CTG-G-TTGG-A----AA--GACAGTAATAGT.......................-T 6197
```

```
                             V2 loop end
H1B.FR.83.HXB2          CAGCTATAAGTTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACGTTCAATGGAACAGGACCA......TGTACAAATGTCAGCA 6954
Env                      S  Y  K  L  T  S  C  N  T  S  V  I  T  Q  A  C  P  K  V  S  F  E  P  I  P  I  H  Y  C  A  P  A  G  F  A  I  L  K  C  N  N  K  T  F  N  G  T  G  P  .  .  C  T  N  V  S
H1A1.UG.85.U455         T--T----GAC-A-T--A------T------AC------------T-----C-----G------------------------------------A------------------G-----GG--CCTGA-----------A---G---.......--C-GG---------- 6399
H1B.US.90.WEAU160       A-------C-----T--A------AT-----AC--------------------------------A--------------------------------------------G------G------A-----------AG-----C.........--A---------- 6964
H1C.ET.86.ETH2220       TGAT----GA--A-T--A------T------AC---A-----A-------------C--T--G--C--------T-------------T--A-------A-----------G----GAG-------A----C---------------.......--CCAT--------- 6357
H1D.CD.84.84ZR085       T-CT----G---A-T--A------T-------C---------A--T--------A-----------------------------------A-----------A--------G-----CG----A-G---------------C---.......--C-A---------T 6496
H1F1.BE.93.VI850        TGAA----G-C-A-T--A------T------AC------------T-------------T-GG--T-----C--T-------------T--A-------A---A---------------G----A-GA--------G-----G---.......--C-AG--------- 6276
H1G.SE.93.SE6165        T-AT----G-C-A---CA------TGT----AC---C-A------T-----------A-T-----C--------------------T--AA--------------T----G----GGG-----GA-----------------------.......-----A--------T- 6378
H1H.CF.90.056           TCAG----G-C-A-T--A------T----------------------T--C--------G--------A--T---------------------T--A-----C---------------G-----C-----A--A---------------TT-.......---------------T- 6285
H1J.SE.93.SE7887        A-AT----TTC-A-T--AC-----T-------------A---A--T-----------G--T---C-A-------------------------------A------------------------G------AT---------------T--.......--C-A--------T- 6267
H1K.CM.96.MP535         TGAA--C-G-C-A-T--A------T------AC-----------A-T--------A---A--------------------------------------A-----C---------------G-----GG----A-G---------G----------.......-----A---------- 6162
H101_AE.TH.90.CM240     TGAG----G---A-T--A------T--T----------AG-----T---------A----------T--------T-----------A-T--A-------A-------T----G------G------AT--------G-----G---.......-----A---------T 6534
H102_AG.NG.x.IBNG       TCAG----G---A-T--A------T-------C------------T------------------------------------------------A------------------G-----GG---AGG--------------G---.......--C-A---------- 6478
H1N.CM.95.YBF30         A-CA----G-C-A-T--A------T---A---CTG-G-----A--T-----T---ACT----------------A--------C-----A--AC-A--C-----C---A-G----------G-AGGA-AC--T-G-----AT---AGC.......-----------G--T- 6547
H1O.BE.87.ANT70         A-TG----CA--A--T-A-------T--A--AC---C--G--A----------------T-----A--------A-----C---------T--A--A--A-A---T--CT-T--G-----C-GC-CAGAA--T-----------CA--.......--C-G---CA-A-CGG 6990
H1O.CM.91.MVP5180       A-CA----T---A--T-A-------T--A--A-T--C-AG----------G------AGT--------C-----------C---------T--AA-A--A-A---C--CT-T--G-----G-C-CAGAC--T-----------C-T-.......--CCAC---A-TTCAG 6989
CPZ.CD.90.ANT           .T-----T-C--ATGGCAC--C------A--ACT--C-----AT-----GA---TCTA-T-----A--------T---------------T--A--A--A-A---T--AT-G-G----G-AG--G--GAT--T-CA---GT---GATG.......-----A-------TCAG 6407
CPZ.CM.01.SIVcpzCAM13   T-CA----CAA---TT-A----------A---C---A--------A--------A--G-----------T-----A-----C---------A--A--A--A-----A-----C---------G--GTA-AC--TTCA--G-A---AA-.......-----G------A---- 6565
CPZ.CM.05.SIVcpzEK505   TCAA----G---A-T--A----------A---CTG-G---------------T--AA-T------------------A---T-----------A--AC-A--A-----C---A-----------G--C-AGAT--T--------T--GGA-.......-----AC-----A--T- 6503
CPZ.CM.05.SIVcpzMB66    -TTA----GA--ATAT-A----------A---C---G--T--A--T--CT----A--C--TC----A--------T---------------T--A-----A-----CC-CT-G--G-------G---A-AT--T-C---G-T-----AG.......-----A------T--T- 6599
CPZ.CM.98.CAM3          .-CA----GAC-A-T--A---C--T---A---CT-----------G-----T---ACC-----------------A--------C-----AA-CC-A----A---A--CA-G-----C-----TGCCA-AC--T-------------A--GGAAGA----AC---A-A---- 6400
CPZ.GA.88.GAB2          .-CG-T--G---A-TT-A---C--T--TA---CA--C-----A--T--------AACC--------A--------A--------C-----A--A-----C-----CC-GT-G-----------G----AGAT-ATCCA----A----AAG.......-----A------A--T- 6359
CPZ.CM.98.CAM5          .-C-----GA--A-T--A------T--TA--ACTG-A--------A-----T--AACT-----------------A--------C-----A--CC-A--------A--AA-G-----------G--C-A-A---T-------------------.......-------------A-A---- 6664
CPZ.CM.05.SIVcpzMT145   T-CT--C-G-----TT-A------T---A---C---C--T--A--T--------ACT---------A--------A--------C-----A--A--A--A-----AC-C------------G---TAGAT-A---A--G-AT-A-A-T.......-----A------A--T- 6484
CPZ.TZ.00.TAN1          T--T---T-T--ACAT-A---C--T---------------T--A-AA---GAT--ATCTA-T-----A-----------CAGA--C-----T--A--A--C-----CC-GT---------GAG--C---AT----CA--G-A-----A-.......--CT-C------TCAG 6598
CPZ.CM.05.SIVcpzLB7     G---------C-T-T--A------T--TA---C---A--C-----T--C------A-T--T-----------A--T---------------T--A-----------TC-----GG-----C-----A--A-AT-----T-GT--T---.......-----AC-----T--T- 6501
CPZ.US.85.CPZUS         .-C-----GAC---T--A---C------A-C-CT--C--------T-----T---ACC-----------------A-----C--------A--AC-A--C-----C--C-----------------G-A--AGAT-----A----AG---GAG.......-----A------T--T- 6998
CPZ.GA.88.GAB1          .-CA----G-A-A-TT-A---C--T--TA---C---A-----A--A--------AACG------------T-----A-----C--------A--A--A--G-----A-----T-----C---G---AGAC--TTCA---A---TAA-.......-------T-----A--T- 7015
MON.CM.99.L1            TCA----T-CA-CCTG-A---C-----A---TA---A--C--A--T---GAG---TCCAAT-A-------G----TT-G------------A--CC-A--C-A---TT-AT-G-G----G--G-CCCCG-C--T-CG---CA---GAGT.......--T-T-----TTCTG 6899
MON.NG.x.NG1            -TAT---T-TC--CT--A------T----Y-TA---Y--C--A-----GA----TCCAAT-A------T-----T--T-----C-------T--C-C--C-A---T--AT-G-G----G--G--CCAK-C----CA--GCAGAACAAT.......--CT-R------TCAG 5515
GSN.CM.99.CN166         TG-A---T-TA-T-G--A------T------TAT--C--T--AAAA---GTC---TCC-----CC-A--GG----------C--C-----T--AC-C--C-A---T--G--------C---G--GTT-AT----CA---GT---CA--.......--CTAC-----GTCAG 6888
GSN.CM.99.CN71          TGA----T-TA-T-T--A---C-----A---TATG-C--G--AAAG---GTC---ACC-----CC-A--GG----------C--C-----T--C-C--C-----T--G--------C---G--GCC-AT----CA---GTG--CAAG.......--T-C-----ATCAG 6879
MUS_1.CM.01.1085        .--T---T-TA-C----AC--C-----A---TA---C--C-----A---GAG---AGCAG---CC-A--GG----G-----A--C-----A--C--G--G-ACAGCT-G---------C---G--GCC-AT--T-C---GCA----GA-.......-----AT-----G-CTG 6852
MUS_1.CM.01.CM1239      .-CT---T-CA-ATT--AC-------------TA---A--T---CAA---GA----AGCAG---CC-A--GG----G-----A--C-----A--T--A--G-A-T-TC-AT-G---------G--GCT-AT--T-CA---CA----AA-.......-----AC-----G-CAG 6824
MUS_2.CM.01.CM1246      .-CA---T-CA-T----AC--C-----T--TTAT--C------CAG---GAG---AGCAG---CC-A--GG-C--A-----A--C--------T--A--A-A-AGTC-CT-------C---G--GCT-AT----CA---CA---CATT.......-----AT-----A-C-G 6959
MUS_2.CM.01.CM2500      -G----CT-TA-C----A---C-----A---TAT-----C--ACAG---GAG--AAGCAG---CC-A--GG-------------C--------A--A--A-A-AGCT-GT-----------G--GCT-AT----CA---CA---CAA-.......-----GG-----A-C-G 6888
RCM.NG.x.NG411          GCAA-G-TTTA-A-GGCA---------AA--AGT--ACA----TT----GA-CCAAA--A--GG--A---T--AGGT--AGG-------A---C-A----A---C--A--G-----CT-G--GCA-AT-AT-CA--CTTT-ATA-C.......----CA------T-CAG 6715
RCM.GA.x.GAB1           G-AA-G-TTTG-A-GGCAC--C--T--AA-TTC---ACA---ATTT---GA-CCAAAG-A--GG--A---T--AGGT--AGA--------T--AC-A--G-----CT-A--GGTC--C--AG----A-AT-AT-CA--CTTT-ATA-C.......----GTT-----T-CTG 6671
DEN.CD.x.CD1            ATAT---T-TA---GGTA---C--TCAA--CAC------C--A--T---GA------TGT------A---T-CAAA---AGGATG-----A--A-----C-----TC-C--T--G------G--C-TC-T-GG-G---GTAT--C-TT.......-----A----A-GTCAG 7005

MAC.US.x.239            T--A-G-T-CA---ACCAC--------T--T--T--CCA-G--T-T---GAC--ACAT-AT-GG--TG-T---AGAT-TAGG--------A--TC-A----A---TT-G--T-G-------G-C-CA-AT-ATTCA--CTTTATG--TAAA...---T-T--G--GGTGG 7381
Env                      R  C  Y  M  N  H  C  N  T  S  V  I  Q  E  S  C  D  K  H  Y  W  D  A  I  R  F  R  Y  C  A  P  P  G  Y  A  L  L  R  C  N  D  T  N  Y  S  G  F  M  P  K  .  C  S  K  V  V
H2A.GW.x.ALI            ---G-G-T-CA---ACCA---C-----A--------C---G--T-A---GAT---CAC-AT-GG--TG-T--AAGGT-TAGA--C-----A--AC------AC--CC-A---G---CG--G---TC-AT-ATTCA--CTTT-C---CAAT...-CT-T--A--AGTAG 7444
H2A.DE.x.BEN            --CA-G-T-CA---G-CA---C-----A---A----C-A-G--T-A---GAT---CAC-AT-GG--TG-T--GAGGT-TAGA--C-----A--AC-G--C-----CC-AT---G-------G---CC-AC-ATTCA--CTTT-A---TAAG...-C--T--A--AGTAG 7463
H2A.SN.x.ST             A-CA-G-T-CA---ACCAC--C-----A--------C---G--T-A---GAC---CAC-AT-GG--TA-T--GAGGT-TAGA--------A--AC-G--------CC-G----G---C---G---CC-AT-ATTCA--CTTT-AG--CAAT...--T-T--G--AGTAG 6890
H2B.GH.86.D205          ---A-GCT-TA-A-G--CC--C-----AA--A-T--CCA-G--T-A---GAC--ACAT-AT-GG--CAGCT-AAGGT-TAGG--------T--CC-G--G---TTTC-A----G-------G---CC-AC-ATTCA--CTTCATG--CAAC...-C-GT--G--AGTAG 7430
H2B.CI.x.EHO            --AA-G-T-TA-A-A--CC-----T--A---A-T--CCA-G-AT-A---GAC--ACAT-AT-GG--TAGTT-AAGAT-TAGG--------T--CC-A--A-----TT-G----G---C---G---CT-AA-ATTCA--CTTCATG--CAAC...---GT--G--AGTAG 7425
H2G.CI.92.ABT96         T--A-GCT-TA---GGCAC-----T--A----------CAGG-AT-A---GAT--ACAT-AT-GG--TG-T--AAGAT-TAGA--------T---C-A--G-----TT-G----G-------G---CT-AC-ATTCA---TTT-C---TAAT...-CT-T--A--AGTGG 6794
H2U.FR.96.12034         T-AA-G-T-CA---ACCAC--C--------T--T--CCA-G--T-T---GAC--ACAT-AT-GG--TG-C--AAGAT-TAGG--------A---C-A--C-A---TC-GT---G-------G---CT-AT-A-TCA---TTTATG--TAAT...-CT-T--A--AGTAG 6923
MAC.US.x.EMBL_3         ---A-GCT-CA---ATCAC--------T--T--T--CCA-G--T-T---GAC--ACAT-AT-GG--TA-T---AGAT-TAGG--------A--TC-A----A---TT-G--T-G-------G-C-CA-AT-ATTCA--CTTTATG--TAAA...---T-T--G--GGTGG 6512
SMM.US.x.H9             T-AA-GCT-TA---ACCA----------AGT--T---CA-G-AT-----GAC---CAT-AT-GG--TG-T---AGAT-YAGA--------A---C-A----A---TT-G--T-GG------G---CC-AT-ATTCA--CTTT-CT--TAAC...-----T--G--AGTAG 6858
SMM.US.x.PGM53          T-AA-G-T-CA---ATCA----------AGT--T---CA-G-AT-----GAC---CAT-AT-GG--TG-T---AGAT-TAGA--C-----A---C-A----A---TT-G--T-GG------G---CA-AT-ATTCA--CTTT-CT--TAAC...---T-T--G--AGTGG 7315
SMM.US.x.SIVsmH635F_L3  T-AA-G-T-CA--CAGCA----------AGT--T---CAGG-AT-----GAC---CAT-AT-GG--TG-T---AGAT-TAGA--C-----A---C-A----A---TT-G--T-GG------G--TCA-AT-ATTCA--CTTT-CT--TAAA...---T-T--G--AGTGG 7388
SMM.SL.92.SL92B         T--A-G-T-CA---ATCA---C--T--AAGT--T--ACA-G--T-A---GAC---CAC-AT-GG--TG-C--CAGATATAGG--C-----A--AC-A--C-A---TT-G--T-G-------G---CA-AT-ATTCA--CTTT-C---TAAT...----GT--G--AGTAG 6817
STM.US.x.STM            ---A-GCT-CA---G-CA---C--T---AGT--A---CA-G-AT-----GAC--ACAT-AT-GG--TG-CT-AAGGT-TAGG--------A--AC-------A---TT-G--C-G---C---G---CA-AT-AT-CA--CTTT-C---TAAT...-CT-T--G--AGTAG 7020
MNE.US.x.MNE027         T--A-G-T-CA---ATCAC--------T--T--T---CA-G-AT-T---GAC--ACAT-AT-GG--TG-T---AGAT-TAGG--------A--TC-A----A---TT-G--T-G-------G-C-CA-A--ATTCA--CTTTATG--TAAC...---T-T--G--GGTGG 6852
DRL.x.x.FAO             A-CA-G-T-TG----C-----C---T-AA----G-----C----AT---AAT----CC---ACA--TATG--GAAAT-CAGA-TG-----T--AC-A--A-A-ATAT-G----GG------G-A--ACTAAAT-TAA-T-GTAAG..........---A----A-A-CAG 6714
MND_2.x.x.5440          TGAG-G-T-CA----TTCA------G-AA----A--A-------A-C--AAC--A-CT---ACA--TAG---GA-AT-TAGG-TG-----A--CC-A--G----TCT-A-----------G-A--ATT-AA---GAC--A-TT-.........---GGG-----GTCTG 6671
MND_2.CM.98.CM16        TTCT-G-T-TA-----CAA--C--T--A-----A--A--C----A---AAT--A-C----ACA--TGA---AAAAT-TAGACTG-----A--TC-A--A-A--TTC-CT-G-GG--CCTAG-A-GATT-AATGTCTCT-A-AA-..........-----C---A---CAG 7044
MND_2.GA.x.M14          AGAA-G-T-TA-----CAC-----T--A-----A--A-------AT---AAT----CC--AACA--CAA---GA-AT-TAGA-TA--------TC-A--G-ACATTC-G----G----CG-G-G--ATT-AAT--AACT-A-TT-..........---G-C------TCAG 6970
OLC.CI.97.97CI12        .-CA-G-T-CA--T--CA------------T--G--A--T----AT----A-T-T-C-GT--CA--T-AGCAAATGG-GA-G-TG-----A--AC-A---AGG-T-C--A----G----------G---AGAA---CC-...............--C-G---CA-TTCAG 6527
GRV.ET.x.GRI_677        .GAT-G-T-CA---T-CA------TGAT-----G--A-A-G-A--T---GAT--AAC--AT-GG--TA-TT-AAGAG--AGA--C-----A--A--A--G-A---TT-G------------G-----GAT-AT-GA--CTTT-CT---AAG...-C-AG-----TTCAG 6940
VER.KE.x.TYO1_patent    .GAG-G-T-CA---TTCA------TGA------T--A-A-G-A--A---GAC--AAC--AT-GG--T-AGT-GAGGT--AGG--------T--A--A----A---TT-G------------G--G-AGAT-AT-----GTATAA--A-AAT...-CT-------ATCAG 7048
VER.KE.x.9063           .GAA-GCT-TA---TTCAC-----TGAT-----T--A-A-G-A------GA---AAC--AT-GG--TGAGC-AAGGC--AGA--C---------A--A--A-----TC-GT-------CC--G--C-TGAT-AT-----TTATAAG-A-AAT...---GT-----TTCAG 7069
VER.KE.x.AGM155         .GAG-GCT-TA---TCCAC-----TGAT-----T--A-AGG-A--T---GAT---AC--AT-GG--TGA-T-AAGAC--AGA--------T--A--A--A-AC--TT-G--T---------G--TG-GAT-ATGCA---TTTAAG---GAA...---T-T-----TTCAG 7049
VER.DE.x.AGM3           AGAA-G-T-TA---TTCA------TGAT-----T--A-A-G-A------GAT--AAC--AT-GG--TGAGT-AAGAT--AGG--C------------A--G-----TT-AT------C--CG--T-TGAT-ATGCA--GTTTAAGA--AAC...---T-T-----TTCAG 6567
TAN.UG.x.TAN1           .G-T-G-T-CA---TTCAC--C---GA---T--A--A-A-G-A--T---GAG--AACT-AT-GG--TA-CT-AAGAT--AGA--C-----A--A--A--C-----T--CT----------GG---CT-AT-AT-CA---TTT--TGTT.........---G------TTCAG 7029
SAB.SN.x.SAB1C          GG-A-G-T-CA---TCCA------T---------A--C-A-G----T---GAT--AACT-AT-GG--TA-CT--AGGC-TAGG--------T--A--A----A---TT-G----G----GCAG---CAGAT-ATTCA---CATAAGG--.........---GG-----A-CAG 7240
MND_1.GA.x.MNDGB1       T-CT-GCT-TA---ATCA------TGAG-----T-A----G-A-A-----AG----G-CTT--GATAAG-...TGT--TTTAGG-----TG--TC-A--A-A--TC--GT---G--A----G-G---TTAAAT---AAT-A-TTG..........---GC---A-ATCAG 6465
DEB.CM.99.CM40          .-CAGGGT-CA---A-CAC-----TGA---T-----C--C----T-T---GA--CTAGCCG----A-A--TT--AAG--CAGG--C-----A--A-----A-A--GTC-A----G----G--G----A-AT--T-----T-----GTT-........-C-AC--C--G-CTG 6694
DEB.CM.99.CM5           ACATGGGT-TA---ACCAC--C----AT-----T--A------T-----GA--CCAGTCG----A----TT--AAG---AGA--C-----A--A-----A-A--GTC-GT---G----G--G-C-GA-AC--T-----T--C--CTTG.........-----C-----G-CTG 6645
TAL.CM.01.8023          ATA----T-TGCA---G-A--C--TTTG-----A--AG-T------T---GA---AC-CAT-A-C-AG-GT----A--T--G--C-----T--A--A--C-----CC-GA----G------G-C--CC-A-GG-CA--T-G---CAAG.........-CAT--CA--TCAG 6289
TAL.CM.00.266           .TT----T-TGCC---G-A--C--TGTG-----T--AG----A--T---GA--AACTCA--ACC--G--T----T-----G--C-----T--A--A--C-----CC-GA----G------G---GCC-A-GG-CA--T-G----AAG.........-CAC--CA--TCAG 6745
SUN.GA.98.L14           .-CA-G-T-TA-----CA---C--TGAT---A-T-------A-A---C-AT----GTATTA-GC-AAATGC-TA-T-CAGGCT------A--T--G--G-A-AT-T-AT---G----G--G-A--ATTAAATGC-ACT-AGAA-..........---A----A-A-CTG 7199
WRC.CI.97.97CI14        .GAA-GCT-TA-A-ATCAC-----TGAG--------A-AG--A-A-----A---A-G--GGCAG--TGA-CAATAT---AGAATG-----T--AC-A--A-CA-TTT-GT-G---------G-G--ATT-AA-TCCTCT-AGCT-..........---GAG-----TCAG 7034
WRC.CI.98.98CI04        .GAA-GCT-TG-A-ATCA------TGAA--T--G--A-A---A-AT----AG--A-GT-GGCAA--TGA-CAATATG--AGAATG-----T--AC-A--A-C--TTT-G--T--------G-GCCCTTAAA-TCCTCT-A-TT-..........---GAG-----ATCAG 7082
WRC.GM.05.Pbt_05GM_X02  A-CT-G-T-CA-A-ATCA------TGAG--T--A--CCAT----AT----A-----C--GGCA---TGAGCAATA----AGAATG-----A--AC-A----A--TAT-AT---G-------G-A--ATTAAAT-G-A-T-AGCT-..........---GAG-----TTCTG 5978
LST.CD.88.524           .-CT-GCT-CA----CCA---C--TGATAGTA-T---CTG--A-A---CAAT----G-GTTA-GTCAAATGC-TATT-TAGGCT------A--A-----C-ACAT-C--T-------C---G-A---CTAAA-TT-TC-G-TAAT............---T---A-A-CAG 6071
LST.CD.88.447           .-CT-G-T-TA----CCAC-----TGATAGTA-T---CAG--A-AT--CAAT----G-GTAA-ATCTAATGC-TATT-TAGGCT------A--A--A--C-A-AT-C--T----G--C--CG-A--ATT-AATTTCTC-G-TAAT............---C---A-A-CAG 6068
LST.KE.x.lho7           A-C--G-T-CA----TCA------TGAT---A-A--A-----A-AT---AAT---GGGTAA-GC-AAATGC-TATT-TAGGCT-----T--A--A--C-A-AT-C-AT-G-G---C---G-AC-ATTAAA-TT-A-T-A-AA-..........---GA----A-A-CAG 7137
SYK.KE.x.SYK173         -TAT---T-CC--TT-CA------T--AAGT--T--AT-CGCA--A---GAG--ACAGA-----C-A--CT----G--T--A--C-----T--AC-G----ACT-AT-GT-G-----C---G---CA-AT--TG-A---GAC-ATGTC.................-A-CTG 6936
SYK.KE.x.KE51           TTAT--CT-T--ACAGCA------T--TAGT--A--GCGCA-T--T---GA---ACAGA-T---CGC---T-------T--A--C-----T--AC-G--A-ACT-AC-GT-G---------G---CAGAA--TG-A---GAT-CTGAG.........---CA------A-CAG 6632
COL.CM.x.CGU1           ---GGCA-CAA---A-GA----GCA-AT-ATACAG----C--A-TA---GAT-T-ACTATAG-G--T---G--AGG-C-GGG-TC-----AG-TC-A--A-ACAT-C--T---G----G--G----A-A--GGG----G------G--........-C-AC-----G-CAG 6361
```

```
H1B.FR.83.HXB2          CAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGA......................TCTGTCAATTTCACGGACAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTA 7100
Env                     T__V__Q__C__T__H__G__I__R__P__V__V__S__T__Q__L__L__L__N__G__S__L__A__E__E__E__V__V__I__R__.__.__.__.__.__.__.__.__S__V__N__F__T__D__N__A__K__T__I__I__V__Q__L__N__T__S__V__
H1A1.UG.85.U455         ---------------------C-A------G-----------------------------------AG---AA--AGG------......................----AA--------AA-------A-------------------TGT--ATC----- 6545
H1B.US.90.WEAU160       ----------------C--------A--------G-----C---T------------------------------CA-----------......................----AA----------------------A------------------TGT---CA-- 7110
H1C.ET.86.ETH2220       ----------C------------AA-----G-------------AT---------T---A---------GT---AC-A---------......................-T--AA---C-G--AA-------C----TA--------------T--TGA------- 6503
H1D.CD.84.84ZR085       -----------------------------G--G-----------------------------------------A-C-----A---......................----AA---C----AA------------T-----------T--T--TCA------- 6642
H1F1.BE.93.VI850        -----------------------AA-----G-----------T--T-----------C-----------GTA-------C---......................---CAA---A--T-AA-T-----A----------------C--T--TGA------- 6422
H1G.SE.93.SE6165        -----------------------A------G------------A--AC-G---------T---------G-A-AA--AA-G-----......................----AA---------A------A-C---GT--------------T--T-A-A----- 6524
H1H.CF.90.056           ------------------------A-----G----------------A--------A--C-----------C--A-CA--------......................A--AAA---A--T-A------A-C----A--------------T--G---C-A--- 6431
H1J.SE.93.SE7887        ----------------C-------A-----G-------------A---C-----------A---------G---CA--A--------......................----AA--CA--T-A--------C----A------------A--T--T-A-A----- 6413
H1K.CM.96.MP535         -----------------------AA-----G-----------------------------C---------------A--A-------G......................----AA---A-T--A--T---A-A----A-----------------T--TGA-A----- 6308
H101_AE.TH.90.CM240     ----------C------------A------G-----------T-----------------------------------A--A----C---......................----AAG--C----AA-------C------------G--C--T--T-A------- 6680
H102_AG.NG.x.IBNG       ----------C-----------C-A--------------------------------------------G---A----G--C---......................----AA---A----AA-------C-----------------T--GCT-ATC----- 6624
H1N.CM.95.YBF30         -T--------C-----------A-A------GA----C-----GT-AA-CC-------A--CT--AATAC---T-GAA-T--T-------.........................-AT--T-G-CAC-GT-ATC-GT-G--G--ATG---TGAGA-A--G 6681
H1O.BE.87.ANT70         T---TACT-----------C---------AC----AGT-----G--AA-A--------G-CA--CT-TA---G-A-AA--AG---G-TG......................G-AAAAG--A-TTT---AGG--GA----AT--C-----GACC--A---T-TA-CC-- 7136
H1O.CM.91.MVP5180       TG--TACT-----------C--C-A----AC----AGT--------AA-AC-G-----G-CA--CT-TAG----A--A--AG-----TG......................GGAAAA---A-T--A--ATCA--A--G-AT--C------ACC--A-----TC--A-- 7135
CPZ.CD.90.ANT           T------T--C--T--------A--C---A-G--GG----ATGGT-A--A--------A-C-TACCA-AC-A-CACTTC----G-A-TG......................AA--GTCGCAAA-AT--ATC--TGCTTGTA-G-T-T-G-A-AGAATT-GA-AACT-- 6553
CPZ.CM.01.SIVcpzCAM13   ----C--T--------------C-AA---------A----C---T--A-CA-T-----G--------TAC----A-CA--AC-G-A---......................GTAAA----GC-T-TA-A---A-CC-TGATTGG--------A--A-G----G-A--G 6711
CPZ.CM.05.SIVcpzEK505   -T-----------C--------A-A------GA----C--A--GT--A-CC----------CT-----AC-TC-A-TA-C-----A---.........................AAT----CA-AA---.........--AC-CT-G--G--AT-A--TGA-AG-A-C 6637
CPZ.CM.05.SIVcpzMB66    -T--G--T-----------C--C--A--T-----C-------GT-T--T--------A-CCT---A----A-G-T-AC---CC-AGAT.........................AGG---G--T-AA-TG--ATGG-C--T--------GA-AT-A---GA-A----- 6742
CPZ.CM.98.CAM3          ----C--------T-----C--C--A-----G---A-T--A---T--A-CC-T-----A---G----C---A-CA--ACTA-TGC----....................................AGG--C-GGT-C-R-T-CC-GA-N--RT-TC-G-ANN-G--C 6531
CPZ.GA.88.GAB2          ----G--T--C--------------A---ACT---A-T-----GT--A-AC-C--------CT-------------A-AG--GG-GTTC...............AGAACTAAGAA-A-G-CAGCTC-A-GTCTCAG-G-C--AG-T---------T-A--A-GGG-AA-T 6514
CPZ.CM.98.CAM5          ----------C--T-----C--A--A---------A-T--A---G--A-CC-T-----A--C-----T-----G-G-CCC-CTG--GCTAGA.........................-G-GAAGATCTA-GACAC----T---T---A----AT-T--AGA-GGGA-T 6807
CPZ.CM.05.SIVcpzMT145   -------T-----------C----AA-----G-C-A-C--C------A-T--G-----G---AC------CA-TC--ACT--GGCC---.............................A-AGACCCTTCA-AG--TTTAGC----A-------C--AGATC-A--- 6621
CPZ.TZ.00.TAN1          T---T--C-----------G---TAT--TA-GA--G-C--AGC-T-A-AC--------GTCC--G-A---------AAC-AA-GC-TAC.........................T-TGT-AAT--CTCAGT-AA--C-C--T--T-----A-AT-T--TGT---AA-- 6741
CPZ.CM.05.SIVcpzLB7     -------C-----T---------AA-----TA-T--T------T--A-CC-G-----T---------------A-AA-CA-G-----G...........................T-CAAGG-T--T.........-AGG-TC----------C--T---AG-A-- 6632
CPZ.US.85.CPZUS         ----G--G-----T-----C--A-AA-----G--GA-T--A--G--CA-AA-----------T-----ACTA--A-T--TACTG-A---.............................-G-AAA...A--TT---AG-C-TT--TC-------AT-CTCAGAGGGA--C 7135
CPZ.GA.88.GAB1          ----G--T---------------AA-----G--TA-C-----GT-A--TA-T-----A---T----T---G-A-CA--ACTG-A---.........................GTG-AG---AAGT-AA-A---A--G-TGT-TGG----------AGTAGA-G-A--T 7161
MON.CM.99.L1            -----AC------T----CT--CCAA--TA-----G-T---TGGT-C-A---G--CA-T-CGGGTAAT-CTCCCA-TAC-AC-G-A-TG.........................ATGAAT--GCAG-AAA-TG-GT---T-GTAG-T-G-T-GGCTAA-C--CT-CA---C 7045
MON.NG.x.NG1            -----AC---C--G-----T--A-T----A-------T---TGGT-T-A-C-T----CT-CCTCT--CAC-A-TR--ACCAC-G---TGATG.........................AATGGGAAA-GAA-TG-AT-C-T-GTA----GGC-TGM-AA--C-CTC--A--G 5661
GSN.CM.99.CN166         -----ACC--C--------C---CAA--CT-----G----ATGG--C-AT--------G-CCTATCAGCC--G-A-TAATACC-GA-T-ATG.........................ATGAATGG-A-G-A-A-TG-AT---T-GTT---GGCT-TGGAGAAG-TTATCA-C-T 7037
GSN.CM.99.CN71          -----AC---C--G-----C---CAA--TT-----G----ATGG--T-ATC-------G-C-TATCAGCC--GGA-CAATACC-GAGT-ATG.........................ATGAATGGCAAA-A-A-TG-AT---T-GT----GGCT-TGG-GAAG-TTATCAGC-T 7028
MUS_1.CM.01.1085        -TAGG--T--C--G-----C--ATTA--CT-G--TG-T---TGG--A-A-C-T-----A-C-TATTT---G-G-A-TAATAC-GCAGT-ATG.........................ATGAATGGGGAT-AAA-TG-AT-A-T-GGA----A-T-TGG-GAA--TTTTAGA--- 7001
MUS_1.CM.01.CM1239      -TACG--T--C--G-----C--ATTA--CT-G--TG-T--CTGGT-A-A-C-T------CATATTT----G-A-TAATAC-GCAGT-ATG.........................ATGAATGGGGA--AAA-TG-AT---T-GGA----AGT-TGG-GAA---TTTAGA--T 6973
MUS_2.CM.01.CM1246      -TAGG--C--C--T----AC--ATTA--GC--A-TG-T---TGG--A-AA--G----T-CCTACCT-CC--GCA-T-ACAC-GC-GT-ATG.........................ATGAATGG-GAT-AAA-TG-AT---TTG------A-T-TGGAGAA---TTTAGG--- 7108
MUS_2.CM.01.CM2500      -TACG--T--C--G----C--ATT----C----TG-T---TGG--A-A-C------T-CCTACTT-CCT-GCA-T-ATAC-GCAGT-ATG.........................ATGAATGGCAAG-AAA-TG-GT---T-G-T---CA-T-TGGAGAA--TTTTAGG--C 7037
RCM.NG.x.NG411          --ACGTC---C-----CATG--A-ATA--AC------C--ATC-T-TGGA--------ATC-A-TAGT--GA-TAGAAC-TGG--CTAT...........................CAA-GAAAACAAAGT---AGG-C-GTT---GGGT--A-TAGAC-TTTTAA-T-- 6858
RCM.GA.x.GAB1           -CACTTC---C------ATG----ATA-TACT--GG-CT-AGGGT-TGGA--------ATCAA-TAAT-T-A-T---ACTTGG--ATAT...........................CAG-GAAGGCAAAGT---AGG-C-GTT---GGTC-CA-TAGTTTTTATAA-T-G 6814
DEN.CD.x.CD1            --AG-GTG------G---AG--ACACA-CA-G-CCAGC--CTGGT--A-CA-G--CA-T-CAG---A-CTTCT---TAATCAT-CACA-ATT...............ATAAGAAACACA-GCAAAGATA-TC-GA-A-T-G----TT--T-CAGAAA--GTT-TGTAA-C 7160
                         V3 loop end

MAC.US.x.239            TCTCTTC---C---AGGATG--GGA-A--CAGACT--T---TGGT-TGGC--T-----A-C-AG-------A-TAGAACTTAT---TAC..........................TGGC--GGT-G---T---AGG-CT-TA--T-GTT--A-TAA-T-TTATAA-C-- 7524
Env                     V__S__S__C__T__R__M__M__E__T__Q__T__S__T__W__F__G__F__N__G__T__R__A__E__N__R__T__Y__I__Y__.__.__.__.__.__.__.__.__.__W__H__G__R__D__N__R__T__I__I__S__L__N__K__Y__Y__N__L__
H2A.GW.x.ALI            -T-CTAC---C---AGGATG--GGA-A-GCA-ACT--T---TGGT-TGGC--T-------C-AGG------A-TAGAAC-TAT--CTAT..........................TGGC--GGT-GA--T---AGA-CT-T---C-GCT--A-CAAAC-TTATAA-C-T 7587
H2A.DE.x.BEN            -T-CTTC---C---AGGATG--GGAAA-GCA-ACT--T---TGGT-TGGC--T-------C-AG-------A-TAGAAC-TAT--CTAT..........................TGGC--GG--GA--T---AGG-CT-T---T-GCT--A-CAA-T-TTATAA-C-C 7606
H2A.SN.x.ST             -T-CTAC-------AGGATG--GGAAA-GCA-ACC--C---TGGT-TGGC--T-------C-AGG------A-TAGAAC-TAT--CTAT..........................TGGC--GGT-G---T---AGA-CT-T---T-GCT--A-CAA-TTTTATAA-C-C 7033
H2B.GH.86.D205          -GTCCTCC--C---AGAATG--GGAAA--CAGTCC--T--ATGGT-TGGC--C-----T-CAAGG-----GA-CAG-AC-TAT--ATAT..........................TGGC--GAA-AA------AGG-CC-T-----GCT--A-TACAT--TATAA-T-G 7573
H2B.CI.x.EHO            T-TCT-TG...TACAGAATG--GGAAA--CAGACC--T--ATGGT-TGGC--C-----T-CAAGG------A-CAG-AC-TAT--CTAT..........................TGGC--GGT-AA------AGG-CT-T-----GCT--A-TTCAT-TTATAA-C-G 7565
H2G.CI.92.ABT96         TCTCTAC-------AGAATG--GGAAA--CA-ACC--T--ATGGT-CGG---T-----A-CAAGG------A-TAGAAC-TAT--CTAT..........................TGGC--GGG-GA------AGG-CT-TA----GTT--A-TAA-T-TTATAA-T-- 6937
H2U.FR.96.12034         TCTCTAC---C---AGAATG-GGG-AA--CA-ACC--T---TGGT-TGGA--T-------C-AG-T-----A-TAGAACTTAT--ATAC..........................TGGC--GGT-GA------AGA-C--TA--T-GCT--A-TAA-T-TTATAA-T-- 7066
MAC.US.x.EMBL_3         TCTCTTC---C---AGGATG--GGA-A--CAGACT--T---TGGT-TGGC--T-----A-C-AG-------A-TAGAACTTAT---TAC..........................TGGC--GGT-G---T---AGG-CT-TA--T-GTT--A-TAA-T-TTATAA-C-- 6655
SMM.US.x.H9             TTTCTTC---C---AGAATG--GGAAA--CA-ACC--T---TGGT-TGGT--C-----T-C-AG-------A-TAGAAC-TAC---TAT..........................TGGC--RG--GAAG----AGA-CC-TA--T-GCT--A-TAA-T-TTRTAA-C-- 7001
SMM.US.x.PGM53          TTTCTTC---C---AGAATG--GGA-A-GCA-ACC--T---TGGT-TGGT--C-----T-C-AGG------A-TAGAAC-TAC---TAT..........................TGGC--AGT-GAAGT---AGA-CC-TA--T-GCT--A-TAA-T-TTATAA-C-- 7458
SMM.US.x.SIVsmH635F_L3  TTTCTTC---C---AGAATG--GGA-A-GCA-ACC--T---TGGT-TGGC--C-----T-C-AGG------A-TAGAAC-TAC---TAT..........................TGGC--GG--ATAGT---AGA-CC-TA--T-GCT--A-TAA-T-TTATAA-C-- 7531
SMM.SL.92.SL92B         T-TC-TC-------AGGATG--GGAAA--CAGACT--T--ATGGT-TGGC--T-----T-C-AG------GA-TAGAAC-TAT--ATAT..........................TGGC--GG-GGAAGT---AGA-C--TA--T-GCT--A-TAA-T-TTATAA-T-G 6960
STM.US.x.STM            TCTCCTCG-----TAGAATG--GGA-A--CA-ACT------TGGT-TGGC--T-----A-C-AG-------A-TAGAAC-TAT--CTAT..........................TGGC--GG--G---T---AGG-CT-TT--T-GCT-GA-CAA-T-TTACAA-T-- 7163
MNE.US.x.MNE027         TCTCTTC---C---AGAATG--GGA-A--CAGACT--T---TGGT-TGGC--T-----A-C-AG-------A-TAGAACTTAT---TAC..........................TGGC--AG--AA--T---AGG-CT-TA--T-GTT-GA-TAA-T-T-ATAA-C-- 6995
DRL.x.x.FAO             ----G--G-----TAGACAT--GCCTG-CAC-A-C--T-GC--TT-TGGC--C-------CAAAGCAT-GTC-G--T-AGT----AGA-ACA...............AGGGGAAAACCAG-C--TCT-A-TC--AAAT-TGTGT-T-GG--TA-TAA---ATGGAACC-C 6869
MND_2.x.x.5440          ----G--G--C---GCACCGC-GCCAG-CAC-A-C--T--AATGT-TGGC--C-----A-C-AAGCAT--TT-T--T-AGC-T---CAGACA......AACCCACGAAAAGGAAAA-ATG-G--TCAT---C--AAAT-TGTATAT-G---GG-TAAA--ATGGGGAT-- 6835
MND_2.CM.98.CM16        -------G--------GCC-T-GCCAG-CACCACC------ATGTGTGGA-CT-G---G-CAAA-CAT--TT-TA-T-A-T----ACAGACA......AATACAAAGAAAGGCAAA-AGG-G--TCAT---C--AAGT-TGTGTAT-G----G-TGA---GTACAACT-- 7208
MND_2.GA.x.M14          -------G-----TG-CCC---GCCTG-TACTA--------ATGT-TGGC--T-----A-CAAA-CATA-TT-T--T-AGT-----TTGGTG......AACCCTCAA.........AAGG-G--TCAT---C--AAAT-TGTTTAT-GG--GA-CAAA--ATGGGGGT-G 7125
OLC.CI.97.97CI12        ----TACT--C--C-CGCC-C-GTCAG--AC---GAGT-G-TTCT-TGGC--T-----T-C-AG-CATA-G-C-AGCAG-C-G---CCTATA...ACAGGGGATTTGAATCATAAACCA-GAGAGGGAACAGGA-TAT-TGTT--GGC-A--GCTGAT--GTATAAGT-- 6694
GRV.ET.x.GRI_677        T---G--T-----TAGATT---C-ATA-TACTA--A-T--AGGGA-AGGA--------T---AG-T-----A-TAGAAC--AG--AT-G.....................CAGAAA-GAGGAAATGATA-TG--A-AGTT-TA----AGT-GA-TAA-TTTTACAACT-G 7089
VER.KE.x.TYO1_patent    T---G--T------GGCTT---G-ATA--AC---GA----AGGGT--T--C-G-----A--CTATCAT--GA-TCGAACCCAG--AT-G.....................CAGAAACAT-GGG-A-ATA-T--CA-AGT-TTG--CT-GT-CA-CAA-C--TATAA-C-- 7197
VER.KE.x.9063           T------C-----TGGACTT--G-ATA--AC---T--T--AGG---A-----G-----A--CTATT-T--GA-TCGAAC-CA---AT-G.....................CAGAAGCAT-GAG-G-ATA-TG--T--GT--TAG-TC-TT-TA-CAA-T-TTACAA-T-G 7218
VER.KE.x.AGM155         T---G--T--C---AC-TT---G-ATA--AC----A-C---GGT---T-A--G-----A--CTATT-----A-TCGAACCCAG--CT-G.....................CAAAAACATGGAG-G-GCA-TG-CT-AGTGTTA--CT-GC-CA-TAA-C-TTATAACC-G 7198
VER.DE.x.AGM3           TG--G--T-----TA-CTTG--A-ATA--AC---GA-T---GG----T----G-----G--CTACT----GA-TCGAACCCAG--AT-G.....................CAGAAACAT-GAG-A-GCA-TG-CT-AGTGTTAG-GT--T-TA-TAAAC-TTACAA-C-- 6716
TAN.UG.x.TAN1           T---TAGT-----TGGATTG--G-ATA-TAC---GAGCT-AGC-T-TGGCA-----------AG------A-CAGAAC--A---AT-G.....................CAAAAGCATGGAG-G-GCA-T---T--GTG-TA----A-C--A-TAAAC-TTATAAGT-G 7178
SAB.SN.x.SAB1C          T-TC-GCT--C--TAGGTT---C-ATA-TACT---AGC--AGG-A-AGGAA-T-----T---TAT-T--C-A-TAGAAC--A---CT-G.....................CAGAAAAATGGCAATT-CA-TG--T--GTT-TT--T-GGC-CA-TAGAT-TTTTAA-T-- 7389
MND_1.GA.x.MNDGB1       ----G--G-----T--GCACT-AGTAG-CAC----AGT-GCTTTT-TGGC--T-----A-C-A-GCATA-G...--A-G--A-T-G-T-CCCATAGATGATAAATATAGGGGCC-A-AGG-A--TCATC-A-GGAAGTTTGT-TAT-AG--G-CAGGA--ATATGGCT-- 6632
DEB.CM.99.CM40          ----GGC-------A--CTC--ACACA-CA-G-C---C--ATGGG-A-AA--T-----GTC-GAT-AG--GAGG-CA-A--AGC--CATATA......ATAAGAAAAGAGGTAAAA-GTG-GG-GCA-A-TGGAAGC-TT--T----G---G-CCGCT--ATACAACT-G 6858
DEB.CM.99.CM5           ----GGC-------A--CTC--ACATA-CA-G-C------ATGGA---AA--C-----TTCAGAT-A----AGG-CA-A--AC--ACATATC...............ATAAGGAAG...GGCC-GGAAA-T--GA-C-TT--TG-C-G-A-C-CA-CA--GTATAA-C-- 6797
TAL.CM.01.8023          -------C--C---G--CCC--C-CCA-CA---C-G----ATGGT--G-TC-------T-CAAAG..........-A-TCAGAC-CAA-T-ATA.....................CA-CAGGCACC--AAA-TG-AT-AGCC-TA--CC-CT--GG-AA-G-ACACAACA-C 6432
TAL.CM.00.266           -------C------G-CCCT--C-CCA--A---C-G-T--GTGGT-AA-TC-C-----G-CAAAG..........-A-TCACAC-CAGGTCATA.....................CA-CAGGC-CCA-A-A-TG-AAG-GCT-TA--CC-CC--GGAAA-G-ACACAACC-C 6888
SUN.GA.98.L14           -TACC-CT-----TA--TAT--G-CAAGTAC---TAGT-GCTTTT-TGGC--T-----A-C-AG-CATA-G-----T-AGT----ACC-ATC............AATAACAAGGTA-GAC-AGGGG-T--GGGA-AAT-TGTATGG-A----GCAGCT--GTGGGGGT-- 7357
WRC.CI.97.97CI14        -----ATG------G-CATG--G-CAG--ACCA-TAGT-GCTTCT-TGGA--T-----A-C-AG-CATA-G--G-GA-A-T-----GA-GTA...ATAGGAAATATAAATCATAAACCA--GGAAGGATCAGGA--ATTTGTGTAT-AG--G-CAGAC--GTATAA---T 7201
WRC.CI.98.98CI04        -----TT---C---G-CATG--G-CAG-TACTA-TAGT-G-TTCT-TGGA--T-----T-CCAG-TATA-GA---G--A-T--G-AGAGATA...ATAGGAAATATAAATCATAGGCCTC--GAAGGTACAGGA--ATTTGTATAT-A---T-CTAGT-GGTATAAG--- 7249
WRC.GM.05.Pbt_05GM_X02  ----TACT-----TG--TT---G-CAG--ACTA-TAGT-GCTTTT-TGG---T-----A-C-AAGTAT--TA---G--A-T-G--AGA-GTA...ATAGGAAATATAAACCATAAACCA--AGATGGAACAGGA--ATTTGTTTAT-A-----CAGCT--ATATAA---G 6145
LST.CD.88.524           -TAC--C---C---G--TAC--G-TAAGTTC---TAGT-GCTTCT-TGGC--T-----A-CCAATCATAC------C-AGT-----CC-CTA.........ACTCCAAAAAAGATG-GAG--C-T-AT-GAGC-AAATTTGTTTAC-AG---GCAGGA--ATGGGGAC-- 6232
LST.CD.88.447           --ACC-CT--C---G-CTAT--G-TTAGTTC---GAGT-G-TTCT-TGGC--T-----A-CCAATCATAC---C--T-A-T-----CC-CTA.........ACTCCAAAAAAGATG-GA---C--CTA--AGCAAAGTTTGTATAT-A----GCAGGA--ATGGGGAC-C 6229
LST.KE.x.lho7           -TAC--C------TGGATAT--GCTTAGTAGT---AGT-G-TTCT-TGGC--T-----A-C-AATCATACCAGG--T-AGC-T---CC-CTA.........ACTCCAAATAAGATG-AAG--C-A-AT-GTGCAAAGTTTGTGTAT-A----GCAGGA--ATGGGG-T-- 7298
SYK.KE.x.SYK173         -----TC---C-----G-A-T-C-ATA--T-G-CTAGC--ATGGT-C-AAC-G-----G-CCTACAA--C-A----TAA---T-GATTTATC.....................AAACAA--GGAT-AAA-TG-GT-AGT--TT--TT----T-CAGAAGCACTTAGAT-G 7085
SYK.KE.x.KE51           ----TTC---C------CCTT-C-ACA-TT-G-CTAGT--ATGGT-T-A-----------CCTATAA--C-A----CAA---T-GGTTTATT.....................AGACAA--GGAT-AAA-TG-AT-AGTG-TT---C----G-CAGAA--ATT-GGAT-- 6781
COL.CM.x.CGU1           ----GTCC-----C----AGT---ATAT-AC---CATGT----TG-AT-AG-C----C-TCAAA--AGTTGTCT--TTGG-CG-AAGAT...................AGAGAGGGA---TGGAAG-AT---TCA-GG-C--TTGAGTATTATTG-T-TCCA-A-GA-A-- 6513
```

```
                           V3 loop start                                                                                            V3 loop end
H1B.FR.83.HXB2             GAAATTAATTGTACAAGACCCAACAACAAT...ACAAGAAAAAGAATCCGTATCCAGAGAGGACCAGGGAGAGCATTTGTT...........ACAATAGGAAAA...ATAGGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATAACACTTTAAA 7252
Env                         E  I  N  C  T  R  P  N  N  N  .  T  R  K  R  I  R  I  Q  R  G  P  G  R  A  F  V  .  .  .  .  T  I  G  K  .  I  G  N  M  R  Q  A  H  C  N  I  S  R  A  K  W  N  N  T  L  K
H1A1.UG.85.U455            A----C------T-C-----TT----T-CAAGA-A--AT-T-C-T-GGTA--GT...-T----T----ACA------CTA-...........GT--C---T---ATA----GG---A-----------------TG-C------AGGG-C-------GA---A--C- 6700
H1B.US.90.WEAU160          ------------------------------...-----------A---AAC-T-A.....------------------T-C--TA-...........----C----G--ATA------G----A----G-------------C---------A---GT------------------ 7259
H1C.ET.86.ETH2220          -------C------G-------G---T---...------G----T--AA-G--A...........-----------ACA-A----CTA-...........G---C----G-CATA------G----A-----------------------GA--A-------------A---C--C- 6652
H1D.CD.84.84ZR085          -----------------G---T----A--G...GA----C----G-CA-C---A..........----A----CA---GC-CTA-...........----C-A-GT-TACA-C-A-G-TA--AG--------T-------T------G---T-------------C---CG 6791
H1F1.BE.93.VI850           C-G---------------------------...-----G---G-T--A-A-T-A..........---------ACA-A----CTA-...........G---C----GC-ATA------G-C--A---A-G--------------------G--A--C-------------C-GG- 6571
H1G.SE.93.SE6165           --------------C--------------ACA-TG-A--GG-T--GAATGGGA...-TT---------ACA-A-G--CTA-...........G---C---TGCCATA------G-C--A-----------------TG---C--A-AG---------AG-GG-C---C- 6679
H1H.CF.90.056              A-C----CA-----C--G--T-----T---...--G----C---T--A-A-T-A........G-----AC-------CTA-...........G---C---TG-CATC------G----A-------------------T---------A--G-C--------G------C- 6580
H1J.SE.93.SE7887           ------GTG---TAC-----T--T------...-----G---G-T--A-AC--G..............-------ACA--TGC-CTAC...........G---C----G--ATA-----------A---G--A-----------------GA-AG-G-T----G----------CG 6562
H1K.CM.96.MP535            C-----------------G--A--------...------------T--A-A---G..............-------A-A------CTA-...........----C---TG-TATA------G----A----G--------C---------G---A----------C-TG-------G 6457
H101_AE.TH.90.CM240        -----C--------C------TC-------...--------C---T--AAC---A..............-------AC---T---CTA-...........-G--C----G-TATA------------A---A-----T-----G-G----A-G--A------------AGT------ 6829
H102_AG.NG.x.IBNG          A-------------C-----T---------...---------G-TG-A-A---A..............--------CA------CTA-...........G---C---TG-CATA-----GG----A-----------------TG-C----A-A--G----------A------C- 6773
H1N.CM.95.YBF30            CC---A-----------G--AGGA--T---...---G--GG-CAGG-G-AG--A............-----T-CT-TGA-----TA-...........-AC----A----ATAG-----G-C--T---------T-C-----TG-CTC--A--A-CT----G-ACCA-TG-GG-- 6830
H1O.BE.87.ANT70            A-C--G-CC---GA------AC-A-TA......GAC-T-C--GAG--GA-A--A........--T---AT-GCCTGG-ACAGCATGGGA......-T-GGG----C-GCAGG-AAC-GCTCA--GGC---TT----C--GTA--A-GCCA-TG-T---GGA--A-TA----- 7288
H1O.CM.91.MVP5180          A-C--G-CC--C-T----GAAGGA-TTGCA...GAGGT-C--GAT--ATA--CA........--T---AT----TGGCGCAG-ATGACACTTAAA-G--GTAAC--TACATC-CC--GATCA--GGT---TT------CATA--A--AGA-TGT----G-A--TG-CC--C- 7296
CPZ.CD.90.ANT              AC-----CA----T------AGGA--T-GGACAGT----...-ATC-A-AA--A..............-------A-TGA-T--CTA-...........-ACG---A--T-GCA-C----G-C-CT--GA----GTTC----CAG-C-A--AGA-GCT----G-AC-AG-ACGT-- 6702
CPZ.CM.01.SIVcpzCAM13      A-TC-G-CG----AG--GGTAGGG--T--C...-------GGC-A-G-A-AG--A..............-------A-TGA-C--CTA-...........-AT--G-ACC-TATAT-T--TG---CA--GA----CTT---CG-AC-C-A-G-GA---C-------G-A--AC-GC- 6860
CPZ.CM.05.SIVcpzEK505      CC---A-----C---------AGGA--T--G...-------GG-CA-G-A-AA--A............-----T--T-TGA----CTA-...........-AC----A---TATA------G-C-CT--G------T-----G-GG-A-AC--GA--...---G-AC-A-TA-GG-- 6783
CPZ.CM.05.SIVcpzMB66       CGTC-C-----C--T---A-AGGA--T--CACA-T--A-GGT...--A-C---A............-----A-TCAGAT----TA-...........GG-----A--CTGTA------G-C-CA--G--G--CTT---CC-GT-G-AC-A-A--GT-----C---T--A--T-- 6891
CPZ.CM.98.CAM3             AGC--A---------------AGGA--T-NC...-GC---GG-CA---A-AG--A..............--------C-TGA-C---TA-...........-AC----A---TATTG-----G---CC-----G----T-----C-G--C-A---GA--GTC---G-CG-A-GAC---- 6680
CPZ.GA.88.GAB2             CCC---------C-GT-----AGGA-------...---G---G-GC---AAACC-T......A-C--------C-A----CT-C...........-AC-CG-A-GCCCTA---------CCCT--GA----GTCC--CC--C---A-G-TA--CT------------T-C---- 6663
CPZ.CM.98.CAM5             A-C--A-T-----TG-----AGGA--T--C...-GC---GG-CA-C-A-AG--A..............---------GT-T-C--CTA-...........-AT----A---CATAG---GG--CCC-----------------TG-C--C-AG--CC-G----TG--A-AG----- 6956
CPZ.CM.05.SIVcpzMT145      A-G----CA----GG-----AGGA--T--C...-------GGTCA---A-AA--A..............------CA-TGA-T---TA-...........-AC----A---TGTAG---GG-C-CT---A-----T-----G-G--A-A-G-GA--C-G---GCA--AG-CC-C-- 6770
CPZ.TZ.00.TAN1             A-TT-A-CG---GA----A-AGGA--------...-------GGTCA-G-A-AG--A........--T------T-TGA-C---TA-...........-AT----A---TGTAG---GG-C-CC--GA----TT-----TCAG-C-A-GC-A---C----T-C-GG-AC---G- 6890
CPZ.CM.05.SIVcpzLB7        C-C--C---------T--GGTAGGA--T--GACA-T--A-GG-...----CCC-A..............--------ACAGCTG--CTA-...........GG--C--A--CTGTAG-----G-C-CA---G-G--T-------G-A----ACCAGA--C-T---T----A-TAC---- 6781
CPZ.US.85.CPZUS            A-T--G-C-----TT-----AGGA-------...---GTGGG--ATG-A-AAC-A..............--------A-TGA-T---TA-...........-AC---CC---GATAG-----G--G-A---G-----------------CTCA-A-CTG-C----G-G--ACAGAG--- 7284
CPZ.GA.88.GAB1             AGTC-A------CAT-----AGGA-------...------GGG-GAGG-G-AG--A........--G------A-TGA-C---TA-...........-AT----A---TGTAG-----G---CC---TCT--CT-C-----G--A-A-G-GA-C-C--------CGA--GG-GG- 7310
MON.CM.99.L1               A-C----C---C-TT--G--AGGA-----A...--T-TT-GG-ATT-A-AG--A..........---G----C-TGA----CTA-...........T-GCA-CTT-T-GTGGG---T----CTC-TA----CT-C--C--AG---A--AGA-CC-----G-G-CTG-CC--C- 7194
MON.NG.x.NG1               A-C--A-C---C-TT-----AGGA--T--G...--CTT--G--ATT-A-AG--T........--TG----A-TGA----CTA-...........T-GCA-ATC-T-GTGGG------C-CY---A-G---T-C---C-GG-A-A--AGA--C-G---G-CGCTG-C---C- 5810
GSN.CM.99.CN166            A-TC---CA----TC-----TGGA--T--A...--C-T--G--ATT-G-AA--A..........---G-G--A-TGA-----TA-...........T-CCA--TG-T-GTAGG--------CT--GA----CT----CCGAC--GA-CCT-AT-GG--------GGG-AA---G 7186
GSN.CM.99.CN71             ACTC---CA----TC-----TGGA--T--A...--CCT-----ATT-A-AG--A..........---G----A-TGA----CTAC...........T-CCA-ATT-T-GTAGGG---G---CT--GA-G--CT----C--AC-C-ACCATA----G---G---TAG-AA---G 7177
MUS_1.CM.01.1085           A-TC---C-----T------AGG-----G...-----T--GG-ACC-T-AA--A..........---G----C-TGA-T--CTA-...........T-GCAGCTT-T-GTAGAT--CG---CAC-CAG---CT----C--A--A-A--AGA--C-G---G--CCAG-A--G-G 7150
MUS_1.CM.01.CM1239         A-CT-A-CC--CTT---G--AGG---T--G...--T-T--GG-ACC---AG--A........--GG----C-TGA-T--CTA-...........T--CAGCTT-T-GTAGG---GG-C-CAC--AG----T-C--C--G--A-AC-A-A--C-G---G--GTAG-A----G 7122
MUS_2.CM.01.CM1246         A-TC-A-CC----TT--G--AGGA--T--G...-----TT-G--ATC-A-AG--A..........---G----C-TGA-C--CTA-...........T-TCAGCTG-T-GTGGAT---G-C-CCC-CAGG---T----C--G--C-A--A-A-TC-----G--CCTG----G-G 7257
MUS_2.CM.01.CM2500         A-TT-A-CA--CCTG--G--AGGA--T--G...--T-TT-G--ACC-T-AA--A..........---G----C-TGT-T--CTA-...........T-TCAGCT--T-GTAGAT---G---CCC-CAG---CT----C--A--C-AC-AGA-CC-G---G-CCC-G-A----G 7186
RCM.NG.x.NG411             AC-G-A-C-----GG-----A-GT----GA...---GTG---G-----TCAT-A......-C-A---AGT-TTC................-TCTCTCT--GGGTGCA-AA--GGCCA-A-GG---TTGG----GAT--...GAT-G-G-T-----AGGTG-G-GG-- 6998
RCM.GA.x.GAB1              TC-G-A-CA--C-GG-----TTCA--T-GA...---GTG---G-G--ATCGC-A......-C-A---AGTCTTT................-TCTC-CT--G-GTAGAGAAG-GACCA-A-GG---TTGG----GAT--...GA--GG--T----CGG-TG-A-GG-- 6954
DEN.CD.x.CD1               A--T-G--G-----C--G--AGGG--TGTA...--CCT---GG--T-AAAC--G......-C-G----A-T--TCA-GAGA...........G-C-CTATT-T-CCA-G-A-GGC---T--GAG----GCA--C--G--AGTAG---GC--C----C--GA--AAGGC- 7309
                           V4 loop start                                                                                            V4 loop end
MAC.US.x.239               AC---G--A----G------AGGA--T--G...---GTTTT-CC-G--ACC--T.....ATGT-T--ATTG-TT--C..............CACTC-CA-CC-ATC-AT-AT-GGCCA-AG--G---TGG---TGGT--...G---G---------GG-TG-AA---- 7667
Env                         T  M  K  C  R  R  P  G  N  K  .  T  V  L  P  V  T  I  .  .  M  S  G  L  V  F  .  .  .  .  .  .  .  H  S  Q  P  I  N  D  R  P  K  Q  A  W  C  W  F  .  G  G  K  W  K  D  A  I  K
H2A.GW.x.ALI               ACT--GC------AG--G--AGGA--T--G...---GTTGT-CC---AACAC-T.....ATGT-----TT-AT---..............CACTCCCAGCC-ATC-ATAA--GACCC--------TGG--CTGGT-C...-A--GCG------GG--AG-CA-GC- 7730
H2A.DE.x.BEN               AC---GCG-----AG-----AGGA--T--G...---GTTTT-CC---AACAC-T.....ATGT----ATT--TG---..............CACTCTCAGCC-ATC-ACAC--GGCCT--G--G---TGG--CCGGT--...G---GC-G-----GGG-AG-CA-GC- 7749
H2A.SN.x.ST                ACCG-AC------AG--G--AGGA-----G...---GTTGT-CC---AACAC--.....ATGT-----TT--TG---..............CACTCCCAGCC-ATC-ATA---GACCC--G------TGG--CTGGT-C...-A--GCG-G-----GG-AG-CA-G-- 7176
H2B.GH.86.D205             TC---AC-C----AG--G--AGGA-----G...--GGTTGT-CC---AA-A-C-......-TGT----ACT-CTT--C..............CATTC-CAGCCTATC-ATAAG-GACCC--------TTGG--CTGGT--...-AG-G---C----CAG-AG-CA---- 7716
H2B.CI.x.EHO               AC---GC-C----A---G--AGGA-----A...-TGGTTGTGCC---AA-A-C-......-TGT----T-TTCTC--C..............CATTC-CAGCCTATC-ATAA--GACCT-A------TTGG--CTGGT--...-A--G---C----CAG-AG-CA--C- 7708
H2G.CI.92.ABT96            AC---A-RA----GG-----AGGG--T--A...---GT-TT-CC---TAC---T.....ATGT-----CT--TC--C..............CACTCTCA-CC-ATC-AT-A--GACCC-----G--TTGG---TGGT--...G---GC---------GG-AG-CA-GC- 7080
H2U.FR.96.12034            AC---GT-C----G---G--AGGG--T--G...---GTCTT-CC-G-AACC--T.....ATGT----ACT--TG--C..............CATTCTCA-GCTATC-ATAAC---CCA-A------TTGG---TGGT--...G-T-G---T-----AGGGG-AA-CC- 7209
MAC.US.x.EMBL_3            AC---G--A----G------AGGA--T--G...---GTTTT-CC-G--ACC--T.....ATGT-T--ATTG-TT--C..............CACTC-CA-CC-CTC-CT-AT-GGCCA-AG--G---TGG---TGGT--...G---G---------GG-TG-AA---- 6798
SMM.US.x.H9                AC---G-GA----G------AGGA--T--G...---GTCTT-CC-G--ACC--T.....ATGT-----TTG-TC--C..............CATTCGCA-CCCATA-AT-AG-GACCA-A---G--CTGG--CTGGT--...G---G-G-------AG-GG-CA-CCG 7144
SMM.US.x.PGM53             AC---A-GA----G---G--AGGA--T--G...---GTCTT-CC-G--ACC--T.....ATGT-----TTGATC--C..............CATTCGCA-CCCATA-AT-AG-GACCA-A------CTGG--CTGGT--...G---G--GT----GAG-GG-CA-CC- 7601
SMM.US.x.SIVsmH635F_L3     AC---G--A----G------AGGA--T--G...---GTTTT-CC-G--ACC--T.....ATGT-----TTG-TC--C..............CATTCGCA-CCCATA-AT-AG-GACCA-A---G--CTGG--CTGGT--...G---G---C----GCG-GG-CA-CC- 7674
SMM.SL.92.SL92B            AC---A--G--C-G------AGGA--T--G...---GTCCT-CC-G--ACC--T.....ATGT----TTTG-TC---..............CATTCGCA-CC-ATA-AT-AG-GGCCA---------TGG--CTGGT--...G---G--------GAG-AG--A-GC- 7103
STM.US.x.STM               AC---G-GC----G------AGGG-----A...---GTCTT-CC-G--ACC--T.....ATGT----CTTG-TC--C..............CATTCGCA-CCCATT-AT-AG-GACCA-A------TTGG--CTGGT--...G---G-G------GAGGAG-AA---- 7306
MNE.US.x.MNE027            AC---G--A----G------AGGA--T--G...---GTTTT-CC-G--ACC--T.....ATGT-T--ATTG-TT--C..............CACTC-CA-CC-ATC-AT-A--GGCCA-A---G---TGG----GGT--...GA--G---T-----GG-GG-AA---- 7138
DRL.x.x.FAO                A------TC----G----AAAGGA--T-GG...-GC-T--T-TCT-CA-CCTCA......-CCT-----TTGTTG---TA-.............CATGGGTT-G--CCAGG-AA----T-A-A-AG--G-ATG-----AT--...GA--G-C-----GGA--GG-----GG 7015
MND_2.x.x.5440             C-GG-A-GG----G---GAAGGGA--T-GA...T-T-TC-TTTCC-CA-CA-GT......TCTA----ATTGCTG--CTA-.............CATGGGTTGG-GCCAGG-AAG---C-T---A-G-GCA-G---C-GT-A...GA--GG------GGAC-GG-----C- 6981
MND_2.CM.98.CM16           C--G-AGTA----G----AAAGGA----GA...T-TGTG-TCTC--CA-C--GT......-C-A-G--ATTGATC---TA-.............-GTGGGTT-G--CCAGGGAA----C-T-AGA---GTATG---C-GT-G...-AG-G-C-----GGA-TAG-CA-GC- 7354
MND_2.GA.x.M14             A-GG-GGTA---GAT--GAAGGGA----GG...T-T-T-GTTTC-G---CA-GT......-C-T----ATT-ATT--CTA-.............CATGGGTT-G--CCAGGCA-G---C---A-A---G-ATG--C-GGT-A...-T--G-C-----GGA-GAG----G-- 7271
OLC.CI.97.97CI12           -----AG-G----TT--GAAGGG---T-GA...-GT-T-GTTTC-G-TAA-TCA......-CTT----TCTG-T--A-TA-.............G--GG-TT-G--CCTTAT--G-----AGATA---GCTTG----GGT--...-AG-G-C-----GGATGG-TGC--T- 6840
GRV.ET.x.GRI_677           AC-G-G-GA--CCG------TGGT--T--A...---GTGTTGCC-G-AACA---.....ATGG-----TT--T---..............CACTCTCAG---TAT-ATACC-GGT-A-A------GTGG--CC--T-C...CA--G-G-T-----AGGGG-A-GG-- 7232
VER.KE.x.TYO1_patent       TC-G-C-CC----GG-----AGGA-----G...---GTCCT-CCGG-AACG--A.....ATGG-G--ACTG-TT--C..............CACTCTCA----TAC-ACATG--GC-T-----G--TTGG---C--T-C...GA--GC--T----GAGGTG-C-GGCG 7340
VER.KE.x.9063              AC-G-C-C---C-A---G--AGGA--T--G...---GTCTT-CC-G-AACC--A.....ATGG----TTT--TT---..............CATTCGCAG---TAC-ATACG--GC-T-----G--CTGG---C-TT-C...CA--GC-GG----GAG-AG-C-GG-- 7361
VER.KE.x.AGM155            AC-G---CA--C-A---G--AGGG--T--G...---GTCTTGCC-G-AACG--A.....ATGG----ATT--TC--C..............CACTC-CAG--GTAT-ATAC--GAC-A--G--G--CTGG--CC--T-C...CAG-GC--T-----AGGAG---GG-- 7341
VER.DE.x.AGM3              AC-G---C---C-A------AGGA-----A...---GTCTT-CC-G-AACA---.....ATGG-----CT--TG---..............CATTCTCAG-GGTAC-ATAC--GGC----------TTGG---C--T-C...CAG-GC--C----GAGGAG-C-GG-- 6859
TAN.UG.x.TAN1              A-G--AGTG----G------AGGA--T--G...---GTCTT-CC-G-AACC---.....ATGG----TCTC-TG--C..............CACTC-CA-C--TAT-ATAC---AT-A---------TGG--CC-TT-C...CAG-GTG-T-----AGGAG-C-GG-G 7321
SAB.SN.x.SAB1C             ACT--C-GA----G------AGGA--T--G...---GT-TTGCCTG-AACC---.....ATGG-T--CTTG-TC--C..............CAT-GCCAG---TAT-ACACC-GGT---A------TTGG--CTGGT--...G---GT--T----GAGGAG-C-GG-- 7532
MND_1.GA.x.MNDGB1          A-G--AG-A---CAC---AAAGGA--T-GG...T--GT-GTG--T-CT-CATCA......-CTA----ATT-TT-----TA-.............CATGGGTT-G--CCTGG-AAG---T-A-AGA---GCATG--C-C-T-C...-A--G-CGT---GGGTTAG-AC-TTG 6778
DEB.CM.99.CM40             AC-T-A-CA---GT---G--AGGA--T--G...--CTACCG-GCC--A-AC--G......-CTA----TCTGT-C---TA-.............--T-CCTTT-T-CAA-GGCTG-GG--A-AGAG-----------GAC-A...-AC-G--GC---GCA---G-AAC--- 7004
DEB.CM.99.CM5              TCTT-GG-----GTT--G--AGGG-----G...--GTAT-G-GCTG-T-A---G......-CCA-----CT-T-C---TA-.............----CCTTC-T-CCA-GGCTG-GG--A-AGAG-----------GG--G...----GT--T----CAGGAG-AG-T-- 6943
TAL.CM.01.8023             ACC--AC-G--C-TT-----AGGA--T-GA...----TC---GATC-T-AG--A......-C-G----TCT-ATG---CAC............-GTCAGATC-T-GCAGGTAA-G--T-A-AGAG---CT-------G--A...GA--G---T----CG-GGG-GA-GG- 6578
TAL.CM.00.266              AC---C-CA--C-T------AGGA--T-GG...----T---GGACT-A-AG--A......-C-G----ACT-ATG---CAC............-GCCAGAT--T-GCAGGGAAG--CC-C-ACAGG--CT-------GC-A...GAG-GC--T----C---AG-CA-GG- 7034
SUN.GA.98.L14              -T----C-G--C-T---GAAAGGG--T-GG...T--CA-GT---T-CTATA-GT.....A-TA----ACTGCTG---TA-.............TATGGGCT-G-GCCAGGTTCT--GT-A-----TG-----G-----AT--...GT--G-C-----GGA-GAG-C--TGC 7503
WRC.CI.97.97CI14           A-G--AG-A---CAT--GAAAGGT--T-GA...T-TTT-GT-TC-G-AAA-TCA......-CTT----TTTGAT--A-TA-.............G--GG-TT-G--CCTCATA-G-----T---A----GCTTA--C--AT--...-CT-GGCTT---GGA--A-TGC-TT- 7347
WRC.CI.98.98CI04           --G--AG-A--C-AT--GAAGGGG--T-GA...T-TGT-GTGTC-G-AAACTCA......-CCT----ACTTAT--A-TA-.............G--GGCTT-G--CCACATA----C--A--GA---GCTTA--C--GT--...-A--GGTTG---GGAC---TG---T- 7395
WRC.GM.05.Pbt_05GM_X02     -CT--AG-A---CAT---AAAGGA--T-GA...-GTCT-GTCTC-G-AA--TCA......-CCT-----TTGATT-A-TA-.............G--GG-TTGG--CCTTATAAG-----T--GA-----TTA--C--AT--...-C--G-C-G---GGAC-T-TG---T- 6291
LST.CD.88.524              ATC----GA----T----AAAGGA--T-GG...T--GA-GTCTCC-CAATC-GT......TC-A----ACTTCTG---TAC............TATGGGTT-G--CATGGGTCT-GAC-A---TTG-----A--C--AT--...GAG-GTC-----GGA-GA-TG--T-- 6378
LST.CD.88.447              ATT----GA----TT--GAAAGGA--T-GG...T--GA-GT-TCT-C-AT--GT......TC-A----ACTTCTG---TAC............TATGGCTTGG--CATGGTA-C-GGC-----TTG--T--G-----AT--...GA--G-C-G---GGA-GA-TG--TC- 6375
LST.KE.x.lho7              ATT----GA--C-T----AAAGGA--T-GA...T--GA-GT-TCC-CAAT--G-......TCTA----ATATTT----TA-.............TATGGCCT-G--CATGGCA-C-GAT-A---TT------G-----AT--...GA--GCC-----GGA-GA-TG--T-- 7444
SYK.KE.x.SYK173            C-G----TA---GA------AGGA--TG-G...T-T-TC----AT--T-AAT-G......-CTG----CTATTTCC-CCCAGTAATA...CAAGGG-A-TTG---ACCGGGA-GG-CGCA-A--GT--CTTC---CGAG-A...-C--G---T----CAG-ATT---T-- 7240
SYK.KE.x.KE51              C--C-GGTG---GA------AGGA--TG-G...T-T-T-----AC----AAT-G......-C-G----ATTTTTC--GCCAGTGATT...CAAGGG-G-TTG---ACAGGGAA-GC-GC--A-AG----TT---C--AG-C...-A--G-G-----GGA--GTT---TG- 6936
COL.CM.x.CGU1              -CTT-AGGG----T----AGAGGA--T-G-...T-TCAC-G--ATC-AAAC-CT......-CCAAT--AGC-AA---CTAC............TATGAGCTT-TTCCTTATTC---AGGC-TTT-T-G-AGG---C-GT--GTACC-ATG-CTG--C-G--T-AGAA--- 6662
```

```
H1B.FR.83.HXB2     ACAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATAATC.....................................................TTTAAGCAATCCTCAGGAGGGG 7326
Env                 Q  I  A  S  K  L  R  E  Q  F  G  N  N  K  T  I  I  .  .  .
```

```
                                                V4 loop start
H1B.FR.83.HXB2      ACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGTACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACT.............................GAAGGAAGTGACACA 7463
Env                 D  P  E  I  V  T  H  S  F  N  C  G  G  E  F  F  Y  C  N  S  T  Q  L  F  N  S  T  W  F  N  S  T  W  S  T  E  G  S  N  N  T  _ __ __ __
```

```
                          V4 loop end                                                                                                                                      V5 loop start
H1B.FR.83.HXB2            ATCACCCTCCCATGCAGAATAAAACAAATTATAAACATGTGGCAGAAAGTAGGAAAAGCAATGTATGCCCCTCCCATCAGTGGA.........CAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAAT..................AG 7606
Env                        I  T  L  P  C  R  I  K  Q  I  I  N  M  W  Q  K  V  G  K  A  M  Y  A  P  P  I  S  G  .  .  .  Q  I  R  C  S  S  N  I  T  G  L  L  L  T  R  D  G  G  N  .  .  .  .  .  .  S
H1A1.UG.85.U455           --A--T----A---T--------G----------T----------G-------C-----------------------CAA--.........GT---A--G---GA------C--------A--A-----------------G----C-AAC................ 7037
H1B.US.90.WEAU160         -----A-----------------------------G-------G-----------------------A------GAA--.........--------------T-----------------T-A---------------------G-AGT..............GA 7610
H1C.ET.86.ETH2220         -----A----A-------------------------------GG-------CG-------------------TGAA---.........AT---A-TG---AG---------C-----A--A---C-G---C-T-----A-CC--A..............GAGCC 6985
H1D.CD.84.84ZR085         --A---A-----------------------------------G-------------------------------TGAA--G.........-------AT-------------------------A-----G---------------GC-.................-A 7127
H1F1.BE.93.VI850          -----T--------TC-G-----------G---------GG-------CG---------A--T-------TGCA---.........A-C----CC---AAC-------------TT-A-----G-----------------..................... 6884
H1G.SE.93.SE6165          -----A--------T-AG-----------G-G-GA---------GG--G---C-----------------------GCA---.........A-C---GA----AAC----GC--------A--AA-----------------G---AAT......AACAATACA-A 7024
H1H.CF.90.056             --T--A--G--------------------G----------------G------CG------------------------CAA---.........A-C----TG---GT-----------------A--AA----G----TT--C-AG-----C..................GC 6925
H1J.SE.93.SE7887          -----AA-A-----T-A-------------G-G-GA--------A-GGAC----C-------C--------------TGCA---.........A-C----CG--CA--------------AT-A-----G---C-T-----A--A--CAGG......GGGAATGGC-- 6913
H1K.CM.96.MP535           -----A--------T--------G--G-----------------------G-----------A--------------TGCA---.........AGT----AC---AGC---------------AA--A----G----------------.................-A 6766
H101_AE.TH.90.CM240       -----A--T-------AG-----G---------------------GG--C----C------------T---------------.........AG-----AT---GT---------------AA-A-----G----------------GT-AAT................ 7160
H102_AG.NG.x.IBNG         --A--T----A-------------G-------G----T------------------C--------------------------CAA---.........AT---A------GAT-----C--------A--A-----------------A--G---................ 7113
H1N.CM.95.YBF30           --T--T--G--T--T--------G----G----G-----TT-----ACA-GG----------G---T-----A--A--A--TC-G---.........GTTC---AC---A-C--C--------T--A---G-TC--GA-TATAG------GCC-................. 7152
H1O.BE.87.ANT70           GG---T--A--T----A-C-G-G---GG-GG---GGTCA---ATA-GG-G-CAGTCG-G-C-C-----A----------AA--T.........A-TC-A-C----ATG-----C--A--T--AA--A-CC--CA--TG---AACACATGGAACAGC............. 7649
H1O.CM.91.MVP5180         GGT--TA-A--T----AGT---G---GC-AG---GATCA---AT---G-G--AGTCGAG---C-----A---------CCC--C.........A-CT-A-C----CAT--C--C--A--T--AA--A-TC--CAGTT----CAACCATGG.................-A 7662
CPZ.CD.90.ANT             ...-TAGCA-AT--------T--G--G--AG-T--TCAT---GGC-T---TTC-----GC--T--CTTAG-C--T-GG--A--G.........A-TG--TCC---A-T--C-GC--A--T--AA-TA-G--G...GA--G-CAAATAT--..................... 7036
CPZ.CM.01.SIVcpzCAM13     --G--AA-A-----------T-G------AG-G--TTCA---AT--G--------GG-G---C-----A-----A--ACCA---.........--T--A-CT---AAC--TCTA-----T-----CA----G---C--...-ACCA-GT-...................... 7182
CPZ.CM.05.SIVcpzEK505     ------T-A-A-----------G-------G----TT-A---ACA-G------------G---T-----T--G--A--T-AA---.........-CT----AC---CTC--T--C--C--T---A-AA--C-TGAC.........................TATACAAAGTC 7086
CPZ.CM.05.SIVcpzMB66      -----TA-T--T----------GG------G---GGC-A-----A-G---------G--G---A-T-CTT--C-----T--A---.........AC---A-AC---ATC--G-----C------A-AT-G--TG-G...............CAGCAA...AAAACAGAC-- 7227
CPZ.CM.98.CAM3            ...TT-T-------T-------GG------G-----TCC---AT--G---G--G--G-G---T--------A--A--T--A---.........GTT--A--T--CA--------CG----T---A-TA--C-TGAG-C--GGCA---C-TC...............AAC-A 7016
CPZ.GA.88.GAB2            TA--TT----------A---T-GG---G-AG----TTCT---AT--G---G--G----GT--C-T------G--T-GA--A---.........ACT--A-CC---AAT--C-CC--------A--C--C---GA-GT-C-GAA-.......................G- 6975
CPZ.CM.98.CAM5            C-A--AA-AT----T-------GG------G-----TCC---AT--G---G--G----G---C--------A--A--A--A--G.........A-C--A-C---CAT---------A--T--A--C--CC-CGA-GC--GA--AAAATGGGAA...............-A 7328
CPZ.CM.05.SIVcpzMT145     -CA-TAA----T--T-------G------AG-G--TTCC---AT-CGT-----G----G---T-T---A--A--A--T--A--T.........A-C----CC--CA-T--T-----C----TA-----C--GAG-T-C--AAGAACCGAGAAGATCAGGGAGAAGATCA 7103
CPZ.TZ.00.TAN1            ...GTTGCA---------C-G-GG--G--AG----TCAT---GGT-T---GTC-----GGG-T--CCTT--C--A-GG--G---.........AC-G-A-A----CAC-----C--C-----A--TA-CA-G---GC---AAAA-AC--CAAT................ 7230
CPZ.CM.05.SIVcpzLB7       T-G--A--A--T--T--G--T-G-------G-----C-A----TA-G---------G-G---A-T---T--A--T--T--A--T.........A-C--CC-C---AAT--C-C------T---T--A-CC--GA-...........................AAACAC-C 7117
CPZ.US.85.CPZUS           ...--A--TAAG--T-------G-------G----TCA----AT--G-----------G---C-T------A--A----AA--T.........GTGC-A--T---AAC--------A--G--AA--A-TC-TGAC-T-AGCATAA-CGCAGTC............AAT-A 7608
CPZ.GA.88.GAB1            ...-TA--G-----------T-G---G--AG-T-GTTCA---AT--GG-------G--G---A-----A--C--A-----A---.........A-C--C-CT--CAAT--T--C-----T--A--T--C-----TTC---CAC-CC-GTGACT...............-A 7628
MON.CM.99.L1              GGGTGG--AG----T-CC----G------GG-G-CACAA---GGAT-T--GTC---GAGT--C--CCTA-----ACGG-AA--C.........--TG-GCAG---A----T-----C--T-CC--C--GA-T---G-G--GTTGTA-C-AAAC................ 7579
MON.NG.x.NG1              GAGTGGT-AA-G---C--C-C-GG---C-AG---CTCA----GGAT-CA-CTCC--GAGTR-C---CTG--A--A-GACAA--G.........--TG-GCAG---A-C--T-----C----CC--CT-GG--GATG-G--GTTGTA-GGCAGC................ 6204
GSN.CM.99.CN166           CAATGG--GG-T----A---T-GG---T--G-C-C-CAA---GGAT-T--GTCT---T----C--CTTA-----ACGGCAG--T.........--T--A-AT--CA----C--C--------T---T-GA--GATG-C-C-ATGTA-G-AAGC................ 7553
GSN.CM.99.CN71            CAATGGA-GG-T--------T-GG---T--G-C-CACAA---GGAT-T--GTCT-GGT-G--C--CTTA-----ACGGCAA--T.........--T--A-AT--A----T--C--C--T--C--C----A--GACG---CGATGTA-GGACAT................ 7559
MUS_1.CM.01.1085          CAATGGA-GGTC--T-----T-G----T--G-C---CA----GGAT-T--GTC---GT-T--C--CTTG-----ACG--AG--C.........-TTG--C-----GAG--T-----A---------T--CA-TGACG---CCATGTA--GAGAT................ 7547
MUS_1.CM.01.CM1239        CAATGGA-GGTC--T--G--T-G---GT--G-C-CGCAA---GGCT-T--GTC----T-C--C--CTTG-----ACGAC-G---.........--TG-GCAG---A-T--T---G-G--T-----C--TA-TGATG---C-ATGTA---AGAC................ 7519
MUS_2.CM.01.CM1246        CAATGGA-GGTC----------GG--GT--G---C-CA----GGTT-C---TC---GT-C--A--CCTG--C--TCGA-AA--T.........--TG--CA----A----CC--G-A--------AT--A--GATG-G-CCATGTACGGCAAT................ 7654
MUS_2.CM.01.CM2500        CAGTGGA-GGTT--T--G----G----T--G-T-C-CA----GGTT-TACGTCG--GT-C--C---CTG--C--TCG-CAA--C.........--T--CCAG---A-G--C---G-G--T--T--C--CA--GATG---CCATGTA-GGCGAC................ 7568
RCM.NG.x.NG411            --GTTTG-T--T--T-TC-CT---ATGG--G-T--TGAT---T-T-C---GTCT-G-AA-G-A--CA-T--A--A-GGCCA-AT.........GCCT-G-----CAACG-C-C-G-G---TACT--T--GC-GAT-TT--CTA-ACAG--AAT................ 7353
RCM.GA.x.GAB1             CCTTTTG-A-------TC-C----ATG--AG-C--TGAT---T-T-C----TCG-GGAAGG-A--CA-G--A--G-GGCCA-AT.........GCGT-A-AG--CAGTG--C-GG-AT-CTAT---T-GGC-GAC-T---CTA-....................... 7300
DEN.CD.x.CD1              -CATATA-GGTG--T-AC----------T-AG-G----CA-----C-C------C-TG-AT--T--CTTA--A--AGCAGAA---.........G------C----CA-C--C---G-GT----CA-TA-TG--GAC-CG----TCCA-CCGAAC................ 7649
                                                                                                                   V5 loop end
MAC.US.x.239              -ATTA-G-G-----TCAT--T-G------A--C----CT-----T--------C---AATG-T---TTG-----A-GAGAG---.........G-CC-C-CG---AAC--C-CAG-G--CA-T--CA--GC--AC-T----T-GAT-G--GGA................. 8028
Env                        N  Y  V  P  C  H  I  R  Q  I  I  N  T  W  H  K  V  G  K  N  V  Y  L  P  P  R  E  G  .  .  .  D  L  T  C  N  S  T  V  T  S  L  I  A  N  I  D  W  I  D  G  .  .  .  .  .  .  .
H2A.GW.x.ALI              -ATTATGCA--G---CAT-----G-----A--T--T--C-----C----C---G---AATG-A---TTG-----T-GGGAA---.........G-GT-G-CC--CAAC----CAG-A--CA-CT--A-TGCT-AC-TT--CACG-A-GGC................. 8070
H2A.DE.x.BEN              -A-TAT......---CAT-----G--G--A--T--T-CC-----T--------G---AATG-A---TTG-----T-GGGAA--G.........G-GT-GGCC---GA-----CAG-A--CA-CA-AA-TGCT-AC-TT--CATA-A---AAAT................ 8083
H2A.SN.x.ST               -ATTATG-G------CAT-----G-----A--T--T-CC-----C--G-----G---AATG-A---TTG-----T-GGGAA---.........--GT-A-CC--CAAC--T-CAG-G--CA-CA-AA-TGCT-AC-TT--C--A--AG-G.................. 7516
H2B.GH.86.D205            -ATTATGCG-----TCAT--C-G---G--A--C----C------C--G--T--G---AAT--A---TTG-----A-GAGAA--T.........G-GC-CTC----AAT--C-C-G-G--TA-C--CA--GCC-AC-TCA-CTCA-AC...................... 8059
H2B.CI.x.EHO              -ATTATGCGT-----CAC--C-G---G--AG-C----C------C--GA-T----G-AATG-----TTG-----A-GAGAG--T.........G--C-CTCC---AAT--C-C-G----CA-C--CA--GCC-AC-TT--CT-GATAG--AAG................ 8048
H2G.CI.92.ABT96           -ATTATG-A--C---CAT--T-GG--G--A--T----CA-----C-------------AATG----CTTG--C----GAGAA--T.........G-TC-G-CG---AAT----CAG-C--CA-C--AA--GCG-AT-T---A-AGAA--GGTCTAAC............. 7447
H2U.FR.96.12034           -A-TATG-G--T---CAT--T-GG--GG-A--T----CA-----T--------G-G-AATG-----TTG-----A-GAGAA---.........G-TC-C-CC--CAAT--T-CAG-A--TA-CA--A-TGC--AT-TT---C-GAT---CAAA................ 7570
MAC.US.x.EMBL_3           -ATTA-G-G--G--TCAT--T-G------A--C----CT-----T--------C---AATG-T---TTG-----A-GAGAG---.........G-CC-C-CG---AAC--C-CAG-G--CA-T--CA--GC--AC-T----T-GAC-G--GGA................ 7159
SMM.US.x.H9               -ATTATG-G-----TCAT--TRG---G--A--C----C------C--------C---AATG--R-TTG-----T-GGGAA---.........G-CC-G-C----AAC--C-C-G-A--TA-C--CA--GC-GA--T----T-GATC---GGC................ 7517
SMM.US.x.PGM53            -A-TATG-G-----TCAT--T-G---G--A--C----C------C-G------T--AATG-----TTG-----T-GGGAA---.........G-CC-G-C----AAT--C-C-G-A--TA-C--CA--GC-GAT-T---CT-GAT-G--GGC................ 7977
SMM.US.x.SIVsmH635F_L3    -ATTATG-G-----TCAT--T-G------A--C----C------C--------C---AATG-A---TTG-----T-GGGAA---.........G-CC-G-C----AAT--C-C-G-A--TA-T--CA--GC-GAG-T----T-GATC---AAC................ 8050
SMM.SL.92.SL92B           -ATTATG-A--T---CAC--C-GG-----A--T--T-C------C-GG---------AATG-C---TTG-----A-GGGAA---.........G-CT-A-CT---AAC--C-CAG-G--CA-C--AA--GC--AT-TT---T-GAT-G-CAAC................ 7479
STM.US.x.STM              -ATTATG-A------CAC--C-G----G-A--T--T-CA-----T--------G---AATG-T---TTG-----A-GACAA---.........G-CC---CT---AAT--C-C-G-G--CA-TA-AA--GC--AC-TT--CT-GAC---CAAT................ 7682
MNE.US.x.MNE027           -ATTA-G-G-----TCAT--T-G------A--C----CT-----T--------C---AATG-T---TTG-----A-GAGAG---.........G-CC-C-CG---AAC--C-CAG-A--CA-T--CA--GC--AC-T----T-GAT-G--GGA................ 7505
DRL.x.x.FAO               CCATATGCAT-C---CAC-----------A---G-TGAT---GCA-G------C-GGAAG--T---ATG-----TG-ATCA---TTTAATAATG----A------A--CA-G--G-C----AAA--T-C--TGAG-TCC-AAAATTGG-GACAGAG......GAAACAG- 7444
MND_2.x.x.5440            CCTTA-A-GT-TC---AT--T-G------AG--GGTGAT-----T--------C--GAA---T---ATG-----TG-GTC---GTTTAATAATG----A------A-TAATG--G-G----AAA--T-C--TGAGGTCC-AAAAAC-G--GATGAC.............. 7390
MND_2.CM.98.CM16          CCATGGA-GT----T-AT--T-G------A---G-TGAT-----T-----------GAAG--T---ATG-----TG-ATC----TTTAATAATG-G--A--------TAATG--G-A----AAA--T-C--TGAGGTTC-AAAGACAG-AGAGGGA.............. 7763
MND_2.GA.x.M14            CCATATA-GT----T-A---C-G---G--AG-GG-TGAT-----T-----G----GGAAG--T---ATG-----TG-GTC----TTTAATAAT--T--A-----CA-TAATG--G-C----A-A--T-T--TGAGGTCC-ACTA-T-G-AGGA................ 7686
OLC.CI.97.97CI12          CCATGGTCA-AG--T---C-G-G-----A-G--G-TGAT---AT-----------G-AG---T---CTT-----ACAACCA---TTCAATAATAG---CC-----..............---CAAAGAG--ACTGA--TGTT-TTG-AA-TACAAAAAGTAGAGAAC..... 7246
GRV.ET.x.GRI_677          -CTTATG-TG-C---CAT---CG----G-AG----TGAT---T-C-CT--CTCT---AAGG-A-----T--A--A-GGGAA--T...........--TT-GGAG---AAC---TCAG-C--G-CA--ATACG-GG---T----TA-AAC--CAAGTCT.............. 7623
VER.KE.x.TYO1_patent      -CGTATG-TG-T---CAT--C-GGTCTG-A-----TGAT---T-T-CCC--TC---GAA--CT-----T--G--A-GAGAA---...........--TT-GCA---CAGG--C-CAG-C--T---A--AC-G-TGAGCTTA--TA-AA--G-AAA................ 7737
VER.KE.x.9063             -CTTATG-TG-T---CAC--C-G-TCTG-C-----TGAT---T-T-CCA--TC---GAA-G-------T--A--A-GGGAA---...........--TC-GGA---CA----C-CGG-C--T--TA--AC-G--GA-CT-A--TA-AAC---AAC................ 7761
VER.KE.x.AGM155           -CTTATG-TG-C---CAT--CCG-TCTG-C-----TGAT---T-T-CTA--TC---GAA--CA-----T--A--A-GAGAA---...........--TT-GCAG--CA-G--C-CAG----T---A--AC-G--GAGCT-A-CTA-AA---CCAG................ 7741
VER.DE.x.AGM3             -CATATG-TG-T---CAT---CG-TCTG-C--T--TGAT---T-C-C-C--TC--GGAA--CC-----A--G--A-GAGAA--G...........--CT-GCA---CA----C-CGG-A--G--TA-C-G-GGAGCT-A--TACAA--G-AAG................ 7259
TAN.UG.x.TAN1             -CATATG----C---CAT--T-GGTC----G-C--TGAC---T-C-CGC-CA-T--GAA--CC-----A--A--A-GGGAA---...........--CT-AGA---CA------CAG-G---TCTA--A-GG--T--CTG---TA-AAC-GCAAA................ 7721
SAB.SN.x.SAB1C            -C-TATG-T--C---CAC----GG---G-AG----TGAT---T-C-CCT-GTCT---AAG-CA--C-----------GAGAA--C...........--CT-GGAG---AAC--G-CAGCC----C----TATG--GA-TTGA-CTA-AAC-G-AAA................ 7917
MND_1.GA.x.MNDGB1         CCTTATGC-T----TCA---T-GG-----AG--G-TGAC---AT-C---------G-AAG--A---TTA--A--T-CATCA---TTTAATAAT--C--A--G---A--CAT-GGG-A----AAA--TAC--TGA--TGC-AAAGATAG--AGTAATGAA........... 7211
DEB.CM.99.CM40            TGGTTGA-G--T--T--G-----T--GT-AG------CT---T-T-CG-----GC--CAT--C---CTA-----G-AAGAA---.........G-GT-G-AG--CAGC---C-C---T-T-CAT-TG-T--TGATGT----CACTA----GGG................ 7350
DEB.CM.99.CM5             -CTTTGA-G--T--T-AG--T--T---T-GG------CA---TTC--G------C--CATG-A---TTG-----A-AGGAA---.........G--T-G-AG---CAT--TTC-G--T-T-CA--TG-T--TGAGGT--C-CA-TA-GGGCAA................ 7271
TAL.CM.01.8023            T-TTT-A-GT-C---CAGT---GG-----G--C--T-G----AGC-------AG--GCT-------CTG--G--A-GAGAG---...........--CC-G--T--CAA-AGC--CG-C-----TA-C--GG-GGAC-TCA-CCACTA---GGGGAAT.............. 6990
TAL.CM.00.266             T--TA-A-GT-----CA-C-G-GG-----G--C----G----AGC-GG--G-AG--GTT------CCT---A----GGGAA---...........--CC-GCAG--CAA-AGT--C--C-----AT--T-GG-GGAT-TCA-CCACCACCGCGGCTCT.............. 7440
SUN.GA.98.L14             CCTTGGGC-AAT---CAC------AGTG-G---G-TGAT---GC-TC---G--C--GAAG--T--CTTG--C--A-CATCA--GTTTAATAACAG---A------G--AAC-GAG-G----AAGCATGG--C---CTG--AA-G-TGG-AGACTGG...AAGGTAAATG- 7929
WRC.CI.97.97CI14          CCATGGAGT-------C---T-G-A-T--AG--G-GGAC---GCAC-T--G----T-AA-G-A---TTG-----AG-ATCA--GTTCAATAAT--T--------C..............--TCAAAGAG--AC-GA--T-TTGTT-AACTGGGAA................ 7738
WRC.CI.98.98CI04          CCTTGGAGT-------CG--T-G-A-T--AG-GG-TGAT---GCAC-T------GT-AA-G-A---TTG-----AG-ATCA---TTCAGGAAT--T--A-------..............---CAAAGAG-GACTGA--TTCTATTCAA-TGGGAA................ 7783
WRC.GM.05.Pbt_05GM_X02    CCTTGGGCA--T----CT--C-G-TC---AG-GG-TGAT---GCATCTA----G--GAA---A--CTTG-----GG-GTC----TTTAACAAT--C----------..............--CCAAAG-G---C-GA-GT-CTCTT-AA-TGGGAA................ 6679
LST.CD.88.524             CC-TGGGC-AAT---CAC------TC-------G-TGAT---GCA-CT--T--G--GAA---T---CTT-----A-CATCA---TTTAACAATAG---A------A-TCAC-GAG-G----AAA--TGG--TGA--TG--AAAAT-GG-ACCACAT......GAAGACTT 6786
LST.CD.88.447             CC-TGGGC-AAT---CAC------TC-G-----G-TGAT---GCATCT-----G--GAA---T---CT------A-CATCA--CTTTAACAACAG---A------A-TCAT-GAG-A----AAA--TGG--TCA--TG--AAAGTT-G-ACCAGAA......GAAAACCT 6783
LST.KE.x.lho7             CCTTGGGC-AAC---CAC------AGC--C--TG-TGAT---GCT-C-A-------GAAG--T---CTA--A--T-CTTCA--TTTCAACAATAG---C------A-TCAT-GAG-A----AAA--T-C--TGAG-TG--AAAAT-GG-ACCACAT......GAGGACTT 7858
SYK.KE.x.SYK173           C-TCGT--GT----TGCC--TCGC-----A--T--TGAT---AGAT-T---A-----TTG--A--CCTG--C--T-CGGCA--G..........--T----A----A-T--C---G-C--T-CAG-AT-GACGGAC-T---ATACTACCCAGGA................ 7577
SYK.KE.x.KE51             T-ATAT---T-G--TGCC---CG---G---------TAT---GG-T-T---ACC--GCT-------TTG--A--T-CAGAA--G..........--C----AG--CA-T--T--------T-CTG-AT--AC-GAT-T---GTACTA-CCAAAT................ 7273
COL.CM.x.CGU1             CCAGT-A-GAAC--T-CC--T-G-A-GC-GG-G--T-GT---GTC-CTCAT-CC-GGATCC------GG--G--TCCAG-A--T..........--TC-ACA---AAC-GGG-G-AACAGCCAG-TA-TGCGTTC-TG-GCACCATAG-AGGA............GAC-A 7001
```

```
                                    V5 loop end                                                                 Rev Responsive Element (RRE) region
H1B.FR.83.HXB2          CAACAATGAGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATAT...............AAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAG...............A 7746
Env                      N  N  E  S  E  I  F  R  P  G  G  G  D  M  R  D  N  W  R  S  E  L  Y  K  Y  .  .  .  .  .  K  V  V  K  I  E  P  L  G  V  A  P  T  K  A  K  R  R  V  V  Q  .  .  .  .  .  .
H1A1.UG.85.U455         .--T-CAA-AAAT----C--------------------------------------------A-------------------G--...............---------------------------C-------------------G-------------------G--...............- 7176
H1B.US.90.WEAU160       AG-G--CC--A--------------------------A------A--T-----------------------------...............-------------------------------------------------------------------...............- 7750
H1C.ET.86.ETH2220       AC-T-GCACAAAA-----A----------A-------------------------------------------------...............-----G---G-----A-G---C-----------------T---C----A--G---------G--...............- 7125
H1D.CD.84.84ZR085       T--T-C-C--AAT--C-C---T------------------------A-----C----------------------------...............----------------------C---------------------------------------G-A...............- 7267
H1F1.BE.93.VI850        .G-G-G-A-CATT----C----------AG------A------AA-----------------C-------G---...............---------G-----------------------------------A---CA-----------...............- 7023
H1G.SE.93.SE6165        T-CA-G----AGT------------------------------------------------------G---...............---AC---------CA--T--C----------------G-----G-------------G--...............- 7164
H1H.CF.90.056           GTCTGCA--AAATT-T-C--------------------------------------G-----------...............--------------------C-G---A------------A---G-------------G--...............- 7065
H1J.SE.93.SE7887        TG-G----GCA-T----C----------AC-------A------A------------------------------...............--------G-G-----G--C-------------------------A-------------G--...............- 7053
H1K.CM.96.MP535         T--T-C-C-TAAT----C--------------------C-----------------------C-G-----------...............--------C-G---------C-----A------------G---C-G------------------...............- 6906
H101_AE.TH.90.CM240     .--T-CG--TAA-----C--------------------A-C--A-A------------------------------...............--------C-----------C-----A------------G-----------------G--...............- 7299
H102_AG.NG.x.IBNG       .--T-G-ACAAAT--A-C------G-------------------------------------G---...............----------------------C----T------------CGT-----A-------------G--...............- 7252
H1N.CM.95.YBF30         .G---CCA--GAAACA--AG-ATAT--CTC--------A-C---GTTA-TCTC------CAA--G--G-----G--C...............--------T-GC--A-----CA-------------AGGT--A--T--A---C-CACA---AGT...............- 7291
H1O.BE.87.ANT70         .-G---CA-CAAT-TA-CA--T-----AAT---G-----C----AA--T-TA-------C-------G-TC--C--CAAAGTAGTAAGGGTA---CCTT-T-GTG-G-C---CAC-C-TA-T---AGGC-AGTCAT--GC-CT---ACTCAT..................- 7800
H1O.CM.91.MVP5180       TTC--CA-GTGAAA-T-CAC-T-----A-T---G----------AA--T-TA-------C-A----G--C--C--C...............---------C-G--AA----T--TA-T---------T--A--AATGTCA---CC-A-AA-AA-CATTCACACCCCTCAC- 7817
CPZ.CD.90.ANT           ....-----AA-T-TT-AAG-GTC-----CT-C-A---TAGCA...---C-A------GCG--G--G-CC-GG--C...............C-G--G---G-G----RT--C--GTC------C--A--A-CA-GN--A--GCC--AAA-AA-ACAACACTCC.......- 7178
CPZ.CM.01.SIVcpzCAM13   .--T---ACTAATA-T-CA--T-----AAT---------C----AAA---T------G------C-C---------C...............------G--GCGC--A---------TCT--G-----A-----A-----A--G...CATACAGTAGGA......GAAAGG- 7327
CPZ.CM.05.SIVcpzEK505   TGGA-C----AAGT-C-C-A-ATAT--CAC---G-----C----CTA--CTC-----GCAG--G--G-----G---...............--G--G--C-GC--A-----CA-------------AGGA--A--T--A---CACACA---ACT...............- 7226
CPZ.CM.05.SIVcpzMB66    G-TG--CA-A--A-CC--G----C-----TT--G-----A-----AA-T--------------------C------...............-----------G----------C--------------------------A------AC------T...............- 7367
CPZ.CM.98.CAM3          --G--TCACCAA-ATC-CGC-ATAC---AC-------A-C---GTA---TTG-----GCT------GC-------C...............--G-----T-GC--A-----CA-------------A-GT--A--T--A--GCACACA--ATCA...............- 7156
CPZ.GA.88.GAB2          --CGGGGA-CAA-ACAGAAG-TTAC-TGTC---G-------------A--T-T------G-------------------...............---A-------------G---C-G--G--G-----A--A--A--------GTACACA---GCAAAAGCA...CTTGAT- 7127
CPZ.CM.98.CAM5          -......A-TA-TATA-CAG-GTAC---TC---G-----A-----A----TTG-----GCT-------A-C-GG---...............-------T-GC--A-----CA-------------A--T-----T--A--GCACAC----GCA...............- 7462
CPZ.CM.05.SIVcpzMT145   GG-TC-GA-TAA-ACATA-G-TT-T-TAAC---G--TA-C----AA--T-T-------------C-----C-----C...............---A---TG-G--AC-----C---------------T--T--AAGT-GA---TAT-CA--AG-AAAACAACACCAT...- 7255
CPZ.TZ.00.TAN1          ..........AATAGTTATAC-CCC-AATTTTCT-CT-TAG-AGAA---T------A-GTA------GCA-G----...............-----G--GG-----C-G--C--GTC---G--T--A-GGCCA-G---A--GCCT-AAA-TA--GCCAATCAT...ACT- 7372
CPZ.CM.05.SIVcpzLB7     A----CA--CAATATC-C----C-T--AAT------G-------CA----TC----------------------C---...............---A-------------------------------G--T----G---G-GA--G--GACT---G-T...............- 7257
CPZ.US.85.CPZUS         -G-T-G-AG-AATATA-CAG-G-TG---AC------------------C--CTTTA---A--A-------C----G---...............--G--G--C-GC--A-----TA-------G-----AGGT-----C--A--GCATACA---A-A...............- 7748
CPZ.GA.88.GAB1          -----G--GTAA-CT--CA--T------AC-------A-C-----AA--T-T-------------GC-C-----G---...............--------TCG---A--G--T---TC---G-----A--A--A----GA--GCATACA---GCAAGACAG...AAAGAC- 7780
MON.CM.99.L1            .............A-CG-GAC-CTTGT-CCCTCG-C-C-GG----T---TCA------TCA--GC-CAG-CG----...............--G--G--GG-------C--C---TCTA-G-----T--A-CA--GC-------GACT-G-GTA............CAC- 7709
MON.NG.x.NG1            .............ACAG--ACGTTGAT-CC-TC--C-CGAG--TCT---GCA---C-CTCA--G--GAGC--G--C...............--G--G--CG-G--C--C--C--GTCT--G-----T--AGCA--CC-A--G---CAT-G--TA...............- 6331
GSN.CM.99.CN166         .............TCT---AA--TGA-CCCGTCG-CC---G--GCT--TGCG---ACTA------G-CC-GG--C...............-----G--GG-G--A--T--GC-CTCGA-G--C--A--TCC---T--A--G-A--AACAC-CC............GCTG 7683
GSN.CM.99.CN71          .............TCC---AAT-TGA-CCC-TCG-CC---G-AGCC--TGCA---AGTA---G-----CC-GG--C...............--G--G--GG-G--A--T---C-TTCGC-G--C--A---TC---TC-A--G---CCACAT-CAGGGGTG......CATG 7695
MUS_1.CM.01.1085        .............TCA---AAT-TGA--CC-TC--C-A--G--TT----GCC---C---TG--GCCTA-G--GGTACAA...............--G--G--GG-G--C--T---C-TTCCA-G--------GG-A-TGC-A-----GCCT--C-CA............CATG 7680
MUS_1.CM.01.CM1239      .............ACAG--AA--TGA--CC-TCC-CGA--G---T----GCC---C---TG---C--A-A--T--C...............--G-----GG----A--C--GC-TTCCA-G--T-----AG---TGC----G--GCCT--T-CC............CATG 7649
MUS_2.CM.01.CM1246      .............ACA--AAA--TGA-CCCCTC--CCA--G-C............----CGC---CG-CTAGACTGAGAGAT...............-C---G---G-G--A--C--TC-TTCC--G-----G--AG-A-TGCGC--G---CCT----CC............CACG 7778
MUS_2.CM.01.CM2500      ..........A-T-T--ATCGATTGA-CCC-TC--CCA--G--CTT---GC----C-C-TG------A-G-G---C...............-----G--GG----A--T---C-GTCC--G-----A--GG-A-TGC-------GCCT----C-............CATG 7701
RCM.NG.x.NG411          ..........ATGACA-CAAATGTCA-CTTGTCT-CT--AG-AG-A--T-TA---GC-GCA-----GGGC-G---C...............--------GG-G---ATG---A-T---TAT-----A--A--T-TC-GA--GTAC-A-ACA..................- 7480
RCM.GA.x.GAB1           .-TT-----CAGT----CAAA--TCA-CCTCTC--CG--AG--G-T--TT-----GC-GCA-----GGGG-G----...............--G-C-A--G----CAG----A-T--CT-T-----A--AG-AAT----A--GTACCA-AC-...................- 7436
DEN.CD.x.CD1            .......AGAA-AC-T--AACATTGA-GCTCTC--CT---G-----C-AGTC------GC------GGCA-G--GG...............---T-----G-GT-AAC----A-T--CC-T--T--T--TTCA-TG-GAC--TAT-AA-GA-CAGAATCGGCAGAGTCTG 7797
                                                                                                                   Rev Responsive Element (RRE) start
MAC.US.x.239            ..........AA-C-A-CTAAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG-----GGGAG-T---...............---T-----G-G--CACT---A-T--CT-G--C-----AG-T-TG-----GTACACTACTGGTGGC......ACCTCA- 8167
Env                      .  .  .  .  N  Q  T  N  I  T  M  S  A  E  V  A  E  L  Y  R  L  E  L  G  D  Y  .  .  .  .  .  K  L  V  E  I  T  P  I  G  L  A  P  T  D  V  K  R  Y  T  T  G  G  .  .  .  T  S
H2A.GW.x.ALI            ..........AA-C---CAAAT-TTA-CTTTA-T-C---GG--GCA--ACTA-ACC--TTA-----GGGGG-T---...............---T-----G-G--AAC----A-T--CT-C-----T--ATCA-A--G----TACTCCTCTACT.........CCAAGG- 8206
H2A.DE.x.BEN            .......CG-A-TC-T-C-AA--TTA-CTTTA-T-C---AG--GCA--ACTG-ACC--TTA---C-GGGAG-C--C...............---T--A--G----AAC----A-T--CT-C-----T--AG-TCAG-GA--GTACTCCTCAACT.........CCAGTG- 8222
H2A.SN.x.ST             ..........AA-C---CAAAT-TTA-CTTTA-T-C---GG--GCA--ACTA-ACC--TTA-----GGGGG-T---...............---T-GA--G--G-AAC----A-T--CT-T-----T--ACCA-T---A---TACTCCTCTGCT.........CCAGTG- 7652
H2B.GH.86.D205          .......A-CAGTACC-C-AAT-TTT---TGA-T-C---AG--TCA--ACTG-AT---TTG---C-GGGGG-C---...............---T-----G----AAC----A-T--CT-C-----T--AG-T-T--GA---TACTCCTCAGTA.........AAACCG- 8198
H2B.CI.x.EHO            ..........AA-CTT-CTAAT-TTA---TGA-T-C---AG--TCA--ACTG-AT-A-TTG---C-GGGGG-C--C...............---T-----G-G--AAC----A-T--CT-T-----T--A-GTAT---A---TATTCCTCAGT-.........ACACCG- 8184
H2G.CI.92.ABT96         ..........AA-C-T-CAAAT-TCAT-TTCA-T-C---GG--GCA--ACTG-ATC--TTA-----GGGGG-T---...............---T-----G----AACT---A-T--CT-T--------A--T-T-------TATTCCTCTGT-.........ACACCG- 7583
H2U.FR.96.12034         ..........AATC-A-CAAAT-TCT-CATGA-C-C---AG--GCA--ACTG-AT---TTA--G--GGGGG-T---...............---T-G---G----AAG----A-T--CT-T--T-----ACCA-TG--A---TATTCCTCAGT-.........ACACCG- 7706
MAC.US.x.EMBL_3         ..........AA-C-A-CTAGT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTA--G--GGGAG-T---...............---T-----G-G--CACT--GA-T--CT-G--C-----AG-T-TG-----GTACACTACTGGTGGC......ACCTCA- 7298
SMM.US.x.H9             ..........AATA---C-AAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG--G--GGGAG-T--C...............---T-----G-G---ACT---A-T--CT-C--C-----A-GT-T---A--GTACACCACAACTGGT......GCCTCA- 7656
SMM.US.x.PGM53          ..........AA-C---CTAA--TCA-CATGA-T-C---GG--GCA--ACTG-AC---TTG--G--GGGAG-T---...............---T-----G-G---ACT---A-T--CT-G--C-----A-GT-T--G---GTACACAACTGGTGCC.........TCA- 8113
SMM.US.x.SIVsmH635F_L3  ..........AAT----C-AAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG--G--GGGAG-T--C...............---T-----G-G---ACT---A-T--CT-G--C-----AG-T-T--GA--GTACACCACAACTGGT......GCCTCA- 8189
SMM.SL.92.SL92B         ..........AAT----C-AAT-TTA-CATGA-T-C---AG--GCA--ATTG-ATC--TTG-----GGGGG-T---...............---T-G---G-G--CACT---A-T--CA-G--T-----AC-T-TG--A--GTACACCACAAGT.........ACCTCA- 7615
STM.US.x.STM            ..........AAT----CTAA--TCA---C-A-T-C---GG--GCA--ACTG-ATC--TTG-----GGGAG-T--C...............---T-----G----AACT---A-T--CT-G--C--A--G--T-TG------TACACAACTAGT.........ACCTCA- 7818
MNE.US.x.MNE027         ..........AA-C-A-CTAAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG-----GGGAG-T---...............---T-----G-G--CACT---A-T--CT-G--C-----A--T-TG-----GTACACTACTGGTGGC......ACCCCA- 7644
DRL.x.x.FAO             AG-TTCCA-C-ATC----TAAATTTATACCGCA--AT--AG--CAAA-TC-G-TC-C-GCA-T-GG-GCAC-T--C...............---T-G--C--GG-G--G---A-T--CT-T--T--A--TG-T-T-C-T---TACAACC---C-GGA.........CAC- 7590
MND_2.x.x.5440          ....----GA-ATATC--AAAATTTATACC-CA--ATTGG--TCAAA-TC-A-AC-C-GCA-TGGGGGC-C-T--C...............---T-G--G--GG-G--T--GA----TT-T--C-----AG-CAT-C-C---CATCATT---CTAAT.........ACA- 7532
MND_2.CM.98.CM16        ...........-A-ATC---AAATTTGTGCC-CA--AT--GG-ACAAA--CGA-TT-C-GCG-T-GG-GCAC-C---...............C--C---------G-G--C---A-T--CT-T--T--T--T--AG-A-TGGCA---TATCATC---CAGAA.........GCT- 7899
MND_2.GA.x.M14          ....--AAGA-A-CTA--AAAGTTT-TGCC-CA--AT--GG-ACAAA-TC-A-TC-C-GCA-TGGG-GCAC-T---...............---T-G------G-G--T--GA-----T-C--C--T--TG-T-T-C-C---TATCATC-A-CAGAT.........GCA- 7828
OLC.CI.97.97CI12        ........-GAA-GC-AG-G-CATTG-T-CC-CC-TC---C-ACGT-A-TC-A-TTGTCTCA-TGGC-AG-G-T---...............---T-G---G-GT-G---TTGA-T--CT-T--T--T--AG--A--C-A--GTAT...................CAGCCT- 7376
GRV.ET.x.GRI_677        ..........GG-CCA--AAATGTGA-CCT-A-TCCTC-GG-AC-CAG--TA---GCGTAC---C-GGGAG-C---...............---T-----G-G--AAC----A-T--CT-T--T--T--AG-T-T--GA---TATACT-GC-CCACA.............- 7756
VER.KE.x.TYO1_patent    ..........AA-AGA-C-AA-GTGA-ACT-A-TCCCC-G--AGAATCT-TC---GCGGC------GGGC-G---C...............---T----GG----CAC----A-T--CT-T---------AG-A-T--G-C-TTATAC--GAGGA.........CATGAG- 7873
VER.KE.x.9063           ..........-ATAGA-C-AATGTGA-ATT-A-TCCCC-G--AGAA-GA-TC---GC-GCA-----GGGACG----...............---T-G---G----CAC----A-C---T-T--------G--T-T--GAC-TTATAC--GAGGA.........CAAGAC- 7897
VER.KE.x.AGM155         ..........AA-AG--CAAATGT-A-ATTGA-TCCCC-G--AGAAAC--TC---GCGGCA-----GGGC-G---C...............---T-G---G-G---AC----A-T---T-T--------AG-A-TC-G-C--TACAC--GAGGC.........CAAGAG- 7877
VER.DE.x.AGM3           ..........AA-AG--CTAATGT-A-ATT-A-TCCCC-G--AGAAAC--TC---GC-GCA-----GGGC-GG--C...............---T-----G-----AC----A-T--CT-C--------AG-A-T--GA--GTATAC--GAGGT.........CATGAC- 7395
TAN.UG.x.TAN1           ..........GAAAG--CTAATGTGA-ATTGAC--CTA--T--GA-A---TA---GCTTA------GGGA-G----...............--GC-CA--G-----------A-C--TT-C--G--A--AG--AT--GA---TAT--T-G--CAACT............C 7854
SAB.SN.x.SAB1C          ..........AA-AG--C-AATGTGA-GTT-TC-CCTC-G--TGAAAG--T----GCT-A------GGGAG-T--C...............---T----GG----CA-G---A-T---T-T--T--T--A-----A-T--GA---TACACT-GC............CCCGAA- 8050
MND_1.GA.x.MNDGB1       ....-CAA-AATGC-A--TAAATTCTTGCCTCCCA-T--A-CCTCCA-TC-A-TTGTTGC-T-TGG-GC-C-T---...............---T----C-----AATG---A-T--CA-------T--AG-T-TG--A---CACACTT-A-CTGAA......CATCAT- 7356
DEB.CM.99.CM40          .............AGT--AAC-CTTA-ACC-TC--C---C--C-----CAGTA------GCA--CC---T-------...............---A-CA--G--G-AA----TA-----T-C--T----T--GCA-TG-GA--GTAC-A--GA-C-GAG......TCGGTA- 7486
DEB.CM.99.CM5           .............A-C--GAA--T-A-ACCCTCCTC------T-A--CAGT------GGCA--TC----C--G---...............--GA-CA--G--G--A----CA-C--GA-G--C--T--A-GTAT--GA--GTAC-A--GA-C-GAG......TCTGCA- 7407
TAL.CM.01.8023          ..........GA-ATC--AAGTGTGGAACCCTCTTC-C-A--C--A--TGC----C--GCA-CGG--GCACGG--C...............--G--GA--G-G--CAC----A-T--CT-T-----A--G--TATC-GA--GTAT-AT-G--C-..........GCAACA- 7126
TAL.CM.00.266           ..........AGT--C--AACAGT-GAGCCCTCGTC-C-G--AC-T---TCG---C--GCA-C-G-T-CACGG--C...............--G--GA--G-G--C-TG--T--T--GTAT--G--G--A--TATC-GA--GTAC-AT-G--CC.........GCTAAC- 7576
SUN.GA.98.L14           ATCA--CATC--A-TTG-AGC-TTC-AGCC-CCTACCA-C-CCTTAA-TC-A-TTGTG--CAC-GG-GCAC-T--C...............---C------G----AGG---A-T--GT-T--T--A--AG-T-AGC-CC-GTAC-CACCC................AGG- 8069
WRC.CI.97.97CI14        .......AGAAATC-AG-GA-GTTC-AA-CTCCTTCCA-C--ACA-A--TC--TTGTGGCA-TGGG-C-A-------...............---T-G------T-AA--ATGA----GA-G-----AGGGG-T-G------GGTT--ACCC-CTGAA......CATCGC- 7880
WRC.CI.98.98CI04        .......A-TAAGC-AG-GGAATTC-AA-CTCCTTCCA-C--ACA-A-TTCA-TTGTGGCA-TGGG-C-A--G----...............---C-G--G-G-C-AA--ATGA----GC-G--T--TGGTG-T-GG-GA---GTGACACCAA-AGAA......CATCAT- 7925
WRC.GM.05.Pbt_05GM_X02  .......A-CGGGC--G-GGAGTTC-AA-CTCCTTCCC-C---CAAA--TCC-TTGTGGCA-T-GG-C-A-------...............---C-G------C-AA--TTGA----------T-----TG-T-G------GGTGACACCT-CTGAC......CATCAT- 6821
LST.CD.88.524           AGGAGGAA-TCTGTCT--TAAGTTTTTGCCTCCCTC-TGGGA--CAA-TC-A-TTGTGGCAAC-GG-GC-C-T--C...............---T-GA-C-G-C-GAG----A-T--CT-T--T--A--ATCA-AGC-TC-GTAC-CACCC............AGGGGA- 6929
LST.CD.88.447           GGGAGGAA-T-TGTCTG--AGATTCTTGCCTCCCTCCTGGGAA-CAA-TC-G-TTGTGGCAAC-GG-GC-C-C---...............---T--A---GGC--AG----A-T--CT-T--T--A--AG-T-TGC-TC-GTAC-CACCC............AGTAGT- 6926
LST.KE.x.lho7           AGGAGGCA-CCTAAGT--AAAGTTCTTGCCTCC-TC-TGGGAA-CTA-TC-A-TTGTGGCA---GGG-C--------...............---T-GA-----T-GA-T---A-T--GT-T--T--T--AG-T-AGC-C--GTAC-CACCC............AGGGGA- 8001
SYK.KE.x.SYK173         ...........--TAC-T-GAA-TTTA--CC-AC--C-A--G--GA---TGT-------GCA--TC-G-TC--T---...............---T-GA-TC-----A-----A-T--CT-T--G-----TG-TCA-CGTC-GTAT-A-T---CT.........AATACC- 7713
SYK.KE.x.KE51           ...........--AC--T-GAATTTTG-ACCTACT-CTA--G--GAA--TGTG------GC---T----TC--T---...............---T-GA-TCGG---A----TA-T--CT-T--G--T--A-GTCA--G-C-CTAT-A-C---CT............ACG- 7406
COL.CM.x.CGU1           TG-TGG-A-TGGGTGTGCT-ATCC-G-A-C-CC-AATTT--A-CAT-CATTAA-T-CTTTG-----GGGA-G---C...............---T-G-------GAGGA--ACTACTTAT-TC--A--AG-TAT---A---TC----AATGTAAATTGGCATCATGGG- 7156
```

```
                             Env gp120 end_Env gp41 start
H1B.FR.83.HXB2               GAGAAAAAAGAGCAGTG......GGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTG...AGGGCTATTGAGGCG 7907
Env                          R  E  K  R  A  V  .  .  G  I  G  A  L  F  L  G  F  L  G  A  A  G  S  T  M  G  A  A  S  M  T  L  T  V  Q  A  R  Q  L  L  S  G  I  V  Q  Q  Q  N  N  L  L  .  R  A  I  E  A
H1A1.UG.85.U455              ---------------T.........C--------A-C-------------A-------T-------A-------G--G---A----------------------------------------A-------G---C-----...---------A----T 7337
H1B.US.90.WEAU160            ----------------GGA...ATGC--------A-------------------------------------------G---------------------------T--------------------A-----------C-----...----------------- 7914
H1C.ET.86.ETH2220            --------------CA.........C--------------------------------------------------A-----A------------------------------G-----------A-----A-G------------...-AA-----A------ 7283
H1D.CD.84.84ZR085            -------------A-A.........C--------A-------------------------------G-----------G-----------------------------G-----------------------A-----A--------------...---------A------ 7428
H1F1.BE.93.VI850             --------------C-.........C-----------------------TC-T--G-ACAGCA-GGAGCA---------G--G--G--A---------------------------------------A-------A-----A--------------...------------A--- 7184
H1G.SE.93.SE6165             ---------------T.........C-G------G-C-------------A-----------------------G--G----A-----------------T-------------------C--------A-----AGG---C------...---------A------ 7325
H1H.CF.90.056                ----------------.........-----G-------CT---------------------------------G--G----A---------------------G-------------------------A-G-----------...--A-----AC----- 7226
H1J.SE.93.SE7887             ----------------..........-------------G-----------------A----------T---------------G----A-----------------T--------G--------C--------A-----A-G-----------...-A------A------ 7214
H1K.CM.96.MP535              ---------------A.........T--------G-----T----A--------------------------------G--G----A------------------------------------------A-----A-G-----------...--A-----A--A--- 7067
H101_AE.TH.90.CM240          ---------G------..........-------------A--A--T----------A-------------------G--G----A------------------------------------------A-----A-G-----------...----------A------ 7460
H102_AG.NG.x.IBNG            -----------G--T.........C-G------G-C-------------A-----------------------GCGG-----A-----------------------------C--C--------C-------A-------G-------------...-AA-----A-----T 7413
H1N.CM.95.YBF30              --------------CCTTT.......C----T--GC----T---------TC-T-----------G---------------G-----A--------------------C-GAC------A-----G--------A--------T-T-C--T--...--A--A--A------ 7455
H1O.BE.87.ANT70              ---------------A.........T-G---ATGC-A---T-G---G-TC-AA-T--------T-----------------GG---CA------G---------A--CACACT--GC--AAG-----------A------G----CC----A...--A--A--AC----C 7961
H1O.CM.91.MVP5180            -G-------------A.........T-G---ATGC-A---T-G---G-GC-AA-T--------T-----------------GG---CAG-----------G-A--CACAGTG--C--AAG-----------A------G----CC-----...--A--G--AC----C 7978
CPZ.CD.90.ANT                --C---------GCA-T..............GCTG--C---T-G--TC-TC-CA-T-----T--C--T--A----------G-----AG--------C----A----GA-T--GY-CCA------T--A--A------GC----C------...CAA--C--A---A-A 7339
CPZ.CM.01.SIVcpzCAM13        --C-G--G-------CCTTT.....CC-G-----AC--------------TC----T----------------A-----------GG--G-C------------------C----G---C-A--G--A--------A--A--------CC--T--...--A--A--A-----C 7491
CPZ.CM.05.SIVcpzEK505        --C------------CCTTT......C----T--GC----T---------TC-T-----------G-----------------G-----A-----------------C-GA-------A-----G----------------T---C------...--A--A--A------ 7390
CPZ.CM.05.SIVcpzMB66         -G-------------CCGTA......T-G--------------------------------------------------G---C--------------------------C-GC----------------------G----C------...--A--C--A------ 7531
CPZ.CM.98.CAM3               ---------------CCTTT......C-G--T--GC----T-----------A----------------------------A---G--GT---------------C-G-----------A--G--------A-----------CC-----...C-A--A--A------ 7320
CPZ.GA.88.GAB2               --TCC----------CCTTT......C-G--G--GC----T-----A---C-T----------------------------A--TG---T-----------A--TC--A-CC-GC-T--G--A-------------A--T--C-------...C-A--C--A--A--T 7291
CPZ.CM.98.CAM5               ----G----------CCTTT......C-------AC-------C-----TC-A-----------------------------A---G--GT----------------C-G--C--G----A--G--------A-----------CC-----...C-A--A--A------ 7626
CPZ.CM.05.SIVcpzMT145        -------G-----CC-T..........C-------AC----T--------C-------------------------------A---G-CGT-----------A---C-G-----G--AA-G--C------------------T---C----A...--A--A--A--A--T 7416
CPZ.TZ.00.TAN1               -GTC--G-----AT......GTG--C------CTG-----T-----A--TC-TA-T-----------T--A-----------G-----AG--------C----------GG---C-C--------T--A--------AC-A--CC----T...CA---C--A--A--- 7533
CPZ.CM.05.SIVcpzLB7          ---------------CA.........GC-G----TG--T------------------------G-----------------A---T-A---------------------------A-----C--------A-------GT-----------...--A-----A------ 7418
CPZ.US.85.CPZUS              ---------------CCTTC......C----T--AC-------G-----TC-T-----------------------------A---G-AGT-----------------C-------------A--G--------A-----------T---C----A...--A--A--A-----C 7912
CPZ.GA.88.GAB1               --C-G----------CCTTC.....TC-G-----AC--------------C----T-------------------------GG--G-A------------------C----G--GC----G--------------A----T---C--T--...-AA--A--A-----T 7944
MON.CM.99.L1                 -G-----G-----......ATTACTT-G--TATGGCC--T--G--C---C-CA-CA-G------G----A----------GG-G-CAG-C-----------T-TC-GTC----C--G-----C--G----G---------------G-G--CC-----...-----CG-AACA--C 7870
MON.NG.x.NG1                 TGAG-G-----CG--CT...ATCAC-T-G---ATGGCC--T--G--C--TC-CA-CA-G------G----G----------GG-G-CAG-C-----------T--C-GTCT--G---G-----G---G----G---------------G-G--CC-----...-----CG-AACT--C 6495
GSN.CM.99.CN166              TG--G--------C.....CTTTC-T-G---ATA-CC---T-GAC---TC-CA-T--G-----T-C---G----------GG---CAG---------A----T--C-GTCT--G---G-----G---G----G---A-----------G-G--CC-----...-------G-A-----C 7844
GSN.CM.99.CN71               TGA----G--G--C.....ATTTC-T-G--CATA-CC---T-GAC---TC-CA-T--G-----TGC---G----------GG---CAG-------------AT--C-GTC---G---G-----G---G----G---------------G-G--CC-----...-------G-A-----C 7856
MUS_1.CM.01.1085             TGC-TGC------C.....ATCTC-T-G--CATAACC--TT-A--A--TC-CA-C----------G----A----------GG-G-CAG-C------------T-TC-GTCT--G---G-C---G---G-C--G---------------GCA---C-----...--A--CG-A--A--C 7841
MUS_1.CM.01.CM1239           T-A--GC------------.....ATATCTC----CATAACC--T--G--A--TC-CA-C--G------G----G----------GG-G-CAG-C------------T-TC-GTC----C--G-C---C--G-C--G---------------GCA---C-----...-----CG-A--A--C 7810
MUS_2.CM.01.CM1246           TCA-GGCT--G---......ATCA-CT-G---ATAACC---T-A--A---C-CA-T--T--G---G----A----------GG-G-CAG-C------------T-TC-GTCT--G---G-----G---G----G------------A---GCT---C--T--...--A--CG-A--A--C 7939
MUS_2.CM.01.CM2500           TCC--GCT--G---......ATAACCC-----ATAACC--TT-G--A--TC-CA-T--T------G----A----------GG-G-CAG-------------T--C-GTCT--G---G-----G---G----G------------A---GCT---C--T--...-----CG-A--AA-C 7862
RCM.NG.x.NG411               A-C----G----T-CCT...TTG-TGC-G--T.......-T--A--T---C-CTCG--------T-CTG-A-----------GG-G-CAG-C-----A--C---T--C-G--T--GC-TG----G---T-------------G--G-------...GACATAG----ACA- 7638
RCM.GA.x.GAB1                A-C-G-----G-T-CCT...TTG-TGC-G--T.......-T--A--T---C-CTC---------T-CTG-A-----------GG-G-CAG-C-----T--C---T--C-G--T--GC-TG-A--G---T----------A--G--CC------...GACATAG--A--CG- 7594
DEN.CD.x.CD1                 TTTC-CGCC--AGGCG-GAAGTGTCTT---TGCT-GG-C--A-A--A--TC-CA-C----------CTGTC--------G-T-GG--CAGGCT-AG--------T--C-GG-T---C-TG-C--G-----------A---C-AG-GC--T--...--A---G----A-GT 7964
                             Env gp120 end_Env gp41 start
MAC.US.x.239                 --A-T-------GG--C...TTT-TGC----G.......---T-G--T--TC-C-C-A-G----TTCTG-A--------G--G--GT----------C-CT---T--C--ACT------G----G-----------A---C-AC-GC--T--...GAC-TGG-CA--AGA 8325
Env                          R  N  K  R  G  V  .  F  V  L  G  .  .  F  L  G  F  L  A  T  A  G  S  A  M  G  A  A  S  L  T  L  T  A  Q  S  R  T  L  L  A  G  I  V  Q  Q  Q  Q  Q  L  L  .  D  V  V  K  R
H2A.GW.x.ALI                 -GA-T-------GT---...TTC-TGC----G.......---T-A--T--TC-C-CGA-------TTCTG-A------A-G--AG-TT-A------T-T-CT---T-TC-GACT------G-C--G-----------A---C-AC-GC--T--...GAC-TGG-CA--AGA 8364
H2A.DE.x.BEN                 -GA-C-------GT---...TTC-TGC----G.......---T-G--T--TC-C-CGA-------TTCTG-A--------GCGG--CC---------T-A-CC---T--C-GACT---C--G-C--G-----------A---C-AC-GC--T--...GAC-TAG-CA--AGA 8380
H2A.SN.x.ST                  -GA-T-------GT--A...TTC-TGC----G.......---T-A--T--TC-CACGA--------GCTG-A--------G--G--CT---------T---CT---T-TC-GACT------G-C--G-----------A---C-AC-GC--T--...GAC-TGG-CA--AGA 7810
H2B.GH.86.D205               -GA-T-------GT--A...ATG-TGC-----.......---T-G--A---C-T-C-ATG-----TTCTG-A----------A-G--TT---------T-A-CT---T--C-GACT---C--G-C--G---------------C-AC-GCCTG--...GAC-TGG-CA-AAGA 8356
H2B.CI.x.EHO                 -GA-T-------...-GTGTACTT-TGC----G.......--TT-G--A---C-T-CGA-G----TTCTG-A--------G--G--CT-----T--T---CT---T--C-GACT---C--G----G---------------C-AC-GC-CG--...GAC-TGG-CA-AAGA 8342
H2G.CI.92.ABT96              A-A-T-------GT--A...TTT-TGC----G.......--TT-G--A--TC-T-C-ATG-----TTCTG-A------A----G--GT---------T---CT---T--C-GACT--G---G----G-----------A---C-AC-GC--T--...GAT-TGG-CA--AGA 7741
H2U.FR.96.12034              -GA-T-----G-G---A...TTT-TGC-T--G.......---T-G--T--TC-C-CCA-G-----TTCTG-A-----T--G--G--GT---------T-A-CT---T--C-GACT---C--G----G-----------A---C-AC-GC--T--...GAC-TGG-CA-AAGA 7864
MAC.US.x.EMBL_3              --A-T-------GG--C...TTT-TGC----G.......---T-G--T--TC-C-C-A-G-----TTCTG-A--------G--G--GT-C-G------C-CT---T--C-GACT------G----G-----------A---C-AC-GC--T--...G-C-TGG-CA--AGA 7456
SMM.US.x.H9                  --A-T-------GG--C...TTT-TGC----G.......---T-G--T--TC-C-CGA-------TTCTG-A--------G--G--GT---------T---CT---T--C-GACT--G---G----G-----------A---C-AC-GC--T--...GAT-TGG-CA--AGA 7814
SMM.US.x.PGM53               --A-T--G----GG--C...TTT-TGC----G.......---T-G--T--TC-C-CGA-------TTCTG-A--------G--G--CG----T-G-T-C-CT---T--C-GACT--G---G----G-----------A---C-AC-GC--T--...GAT-TGG-CA-AAGA 8271
SMM.US.x.SIVsmH635F_L3       --A-T--G----GG--C...TTT-TGC----G.......---T-G--T--TC-C-CGA-------TTCTG-A--------G--G--GC---------T---CT---T--C-GACT--G---G----G-----------A---C-AC-GC--T--...GAT-TGG-CA--AGA 8347
SMM.SL.92.SL92B              AGA-T-------G---C...TTT-TGC----G.......---T-G--T--TC-C-CGA-G-----TTCTG-A-----------G--GC-A--------T-CT---T--C-GACT------G----G---------A--A---C-AC-GC--T--...GAC-TGG-CA--AGA 7773
STM.US.x.STM                 --ACT--G----GG--C...TTC-TGC----G.......---T-G--T--TC-C-CGA-------TTCTG-A-----------G--GC-------------CT---T--C-GACT------A----G-----------A---C-AC-GC--T--...GAC-TGG-CA--AGA 7976
MNE.US.x.MNE027              --A-T-------GG--C...TTT-TGC----G.......---T-A--T--TC-C-C-A-G-----TTCTG-A--------G--G--GT----------C-CT---T--C-GACT------G----G-----------A---C-AC-GC--T--...GAC-TGG-CA--AGA 7802
DRL.x.x.FAO                  -GC-G--G----GT-CA...GTC-TGC-T--G......A-------CC-----A-CTTG-----TTC-G-A---T--T---TGG-G-C-G-CA----T--C----T--CAGGCT--GC--A----G--G---G-------A--GC-AC-------...C-CCTGG-G--ACA- 7748
MND_2.x.x.5440               -GC---------G--C-...GTCTT-C-T---.......A-G--C--CC--C-A--TTTG-----TTC-G-G------T---TGG-GG--G-C-----T--C----T--CAGACT---C--AA---G--T---G-----------GGT-C-------...--CCTG--A--TCA- 7690
MND_2.CM.98.CM16             -GC-G--G--G-G--C-...GTCTTGC-T--G......A-GT-C--CC--C-A--TTTG-----TTC---G------T---TGG-GG--G-C-----T--C----T-TCAGGCT--G---AA---G--T---G-----------GGT-C-------...--CCTG--A--TCA- 8057
MND_2.GA.x.M14               AGC-G--G----G--CA...GTCTTGC-T---.......A-G--C--CC--------CTTG-----TTC-G-G------T---TGG-GG--G-C-----T--C----T--CAGGCT--G---AA---G--T---G-----------G-T-C-------...--CCTG--A--TCA- 7986
OLC.CI.97.97CI12             --A-T-------G-...GCATTG-TGC-G--G.......--AT-G---G-TC-C--CCT-------TC-G-A------A---TA----C--TA-----T--C---T-TTCG--T--GC-TG-G--G-------G---------GC-CA------...GA-ATG------AA- 7534
GRV.ET.x.GRI_677             ----------G-TGCCATTC...-TGC----G.......--T--A--C------------T--T----CTG-A-------------GG---CA----------A--C---T-TC-G--T--GC-TG----G---T-------------G---C-------...GC----G-G--ACA- 7914
VER.KE.x.TYO1_patent         -GC----G----TC...CCGTTT-TGC----G.......---T-A--C--------G--T--T----CTG-A-----A-----GG-G-CAG-C---------C---T-TCAG--T--GC-TG----G---C-------------G---C-------...GC----G-G-----T 8031
VER.KE.x.9063                --TC----C---TC...CCGTTT-TGC----G.......-----A--C-----A-----T--T----CTG-A-----A-----GG-G-CAG-C---------C---T-TCAG--T---C-TG----G---A-------------A---C-------...GC----G-G-----T 8055
VER.KE.x.AGM155              -GC-----C---TC...CCGTTC-TGC----G.......-----A--C-------------T--T--G-CTG-A-----A-----GG-G-CAG-C---------C---T-TCAG--T---C-TG----G---T-------------G---C-------...GC----G-G-GA--T 8035
VER.DE.x.AGM3                --AC---GC---TC...CCGTTC-TGC----G.......-----A--C-----A-----T--T--G-CTG-A-----A-----GG-G-CAG-C---------C---T-TCAG--T---C-TG----G---C-------------G---C-------...GC----G-G-----T 7553
TAN.UG.x.TAN1                -------G--G-TGCCCTTC...-TGT-G--G.......-----A--C--T-----G-----T---GCTG-A-----T---A-AG-G-CAG--T-------C---T--CAG------C-TG-A--G---T-------------G---C-------...GC---AG-----CA- 8012
SAB.SN.x.SAB1C               --C-------G-TG...CCATTC-TGC----G.......-----A--T--------------T---GCTG-A-----T-----GG-G-CAG-C--------C---T-TCAG--G---C-CG----G---C-------------A---C----T...GC----G-A---CA- 8208
MND_1.GA.x.MNDGB1            A---G--G----G--CA...GTAAT-C-T--T......A-------TC-GC-CTCGCTG------TC-G-G------T---TG--GG--G-A-----T--C--AT-TCAGTCT--GG--A----G------G-A--A--A--AC-G---T--...-A-CTC--A---CAA 7514
DEB.CM.99.CM40               --C-C--G--G----CAGGAATT-CTT-T--T......C-GG-A-CA--TC-CA-TA--------TGCTG--------------G----CAG----------------T-TC-CTCT--GC--AG---G-----------A---C-AG-GC--T--...-A----G-G-----T 7647
DEB.CM.99.CM5                --C-------G--G-C-GGAATT-CTT-T--T......C-AG-G-CT--TC--A-TA--------TGCTG--------------G----CAG----------------T--C-CTCT--GC--AG---G-----------A---C-AG-GC--T--...-A----G-G-----T 7568
TAL.CM.01.8023               A---G--------C-C-CCTCTG-CTT-G---.......---A-A--A--TC-CA-CA-------C-CTG-C----------TGG-G-CAG-C---------G---T-TC-TTCT--GC--AGC--G---------------G-GC-CC------.......-----A-----CAT 7287
TAL.CM.00.266                AGC-----C----G-CCCCGCTA-CCT----C.......---A-A--A--TC-CA-CA-------C-CTG-C----------TGG-G-CAG-C---------G---T-TC-CTCT--GC--AG---G---------------G-AC-CC------.......-----A-----CAT 7737
SUN.GA.98.L14                AG-----G--G----CACCTGTT-CCC-------......GCTT-G-CAC-TC-CA-T--T--G--T-CTG-A------TT--TA--G-C--TA--A--T--C------CA-GC-G-C---CAA--G---T-------------GC-G-------...GT-CT-G-----AA- 8230
WRC.CI.97.97CI14             ----G-------GT...GCCATTATTC-------......A-T--A--A---C-C--GCTG-----TCTG-A------A---TA----CAGTA------A-C---T-TCAG--T--GC-TG----G---------------A--G--CC-------...GA-CT-G-A---AA- 8038
WRC.CI.98.98CI04             ------------G-...GCTATT-TCT-------......A-T--A--A---C-C--GCTG-----TCTG-A------A---TA----CAGTA------A-C---T-TCAG--T--GC-TG----G---T-----------A--G--CC-------...GA-CT-G-----AA- 8083
WRC.GM.05.Pbt_05GM_X02       --------C---G-...GCAATT-TTT-------......C-T--A--C---C-C--CCTG-----TCTG-A------A-G-TA----CAGTA------A-C---T-TCAG--T--GC-TG----G-----------------G---C-------...GA-CT-G-A---AA- 6979
LST.CD.88.524                -GC---C--------CTCCACTA-CTC----T......GC-T-G--AC-TC-CA-T--T-----T-CTG-A------TT--TA--G-C--TA--A--T--C------CA-GT-G--A-ACAA--G---T-------------GC-G-------...GT-CT-G-----AAA 7090
LST.CD.88.447                AGC-G--G-------CACCGCTG-CTC----T......GC-T-G--AC-TC-CA-T--T-----T-CTG-A------TT--TA--G-C--TA--A--T--C--A---CA-GT-G--A-ACAA--G---T-------------GC-G-------...GT-CT-G-----AAA 7087
LST.KE.x.lho7                -GC---CC-------C-CCTCTG-CTT-------......GC-T-G--AC-TC-CA-T--T-----C-CTG-A------TT--TA--G-C--TA--A--T--C------CA-GC-G-C---CAA--G---T-------------GC-G-------...GT-CT-G-----AAA 8162
SYK.KE.x.SYK173              -G--------G----CACCATTG-C-C-T--G.......-----G--A--AC-CTCG--T--G--C-CTG-A-----G-GC--TG-G-CAG----A--AC-C---T-TCAGACT--G---G----G---------------C-A--G-------...GA---AG-------A 7874
SYK.KE.x.KE51                A-C-G----------CACCATTG-C-C-T--G.......-----G--A--AC-CTCG-----G--T-CTG-A-----GAGC--TG-G-CAG----T--AC-C---T--CAGACT--G---G----G---------------C-A--G-------...GA----G-A-----A 7567
COL.CM.x.CGU1                --C----G---..........GGCATCT-T-C-TTC-CCA--T-A-CCC-----A-T-GG--C--TGCTG-A------T----G--TG--G--------A-C-----TCA-TCCC-TAATGGAA-GGCTAGTGCCAGCAGC---CGAA--T--CTG-AACTGG-C---A-- 7317
```

```
                                                                                                                                                                     RRE end
H1B.FR.83.HXB2          CAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATC 8077
Env                      Q  Q  H  L  L  Q  L  T  V  W  G  I  K  Q  L  Q  A  R  I  L  A  V  E  R  Y  L  K  D  Q  Q  L  L  G  I  W  G  C  S  G  K  L  I  C  T  T  A  V  P  W  N  A  S  W  S  N  K  S
H1A1.UG.85.U455         --------------A-------T-----------T--A---------------G---------------------C---------------A--A-------C-----------G-----C---------A-------C-----CT----------------- 7507
H1B.US.90.WEAU160       -------------CG-----------------------A---------------G-------------------------G-------------------------------------C---------A------------------------------G--- 8084
H1C.ET.86.ETH2220       ---------A-------G-----G-----------T-------------A-----G--------A-A-----C------G---------------A----------C---------------------C----------------------CT---------------G-- 7453
H1D.CD.84.84ZR085       ---------------------G-----------T--A-----------------------A----------G---------------A--A-----------------------A-------------A----A--C-----CT---------------G--- 7598
H1F1.BE.93.VI850        ----------------G-----------------T--A---------------G-------------------------------------A-----C-----C-----------------C---------AA-----C-----CT------------------ 7354
H1G.SE.93.SE6165        --G------------------------------T--A---------------C---------A------------------------A----------C-----------------C---------AA-----C-----C-T----------------- 7495
H1H.CF.90.056           -G-------A------G-----G----------T--A---------------G---C-----------------G----------------------C-----------------C---------AA------------CT------------------ 7396
H1J.SE.93.SE7887        ---------------A-------T---------T-------------------G-------------------------------A-----A---------C-----------------C---------AA-----C------------------------ 7384
H1K.CM.96.MP535         -----A--------A--G-----G----------T--A------G-----------------A------------------------A----------C-----------------C---------AA-----C-----CT----C------------.. 7235
H101_AE.TH.90.CM240     --G------------------------------T--A---------------G------------------------------A--T----A--AC-------C-----------A-------------C-----CT-C-C-----------G--- 7630
H102_AG.NG.x.IBNG       -----A--------A-------G----------T--A---------------G--------C--------------GA---------------A--A-------C-----------------C---------A----A--C-----CT----C-----------A- 7583
H1N.CM.95.YBF30         -----A--T-----------T--A---------T--A-------------A-G----T---A-A-----------T-G------G--AA----AA-TC-A-----C-----A------ACA--A---TAT--CA----------------AG-C------C--C--TA- 7625
H1O.BE.87.ANT70         --G-----AT--C--AGG--AT-TR-A-----T----GA--A----GA--TC-CC-G--A--CT-A----CC-TA---C--A----G--A-----AA-CC-A-----C--TAAA-----G--AG-C---TA---AT-A--AAAA------AGA-CA----TAGGA..... 8126
H1O.CM.91.MVP5180       --G--A--CT--C--AGGT-AT-T--A-----T--T-GA--A----GA--TC-CC-G-AA--CT-A----CCCTTA--C--A----G--A-G---AAACC-A-----C--TAAA---------A--C--TTA---AT-A--AAAA-----CA-ATCA---TCAGGA..... 8143
CPZ.CD.90.ANT           --G--A------C-A--G---T-G--------AG-A--A--A---------------G--T--A--C--G-AG-------GA---------A--AT--A-CC-C---------TG---AC--GG-G-CC--TCA----A-G-----------AA-TCC---GTA--CTTCA- 7509
CPZ.CM.01.SIVcpzCAM13   -----------A--------A-----------AG-A--A--A-----A--C---T-A--T-----A---------------C-------G---A-T--A--CC-------C-----------GTCT--C--TTAT--CA-------------CAAA-CA---TCAGGC..... 7656
CPZ.CM.05.SIVcpzEK505   -----A--CT--C--------T----T---------T--A---------A------G---T---A-A------------G-G---C--G--AA----A--CC-A-----C-----A------TCTG-G--TTA---CAA-------------A-C-C------C--C..... 7555
CPZ.CM.05.SIVcpzMB66    ---------------------T-------------T--A-----T------C-GG-T--T--------G--------------------------T-A--AC-A-----C-----------------------T-------------CA---CC----C---..... 7696
CPZ.CM.98.CAM3          ---------T-----------T----T---------T--A--------A------G-T--T-----A-----------C-GA-----G--AA----A--CC-A-----C-----A------GC------TTAT---A--------------AA--C-------GC---T.. 7488
CPZ.GA.88.GAB2          --G--A--C---C----G--AT--------------A--A-----------G--GG----T--------------T-----------G--AG----A-CAC-A-----C------------ACTG-G---TA-T--A----A--C------A--TCC----AC-G---CAA 7461
CPZ.CM.98.CAM5          -----A--T------------T----T-------------A--------A-------G-T--T-----A-----------------------G---A----AA-TC-A-----C-----A-----GC---C--TTAT--CA-G---------------C-------GC---C.. 7794
CPZ.CM.05.SIVcpzMT145   -----A-----------G--TT-T--------A-----A-----------G---G-T--T----------G------G-----G--A-----A--TC-A-----C--TA-A------ACT-----TC-----------A-GC-----CAAG-CA---G-A----TT.. 7584
CPZ.TZ.00.TAN1          -----A--CT-------G---T-T--A---------T---------------C-----G--T--A--A--G-A---A---GA--C-----------AA-CC-C-----A--TG--AAC---T-GG-G--TCA--G-AG------A-----CCTC-CC---GC-G-AG-T-- 7703
CPZ.CM.05.SIVcpzLB7     --G--A--CT--C--------T---------------------A--------A------G----T---A-C-----------------------A-----A--A--A--------------A-----------C---------T-------A-----AGA-C---------C..... 7583
CPZ.US.85.CPZUS         ---------T-A--------AT----T---------T--A-----T--A--C---G-A--T-----------------T-------G---A----A--TC-G-----C-----A------ACA-----TTAT--CA-----------------A--CC--------C--C.. 8080
CPZ.GA.88.GAB1          --------C--A--------AT--A-------AG-A--A--A--A--A--C---T-G--T-----A-----G-----GC-------G---A-T-----CC-G-----C------------GGCTG-C--TTAT--CA-G-----------CAACTC----CC-GGG...AG 8111
MON.CM.99.L1            ------AGCT-AC-A--G--T--T--T-----AG------------A-----CC-CC-GAC------A--G-AG-T-A-------C---ACAT-G--AAATGCA-----A--TG-CAAT---GC-G-G--TCA---AA-A-----G-----CAA-TC-.............. 8026
MON.NG.x.NG1            --G---AGCT-A--A--G---T-T------------------A--AT-A-----CC-CC-GAC------A---AG-T-A-------C---ACAT-GT-AAATTCA-----A--TG-CAATCGCGCAG-R---CA---ACAA----TA-----CAA--CG.............. 6651
GSN.CM.99.CN166         ------AGCT-A--A--G-CAT-T--A---------A--A-----A--AA-CC-CC-GTCAT-AC-A--G-A---TT-G-G-------ACCA-T--CCAAGCA-----C--TG-GAAC-GG-CA--C--TCA---CATA--A---------CA-CTCA.............. 8000
GSN.CM.99.CN71          ------AGCT-A--A--G--AT-T--A---------A--A-----A--A--CC-CC-GTCTT-AC-A--G-A------G-G-------ACAA-T--CCAAGC------C--TG-GAATC-G-CA--C--TCA---CATA--A----------ACTCA.............. 8012
MUS_1.CM.01.1085        -----A------A--A--GT-GT-T--T---------T--A-----A-----TC-CC-GACA--GC-T--G-AG-T-A--------C---GCC--T---AATC-C-----A--TG-AAAC-GG-AG-----TCA---AGA-----A-------A-TCA.............. 7997
MUS_1.CM.01.CM1239      --G--A---------A--GT-GT-T---------T--A--A-----A-----CC-CC-GACA--GC-T--G-AG-T-A-----------GCCT-GT--AACC--------A--TG-GAAC-GG-AG--C--TCA---AGG-----A-------A-ATCA.............. 7966
MUS_2.CM.01.CM1246      -----A------AC-A--GT-GT-T--T---------A--A--A--A-----TC-CC-GACA--AC----G-A--T-G-G-----C---GCCA-A--AAATC-G-----G---G-CAAT-GG-AA--C---CAT--CT-G-----A--------A-TC-............. 8095
MUS_2.CM.01.CM2500      -----A---------A--GT-GT-T--T---------A--A-----A-----CC-CC-GACA--AC----G-A--T-A-----------GCCT-GT-AAATC-G-----A--TG--AAC-GG-AG--C--TCA---CAGG-----A------AACTC-............. 8018
RCM.NG.x.NG411          --G---G-A---C-AA-G-----C--------G-CT--AA-C--------GC-TG--ACT---A-T--G-A------GC---C---TCT--AT-AAAT-CA-TT--C--TG-AT-G-G--AAG-G--TCA---AA-G---GAG----TATA----CAA........... 7797
RCM.GA.x.GAB1           --G---A-----C-AA-G-----C--------A-CT--AA-T--------GC-TG--ACT---A-T--G-A------GCA--C---TCT--AT--AAT-CA-TT--G--TG-CT-G-G--AAG-C---CAT--A-TG-----G---CATTC................. 7747
DEN.CD.x.CD1            --CAGTGCA--AC----GT-GT----A--------T--AA-C---A-C--TC-CC-GACT--AA-A--G-A---T--T--A--C---GCA--G--AAATGAA-----A--TG-CT-G--G-AA-----TCA---CT-A--AGAA----GCAA-CAA............. 8120
                                                                                                                                                                     RRE end
MAC.US.x.239            -----AG-AT------G---G--C--------A-CA---A-C------A-T--GG--ACT--CA-C--G-AG---T-------C--GGC--AG---AATGC------A--TG-GTTT-G--AAG-C---CA----A----A--A---CCAAA-GCAA-T........... 8484
Env                      Q  Q  E  L  L  R  L  T  V  W  G  T  K  N  L  Q  T  R  V  T  A  I  E  K  Y  L  K  D  Q  A  Q  L  N  A  W  G  C  A  F  R  Q  V  C  H  T  T  V  P  W  P  N  A  S  .  .  .  .
H2A.GW.x.ALI            ------G-AA------G---G--C--------A-CG--AA-T-------------G--ACT---A-C--G-A---T-------C--GGC--GG--AAATTCA----A--TG-ATTT-G--AAG-C---CA----A----A--A---GTAAA--AC-CC........... 8523
H2A.DE.x.BEN            -----AG-AA------G---G--C--------A-CG--AA-C-------------G--ACT---A-C--G-AG---------C----GGCA-AG--AAATTCA-----A--TG-GTTT-G--AGG-C---CA----A----A--G---GTAAA-GAC-CT........... 8539
H2A.SN.x.ST             -----AG-AA------G---G--C--------A-CA--AA-T-------------G--ACT---A-C--G-A---T-------C--GGC--AA--AAATTCA-----A--TG-GTTT-G--AAG-C---CA----A----A--A---GTAAA-GACACC........... 7969
H2B.GH.86.D205          ------G-A-------GG--G--C--------G-CG--AA-C-------------G--ACT--CA-C--G-A-----------------GGCA-AA--AAATTCA-----A--TG--TTC-G--AGG-C--TCA---GA----A--A---CCAAA-GAAACC........... 8515
H2B.CI.x.EHO            -----AG-A-------GG--G--C--------G-CG--AA-C-------------G--ACT--CA-C--G-A---T--C--A--C--GGCA-AA--AAATTCA-----A--TG--TTC-G--AGG-C--TCA---GA----A--A---GTAAA-GAA-CC........... 8501
H2G.CI.92.ABT96         ------G-AA------G---G--C--------G-CG--AA-C------A-T---G--ACT--CA-C--G-A---T---------C--GGCA-GG--AAATTCA-----A--TG-GTTT-G--AGG-C---CA----A----A------G----CCT-G-TGC-------A- 7911
H2U.FR.96.12034         -----AG-AA------GG--G--C--------A-CT--AA-C------A-T---G--ACT--CA-C--G-A-----------------GGCAAGT--AAATGCA-----A--TG-CTTT-G--AGG-C---CAT---A--------A----TAAACGACACA........... 8023
MAC.US.x.EMBL_3         -----AG-AT------G-T-G--C--------A-CA---A-C------A-T--GG--ACT--CA-C--G-AG---T--G----C--GGC--AG---AATGC------A--TG-GTTT-G--AAG-C---CA----A----A--A---CCAAA-GCAA-T........... 7615
SMM.US.x.H9             -----AG-AT------G---G--C--------A-CT---A-C------A-T---G--ACT---A-C--G-AG----------------GGC--CRN-AAATTCA-----A--TG--TTT-GG-AGG-C---CA----A-K--A--A---CCAAA-GACACA........... 7973
SMM.US.x.PGM53          -----AG-AT------G-T-G--C--------A-CT---A-C------A-T---G--ACT---A-C--G-AG----G-----------GGC--AG--AAATTCA-----A--TG--TTT-GG-AGG-C---CA----A----A--A---CCAAA-GCC-CA........... 8430
SMM.US.x.SIVsmH635F_L3  -----TG-AT------G---G--C--------A-CT---A-C------A-T---G--ACT---A-C--G-AG-----G----------GGC--AG--AAATTCA-----A--TG--TTT-GG-AAG-C--TCA----A----A--A---CCAAA-GAC-CA........... 8506
SMM.SL.92.SL92B         -----AG-AT------G---G--C--------A-CT--AA-T--T---A-T---G--ACC--CA-C--G-A-----------------GGCA-AG--AAATTCA-----A--TG-ATTT-GG-AAG-C---CA----A----A-TA---CCAAA-GACA-T........... 7932
STM.US.x.STM            -----AG-AT------G---G--C--------A-C----A-C------A-T---G--ACT---A-C--G-A-----------------GGCA-AG--AAATTC------A--TG-GTTT-G--AGG-C---CA---CA----A--A---CCAAA-GA--CT........... 8135
MNE.US.x.MNE027         -----AG-AT------G---G--C--------A-CA---A-C------A-T---G--ACT--CA-C--G-AG---T-------C--GGC--AG--AAATGC------A--TG-ATTT-G--AAG-C---CAT---A----A--A---CCAAA-GCAA-T........... 7961
DRL.x.x.FAO             ------G-GT-AC-CA-GT-A--TA--------G-A---A-C--------CC-CC--ACA---C-T--GGA----A--GGA--C---GCAA-GT--A-CC-C-----A-----CTTTGC--AAG-A---CAT--AAA----GTA---CCCAA-GAG-CA........... 7907
MND_2.x.x.5440          --CTCCG-GT-A--AA----A--TA-------TG-A---A-T--T----TCC-CC--ACA--CT----GGA----G--GC---C---TCAAGG--CTC-G-A-----A--T--ATTCTCC-AAG----TCAT---AG---AAAA---CCCAA--A-A-T........... 7849
MND_2.CM.98.CM16        --CTCCG-GT-A--AA----A--TA-------TG-A---A-T--T-----CC-CC-TACA--CT----GGAC--TG--GCT------GCAAGA--ATCC--G-----C-----ATTCGCC-AGG-A---CA---ACA---A--A---CCAAA-GAC-CT........... 8216
MND_2.GA.x.M14          --CTCCG-GT-A--AA----A--TA-------TG-A--AA-T--T-----CC-CC--ACA--CT----GGAC--TG--GC---C---TCAAGA--A-CAG-A-----A-----ATTCTCC-AAG-G--TCAT---AA---T--G---CCAAA-GAAA-T........... 8145
OLC.CI.97.97CI12        --G--AG-AT--C--A-G----TA--------AG-G---A-C--------TC-CC-G----GA--A--G-A------C--A----GGCA--A--AAAACAA-----A--TGAATTT-TG-AAG-G---CAT---A-A-----A-----CTAC-C-.............. 7690
GRV.ET.x.GRI_677        -----A--GT-----A-G--G--CA-T-----TG-G--AA-C---A-T--CC-CG--ACA---C-C--G-AG------G-------GGCA-GG--AAATTCA-----A--TG-GT-G----AAG-A--TCA---CA-A-----A-----GTA--A-AAC........... 8073
VER.KE.x.TYO1_patent    --------GA-----A-G--G--CA-T-----TG-T--AA-C---A-T--CC-CG--ACA--CC-T--G-AG------G-------GGCA-GA--AAACTCC-----G---G-AT-G----AAG-A--TCAT--CA-A---GAG---CCCTGG-CAAAT........... 8190
VER.KE.x.9063           --------AA-----A-G--G--CA-T-----TG-G--AA-C---A-T--CC-CG--ACA---C-T--G-AG------G-------GGCA-GAT--AACG-C-----G---G-AT-G----AGG-C--TCAT--CA-A--A--A---C-ATGGCAAAAC.......... 8214
VER.KE.x.AGM155         --------GA-----A-G--G--CA-T-----TG-G--AA-C---A-T--CC-CG--ACA---C-T--G-AG-----GGC------GGCA-GGT-AAACGC------G---G-GT-G----AAG-A--TCAT--AA-A--A--C----CGTGG-A-AAT........... 8194
VER.DE.x.AGM3           --------GA-----A-G--G--CA-T-----TG-G--AA-C---A-T--CC-CG--ACA---C-T--G-AG------G----C--GGC--GGT--AATGC------G---G-AT-G--G-AAG-C--TCAT--AA-C--A--G---C-GTGG-A-AAT........... 7712
TAN.UG.x.TAN1           --G-----GA--C-AA-G-----CA-T-----TG-G--AA-C---A-T--CC-CG--ACT--CC-T----A-------G-------GACA-GG--AAATT-G-----A--TG-ATTC----AGG-G--TCA----A-G-----G----C-TTC-ACAAT........... 8171
SAB.SN.x.SAB1C          --G-----AA--C-AA-G-----TA-T-----TG-A--AA-C---A-T--CC-CG--ACT---C-T--G-A-------G-------GGCA-GGT--AAC--------A--TG--TTC-GG-AGG-G--TCA---GA-G--C-T------GTA--A-AAC........... 8367
MND_1.GA.x.MNDGB1       --GTCTG-A--C--AA-------CA-A-----AG-A---A-TT-A---A-TC-CC-GACCAG-T----G-AT--TA-C-----C---GCTT-G---TCTCAA-----G--T--AT-GGC--AGG-G--TCAT---T----AGAG----C-AA--CAA-C........... 7673
DEB.CM.99.CM40          --CGG---G-----AAC-T-A----C------AG-G-GAA-C---A-CA-CC-CC-GACT---A-A--G-AG---T----A--C---GCAAAAT-AAATGAA-----C---G-CTTC----AG--C--TCAT--CA-G--C---------AACTC-CT-.......... 7806
DEB.CM.99.CM5           --CGG----------TC---A----CA-----AG-G-GAA-C---A-CA-CC-CC-GACT---A-A----AG--------A-----GAGCAAGT-AAATGAA-----A--TG-CTTT----AA--C--TCA----A-G--C--------CCA--CA---.......... 7727
TAL.CM.01.8023          --G--A--C-----A---T-A-----G-------------A-T---A-T--CC-CC-GACT--GC-T--G-A----T--G----C---GCCAGGT-AAATTCA-----A--TG--T-G----AG--C---TA---CT----A--G-----AAG-CA............ 7443
TAL.CM.00.266           --G--A--C-----A---T-A-----G-------------A-T---A-T--CC-CC-GACC--GC-T--G-A----T--G----C--GGCTAAG--AAATTCA-----G---G-AT-G----AG--C---TA---CT-G--A--G----CAAA-CA............ 7893
SUN.GA.98.L14           --G--AG-G--AC-AAG------CA-A-----AG-G---A-T-----A---C-CC-GACA--CC-A--GGAG--TG-GC-A--C---AGC--TT-A-CTTC------G--TCAAT-G----AAG-A---CAT---AA------C------TA--A-............. 8386
WRC.CI.97.97CI14        -----AG-G--------G-----CA-------GG-G--AA-C--------TC-CC-GTCA--CA-T---G---T-T--G-A--C--GTT-AAAT--AAACAA-----A--TGAATTG-C--AAG-A---CA---C-----T---------CAGTC-............. 8194
WRC.CI.98.98CI04        -----AG-AT-------G-----CA-------GG-G--AA-C--------TC-CC-GTCA--CA-T---G---T-T--G-A--C--GTTAAAAT--AAACAA---------TGAATTG-C--AGG-A---CAT--CAA---A----------AGTC-............. 8239
WRC.GM.05.Pbt_05GM_X02  -----AG-G--------G-----CA-------GG-T--AA-C--------TC-CC-GTCA--A--C---G----T--G-A--C--GGT-AAAT--AAACAA-----A--TGAATTG-C--AAG-C--TCA---A--A--C------G-G-AGGCA............. 7135
LST.CD.88.524           --G--AG-G---C-AAG------CA-------AG-G---A-T-----A---C-CC-GACA---A-T--GGAG---T-G--A--C---GCA--A--T-CCTC------G--TCAGT-G----AAG-G--TCAT---AA---A--A------TA--A-............. 7246
LST.CD.88.447           --G--AG-G--C--AAG------CA-------AG-G---A-T-----A---C-CC-GACA---A-T--GGAG-----------C---ACT--G--T-CTTCC-----G--TCAAT-G----AAG-A--TCAT---AA---AGAA-----CTA--A-............. 7243
LST.KE.x.lho7           --G--AG-A-----AAG------CA-A-----AG-G---A-T-----A---C-CC-GACA--CC-T--GGAG--TG-G--AC-----GC------T-CCTC------G--TCAAT-G----AGG----TCAT---AA----GAG----CATA--A-............. 8318
SYK.KE.x.SYK173         --G--A--C--C--AGGCT-A--T--------TG----AA-C--TA-C--CC-GC--ACA---T-A--G-CG--T--C-GA--C--GGCCA--T--TCA-A------A---G-ATTC--G-AG--C---CAT--A--A---A-G---G-GAAAGCA--TG-A--C--C-- 8044
SYK.KE.x.KE51           --------C--C--AGGCT-A--G--------TG-G--AA-C--TA-C--CC--C--ACT---C-A--G-CG--T--C-GA--C---GCCA-TA--TCA-A------A---G-ATTT--G-AA-----TCA---A--A--AA-G---C-GCAAGCA--TG-C-----C-- 7737
COL.CM.x.CGU1           --GTCTGCA---C-A--G-----G--------GG-T--AA-C-------T-C--G-TGCCA--A-A---G-T--TT-GG----A---GCAAAG--T-CC-GCATT--G--TG-AAAC-TG-AA-----T-GA---AT------C-----CAAG-CC---........... 7476
```


PLV Complete Genomes

```
H1B.FR.83.HXB2             T.............................CTGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATA 8217
Env                        __.__.__.__.__.__.__.__.__.__._L__E__Q__I__W__N__H__T__T__W__M__E__W__D__R__E__I__N__N__Y__T__S__L__I__H__S__L__I__E__E__S__Q__N__Q__Q__E__K__N__E__Q__E__L__L__E__L__D__
H1A1.UG.85.U455            -.............................-A---GG-C--A------A---T-------C--C-A-----G-A---------GT-G-------G--A-----T-TCAAC-------G-------G-------G------------T---C------C---G--C- 7647
H1B.US.90.WEAU160          -.............................-A---TT-C--------A---T------------------G-------------------G-------T--AA----------------G-----A-------A-----G------------------G---- 8224
H1C.ET.86.ETH2220          -.............................-AA--GG-------G--A---T---------C------T----------GT----T--GA-A----T--AAT--GC-----TC---------------G--C--A------A---T--C-A-C---G--C- 7593
H1D.CD.84.84ZR085          -.............................G-A--GT-T-----GGGA---T---------C------A--------G---------GG-------T--AA---G-----------------------T--------------------A-----------A-----G--C- 7738
H1F1.BE.93.VI850           -.............................-A---G---------A---T------------------A-AG--G---G------T---A-A----T--AAA--G-------G--A-----------G--------------------------C------C- 7494
H1G.SE.93.SE6165           -.............................TATA-TG-A-----G--A---T---T-----A--A----A--G-----C-----------TA-CA----T--AG-C--C---------------A--G--------G------------------C-------CC--G--CC 7635
H1H.CF.90.056              A.............................-A-AGTG-A--C---G-CA---T---T--------A-----T-A-C------G-----------GAGGA----T--AGG--GC-----TC-------C------G---------------G--C-----A-C---G--C- 7536
H1J.SE.93.SE7887           -.............................TAT--GG-C------G-GA---T---------AC-A----A--G-------------------G-AA-----T--AGTC-----------G-A-----------G----T-------A---C-------C---G--C- 7524
H1K.CM.96.MP535            ...........................TCTTG----G---------CA---T------------------A-A--G---GG-------T--GA-AC----T-TAAG--------------A---C-----G------------------T-------C---G--C- 7377
H101_AE.TH.90.CM240        -.............................T-T---G---------CA---T--A----A--A----G---------G----------A-CA---T-TGAGA--C--AC---------G-----A--G--C-G-------A-G--T--G--A-----G---- 7770
H102_AG.NG.x.IBNG          -.............................T-TA-TG-C--A--G--A---T--------AC-A----G-A--------G----------GA-A----T-TAATC------------------G----G--G-------------------C--G----C---G--C- 7723
H1N.CM.95.YBF30            CTCT..........................TAT--TACA--C------A-TTTA------CAAC-A----TGAGA--G-A-GA--C--TT--G-TG-C--TTTTGGAC-T--A---C-GG-A---G-A--A--GA-C-CA-----GA--TC-C-C---------G---C 7768
H1O.BE.87.ANT70            ..............................AAC---AGC------G-CAC-TTA--A---CA---A----TC-GC-G--A-G---CATA-GCTC-ACC---T-TGAGGA---AC--A-GG-A---GTA-----G---C-A-----GA--A-G--GC----G------G 8265
H1O.CM.91.MVP5180          ......................AGATATAATGAT--CAGT------G-CA--CTT--A---CA-C-A------CA-C-C--A------GTA-GCTC-A-T---T-TGATGA---AC---C-G-A---G---A-G-----------T-A---C---G-----GC-----G 8291
CPZ.CD.90.ANT              GCAAACATGT...GCAAAGAACAGCAGTGATA-AC--TGT------G-AA-T-T---A---CAA--A---------TT-G-AC-G----CA---G-ACAG---T-TAATA-CT-AC--AT-G-A--TG-G--A-----G-GA---A--A-G-------AT---C----C- 7676
CPZ.CM.01.SIVcpzCAM13      ........................AAGTCTA--AGTG-C--------CA-TTTA--A---CAAC-A-------A-TTG--C-CT---------G-GAC---TTTTGGT---T-A-----GG-A---TCA--A-----G--A------A-G--C--------G--G---C 7801
CPZ.CM.05.SIVcpzEK505      ......................AATAACTCCTAT--CACA--A---GGAA-T-T---A---CAAA-T-----TGAGC--G-A-GA-----TT--G-TG-C--TTTTGGGC-GC-A--GC--G-A---G-G--A---AGC-TA------A-GTC-C-C---------G---C 7703
CPZ.CM.05.SIVcpzMB66       ........................AAATCTTAT--TG-T------T--A---T--A-----C-A------A---G--G---------GATG-T--CT-TAATC--T-A---A-GG-A----C------G-----T------A-----C-CC-A--------C- 7841
CPZ.CM.98.CAM3             ........................ACATCCT-T--TG---C------A-TTTA--A---CAA--C-------A-AGGG-A--A-----TT--G-TG-C--TTTTAG-C-T--A---C--G-A---G-G--A---A-C-CA----GA--AGTC----------G---C 7633
CPZ.GA.88.GAB2             A..........................TCTTAT--GG-C-------G-A-TCT------CA-C-A----T-AGTT-G-AG-A---C-T---G-AACC--TTTTCAA---T-AC-GAG-G-A--TG-G--A--GA-T-GC------A---GC----------G---C 7604
CPZ.CM.98.CAM5             ........................ACATCCTAT--TG----C------A-TTTA--A---CAA--T-----T-A-A-GG-A--A------T--G-AG-C--TTTTAG-C-T--A--GC--G-A---G-G--A--GA-C-CA-----A---TC-T---------G---C 7939
CPZ.CM.05.SIVcpzMT145      ........................AGTGACTATC-GGTT---------A-TTAC--A---CAAC-A-----T------G-A--T---------G-AC-C--CT--A-A--GC-A--G--GG-TA-T-CA----G--G--A------A-G--G-----------G---T 7729
CPZ.TZ.00.TAN1             -ACAAAGTGCAATCACAGTGATGCAAAGTACTAT--CTGT--A-----CA-TTT---T---CA---A----TC--TT-G-AG-A--C-CT---G-AACC---T-----C-GT-A--GA--G-A----CA--A--G--G--A--CA-------G--G--A---------C- 7873
CPZ.CM.05.SIVcpzLB7        ........................AAAACATATA-TG-A-----G-CA---T--A-------A-------G---G-A-GA---------GAAA-T--TT-TGGG-----A--GC--G-A---G-------G-----T-----GA--A--C-T--A---C-G---C 7728
CPZ.US.85.CPZUS            ........................CTCTCCTAT--TGCT-----GGCA-TCTA--T---CAA--A--------A--G-A-GA--C--TT--G-TACT--TTTTAGTC-T--A---C--G-A---G-A-----GA-T-CA------A--TC-C-C---------G---C 8225
CPZ.GA.88.GAB1             C.......................AATTCCACA--TG-C-----GGGA-TCTA--A---CA-C-A----T-A-TT-G-GTCT--C------G-GAA---TTTTGGTC-GT-A--G---G-A--G---G-A---TCA--A-----G--A------AGG--CC----------G---C 8257
MON.CM.99.L1               ...................TGGGCAAAAGGT-ACTTC-CTGAG--G-CA-T-T---A---CAAC----AGTGAGTT-G-AG-T---G------TGACC--T--GCAGC-CT-G----CTG-----G-G-------GC--A--CC----T--G---A--A-GCCG-GAC 8177
MON.NG.x.NG1               ...................TGGGCTAAAGGC-ACTTC-CCGAG--G-CA---T---A---CAAC-A--AGT-TGCT-G--G-----G-T--GGCAC-T--C--GCADC-CT-G--G--GG-T--GG-G--A----G---A----C---T--GC-CA--A----G-GAC 6802
GSN.CM.99.CN166            ...................TGGGCTAATGGCTC-CTT-CAGAC---G-GA---T-------CAAA-----AGT-TGTT-G--G-G---G-----TATAC------GCAG--GT-A--GC--G-AA-CC-G--A----C-TCA--CCTGA-T--GC-GA--A-G--GTCC- 8151
GSN.CM.99.CN71             ...................TGGGCTAAGAACTC-ACG-CAGA----G-G----T-------CA-------AGT-AGTT----G-A---G-----TATACC-----GCAA---T-A---A-TG-AA-CC-T--A---TC------CATGA-T--C-----AA-G---TCC- 8163
MUS_1.CM.01.1085           ...................TGGGCAAATCACACCC-G-CAGG---G-AA-T-T---A---CAAC----AG---GTTGG-AG-----G-T--T-CAACT----GGAG---T-G---TTGG-A----GA-----G---G-----C-G--TA-----CAAA----GTTAG 8148
MUS_1.CM.01.CM1239         ...................TGGGCCAATCATACCTT--CAGGC--G-AA---T------CA-C-----AG--AGTTGG-GG-T---G-T--T-ATACC----GGG-C-CC-A----C-G-A-----G--A--G---G-A--CC----CA----GCAAA-----TTAG 8117
MUS_2.CM.01.CM1246         ...................TGGGCAAACCACACCC---CAGGA--G-CA---T------CAAC-----AG----TT-G-GG-----G-----GCAG-C--C--TGGTC--T-G---ATGG-T---C-A--A--G---G-A--CC----CA-G--GCA-A-GC-TTTAG 8246
MUS_2.CM.01.CM2500         ...................TGGGCTAATCATACCC-G-CAGAA--G--A---T------CA-------AG---GTTGG-GG-----G----TGCAACC--C--AGGTC--T-A--GTTGG-T--G--A-----G--GG-A--CC-G--TA-GC-GCAAA-GC-CTTGG 8169
RCM.NG.x.NG411             .................................AC--CAGAG----CA-ACA--A---C-T-------A---A-T---TCA-GA-TAGA-G-AAAC--CTCAGTAGCTT-AC-G--TG-A---G-A-----T---CGT----TG--T--C---GAAA----GA--T 7933
RCM.GA.x.GAB1              ...........................AACAAAACG-CTGAG---C-GA-AGAAT-A---T--C-------A---A-T--CTCTT---TAGAGGCTAAC--TACAATAGC-T-AC-G--GG-C--GG-T--A--T--G--A----TG--T-----GGA-A-----AG-- 7889
DEN.CD.x.CD1               ..........TCAATACCTGGATTGGCATTATACC-G-CTGAC-TCC-GA---T---A---CAAC-A----G-C---C--AGCAGC-CTATTTG--AATG-GTCAGA-GCCT-GA-GA--G-A--GGCA-----G----GA----TG--C---G--CAATCC---T--G 8280

MAC.US.x.239               ..............................--AAC--CA-AG----CA-TGA---T---CAA-------GC--A-GG--G--TTC-TGGA-GAAAAT---ACAG--C-CC-A--G--GG-A---TT--A-----G-----CATGT-T------CAAA-G--GA--- 8623
Env                        __.__.__.__.__.__.__.__.__.__._L__T__P__K__W__N__N__E__T__W__Q__E__W__E__R__K__V__D__F__L__E__E__N__I__T__A__L__L__E__E__A__Q__I__Q__Q__E__K__N__M__Y__E__L__Q__K__L__N__
H2A.GW.x.ALI               ..............................T-AA---CTGA----G-CA---T---G---CAA-------ACA-C--G-CCGTT-CCTAGAGGCAAAT--CAGTGAACAGT-A---CGGG-A----TT--------------ACGT-T---C--CAAA-----A--- 8662
H2A.DE.x.BEN               ..............................T-ATCG-CTGAC----AA-T-T---A---CA---------G-A-C--G-CCG-T-CCTAGAGGCAAAT--CAGTCAAAGTT-A------G-C----TT--A------------ATGT-T------CAAA-----A--- 8678
H2A.SN.x.ST                ..............................T-AACG-CTGA------CA---T---A---CA---A------GCA-CG---CCG---CCTAGAGGCAAAT--CAGTGAAAGTT-A---C-GG-A----T---------------CATGT-T---C--CAAA-----A--- 8108
H2B.GH.86.D205             ..............................--CAC--CA-A-------A---T---A---CAAC-------A-AGC--G-CC-TTTC-TGGATGCAAAT---ACGG-TC-GT-A-----GG-T--G-TA--A--G--------CATGT-T------CA-A--A--A--- 8654
H2B.CI.x.EHO               ..............................--TA-G-CAGAC------A---T---A---CAAC-A-----G--GC--G-CCG-TTC-TGGATGCAAAT---ACAAAA---C-A-----GG-A--G-TA--A-----------CATGT-T--G---CA-A-----A--C 8640
H2G.CI.92.ABT96            A.............................T----G-CACAA------A-T-T---A---CAA--A-----A-AGC-G--C---TT--TGGAGGA-AAC--CACACGGC--T-A------G-A-----TT-----------A--ATGT-T--G---CA-A-----A--- 8051
H2U.FR.96.12034            ..............................T-AAC--CT-A----G--A-T-T---G---CAA--A-----GGA-A-GG-G---T-C-TGGA-GAAAAT--CACACAGC--T-A----C-G-T----TA--A-----G--A--CATGT-T------CA-A-----A--- 8162
MAC.US.x.EMBL_3            ..............................--AAC--CAGAC-----CA-TGAT--T---CAA--------GC--A-GG--G--TTC-TGGAGGAAAAT---ACAG--C-CC-A-----GG-A----TT--A-----G-----CATGT-T------CAAA-G--GA--- 7754
SMM.US.x.H9                ..............................T--AC--CT-RC-----CA-T-T--YT---CAA--------A-A-C-GG-G---TTCCTAGAGGCAAAT---ACTSAA-Y-T-G------G-A----TT---------M-----ATGT-T-----GCAAA-----A--- 8112
SMM.US.x.PGM53             ..............................T---TG-CT-A------CA-T-T---T---CAA--A-----A---C-GG--G--G-CCTAGAGGCAAAT---ACTCAGGC-T-G------G-A----TT---------------ATGT-T-----GCAAA-----A--- 8569
SMM.US.x.SIVsmH635F_L3     ..............................T---TG-CT-A----G-CA-T-T---T---CAA--------A---A-GG--G--TTCCTAGAGGCAAAT---ACTCAA---T-A------G-A----TT---------------ATGT-T-----GCAAA--C--A--- 8645
SMM.SL.92.SL92B            ..............................--T-TC-CAGAC-----CA-T-T---A---CAA--A-----G-A-A-GG--G-ATTCCTAGAGGCAAAT---ACTCAAA-GT-G--G---G-A-G-CTA-------G--A--CATGT-T------CA-A----GA--- 8071
STM.US.x.STM               ..............................T---T--C-GA------CA-T-T---A---CA---A-----G---A-GG--G--TT-CTTGAGGCAAAT---ACACAAC-GT-A------G-T---GTT--A--G---------ATGT-T--G--GCAAA--C--A--- 8274
MNE.US.x.MNE027            ..............................--AAC--CA-A------CA-TGA---T---CAA--------GC--A-GG--G--TTC-TGGAGGAAAAT---ACGG--C-TC-A-----GG-A----TT--A-----------CATGT-T------CAAA-G--GA--- 8100
DRL.x.x.FAO                ..............................G--ACT-CA-A-----CCTCAGA---A------------C-A-A-AG-G-AG-T-G-ATAT---A-AAC--TACTCTTGATT-AC-GA-GG-AT-TG-A--AG--C--------ATCTTT--G---CAAA------GGG 8046
MND_2.x.x.5440             ..............................A-A-TG-CC-A-----CCTCAGA---A---C----A---------AG-G-A----GCATAGTG-CAAAT--GAC-ATAGACT-AC-GAGGG-AT-TG-ATTAG--C-G-G---CATCTTT--G--GCAAA------GGG 7988
MND_2.CM.98.CM16           ..............................A-AAC--CA-A-----CCTCAGA---A---C----A-----T-A-AG-G-A-CAGC-CTA--TGA-AAT--GACAGTAAACT-AC-GA-GG-AT-TG-ATT-G-GC--------AT-T-T--GC--GAAA------GAG 8355
MND_2.GA.x.M14             ..............................A-AACT-CT-A-----CCTCAGAA--A---C----A-----T---AG-G-G-CAGC-AT---T-A-AAT--GACAAT-GATT-AC-GAGGG-AT-TG-ATT-G--C--------ATGT-T------CAAA----G-GAG 8284
OLC.CI.97.97CI12           ..............................-A-TC--CAGA---GGCA-TGAT--T---T--A-----C-A--GA--G-AG-T----TA---CAGGA---G-TGA-AG-T-GAC-AT-G-AT-TC-AG----GC---G----C--T-T--G--GCA-AGG--TC-AG 7829
GRV.ET.x.GRI_677           .................................ACT-CT-AG---G-CA-T-T---T---T---------G---C------TGCC-TGGA-G--AAC--ACTCAAC--T-G------G-A-----T---G--TC-------CTGG-TCTG-ACCA-A----G---G 8209
VER.KE.x.TYO1_patent       ..............................-G-ACT-C-GA----C-AA-T-T---T---T---------A---C---AGCTG---TGGA---AAC--TACGAGACA-T-A-TGA-GG-TAG-G-A--AG-G---------CT-G-T-CC-ATCA-A-G---AC-- 8329
VER.KE.x.9063              ..............................A--ACG-CC-A----C-AA---T------T---------A---C-G---GGAG-G-TGGA-G--AAC---ACAGAACAGT-A-T-A--G-AAG-G-A--AG----G--A---TT-G-T-CC-ATCAAAG---GACAT 8353
VER.KE.x.AGM155            .................................AC--CAGAG------A-T-T-------T----------A-A-C-G--AG-AGGA-TGGAGG--AAC---ACAAAACA-T-G---C-GG-AAGGG-A--AG-G---------TTGG-T-CT-ATCAAA-G--GTCAG 8330
VER.DE.x.AGM3              ..............................AG-ACC-CTGA-------A-T-T---T---C----A-----A---C-G--ATCGT---TGGA-G-TAAC---ACAA-ACA-T-A--G---G-CAG-GCA---G-G--G------TTGG-T-C--ACCAAA-----AG-- 7851
TAN.UG.x.TAN1              .................................AC--CAGAC---G-CA-T-T-------CA---A-----G--CC----A-CTGCC-TGGA-G-AAAT--TAGTA-TACTC---TCA--G-AT-TG-G--AG-GC-G--A---ATGG-TACT-ATCAAA-G----G-G 8307
SAB.SN.x.SAB1C             .................................ACT-CCGAC---G-AA---T---G---CAA--A-----G--GC----AG-A--G--TGA-GCTAAC--CAGTAGGA-TT-A--GC--G-A--TG-A---G--C--------TTGG-TAGT-ATCAAA----G-TGT 8503
MND_1.GA.x.MNDGB1          ..............................A-CACT-CA-A-----CATCAGAA--T----A---A-----G-C-AG--C-G-TT--CTGCA-CAAAAC--TACAGAAA-GT-AA--C-GG-AT-TG-T-GAG-GC---GA--CAC-T-T------CA-A-G----GAG 7812
DEB.CM.99.CM40             ..............................GA---T-CAGAC---G-CA-T-T-------CA---------A-TGA-GG-AGCA-----T---GATGAGTGGG-AGGAGCTT-AC-GAG-G-A---G-G-----G--G-G-----T---T-CCC--CAATC-C--C-AG 7945
DEB.CM.99.CM5              ..............................GGA----CTGA-------A-T-T-------CAA--------A--GA--G-AGCA--C--T---GATGA-TGGG-AGGAGCCT-AC-GAG-G-A--GG-A--A-----G-GA----TG--T-CC---CA-TC-C-GAC-G 7866
TAL.CM.01.8023             ...................TGGACAAACTCCACAA-C-CTGA----C-GA---T---A---CAA--------G-AGTTGG-AG-----GC-T--GA-ACT--CACAGTA--GT-AC-----G-A---G-A--A--G----GA----TC--T--G--GCAAA-G---A--G 7594
TAL.CM.00.266              ...................TGGTCGAATTACACA--T-CTCAA---C-GA-T-T---A---CAA--A-----G-TGA--G-AG-----C-----G-GC-C---TCACAG--GT-AC-----G-A---G-G--A--G----GA----TC--T--G--GCAAA-G---A--G 8044
SUN.GA.98.L14              ..............................A-CACG-CC-A-----C-A-GGAC-----------A-----T--GC--G-A--A-TG--TGATGATAAT---ACAG-TC-TC-CC-G--GG-AT-TGTAACAG-GTT-G-----C--A-CA-G--CAAAC-GC--C-AG 8525
WRC.CI.97.97CI14           ........................TGGAATGCAACT-CA-A----G-AA-TGAA--T---CAAC-T---T-T-A-AG--CAG-TG-ACTGCAGGCAAAT---AGTG-AA--T-A-T-ATTG-C--TG-T--AG-GC--GCA---ATGTTT--G--GCA-A-G---TC-G 8339
WRC.CI.98.98CI04           ........................TGGAATGCTACT-CA-A-------A-TGA------CAAC-T---T-T--GAG--CAG-TG-A-TGCAGGCAAAT---AGTGTAC-CC-A-T---GG-T--T---ACAG--C-------CATGTTT------CAAA-G--GC-AG 8384
WRC.GM.05.Pbt_05GM_X02     ........................TGGAATA--AC--CTGC-------A-TGAA--A---CAAC-A---C-TTT-AG--CAG---CC-TG-TG--TAAT--CTCAGATA-GT-A-TGC--G-T--TGCT--AG--C--TCA---ATGTTT--G---CA-A-----AC-G 7280
LST.CD.88.524              ..............................G-TACT-CT-A-----CAAGGGAT-------A--A----A---C--G--GGATC--TAGA-GCAAAC--CAC-A-AC-GT-GC----GG-TT-C-CAACAG-GTT-G-----AGGA-TA-C--TAAAA-G---C-AG 7385
LST.CD.88.447              ..............................A-AACT-CT-A-----CAAGGGAT-------T-----------C--G--GGAGTA-TAGA-GCAAAT--CTC-A--C-GT-GC-----G-CT-T-C-ACAG-GTT-G-A---AG-A-T-CC--CAA-A-GC--C-GG 7382
LST.KE.x.lho7              ..............................A-CAC--CC-A-----CCA-GGAT------GA--A-----ATC-A-GG-AGCA-T---TGAC-AGAAC--TACT--TC-GT-AC-----G-AT-C-C-ACAG-GCT-G-----C-GA-TA----TAA-A-G---C-AG 8457
SYK.KE.x.SYK173            CAAC....................TTCTGT-CCA---CACAG----AA-T-T---A---CACAGA----ACA---GG-AG------TG--TGA-CAT--TG-TGGA--GC-CAG----G-A---G-A--A--G--G-GA----TG--T--TC--ACAA-G---C-GG 8193
SYK.KE.x.KE51              -.......................CGTTGC-CAAC--CCCAA--G-AA---T---A---CATACA----AC-GC-GG-AG-T----TG---GA-CAC---G-TAAT---C-GAG----G-T--GG-G--------G-------TG--T--TC-TACTA-----C-AG 7883
COL.CM.x.CGU1              ...........................GGAGA---GG-TCCA---C-GA---T-------AAC-----C-TGAGCG-G-A-GA---------GATA-----G-GG-AGACT-A-T---GG-AT-TG--TTA-----GG-A------A-GA----GGCA-------GGG 7618
```

```
                                                                                                                                                          Tat Rev intron end Tat Rev exon 2 start
H1B.FR.83.HXB2         AATGGGCAAGT...TTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCAC 8384
Rev exon 2                                                                                                                                                                           N  P
Tat exon 2                                                                                                                                                                            P  T
Env                    K  W  A  S  .  L  W  N  W  F  N  I  T  N  W  L  W  Y  I  K  L  F  I  M  I  V  G  G  L  V  G  L  R  I  V  F  A  V  L  S  I  V  N  R  V  R  Q  G  Y  S  P  L  S  F  Q  T  H
H1A1.UG.85.U455        -G-------A-...C----A-------------CT------------------T-G-C----TG----A--------------AA----A---------------A----G---------A-C--------------------C-----TC-G-------------TT 7814
H1B.US.90.WEAU160      ------------...-------C-------G-----T----C--------------------A-C---------------------------A--------------------A-----------C------A---------------------C--------G--A------------ 8391
H1C.ET.86.ETH2220      ------A--A-...C-----------------------------------------------A----------------G-----TG--A--------------A---------G----------------------------------C------------------------TT 7760
H1D.CD.84.84ZR085      -G-----G---...----------------G-------C-A-----------------------A---------------------------A-----------------------A-----------------A-----------------G--C-----T---------------T- 7905
H1F1.BE.93.VI850       ------------...C--------------G-----T----C---------------------A----------------G--------A----C------------------------------------A-----------A-------C-----T--G--A--A-------TT 7661
H1G.SE.93.SE6165       -G----------...---------------GG--------GA-----A-----------A----T-----------------T--AA------------------------G----------------A------------------C-----T--G--A----------TT 7802
H1H.CF.90.056          ------------...C------C-------G-----T--C--------------------A------------------------T--AA--------------A---------G----------------A------------------C-----T--G--T----------TT 7703
H1J.SE.93.SE7887       -G---A---A-...---------------------T----C-----------C--C---A-------------A-------T--AA---------------A---------G---G-------A----------------C-----T--G--A----------TT 7691
H1K.CM.96.MP535        ------------...C-------------G---------A-----A-------------A-------------A-----------A----------------C---------G------G------A--------C-----------C-----T--G--A----------TT 7544
H101_AE.TH.90.CM240    ------------...C-------------G------------------------------A----T-----------------T--AA---------------A---------G---------------A------------------C-----T--G--T--C-------CT 7937
H102_AG.NG.x.IBNG      -G----------...C-A-----------G------------------A---------G-A----T-----------------T--AA----------------------------G---A-----A-A------------------C-----T--G-----C-------TT 7890
H1N.CM.95.YBF30        ------AC---...C------GC------GGT--T-----A-----------------A--GCT-------------C----A-T-----CA-------CA-AAG-ATA---A-AA-----A-AGCA-------------------T--CC-T-----G-------TT 7935
H1O.BE.87.ANT70        -------CTC-...A-T---------C--G-------T--A---T---------------A--GCA-----C----------CAC-A-----GG-G---G-TA-CA-GATA------AA---------A-ACA-------------CA---CC-C-----A----T--C- 8432
H1O.CM.91.MVP5180      -------CTC-...C-T------------G-------T--A---T---------------A--GCT-----C-----G----CAC-AA-----A-----G-TA--A-GATAA-----AA-C-------G-ACA-------------CA---CC-C-----G----T--CT 8458
CPZ.CD.90.ANT          -----AGCTCA...--A------------G--------C-A-----A-------------A----T--T---------------CTA-T-----AC--------TT-GC---TAT--G--AG-TGCT-A-GA-AG---------------CAT--TC-G--A------...... 7837
CPZ.CM.01.SIVcpzCAM13  ---------TC-...---------C-----G-------C--C------------C------A-T---C-T---GC---------GA-TA---------------TA-CA-GT-A--TG-C--AG-GA-C-GG-----C---------T--C-----CC-C-----G-------TT 7968
CPZ.CM.05.SIVcpzEK505  -----T------...---------C-----G----T-----A------------C------A-T----------G------C-----A-T-----CA----------A-AAGCATAA--A-G-----G--AGCA-------------------T--CC-C-----G-------TT 7870
CPZ.CM.05.SIVcpzMB66   -G-----G--C...C-C------------G--------TCA-------------------A----T--C--A------------A--AA--------------------------T-G-----A-----A---------------------------TC-G--A----------TT 8008
CPZ.CM.98.CAM3         -----T------...--A------------G-T--T--CCGC------------C-------G--T---------------C-------A-----CA-----------AGGG---A-CA-A--AT-T--AGCA-AG---------------C--T--CC-C-----G-------TT 7800
CPZ.GA.88.GAB2         -----T-CTCA...--A------------G-----------C-----------------A-G--T-----CT----G-----GG-TA-------------T--AA-A-G----G-AAA----A-A-GA---AG-------------C-----T---------------------TT 7771
CPZ.CM.98.CAM5         -----T------...--A----G------C--T------TC-G---T---------------A----TT-G--AG------C---T--AA-------T---C--A-AGGAAT---TA-G--AG-TA-AGCA-AG-------------C--T--CC-C-----G-------TT 8106
CPZ.CM.05.SIVcpzMT145  C------T-A-...--A----G-------G-----------------T-------C-----GA-G---T----AG-----------A-TA----C---------TGC-----AA--GGA--AC--A-C-----G------A--------C-----T--G--A--G-------TT 7896
CPZ.TZ.00.TAN1         -----AGC---...C-T---G--------G-T------C-A-----------------A--GCT-----C--------C---A--A-----AC-T-----TC-CA-GTT-A--G--AA-G-----T--GCA------------T---A-----CC---TT-CA----T--CT 8040
CPZ.CM.05.SIVcpzLB7    -T---A-----C...C-------------G------GTC-C-----------C------A---------T-------A------------AA---TG-GT--G---A----------G--CG--------A--------C------------C-----TC-G--A----------TT 7895
CPZ.US.85.CPZUS        -----T------...C-A-----C------G-T--T--C--C-----------------A----TT----AG------C-A----A-----AA-C-----T--AGG--TGA--T-----AC-----AGCA-A-------------T-----T--CC-C-----A-------TT 8392
CPZ.GA.88.GAB1         -------C--C...------------------G----------A------------------A-T---C-T---GC-----------AA-CA----G----------TA--A-GA-A--TT-C--AG-----C-GG--------------C--C-----TC-C-----G-------TT 8424
MON.CM.99.L1           ------ACTTC...C-------------G-----CT-C--A---T-------C--C---A--------T-AG------C--CT---A-----C---------C--A-GTTCA----AGGAG-TA-C-G---GT-AG-----------C--T-TTC-T--T-C-----T--CT 8344
MON.NG.x.NG1           ------ACT-G...--A------------G-T--CT-C--A-----T--------T--GA-C---------AG-G----C--CT---------CC----GG-GC-CA-GTTCA-C--AGGG--CA-TCG---G--AG-------T--C--T-TCC-C--TCCA---...... 6963
GSN.CM.99.CN166        ------ACTCC...C-T----G-------G-T--TT-CG------A-A----C--T--GA-C--TG----AG----T-C--C---AA---C--------------AA-GTT-A-C---AA---GC-T-GA--GA-----------C--C-----TC-C--CCCG----T--TT 8318
GSN.CM.99.CN71         ------ACTCC...C-T----G-------G-T--TT-C--------C------C--T--GA-T-----------G----T-C--CT--A-----C---G------A-AA-GTT----T----AA----GC-T-GA--G-------------C--C-----TC-C--CCCG----T--TT 8330
MUS_1.CM.01.1085       -G----ACTCC...C-C---G-A------G-T--CT-C--A------C---------C---A-C---TGC----G-----T-C----TC-A--GCT---T--GC----CA-GTT----GT-AGGA----C-A-GG---------------C--C--T--CC-T--TCC-----T--TT 8315
MUS_1.CM.01.CM1239     ------ACTC-...--A---G-A------G-----CT-C--G-----T---------T--GA-C---TTGT----G------C----T--A--GCT---T------T-CA-GTT----G--AGGC----C-T-GG--GA-C---------C------T--CC-T--TCCA----T---T 8284
MUS_2.CM.01.CM1246     ------ATTCC...C-T---G-A------G-----TT-C--A-------C-----G-C--GA----TTGC----G------C------AA---C----------------A-GTT-A--T-AGGA--GC-T-GA--GA-----------C--C--T--CC-C--TCCA----T---T 8413
MUS_2.CM.01.CM2500     ------ATGCC...C-C---G-A------G-T--TT-C--A-----C-TC---G-T--GA----TTGCT-AG------C------AA---CA---------TA-AA-GTT-C-G--AGGG----C-C-GA-AG-------------T-----T--TC-TC-TCCA----T---T 8336
RCM.NG.x.NG411         C-----G-GAC...A--CTC-G----C-A-----GGATTGG---T-AAA--------GGA--GGA--TT-C--TA-TTT---AA-AA----C----------A-C---TTAC-CTGGAGCTGT--T-G--AGC-----GCA--T--CA-G--T--GCTC-C--CCC-A-CT 8100
RCM.GA.x.GAB1          -T----G-GA-...GCA-TC-G----C-G--TC-TGACTGG---A--CAA--------A--GG-T-CT-T------AT---TA-CA----A------G---C--GGCTGT--TGGAA-TGTC-T-G--ATC--------A--G---AGG--TC-C......TCCC-A-C- 8050
DEN.CD.x.CD1           -T----AT-A-...---------------G-TT---GC--A---T-C----GG--T-G---GG-AG-GTATG--A---C---G--A---ATA----------A-AA-GTT-A-CA--AA-G-GT-A-G---GC--T-TAGA--C--C--T--TC-T-T-CAGATCC-TAT- 8447

MAC.US.x.239           GC----ATGTG...--TG-C---------G--C-TG-TTC----A-AAA-------C---ATGGAG-TTAT-----T-T---AG-AA--CTG------------GA-CTA-A--G-ACAA--GC-AGC--AGT-A---------G---AGG---G-G-TC-C-TCCC-A-C- 8790
Env                    S  W  D  V  .  F  G  N  W  F  D  L  A  S  W  I  K  Y  I  Q  Y  G  V  Y  I  V  V  G  V  I  L  L  R  I  V  I  Y  I  V  Q  M  L  A  K  L  R  Q  G  Y  R  P  V  F  S  S  P  P
H2A.GW.x.ALI           GC----ATGT-...--TACC--C---C--G--T----CGCC---G-CAA------TC---ATGGAG-TTAT---A---T---AA-A----C-C-T---------AA-ATA----G-GCAA--GT-A-G----C-C---A----C---AGG--TG-T-TC-CCTCCC-T-C- 8829
H2A.DE.x.BEN           GC----AT-T-...C-TG-C--C-------G--T----CTCC---G-CAA------TC---ATGGAG-GCAT-------TG--AA-AA---C------------CAA-CTA----G-GCAAT-GT-A-G----T----AA----C---AGG--TG-T-TC-C-TCCC---C- 8845
H2A.SN.x.ST            GC----ATGT-...--TG-C--C-------G-TT----CTCC---A-CAAA-----TC-G-ATGGAG-TTAT-------T---AA-AA---T------------AA-ATA----G-ACAA--GT-A-G----C----AA----C---AGG--TG-T-TC-C-TCCC---C- 8275
H2B.GH.86.D205         GC----ATGT-...--TG-T---------G--C-C--CTCC---A-AAAA-----CC-T---GGAC--TAT-------C---G--A----TG------------AG-ATACA-TG-GCAG--GC-AGCA---C-----A----C---AGG---G---TC-CCTCCC-T-C- 8821
H2B.CI.x.EHO           ------AT-T-...--CA-T---------G--T-C--CTCC---A--GCA--C--C-GG---GGAT--TAT------AT---AA-A----TA-----------CAA-ATACA-TA-ACAG--GC-AGCA--GC-----A----C---AGG---G---TC-CCTCCC-T-C- 8807
H2G.CI.92.ABT96        GC----ATGTA...--TG-C---------G--C----CTCC---G-TAAA--CG-TT-C---GG-C-TTATG----G-C---AG-AA---TG------------CA-ATA----G-ACAA---C-AGGA---C-------A--G---AGG--TG-T-YT-C-TCCC-T-C- 8218
H2U.FR.96.12034        -C----AT-TC...--TG-T--C------G-TC-C--CTCA---G-TAAA---G--T-C---GGAC-GTATG------C---AA-AA--ATAC-T---------GA-ATA----G-ACAGC-GT-AGG---TC-----A-A--T---AGG---G-T-TC-C-TCCCA--C- 8329
MAC.US.x.EMBL_3        GC----ATGTG...--TG-C---------G--C-TG-TTC----A-AAA-------C---ATGGA--TTATG----T-T---AG-AA--CTG------------GA-CTA-A--G-ACAA--GC-AGC--AGT-A---------G---AGG---G-G-TC-C-TCCC-A-C- 7921
SMM.US.x.H9            GSK---ATNT-...--YG-C------K--G--C-T--TTY----A-AAAA------C---ATGGTG--C-------TCT---AG-AA---A-------------GA-ATA----G-GCAG--GT-AGC---GT-A--A-----T---AGG---G-G-TC-C-TCCC-T-C- 8279
SMM.US.x.PGM53         GC----AT-T-...--TG-C---------G--C-T--TTC----A-AAAA------C-G-ATGGTG--C-T-----TCT---AG-A----G-------------GA-ATA----G-GCAA--GA-AGC---GT-A-----A--T---AGG---G-G-TC-C-TCCC-T-C- 8736
SMM.US.x.SIVsmH635F_L3 GC----AT-TC...--TG-C---------G--C-T--TTC----A-AA-A------C---ATGGTG--C-A-----TTT---AG-A----G-------------AA-ATA----G-GCAA--GC-AGC---GT-A--A-----T---AGG---G-G-TC-C--CCC-T-C- 8812
SMM.SL.92.SL92B        GC----ATGTG...--TG-A---------G--C-T--CTCC---G-AA-A--C---C-G-ATGGAG-CT-TC----CAT---AA-A---TTG--------------A-CTA----G-ACAA--GT-A-G---GT-A---------T---AGG--TG-T-TC-CCTCCC---CA 8238
STM.US.x.STM           GC----ATGTA...--TG-C---------G--C-T--CTC----G-AA-A--C---C---ATGGAG--TATT----TAT---A--A---ATG------G---CAA-CTA-A--A-GCAAT-GT-AGC---GC-----A-A--T---CGG--TG-G-TC-C-TCCC-T-C- 8441
MNE.US.x.MNE027        GC----ATGTG...--TG-C---------G--C-TG-TTC----A-AA--------C---ACGGAG-TTAT-----T-T---AG-AA--CTG------------CA-CTA-A--G-ACAA--GC-AGC--AGT-A-----A--G---AGG---G-G-TC-C-TCCC-A-CT 8267
DRL.x.x.FAO            -T-TAA--TCA...-G-GCT--C------G-TT-C--TTGG---TCTAAA-----T--GA--GGAT-TT-C------ATG-CAA-AA----A---------TT-AGCA---T--TGGAGC-C---AGG---GT----------T---CGC--TC-T...CC-T-T-T-TTT 8210
MND_2.x.x.5440         -C-TAAACTTCCATGG-CTA-C-G----CG--C-T---TGG----TAA----G-T--GA--GGAC-GT-AG------T--TTA-CA----A---------GT-AGC-TGCT--TGG--AG--T-AGGG-AGT---------------CGC--TC-T...CC-T-TGT-TT- 8155
MND_2.CM.98.CM16       -T---A-TTCC...-G-GCT-GC-----CG--T-T--GTGG-----TAAA---G-T--GA--GGTT--C-A------AT--TGA-TA---TG-----G--TT-AGC-TGCT--TGG--AG-CT-AGGA-AGT----------T---CGC--TC-T.........C-TT-T 8513
MND_2.GA.x.M14         -T-T-A-TTCA...-G-GCT-GC-----CG--C-C--GTGG-----CAAA---G-T--GA--GGA---T-----A--ATG-TAG-AA----AC-T-----TT-AGC-TG-T--TGG----C-A-AGG---GT----------T---CGC--TC-T...CC-T-TGT-TTA 8448
OLC.CI.97.97CI12       ---TAAGTGA-...-G-ACA--A------C--TT-GG-T------A--ACT---G-----A--GGAC-TT--G----GTT--CA-AT--G--AC-C-AT-----AC-A-GAT--G--AGG-AGA-T-GG-TTA-GC-------T---GTGT-T--GCA-C-C-CA...... 7990
GRV.ET.x.GRI_677       -T---T--G-G...--C---TCA-----CTCAC-GT---C----T-AG-C---G-T---A--GGAT-TT-AG-GA---TGATTA-TC----A------T-T-CA-GG-TAT--TGGGGATGTA-C-GA-ATA---------------AAT--TC-CC-C............ 8364
VER.KE.x.TYO1_patent   GT---T--GA-...--C---TC------CG-TT-CT----A-----TAACAT-T-----A-GGGAT-TT-AG--A---T---AA-AA----G------T--C---ACA-A---TA-GGATGTA-AGTG--G---------------GTT--TC----TCCA......... 8487
VER.KE.x.9063          CT---T---AC...--C---TC------CG-TT-CT----A-----AAACAT-C-T---A--GGAT-CT-AG------T---TA-CA----AC-T--GT--C-G-A-A-AA--TA-AG-TGCA-CGG---G------------C------T--C-----TCCA......... 8511
VER.KE.x.AGM155        -C---T-G---...--T---TC------CG-TT-TT----A------AACAT-T----GA--GG-T-TT--GC----AT----G-TA----G------T-GC---ACA-AT--TA-A--TGCA-AGC---G-----------T--C--T--T-----TCC-......... 8488
VER.DE.x.AGM3          GT---T--GA-...--C---TC------CG-TT-CT----G------AACAT-C-----A--GGAT-TT--GAT---CT----TA-TA----A------T-GC---A-A-A---TA----TGCA-AGC---G-----------T--C--T--TC-T--TCCA......... 8009
TAN.UG.x.TAN1          -T---A-TTC-...-G------CATC---G--G-CT--TCA---T-C----GG--T----GGGGAT-TTAT------AT---AC-TA--TTG--T--G--G-C--GGCT-A-TTGGGGATGCA-AGC---G-----------T--C-TT--T--G--TCC-......... 8465
SAB.SN.x.SAB1C         CT---T--GAC...--T---TCG------G--C-C-----A---T-TG-A-GG--G--GA--GC---C---G-TA---C-----A-CA---TAGCT---G-TC-GC-A-TAA--A-AGGA--TC-T-GG-AGT----AA-A--G---G----TC-C--C-CC-TCC--TCT 8670
MND_1.GA.x.MNDGB1      -CCTTA--TC-...-G-GCA-G-------G--T-T--TTGG---G-TCAA--CT------GGGGAG-TT-CT----GTT---AA-TA----A---------TT-G--A--CT--TGGAA--C-A-A-G---GT-----------C---CG---TG-C-TT-CA......... 7970
DEB.CM.99.CM40         -T----ACTC-...--A------------G--C-GT-C-GA---T-C----GG--T-G----G-AG--TAC---A---CT-CT--AA--TTGC----------C-A-GTT--G-G-GAA---T-GA-G--AGT-GT-T-GA---------C--T--GCA-ACC-TT----CA 8112
DEB.CM.99.CM5          -T----ATTCA...--A------------G--C--T-C-GA---T-C----GG--T-G----G-AG-GTAT---A---C--CT---A--TTGC-T--G----CAA-GTT--GGG--AA---C-G--C--AGT-GT-T-GA---------C--T--GCA-ACC-TT----CA 8033
TAL.CM.01.8023         -C----AC--C...C-C----GC-----C--TT-GT-TGC----T-TA-A-GGC-T-G-A--GCAG----AG------CTA----AA-TTTGC-T-----T--AA-GTACA-TA-CAA-G--T-A--GTTGC-C---------C--C-T---TC--C-C---A------CT 7761
TAL.CM.00.266          -C----AC---...C-C----GC-----C--TT-GT-TGCC---T-TA-A-GGC-T-GGA--GCAG----AG------CTA-TC-TA-CTTAC-T--G--T--AA-GTACA--G--AA-G-GC-C--GTTGC-C---------C--C-T---TC-TC-----A------CT 8211
SUN.GA.98.L14          ---TTAAC.......--C----G------GG--C-GT-TC-A---T-TCTA--C--------ATGCAG-CC----TA----CATTA-AA---CAGCT-----TC--AG-TT-A-CA-ACAACAGA-CT----GA-GT-T----------AGGGTG--G--CCC-TCTG-TT-T 8689
WRC.CI.97.97CI14       -T---A-TGC-...-G-G-T-GC---C-CC-ATGGTTT---ATCTG-AAC-GGT------GGGGAT-GT-AG-TA-----ATAG--------A---------TT-AGCATG-T--A-GAA--GTA-T-GG-T--C----------T--C--TAAGC-TATA-C----...... 8500
WRC.CI.98.98CI04       -T---A---CA...-G-G-C-G----C-CC-GTGGTTT---ATTTG-AAC-GGT------GGGGAT-GC-T--TA-----ATAG-A-----CA---------T-AGGATGGC--A-GA---GCA-T-GG-T--C----------T--C--TAAGC-TATA-C----...... 8545
WRC.GM.05.Pbt_05GM_X02 -T---A-TCAA...-G-G-TTCA---C-CTCATGGTTT---ATTTG-AAC-GGT------GGG-TC-TT-AG--A-----TTAG-A-----AA-----G-TT-AGGATGCT--T--A---GTC-T-GG-T--C----------T-----TAAGC-TGT--C----...... 7441
LST.CD.88.524          -T-TTAAT.......--T----G----A-GG--C-C--T-C---T-CCAA-----T--G-ATGCAG-GC-T---A-CAT---TA-AA--ATAC-------TC-AAGCTT-A-CA-ACAA-GT--AGTG-AGA-GT-TAG---C---AGGGTTC-CG-CCCCTCTG-TT-T 7549
LST.CD.88.447          -G-TCAAT.......--T-----------AG----CTT-TCA---T-CCAA-----T--G-ATGCAG--C-A---A--AT---TA-TA---T-------G---AAGCTT-A-TG-GCAA-ATA-AGTA-A-A-GT-TAG-------AGGGTTC-TG-CCCCTCTG-TT-T 7546
LST.KE.x.lho7          ---TTAAC.......--T----G------GG-T--CT--C-C---T-TACT---G-T--G-ATGCAG-GT-A--TA--CTT-TAA-CA----G-------G--T-GAGCTT-A--A-ACAA-AT--AGTA-A-A-GT-TAG---G---AGGGTGC-C--CCCCTCTGTTT-T 8621
SYK.KE.x.SYK173        ------ACTCG...--------GC------G--T-GT-G--A---T-C-TT---C----GA--GGAT-TTATG--A-------CT--A--GCT------GC-----AG-TT-AGTG-AGGA--CA-A--G-ATT-GTT-GGA--C---GTT--TA-T-T-CAGA-CC-TAC- 8360
SYK.KE.x.KE51          ------ATTCC...C------GC------G-TC-GT-G--A---T-CCA---CC-T--GA--GGAT-CT-TGCCA---C--CTA-A--GATA------G--C--AG-TT--CTTGGGGC-----A-GA-ATA-GTT-GGA--C-----T--TC-T-T-CAGA--C-G--G 8050
COL.CM.x.CGU1          -T---A-T-A-...-G--TTTCAG-C---GG-C--TTC--CATCT-TAA----G--CTC-ATGCTGC-TATG-T---------T--AA----AC-T-----TA--A-G-TA---A-AG--TGCA-A-GAG---C-TTCAGA-TTA-GGG-TTTCAGCA-A-AGGC-GGACT 7785
```

```
                                                              Tat premature stop in HXB2                                Tat end
H1B.FR.83.HXB2          CTCCCAACC.....................CCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGAC...AGATCCATTCGATTAGTGAACGGATCCTTGGCACTTATCTGGGACGATCTGCGGAGCCTGTGCCTCTTCAGCTACCA 8530
Rev exon 2              P  P  N  .  .  .  .  .  .  .  .  .  P  E  G  T  R  Q  A  R  R  N  R  R  R  R  W  R  E  R  Q  R  .  Q  I  H  S  I  S  E  R  I  L  G  T  Y  L  G  R  S  A  E  P  V  P  L  Q  L  P
Tat exon 2               S  Q  P  .  .  .  .  .  .  .  .  .  R  G  D  P  T  G  P  K  E  *  K  K  K  V  E  R  E  T  E  T  .  D  P  F  D  *
Env                      L  P  T  .  .  .  .  .  .  .  .  .  P  R  G  P  D  R  P  E  G  I  E  E  E  G  G  E  R  D  R  D  .  R  S  I  R  L  V  N  G  S  L  A  L  I  W  D  D  L  R  S  L  C  L  F  S  Y  H
H1A1.UG.85.U455         GC-----T-.....................--AGA---T-T--G------A-G-A---C-----------------CA--G--A----..------G----C-------G-----T---A--GA--GC---------C-------A-------------------- 7960
H1B.US.90.WEAU160       ------G--.....................------------------------------C-----------------------------..------GGAA--------G-T----T---AA-----------T---C--A------------------CT------ 8537
H1C.ET.86.ETH2220       A-----CA-.....................------------------T--G------C-----------------CA--G-------..--------A--C------------T---A---A--T-T--------C-----------------------T----- 7906
H1D.CD.84.84ZR085       ------G--.....................-------------------A-------------------------CA--------GT...-C-----------G---------T--CA---------------------A---A------------------- 8051
H1F1.BE.93.VI850        A------G-.....................--------------------------C------G---------CA--G--A---..------G-G--C-------C----T---A--T---GC-------------A---A------------------G 7807
H1G.SE.93.SE6165        AC--ACCA-....................-A-----A--------------------C------G-----C--CA--G----G-..------AG-G---------G-----T---AC-G--------------C----------------------------- 7948
H1H.CF.90.056           G----G-A-....................--AC--------------------C-------G-----C--CA----------..-------G-GA----------------T---AC--G--G---------C--C-----------CA-------------G 7849
H1J.SE.93.SE7887        A------A-....................--A-C--A-G-----------G-----C-------G---------CA--G---AC-..------G-------------------T-C-A--T---GC---------C-------A---------------T----- 7837
H1K.CM.96.MP535         A--------....................T--------G-A-----A----------C------------------CA----A-A--..------AG-----------G---C-T---A--G---GC-----------------A------------------G 7690
H101_AE.TH.90.CM240     TC--ATCAT....................-A--A--A---------A----------C------G-----C--CA--G--------..------AG-G----------G-----T---A-----GC-----------A---------------------------- 8083
H102_AG.NG.x.IBNG       AC--ACCA-....................-A-----A--------A-----A---C-------G---------CA----A---------..------G-G----------G-----T---A-----GC--------------------------------------- 8036
H1N.CM.95.YBF30         A-------A.....................G-A--------A--------A---A-C----G--GC-T-------CA--------G-..------G-G----------G-----T--CA--T---G-------G--C--C----A-----TGA------CT------ 8081
H1O.BE.87.ANT70         AA--ATCA-....................-AAGA--A-G-A-GA-C---A-G-A---C--G--G--G---------AGA--G---GCC-...--G-GG--A-CC-CGCC-C-A--G-T----C----GT-G-ACACG--C--CA-A-CAA-AATAT-G-GG-CT------ 8578
H1O.CM.91.MVP5180       G-----CA-....................-G-CA--A-G-A--A-C---A-G-A---C--G---------------AG--------GCC-...-AG-GG-CAGCC--GCCACCA----T----CA--AGT-G-ACACG-----CA---CAA-AAT-T-G-GG-CT----- 8604
CPZ.CD.90.ANT           A----T---CAAAAC...............-A-CA--AT--A--GCA---A----A----AG-------------AGA-A------GAT-...--G-GG-GGGCC--GCA-C----G-T---C-----CT-G----TG--C---AC----A-AAT--AG-GG-T------ 7989
CPZ.CM.01.SIVcpzCAM13   A-----G--....................-A-----A---A-GA-C--T--G--A--C--GC-G-------------CC--G---C-G-..------G-GA-------A-CA--G-G--G-C-G--C------------C--C--C-A----GT-A---GG--------G 8114
CPZ.CM.05.SIVcpzEK505   A-------G....................A-TC-------A--------A---A---C---------C-C--------CT-----AC-G-..------G-G--C--------G-----T---A--T---GCG-----G--CT-CA---A---TCTA------CT------- 8016
CPZ.CM.05.SIVcpzMB66    A-----G--....................-G-----AT-GA--------A---A---C--------G----------CC-----AC-T-..------A-----C--------G-----T---A-----GC----A----C--A---GA------T------CTT----- 8154
CPZ.CM.98.CAM3          A----G--A....................A-A----A----A--------A-G-----C---------C-T--------CC--G--A--G-..-------------C--------G-----T---A--T-----------A-----C----A---CCTG------T-T----- 7946
CPZ.GA.88.GAB2          A-----G-A....................A-AG-------A--AGT---A--------C-G--G--G---------CA--------G-..------TG-GA-------C-G----TT--A--T--CTG---------CA-C----A----ACGA-----CT-----G 7917
CPZ.CM.98.CAM5          A----G--A....................A-A----A---A-------GA-G-----C---------C---------CC--G----G-..---------CGA----G--C-G---G-T---A--T---G-------A-----T----A---CCTGA-----TT------ 8252
CPZ.CM.05.SIVcpzMT145   A-----G--....................AATCA------A---G---T--G--A--C----A---G-------------AC-T-----G-..------A--CA----G----G---G-T----C-T---G-------A-----T----A---CCT-AG----CT------ 8042
CPZ.TZ.00.TAN1          AC--A-G-G....................GA-CA--AT--A--ACA---A-G------C-C--G--G---------GC-------ACAT-...--G-GG-CG-CC-CGCCAGCA----T---CAGTA-CG-------G--C--CA---A---CCT-A---GG-TA----- 8186
CPZ.CM.05.SIVcpzLB7     TC---C-T-....................----------T--CTCA-----G-----C-------------------CA--G--C-----..------AG-GA----GC--G-------T---A------GCT---A----C---A-GA-------------CTT----- 8041
CPZ.US.85.CPZUS         T-------A....................A-TC---A---A--------A-------C---------GC-C----A-A-C-----AC-T-..------A-CGA--------C-G-------T---A------G-------A-----T----A---CCTGA-----CT------ 8538
CPZ.GA.88.GAB1          A-----GT-....................-A-----A--AA-GA---T--G--A--C--C---G------C--CA---------AG-..------AG-GA--------G-A--G-G---C---------------------C--C--A-A----G-GA-T-GG--------- 8570
MON.CM.99.L1            A-----T--....................-ACGC-----AGCCA-C-------C---C--GC-C--GC-----C--CG-CAG--ACTC-..-------GCCAGC--ATC----A----TT--TA-CA-----------A-----CA---A---TGT-----GG-CA----- 8490
MON.NG.x.NG1            ...ATCC-TACCCAAGTCCCTCGG.............-A---A-GAGC---A----A--C--G----GC-----------CA--G---GA--AGGTCRA-AG-CTATGCGAGAGGATTT-TAACCATTA--TGGACA---TGC-GAGACCTAA-AGT-TGGAC-TT-C-G-T 7118
GSN.CM.99.CN166         A-----CAA....................G-AGT--A--G--GTCC---A--C-C--CC-------GC-T---C-C-CA--G--AG-T-...--G---G-GA---ATC-A-CA-----T--C-AGT--CT-G--------C---A---A----GTGA-T-GG-T------ 8464
GSN.CM.99.CN71          AC----CAG....................-AAGT--A--G---TCC---A--C----C-AG-A----C-----C-C-CA--G--AC-T-...--G---G-CA---ATC-C-CA-----T--C-AGT--CT-G--------C---A----T---GTAA-T-GG-T------ 8476
MUS_1.CM.01.1085        A---G-C-G....................-A--C------A--AGC---A-------C-----G---C-----AGCG--AG--------..------G-GA---ATC---G-----T---CT-T--CT----------C--C--C-A---AGTGA-T-GG-------- 8461
MUS_1.CM.01.CM1239      AC--A-C-T....................--AGA--C-G-A--GCC---A-------C-----G--GC-T---CAGC--C-G---GAC-..------G-G----ATC---------T-----T-G--GT-T--------C--C--C-A---AGTGA---GGTTT----- 8430
MUS_2.CM.01.CM1246      AC--G-.................................AGTC--C-GA-C----GC--CAA-A-A-A-C---A-A--G-...GCTGGA-GAA-CGGCAG-C--A--------T-G--GT-T-CA--A--C--C-AC-A---AATGA-T-GG-CT----- 8538
MUS_2.CM.01.CM2500      AC--G-G--....................---GA------AAGA-C-------A--C--------G------C-CC----G---T-T-..------G-G--C-ATC---------T-----T-T--------------C--C--C-A---ATAG-G-GG-------- 8482
RCM.NG.x.NG411          TGTTATCATCAGCAGATCCCTATCCGCAAT......C-GGAG--ACA---AATCAA-GA--G----A----A--CAGTG----AG-C-GT...G--CA--AATAT-ACAACTGGCAGAGAGGA..................TATTTTTATA-CCTGA-AC--C--CCGAT 8243
RCM.GA.x.GAB1           TCTTATGTTCAGCAGATCCATATCCACAAC...-C---GGAACCGCAAA-TCC----GA-A--AG-----AC--T-GAGA---AG-T-G-.....AA--AGTACAACAATTGGCTGAGAGAA..................TATT-CT-GA-TCAA--GA--CA-CCGTT 8193
DEN.CD.x.CD1            -GA...........................--CCA-AAGG-A-CA-AT-AGCC--A-GA--G-----G--AAA-T-GAGAC-G--ACAGGCTC-AG-GGTGG-A--ATCC--GA-----T----AGTA-C--T-----A-----CTCACAG---CTGACG-GG--T----T 8590
                         Tat Rev intron end  Tat Rev exon 2 start                                                               Tat end
MAC.US.x.239            TCTTATTT-CAGCAGACCCATATCCAACAG......-AC--G-CACT---AACCA--GA--GCA-----A-A-AC-GTG----AG-C-GT...G-CAA--GCTCC-GGCCTTGGCAGATAGAATATA--CAT-TCCTGATC-GC-AACTGA-AC-------G-CT-GG-T 8951
Env                      S  Y  F  Q  Q  T  H  I  Q  Q  .  .  .  D  P  A  L  P  T  R  E  G  K  E  R  D  G  G  E  G  G  .  G  N  S  S  W  P  W  Q  I  E  Y  I  H  F  L  I  R  Q  L  I  R  L  L  T  W  L
Tat exon 2                                  P  I  S  N  R  .  .  .  T  R  H  C  Q  P  E  K  A  K  K  E  T  V  E  K  A  V  .  A  T  A  P  G  L  G  R  *
Rev exon 2                                 N  P  Y  P  T  .  .  .  G  P  G  T  A  N  Q  R  R  Q  R  K  R  R  W  R  R  R  .  W  Q  Q  L  L  A  L  A  D  R  I  Y  S  F  P  D  P  P  T  D  T  P  L  D  L  A
H2A.GW.x.ALI            GGTTAC-T-CAACAGATCCATATCCACAAG......-AC-AG--ACA---AACCA--GG----AC-----AA-AC-TTG------AC-TT...G-GGA--GATTG-GGCCCTGGCCGAT-GCATATT-ACAT-TCCTGATC-AC-T-CTAGCTC-------G-T-GGG-T 8990
H2A.DE.x.BEN            GGTTATCT-CAACAGATCCATATCCACAAG......-AC-GG-GACA---A-CCAACGA--G-AC-----AA-AC-TCG---G-GACAGT...G-T-A-GACTTG-GGCCTTGGCC-ATAAACTATG-GCAG-TCCTGATC-AC-TACTGACTC-------G-T-GGG-T 9006
H2A.SN.x.ST             GCTTACTT-CAACAGATCCATATCCACAAG......-AC-GG--ACA---A-CCA--GA----AC-----AA-AC-TTG--A----C-TT...G--GA--A-T-G-GGCCCTGGCCGATAAGATATA-ACAT-TCCTGATC-GC-A-CTGA-TC-------G-A-AGA-T 8436
H2B.GH.86.D205          TCTTAT--TCAACAGATCCCTATCCGCAAG......-AC-GG-GACA---A-CCAACGA----AC-----AA----GTG--A--GAC-GG...GAC-A--GATCC-GGCCCTGGCAGATAGAATACA--CAT-TTCTACTG-GC-A-CTGAG-AA------A-TT-GG-T 8982
H2B.CI.x.EHO            TCTTAT--TCAACAGATCCCTATCCGCAAG......-AC-GG-GACA---A-CCAACGA----AC-----AA----GTG--A---AC--G...G-C-A--GATCT-GGCCCTGGCAGAT-GAATACA-CCA--TCCCGAT--GC-AACTGAG-GA------G-TT-GG-T 8968
H2G.CI.92.ABT96         TCTTATGTTCAGCAGATCCATATCCGCAAG......-AC-AG--ACW---AACCAA-GA----AT------A-A-AGTG--A---A--GT...G-C-A--GATCT-GGCCTTGGCAGATAGAATATA--CAT-TCCTGMTC-GC-A-CTAGG-AA------G-CT-GG-T 8379
H2U.FR.96.12034         TCTTATGTTGAACAGATCCCTATCCAACGG......-AC-AG--ACA---AACCAA-GG----AT-----AA--C-CTG-------C-GT...G-C-A--GATCT-GGCCTTGGCAGATAAGATATA-ACAT-TCCTGATC-AC-AACTGA-TC----G--G-CT-GG-T 8490
MAC.US.x.EMBL_3         TCTTATTT-CAGNNNACTCATACCCAACAG......-AC--G-CACT---AACCA--GA--GCA-------A-AC-GTG---AG-C-GT...G-CAA--GCTCC-GGCCTTGGCAGATAGAATATA--CAT-TCCTGATC-GC-AACTGA-AC-------G-CT-GG-T 8082
SMM.US.x.H9             GCTTRTGTTYAGCAGATCCCTATCYARACG......--C-AG--ACT---AACCAA-GA--G---------A-AC-GTG---R-----GT...G-CAA--GATCT-GGCCTTRGCAGATAGAATATA--CAT-TCCTGATC-GC-A-CTGA-AC-------G-CT-GG-T 8440
SMM.US.x.PGM53          G-TTATGTTCAGCAGATCCCTATCCACAAG......--C-AG--ACA---AACCAA-GA--G---------A-AC-GTG---------GT...G-CAA--GATCT-GGCCTTGGCAGATAGAATATA--CAT-TCCTGATC-GC-A-CTGA-TC-------G-CT-GG-T 8897
SMM.US.x.SIVsmH635F_L3  GCTTATGTTCAGCAGATCCCTATCCACAAG......-AC-AG--ACC---AACCAA-GA--G---------A-A--GTG---------GT...G-CAG--GATCT-GGCCTTGGCAGATAGAATATA--CAT-TCCTAATC-GC-AACTGA-AC-------G-CT-GG-T 8973
SMM.SL.92.SL92B         TCTTATCATCAGCAGATCCATATCCAGCGG......-AC-AG--ACT---A-CCAA-GA--G-AC------A-A--GTG--A--G---GT...G---A--GATTG-GGCCTTGGCAGATAGAATATA--CAT-TCCTGATC-GC-A-CTGA-AC-GA-A--G-CT-GG-T 8399
STM.US.x.STM            TCTTGTCGTCAGCAGATCCCTATCCACAAG......--C-AG--ACA---AACCAA-GA--G-AC-----AA----GTG---------GT...G--AT--ACTCC-GGCCTTGGCA-ATAGAATATAC-CAT-TCCTGATC-GC-A-CTAG-AC-------G-CT-GG-T 8602
MNE.US.x.MNE027         TCTTATTT-CAGCAGACCCATATCCGACAG......-AC-AA-CACT---AACCAA-GA--G-AC------A-AC-GTG-------C-GT...G-CAA--G-TCC-GGCCTTGGCAGATAGA-TATA--CAT-TCCTGATC-GC-AACTGA-AC-------G-CT-GG-T 8428
DRL.x.x.FAO             AAGGACG-ATACCGCCCAGAGACG......A-C-ACTAC-A-AGACA---A--CA--GA-A---G----AGCA---C----G-GAGTGG...-ACAT--GATC-GA-CCCTC-AAGC-AGGAT-CTCC-AAGCTTGGAGCGAAGAA......ACAG-AGGG-C--GGTT 8365
MND_2.x.x.5440          AAGGG-GA-TATCTCCGGCCCCAC............AAC-T-A-AC----A--CA--GA--G--G----AGCC---CTT---A-A-C-G...-ACAT--AGTC-GAGAGCTC-A--CAAGAAT-TAGG-AGCCTTGGA-G-CAGA-..........-AAG----GAG-TG 8301
MND_2.CM.98.CM16        G--TTC-AGGGAGACTATCTCCGTCCCCAC......AAC-T-A-ACA---A-GCAA-GA-AG--------AGCCC--CTC---A-A-C-G...--TA---AATC-GACAGCTC-AA-GTAGAATTTGGG-AGCCTTGGAGCAAAGA-..........-AAA----AG-TTG 8665
MND_2.GA.x.M14          AAGGG-GA-TATCTCCGACCTCGC............AAC-T-AGACA---A--C-A-GA-AG-----G--AGCTCA-CTT---AGA-C-G...-ACAT--AGTC-GAGAGCTC-AAGAAAGAATTTGGC-GACCTTGGAGG-CAGA-..........-AAA-A--GA-TTG 8594
OLC.CI.97.97CI12        ............TTTTCAGATACCCTC......CA--ACAAG--ACA---A--CAACGG-----G-----ACA-C---TTCTT-GAGAG-......A--T-CAACGAGTG----T-TA--ACA-AGTG--CAG--ACATTGAGAGTA-ATTG-A-AGA---ACTA-CA-- 8136
GRV.ET.x.GRI_677        ............CAGATCCATATCCACAGT......TC-G-G--AC----A--CAACGG--G--GGC---ACA---GTG----A--CAG-AGC--CAAAT-GAT-AG-T--C-G-A-GAG-CCT--ACACCT-C-AGGATCAACAACT-GTG-CT-AA-----AGAG-TG 8516
VER.KE.x.TYO1_patent    ............CAGATCCATATCCACCAA...GT---GAAG-GAC----A--CAACGCC--C--GCC---A-A--GTG------ACAG-AGG-TCAAAT-AGAG-CTTG-T-GAA-GA--CCAAGAGC-GG--TATGC-GT--ACCGC-TG-CTGACGAGGCT-A--.. 8640
VER.KE.x.9063           ............CAGATCCATATCCACCCG......T-GAAG-GAC----A-GCAACGC--G---GCC--AA-A--GTG----G------AACGACAG-TCAGAT-CTTG-C-GAA-GGA-CT-G-ACG-GACAAA-GAGAGGCA-TT-GAGCAAA-G---G-C-A-TTG 8663
VER.KE.x.AGM155         ............CAGATCCATATCCATCCG......T-GAAG-GACA---A--CAACGC--G---GCC--AA-A--GTG--AGA-C--G-AAA--CAAATC-ACGCATTA-C-GAA-GAA--T-GGGGACGAGACA-GAGGACCA-TT-GTGCA-G-AA--G-C-A-TTG 8640
VER.DE.x.AGM3           ............CAGATCCATATCCACCCG......T-GAAG-GACA---A--CAACGC-----GGCC---A-A--GTG------AGCG-AAG-ACAG-TCCGAGCCTTG-C-GAA-GAA-CT-GCACAGCAGA-TGGA-GAGCAACT-GTGCAAG-GA--G-C-A-TTG 8161
TAN.UG.x.TAN1           ............CAGATCAATATCCGCCTG......--GAGG--ACA---A--CAACGC--G--G-----ACAA---TTCCAG---CAG-...---GA--AGTCGCCGCCCTCA-TGAAAGAAT-TT-AT-GCCCA--AGAGGAG---T--AAGCGGAGGAG--AGCTTG 8614
SAB.SN.x.SAB1C          TCT-ATTAA...CAGATCCATATCCACCTG......-ACAAG-GACA---A--CAA-GA------GC----CA-C-GTG-GCTA-ACAG-GAC--CAGGTCATACACTTG-C-GA--GAA--CTTG......................................--AC---T 8795
MND_1.GA.x.MNDGB1       ............CAGGACTGCCAGCAGAAC...CTATAC-G-A-AC----A--CAACGG---------A-CAACAGCTT---ACT--GA...GAGCA--ACTCCGAGAACTTG.........AAGGAAGAA-CCTTAA-CAGATCCTTGA-AGAGGA-C-G-C-AG-TT 8113
DEB.CM.99.CM40          AAG...........................GA-GA--A-GTGCTTC-A-T--C--A-GA--G--GCAG-A------GAGAC-G---G-C-...--G-GG-CAGC---GCC--GA----T---CA----T---CAGT--C--------G-----C-AACAA---CTGACG-GGCT------ 8252
DEB.CM.99.CM5           AAG...........................GA-GT--A-GTGCTTCAA-T--G----GA--G---CA--A------AG-C-T--AGAG-...--G-GG-CGAC--CGCCC-GA----T----A----T----AGG--C-----------CA-CTCACAG--TCTGACG-GGCT------ 8173
TAL.CM.01.8023          A----G-G-....................AAC----A---A-GAGT---A----A-GA--GTAC---------A--CCC-----GAG-...--G-GG-GAAC--CGCCTC-A----T---T-----T---T--CA----------G--C--CAAC-----CCT-AG--A---T------ 7907
TAL.CM.00.266           A----G-G-....................AAC----AT--A--AGC---A-G---GAC-CTAC---------A--CA--G---CT--...-AG-GG-ACA---CGCAAGCA----T-----A-----T----AGCA----T-----A------ACAG-T---AT-AGA-A--T------ 8357
SUN.GA.98.L14           G-TGA-CAGGATTGGCTCCAGGAAACCTGC...CCCAA---ACAGACAAA----A-GA-----GAC--AGAA---A---AT-......TAC-TCAA-T-AGAGCAGTCC--GAA-GAA-CTTTG-CACCACCTTGGACAG-AGACT-GGACGAG-CT--G--AG--TC 8850
WRC.CI.97.97CI14        ............CAGAACCACCTCAACAAC...TT--A-GAG-GACA--AA--CA-CGA-----G-AT--AA-A---TTACA----TATAGAA-TCAA-GA-T-GCAGTAC-GGA--GAA-CT---A-AG---TCTTGA-CAGCTT-CTAACC-CA--A------ATTTG 8655
WRC.CI.98.98CI04        ............CAGAGCCACCTCAACAAC...TT--A-GAG-GACT---A--CA--GA--------C--AA-A---C-ACAGT-CTTTAGACGTCAG-GA-T-GCAGT-C--GAA-GAA-CT---A-A-C--T-TTGA-CAGCTT--TAAAC-T-T-A-CTGCAGT-TG 8700
WRC.GM.05.Pbt_05GM_X02  ............CAGAGTCATATCCTGAAC...CT--A-GAG-GACT---A--CAACGA----AC-----AA-A---TTTCATAGCCATAGAAGA---AGACT-GCAGT-C--GAA-GAA-C-CTGA-AC---TCTGGA-CAGCTT-GCAACC-TA--T-A----ATTTG 7596
LST.CD.88.524           G-TGA-CAGGACTTCAAGTGGGAGAACGGA...GA-A-C-AA--ACA---A--CAA-GA--G---GAC--AGAA---C-----A-ACAT-TAC-TCAA-T-AGAGCAGTGC--GAA-GAA-CTTTCAGGCCACC-TGGA-CG--GACT-GAACGAG-CT--ACAGG--TC 7716
LST.CD.88.447           G-TGA-CAGGACTGCAAGTGGGAGAAAAGA...GA-AAC-AA--ACA---A--CAA-GA--G---GAT--AGAA---C-A---A-ACAT-TAC-TCAA-T-AGA-CAGTAC--GAA-GAA-CTT-CAC-CCACC-TGGA-CG--GACT-GAGCGAG-CT--ACAAG--TC 7713
LST.KE.x.lho7           A-TGA-CAGGACTACAAGTGGGAGAAAGAA...GA-AAC-AA--ACA---A--CA--GA-----GA---A-C---T-C---G-CTAT-TAC-TCAA-T-AGAGCAGTGC--GAA-GAA-CTT-CAGGCCACT-TGGA-CG--GACT-GAACGAG-CT--ACAAG--TC 8788
SYK.KE.x.SYK173         ..............................CA---C-GGA-GGAC--A-GCAA-CCC-C---C----AA-A--GC--T-GAGACAGA...GA-GGAC-CAACG-CTCT-C-TTCAG-AGA-A-TC-T-GA--C-ATC-T--GAAGCAGGACAG-AGC-TT-GAGGAC 8494
SYK.KE.x.KE51           -C---T---.....................CTCTAC-A-A-GGCC--A-C---GCC--------G-ATA-A-ACAGATACAGACA-T...TACAAGCCCA-CG-CTCT-CGTTCAG-AGA-A-T--T-GA-AC-ATCG--CGCAGAAGC-CAGAAAC-TT-GA--AG 8193
COL.CM.x.CGU1           AATGTGT-TTCA..................-A--TCC-TGG-C-AGAC-GG--GT-GGA-C-GCC---CAAC-C---AGA-ATACACAGA...GAGCAAGAAAACAGCAGC-G-A---TT-G-............................................-G 7889
```

```
                                                                                            Rev end (TAA) in some lineages
H1B.FR.83.HXB2          CCGCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTGGAACTTCTGGGA....................CGCAGGGGGTGGGAAGCCCTCAAATATTGG................................TGGAATCTCCTACAGTATTGGAGTCAGGAACTAA 8646
Rev exon 2               P  L  E  R  L  T  L  D  C  N  E  D  C  G  T  S  G  .  .  .  .  .  .  .  T  Q  G  V  G  S  P  Q  I  L  .  .  .  .  .  .  .  .  .  .  V  E  S  P  T  V  L  E  S  G  T  K
Env                       R  L  R  D  L  L  L  I  V  T  R  I  V  E  L  L  G  .  .  .  .  .  .  .  R  R  G  W  E  A  L  K  Y  W  .  .  .  .  .  .  .  .  .  .  W  N  L  L  Q  Y  W  S  Q  E  L
H1A1.UG.85.U455         ---A-----------CGC----------G----AGC---------------CGCAGCAGCCTCAAGGGACTGA-ACT---------G-G-----------CT-..............................---------T--G-T-------G---G-------- 8097
H1B.US.90.WEAU160       --------TC---------------C--A-------------------.....................------------------T-----------T................................-----------G-----------C--------- 8653
H1C.ET.86.ETH2220       ----------------A---------C-G------CA---------------CGCAGCAGTCTCAAGGGACTA-AG------------A----T------CT-..............................G-A-GC--TG-G---------G---T---G---- 8043
H1D.CD.84.84ZR085       ------------A---A---A----C-G--------------------.....................-----------------------------CT-..............................-----------G--A----------GA-----G- 8167
H1F1.BE.93.VI850        --A-----------CA-A--A----C-G-------------CAGGGGCCTG.....................A-G--------------T---------CT-..............................G--------ACG-G--------------------- 7923
H1G.SE.93.SE6165        --------------CCA----------G------CA---------------CGCAGCAGCCTCAAGGGACTGA-ACT-----------GT--T-------T-..............................---------T--GTT-------G-CAG-------- 8085
H1H.CF.90.056           --T-------------------A------GT-----C---------------G.....................A-A------A----------------CTC..............................---------T-----A--C---G-A---------- 7965
H1J.SE.93.SE7887        ---A-----------CG---------C-G------C-----G-AC-------.....................-T-----------GAT----------CT-..............................GT---C--TG--TG-------G-G---------- 7953
H1K.CM.96.MP535         --AA------A-----A----A-----G--------CC-----AGGGGACTG.....................A-GG----T------------------CT-..............................-----C--TG----------------------- 7806
H101_AE.TH.90.CM240     ---G-----------AC--------C-G------CG---------------CACAGCAGTCTCAAGGGACTGA-AC------------G----------CT-..............................G-------T--GTTA------G-C---------- 8220
H102_AG.NG.x.IBNG       ---A-----------A---------C-G------C----------------CACAACTGTCTCAAGGGACTGA-ACT--------G-------------CT-..............................---------TA--TCA--C---GT---------- 8173
H1N.CM.95.YBF30         --------C-----C---------AC-G-G-----C-C-------C------CAGAGTCTCAGCAGGGGACTG-AACTACT-AAT---CT-AGA-C-C-C-T-..............................---GGAA-A--TGCA------G-AA-A--GT--- 8218
H1O.BE.87.ANT70         --T------CA-----GCA-CAGGGA-CCA--A-G-GA-CAGCTA----A-GCTTGGACTGTGGATCCTAGGG-AG-A-ATAATTA-T-TTTG--G-AT---T..............................GCAGC-G-AAC---A--C---CTA--A---T-GC 8715
H1O.CM.91.MVP5180       --T------CA-----A-A-CAGGGA-CCG----C-GA-C--CTAC------CTGGGACTGTGGATCCTGGGA-AA-A-ACAATT-----TTGT-G-CT---T..............................G-AGC-G-AA-G--A------CTA--A---T-G- 8741
CPZ.CD.90.ANT           GAT--GC---AC--GT-------ACC-TTG-GCAG-CC-CC---AC--CT-CAGA.....................ATT-CTTTCA-ACTGTG-AA-C-TCTGGA-..............................AAC------AGCACCCTC---...ACAAT-A-C- 8105
CPZ.CM.01.SIVcpzCAM13   GA------C-AG---GC--GC-----TTG---ACAGT-----AAGT-----CAACTGTTAATCAACATCCTCA-ACT-ATACA---GAAA--G-CTCTACT-..............................A--GGGA-AA-C--A------G-AAGA---T--- 8251
CPZ.CM.05.SIVcpzEK505   --------C-----GT----CA--CC-G-G-----CCT-----------A--CAGAACATCCACAAGGGACTG-AACTACTTAAT---CT-AGA-T-----T-..............................---GGAA-AA--GCT--C---G-AAG---G---- 8153
CPZ.CM.05.SIVcpzMB66    ---A--------------------A--GCT-----CAT-------GG-----CAGACCCTCCTCAAGGGACTGA-AC----AC----------A-T-C-CCT-..............................A--GGCA-T-----A------G-A-----G---- 8291
CPZ.CM.98.CAM3          ---A---------------A-----AC-GGG-----CAT-----AAC------CAGAGCCTCAACAAGGGACTG-AACA-CT-A--A-CTT-AG--G---CCT-..............................---GGAG-GA--ACC--C---G-AAG---G---C 8083
CPZ.GA.88.GAB2          --T---------T-GT----C-G-CC-TTG---CC-GC------AAT-----CAACAGGTCCTGAGAGGACTGA-ACTACTCA----GCTT--G-GCC-GCT-..............................AAAGGGA-AGG---A------CTG--A------- 8054
CPZ.CM.98.CAM5          ---A----T-----------CA--CC-GTG-----CA---C--A--------CAGAGCATCAACAAGGGACTG-AACA-TT-AA---CTTTAGT-C---CCTA..............................AA-GGAG-A---ACA--C---G-AAGA--GA--C 8389
CPZ.CM.05.SIVcpzMT145   --AA----------GTGCA-CTC-GA-TTG---AC-AC----GA-A------CAATACAGCCTCAGGGGAGTA-AACAAATAG-TACTCT---AC-CC-ACT-..............................A-AGGGAC---G----------CAACA---A-C- 8179
CPZ.TZ.00.TAN1          GAC---TCA-A----CA---G---CC-CTG--TC-GCC--C--GCA---AA-CAGGGGATAATCAGCTTGGCA-A---CCTAGTAAT--TG-AT-G-AC-ATCATAGTAGGAGTTAGACAGATC............ATTG-GTGGAGCAGTA--ACTTA-GCTAGCT--- 8344
CPZ.CM.05.SIVcpzLB7     ---A----C-------G----A--A-CG-G-----CAT----GA--G-----CAATACACTCTGAAGGGACTGA-ACT--T------------ACT----CT-..............................CA-GGAA-T------------G-AA----G---- 8178
CPZ.US.85.CPZUS         ---A---GAG--------A--A---C-G-G-----CA--AC--A-C------CAGAACATCAACAAGGGACTG-AACT-TT-AAT---CT-AGAGC-CGC--C..............................---GGGG-AA-CGCT--C---GCAAG---G---- 8675
CPZ.GA.88.GAB1          GA------C-AG---GC--GC-A---GTG---ACAAC-CA-GACGT-----CATCTAATACTTCACAGCCTCA-ACT-CTAC----GAGA--GTGCCTGCT-..............................G--GGAA-TA-T--A------G-AA----G---- 8707
MON.CM.99.L1            GAT-----A-----C-G-A--AG-GA-CTACC----CC-CC-GAGAG--TCCCAGAGGCTGCCACCCCTCCTG-A--TAC-CCT-CTCCAA--GTGGG-AA-CCTTAGACGCTTGCTAGCCTACTGCCAGTAT.....................G---TC--A-----CC 8639
MON.NG.x.NG1            ATTGAG-GACTC-GGCT-GAC---CTACCG---CC-CCA-AG-G-G-GCA-CCACCTGATCCCTCTTCTCCGCGA-CT-T-CC-AC-GCTT-GGG--GCCA-TAGCCGCCTTCTT........................GCCTA---G------G--CTC--A--GY-CC 7264
GSN.CM.99.CN166         GAT----GCCAG----GC--G--C-C-CCGTG----C-G-A-GT-C---T-GGAACAGCTGCAGAAAGCAATC-AG-AA-CAATCC-GCAAGCG-G-C-GCTCAGGGAAGTAGCAGCCAGAGTG............ATAGCCTA-A-TAGC---G---TC--A-----CC 8622
GSN.CM.99.CN71          GAT----GCCAC----G---G--C-C-CCA------C-GCA-GT-C---T-GGAGCAG............CTG-AG-A--C--TCC--CAAGCG-G-C-GCT-GGGGAAGTGGCAGCCAGAGCA............GCAGCCTATA-TAGC---G---TC--A-----CC 8622
MUS_1.CM.01.1085        GAC---A-T-AG--C-GC--CC-CCC-TTG-T--C-AC-TC-GAGAGCA-TCCACCTGATACAGAAGGGGTCAATA--ATT----TTCA-G----GGGCACT-ATTGAGCCTCTCAGGGAAGCCTTGCAAAGAGCCCTCGC-TG-T-----C--G---TC--A--G--CC 8631
MUS_1.CM.01.CM1239      GACA--AGTCAG--C-G-A-CA-CCC-TCGCT----AC-CC-GAGGGCA-CTCACCTGATACAGAAGTGGTCAGTAC-ACT----ATC-TGA-T-GGG-ACAAGCCCGCCGCTGCAGTGAACTCTTAGCGCGCCTC-TAGCATA-A-T--A---G---TC--A-----CC 8600
MUS_2.CM.01.CM1246      GAT-----CCAG--CCGC--C--ACC-CTG-T--T-GACTC--AGAA-CTACCACCACCTGACAAAGGGGCTAGTA-ATCC----ACTCTG--T----CC-TTAGGGAAGCCTTAGGGCGCTTCATCGCACAC............GC---A---G---TC--AA----CC 8696
MUS_2.CM.01.CM2500      GAT-------AG--CT----CA-ACC-TTG-T--C-GA-CC-G--GA-A-TCCATCACCTCCAGAAGTGGAGCAT-TCTCT----CTCAGGGCG-G-GCAACCAGCACCCACCTTCTGGAAGTCCTTAGACGCTTGCTTGCATAT--T--A---G--CTC--A---T-CC 8652
RCM.NG.x.NG411          AGAGA-CCTTTG-AGGAGGC-CTA-CA--TCT-C-GCAACTC---CTCA-TGATCTACCAGAGCCTCCAGTCAAT-CCTTCTC-AATCCTT--TC-T-TGC--TGGATAGGAGCT...........................AAG-----AG--G--T-G--A--GT-C- 8386
RCM.GA.x.GAB1           GA--AG--TTTGGACG-AGC-ATCGCAG-TTT-C-GAAGCTGCAGCTCAAT-ATCTTCCAGAGCCTCCGGTGGATTTTA-CT...........................................AAGA---A--G--T-G--A--GT-C- 8303
DEN.CD.x.CD1            GAC-------A-----G-G--TC-GA-T-G--CCGCCTGTAGT-GGA-TT-GGACTTTGTA.............................-AT-GGC-CA-CTTAATTTTGCTTGGC.......................TA-T-C--A---G--CTC--A-----CC 8706
                                                                                                                     Rev end                                             Nef start
MAC.US.x.239            ATT-AGC-ACTG-AG-AC----C-ATCG-GAGTATACCA-ATC--C-AACC-ATACTCCAGAGGCTCTCTGCGAC-CTACA-A--ATTCGAGAAGTCCTCA--ACTGAA..........................CT--CCTA------A---G--T-GAGCT-TT-CC 9094
Env                       F  S  N  C  R  T  L  L  S  R  V  Y  Q  I  L  Q  P  I  L  Q  R  L  S  A  T  L  Q  R  I  R  E  V  L  R  T  E  .  .  .  .  .  .  .  .  .  L  T  Y  L  Q  Y  G  W  S  Y  F
Nef                                                                                                                                                                           M  G  G  A  I  S
Rev exon 2               I  Q  Q  L  Q  N  L  A  I  E  S  I  P  D  P  P  T  N  T  P  E  A  L  C  D  P  T  E  D  S  R  S  P  Q  D  *
H2A.GW.x.ALI            GTA-AGC-TCTG-AGGGA---AC-ATCC-G--TCTCCCC-ATC--C-AACCGATCTTCCGGAGTCTTCAGAGAGCGCT-ACAACAATCAGGGA-TGGCTGA-ACTTAAA.........................GCAGCCTA---G------G--T-CG--TGGA-CC 9133
H2A.DE.x.BEN            ATA-AAC-TCTG-AG-GA---AC-ATCC-A--ACTCCCC-............ACCCGCCGACTGATCTCCCAGA-TCTAACAGCAATCAGGGA-TGGCTGA-ACTTAAG.........................GC-GCC-AA--G--A---G--T-CG--TGGA-CC 9137
H2A.SN.x.ST             ATA-AAC-TCTG-AGGGA---AC-ATCC-G---CT-CCA-ACC--C-AACT-ATCTCCCAGAGTCTTCGGAGAGCATT-ACAGCA-TCAGAGA-TGGCTGA-ATTTAAC.........................ACAGCCTA---G--A---G--G-CG--TGGA-CC 8579
H2B.GH.86.D205          ATA-AACG-CTG-AG-AC---AC-GT-G-A--CCT-CCAA.......................ATCCTCCAC-AA-TCTCAACCA-TCT--AACC-CTCA--CTCCCG.........................GTTGCGTA---C------G---TCAGCTGGT-CC 9101
H2B.CI.x.EHO            ATA-AGCG-CTG-AG-AC---AT-ATC--A--CCT-CCAAACC--C-AACC-GTACTCCAACCGCTCAGGCTT-CACCT-CA...........................................TA---C-GA---G---TCAGCTGGT-CC 9078
H2G.CI.92.ABT96         ATA-AGC-ACTG-AG-GC---AC-ATCC-G--TCY-CCA-ACT--C-AACCGCTGTTCCAAAGAATCTCCAGGAC-CTCCAAGCAATTAGAGAAC-TCTCC--CTAGAA.........................GCAGC-TATT-CAGC---G--TTC-GCTGG--CC 8522
H2U.FR.96.12034         ATA-AGC--CTG-AGGGA----C-ATCC-G--TTTGCCA-AGC--C-A-CCGTTGTTC.....................CA-A-CATCAGAGAACGTCTGCACCTTGAA.........................ATAGC-TA------A---G--T-G---T-TT-C- 8612
MAC.US.x.EMBL_3         ATT-AGC-ACTG-AG-AC----C-ATCG-GAGCATACCA-ATC--C-AACC-ATACTCCAGAGGCTCTCTGCGAC-CTAC-AA--ATTCGAGAAGTCCTCA--ACTGAACTGACC.........................TA------A---G--T-GAGCT-TT-CC 8225
SMM.US.x.H9             ATT-AGC--TTG-AGGGAT-G-C-AT-G-G--NCTGCCAAATC--C-A-CC-GTGCTCCAGAGCCTCTCAAGGACGCT-CA-A-A-CCCGTGAAGTCATCA-AGTTGAGATAACC.........................TA----------G--T-G-GTT-C--CC 8583
SMM.US.x.PGM53          GTT-AAC-ACTG-AGGAC--ATC-CT-G-GA-TATACCA-ATC--C-AACC-GTGTTCCAGAGGCTC.........TTACA-A-AATCCGAGAAGTCGTCA--ACTGAG.........................ATATCCTA---G--A---G--T-C-GCTGGT-CC 9031
SMM.US.x.SIVsmH635F_L3  ATT-AGC--CTG-AGGGAT-G-C-AT-G-G--TCTACCA-ATC--C-AACC-GTGCTCCAGAGACTCTCAAGGACGCT-CAAA-A-TTCGTGAAGTCATCA-AATTGAA.........................ATA-CCTA------A---G--T-GAGCT-TT-CC 9116
SMM.SL.92.SL92B         ATA-AAC-ATTTGACT-G---AGCATCC-GAGCCTACCA-A-C--C-AACA-CTGTGCCAGAGACTCTCAGAGAT---TCA-CC-ATTAGAGAGCTTGTCA-AAGAGAAGCAGGC.........................TATA-C-GT---G--T-GA-TT-CT-C- 8542
STM.US.x.STM            GTA-AAC-ACTT-AGGGCT-GTC-ATCC-G--TCTACCA-ACC--C-ACCC-ACCTTCCAAAGGATCTCCAGGATACT-CA-A-AATTAGAGAAGTCGTCA-ACTAGGA.........................GCAGCCTATT--------G--T-CATCTGGA-CC 8745
MNE.US.x.MNE027         ATT-AGC-ACTG-AG-AC----C-ATCG-GAGCATACCA-ATC--C-AACC-ATATTCCAGAGATTCTCCACGAC-CTACA-A-A-TCCGAGAAGTCCTCA--ACTGAA.........................CTA-CCTA------A---G--T-GAGCT-CT-CC 8571
DRL.x.x.FAO             GAAGGAATCGAGAGG-TA-A-CTGGC-G-A--ACG-GAAA-C-G----T-A-TATGGGTGGTCAGAA.......................................----A-G-AGCGG-A--AGGCAT-TGC. 8459
MND_2.x.x.5440          G-TGAAA--GTCGACG--T-AC---TGGCT--A--ACC-CC--GCAG-AATTGAATATGGGTGGCAAGAG.......................................---AA-GCAGCGG-AGCAGCATT-TATC 8399
MND_2.CM.98.CM16        G-TGAA--CTTCAAGGGG--ACG-GTGGCT--A--A-C-AC--GCAG-AATTGAATATGGGTGGCAAGAG.......................................---AA-ACAGCAG-AG-AA-AAT-T-T- 8763
MND_2.GA.x.M14          GTTGAA-GAGTCAAC-A-T-AC---TGGCT--A--A-C--C-GGCAG-AATTGAATATGGGTGGCAAGAG.......................................---AA-GCAGCGG-AGCAAG-AT-TGCC 8692
OLC.CI.97.97CI12        --TG-G--AG-GAG--G-GAGAG-CACTCGT...........................................TGGATC...CA--AGAGCTTG-A-C 8189
GRV.ET.x.GRI_677        -A--------ATAAGGACT-G-TG-TAC-AC-TCTGCC--ACC--C--CAT-TTCATCAGGACAGCAGTGGGA.......................................TA---G------G--CTC---C----CC 8617
VER.KE.x.TYO1_patent    ................ACT-G-C-ATAC-AC--TTGCC--ACC--C--CATTCAGCTTAGGAAAGCTTTTCAA.......................................TA---G--A---G--CTCGCA-----C- 8725
VER.KE.x.9063           G-T--C--CGTT-AGCAC--G-C-ATAC-AC-CCT-CC--ACC--C--AATTCATCTTAGAGGCGCTTGGCAG.......................................TA-A-C--A---G--CT-GG------C- 8764
VER.KE.x.AGM155         GTT-----C-CT-AGCA---G-T-ACAC-AC--CTGCC--ACC--C--TCTCAAGCTTAGAAGCGCCTGGCAG.......................................TA-T----A---G--CT-GG---G--C- 8741
VER.DE.x.AGM3           GT---C--TCAG-AGCA---G-C-ATAC-AC--TTGCC--ACC--C--A-TTCATCTTAGGAGCGCTTTCCAG.......................................TA-A----A---G--CT-GG------C- 8262
TAN.UG.x.TAN1           G--GCA-CATTTGACCAAT-G-TGCT-G--A-TC-GCAGTTGGT-AT--A--CTTTACCAGATCCTCCGCAGGA--CTCAC-ACTCTTCTT-AACTGCT-A-ACAGGAG.........................--CC-ATA-A-T------G--T-G---C--T-C- 8757
SAB.SN.x.SAB1C          GT-TCAC---TTGA-CACT-G-C-CAGG-AT...T-GACCAGTTGGT-CTCGACAATCTTCAGCAACCTCCAT---T-CCTCCA---CAT--AAC-GAGAACCAGACAGCTAACAGCT........................-AT--TG-A---G--T-G--A--G--C- 8938
MND_1.GA.x.MNDGB1       TGC-AG-GAGTGACC-A-T-G-C--TGGCT--AA-A-C--AG-GCAGCAATTGAATATGGGTTC.................................CTCGCAGTCCAAGAAGCG.........ATCAGAAGC-T-GG-TCG----TCGTCA-CTT-GC--C-A-T- 8238
DEB.CM.99.CM40          GAC---TC-CA---GCA-T-G-C-GA-T-GA--AGGCT-T---T-A---CA-GAAGGAATATATAGACTGTCATTTGCTCTT---C--AGA--G-G-CTC-T-GGGGAACGCTTTAGAGTGCTT............G--GGGTA-T---------G--T-GA-A---T-C- 8410
DEB.CM.99.CM5           AAC------GA---GCGC--G-C-GA-T-GA--AGGCT-T--G--C---CA-GAAACAATATATAGACTTTCATTTGCATTT---C--AGA--G-G-CTC-T-GGGGAACATTTAAGAGTGGTC............---GG-TATTGT--A---G--T-GA-A---T-C- 8331
TAL.CM.01.8023          GAC-------A-----G-T-G-C-GAGTTG-GACC-CC--A-G-GC---AA-GCGACCACCCAACAGGCAGCG-AAGCCACA-AC-CCTA---G-G-AT-AAATTCACAGCAGCAGTA.....................GCAACAA-C------G--TTCAGT---A-CC 8056
TAL.CM.00.266           GAC---AGCCA-----GCT-G-C-GA-TTG-GACC---C-A--GC-GC-AA-GCCAAGAGCCAACAGGTAGCG-AAGCCACA-AC-CCTT---ACG-CTCCACCTGACAGCAGCAGCAGCGACTCTCCAGTAT.....................G--TTCAGT---A-CC 8506
SUN.GA.98.L14           T-TG--AGTGACTC-T---AA-TGGC-G-AAGCTTGG-GAAT-G-C--A-C-CAAAACATCTACCACCTGCTAA-TTTCTT----C-CCTA--T-CGAC--CCTTCCAT...CATGGGCAACGCCTTTGGCAGACCCTC-GAGGTTGG-T-GG--C-CAC-TTATT-AGG 9017
WRC.CI.97.97CI14        --AACACCTCCTTCACA-TG--TGGAGTCT-CTTTCGTGCCTC--GTG-T-GATAAACGATGGGTGG........................AGTCTTCTCAAAGACATC.........................................AG-G-A--AGCAT-TGG-G 8759
WRC.CI.98.98CI04        --AAACCCTCCTTCA-A-AG--TGGAGTCTCCTTGCGTGCCTCT-CTG-T-GATCCAGCATGGGTGG........................AATCTTCTCAAAAGAATC.........................................AG-G-AGCAGC-C-TGC-T 8804
WRC.GM.05.Pbt_05GM_X02  --TGAGCCTCCTGCAG--AAGCTACAGCCTTCT--GCTATCTGTGGA-A---ATACAGAATGGGTGG........................AGTCTTCTCAAAAGACTC.........................................-G-G-AGCAGC-T-TGCTG 7700
LST.CD.88.524           A-TG--TGTGACTC-C--GAA-TGGC-G-A-GC-TGG-GAAT---G--CCTCACCCTCGTGTGGCAATTCCTCACTT--CT-G-AC-TATAG-G-TCCT--CTTACCAC...CATGGGCAACATCTTTGGCAAACCCTC-GCAGAT-C-T-GTGGA----GCTC-GAA-- 7883
LST.CD.88.447           A-TG--TGTGACTC-C--GAA-TGGA-C-A-GCCT---GAATT-----CCTCAGCCTTGTGTGGCAGTTCCTTTCGT--CTAG-CC-TTTAG-G-TCCT--TCTTCCAA...CATGGGCAACATCTTCGGAAAGCC--C-GCAGA-GGGT-GTGGAA-ACGCTC-GAAG- 7880
LST.KE.x.lho7           G-TG--AGTGACTC-C--GAA-TGGC-C-A-GAAGGG-GAAT---C--CCT-TCACTAGTGTGGCAGAGCCTCA-TT--CT----C-TTTGT-G-TCCT--TCTTCCAA...AATGGGCAACGCCTTTGGCAGACCA-C-GCAGATGGAT-GTGGAA-ACGCTC-GAAG- 8955
SYK.KE.x.SYK173         AGT--GC--CAG---CAGA-CTC-GA-T-GACA-C--ACTATCACCTG----..........................................................T--A-CAGC---G--TTCA-C--G--C- 8574
SYK.KE.x.KE51           AGTT-C-GAGAG--GCAGA-CAC-GA-T-GAG-AC----AATTGCCTG----..........................................................T--A-CAGT---G--GTCACA--G--C- 8273
COL.CM.x.CGU1           ATTGA--GA---G-GGTGGAGC-GCT-GTT-C-TCA-TGCAGCAA-GG-ATCCTTACT...............................GTCCTGCCTCAACATCCGCCAACAACCTCA............GAC--GAGGAGGAGTGC----.....---ACT--TC 8008
```


PLV Complete Genomes

```
                          Rev end                                                                                                                                                     Env gp41, gp160 end
H1B.FR.83.HXB2            AGAATAGTGCTGTTAGCTTGCTCAATGCCACAGCCATAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAGTAGTA.........CAAGGAGCTTGTAGAGCTATTCGCCACATACCTAGAAGAATAAGACAGGGCTTGGAAAGG......ATTTTGCTATAAG..... 8796
Rev exon 2                  E  *
Env                       K  N  S  A  V  S  L  L  N  A  T  A  I  A  V  A  E  G  T  D  R  V  I  E  V  V  .  .  .  Q  G  A  C  R  A  I  R  H  I  P  R  R  I  R  Q  G  L  E  R  .  .  I  L  L  *
H1A1.UG.85.U455           -A-T-------A---CT-----TG-C--TGT---AG-----------G-T---T---------------A---G-..........---AC-AT-G-----------T-A---------------C-----A-----A---------.....GC---------C..... 8247
H1B.US.90.WEAU160         ----------------T-----T-----------T--------------------------------A---G..........---A--A--------------T---------------------------A---------.....GC-------------..... 8803
H1C.ET.86.ETH2220         -A--G------A---ATC----T---A-------A----T-----G---A--A--------AT--------T--A--..........---A--AT---G-----T-CT--A----------------------------GCA.....GC-----A----A..... 8193
H1D.CD.84.84ZR085         -----------A--------G-TG---A----A-----------------------A--------CA----.........AGGA---------A----G-G-TT--------C-C---------------T--A---------.....GC---------A..... 8317
H1F1.BE.93.VI850          -----------A-C---C--T-T---A-------A----T---------A-----------AA----------TT-G..........---A----G-------G---T-A--------------------GCA----------.....GC------G---A..... 8073
H1G.SE.93.SE6165          -----------A----T-----TG--A-AGT---A-----------A-TT-----------------------C-..........---A---------------C-T-A--------C---G--------A-----------A.....GC---------A..... 8235
H1H.CF.90.056             -----------A--GAT-----T--CA-------A-----------------A------G--A------T-A---G..........---A------G---------T-------------------------T-----A.....-GC---------TA.... 8116
H1J.SE.93.SE7887          -----------A----T-----T---A-------A--------------A--A--------AA-C------A---C-..........---A------T----------TT-----------------------T--A-----A.....GC----------A..... 8103
H1K.CM.96.MP535           ----------CA----------T---A-------A-------------GA--A--------AA--------A---G-..........---A------T-------T-A-TT---------------------TC-C-----A.....GC---A------A..... 7956
H101_AE.TH.90.CM240       -A-T-------A--TCT-----TG----T-----A-------C---G-G-T-------C----------------C-..........-----------G-----C----T--------------C-----------A---------.....-C---------C..... 8370
H102_AG.NG.x.IBNG         -----------A--AT------T---A-A-T---A----T------A-CT----------A-C-------A---G-..........---A---TGG-------------A-------------C-----------C---------.....GC---------C..... 8323
H1N.CM.95.YBF30           G-G--------A-C--------T---A-A-----T--T-T------A--A--A--------A---------T---C-..........---A--ATAG-A--G-GA--ATTA--------------C-----A---C-A-----A.....GCAC--A--------..... 8368
H1O.BE.87.ANT70           -----------ACA--------AG-CA-ACTT--AG-G--------CA-TT----T--CG-CA-A--C-C--GGA--..........---A--ATAG-A-C--GA-----TA----C--A--G-----T-----------A-----A......-G---AT-G---C..... 8865
H1O.CM.91.MVP5180         -A---------ACA-A-C----TG--A-T-TT--AG-GT----T--CA-TT----T--CG-CA-C--CTT--GTC--..........---A--ATAG-ACA--GAT-C-TT-----C--A-------T-----A--TGCA-----A......--C--AG-G---C..... 8891
CPZ.CD.90.ANT             GA-C-GAGATCA---AGAACA-TG-CAGACTT--T--TTGG----GG--AAAA--------CA-AC-TCT--CTC-C..........---ACTATAGTC---ATC--AA-GG-AG-------GC-C--C-----A--G------TT......GCA--AAAT---AC.... 8256
CPZ.CM.01.SIVcpzCAM13     GA-C-------ACA-----AT-AG----A-----T--T-------GA--A--C--------AA-C---AG--C---G..........---AT--TA-TC---ATC--AG--A-------CC-GC-C---------A---C-A--G--A......-C---A--T---ATCTTG 8406
CPZ.CM.05.SIVcpzEK505     -A-T-------A-A-A------TG--A-T-----TG-T--------A--A-----------AA-C-----GC----G..........---A--ATAG-AC-G-GA--A-TA-----------G-----C-----A--TC-A-----A.....GCAC-TT-----A..... 8303
CPZ.CM.05.SIVcpzMB66      -A-C------CA-A--------TG-CA-A-----T-----------A--A--A--------AA--------A---C-..........--GA--TT-G------GC--A-TGA----C-----------C-----A---C-T-----A.....GCA--AT-----A..... 8441
CPZ.CM.98.CAM3            -A-C------CA-A-----AT-AG----A--------T-----------A--A----C--AA--C-------T-C-..........---AT-ATAG------GC----TG--T-----A-------T-------T--A-----A.....-GC---------C..... 8233
CPZ.GA.88.GAB2            GA-C-------A-----C----TG--A-T-----A--TAT------C--AA-A--T----CAA-A----------C-..........ACTA--ATAG-G---GC--ATTG--T-----A-------C--G--A--T--A-----A.....GC---A---G--TAAA.. 8207
CPZ.CM.98.CAM5            -A-T-------ACA-----AT-AG--A-A-----A--T---------A-T--A--------AA-------GT-T---..........---AT--TAG-G---GC----TA--T---------G-----T-----A--T--A-----A......-G----T-----C..... 8539
CPZ.CM.05.SIVcpzMT145     ----------CA-A-----AT-AG--A-T--C--T--T-------GA--A--A--T----AA-CT------C-A---..........AC-C-CTTGG-C--G-GC---TTA--T-------------T-----A-----A-----A.....GC------T-------..... 8329
CPZ.TZ.00.TAN1            GAGT-TTGCTAA-ACAAGCCA-AG-CAGACTT--T-ACTTTAC--GGTG-T--------TTAA-C-------G---G..........GTTTACATAGCC--G-GA--CA-AA-T--T-------------T--------TC-----CTA.....GCC--AAAT---A..... 8494
CPZ.CM.05.SIVcpzLB7       -A-C------CAC---------TG--A-A-----A-----------------A--------AA-C------A-----..........---A--AT-G------GA--ATTA--T-------------G--T-----A--T--C-------.....GC--AT-----A..... 8328
CPZ.US.85.CPZUS           -AGT-------ACA-----AT-AG--A-A-----A--------------A--A--T-----AA--------T--AC-..........AG-A-GCTC-T-CT--G----ATA--------A----------T--G--A---C-A--G--A......-GC--AT--------..... 8823
CPZ.GA.88.GAB1            -A-T-------A-A---C----TG----A-----T--T----------A--A--T---------AA-A-------CTT-T..........----TTA-ACT----AT----A-AA----C-----GC-C---------A--GC---------A.....GC---A--T----CACTG 8862
MON.CM.99.L1              -AGCAGCA-TCACCTCTC-C---G-C---CTT--T-GGTTCAC-ATC-TCT----G---GC-C-AC-TC-T-CT-G-.........GG-C--CTA--G-----A--CGTGGCT--C--C---C--C---C-CC-------AGC-----TA.....T-CC-TAAC-G-...... 8788
MON.NG.x.NG1              -ARCAGCCTGCACAG-AGCCA-TG-----CTC--ACGGTTCAC--TCATCT-------CGC---CC-TCGCC-G-G-.........GG-C--TTA--GC---GAC-CGTGGCTG-C--C---C--C---C-CC-------GGC---G-TT.....C-CC-TAAC-G-...... 7413
GSN.CM.99.CN166           -AGCAGCA--AACAG-AA-C---G-CAG-CTT--T--CTTCACGTGGA-CT-------GGCC--AC-TC-C-CTTGT.........AG-C---TA--GC---AAT---TGGCT--C--CC--C--C---C-----A--GGC---G-TA.....C-A--TAAC-G-...... 8771
GSN.CM.99.CN71            -AGCAGCA-TCTCAG-AA-C---G-CAG-CTT--TG-CTTCAC-TGGA-CT-------GCCCC-CG-TC-GACT--T.........GG-C---TA--GC---AAT-CTTGGCT--C--CC--C--C---C-------GGC---G-TA.....C-AC-TAAC-G-...... 8771
MUS_1.CM.01.1085          -A-CAGCCAGCACAGCAG-C--TG--CAGCTC----CCTTCACCTGGA-CT----G--A-CA--AC-TC---CT-GT.........-G-A---TA--GC---AAT-CTTGGCT--C--C---C--C-CG-CC-G-----GGC---G-TA.....C-CC-TAAC-G-...... 8780
MUS_1.CM.01.CM1239        -AGCAGCA-GAACAGCAGCC--TG-CCAGCTC-----CTTCACCTGGA-CT-------GCCCA-AC-TC-GA-T-GT.........AG-C---TA--G----AAT-C-TGGCT--C--C---C--C---C--CC-------GGC-----TA.....C-CC-TAAC-G-...... 8749
MUS_2.CM.01.CM1246        -AGCCGCA-TCTCAG-AGC----G-C---CTT--G--CTTCAC-TGGA-CT----C--ACC-A-AC-TC-GT-TAGT.........---AA-ATA--G---AAAT-C-TGGCT--C--C---C--C---C--CC-C-----GGC-----TA.....C-CC-TAAC-G-...... 8845
MUS_2.CM.01.CM2500        -AG-AG-A-TCACAGCAGCCA-TG---GGCTC-----CTGGACC--AA-CT-------CGCC--AC-TC----TAGT.........A--CA-CTA-AA-A---AT-CTTGGCT--C--C-A-C--C---C--CC-C-----GGC-----TA.....C-CC-TAAC---...... 8801
RCM.NG.x.NG411            -AGGCTTCAGCAGCT-GC-TGCGG-GATGG-GAGG-CTAAT-CTTACT-CACCTGGAGAG-AC-ATGC-CG--TTGC.........AG--ACC--GCAG--TGGCCAGC-ACGT-GTGC-----G--C--G--------C--G--A......T--C-TAAC--G...... 8535
RCM.GA.x.GAB1             -AG-ATTCAGCAGCT-G--TGCTG-GATGG-GCTTCA-AAT-CTTACT-CACCTGGAGAG--T-ATGT-CG--G-C-.........-G--ACTT-GC-G--TGGCCAGC-ATGG-GTGC-----G--C--G------C-C--G--A......C---GTAAT--G...... 8452
DEN.CD.x.CD1              -A-CAG-CCTCAGGGAT---GGG-C-AG-G-CAT-CA-CA--G-AGG-C-ACAG-G--AGT---AT-G-CG-CTC-T.........ACCC----AGCA----AAG-GGTGGC---C--C---C-G--C---------C-T----TT.....G-CC-CAAC--G...... 8855
                                                                                                                                                                              Env gp41, gp160 end
MAC.US.x.239              -TG-GGCG-TCCAGGC-G-CTGG-GAT-TG-GA-AGAGA-TC-T--G-GC-C-TGG-GAGACT-ATGG--GACTC-T.........AGGA---G-G-A---TGG--A-T-GCA--C--C--G--G--T----A--GC-T--GCTC......-C-C-CT-G-G-...... 9243
Env                       H  E  A  V  Q  A  V  W  R  S  A  T  E  T  L  A  G  A  W  G  D  L  W  E  T  L  .  .  .  R  R  G  G  R  W  I  L  A  I  P  R  R  I  R  Q  G  L  E  L  .  .  T  L  L  *
Nef                        M  R  R  S  R  P  S  G  D  L  R  Q  R  L  L  R  A  R  G  E  T  Y  G  R  L  L  .  .  .  G  E  V  E  D  G  Y  S  Q  S  P  G  G  L  D  K  G  L  S  S  .  .  L  S  C  E##  .
H2A.GW.x.ALI              -AG-AGCGTTCCGGGC---TGCA-GGATTG-GAGAGAGA-TC-TA-AA-CACCTGGAGAGACT-GTGG-GG-C---G.........--GT-G-TCG-G--GAGG--A-T-GCAG-C--A--G--G--C--G-----GGCA---TT.....GCCC-C--G-G-...... 9282
H2A.DE.x.BEN              -AG-AGC-TTCCAAGCA--CGCG-GGA-T--GAGAGAGA-TC-T--G-GC-C-TGG-GAT--T-ATGG----C--CG.........-G-C-CATCG-G--G-GA--A-T-GCAG-T--A--------C--G-----AGCA---CTC.....GCCC-C--G-G-...... 9286
H2A.SN.x.ST               -AG-AGCGTTCCGAGC---CGCG-GG--T--G-GAGAGA-TC-TA-AA-C-CCTGGAGAG-CT-CTGG-GGAC-C-G.........GG-CA-AT-G-G--G-GA--A-TTGCAG-C--A-----G--C--G-----GGCA----TC.....GCCC-C--G-G-...... 8728
H2B.GH.86.D205            -AG-AGCGCTCCGAGCGGCAGCA-GG--T--G-GAGA-A-TC-T--GAGT-CCGG---G-CCT-ATGG--G-CTC-C.........AGGA----GGCG-----A--CAT-GCG--C--C--G--C--G-----C-----A--AC-T---CTC......-CCC-CT-G-G-...... 9250
H2B.CI.x.EHO              -AG-AGCAATCCAAGCAGCAGC--GG--TG-G-GAGAGA-TC-T--GAGC-C-G-GAGG-CCTCGTGG-G---CC-C.........AGGA----GGCAG---AG--CATTGCA-----C--G--G--G--C-----A--GGCT---CTC.....GCCC-CT-G-G-...... 9227
H2G.CI.92.ABT96           -AG-AGCGTGCACCGCAGCAACT-GG--TG-GCAAGAGA-TC-TA-AAGCAC-TGGAGAGCCT-ATGGA--ACTC-G.........GG-A-G-T-G-A----GG--A-T-GCA-----G--------C--G--A--GC-T--GCTC......-C-C-CT-G-G-...... 8671
H2U.FR.96.12034           -AG-AGCATTCCAAGCG--CGGG--A--TG-GCGAGAGA-TC-TAGCCGCAC-GGGAGAGAAC-ATGG--GACTC-G.........GGGC-G-T-G-A---TGGC----AGCT-----A--G--G--C--G--A--A--T--GCTC.....GCGC-TT-G-G-...... 8761
MAC.US.x.EMBL_3           -TG-GGCG-TCCAAGC-GGCTGG-GAT-TG-GA-AGA-A-TC-T--G-GC-C-TGG-GAGACT-ATGG--GACTC-T.........AGGA---G-G-A---TGG--C-T-GCA--C-----G--G--T--G--A--AC-T--GCTC......-CGC-CT-G--------..... 8374
SMM.US.x.H9               -AG-AGCA--GCARR---G-TGG--ATTTGYRCGAGAGA-TC-T--AAGC-C-TGGAGAGACT-ATGG--GACTC-G.........GG-A-G-T-G-A--G-GA--A-T-GCA--C--MR--C-C--C--G--A--GC-T--GCTC......-C-C-CT-G-G-...... 8732
SMM.US.x.PGM53            -AG-AGCA-TCCAGGCAGC-TGG-GG--TCGG-GAGA-A-TC-T--AAGC-C-TGGAGAGACTCATGG---ACTC-G.........GG-A-G-T-G-A---TGGG-A-T-GCA--C--G--G--T--G-----AC-----CTT.....GCAC-C--T-G-...... 9180
SMM.US.x.SIVsmH635F_L3    -AG-AGCA--ACAGGCG-G-TGG--A-TTG-GCGAGAGA-TC-T--GAGC-C-TGGAGAGACA-ATGG--GACTC-G.........GG-A-G-T-G-A----GG--A-T-GCA--C-----GC-CG-C--G--A--GC-T--GCTC......-C-C-CT-G-G-...... 9265
SMM.SL.92.SL92B           TCG-AGCCTGCCAGGAAGCATGG-GAT-TG-GCAAGA---TA-T-TC-GC-C-TGG-GACTCA-ATGG--GACTC-G.........GG-C-G-T-G-A----GG--CGCAGCA--C--A--G--G--C---------C-T---CTC......--GC-TAAC--G...... 8691
STM.US.x.STM              -AG-AGCA--GCAAGCAGCATGG-GA--TG-G-GAGAGA-TC-T--GAGC-C-GGGAGAGACT-ATGG---ACTC-T.........GG-A-G-T-G-G---AGG--CG-GGC---C--A--GC-C--C-----A--GC-T---CTC......-CAC-CT-G-G-...... 8894
MNE.US.x.MNE027           -AG-AGCG-TCCAAGT-GCCTGG-GAT-TG-GA-AGAGA-TC-T--G-GC-C-TGG-GAGACT-ATGG--GACTC-G.........GG-A-G-T-G-A---TGG--A-T-GCA--C-----G--G--C-----A-AGC-T--GCTT......-C-C-CT-G-G-...... 8720
DRL.x.x.FAO               ...-CGCT-CTACAAGAGG-TGCGC-GAGG-T-TGGCGC-A--GGAGGC-ACTTGG-CTATCATCAGCTAG--CCC-C.........-G--CCCT-GC-CAG-AAG-GGCAGC---T--C--G--G--T-----A---GCA--GGTC.....C-C--CAAC--G...... 8606
MND_2.x.x.5440            -AGTATTACAA-GCTTTGC--AG-GGCTATGGAG--G--GGTACCAACTT--ACT-TCAG-T----G--G--C--C-............-CCTT-G-C--G-GC--CT-GA----C---C-CC-C---C-CC-C--C-----A--TGCT--GGCC.....C-AC-CAAC-G-...... 8545
MND_2.CM.98.CM16          -AGTACTACAA-GCTATGC--AGGGGCTATGGAG--G--GGCACCAATG---ACT-TCA-CA-C-GC-TGCT-CAG-............-CCAT-GC---G-GC---AT-A----C--A--------TC-------AGCA---GTC.....C-CC-CAAC-G-...... 8909
MND_2.GA.x.M14            -AGTATTACAAAGCTTTGC--AG-GGCTATGGAG-GG--GGCACCAACTT--ACT-TCAGTTA---G--G--C--CG............-CTAT-G-A--G--C---G-GA---C---C-GC-C---C-GC-C--C-----A--AGCT---GTC......C-GC-CAAC---...... 8838
OLC.CI.97.97CI12          -CGTCT-GAGAACCTTAAAA-AG--GAGACA-.................CATGGGATC-ATT-GCAGC-GAAGAA-..........--GA--AGGG-A--TC--C-A-AGA-GGCT-AGGAGCA-T-GG-GAGA-AAGGA-C-GAA...ACA-C-GGAG-C-G-...... 8323
GRV.ET.x.GRI_677          -AG-GGCA--AACAG-GC-TGCTC-A--TCTG--G-GG--T-CGAGG--A-CCTGG-GC--AC-GGGT-CTA-T--C..........-G-TCC----A-C-G--AG-CAT-A---GT--A------G-GC-G--A---C-T----AA.....G-CC--GGG--......... 8766
VER.KE.x.TYO1_patent      -A-CCG-C--ACAAGAAA-A---C-AA-TCTG--AGGC-TT-C-CAAA-C-CATGTC-CCA-A-ATGGCTT-CTTGC.........AG-TCC----A---GAAC--CGT-A---GT--A------G-GC----A---C-T--GGAA.......--CC-TAAT--G...... 8874
VER.KE.x.9063             -A-CCGCA--ACAAGAAGCA-GTCTC--TGTG--ACGCTTT-C-CAAA-C-CAGGTC-CCAAA-ATGGCTT-CTTGC.........AG-TCCA---A---GCAC--CAT-AG--GT--A------G--C----A--GC-T--GGAA.......--CC-TAAT--G...... 8913
VER.KE.x.AGM155           -AGCCGCA--ACAAGAAGCAG-T--CCATCTG--G-GCTTT-C-CGCA-C-C-G--C-CCA-A-ATGGCTT-CTTGC.........AG-TCC----A-C-G--A--CAT-A--TCT--A------G-GC----A--GC-T--GGAA.....G-CC-TAAT--G...... 8890
VER.DE.x.AGM3             -AGCCGCA--ACAAGAAGCAG-TGTC--TTTG--ACGCCTT-C-CAAA-C-C-GGCT-CCA-A-ATGGCTT-CTTGC.........AG-TCC----A---G--A--CAT-A--TCT--A------G-GC----A---C-T---G-A.......--CC-TAAT--G...... 8411
TAN.UG.x.TAN1             -AG-GG-A--A-CA--G-CTT-TG-G--TTTG--G-GC--T-CGCAAAGC-CC-GTCG--C-C-ATGGA-T-CTTGC.........AG-TCC----A-C-G--A--C-T-G-ACAT--A--------GC----A-AAC-----C--......TGG--CAAC--G...... 8906
SAB.SN.x.SAB1C            -AGCAGCA--A-CGTCAC-CACTGTG--TGTG-T-CA---T-CGA-AAGC-CC-GT---TCA--ACG-C-T-CTTGC.........AG-TCCAT-GT---G--AG-CAT-GCACAT--A--------GC-C--A-AGC-T--G--A.....TGG--TAAT--G...... 9087
MND_1.GA.x.MNDGB1         GTTGG--G--C-G-TA-ACCGGATGGCTACA---A-ATAG-A-CTTCACA---TGCAGAGAA-CA--CATT-CT-CG..................G-G-CGTGC--AT-G-------TC-----G--T-----AA-TGCA---C-A.....CCCC-TAAC---...... 8378
DEB.CM.99.CM40            -AG---CCATCACGT-GGCAGGTGGG-AGCTTTGGCAGTGGC-CCAG-G-ATTG-T--G-T--CGT--C-G-G-C-T.........GTCA----AG-AG----A----T-AGAG-G---GCT-----C--------AGCA--GCTC.......--CC-CAAC-G-...... 8559
DEB.CM.99.CM5             -AG-C-CCATCA--T-GGCAGGGC-A-AA-TCTGGGAGTGGC-CCAA-G-ATTG-TC-GGT--CG---C---G-C-T.........GTCT-G-GAG-C---AACC---T-GC---T---GC------C-----A--AGCA--GTT-......T-AC-CAAC-G-...... 8480
TAL.CM.01.8023            -AGC-GC-TGG-GAGA--GCG---CA--GCTT--A-GGTTCACCTGGCGCT----GTC-G-T--AT-G-CG-CT-CG.........-G-T-G--AGTGCAG--A---GCAGCG--C--GC-G----CG--G--A--GC-T---TAT.....TGCC-TGCT-G-...... 8205
TAL.CM.00.266             -AGC-GC-TGGCGAGA-GGCA---CA--GCTT--A-GGTTCAC-TGGTG-T----GA--G-T--AT-G-CG-CT-CG.........-G-T-G--AG--GA---G--CGCAGCT--C--C--GC-C-CG--G--A--AC-T---TAT.....TGC--TACT-G-...... 8655
SUN.GA.98.L14             CT-CG--CAGGA-C--GAACGCGTGCAGA--CAG--GG-AGAG--TATCATC-T-TC-G-A--CA-G-AGTAG-ACCTTTAGTCTCGGCA--AAA--G--GC-C-AAT-GA-T-G-A-A---G---GT---ATTCC-AGT--G-CCACAGAG-CCACTT-G--...... 9181
WRC.CI.97.97CI14          C--TAC-CAAGAAG-TTGCAAAG-GCAGG-AG-AG-ATA-GCAG--A----TAT---CA-CT-GC-GCAG-AAGAC-.........GC--C-AGCA----ATC-GGATG-TTGCTA-C---CTCTGCGC-AGCTTTGATA--T-CAGAAGA-GC--CAGTG--AGAAGA 8920
WRC.CI.98.98CI04          GC-TAC-CAAGAAG-TTACAAAG-GCAGG-AG-AA-C-A-G-AGAAA---ATAT----A-ATAGC-ACAG--AGAC-.........AC---TAGCA-----TCTGGAAA-TT--TA-C---CTCTGCTC-AGCCTTAATA-G--CAGAAGA-GAC-CT-C--------...... 8959
WRC.GM.05.Pbt_05GM_X02    C-CTACAG-AGAAG-TTGCTGCGCG-CGGCAGAAA-G-AAG-AGAAG---ATATACACGCCT-GC-GCAG--AGAC-.........AGCA--AGCA-C--CATC-GGG-AATT-CTG-C---CTCTGC-C-A-CCTTAATCCGT-CAGAAGAGGACA--AG---...... 7855
LST.CD.88.524             CTGCG--CTGGA-C--GAACAAG-GCAGA-GGAA--G--G--AGCTATCAAAA--TA--G-AACA-ACATTAGGT-G.........GG-CA-AACA-A---AGATGG...AGAT-CAGATTC----GTG-G-TTCC-AGT--G-CCACAGAG-CCACT--G--G...... 8035
LST.CD.88.447             TTGCG--CTGGAGCT-GAACAAG-TCAGA-GGAA--G-AG--ACATATCA-CAG-TC--GCAAGA-ACCTCAG-C-G.........GG-CA-AAGAAG---TGG-GATT-AGAT-CAGAG------GTG-GTTTCC-AGT--G-CCACAGAG-CCGCT--G--G...... 8035
LST.KE.x.lho7             TTC-G--CTGGC-C--GGA-AAATGCAGA-GGAA--G--G-CAGTTATC-A-C--T-----AAAA-ACATCCA-T-G.........GGCAA-AAGAAG---TGG-GA-T-AGAT-CGGAG------GTG-GATTTC-AGT---GCCACAGAG-CCGCTT-G--G...... 9110
SYK.KE.x.SYK173           -A-TCGCA--A-C-TCGC-CGGA-GG-AA-T-CG-GATTGG--G--G-CTATTTGGC-AGCAA-ATAC-CC-CTACC.........AG-C---T-GTCGAG--CG--GCAGC-T-G--GC-G--GC-TC----A--------G-TC......TACC-TAAC-G-...... 8723
SYK.KE.x.KE51             -AG-AGCA--CA-CTCGC--GGACGA-AAGTG--GCACCAGA-G-TG-CCATTTGGC-AGCCC-AC-C-CCTAT-CC.........AG-C---T-GC-GAGAA-G-CGCAGC-T----AC-----C-TC-------GC----G-TC......TACC-TAAC-G-...... 8422
COL.CM.x.CGU1             G-T--CTA......TTGG-A-C--GG-GTC--AG-C-TCAGA-TCT----ACACTTC-G---TGGC--CG-TCT-CG.........GCGC-T-GC--GC-TCGAGCG-CTG-ATAT-TGC--G--TGG-T-T-T-A-CGTCCCCAA......GG-CCTGCC---...... 8151
```

```
                                                  Nef start
H1B.FR.83.HXB2      .......................ATGGGTGGCAAGTGGTCAAAAAGTAGTGTGATTGGATGGCCTACTGTAAGGGAAAGAATGAGACGA...........................GCTGAGCCAGCAGCAGATAGGGTGGGAGCAGCATCTCGAGACCTGGAAAAACA 8915
Nef                                   M  G  G  K  W  S  K  S  S  V  I  G  W  P  T  V  R  E  R  M  R  R  .  .  .  .  .  .  .  .  .  .  A  E  P  A  A  D  R  V  G  A  A  S  R  D  L  E  K  H
H1A1.UG.85.U455     ...........................................AG--CAGAG-G-A-------GAG--T--A----------GA-..-..-..-..-..-..-..-..-..A--CCTG--------A-AG-A--A-------T-----A---TT-A--T---T- 8366
H1B.US.90.WEAU160   ...............................TA---------C-----GT-G---------G--A--------------A-------------------------------AG------------T----------T---C------ 8922
H1C.ET.86.ETH2220   ...............................G---C-AT------T-----CCAG-A--------G--A----A-------A-------------------------CT-----------GG-A--A--------G---------T-A--C---T- 8312
H1D.CD.84.84ZR085   ...............................A--------------A-AG---------G--A--------------A---AA-...ACTGATCCAAGGGAAAGAAGAAGAC-----------------G-------------T----------------G--- 8460
H1F1.BE.93.VI850    ...............................G----------------C---A-AG----------G-----G--------------A-........................A-CCCTA----------AG------------TG---------T-A--C-G--G 8192
H1G.SE.93.SE6165    ...............................A-----------------CA-AG----------CGAG--------------A---AAC........................A-CCCTA----------AG-A--A-------T-----A---TT-A--T-GG-- 8354
H1H.CF.90.056       ...............................A-----A-------------GA--GG---G---T-----A---------------G-------------------------------A----T------AG---A-------TG--------TT----T-G--G 8235
H1J.SE.93.SE7887    ...............................GAA---A-------------.................CAG----------------G-------------------------GCTC-C-CT------------G-A----------TG----A---TT---CT--G-- 8213
H1K.CM.96.MP535     ...............................------------------------A-AG----------G--A-------------------GCACGACCAGCAGCAGACAGGGTGGGAA-AC-A-----------CG------------T-----A--------CT-G--- 8102
H101_AE.TH.90.CM240 ...............................AA-T---------------CA-AG-G---------CAG--C------A--A-AG-A-........................A--CCT------A----AG-A--A-------T-----A---T--A--T------ 8489
H102_AG.NG.x.IBNG   ...............................-----------------C--CA-AG-G---------AG--T-T-A-----------A-........................A-CCCTA--------ACAG-A--A----------T----A---TT-A--T-G-- 8442
H1N.CM.95.YBF30     ...............................AAAG-TT--------G--C--CC-AG-A--------AGAAA-C--A----------A-------------CAAACGCAAGAACCA--A-TAGAGC------TAG-A-CA-------T----A---T--A-CT--T-G 8502
H1O.BE.87.ANT70     ...............................AAA-GCA-T-AG----G---AAT-TGAG------G-AG-A-----A----------A----------.ACTAGAACTTTCCCTGAGT----A---TGC---CC-G-A--A---CAGATC--CA-G--AT-A-C-GCTAG 9002
H1O.CM.91.MVP5180   ...............................GAATGCA---AGC-----C-AAT-TGCA-----T-AGAA----A--T----------A-----------TCCTCCTCTGATCCTCAAC-A---TGT---CC-G-A--A-----T-TC--CA-G--GT-A-C--CTAG 9025
CPZ.CD.90.ANT       ...............................------TCTGCA-----T----TC-AGTG-G-----GCA.............AC--GC---T--GAA-.......................ATACA---AA--AATC-----GATA-A--GC-CTGTGGAAAC--AT---CC-GTAG 8369
CPZ.CM.01.SIVcpzCAM13 CTTTAAAATAGTCTTGCTTTAAG.....A-CTC------T-------CC-AG-A--------GAAA-C--AA------A-A-GAG.......................----CCAAT----------AG-A--A----AG-TT---AAG-T--A-----T- 8548
CPZ.CM.05.SIVcpzEK505 ...............................AAAAGTT--------G-----CC-TG-G-G-----ACAA--C-TA--T--------A-.......................CAGC-AGATC-------AG----A------TT----A---T--A-CT--T-G 8422
CPZ.CM.05.SIVcpzMB66 ...............................A----------------C---C--G-G--C-----AGAG-----A--C---C-T--GA-G.......................A-AC-AA-T--------AG-A--A---C---T---A--G-----A-C--G--- 8560
CPZ.CM.98.CAM3      ...............................CAA--G----------------CC--G-A--G-----AGA---C--A------C-A--CGA-.......................CTCA-C-TGCTT---CCAG-A--A---C---TC--A-A----T-A--GGC--- 8352
CPZ.GA.88.GAB2      ...............................CAA---A-----T-----C--CA-TG-A--------GCAG--C--A--G---C-T--G--G.......................A--C--GA---------AG-A--A----A--TG----AG--TT-A-C--GG-G 8326
CPZ.CM.98.CAM5      ...............................AA------------------C--G-G-----G-AG----T--AC---G--A----A-.........................A--CCA---------TCAG-A--A---C---TC--A-A---T--A--GGC--G 8658
CPZ.CM.05.SIVcpzMT145 ...............................A-C-----------------CC-AG-A--------AGAAA----A--G-----A-A-A-.......................---CCAA----------G-A--A--G-A--T---A-A----A-A-C--GC-G 8448
CPZ.TZ.00.TAN1      ...............................AAA--TA-TTGGT-G-T-GCC-.......-GGCC-GG-AA-CC-TC---GATC-T.............................CA-A-CA-CT--AGT--GCCT--A---CAG--C--A-A------CC-G--TA- 8601
CPZ.CM.05.SIVcpzLB7 ...............................AA----A-------------CC--G-A--------GAA--T--AAG------T----A-.......................-AAACTGT---------AG-A--A---C----C--A-A------A-CT-GG-- 8447
CPZ.US.85.CPZUS     ...............................-----------AA--------T--G-----CA-AG-A--------AGAA--C--AA-C---T------A-.....................ACTCAGA-AACAG----------AG-A--A---C-T-TC----A----T-A-C-G---- 8948
CPZ.GA.88.GAB1      CATTAAG.......................AAC---A-----T--------C--G-A--------GAG--C--AAG------A--GGA-.......................---CCAA---------GG-A--A----A--TT--GAAG-----A----G--- 8988
MON.CM.99.L1        ..................................................................................................................................................... 8788
MON.NG.x.NG1        ..................................................................................................................................................... 7413
GSN.CM.99.CN166     ..................................................................................................................................................... 8771
GSN.CM.99.CN71      ..................................................................................................................................................... 8771
MUS_1.CM.01.1085    ..................................................................................................................................................... 8780
MUS_1.CM.01.CM1239  ..................................................................................................................................................... 8749
MUS_2.CM.01.CM1246  ..................................................................................................................................................... 8845
MUS_2.CM.01.CM2500  ..................................................................................................................................................... 8801
RCM.NG.x.NG411     ..................................................................................................................................................... 8535
RCM.GA.x.GAB1      ..................................................................................................................................................... 8461
DEN.CD.x.CD1       ..................................................................................................GTTACTG-AGC-TGAT-AG-AC--AGAAACCTTT-T--ATG-AGA-...GAAATC-TC-- 8912

MAC.US.x.239        ..................................................................................................GGGACAGAAAT-CAATCAG-G-C-GTATA--AATA-TC---GGA--A---CA-CTG--G- 9303
Env
Nef                 _.__.__.__
```

```
H1B.FR.83.HXB2          TGGAGCAATCACAAGTAGCAATACAGCAGCTACCAATGCTGCTTGTGCCTGGCTAGAAGCACAA......GAGGAGGAGGAGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGG 9079
Nef                      G  A  I  T  S  S  N  T  A  A  T  N  A  A  C  A  W  L  E  A  Q  .  .  E  E  E  E  V  G  F  P  V  T  P  Q  V  P  L  R  P  M  T  Y  K  A  A  V  D  L  S  H  F  L  K  E  K
H1A1.UG.85.U455         -------G-------C--T------T--T----T-----CAG------------G-----G---......-----A-GA--C--A--C--C-----T-GG--A--------------------------T---------T-T-----C---TT----------------- 8530
H1B.US.90.WEAU160       ----------------------GAG......-AT-------A---C-T----T--A--------......-----T---------------------T-G---------------------------------GCAT----------------.------------ 9079
H1C.ET.86.ETH2220       ---G---C-T-----C-----C---C-C--C-AT---C---A-----------G----G---GAG...-----A--A----A--C---------G-----------------------------------T--------AT-C-----C----T------------------ 8479
H1D.CD.84.84ZR085       ---G-----------------------A----T-------C--------G---------GGAA...-----C---------C---------G---------------------------------------------------------------------G-A------ 8627
H1F1.BE.93.VI850        G--------T-----C-------TAG-A---T---C---AC-TG-------G---------......---------A-A--C---------G------------G-------------------C--T----------------C--T----------------- 8356
H1G.SE.93.SE6165        ----------------C-------------C-AT---C---A-----------G--------GAG...-----CTCA--A--A--C---------G--A--A-------G----------G-------------TT----GT---T-T-----C---TT----------------- 8521
H1H.CF.90.056           C--G---G------T--AT---------T----T---CG--A-GCC--------G---------GAG...--C-G-----A--A--C---------GG-------------------------------C--T---G---T-T-----C----T----------------- 8402
H1J.SE.93.SE7887        ---G-----------C-----------------A-------AC-----------G------------......ACA-----------A--A---------A-------A---A--------------------A-------G-----T---------TT----------------- 8377
H1K.CM.96.MP535         ---G---G------C--------TT--CAC-AT---C---A-----------G--------G......-----A-----A--A----------G-----------A--------------------T---------T-T-----CG--TT----------------- 8266
H101_AE.TH.90.CM240     ---------A--------T---......ATAGAT------A-----T-----GAG-------......-----C--------A--C--------TG--G-------C-----------------T---G---T-T----------TT----------------- 8647
H102_AG.NG.x.IBNG       ---------------C--T---------CAA--T--CC---A-----------G-------------......-----T---A-T--A--C--------GG--A-----G-----G-------G-------------T----G-----C-----C--------------------- 8606
H1N.CM.95.YBF30         A--G--C-----C-TA--A-----TAG--AC-AT----AAAG-ATA--T---------------GAA...--A--A----A--A--C--------ACGC----------A-----G-----A--C--T--ACAG---T-T------TC-TT------------T---- 8669
H1O.BE.87.ANT70         A----GG--AC-----TC-C----TC-TCAA-A------A--CCT---A-TC------AGT--C.....C-A-----A--A--A-----------AG---------A--G---C---G--------G-------C--T--A-G---AT-T--C--C---TT----------------- 9166
H1O.CM.91.MVP5180       A--G-G---AT-----TC-C-C--TC-TCAA-A------A--CCT---A-TC-----CAGC--C.....A-A--T-----T--A--C--C-----A-G------A--G---C---G--------G-------C-TT--A-----CT-T--C--C---TT----------------- 9189
CPZ.CD.90.ANT           A--G---C-A--CTCCTCT-C--T--GTA-AGAA--A-A--TAATCA---ATTC-----AT--T.....ACA-----A-G-ACA--C-----T--G-G------A--C--CA-G-----C---G-----C-----AG-A--ATT---------C--GTCATGN--------------- 8533
CPZ.CM.01.SIVcpzCAM13   ------T-----T-----AC-----C--AG---T---CAAA---TA--A---T----T-A-ATGAAACATC-T--A--A--A-----C--------T-G---A-----G--A-----------------------A-G---CT-T--C---TC-----------G--G----- 8718
CPZ.CM.05.SIVcpzEK505   A--G--T--T--C-TA--A-----TAAG-AG-AT---CAA--CCTA--T----------AG--......A-A-----A-CA--A--C--------ATGT-----------A------------------A-----A-A---T-T------TC-TT--------G--------- 8586
CPZ.CM.05.SIVcpzMB66    ---------A--C-----A--C---T-CCAG--A----AGA-CCT-------T-------A-GT-......C-AA-T--A--A--A--C--------A-GG-----A------C------------------C-TT----G---T-T-----C--------------------- 8724
CPZ.CM.98.CAM3          -------G----C-C---GC-C---C--CAG-AT---CAGA--CTG------T------A-ATGCAAGAAC-CA-T--A--A--A--C--------T-G------A------C-T-----T-----T-----C-TT--A------T-T---T-A---------------A------ 8522
CPZ.GA.88.GAB2          ---------A--TTCC--AC----CC-TCAG--T---CAAA--CT---A----------A-ATG.......AC-------T--A--G-------A-G-------------------------------A--T--A-----CT-T-----C--------------------- 8490
CPZ.CM.98.CAM5          ------------C-C---A---CAC...A-A-AAT--CAGA-CCTG------T------A-ATGGAAAATC-T--A--A--A--A--C--------T-G---C---------CC--------------------A------CA----T-T---T-A--------------------- 8825
CPZ.CM.05.SIVcpzMT145   -------G------T---A-----CC--CAA-------AGA---TA--T----------A-ATG......C----T--A--A-----A----------A--A--G--C--------T---------T-----C-TT----GT--CT-T-----CTCT------------------ 8612
CPZ.TZ.00.TAN1          A----GTC----T-C--A--CCCT--GTA-CT-AGCA-A--TG-TA-AA-ACTCT-C--AC--T......ACT--A--A--A--A-------------G---AGCA-----CA-G-----C---G-----C-----AG-G---CT---AA-------GTCATGG--C------------- 8765
CPZ.CM.05.SIVcpzLB7     ---G-----T-------T-----CC--CAA--T----AGA-CAT-------T----G-AG--......A-------A--AA-AT-G--------A-A-----------A-----G---------G----------TT--A-AG---T-T------G--T----------------- 8611
CPZ.US.85.CPZUS         ---G--T-----T-C---A------C--CAG-A---CCAAA--CTG--A---T----T-AGATGACTAATC-T--AAGT--A--A--C--C--T--TCGC-----A-----A-----G--C-------G--C------------ACA---AT-C---T-AG--TT----------------- 9118
CPZ.GA.88.GAB1          C-----T-----T-----G--C--CC---AG--T---CAAA--CTA--T---------G-A-ATG.....--CA-T--A--A--A--A--C-------A-G--A-----T--AACG-------------T--A-----T-T------TCA------------------- 9152
MON.CM.99.L1            ............GG-C--A---GGCAA-CAACG-CGGCAAC-----------G--C-AC...GACGAC--C--C--C-GT--C----G------A-GG--C-GA--T--GC-TC-CGACCCT----GG---CTCATGA-G--C--CTCT--T-A---G---------- 8934
MON.NG.x.NG1            ............GG-GC-G--CCACCGCT-A-A-CGCCGCC-AACCA-TGTCG-T--T-ACGCCGACGAG--CC-A--TTT-CCT--C-G---G--TCGC---AT---G--AC-T---GACCCT--A-GG---ATCATGA-G--C---TCT--T---------------- 7571
GSN.CM.99.CN166         ..........................--AGC-A--AC-CAAGA-A-GAA-C--C................-----T-T--ACCATCA-GC-----A-G---A-G------AC-T---GACCCC--C-----A-TCATG--G----CTCT--------G---------- 8899
GSN.CM.99.CN71          ..........................--AAT-ACCCT-CAGGCAAAAT--CA-C..................--TT-T--ACCATCA-GC-----G-G---A-----G--GC-T---GACCCC--C-----ACTTATG--G----CTCG--------G---------- 8896
MUS_1.CM.01.1085        ............-G-AGC--G--A-AG-C--GTG-TA............................GAA---C-T---...CGTACC-GC--G--GCGT--GAG---G--AA-CC-TGACCCC--C------CTGATG--G---A-GTCG--T--C-------------- 8905
MUS_1.CM.01.CM1239      ............-G-A------GAGCGGC-ACT--TC............................GAA--AC-T---...CGTTCA-GC----GCGG--G-GA--G--CC--C-GGATCCC--A------ATGATGA-T--C--CTCT--T-A-------------- 8874
MUS_2.CM.01.CM1246      ............--AC-AGGGCCGCAGGCAA-T--TA............................GAA--AC-T---...TGCTCC-GC----ACGC--GA-A--C--GA-TC-TGATG-C-AC-----ACTCATGA----C--CTCT----A---G---A-----A 8970
MUS_2.CM.01.CM2500      ............GG-A--T--C-A-AA---A-T--TA............................AAAA-AC-TA-A...CG-TCG-G-----ACGC--G--ACAC--AA-CCAGAACCCC--C--T---TT-ATAA--A-C--CTCT--T-----G---------- 8926
RCM.NG.x.NG411          .........AGG-G-GTTT-CA-T--A-A--CAGC-GT-A-TGGA---AATAGATTGGTATG--GATACA--T--TACCTT---A--G--------G-G---A--A--C--CA---G------T-------ATT---AA-------ATCT------------------- 8696
RCM.GA.x.GAB1           -AC----GAA--GCAAGCTTCA-T--AT-AC...-T--ACTGG-A--AAGATACT...............--T--CACCTT---A--A--------G-A------A-----AC-T----------G-------CT---AA----CA-GTCT------------------- 8613
DEN.CD.x.CD1            ATC-T-TCCAG--CTGGA-CC-CA-CATC-A-GA.........ATACA-CAATAT-TG-T-G-CCATGAT--T--A--A--A--A--C--------ACGG--A-GACG---AC-GCACG--CC---C--T--A-ATATGA-----A-GTCT--T-----G---------- 9073
                                                     Premature stop in SMM239

MAC.US.x.239            GA---A--AATT-GCATA--GA-A-CA-AA--TGG---A-ATA...-ATGA-TA-...............--T--T--CTT---A--GG-AT----G-GG--AA-A--T--CC-----A------G------ATTG--AA----CA-GTCT--T---A------------ 9455
Nef                       R  E  K  L  A  Y  R  K  Q  N  M  D  D  I  .  D  E  *  .  .  .  .  .  D  D  D  L  V  G  V  S  V  R  P  K  V  P  L  R  T  M  S  Y  K  L  A  I  D  M  S  H  F  I  K  E  K
H2A.GW.x.ALI            AA---AG-AAGA-TTGTA---GCA-CAGAA--TGG---A--TA...-AT-TAGAT--T..............--T--TAGCCTA--A--GG-CT-T-------AAGA----AA------A-------A----ATTG--A------A-GTCA--T--AA---------G-- 9497
H2A.DE.x.BEN            AA--CAG-AAGATTTGTAT-GGCAGCA-AA--TGG---A--TA...-AT-CTGAT..............--T--T--CCTAA-A--AG----T--T-----AAGA-----ACGG---GA------C--T--ATTG--AA-----A-GTCA--T---A----------A- 9498
H2A.SN.x.ST             A--G-AG-AAGGCTCGTA---GCA-CA-AA--TGG---A--TA...-AT-CAGAT..............--T--T--CCTA--A--GG-C--T-------AAGA-----A------GA------A--T-G-TTG--AAG----A-GTCA--T--GA------------ 8937
H2B.GH.86.D205          --CCCGGGGGG--GAAG-GGGAGGGCA-CAAGATGCA-A--AA...AGTGATGAG..............--C--T--A-T--GA-CCA-ATGTAAG--C--AATA-----AC-G-------------A-----ACT---A-----CA-GTCT--T---A---------C--- 9411
H2B.CI.x.EHO            GACTTTGGGAG--GAAG-GGGAGG-AA-CAAGATTCA-A-......-AGGATGAT..............-----CA-T--A--A--GG-C-GT--A-G---CGG---C--AC-G--G---------A-T---ACT---A-----CA-GTCT--T------------------ 9385
H2G.CI.92.ABT96         AA------AGTT-GCATA--GGYAGCA-AA--ATG---A--TA...-ATA-TGAT..............--TA-T--CTTA--A--AG-A------TAC--CA-A--C--GC-----GT-----G------ATTG--AA----CA-GTCA--T---A------------ 8883
H2U.FR.96.12034         GAA--AT-AATT-CAATAT--GCA-CA-AAC-TGG---A--TA...-ATGAAGA-..............--TA-T--TTTA--A--AG-AG----A-G---AAGA---G-AC-G---A--------------ATTG--A------A-GTCA--T---A----------A- 8979
MAC.US.x.EMBL_3         AAAG-A--AATT-GCATA--GA-A-CA-AA--TGG---A-ATA...-ATGA-GA-..............--T--T--CTT---A--GG-AT----G-GG--AA-A--T--CC-----G-------------ATTG--AA-----A-GTCT--T---A------------ 8586
SMM.US.x.H9             A-------AATT-G--TAT-GACA-CA-AAC-TGG---A--TG...-ATAATGA-..............--T......TTA--A----G------G-GC--AAGA--C--CG-----ATC-----A-----ATTG--AA----CA-GTCT--T---A------------ 8938
SMM.US.x.PGM53          GA---A--AGTT--AATAT-GACA-CA-AAC-TGG---A--TA...-ATGATGAT..............--T--T---CT-A-A--AG-AT---ACAC--AA----G--A------G-C---T-A-----ATTG--GA----CA-GTCT--T---A------------ 9392
SMM.US.x.SIVsmH635F_L3  AAA-----AATT-G--TAT-GACA-CA-AAC-TGG---A--TG...-ATGATGA-..............--T--T--CTTA--A---G-CT---GCAC--AA-A--C--C-----GG-C-----A-----ATTG--AA----CA-GTCT--T---A------------ 9477
SMM.SL.92.SL92B         AA--CA--AGCA-CAGTAT-GGCAGCAGTG--TGG---A-ATA...-ATGA-GAT..............--T--T---CT--CT--AG-AGTT---TGG--AA-A-----CC-----G-----T-A-----ATTG--AA----CA-GTCA--T---A------------ 8906
STM.US.x.STM            AA-C--T-AGCT-GAATAT-GACAGCA-AAC-TGG---A--TA...-ATGA-GA-..............--T--TA-TCTA--A--AG-AG---ACAT--AAGA--C--AC-T--GGA-------------TTG--AA----C--GTCA--T---A-----AGC---- 9106
MNE.US.x.MNE027         GA-G-A--AATT-GCATA--GA-A-CA-AA--TGG---A-ATA...-ATGA-GA-..............--T--T-GCTT---A--GG-A-----G-GG--A-GA--T--C-------G-------------ATTG--A------A-GTCT--T---A------------ 8932
DRL.x.x.FAO             -CATATGCAAG--GCAGCAGCAGAGAA--A-GT--CCCT-C-CACA-AAGAAGA-C--C-CTC-GAAGAA--A--A--A--A--A-----------ATGT-----ACGC----------GAGCCA--C-----A-ATCTA-----CT--TC-------------G-------- 8791
MND_2.x.x.5440          AA-GCACCAGCAGGA-GTA-C-CTCC-CT-AGAA--A-AGCAACCAAGTGATGA-..............-----A---.....--A--C----------TAC--AAC-CAG--AG-GCCTGA-GCA--C--------ACTTGA----CA-GTC-------------------- 8706
MND_2.CM.98.CM16        ..........................................-GAAGAG..............--A--AAC-.......--C--------ATAT--A---TGC--AG-----GA-CCC--C--------ACTTG--T---A-GCC-------------------- 9010
MND_2.GA.x.M14          .....................................AGTA-ATGATGAGC--..............A-A--A--A...AG---C--------TTAT--CACACAG---G-GC-TGA-GCA--C-----AAACCTGA-----A-GTC-------------------- 8951
OLC.CI.97.97CI12        G-C-AT-GAGT-TTTA-TG-C-GGG-AC.....................ACTGAC..............AC---T--TCT--AA--A---------G-AG---------G--AG-GC---A-GCC-----T---A-TCTAA-T--G--CTCG--T----------------- 8478
GRV.ET.x.GRI_677        C--CCGCCG-CTGCAAGAAGGAGACA-CTT-...G---AGTGG...-ATGATGAT..............--A--A--A...--A--C--C--T--GCA-----GA--C--C-------AG-----C--T--ATT---A--G--CT--TC--------------TC------ 8930
VER.KE.x.TYO1_patent    G---AGG-GGCTTCAAGAAGGAGACA-CTGG...G-A-AGTGG...T-GGATGAT..............-----A--A......-----A---------G-G---AAGA-----C-----G-A-----------T--ACTT--A--G---T--TCG-------------------- 9047
VER.KE.x.9063           ----AG-CG-CTGCAAGAAGGAGAC-A-TGG...G---AATGG...T-AGATGAT..............--A--T--A......-----A-----T--A-A---AAGA-----A--G--G-A-------------ATT---A--G--CT--TCG-------------------- 9086
VER.KE.x.AGM155         A---AG-CGTCT-CAGGA-GGGGACCA-TGG...G---AATGG...T-AGATGA-..............--A--T--A......-----A---------A-G---AAGA--G--AC------A---A--A-----ACTT--A------T--TCG-------------------- 9063
VER.DE.x.AGM3           A---AG--AATT-CAAGAAGGAGATAA-TGG...G---AATGG...T-TGATGA-..............--A--T--A...--A--A---------A-G---AAGA--G--GC------A------C--T--ATT---G--G--CT--TCG-------------------- 8584
TAN.UG.x.TAN1           A---AG--AATTGGCAGCTG--GACA--TGG...G---ACTGG...-AACCAGAG..............--A--A--A......--C---------A-G---A-GA--G----------A------A--T---TTG--A------T-CTC---------------------- 9073
SAB.SN.x.SAB1C          A--T-T--AA-TG-AACATC-AGG-AG-CAG...CCCT-GTGG-A--ATGA-GAT..............--A--A--A...--A--C---------A-GG---TGCT----AC--C--G-------A--T---TTG--AA----CT--G--------------------- 9263
MND_1.GA.x.MNDGB1       ...............AGATG-G-A-CTT-ACTTA-CACAGCAGGA-C-AGAAGAG..............-----A--A...--T--A-----T--GTGT-GC--A--TT-C------GTGCCAT-A-----A-ATCTGA----CT-CTCT--T---A------------ 8515
DEB.CM.99.CM40          AA-TC-T-GA-G-GAA-TA--GTGCA-CAA--AGC-ACAACAAGACTTAGAT..............C--ACA--A-CA--A--C--C-----ACGC--C-------T-CA-CC--GACCCC--A--T---TT-ATGA----CTACTCT--T-----G----------- 8720
DEB.CM.99.CM5           A-ACC-CCA--G-GAG-TA--ATGCA-TCAG-GAGCCCAG-AG---TATGAT..............-----AC-T-G-A-A--C---G-T----G---AAG---C--AA-T--GGA-CCC--------ATTGATGA----CTACTCT--T-A------------- 8647
TAL.CM.01.8023          .--GAAGGCACTCTCGGAAGGAGA--TCTGGTATG-AAAG-AGCAA-AAGAAAAT..............--CAT----CTT--A--C---------A-GG-----A--T--A--G--G--CTGT-A------ACTG-G-A----CT-CTCGTT------G---------A- 8393
TAL.CM.00.266           .--CAAGGCATT--CACAAGGAGAG-TTTGGTATG-AAAA-AACAA-AAGAAAAT..............--CAT---ACTA-----C--C--T--G-GG---A-A--T--G--G---A-CTGT-----T--ACT--G-A----CT-CTCGTT------G----------- 8843
SUN.GA.98.L14           ...........................................................................................GCCA--A--T---CAGCTAA--------CTCA--T---A------------ 9231
WRC.CI.97.97CI14        ....................................-AGTGATGA-..............--T--A-GA...--A-----------A--T-----A--T--CA------A-CCA----TT--ACTGCTA-----C--CTCG--T--A--T----------- 9030
WRC.CI.98.98CI04        ..................................A-AAGA-AGT-CTGAT..............--T--A--A...--A--C--------TCAT--------T--AA------A-GCT--A-TT--ACTGCTA---------CTCG--T--A--T----------- 9072
WRC.GM.05.Pbt_05GM_X02  ..................................A-AAGACT-A-CTGAT..............--T--T-CA...A-A--A--C--T----G---A--A--T----------AGGCA--A-T----ATCTTA-----C--CTCT--T--A--T----------- 7968
LST.CD.88.524           ...........................................................................................CTCT--C--T--AA-CTTGA----C--CTCA--T---A------------ 8085
LST.CD.88.447           ...........................................................................................CCCT--C--T---A-TCTGA----C--CTCA--T---A------------ 8085
LST.KE.x.lho7           ...........................................................................................CCCT--C--T---A-TCT-A----C--CTCA--T---A------------ 9160
SYK.KE.x.SYK173         A---CA--AATTCTCACCTGGAGATAT--TG...C-G.........-ATGAAGG-..............--CACA-GTCTC--A-----C--T--GTGT-----AAC---AC-C---A--T-A--C-----ATTG--AA----C--CTCT--T---A-----A-C---- 8911
SYK.KE.x.KE51           G--GAA--AATGG-A-GCGGGAGACATCTG-...C--.........-ATGA-GG-..............--T-----TCTA--A--C-----T--TTT---A--A-----C--G-----CT-A--------ACTT--CA-------CTC---T---A-----AG----- 8601
COL.CM.x.CGU1           -A--CATCCAGATGTATATT-CCACAAC.........C-A-AACCA-AGGACG--..............--A--T-TA--A-CA--A-G-TAT--G-TGT---GAC-G--AG--C---AGGC---CATT---CT-CTG--G--CA-CTCTTG------G---------- 8318
```

```
                          3' LTR U3 start                                                      Nef premature end in HXB2
H1B.FR.83.HXB2            GGGGACTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGG.....................GTCAGATATCCACTGACCTTTGGATGGTGCTACAAG 9228
Nef                       G  G  L  E  G  L  I  H  S  Q  R  R  Q  D  I  L  D  L  W  I  Y  H  T  Q  G  Y  F  P  D  *  Q  N  Y  T  P  G  P  G  .  .  .  .  .  .  .  V  R  Y  P  L  T  F  G  W  C  Y  K
H1A1.UG.85.U455           ----------T---T----------A--G-A---------A---------------G----T-----------A-T-----------G-----T-----------------A.....................A------C-----A--A----------------- 8679
H1B.US.90.WEAU160         ------------------T---------A-------------------------G-----------------T-----------G------------------------.....................ACT-----C------TG-------------T---- 9228
H1C.ET.86.ETH2220         --------------T-----T-----A-G-A---G--G--G--T------T-----G----TA-----------T-----------G--A------------A------.....................-----------------------------T---- 8628
H1D.CD.84.84ZR085         -----------------G--T--------G-A---------------------T---G-----------------T-----------G--A----------------------.....................A-----------T-A--------G-------T-G-- 8776
H1F1.BE.93.VI850          --------------T-----T-----A-G-A----GG-----C-------------G-T-------------------------------G-------------------------.....................A-------------------G----------T---- 8505
H1G.SE.93.SE6165          ----------T---------T-----A-G-A---------G--------C------G----A-T--------A--------A----G------------G----------.....................ACT----TC--C------------------T---A 8670
H1H.CF.90.056             ----------T---T-----T-----A-GCA------G--C---------T-A---G----TA-------------------------C-G-------------------------.....................-AG----T---C-----------G-------T---- 8551
H1J.SE.93.SE7887          ----------T---T-----T-----A-G-A---------G---------------G-TC-TA-----------T-----------C-G--A----------------------.....................A-------C---T----A------------------ 8526
H1K.CM.96.MP535           ----------T---------T-----A-G-A---------A---------------G-T--T-----------T-T-----------G--A----------------------.....................A-------C---------------G------------ 8415
H101_AE.TH.90.CM240       ----------T---------T-----A-G-A---------G--------CT-A---G----TA-T-----------T-----------G--A-----------------------.....................A------TC-----TGT--------------T---- 8796
H102_AG.NG.x.IBNG         ----------T---T-----T-----A-G-A---------A---------------G----TA----------A-T------A----G-----------------------.....................ACT----TC--------------G-------T---A 8755
H1N.CM.95.YBF30           ----------------G--TGG---AG--A---G--------T--A--C--C-----G--T------------AT-C-------C-G---T-------------------A.....................A-T-----C--CG-A-----------------T---A 8818
H1O.BE.87.ANT70           -A------------T-----T-------T-A---GC---A-----G-----T---G-G--TA----T--G--A-T-----------G------------------A-----A.....................AC---G-TC--------A----------TG-TT--A 9315
H1O.CM.91.MVP5180         -A--------T---T-----T-------T-AG---GC---A-----G-----C-----A--T-----T--G--A-T-----------G----TGT--------G--A-----A.....................CCT----TC--------A----------TG-TT--A 9338
CPZ.CD.90.ANT             ----------TN--------T-NN-A-TT-A----GC--C------A---A-------G-TTA----C--G--AAT-----------G-----T------GA--A------A.....................--T-----------CTGTAGA---------T-TT--- 8682
CPZ.CM.01.SIVcpzCAM13     --------------T--G--T-------G-A---G-----A---G-----A--T-A-----C---G-T--T--T---------T---------C-G-----------C-----A-----A.....................-AG-----C----CCTG--------------TT--A 8867
CPZ.CM.05.SIVcpzEK505     ----------------G--TGG--AACC-AG--------A--T--A--C--C---G-T--T--T--T----AAT---------C-G-------------T-----A-----C.....................--------C--C------------------T---- 8735
CPZ.CM.05.SIVcpzMB66      ----------------G--T-----A---AG--------G--------------T---G----T--T--------A-T-----------G--A-------------------------.....................A---------------------C-------T---A 8873
CPZ.CM.98.CAM3            --------------------T-----AT--NNNNNNNNNNNNNNN-G----ACG--TA---T--------G----T-----------G-------------------------A.....................--------C-----A--A-A---GG------TT--A 8671
CPZ.GA.88.GAB2            --------------T--G--T-T---A-G----------G-----G--C--C---G-T--T--T--C-----A--------------G-A-T-----T--T--T--A-----A.....................--GC-T-----GT------A-----------T---- 8639
CPZ.CM.98.CAM5            --------------T--G--T-----AC--A---G-----G-----A-----C---G-T--T---------G--AATT--------C-G-----T--------------------A.....................--------C-----A--A-------------T---- 8974
CPZ.CM.05.SIVcpzMT145     --------------T----CTGG--AAGG-A------G--G---------T-A---G-T--T------------T---------C-G--A--------C--G--------A.....................--GC-C--C-----A-----C--G-------TT--A 8761
CPZ.TZ.00.TAN1            ----------T------T--TT---T-C--A----GC--CC-----A--CACC-----G--TA-T-----G--TGT---T--A--C-G-------------C--T--A-----A.....................A-------C------TGTAGG-------TA-TT--- 8914
CPZ.CM.05.SIVcpzLB7       --------------T--G--T----TA---AG--------A--------C--CG--G-T--T-----------AAT---------C-G-----T-----C-----------------.....................A-T----------A-A--------CG------T---- 8760
CPZ.US.85.CPZUS           --------------T--G--T-T---AG-----------G-----A-----C---G-T-----T--------AAT---------C-G-----T-----C--------T--A.....................--A--------C--A--A-A---G-------TT--- 9267
CPZ.GA.88.GAB1            --------------T--G--T-----AGG----------G---------C--C---G----T-----------------T---------C-G---------------A-----A-----A.....................ACA----TC-----CTGT-------------T---A 9301
MON.CM.99.L1              -A-------G--A-A-GT--T---GTG--GAC-----TAGA-AGA-A--G-AA-ATGC---T-TGGA-TGG----TAA----A-GA-G--TAC-A--T--GGA----------.....................-----G--C--GACT-TGCCC--G-TCCTG-GGTGC 9083
MON.NG.x.NG1              -A--------TAA---GT--T---GTGTWGACC-Y--TC-A-AGT-A--G-AA-ATGC---TGT--TGTGG--ACTTG-G----GA-G--TAC-G-TT--CY-C--A-----C.....................--GC-G--C---ACC-T-CC---G-TCCTG-GGTGC 7720
GSN.CM.99.CN166           -A--------G--CA-GT--T---GTG-GGAC--G--CC-G-AGT-G--A-AA-ATGC---T-TGGAGTGG---CTGG-G----GA-G--TCTCA-T---T-----C-----A.....................AC-C-C-----TAC--TACCA----TC---AT-TGT 9048
GSN.CM.99.CN71            -A--------G-CCA-GT--T-T-GTG--GAT-----CC-A-AG--A--GAGC-ATTG---TT-TGAGTGG---ATTG-G----GA-G--TCC-G-GG--T-----A-----A.....................A--C-C--C---ACT-TGCCA--C-TC----GGTGC 9045
MUS_1.CM.01.1085          -A------------A-GT-CGT-AGTAC-GA---G-TCC-G-AGT-G--GACC---GC---TGTGGAGTGG--AATAA-A--A-GA-G--T--TG--T---GAT--C-------.....................ACTC-G-----CAA--TACCA--C-TCCTA-GGTGT 9054
MUS_1.CM.01.CM1239        -A------------A-GT--T-T-GTG-GGA------CC-A-AAT-G--GACT-AT-C---TGTGGACTGG--AATTG-A--A-GA-G--TAC-G-TT---GAG--C--G---.....................AC-C-G--C---ACACTGCCA--G-TCCTG-GGTGC 9023
MUS_2.CM.01.CM1246        -A-----A-----AA-GT--GTG-G-AT-GAC----TGC-A-AA--A--GACT-ATGC---TGTA-A-TC-AAAATAG----AAGA-CA-TA-TA-----GGAT--C-----C.....................AC-C-T--C----G-CAACCA--G-TCCTG-GGTGC 9119
MUS_2.CM.01.CM2500        -A------------A-GT--T--AGTG-GGA------CC-A-AAT-G--GACC-ATGCA--TGTGGAGTGG--TCTTA------GA-G--T--TG------GA---C--T--T.....................--TC-C--C--G--T-TGCCA--TGTC-TG-T-TGC 9075
RCM.NG.x.NG411            --------------A-TTA-TT---AG----G---AGGCGC--TT-A--C--C-AC--GG-GA-TGA---T--AATAA-A--A----G------T-----T-------------.....................A-A---------ACA-TG-----C---CTA-GGC-A 8845
RCM.GA.x.GAB1             --------------A-TTA-TGGAGTATC---------AGA--AT-G---A---ACC-GG-AA-TGAG--T---ATAA-A-------G---A--------T-----------A.....................A-A---------ACACTG---------CT--GGC-A 8762
DEN.CD.x.CD1              -A------C-G-ATA-TT--T--AGT-C------G--T-C---TT-A--C--C-ATGC-C-GA-TGA---T----T-A-AA-A-GA-G-----C-------GAT--------C.....................A-TC-G--C--G--AGAG-----G-TCCT--GG--A 9222
                          3' LTR U3 start
MAC.US.x.239              --------------A-TTA-T--AGTGC---------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-A-A--A----G----G-T-----CT----A-----A.....................A-T-----C---AA---A-----C---CTA-GG--A 9604
Nef                       G  G  L  E  G  I  Y  Y  S  A  R  R  H  R  I  L  D  I  Y  L  E  K  E  E  G  I  I  P  D  W  Q  D  Y  T  S  G  P  G  .  .  .  .  .  .  .  I  R  Y  P  K  T  F  G  W  L  W  K
H2A.GW.x.ALI              --------------A-GT--T--AGTG-G-------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G----GATAA-T--A----G---------T--T-AT--------A.....................A-A--G--C--GA--TT------G---CTG-GG--- 9646
H2A.DE.x.BEN              --------C-----A-GT--T--AGTAGG--G-----TAGA-----A--CA-A-ACC-AG-AA-AGAGG----GATAA-A--A----G------T--T--T-AT--------A.....................--A--G--C---A--TA---C--G---CTG-GG--- 9647
H2A.SN.x.ST               ----------------GTA-T--AGTG-T--G----GTAGAG----A--CA-A-ACT-AG-AA-GGA-G-G--AATAA-TGGA--C-G---------T--T-AT--A-----A.....................--A--G------AA-TT------G----TA-GG--- 9086
H2B.GH.86.D205            --------------A-GTA-T-TAGTG-G--------CAGA--AT-A--CACA-ATT-TG-GA-TGA-G-----ATTG-GT---GC-G---A-----T--T-AT----------.....................A-A--G-----CAAATA------T---CTG-GG--- 9560
H2B.CI.x.EHO              ---A----------A-TT-CT-TAGTG-G-----G--TA-A--A--A--CACA-ACT-AG-AA-TGA-G-----ATTG-GT---GA-G---------------AT--------A.....................--A--------CAA-TT------C---CTA-GG--- 9534
H2G.CI.92.ABT96           --------------A-TTA-T--AGTG----------TAGA-----A--CA-A-AC--GG-AA-GGAG-----TATAA-T--A----G------T------G----A-----T.....................A-A---------R----G-----C---CTA-GG--A 9032
H2U.FR.96.12034           ----------T---A--T--T-T--T--G--------TAGA---T-A---A-C-ACT-AG-AA-TGAGG-G--AATAA----A----G------T-----CT----A-----A.....................ACT----GG---A--TG-----------CTT-GG--- 9128
MAC.US.x.EMBL_3           --------------A-TTA-T--AGTGC---------TAGA---T-A--CA---ACT-AG-A..........---AT-A-A--A----G----G-T-----CT----A-----A.....................A-T-----C---AA---A-----C---CTA-GG--A 8726
SMM.US.x.H9               --------------G-TTA-T--AGTATT--------TA-A--AT-A-----R-ACT-AG-AA-RGA-G----AATAA-A--A----G---A--T------G----------A.....................A-------C--TA--TT-----------CT--GG--A 9087
SMM.US.x.PGM53            --------------A-TTA-T--AGTG-G--------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-AC---A----G------T------T----A-----A.....................A-------C---A--TT------C---CTA-GG--A 9541
SMM.US.x.SIVsmH635F_L3    --------------A-TTA-T--AATG-G-A------TAGA--AT-A---A---AC--GG-AA-GGA-G----AATAA-A--A----G---A--T------TT---------A.....................ACT-----C--TA--TA------G---CT--GG--A 9626
SMM.SL.92.SL92B           ----------T---A-TTA-T--AGTG-G--G-----TAGA-----A---A-A-ACC-AG-AA-AGAGG----TAT-A----A----G------T--T---AGT--------A.....................A-A--G--C---AA-TT------G---CT--GGC-- 9055
STM.US.x.STM              --------------A-TTA-T--AGTG-G--------TAGA--AT-G--CA---ACT-AG-AA-AGA-G----GATAG-T--A----G------T------G----A------.....................A-A-----C---AA-CAG---------CTA-GG--A 9255
MNE.US.x.MNE027           --------------A-TTA-T--AGTG----------TA-A---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-G-G--A----G----G-T-----CT-G--A-----A.....................CCT-----C---AA---A-----C---CTA-GG--A 9081
DRL.x.x.FAO               --------A-----A-TTGGTGG--T-TC------G--ACA---T-G-----C-ATGCTC-AA-TGA-TGG----T-A-AAAA-GC-G---A-GT-----CAAG--A--T--C.....................A----G--C---AA--TG-----G-TCCTG-T---- 8940
MND_2.x.x.5440            --------------A-TTGGTGG--TA--------G----G---T-G--C--T-ATGCAC-AA--GAGTGG--ACTAA-A---A---G-------------AAG--A--G---.....................A-A--G-----GAA---T--C--G-TC-TG-T---- 8855
MND_2.CM.98.CM16          --------------A-TTGG--T--TA--------G----G---T-G--CT-A-ATGCTC-AA-TGA-TGG--A-T-A-T---ACC-G---ATCT--T--TGAT--C-------.....................A-A-----C--GAA---A---A-G-TC-TA-T---- 9159
MND_2.GA.x.M14            --------------A-TTGGTT---TAG-------G----G---T-AA-C----ATGCTC-GA-TGA-TGG--A-T-A-A--A----G---AG-A-----TT------------.....................A-A-----C--GAA--GA-----G-TC-TA-T---- 9100
OLC.CI.97.97CI12          ----------G---A-GTA-TGG--A-CT-----GG-GA--T-GA--T------ATGCT-TGA-TTGTTGG--AATAA-AGG-A-C-TC---GCAACAGAGGAG-TAGA--A-...............GGCACGT-A--GAGA--TAA---GCCA--G----TA-GGG-T 8633
GRV.ET.x.GRI_677          ----------T---A--TA-T----TG--------G--A-G-----GA--T---ATGC--TGA--GAGTGG--AATAA-AGA-----G---AGCT---T-------C--G---.....................A-A--G--C--GAGAGT------C-TC----TT--- 9079
VER.KE.x.TYO1_patent      -A--------T---A-TTA-T-----G-T--G---A-TA-G-----GA------ATGCTCTTA-TGA-TGG--GATAA-TGA-----G-A-TGC--GGT-GAAG--A-----A.....................A-A----TC--TAAATG------G-TC----TT--- 9196
VER.KE.x.9063             -A--------T---A-TTA-T-----G----G--GA-TAGA-----CA-C----ATGCTCTTA-TGA-TGG--AAT-A-TGA-----G-A-CGC--GGT--A----A-----C.....................--A--------C-GCTG------C-TT----T---- 9235
VER.KE.x.AGM155           -A--------T---A-TTA-T-----G-T--G---A-TA-G------A------ATGC-CT-A-TGA-TGG--AAT-A-TGA-----G-A-CGCA-GGT--AA---A--T---.....................A-A-----C--GAG-TG------C-TC----T---- 9212
VER.DE.x.AGM3             ----------T---A-TTA-T-----G-C--G--GA-TC-G-----AA-C----ACGC-CT-A-TGAGTGG--AAT-A-TGA-----G-A-TGCT-GGT--GA---A-----A.....................A-------C---AGATG---C--C-TC----TT--A 8733
TAN.UG.x.TAN1             -A--------G--CA--TA-TGG--A-C--AG---G-GC-A-----GA----C-ATGC-CTAA-TGA-TGG---ATTA-AGA-----G---AGCA--TT-C--T--------C.....................AC---GA-G--C-GATG---C--C-TT----TTG-- 9222
SAB.SN.x.SAB1C            --------------A-TTA-T--AGTG----G---A-GA-G---T-G-----C-ATGCTCTTA-TGAGTGG---ATAG-AGA--GA-G------T--T---GAT--------T.....................AC---G--C---AA-TG--------------TT--A 9412
MND_1.GA.x.MNDGB1         ---------G----A--TA-T-TAG-AGG------G----A-----A-----C-ATGCAG-GA-TGAGTGG--A-TTGAA----GA-G---AC-G--T--GA----T-----A.....................AC---------TAA---A-------TCCTG-TT--- 8664
DEB.CM.99.CM40            -T--------G-ATG-TT--T----TGCC-----G--T-C---TT-A--G--TCATGC-C-GA-TGA-TGG--AATTA-A--A-GC-G--TAC-G------AAG--A--T--A.....................--T--G------A-TTA---C--G-TCCT--TT--A 8869
DEB.CM.99.CM5             -T--------G-ATA-TT--T----TGCT--------T-C---AT-A--AA-CCATGC-C----TGA-TGG--AATAA-A--A-G--G-TT---G------GA---C-----A.....................CCA--G------ACATA------G-TCCTA-TT--- 8796
TAL.CM.01.8023            --------A-G---A--T--T--AA-A-C-----G--T-CA---T-AA-C----ATGC-C-TA-TGAGTGG---ATAC-G--A----G----TCA------GA---------A.....................ACA--G--C---T-ATG---C--TATC-TG-GG--- 8542
TAL.CM.00.266             -A------A-G---A--T--T--AA-A--C-T--G--T-CA---T-GA------ATGC-C--A-TGA-TGG---ATAC-G--A--C-G--------------GAG--A-----A.....................ACA-----C---T--TG------TATCCTG-GG--A 8992
SUN.GA.98.L14             -T---------------CTGGT-----AG--CC---G----A---T-G---T---ATGCAG-AA-TGAGTGG----T-A-AA---GC-G---AG-T-----CAA------T--A.....................--G---------AA-G-------C---CTA-GG--- 9380
WRC.CI.97.97CI14          -A--------G-ATA--T--TGG--AGCG------G----A--AG-GA-C--C-ATGCAC--A-TGA-TGG--GCTGA-AGA--GC-G---AGCT------GA---A-------.....................A-A----GG---AA---A-----G-------GG--A 9179
WRC.CI.98.98CI04          -A--------GTATA--T--TGG--T-CT------G----A--AG-GA-C--T-ATGC--T-A-TGA-TGG--ACTGA-AGAA-G--G----GCA------GAG--A-----C.....................A-AC-C-GG---AA---A-----C-----T-GG--A 9221
WRC.GM.05.Pbt_05GM_X02    -A--------G-ATA--T--TGGAGT-CT------G----A--AG-GA-CT-A-ATGC---T--TGA-TGG--TATAG-GGAG-GA-G---AGC------CGAT--A-------.....................---C-C--C---AA---A-----C-------GG--A 8117
LST.CD.88.524             -T---------------TA-TGG-----G----GA-T--C---T-AACA--C-ACTGTG-GA-TGA-TGG--TCT-A-TGGA--C-TCAT----------GAT--A-----A.....................ACTC-T--C-----CG-----------CTG-GGC-- 8234
LST.CD.88.447             -T---------------TA-TGG-----G--------G--C---T-ATG---T-ACTGTG-GA-TGA-TGG--TCT-A-TGGA--C-TCAT--CT------GAT----------.....................ACGC-T-TC---T-AG-A----------TG-GGC-- 8234
LST.KE.x.lho7             -T---------------TA-TGG-----G--------G--C---T-AAT---C-ACTGTG-GA-TGAGTGG--TCT-A-TGGA--C-TCAT----------GAT--A-----A.....................ACA-----C--T-----A----------TG-GGC-- 9309
SYK.KE.x.SYK173           -A------C-G--CA-G-A-T-T-GTG---AG---G-T--A--T---C----C-ACC-AC-GA-TGAG--T--AAT-A-A...--CAG-ATC--T-----CT----------A.....................AC-C-G--C--GT-A-T-----------TA-GGG-A 9057
SYK.KE.x.KE51             -A------C-G--CA-GT--TT--GT----AG-----T--G--T--AC-G----ATT-AA-GA-TGAG--T--TGTGA-AGA---C...ATC-C---T--GAGT--------A.....................ACA-----C--G----T---C--G----TA-GGG-A 8747
COL.CM.x.CGU1             -T--------G---A-T---AGAA---CTGAG---G-T--CC-GA-A--G-A--ATGC----ATTGAGTGG----G---AAAA-G--G-TT-G-G--TGA-GAT-A--TG--AGAGGATGGAGCATTAAAAGAG-A----A-G--T---GTTGC-------CTT-GG--A 8488
```

```
H1B.FR.83.HXB2         CTAGTACCAGTTGAGCCAGATAAGATAGAAGAGGCC...AATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGAC.........CCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCAC...GTGGCCCG 9383
Nef                     L  V  P  V  E  P  D  K  I  E  E  A  .  N  K  G  E  N  T  S  L  L  H  P  V  S  L  H  G  M  D  D  .  .  .  P  E  R  E  V  L  E  W  R  F  D  S  R  L  A  F  H  H  .  V  A  R
H1A1.UG.85.U455        --------------T----C-G-AG----G--A--T...-C-GG-----------A---TC-AC--------A-AT---AA-----AG-A-----T.........GAA---A----------AT-----A---------TAC---G-----AA-A---...AGA--TTA 8834
H1B.US.90.WEAU160      -----------G--A-----G---G-------A--T...---G-----------A---------------A-------------------------.........-AT---A------A---AT-----A---------AAA----------------...--A----- 9383
H1C.ET.86.ETH2220      --------------C---AG-G-AG-G-----AAT-...---G-----------A-T-----C-G------C-----------A-----G--T.........GAA--C-------A---A------A--------T-A-------CGCAGA---...A------- 8783
H1D.CD.84.84ZR085      --------------T-----GGT-G-----A-----...-C-G-------AG---A-T-----C-------A-CT---A-----A-----G--.........GA----AG----------T------A---A-----A-A--------G-G---...AA----AA 8931
H1F1.BE.93.VI850       --------------T-----AG--G-G---A-----...---G-----------A-T-------------CA-----AA----A-----G--T.........GAA--C--------AC-GAG-----A---------A----G---C-GAGA---...A-A---A- 8660
H1G.SE.93.SE6165       ---------A-G--T----CAG--G----G--A---...-----------------A---TC-A-------CA-CT---A-----A-----G--.........GAA--C---------C-G-T------A--------TA--------CGGAGA---...A-A----- 8825
H1H.CF.90.056          -----------AA-T---C-GG--G-----C-----...---G-----------A-------C-------CA---------------A-----G--T.........GAC-G---------C-GAT-----AA--------T--A--------GACA---...T------- 8706
H1J.SE.93.SE7887       --------------T---AGCG-AG----G--A--T...---G-------A---A-T--C----------CA-AT---A------A--A-G--T.........GAA--A-----------C------A----------TC-------CGGAGA---...A-A----- 8681
H1K.CM.96.MP535        --------------T----CAG-AG--------A--...-CAG--------G---A-T--C----------CA-A-A--A------A-----G--T.........GAAC-C------A-A---AT-----A---------TTC-------CGAAGA---...--A----- 8570
H101_AE.TH.90.CM240    --------------C-A-AGAG-AG----G----A-...--C--------A---A-T--C----------CA------A------A--A-G---.........GAA--A---------C-GAT-----A---------TGC-------CGAA-A---...--A----- 8951
H102_AG.NG.x.IBNG      ---------A-G--T----CAG--G----G--A---...---G-----------A---T--A--------CA-CT-T-AA----A-----G---.........GAC--T---------C-GATC-----A--------TA-A--G-----CAGA---...ACA--TA- 8910
H1N.CM.95.YBF30        ---------T-GTCAG-T--AG-AG----------T...---G-------C----ATGC-C-C--------CA-AT-T-AA-----AGCA-----T.........GATC-T-A---------G-T----C-A---------TC-------AGAAGA--T...--A--AA- 8973
H1O.BE.87.ANT70        -----------GTCAGA---AG--GC----AGACTAGGA----C-T-T----GGG-T-ATC-CC-G--T--A-CAT-TGCC-----AT-T--A--T.........A-AC-T-A----A-AC-GAT-----A------T--ATCT----GCAACACC--T...--T--TAT 9473
H1O.CM.91.MVP5180      --G--------GTCAG----AG--GC---GAGACTGGGT----C-AAT--AG-TG-T--TC-TC-----T--A-CTT-TAAT-----AGCT--G--T.........G-AC-CG-G--GA-AC--A-A---CA------T--ATCAT---GC--AACA--T...A-A----T 9496
CPZ.CD.90.ANT          T----------A--AT--CCACCTGAT...............-AT----GG-A--TAC-CC-C--------CTT-TACA-----AGAT-GA---.........--AC-T-A---GA-C---AGA---GAA-----TGCAA----GATGAGAAGA---...A-T--AA- 8822
CPZ.CM.01.SIVcpzCAM13  --------TT-GACAA-T--GC-AG-G----CA---...---G-G-----T----A-T----AC-G------A-AT-T-A------A-----G--T.........GAAAGT-A-------C---ATA---C-C----------A-A--G--T--AAGA--T...A-A--AA- 9022
CPZ.CM.05.SIVcpzEK505  ---------C-GTCA--T--GG--G-------A--T...---C-------C----ATGT-C-------T--CA--T---AA------A-----A--T.........--A--C-A-------C-G-T----C-C----------TC-------AGAGTA---...AGA--AA- 8890
CPZ.CM.05.SIVcpzMB66   ---------T-A--A-----AG-AG-G---AGA--A...---G-------C----A--TA--------T--CA-CT---A---C--ACAA--G--T.........GAA-CA---------C---TC----CA------------G--T---C-GA-G---...AGA--T-- 9028
CPZ.CM.98.CAM3         --G--T---C-GACA-----AG-AG-G---C-----...---G-------T--T-A--T-C-AC-G-----CA-CT---A------A-----A--T.........G-AC-TG-------C--G-TC-----A---------A-G--G--C--AAGA---...A-A---A- 8826
CPZ.GA.88.GAB2         -----G-----ATCA-----GGCAG--------A---...--C--G------G-T-ATCT-C-A-----T--CC--T--ACT-----AT-T--G--T.........GAA----AG--G--TC-GATA----A---------T-AA-----CC-G-GG--T...C-T--AA- 8794
CPZ.CM.98.CAM5         --G-----TC-GACA-----GG-AG-G--GA-A---...--CG-------C-C--AT-TAC--C-T-----CA-AT---AA-----A-----G--T.........GAAC-TG-------C-----TC-----A---------A-G--G--TC-GA-A---...C-A---A- 9129
CPZ.CM.05.SIVcpzMT145  --G--C--TT-GCCA--T--GG--G-G--GA-A---...---G-------A----A-T--C--C----T--CA--T---A------A-----G--T.........GA---A--------AC-G-TA----CT-----T---AAG--G--TC-CAGA--T...C-A--AA- 8916
CPZ.TZ.00.TAN1         T-G-----G--A--C..............CCAC-A...G-AG-T-AT-----G-A--T----C----T--A-CCT-TAGC-----A-CTACC--T.........--A--TG----GACTC-GATC---C-C----------A---------AGAAGG---...A-A---A- 9054
CPZ.CM.05.SIVcpzLB7    ---------C-AACC-----AG--G-G--G-----T...-----------------ATCTAC-AC-G-----AA-AT---A------A-----A--T.........GAA--C-A-------C-GATA---AA-A---------AG--C--CC-AAGA---...A-A--T-- 8915
CPZ.US.85.CPZUS        T----C--TC-CACAGA---GG-AG----GC-----...---------------C--A--TAC--C-G-----CA--T---A------A-----A--T.........GAAC-TG-C----C---ATC---CA---------CTGAA--G--TCGGAGA---...AGA--TAA 9422
CPZ.GA.88.GAB1         -----G--CC-GACAGAG--GC--G-----C-A---...---G-------T----A-T--C----G--T--CA-AT-T-A-------------A--T.........GA---C-A---G--C--G--TC---C-C----------A-G--G--C--AAGA--T...A-T---A- 9456
MON.CM.99.L1           T-GCGC--G--A-CC...ACC-C-GAG--TTCC-AG...G-GGG--AT---G--TT-CT-C-AACC--T--G-CCTAT-AAGGCC-------G---.........--CC-T---C-GT-C--G-TC-TCTCA--CTG-TC-AAG--G---G-CA-GAG-GGCAGACAG-T 9238
MON.NG.x.NG1           --GCGT--G--A-CC...ATCC--GA---CTC--AG...G-CGG--AT--YG-GTT-CT-C--ACC-----G-CCTA--A-GG-C--GAT--A---.........--CC-T-AGC-GT-C--G-TC-TYTCC--CTGTTCTAAG--G-G-G-CA-GAG-GGCCCTCAG-T 7875
GSN.CM.99.CN166        --GCGG-----A-CC...ACA-CTGAG--CTCT-AG...CC-GGG-AT--TG-ATATCTGC--ACC-------CCTAT-A-GGAC---GT--A---.........-A-C-T-AG--GT-C--G-TC-TCTCC--CTG-TC-AAG--G--CA-AA-GTCAGGCA-ACAG-T 9203
GSN.CM.99.CN71         T--CGG-----G-CC...ATG-CTGAG--CTCT-AG...CCCGGG-AT--TC-ATATCTG---AAC--T--C-CCTAT-A-GGACA-CAA--G---.........-ACC-T--G--GA-C--G-TC-TCTCC--CTG-TCTA-G-----C--GA-ATCAGGCTG-CAGAT 9200
MUS_1.CM.01.1085       --GAG-GA---G-CC...ATG-C-GAG-GCTCT-AG...G-GGG--AT--CC-ATA-CT-C-AGAC-------CCTTT-A-GGGA-AGAA--A---.........-ACC-T-AGC-GT-CC-G-TC-TCTCC--CTG-TC----T-G--CC-GA-GTCAGGCT-ACAA-T 9209
MUS_1.CM.01.CM1239     ---AGGGA---C-CC...ATA-CTGAG--CTCT-AG...G-GGGG-AT---G-ATATCT-C-CACC--TG-A-CCTA--AAGGCAA-GA---G---.........--CC-T-AGC-GT-C--G-TC-TCTCA--CTG-TCT-----C--CA-GA-GTCAGGCAG-CAATT 9178
MUS_2.CM.01.CM1246     T-GAG---G--A-CC...ATG-CAGAG--GTC--AG...CCAGG--AT--CC-GTA-CT-C--AC------G-CCTATG--GG-C--GAT--A---.........-ACC-C-A---GT-CC-G-TC-TCTCC---TG-TCTA-G--G--CC-AA-GTCAGGCCC-CAG-T 9274
MUS_2.CM.01.CM2500     T--AGG--T--GC-C...ATA-CTGAGA-CTCA-AA...CTGGGG-AT--TG-ACATCT-C---GC--T--C-CCTT-A-TGGGA--GAA--A---.........--CC-T-AG--GT-CC-G-TC-TCTCC--CTG-TC------G--TC-CA-GTCAGGCCGCCAGTT 9230
RCM.NG.x.NG411         T----------A--T...-TC-GTGAT----CAA-A...G-G--T..-----GGCATT-TC-CC-G--T--A-CAGAAACTAG---------C---.........--CTG-G-G--G--A----CA----A----A-ACC--T-T-----G-AG--T-....CA-GGTA 8994
RCM.GA.x.GAB1          T-G--G-----A--T...-TATCTGAT----CAAGA...G-AG-T..--AG-GCAT--T---C-----T--A-CAGAAACAAG---------A---.........--ATG-G-G--G--C--G-CC----A----A-TCCTATG--G---G-AG--T--...A-A-G-TA 8911
DEN.CD.x.CD1           T-G--C--G--GACT...ATAG--GAG---T-T-A-...-----G...---G-A-ATT-TC--C-G--TGACAGATATGA-GG-CAACAA-C----.........--TTG------GACCC-G-TA-----A------CCTGAAT----C-GGTG-T-T...AAA...GC 9368

MAC.US.x.239           T----C--T--AA-T...-TATCAGAT--G-CACAG...G-GG-T...---G-GCATTAT--AA-G--T--A-CTCAAACTTCCCA-TG-------.........--TTG-G----G--TC---CA----A------TCCAACT--G--C-ACAC-T-T...-A---ATA 9753
Nef                     L  V  P  V  N  .  V  S  D  E  A  Q  .  E  D  .  E  E  H  Y  L  M  H  P  A  Q  T  S  Q  W  D  D  .  .  .  P  W  G  E  V  L  A  W  K  F  D  P  T  L  A  Y  T  Y  .  E  A  Y
H2A.GW.x.ALI           -----------A--T...-TCCCACA-----G--AG...G-C-CT...----CTCA-T--C--C-------A--ACAAACAAGCA--CAT-------.........A--C-TG-G--GACA----TT-----A------CCTAAG--G--TCA-G--T--...AAA---TT 9795
H2A.DE.x.BEN           ------T----A--A...CTCTCACA---G-CA-AG...G-AG-T...----GC--A-T----AG-------A-CACAAACAAGCA-ACAT-----T.........GA-C-TG-G--GACA----T----CA-------TC-ATG--G--C-A-A-CT---...AA----TT 9796
H2A.SN.x.ST            -----------A--T...-TCCCACA---G-GA-AT...G-C-GT...----CTCA-T----AG-G--T--A-CACAAACAAGCA--T-T-------.........---C-TG-----ACA----TT------------CC-ACG-----T---AGCT--...-A----TT 9235
H2B.GH.86.D205         --G--------A---...-TGCCAGC--CGACCCGA...G-GG-G-AG--A-C-CATT--C-AA-G-----G-CACAGA-CTCCTCATG-------.........ATCC-TG-G--GACTC-TATC---CA------TTC--T---G----A-G--T-T...-----TTT 9712
H2B.CI.x.EHO           -----G---A-AA-T...ATG-TAGC----CCA-AG...G-CG-G...--A-C-CATT-TC--G-G--T--A-CACAGACCTCC-CATG-------.........--CC-CGA---GACCC-T-TC---CA-------TC--T--------A-G-CT-T...------TT 9683
H2G.CI.92.ABT96        -----------CAC-...-TCTCTGAT--G-CACAG...G-GG-C...----CTCATT-----G----T--A-CACAAACTAGCCCATG-------.........-------CC--------G-C-----A-------CCTACAT----C-A-G-CT--...AG---ATT 9181
H2U.FR.96.12034        ----A------A--T...-CA--TGATA---CACAG...G-GG-T...----GGCTGTATC--G-GGGTT-A-CACAGACCAGCT-CGA---G---.........-ACTG-G----G-CC----TA----A------TTC-TCT--G----A-AGCT-T...CA---ATT 9277
MAC.US.x.EMBL_3        T----C--T--AA-T...-TATCAGAT--G-CACAG...G-GG-T...----GGCATTAT--AA-G--G--A-CTCAAACTTCCAA-TG-------.........--TTG-.----G--TC-GAC-----A------TCCAACT-----C-ACAC-T-T...-A---ATA 8874
SMM.US.x.H9            T---WC--G--AA-T...-TCTCAGAT----CTCAG...G-AG-C...----CACATT-----A-G--T--A-CACAAACTTC-CA-TG-------.........--CTG-G----G--AC-G-CA----A------TCCARAGT----T-A-R-CT-T...AA---ATT 9236
SMM.US.x.PGM53         T-G--C-----AAGT...-CCTCAGAT----CTCAA...G-AG-C...----CACATTA--AG-G--T--A-CTCAGA-ATCCCA-TG------T.........--CTG-G----G--TC-G-CA----A------TTCA-AA-----C-ACAGAT-T...-A---TTT 9690
SMM.US.x.SIVsmH635F_L3 T----C-----A--T...-TCTCAGAT----CCCAG...G-AG-C...----CACATT--C--G-G--T--A-CACAGACT---CA-TG-------.........--CTG-G----G--AC-G-CA----A------TCCAGAAT-G--T-A-AGCT-T...AA---ATT 9775
SMM.SL.92.SL92B        --G--G-----A--C...-TCTC-GAT--G-CAAA-...---G-T...----C-CATT-----G----T--A-CCCAGACAT--CA-CATTCA---.........--TTG-G--A-G--CC---CC----A----A-TCCT-A------C-ACACAT-T...-AA--ATT 9204
STM.US.x.STM           -----------A--T...ATGTCA-AT--G-CACAG...G-GG-T-AT-G--CACATTATC--G-G--T--A-CACAGACA---CA-TG-------.........--TTG-G-------AC-G-TT----AA-----TCCA-TAT----CCA-AC-T-T...-A----TT 9407
MNE.US.x.MNE027        T----C--T--AA-T...-TATCAGAT--G-CACAG...G-GGGT...---G-G-ATTAT--AC-G--T--A-CTCAAACTTCCCA-TG-------.........--TTG-G----G--TC---TA----A------TCCAACT-----C-ACAC-T-T...-A---ATA 9230
DRL.x.x.FAO            --G--C--T--G---...-TAG-TGA---TCT--TA...---C-G...C--TGT-A---GC----GA--T-CAGCCAGA--GG-CCT--T-------.........--A--AG-----AGAC-CAT----CA------TCCAG----G----A-ACATTT...CA---TT- 9089
MND_2.x.x.5440         --G--C--G--G-CA...-TGTC-CCCCCCCT--AG...G-GG-T...--ATG--A---GC--C--A--T--TCTCAAA--GGAATCCA---A---.........--A---G----GAG---GAT-----A------TTCAG-G--T--T-A-ACGTT-...TAT--T-C 9004
MND_2.CM.98.CM16       --TTGC--G--G-CA...-TGCCTCC---CC---AG...--C--T...---TGT-A--AA---C-C--GT-CTCTCAG---GGAATCCA---A--A........T--TTG-G----GAG-C---T-----A-------TCTG----T--T-ACACCTTT...TAT----C 9309
MND_2.GA.x.M14         T--A-C-----AC-A...-TACCACC---CC---AG...--CC-G...---TGT-A---GC--C-TA--T-CTCTCAG--AGGAATCCA---A--A........T--ATG-G-G--GAG-C-GAT-----A-------CCTGAAT-G--T-A-ACATTT...TAT----C 9250
OLC.CI.97.97CI12       --TAGGAG---G--A---TTGCCTGAGACCCTT-AAGAG-GAGGG............................................................................................................. 8678
GRV.ET.x.GRI_677       -----C-----G--C...CTGC-TGAG--G-CACG-...--CTGT...----GACA-T-TC--A-G--T--A-CACAGA--GGG-AAGATCC---TGGAATA...GATC-TG-------C--G-TC----A-------CCGAAGT-G--GG-GG-GT--........... 9226
VER.KE.x.TYO1_patent   -----G-----G--C...TTAC-TGAG---CACAA...-CATGT...--A-GACATT--C-AG-C--T--A-C-CAGA--GGA-AAGATCCA--TGGTATC...AGCC-TG---GA-C--G-T-----A------TCCTATGT-G---A-A--GT--.......... 9343
VER.KE.x.9063          --------G--A---...ATGC-TGA---G-CA-AG...-CCTGT...----GACATTA---GCG--T----CACAAA-AA-A-A-GACCCC--TGGCATC...AGTC-TG---GACC--G-TC----A-------CCTATGT-----G-G--GT-T.......... 9382
VER.KE.x.AGM155        --------G----CC...CTGC-TGAG----CA-AA...-CATGT...--A-GGCATT-----G-------A-CACAA-------AAGACCC---TGGTATA...AATC-TG-----A-A--G-CA----A------TCCAATGT-G--TG----GT--.......... 9359
VER.DE.x.AGM3          T-G--------A--C...CTGC-TGAG----CA-AG...-C-TGT...----GACATT--C--G-G--T--A-CACAAG--AGG-AAGACCC---TGGAATC...AACC-TG-------C--G-TC----A------TCC-ATGT-----G-C--AT-T...-ACC-TAA 8888
TAN.UG.x.TAN1          T----G--G--A--T...-TCTCCCAG---CCCAG...G--..........----GACA-T-T---C-C-------CCCAGA-TG-AT--GAAAGC----.........--TTG--A---GACA--G-T-----AA---A-TCCA-T-------G----GT--.......... 9360
SAB.SN.x.SAB1C         --G-----T--G--C...CTGTCAGAG--G-CTAAG...----GT...------CA-T--C--C-G--T----CTCAAG--GCCTATGA-------.........G-CTG--A---GACCC-G-T-----A-------CCA-TA------G-GG-CT-T...--T--TT- 9561
MND_1.GA.x.MNDGB1      --G-A------GAGCAG--C--TAGG---T---TAT...GCAGCT..........----ATCATC--------T-CTCCCAGT-ATG-CCTCA---A--T.........--A--AG---GACCC-CAT----TCTGGGAC-CT-AT-T-GC-TA-GAC-TTG...CA-CATTA 8813
DEB.CM.99.CM40         --G--G-----A--A...ATAGCAGACCC---TTAT...G-G--T-AT----GG-A--T---------TGA--CCCA--A-GGGAT------A---.........--CC-T-----GAG---G-TT----A----------ACT--G--C-A-TG-T-T...AA---AG- 9021
DEB.CM.99.CM5          -----G-----A---...ATAGCAGATCCT--CTAT...G-G--T-AT----GA-AT-T-C-C-----TGA--CCCA--AAGGACAAGCA--A---.........--CC-T-AG--GAG-C---TA----A------TTC-TC---T--C-A-TGCT-T...AAA--AG- 8948
TAL.CM.01.8023         --GTGC--C--G---...ATCC-CGAG--C--CA-G...G-GG-C...-GAC--...CT-C-CA-G-----A-CATATGACGG-CA-CA---A---.........--CTG-G-C----CCC-G-TC---GTC-----TGAGAAG--G--C-ACACC-CAGGA-C-AAAAT 8691
TAL.CM.00.266          --CTGC-----A---...ATCC-TGAG--T--CA-G...G-GG-C...-GAC-T...CT-C-CCCG--T--A-CATATGATGGGCA-GCA--A---.........--ATG-G-----TCCC-G-TC---GTC-----TGAGAAG--G--C-ACACC-C-GGA-C-AAGAT 9141
SUN.GA.98.L14          ----C------CACC...AT-G-TGA---CAG--AT...CC---TCATCCTTGTCAAGC-C-AC-G--TT-CAGCCAG-AAGGG-TT-AT--A---.........--TTG-G----GAGAC-TATC----CA-----TCC-ACT--G--C-A-G--TTT...AG----AT 9532
WRC.CI.97.97CI14       T-G--T-----AC-C...TTC-GTGA----TCA-AT...-----C-...--CCCATA-TATGA-AG-A-TAAGT-ACTG-AT-CA-CATGT-C---GCCAGAGGAT--CTG-G-----CATC-G-TC----A-------CCTAAG--G--C-G-G--TTTGTA-CA-GAA- 9340
WRC.CI.98.98CI04       T----G--T--G--A...TT--GTGA---GTCTAAT...G-C-GC...TCTC-GTA-TA-GA-AAGA--AAGC--TTA-AT-CA--CTGCC-A--ATCAGAGGAT--TTG-G-----CAC--G-TC----A-------CCT--A-----C-G-G-CTTTGTG-CAAGTA- 9382
WRC.GM.05.Pbt_05GM_X02 T-G--T-----GC-C...TTC-GTGA---GTCTC-T...--CTC-...--CC-AGG-TTTGA-AAGA---TGC--CTA-AT-CA-CAG-A-C---GTCAGAAGAT--ATG--AG---TAC--G-TC----A------TCCTA-G--G--T-G-G--TTTGTG-CT-GGA- 8278
LST.CD.88.524          T-G-AG-----A-CA...TG-G-TGAGTTTA-A-AT...CCCTC--AT----GACAAT-T---C----TT-AAGCCAGT--GGG-TCC-A--A---.........--CTG-G-----AG-C-CATC---CAC---A-TCC-ATGT-----G-GG-CT-T...--A----T 8386
LST.CD.88.447          --G-A------G-CC...TG-G-TGAGT-CACA-AT...CCCTC--ATT-C-GACAGT--C--C-G---T-AAGTCAG---GGG-TCCA---G---.........--CTG-G-----AG-C-TATC---CAC---A-TCCTATG------G-GG-CT-T...--A----T 8386
LST.KE.x.lho7          T---A------G-CT...TG-G-TGAGT-TA-A-AT...CCCTC--AT----CACAAT--C-T--G--TT-AAGCCAG---GGG-TCC----A---.........--CTG-G----GAG-C-TATC---CAC--CA-TCCAATGT-G---G-GG-CTTT...A-A----T 9461
SYK.KE.x.SYK173        -----T---AA----...ATAG--GG-TGCTT--AG...T--G-G...---C-T--ACT-C-CC-C--T--A-CCTCTGGA--G----GTTCGAG-ATGGGAGAA--TC-TGTG---C-TCA-CCCCCACCTGGCT-T-CA-CAGGTTGGGAGATGGC-...CGCTTG-A 9215
SYK.KE.x.KE51          T-G-CC--TAA---A...ATAG-AGGGT-CTT--TA...G--G-G...--AG----ACTGA--A-G-----A-CAGCAGGAGTG---GCTTCA--AGAC......--AC-T--G--GAATC-CAT-----AC--CA-TCCA-A-T----C-A-ACA-CAGGCTG--AAAT 8902
COL.CM.x.CGU1          T----CTACA-A---...C-ACTT....................................................................................................G-AGAGTAT----ACAGCT-TGACT--AGT-T 8541
```

```
                                          TCF-1 alpha binding
                                                        Nef end
H1B.FR.83.HXB2           AGAGCTGCATCCGGAGTACTTCAAGAACTGC...TGA..................................................................................CATCGAGCTTGCTACAA......... 9434
Nef                       _E__L__H__P__E__Y__F__K__N__C__.__*_
H1A1.UG.85.U455          T---------------T--AT--AG--.......---.....................................................................AAGACTGCTGA--CA--AG-----GAC-......... 8893
H1B.US.90.WEAU160        --------------------T--AG----.....---.................................................................................-------T--T------........ 9434
H1C.ET.86.ETH2220        C-----A-----------T-A---AG----....---.................................................................................--CA--CGGGA--TTCCGCC...... 8837
H1D.CD.84.84ZR085        ---AAA-T--------------------......---.................................................................................---------C................ 8975
H1F1.BE.93.VI850         ----AGA---------T--A-C-AG--.......---.....................................................GACTGCTGACACAGAGATTGCTGA--CA--AGAAT---A--.......... 8732
H1G.SE.93.SE6165         ------------------A---AG----......---.................................................................................--CA--AG-----GAC-AG....... 8878
H1H.CF.90.056            --TAAA-------------...--AG----....---.................................................................................--CA--AGA-.--GA--.......... 8753
H1J.SE.93.SE7887         ----------------TT-A---AG----.....---.................................................................................--AA--AG--T---GCG......... 8732
H1K.CM.96.MP535          --------------C----A---AG----.....---..................................................................................................... 8604
H101_AE.TH.90.CM240      ---A-A------A-------AT--AG----....---.................................................................................--AG--AG--T--.ACT......... 9001
H102_AG.NG.x.IBNG        ------------------A---AG----......---.................................................................................--CA--AG-----GAC-......... 8961
H1N.CM.95.YBF30          ----------------TT-A-----------...---.................................................................................--AG-GA---TACTGCTGACAA.... 9029
H1O.BE.87.ANT70          GATAACT--C--A---CT----C----GGA-....-A-.....................................................................AAACTGCTGACCT--AGA-TGCTG-CACTGT.... 9538
H1O.CM.91.MVP5180        GC-AAA---C--A---CT----CCC--G.......-A-...........................................................................................CTGACACTGC.......... 9537
CPZ.CD.90.ANT            ---AAGA--C--T--A--T----GAG--.......-A-.......................................GACTGCTAACAGGCGGATATAAAACCGCTGACAG-G-AGA----TGGTCT......... 8903
CPZ.CM.01.SIVcpzCAM13    ----AAA-----------T-A---AG--......---....................................................................CTGGCGCATGCG-C-AAGAAG--..CTG-CTGTGC... 9086
CPZ.CM.05.SIVcpzEK505    ----------------T--A-C------......---.................................................................................--A....................... 8927
CPZ.CM.05.SIVcpzMB66     ------A-----------A---AG-T---.....---.............................................................................CAA-GGACTTG--..CTG............ 9078
CPZ.CM.98.CAM3           ---A-AA-----------T-A-C-----......---....................................................................TGCCCAGCGCCTGCG-AGTAGAAA.CTGCTGACAA... 8891
CPZ.GA.88.GAB2           ---AAAA-----T-----T-A--GAG--......---...............................................................AACCGCAACTGGCGCATGCG-TCT..A-AACTGCTGAGCCTGA 8866
CPZ.CM.98.CAM5           ---A-AA-----------T-ATGGAG--......---.................................................................CACAACCCAGCGCCTGCG-AGTAGAA..CTGCTGACAA... 9193
CPZ.CM.05.SIVcpzMT145    ----AAA--------A--T-A--GAG--......---..................................................................CCACAGCGGCGCATGCG-ACT.AA-AACTGCTGATCAACA 8986
CPZ.TZ.00.TAN1           ---AAGAT----------------AT-A..........................................GGACTTCCGGGTGCCATGACTCAGAACTGC-GACAGAG-ACTTTTGGACTC... 9133
CPZ.CM.05.SIVcpzLB7      ----------T---------A-------......---.................................................................................--A....................... 8952
CPZ.US.85.CPZUS          ------------A---------CG----......---..............................................................TGGCTGTAACCGCGCAGGCG-A-TAAAA..CTGCTGACTG... 9491
CPZ.GA.88.GAB1           ---A-AA-----------A---AG--........---...............................................................CTTTAGACTGGCGCATGCG-ACAAGAA..CTGCTGACTCTGC 9527
MON.CM.99.L1             --CA-A--T--A-C-AG-GGAGCG---GAAGCGC-T-GCCGCAAACCGCATCCTCTAG.........................................................................-ACTCTGCCC... 9306
MON.NG.x.NG1             ---TGA-TTA-A-C-AG-AGAACG---GA-GCGC-T-ACCGCAAACCGCATCCTATAG....................................AACTCTGCCC........................................ 7943
GSN.CM.99.CN166          ---C-A--TG-A-C--G-GGAGCG---GATGCGCCT-ACCGCAAACCGCTTCCTATAG.........................................................................-ACTCTGCCC... 9271
GSN.CM.99.CN71           GA-C-A--TG-A-C--G-GGAGCG---GAAGCGCCT-ACCGCAAACCGCTTCCTATAG.........................................................................-ACTCTGCTC... 9268
MUS_1.CM.01.1085         ---C-AAATG-A-C--G-GGAGCG---GAAGCGC-T-GCCACAAACCACATCCTATAG.........................................................................-AC-TTGCCC... 9277
MUS_1.CM.01.CM1239       ---T-AAATA-A-C-AG-AGAACG---GA-GCGC-T-ACCGCAAACCGCATCCTATAG.........................................................................-ACTTTGCCC... 9246
MUS_2.CM.01.CM1246       TA-T-AAATA-AAC-AA-AGAGCG---GA-ACGC-T-ACCGCAAACCGCATCCTCTAGAACTCT......GCTA.............................................................................. 9342
MUS_2.CM.01.CM2500       -A-C-A-ATG-A-C--G-GGAGCGA--AC-GCGC-T-GCCGCAAACCGCATCCTCTCGCAGTCT......CCATAA....................................................CTAC............. 9304
RCM.NG.x.NG411           TAGA--------T----T---TGGAG-GA-A...AAGAACAAGACCAAGTAA.......................CCACTTCCTCTGGAGTTGCCTTGGAAACAGCG-AT--C-A--GAG-AGATGCTGA 9098
RCM.GA.x.GAB1            TAGA--------A----T---TGG-G-AA-GAAGAACAAGACCCAGTAA...............................................CCACATCCTCTGGGGTTGC-T-G-TAACCAGGCAG-AGAATCTGC 9005
DEN.CD.x.CD1             --GC-ATA-GAA-CTTC-AACTTCAC-GCATAATAACTACTGCCAGAGAGCTAAGCCCAGGAAG......TAG...........CCTAACCGCAAACCACATCCTACTGCAGAACTTGCTGAGAGGCTCA..................................... 9481
                                          TCF-1 alpha binding                                                                  Nef end
MAC.US.x.239             T-TTAGAT-C--A--AG-G--TGGA-G-AAGTCAG-CCTGTCAGAGGAAGAGGTTAGAAGAAGG......CTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAAGGAAACTCGCTGA................................. 9868
Nef                       _V__R__Y__P__E__E__F__G__S__K__S__G__L__S__E__E__E__V__R__R__R__.__.__L__T__A__R__G__L__L__N__M__A__D__K__K__E__T__R__*_
H2A.GW.x.ALI             TATT--A--C--A---G-A--TGG-T--AAGTCAG-CCTGCCAGAAGATGAGTGGAAGGCAAGA......CTGAAAGCAAGAGGGATACCATTTAGTAAGAACAGGAACAGCTGA..................TTTGGTCAGGG--GGA--TAACTACT.GAAAACAGCT 9940
H2A.DE.x.BEN             CACT---T-C--A--AG-G--TGG-C--AAGTCAG--TTGCCAGAGAAAGAATGGAAGGCAAAA......CTGAAAGCAAGAGGGATACCATATAGTGAATAA................CAGGAACAACCATACTTGGTCAAGG--GGA--TAGCTACT--GAAACAGCT 9944
H2A.SN.x.ST              TATT-GAT-C--A---G-G--TGG-T--AAGTCAG-CCTGCCAGAGGATGAATGGAAGGCAAGA......CTGAAAGCAAGAGGGATACCGTTTAGCTAA...............AAACAGGAACAGCTATACTTGGTCAGGG--GGA--TAACTA---GAAAACAGCT 9383
H2B.GH.86.D205           CA-TAG-TT---A--AG-G---GG-T--CAGTCAG-GCTACCAGAGGAGGAGTGGAAGGCTAGA......CTAAAAGCAAGAGGGATACCTACAGATTAG..........GCAAGAGACAGCAGCATAAACAGGAACTAGCTGACA-TGCAC.AAG-AGGAAACTAGCA 9864
H2B.CI.x.EHO             CAGCAG-TTC--A---G-G--TGG-T-TCAGTCAG--ATGCCAGAGAAAGAGTGGAAGGCTAAA......CTGAGAGCAAGAGGAATACCTACAGAGTAG..........ACAGGAAATAACAGCACAAAGAGGAACTAGCTAACA-TGCA-A-AG-AGGAAACTAGCT 9836
H2G.CI.92.ABT96          TATAT-------T---G-G----G-TGGAAGTCAG-GCTACCAGAAGCAGTATGGAAAGAGAAA......CTAAAACAGAGAGGACTGCCTATAGAATAA.........................................GCAGA-AG-GGACAAGCT 9293
H2U.FR.96.12034          TATTAA-TGC--A--AG-G--TGG--GTAAGTCAG-TCTGTCAGAGGAAGAGGTTAAGAGGAGG......CTAACCGCAAGAGGCCTCCCAGTAAAGAATTGCTGA..................................CAGA-AGGAAACAAGCT 9394
MAC.US.x.EMBL_3          T-TTAGAT-C--A--AG-G--TGGA-G-AAGTCAG-CCTGTCAGAGGAAGAGGTTAGAAGAAGG......CTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAGGGAAACTCGCTGA...G..ACAGCA............................ 8996
SMM.US.x.H9              T-TTGA---C--A--AG-G--TGGT-GTCAGTCAG-CTTGTCAAAGGAAGAGGTGCAGAGAAGG......CTAACCGCAAGAGGCCTTTTAAAAATGGCTGACAAGAAGAAAACAAGCTAG........................................ 9351
SMM.US.x.PGM53           CATTAGA--C--A--AG-G--TGGC-GTAAGTCAG-CTTGTCAGAGGAAGAGGTAAAGAGAAGG......CTAACCGCAAGAGGCCTTTTAAAAATGGCTGACAAGAAGGAAACAAGCTGA........................................ 9805
SMM.US.x.SIVsmH635F_L3   TATTAA-T-C--A--AG-G--TGGT-GTAAGTCAG-CTTGTCAGAGGAAGAGGTAAAGAGAAGG......CTAACCGCAAGAGGCATTTATAAAATGGCTGACAAGAGGGAAACAAGCTGA........................................ 9890
SMM.SL.92.SL92B          T-TAAG---C--A--AG-G--TGG-TGGAAGTCAG-CTTGACAGAGGAAGAGGTTGAGAGAAGGTTAGCTAACAAGCCAAAACCGCAGAAGAAGATGGCGGATAAGAAGGAAACAAGCTGA........................................ 9325
STM.US.x.STM             T-TTAGA--C--A--AG-G--TGGTTCTAAGTCAG-CTTGCCAAAGGAAGAGGTTGAGAGAAGG......CTAACCGCAAGAGGCCTTCTAAAGATGGCTGACAAGAAGGAAACAAGCTGA........................................ 9522
MNE.US.x.MNE027          TATTAGAT-C--A--AG-G--TGGA-G-AAGTCAG-CCTGTCAGAGGAAGAGGTTAGAAGAAGG......CTAACCGCAAGAGGCCTCTTAAAAATGGCTGACAAGAAGGAAACTAGCTGA........................................ 9345
DRL.x.x.FAO              TATT-AC-----A---G-G--TGGAC-TGTTAAAA-TATAAAGTCAGTAGAAGAGCCAGGT..................................TGTTGCAAGAGGAAGTGGTGGCAGTTCAAGCCAACCACAGAAGGTTGTC-TGGA-AC-ACCTTCAGAAGT 9220
MND_2.x.x.5440           TATAA-TAGG--A---G-A-A----TGTGTG.................ACAAGTCTGAGCTAT......GAAGCCTACAAGAAGGAGGAAAAGCCTGACTGCTGCAAGAGGAAGTGGTGGCAGTTCTAGCTAACCACAAAGGGTTG-T-TAAGCAACTGCTAACTCTGG 9150
MND_2.CM.98.CM16         CATAAA-AGA--A-GTG-A------C-TGTG...CA-AGCATCAGCCATATTGCTTATGCTAAG......GAGCATAAGCCAGAGTGCTGCAAGAGGAAGTGGTGGCAGTTCTAG...............CCGACCACAGAAGGTTG-C-TGGA-A-GACCTGCAGAAAT 9455
MND_2.GA.x.M14           CATAAAATGC--AA--G-A-A--G-C-TGTT.................ACAAGCCTCTCCTAT......ACTGCCTACCAGGAG......AAATCAGACTGCTGCAAGAGGAAGTGGTGGCAGTTCTAGCTAACCACAGAAGGTTG-T-TGGA-A-G-TCTTTAATTCA 9390
OLC.CI.97.97CI12         ............................................................................................................................................... 8678
GRV.ET.x.GRI_677         ........-GC-----CATG--T---G--ATGCACGA-CATGCAAAGCGCTAG..........................................TGTCAGCACTTTGCGGTT................................... 9289
VER.KE.x.TYO1_patent     .......G-C--CA-TCGGGAGT-CTTTACTGACATGCATGGGCTGGTGAAGAGGAAGTAG......................CCAGACCGCAAGCCTGCGGTTAGAACATCACCATGGAGATGAC-TTAAAAACTGCTGAC..... 9456
VER.KE.x.9063            .......G-C--AA-CAGACAGT-TTTTGAAGACATGCATGCAATAGTGAAGAGGAAGTAG.......................CTAACCGCAGGCTTGTGGTTAAGCCGTTGCCAGGGAGATGAC-TT-GAAACTGCTGACAG... 9496
VER.KE.x.AGM155          .......G-C--CTCAAGGGAGT-CTTTACAGAC-T-TATTCAACAGTTGGTACAGGAAACTAG......................CCGACCACAGGCTTGCGGTTTCCTGGTTGCCTAGGAGATGAC-TTAA-AACTGCTGAC..... 9474
VER.DE.x.AGM3            CAGAAAAT---TCACTG--A-GC-TG-TCTT...G-CAAGAGGAAGTAG..........................................................CT.AACCGCA.GGCTTGTGG-TA---ACATC--C-TGGTGATGA 8976
TAN.UG.x.TAN1            .......A-C-----C......-GCTTTAAGGACATGCATGGTCTTGTTAAGAGGAAGTAG.......................CCTAACCGCAGGCTTGCGGTTAAGCTAACTAGGCCGTTGCCTAG-AGATGACAT-AAGAAGTGC 9472
SAB.SN.x.SAB1C           GAG------C--A---C-GG-GCC-TCAGCA...CA-GGCTAA.........................................................CCACAAAACCACATCCTATGGAGTTGTCATG-TGATGACATTAAGAACTGC 9652
MND_1.GA.x.MNDGB1         -C-TA-CACC-T-AGTGTT-CA-TA-GGCTAAGAGT-TTGAACATCTGCCATTTTGGAAGAGGA......AGTAG...................................................CCTAACCGCAAAACCACATCCTAC---AGA-CTGTAGTTGC 8923
DEB.CM.99.CM40           CC-TGCTG-A-AAAGAG-GCAT-CT-GAA-ATGTATGTTCCCTAAGAGGAAGTAG.................................................CTAGCTTAACCGCAAACCACATCCTA--G-AGCGCCCG............. 9114
DEB.CM.99.CM5            -C-TGAAGCC-ATA-ACCA...G--C-TAA-AGAA--TGCATGTTCCCTAAGAGGAAGTAG...........................................CTAGCTTAACCGCAAACCGCATCCTCTTG-AGAGAGC-C............. 9045
TAL.CM.01.8023           G-CTGA-T--GACAG--TGGAG-G---GAAGAAGGAGCTGCTAGCTCCACCGCAAACCGCATCC......TCTTGA...........................................................................AAG 8764
TAL.CM.00.266            G-CCGA-TGGGACAG--TGGAG-GAG-GAAGCGGATGCTGCTAGCTCCACCGCAAACC.................................................................................GCATCCTCTTGAAG 9213
SUN.GA.98.L14            -C-AAA---C--T--AG-G------C-TGTTACTA-TCTGCAGTGGGAGGTGTAG..............CTTAACCGCAAACCGCA.....TCCTCTTGCATCGCC..TAGGCAACGGGGCTAGCGCATGCGCGCTAGAGTA.......................... 9653
WRC.CI.97.97CI14         -CTC-CTGGA-A-TCAAT-GC--CTGGGAAAGAGGCCATTGAGAGTCTCTATAAG...AACAAG......AAGCAGTGCAAGTAG.........................................................-GACATG... 9423
WRC.CI.98.98CI04         -CTC-CTGGA-ATATAATAGC-C-AGGAGAAGATGCTATTGAGACCCTCCATAAAGAGAAGCTG......AAGCGGCACAAGTAG.......................................................................... 9461
WRC.GM.05.Pbt_05GM_X02   -CT-GCCG-CGGTC--ATAGC--CTGGTATTGAGACCCTC......TATGAGATG...AAGAAG......AGAAACAACAAGTAG.........................................................-GACTTG... 8355
LST.CD.88.524            GAGAAA---A--A-CAAGTG-TC-AGCAACTGCTATGAGATTTAACTGCGAGAGGAAGTAG.........CCAAACCGCAAACCGCAGAACATCCTCTTGCAGCCCGGTTGCCCTAGCAACCGGGCTAGCGCATGCGC............................ 8515
LST.CD.88.447            CAGAAAA--A--TA-AAG-A-AC-AGCA-CAGCTATTGGCTTG......CAGAGGAAGTAG.........CCAAACCGCAGACCGCAACACATCCTCTTGCAGCCTGGTTGCCAGGGCAACCAGGCTAGCGCATGCGC............................ 8509
LST.KE.x.lho7            CA-AAA---G--T-C-A-GA-AC-A--TACTGCT-TCGCTTTTGACTGCAAGAGGAAGTAG.........CCAAACCGCAAACCGCAGCACATCCTCTTGCAGCCCGGTTGCTAAGGCAAC...CGGGCTAGCGCATGCGC....-AT-GCT-GCTG-...CGAGCA... 9609
SYK.KE.x.SYK173          GTTAGA-AGA-AAACAGGAAAACCAC-AGAG.........CTGCAGTCAGCACTTAGTAAGAAC......ATCTCCTAG.................GAGACTCCATGGTGACAAGCTCGGCCCAGGGGA................................ 9312
SYK.KE.x.KE51            G-CT-GT--G-AATTAG-GAGAC-A-CAG-AAAGC-CTAGCTCACCGCAGAAACCACAGGGTAC......ATCTCCTAG.................GAGACACCTAGGAGACAGAGACCGCCCAGGGGAGGAGTCGGGGCGGGGGAG-CT-ACC-A--CTGCATAAAAG 9048
COL.CM.x.CGU1            GTT-AGCGT-A-AAGCAGGAAG-----GAAACCACA-CAGGTTGCCATAGAGATGGTGGACTGA...............................................................................GTTGCT 8611
```

```
                              NF-κ-B-II                        NF-κ-B-I           Sp1-III                              Sp1-II    Sp1-I                    TATA box
H1B.FR.83.HXB2              ..................GGGACTTTCCGCTG..................GGGACTTTCCAGGGAGGCGTG.................................GCCTGGGCGGGACT.GGGG.AGTGGC.GAGCCCTCAG.ATCCT....GCATATAAGCAG..C 9521
H1A1.UG.85.U455             ..................----------A---..................----------GA--GA-G.--T.................................G-TT-------AGT-.----.G------.T---------.--G--....-------------..- 8981
H1B.US.90.WEAU160           ..................----------A---..................----------G---------.................................AT----------A.----.------.----------.--G--....-------------..- 9521
H1C.ET.86.ETH2220           ..................----------A---..................----.-G-------AG-AG-G.................................-T-----------.----.------.C-A-------.--G--....-------------..- 8921
H1D.CD.84.84ZR085           ..................................................................................................................................................................... 8975
H1F1.BE.93.VI850            ..................----------A---..................------------A-G-CG--CC.................................AGAG--.-------.----.------.TCA-------.--G--....-------------..- 8819
H1G.SE.93.SE6165            ..................----------A-CT..................----------G-------C-G.................................------A--AGT-.----.------.T-A-------.AG--....-------------..- 8965
H1H.CF.90.056               ..................--------------..................----------G---------A.................................T--------GA-.----.------.C-A-------.--G--....-------------..- 8839
H1J.SE.93.SE7887            ..................--------------..................-------------GA-GAGTG.................................-----------GT-.----.------.T-A-------.--G--....-------------..- 8820
H1K.CM.96.MP535             ..................................................................................................................................................................... 8604
H101_AE.TH.90.CM240         ..................A-A---.-------..................-------------GA-GTGTG.................................GC-G.-----AGT-.----.---A--.T-A-------.--G--....-----A-------..- 9088
H102_AG.NG.x.IBNG           ..................----------A---..................----------.---GA-GTGT.................................G-TT---A--AGT-.----.------.T-A-------.C-G--....-------------..- 9048
H1N.CM.95.YBF30             ..................--------ATA--TG....GGGACTTTCCGCCA.-------------.A-GTGT.............................GGTT..GGG--AGT--CT-.-.........................G--....-----A-------..- 9120
H1O.BE.87.ANT70             ..................--A-------AGCAAAGACTGCTGACACTGCG..------------T-G-AG-GA...............................-A---G-C-GT-C.---G-----TA-.-------A-.G--....-------------..- 9642
H1O.CM.91.MVP5180           ..................----------AGAC....TGCTGACACTGCG..------------C-T--GAG-GA..............................TAA--G-C-GT-C.---G-----TA-.-------.--G--....-------------..- 9638
CPZ.CD.90.ANT               ..................----------A--AT.................----..............................................GT-A--AT-G-C-GGCA-T--G-----TTTC----G-CT..G--....-------------..- 8978
CPZ.CM.01.SIVcpzCAM13       ..................----------AA....................----------G-------G-...................................TC.-AG---A-T-GT-.-----.-----TTTC--------..AG--....-------------..- 9172
CPZ.CM.05.SIVcpzEK505       ..................---------TA--TG...GGGACTTTCCAGCA.----------G------........................................T-T-GT-T----.-----TTG.-------..AG--....-------------..- 9018
CPZ.CM.05.SIVcpzMB66        ..................----------A...CA................---GA-----G------....................................GT-A-CAT-G-C-GGGGA-A-G-----TT.--------.-AG--....-------------..- 9165
CPZ.CM.98.CAM3              ..................---------TAA....................----------A--G--G..TG...................................GT-CA-G-C-GGACT.--GC----TTTTA-------.-AG--....-------------..- 8976
CPZ.GA.88.GAB2              ACTTTTACAA........----------AA.............................GG..........................GGGGTGG-TTCAA-G-A-GAGTC---GT-----TTT--------A..G--....-------------..- 8951
CPZ.CM.98.CAM5              ..................---------TAA....................----------A--G--A..-G...................................-TCA--GC--GACA..--GC----TTTCAGTT-TCAG-................................ 9261
CPZ.CM.05.SIVcpzMT145       TTTGCTGACAA.......----------A....................A-------------G--TTG-GT..........................TTGGGGCGTAACT--G-C-GAACT.--G-----TTG.--CA-CA-AG--....-------------CG 9094
CPZ.TZ.00.TAN1              ..................----------AA.........................................................TGTGGGTG-TTACT-G-C-GGACA---G-----TTTT.G--.-GCTGAG--....-------------..- 9207
CPZ.CM.05.SIVcpzLB7         ..................-A------T-TGACCA................----------CCA--G.....................CTTCCGGGGAGGCGT-T-AA--GA--GT-T----.-----TTT-.--------.-AG--....-------------..- 9054
CPZ.US.85.CPZUS             ..................A--------TAA....................----------A--GACGT-C.C..........................AAGGGGGT-GG-CA-G-C-GAACA.--GC----TTTTA------A.-GA--....-----A-------..- 9587
CPZ.GA.88.GAB1              ..................----------AA....................---------G-----AG-.T...................................G-T-G--AG-C-GGGTT---GT-----TTT---------.-G--....-------------..- 9614
MON.CM.99.L1                ..................----------A.........................-T-ACCATGACT-C-G-GCCCGGTTGCTGCGGCAACC.GGGCGGTCCAAAGGACTTG-T-G-T-GACCC---G------TTG.-----TCGT-T.....CA----A-------..A 9423
MON.NG.x.NG1                ..................----------A.........................-TTGCC-TGACT-C-G-GCCCGGTTGCCGAGGCAACCGGGGCGGACTCGTGGGAGGGACTG--AGG-AC---G-----......-T-ACCCTCAGATCCT-----A-------..A 8061
GSN.CM.99.CN166             ..................----------A.........................-T-ACCATGACT-C-G-GCCCGGTTGCTGAGGCAACCGGGGCGGACTCGTGGGAGGGACTG--AGG-.C---G-----TCTA---CGCTAT.--A....-----A-----A..A 9389
GSN.CM.99.CN71              ..................----------A.........................-TTACCATGACT-C-G-GCCCGGTTGCTGAGGCAACCGGGGCGGACTCGTGGGTGGGACTG--AGGC.C.--G-----TCTA---CGCTAT.--A....-----A-------..A 9385
MUS_1.CM.01.1085            ..................----------A.........................-T-ACCATGACT-C-G-GCCCGGTTGCTGAGGCAATCGGGGCGGACTCGTGGGTGGGACTG--CGG-ACT--G-----......-T-ACCCTCAGATTCC-----A-------..A 9395
MUS_1.CM.01.CM1239          ..................----------A.........................-T-ACCATGACT-C-G-GCCCGGTTGCTGAGGCAACCGGGGCGGACTCGTGGGTGGGACTG--AGG-ACT--G-----......-T-GCCCTCAGATTTC-----A----CA..G 9364
MUS_2.CM.01.CM1246          ..................-A---C----A.........................-T-ACCATGACT-C-G-GCCCGGTTGCCGCAGCAACCGGGGCGGACTC.AGGGCGGACT-G-A-GGACCT-AG-----TTG.-------ATC-A.........-----A-------..A 9459
MUS_2.CM.01.CM2500          ..................------A-AA-...............................GCCAGGTTGCCTGGGCAACCTGGGCGAACTCGTGGGTAAAACTG--AGGAGC---A------TT-.-------A--T-.........-----A-------..- 9405
RCM.NG.x.NG411              ..CACAGAA.........----------A---GCGCCTGCGCGCTGGTGTAA----------..........................................AGACT-A--T-GGAG----GT-----TTG.-----TCCT.G--....-------------..- 9203
RCM.GA.x.GAB1               TGATGCAAAA........----------A---GTGCATGCGCACTGGGGAA.----------...........................................GGGAT-AC-TGGGA----G-----TCAG.-----TCCT.G--....-------------..- 9111
DEN.CD.x.CD1                ..................----------A.-.............GTTGCCTA-CA--CC-GGGC---C.............................TCAGGGGA-GGCCT-G-A-GTCT.CT-GGAGT-GCTT.A--CGCTTC-A--.........-----A---T-..A 9579
                             NF-κ-B-II                                     NF-κ-B-I                                                                               TATA box
MAC.US.x.239                ..AACAGCA.........----------A-AAG.................---.A-G-TA....................................CGGGGAGG-ACTG----GGA-CC-GTCG--AACG.---A-TTTC.TTG....ATG-----AT-TCA- 9960
Nef
H2A.GW.x.ALI                GAGACTGCA.........----------AGAAG.................---G--G-AACC..................................AGGGGAAGGACAT----GGA-CT-GT-G--AACG.----.TCAT-.-TC....CTG-----ATGTAC-10035
H2A.DE.x.BEN                GAGGCTGCA.........----------AGAA..................---G--G-AACC..................................AAGGGAGGGACAT----GGA-CT-GT-G--AACG.------TAC.TTA....CTG-----ATGTAC-10040
H2A.SN.x.ST                 GAGACTGCA.........----------AGAA..................---G--G-TACC..................................AGGGGAGGGACAT----GGA-CC-GT-G--AACG.------TAC.TT-....CTG-----ATGTAC- 9479
H2B.GH.86.D205              GACACTGCA.........----------AAAA..................-A-G--G-AACA..................................TGGGAGGA-CAA----GGG-TT-GTTG--AACG.---ATTAA-.T--....CTG-----ATGTAC- 9959
H2B.CI.x.EHO                GATACTGCA.........----------AGAA..................---G--G-AACA..................................GTGGGAGGGACAT----GGA-CA-G.-A--AACG.---ATCAAG.A--....CTG-----ATGTAC- 9930
H2G.CI.92.ABT96             GACACAGCA.........----------AG-A.................T--G--G-AACA--G-AG.....................................GTACTAT----GGGACT-GTTG--AACG.---C.-AA-.T--....CTG-----AT-TAG- 9388
H2U.FR.96.12034             GAGACAGCA.........----------A-CA................GA--GAGG-AACCA-G-AG......................................GA-CAA----GGA-CCA.T-G--AACG.---.TCAA--GA-....CTG-----ATGTAC- 9489
MAC.US.x.EMBL_3             ..................----------A-AA...................................................................GGGGATG-TATG----GGA-CC-GTCG--AAC-.---A-TTTC.TTG....ATG-----AT-TCA- 9070
SMM.US.x.H9                 ..GACAGCA.........----------A-AAA.................---GA-G-TAC...................................GGGGAGG-TCTG----GGA-CT-G.C---AACG.---A-TTAT.T--....CTG-----AT-CAA- 9443
SMM.US.x.PGM53              ..GACAGCA.........----------A-AAA.................---GA-G-TA...................................CGGGGAGG-ACTG----GGA-CT-G.C---AACG.---A-TTAT.TT-....CTG-----AT-CAA- 9897
SMM.US.x.SIVsmH635F_L3      ..GACAGCA.........----------A-AA..................-----------CAAG--GC--TCA.......................TGGGGAGG-ACTG----GGA-CT-G.C---AACG.---A-TTAT.T--....CTG-----AT-CAA- 9995
SMM.SL.92.SL92B             ..GACAGCA.........----------A-AA.................G----G-G-A--A--G..........................................A-GAGTCTG--CGGGG--AA..CG--AACG.---CT-AAC--TG....TAG-----AT-TAC- 9418
STM.US.x.STM                ..GACAGCA.........----------A-AA..................---G--G-AACA...................................GGGGAGG-ACT.----GGAACT-GT-G--AACG.----TTTAC.TT-....CTG-----AT-CAA- 9614
MNE.US.x.MNE027             ..GACAGCA.........----------ATAA..................---GA-G--...................................ATGGGG-GG-ACTG----GGA-CT-GTCG--AAC-.---A-TTTC.TTG....ATG-----AT-TCA- 9437
DRL.x.x.FAO                 GCTGACTGAGAA......----------A-A-GAGCATGCGCTCTGTGGC.A----------..................................AGGGTGAC-T-G-TTGGGG-A.......---TCTG.-------CT.G-A.......-----A------.CT 9330
MND_2.x.x.5440              A.................----------A---GTGCATGCGCACTGGACT.G----------.A-ACT-ACGTCGGG...........................................T----G------CAG.---AA...C.-TGGC.T------------.CT 9249
MND_2.CM.98.CM16            GCTGACTCTGGA......----------A--AGTGCATGCGCACTAGACT.G--------T-CA-ACT-A........................................................................................... 9518
MND_2.GA.x.M14              GCTGACTCTGGA......----------A--AGTGCATGCGCACTGGACT.G--------T---ACT-A---CGGG..........................................T----G-----TCAG.---CA..T..-TGAC.T------------.C- 9500
OLC.CI.97.97CI12            ..................................................................................................................................................................... 8678
GRV.ET.x.GRI_677            ..................------------CA...................----------CA-T--G.......................TGGATCGGAGGCG-TACA--G-C-GTAC.T--G------TTTC------AG-G-TG.........-----A-------..A 9387
VER.KE.x.TYO1_patent        ..................----------AGC-AA................----------A--C--G.........................................ACAT-G-C-GTCC.---G------TTTA-------.-A--.........-----A-------..A 9541
VER.KE.x.9063               ..................----------AA....................----------...................................GGGCG-G-CAT-G-C-GTAC.---G-----TTTTA-------.-G--.........-----A-------..A 9577
VER.KE.x.AGM155             ..................----------AGCAC.................----------A--C--.......................................GACAT-G-C-GTAC.---G------TTTA-------.-G--.........-----A-------..A 9558
VER.DE.x.AGM3               CATTAAGAACTGCTGAC..----------AGCAA.................-------------.........................................C-GG-CATG--CGG-AC---G------TTTA-------.-G--.........-----A-------..A 9077
TAN.UG.x.TAN1               TGATCAGCA.........----------AG...................................................................GGTG-TGC-A-G-C-GTACTT--G------TT.C-------.-G--.........-----A-------..A 9550
SAB.SN.x.SAB1C              TGACTGAGA.........----------AG....................................................................GGT-GAGACTG--CGG-ACT--G------TTG.---CGT-CT.G-.....CA-----------..- 9730
MND_1.GA.x.MNDGB1           TTGGCAACCTGCTT.....A-C-ACC-GGA---GCGCTTGCGCGCTAGGA..A-----------AA...............................................CAGG-A--GG-A---TCG.---CA..TG.-TG....CT-----------..- 9023
DEB.CM.99.CM40              ..................----------A.........................................GTTGCCTAGGAAATATCATGGCAACCAGGGGA-GGCCT-G-T-GTAC.---G------CA..G--CGCTG-A---....A----A----C-..A 9210
DEB.CM.99.CM5               ..................--T-------A.........................................GTAGTCAAGGAGACTCCATGGTTACAGGGGAG.GGACT-G-C-GTAC.---G-----TTTG.----GCTGT-A--....-----------T-..A 9141
TAL.CM.01.8023              GACTC.............-----------........................................CGCTGCTAAGGCAACGGGGGACGGACTGGGGCG-GGA-C-G-A-GAGCT.--G------TA..-----............................ 8845
TAL.CM.00.266               GACTC.............----------A........................................CGTTGCTAAGGCAACGGGGGACGGACTGGGGCG-GGA-C-G-A-GAGTT.--G------TTG.-------A-.--A....-----------------..A 9315
SUN.GA.98.L14               ..................-----------GAC..........................................................................AGGG..A--GG-A---TT-.---A-..TT.A--....-----------------..- 9707
WRC.CI.97.97CI14            .......TTGCAGTAGGAA-T-G----G--G-TAAGCTACA..........................TCCTGAGACTGTGTACCAACGGACTTA.................................................................... 9484
WRC.CI.98.98CI04            ..................................................................................................................................................................... 9461
WRC.GM.05.Pbt_05GM_X02      .......TCAGA...GGAA-CTGG---G--G-TAGGCTAGTA.........................TCCGCTAGCTGTGTACCAACGGCCTTAGCAA--AT-CTAATGAACT---ACTCT-GGAG-AGACTT-TGGTT-TGCCCT---A---T--ATG 8479
LST.CD.88.524               ..................ACA-GC--G-TGAACAGC..............A----------.........................................--A----CGGG--A-G.T-T-.........---A-..TT.G--....-----------------..- 8588
LST.CD.88.447               ..................A-------G-TGATGAGC..............A----------.........................................--A----CGGG--A-G.T-T-.........---A-..TT.G--....-----------------..- 8582
LST.KE.x.lho7               ..................----------AGGACGGGCGG...........----GG--TG.......................................................................---A-..TT.G--....-----------------..- 9663
SYK.KE.x.SYK173             ....................................................--AG-C-GGGC---G-AA-GCTCA.........................................................----.....A--.........-----A----C-..A 9356
SYK.KE.x.KE51               CCGATGCTCTGTAGCTATCC................................................................................................................................................. 9068
COL.CM.x.CGU1               GAGACAGCA.........-----------T-C..................-----------A-TT--GAG-..........................-A-CT--GC---GGGAA--GC---..........----GCTA-A.--GCTT-T---A---TCAG.A 8709
```

```
                                         TAR element start
                              3' LTR U3 end 3' LTR repeat start
H1B.FR.83.HXB2            TGC....TTTTTGC.CTGT..ACTGGG..TCTCTC...TGGTTAGA.CCAGAT.CTGAGCCTGGGAG.CTCTCTGGC.....TAA...........................CTAG..................................................GG 9592
H1A1.UG.85.U455           ---...----C--.----.------.------...-T------.------.C-----------.---------......-G...........................-G--..................................................-- 9051
H1B.US.90.WEAU160         ---...------.----.------.------...-------.------.------------.---------......-G...........................----..................................................-- 9592
H1C.ET.86.ETH2220         ---...----C--.T---..--C---.------...-A-G----.------.------------.---------......-T...........................--G-.................................................-- 8992
H1D.CD.84.84ZR085         ............................................................................................................................................................. 8975
H1F1.BE.93.VI850          C--...----C--.----.------.------...-A-----.------.T---------G-.---------......-G...........................-G-..................................................-- 8890
H1G.SE.93.SE6165          C--...--C-C--.----.------.------...-T-----.------.T-----------.---------......-G...........................-AG-.................................................-- 9036
H1H.CF.90.056             ---...----C--.T---.------.-----T..-------.------.------------.---------......-G...........................---A.................................................-- 8910
H1J.SE.93.SE7887          ---...----C--.T---.------.------...-T-----.------.------------.---------......---..........................----.................................................-- 8891
H1K.CM.96.MP535           ............................................................................................................................................................. 8604
H101_AE.TH.90.CM240       C--...----C--.T---.------.------...-T---------G-.-----C-----.---------......-G...........................-A--.................................................-- 9158
H102_AG.NG.x.IBNG         ---...--C-C--.----.------.------...-T-C----.------.------------.---------......-G...........................-GGA.................................................-- 9119
H1N.CM.95.YBF30           C--...--C-C--.T---.------.------...-T-C-G--.------.TA---T-------.-ATAT---G................................................................................ 9182
H1O.BE.87.ANT70           C--...---C---.T---..--C---.----GG...-TAGAG--.----G-.-------C-----.---C-----....CTC...........................---C..............................................T- 9713
H1O.CM.91.MVP5180         ---...---CC--.T---..--C---.---TAG...-TAGAG--.----G-.-------C-----.---C-----....CTC...........................---C..............................................T- 9709
CPZ.CD.90.ANT             ---...---GC--.T---..TAA---.C-----G...CC.-AG--G------.-A-----------.---------......---TTTTT.....................GGCAG..............................................-A 9055
CPZ.CM.01.SIVcpzCAM13     C--...---C---.T---.------TC--..-T...CAC.-G--.------.-------C-----.-------C-......-G...........................-GA..................................................-- 9240
CPZ.CM.05.SIVcpzEK505     ---...-GC-C--.T---.------TC-----T....-----------.TA-----------.---------......-G...........................-A-..................................................-- 9088
CPZ.CM.05.SIVcpzMB66      C--...---GC--.T---.------TC-----T....-----------T------------.---------......-G...........................--A..................................................-- 9236
CPZ.CM.98.CAM3            C--...-C-C---.T---.------.------..T--C.----.------.------------.--------T....C-G...........................G---...............................................A- 9047
CPZ.GA.88.GAB2            ---...-C-C---.T---..T-----.-----T...A--CAGA.------.------------.---------.....-G-...........................---................................................- 9020
CPZ.CM.98.CAM5            ............................................................................................................................................................. 9261
CPZ.CM.05.SIVcpzMT145     C--...-C-C--.T---..CAA---TC-----T.....-----.------.------------.---------......---CAAG..........................................................................-- 9164
CPZ.TZ.00.TAN1            ---...---GC--T-----..AAA--..----TG.CCTAA-CTG.------.------------.---------......-G...........................-G-CTGGC.......................................TA-A 9284
CPZ.CM.05.SIVcpzLB7       C--...-C-GC--.T---.------TC-----T....-----------T------------.---------......-G...........................--A..................................................-- 9125
CPZ.US.85.CPZUS           C--...-C-C---.T---.------.------...TC-C-AGA.------.AA----C-----.---------......-G...........................-GTA................................................- 9656
CPZ.GA.88.GAB1            ---...---C---.----.------.-----T..TCAC.-G--.------.T---------------.-----AC-.....AGT...........................G-A..................................................-- 9684
MON.CM.99.L1              C--...CC.........---.AG--ATCCT--GGAGTC....................................................................................................................... 9448
MON.NG.x.NG1              C--...-C.........---.AGM-ATCCGC-GGAGTC.........................--.-----GTCT..............................CCAC--CGGCCT.............................GGGAGCCCTTGGA- 8123
GSN.CM.99.CN166           C--...-C.........---.AG-GACTCT--GGAGTC....................................................................................................................... 9414
GSN.CM.99.CN71            C--...-C.........---.AG-GACTCT--GGAGTC....................................................................................................................... 9410
MUS_1.CM.01.1085          C--...-C.........---.AG--ATCCC--GGAGTC....................................................................................................................... 9420
MUS_1.CM.01.CM1239        C--...-C.........---.AG--ATCCC--GGAGTCAGTTC---.CT-C-CA.GC-GC.-------C-CTCGGAA-.......--G.GTAA.AGCGGCCTGAGAGCCCTTGCCTAGGAGAACTG.............................GTTCTCGGCTTGCT 9479
MUS_2.CM.01.CM1246         C--...-C.........---.AG--ATCCAA-GGAGTC....................................................................................................................... 9484
MUS_2.CM.01.CM2500        ---...-C..........G--ATCCAGACGGAGTC.......................................................................................................................... 9430
RCM.NG.x.NG411            ---...A---C--T----.A--CA-........------CCTG--AGGCTA------.----------TTGT-------T....A-GTCTCTAGGAGACT.AGCTTGAGCCTGGGTGTTCGCT.....................A-T 9311
RCM.GA.x.GAB1             ---...-C-GC--T.---AA-AC---.------CCTG--AGGCTA--G---.----------T-.T-------T....A-GTCTCTAGGAACTCCAGCTTGAGCCTGGGTGTTCGCT...................GGT- 9221
DEN.CD.x.CD1              ---...........T----.A--A-TCGG-T.........................-C---.-----CCAGCGGCCTGG.........................................................................-A 9622
                                          TAR element start
                              3' LTR U3 end 3' LTR repeat start

MAC.US.x.239              ---...A---CGCTC.---.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................................................GCTA-10064
H2A.GW.x.ALI              C--...TGC--GCAT.---.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................................................GCTA-10139
H2A.DE.x.BEN              C--...T-C--GCAT.---.ATTCA-TC.G----GC.GGA-AGGCTGG-----.-GAGC----A---GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................................................GCT--10144
H2A.SN.x.ST               C--...TAC-CGCAT.---.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................................................GCTA- 9583
H2B.GH.86.D205            C--...T-C-CGCAT.---.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................................................GCTA-10063
H2B.CI.x.EHO              C--...T--G-GCAT.---.ATTCA-TC.G----GC.GGA-AGGCTGG-----.-GAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................................................GCTA-10034
H2G.CI.92.ABT96           ---...A---CGCTC.---.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAGC--C-----GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................................................GCTA- 9492
H2U.FR.96.12034           C--...TC---GCAT.---.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................................................GCTA- 9593
MAC.US.x.EMBL_3           ---...A---CGCTC.---.ATTCA-TC.G----GG..CA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................................................GCTA- 9173
SMM.US.x.H9               ---...A---CGCTC.---.ATTCA-TC.A----GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................................................GCT-- 9547
SMM.US.x.PGM53            ---...A---CGCTC.---.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................................................GCTA-10001
SMM.US.x.SIVsmH635F_L3    ---...A---CGCTC.---.AGTCA-TC.G----GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................................................GCTA-10099
SMM.SL.92.SL92B           C--..TT-GCCGCTC.---.ATTCA-TC.G----AC.GGA-AGGCTG------.AGAGC--C-A---GT------CAGCACT-........AGTAGGTGAGCCTGGGTGTTCT--.....................................................ACTAA 9523
STM.US.x.STM              C--...T-A-AGCTC.---.ATTCA-TC.A----GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................................................GCTA- 9718
MNE.US.x.MNE027           ---...A---CGCTC.---.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................................................GCTA- 9541
DRL.x.x.FAO               ---...T-AC-GCATG-...A-AC--TC.----C-..TG--AGGCTA------.----------T-.T-------TG....AGTCTGCTGAAGGAGGCTACCTGCTCGAGCC-CG-GTGGTTCGCAGGGCCTTGGCCTTGGGCTCTGGGTAGCTCGGTCAGGAGGCTTC 9484
MND_2.x.x.5440            ---...T-GCCGC-.----.A-G--A-TC----C-T..GA-AGGCTG------.TGAGC.------..TTGT---CT.GG.TGAGTTCTTG.AAGGAG..TGCCTGCTTGTAGC-CT-GGCGGT.............................TCGCAGGCCCCT-- 9369
MND_2.CM.98.CM16          ............................................................................................................................................................. 9518
MND_2.GA.x.M14            C--...T-GCCGCT.----.ATT.-A-TC----C-T......................................................................................................................... 9529
OLC.CI.97.97CI12          ............................................................................................................................................................. 8678
GRV.ET.x.GRI_677          ---....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA.----C-..A---------T-.T--G----T.......-GCCTA...ACCCGGT.TGGCCA.CCGGGGG--A..............................................-- 9483
VER.KE.x.TYO1_patent      ---....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA.----C-.----------T-.T--G----T.......-GCCTA...ACCTGGT.TGGCCA.CCAGGGG--A..............................................-- 9637
VER.KE.x.9063             ---....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA.----C-.----------T-.T--G----T.......-GCCTA...ACCTGGT.TGGCCA.CCAGGGG--A..............................................-- 9673
VER.KE.x.AGM155           ---....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA.----C-.----------T-.T--G----T.......-GCCTA...ACCTGGT.TGGCCA.CCAGGGG--A..............................................-- 9654
VER.DE.x.AGM3             ---....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA.----C-.----------T-.T--G----T.......-GCCTA...ACCTGGT.TGGCCA.CCAGGGG--A..............................................-- 9173
TAN.UG.x.TAN1             ---....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA.----C-.-------------.T--G----T....G-GCCTA...ACCTGGTCTGGCCATCCAGGGG--A............................................GA- 9649
SAB.SN.x.SAB1C            ---.......---.GCGC----AA-ACA-.......---GC.T---AGGCTG------.----------T-.T-------T....A-GTCTG..AACCAGCT.TGAG...CCTGGGTGTTCGCT....................................GGG-A 9833
MND_1.GA.x.MNDGB1         ---...A---CGCT..----.T-G--AGTCTCTA-TACA-AGGCT-A...............---..TTGTATCTCTGA.GC.................................................................................. 9080
DEB.CM.99.CM40            ---...........T----..-GCTATCCA............................................................................................................................... 9227
DEB.CM.99.CM5             ---...........T----..-GCTATCCA............................................................................................................................... 9158
TAL.CM.01.8023            ............................................................................................................................................................. 8845
TAL.CM.00.266             ---....-C.........---.AG--ATCCT-T..G-TC...................................................................................................................... 9338
SUN.GA.98.L14             ---....-C--C--C.---..T-G--AG.------..GC-CAGAG-..AT--GCGAC--C----A---.G----ACCT....AG-GCTGGCTTGTCGCTCCTGAGTGAGCACTTGCTAAG...........................CACCTCAC 9817
WRC.CI.97.97CI14          ............................................................................................................................................................. 9484
WRC.CI.98.98CI04          ............................................................................................................................................................. 9461
WRC.GM.05.Pbt_05GM_X02    CT-TGTAGC-A-CTGG-TGCAGTCTTCTGC--GAGGCTA---GCCTGAGT-G-GG-CCCA--T--CCTAG----CTTGGTGCCTGAGTGGTGGTCCCACTTGGCCAAGAGACT-AG-TAGGAG.............................CTGACCACGCTT-C 8616
LST.CD.88.524             ---....-C-................................................................................................................................................... 8594
LST.CD.88.447             ---....-C-................................................................................................................................................... 8588
LST.KE.x.lho7             ---....-C-CC--T.---TCGGGA-..TC--TG-TCTG--GA...C-AGC-G.TC-CT------T-AGCACACT-GT...CAGCTAAGC.AGTCAGCCCTGGGAGGCTACA.....CTTAGC.............................ACCCACCTGGGCCT 9780
SYK.KE.x.SYK173           ---....-C.........---AGCTATCCGCGGAG-TGCAACTA.........................G------CCAGCGGCCT...GGG.AGCCCTGAGAGACTGGTAGAGCGGCCC.GGGA.............................GCCCTTGCCAGTAA 9452
SYK.KE.x.KE51             ............................................................................................................................................................. 9068
COL.CM.x.CGU1             GC-.....-CGAGAGCT-CGAGG-..................................................................................................................................... 8728
```

```
                           TAR element end┤                                                                                                     Poly-A signal
H1B.FR.83.HXB2             AACCC.....ACTGCTTAA......................................................................................GCCTC...................................AATAAAG...CTTGCC.......TT 9626
H1A1.UG.85.U455            -----.....---------......................................................................................-----...................................-------...------.......-- 9085
H1B.US.90.WEAU160          -----.....---------......................................................................................-----...................................-------...------.......-- 9626
H1C.ET.86.ETH2220          -----.....---------......................................................................................-----...................................-------...------.......-- 9026
H1D.CD.84.84ZR085          .......................................................................................................................................................................... 8975
H1F1.BE.93.VI850           -----.....--------.......................................................................................-----.............................................................. 8903
H1G.SE.93.SE6165           -----.....---------A.....................................................................................-----...................................-------...------.......AA 9071
H1H.CF.90.056              -----.....---------......................................................................................-----...................................-------...------.......-- 8944
H1J.SE.93.SE7887           -----.....---------......................................................................................-----...................................-------...------.......-- 8925
H1K.CM.96.MP535            .......................................................................................................................................................................... 8604
H101_AE.TH.90.CM240        -----.....---------A.....................................................................................-----...................................-------...------.......-- 9193
H102_AG.NG.x.IBNG          -----.....---------......................................................................................-----...................................-------...------.......-- 9153
H1N.CM.95.YBF30            .......................................................................................................................................................................... 9182
H1O.BE.87.ANT70            -----.....G--CG----......................................................................................CG---...................................-------...------.......-- 9747
H1O.CM.91.MVP5180          -----.....G--------......................................................................................CG---...................................-------...------.......-- 9743
CPZ.CD.90.ANT              G----.....----...........................................................................................-T-............................................................... 9068
CPZ.CM.01.SIVcpzCAM13      -----.....----.........................................................................................AA-----...................................-------...------.......-- 9274
CPZ.CM.05.SIVcpzEK505      -----.....---------......................................................................................-----...................................-------...------.......-- 9122
CPZ.CM.05.SIVcpzMB66       -----.....---------......................................................................................-----...................................-------...------.......-- 9270
CPZ.CM.98.CAM3             -----.....---------......................................................................................-----...................................-------...------.......-- 9081
CPZ.GA.88.GAB2             -----.....---------......................................................................................-----...................................-------...------.......-- 9054
CPZ.CM.98.CAM5             .......................................................................................................................................................................... 9261
CPZ.CM.05.SIVcpzMT145      -----.....---------......................................................................................-----...................................-------...------.......-- 9198
CPZ.TZ.00.TAN1             G----.....G--------......................................................................................CG-.TC..................................-------...-C----........- 9316
CPZ.CM.05.SIVcpzLB7        -----.....---------A.....................................................................................-----...................................-------...------.......-- 9160
CPZ.US.85.CPZUS            -----.....---------......................................................................................-----...................................-------...------.......-- 9690
CPZ.GA.88.GAB1             -----.....---------......................................................................................-----...................................-------...------.......-- 9718
MON.CM.99.L1               .......................................................................................................................................................................... 9448
MON.NG.x.NG1               CTAGGTAAAGCGGC--GGGAGCCCTTGCCTAGAGAGGTGCCGGCTCTCGGCTCGCTTGCTGAATAGGCTTTACC...............................AG-C-...................................-------...G-GCTTGAAAGCAAA 8223
GSN.CM.99.CN166            .......................................................................................................................................................................... 9414
GSN.CM.99.CN71             .......................................................................................................................................................................... 9410
MUS_1.CM.01.1085           .......................................................................................................................................................................... 9420
MUS_1.CM.01.CM1239         TTG-TTCGAAG--TA--G-GCCCACAATA............................................................................................................................................. 9508
MUS_2.CM.01.CM1246         .......................................................................................................................................................................... 9484
MUS_2.CM.01.CM2500         .......................................................................................................................................................................... 9430
RCM.NG.x.NG411             TCT-TGAAGGG--TGC-GGGAGCGCCTTGGCTGGGTCTTGGTAGACCTTCAGCTGGGTACTTGGCCAGTCCCAGTGGGAGTAGGCTCCGCTTGCT..TTCT.GGTTAA-TGCTC..........................------C....-.--ATGCTCTTAG 9442
RCM.GA.x.GAB1              TCT-TGAACAGGCTTGCTGGGGTGCCTCTCGCTCTTCGGGTAGACCGCCAGTTGAGGCTCGGCCGGCCTCAA..CGGGAGAGATCACCGCTTGC...TTATAGCCTTGAAGCTC..........................-------C......ATGCCAG..-- 9347
DEN.CD.x.CD1               GC--TGAG..-GGC----GCAGCGGCCTGGGAGCCCTCTAAGGTATTGCAACAAGTACC............................................................................................................... 9679
                                                                                TAR element end┤                                                     Poly-A signal
MAC.US.x.239               -CT-TC.ACC-GCA---GGCCGGTGCTGGGCAGAG.......................................TGACTCCACGCT.TGCTTGC...TTAA.AG.C----TTC................................-------...--.---AT..T..--10148
H2A.GW.x.ALI               -CT-TC.ACC-G-----GGCCGGCACTGGGCAGA........................................CGGCTCCACGCT.TGCTTGC...TTAA..AGA----TT.................................-------...--.---AA.....--10220
H2A.DE.x.BEN               -CT-TC.ACC-G-A---GGCCGGTACTGGGCAGA........................................CGGCTCCACGCT.TGCTTGC...TTAA..AGA----TTC................................-------...--.---AG.....--10226
H2A.SN.x.ST                -CT-TC.ACC-G-----GGCCGGCACTGGGCAGA........................................CGGCTCCACGCT.TGCTTGC...TTAAA.AGA----TT.................................-------...--.---AG.....-- 9665
H2B.GH.86.D205             -CT-TC.ACC-GCA--AGGCCAGTGTTGGGCAGA........................................CGGCTCCACGCT.TGCTTGC...TTAA..AAC----TT.................................-------...--.---AA.....--10143
H2B.CI.x.EHO               -CT-TC.ACC-GCA---GGCCGGTGCTGGGCAGA........................................CGGCTCCACGCT.TGCTTGC...TTAA..AAC----TT.................................------A...--.---TA.....--10115
H2G.CI.92.ABT96            -CT-TC.ACC-GCA---GGCCGGTGCTGGGCAGAG.......................................CGGCTCCACGCT.TGCTTGC...TTACAACGA----TT.................................-------...--.---AAA....-- 9577
H2U.FR.96.12034            -CT-TC.ACC-AAAG........................................................................................................................................................... 9607
MAC.US.x.EMBL_3            -CT-TC.ACC-GCA---GGCCAGTGCTGG.CAGAG.......................................TGGCTCCACGCT.TGCTTGC...TTAA.AG.A----TTC................................-------...--.---CA..T.T-- 9257
SMM.US.x.H9                -CT-TC.ACC-GCA---GGCCGGTGCTGGGCAGAG.......................................TRGCTCCACGCT.TGCTTGC...TTAA.AG.A----TTC................................-------G..--.---AG.....-- 9631
SMM.US.x.PGM53             -CT-TC.ACC-GCA---GGCCGGTGCTGGGCAGAG.......................................TGGCTCCACGCT.TGCTTGC...TTAA.AG.A----TTC................................-------...--.---GC.....--10084
SMM.US.x.SIVsmH635F_L3     -CT-TC.ACC-GCA---GGCCGGTGCTGGGCAGAG.......................................TGGCTCCACGCT.TGCTTGC...TTAA.AG.A----TTC................................-------...--.---AT.....--10182
SMM.SL.92.SL92B            -CT-TCAACC-GCAACAGGCCATTGCTGGGTAGAC........................................GGCTCCACGCT.TGCTTGC...TTGA......---TC.................................-------..G--.--AAC.....-- 9602
STM.US.x.STM               -CT-TC.ACC-GCA---GGCCAGTGCTGGGTAGAG.......................................TGGCTCCACGCT.TGCTTGC...TTAT.AT.AT---TTC................................-------...--.---AA..T..-- 9802
MNE.US.x.MNE027            -CT-TC.ACC-GCA---GGCCGGTGCTGTGCAGAG.......................................TGGCTCCACGCT.TGCTTGC...TTAA.AG.A----TTC................................-------...--.---AT.....-- 9624
DRL.x.x.FAO                TGGAGGTAGAT-ACTAGGGGCTGCTCTGGTCTTGGTCTGGTAGACCTTCGCTACTTCTAGTGGCCATAGAAGTAGGGGAGTAGGCACCGCTTGC...TTAT..CCTAA-TGCTC..........................-------C..TA..--CA.....-- 9609
MND_2.x.x.5440             CTTGTGGCTCTGG-TAGCTCG.TCAGGTG..TTCTGGAAAGGCTTACGGGAGGACACTCTTGCTTGGTCTTG..ATAGACCTCTAGCAATCCCTTAGGCCTGGAGAT.-GTGGGATTGACTACCGCTTGCTTGCTATTATTGCCC--------....C-TA-CTGATTAGA 9529
MND_2.CM.98.CM16           .......................................................................................................................................................................... 9518
MND_2.GA.x.M14             .......................................................................................................................................................................... 9529
OLC.CI.97.97CI12           .......................................................................................................................................................................... 8678
GRV.ET.x.GRI_677           -CT--.....TTG----C-TAT...................................................................................AG---...................................------C...C---TCGC....-- 9522
VER.KE.x.TYO1_patent       -CT--.....TTG-----GAA....................................................................................AG--....................................-------........-----TG..CA-- 9675
VER.KE.x.9063              -CT--.....TTG----GGAA....................................................................................AG--....................................------AC..T-----TG..CA-- 9713
VER.KE.x.AGM155            -CT--.....TTG-----GAA....................................................................................AG--....................................------........------TG..CA-- 9692
VER.DE.x.AGM3              -CT--.....TTG-----GAA....................................................................................AG-.T...................................------T..---CG-TG..CA-- 9212
TAN.UG.x.TAN1              CCT--......T------GAA.......................................................................GGCTATT.........G--...................................-------...G-----TG...C-- 9691
SAB.SN.x.SAB1C             -GG-AAGCCT-AGAGCACTCTGGTGAGGTCTTGGTAGACCCTCGCCTGGCGACTGGCCATTGCCAGTAGCAG..AGACTCC......GCTTGC...TTGCTT.GATTG-CTC..........................-------T.------........-- 9949
MND_1.GA.x.MNDGB1          ................................AGATCCCCTTAGAGCAAGGACCAGAGTCCTGAGTGACTGG..GTCTGAGCACCTCACTCGGGGCTGATCACCTCGAGGTAGTGGAACTCCTTGCTTGCTTGCTATTGTCTTC.-------T.AAC-TAGAA..TTAGA 9212
DEB.CM.99.CM40             .......................................................................................................................................................................... 9227
DEB.CM.99.CM5              .......................................................................................................................................................................... 9158
TAL.CM.01.8023             .......................................................................................................................................................................... 8845
TAL.CM.00.266              .......................................................................................................................................................................... 9338
SUN.GA.98.L14              CCGAGCCTGAT-A----GGGGAGCTAGAGG............................................CTCCTTGCTTTGCAGTACAAG.........C-T--...................................-------...-C----.......-- 9888
WRC.CI.97.97CI14           .......................................................................................................................................................................... 9484
WRC.CI.98.98CI04           .......................................................................................................................................................................... 9461
WRC.GM.05.Pbt_05GM_X02     TTGTTAGGGTGT-TGCAG-T.....................................................................................................................--------...AG-CTTTTGCTTAGC 8658
LST.CD.88.524              .......................................................................................................................................................................... 8594
LST.CD.88.447              .......................................................................................................................................................................... 8588
LST.KE.x.lho7              G-T-AGCCTAGGGAGC---G.................................GGCTCCTAGCTTGCTACTT..TAAA...GCCTTC..............................................................--------.....AA-TTGCTCATTG- 9849
SYK.KE.x.SYK173            -GG-TAGTGAGG-C-GGGTTTGGCCGCCCGGTCTTCGGCAGCTCTCCGTTGCTT............................GCTTTGCTTACGC.TGATCAGCC-.A...................................-------T.GTGGTAAACGCAAG-A 9554
SYK.KE.x.KE51              .......................................................................................................................................................................... 9068
COL.CM.x.CGU1              .......................................................................................................................................................................... 8728
```

```
                              Extensive secondary structure in this region?
                    3' LTR R repeat end 3' LTR U5 start                                                                   3' LTR U5 end
H1B.FR.83.HXB2            GAG..TG.CTTCAA.GTAGTG...TGTGCCCGTC.TG.TT..GTG.TGACTCT...GGT..AA.CTAGAGA..............TCCCT..CAGA...................CCCTTTTAGTCAGTGTGGAAAATCTCTAGCA..    9719
H1A1.UG.85.U455           ---..GTG---G--.------..----------.--.--..A-T.-------..---.--.-------...............-----.----....................--AC-A---A-T.GTGTA-------------..    9178
H1B.US.90.WEAU160         ---.--.------.------..----------.--.--..---.-------..---.--.-------...............-----.----....................--A----------------------------..    9719
H1C.ET.86.ETH2220         ---.--..............................................................................................................................................    9031
H1D.CD.84.84ZR085         ....................................................................................................................................................    8975
H1F1.BE.93.VI850          ....................................................................................................................................................    8903
H1G.SE.93.SE6165          -CC.................................................................................................................................................    9074
H1H.CF.90.056             ---.--.---T.........................................................................................................................................    8953
H1J.SE.93.SE7887          ---..G-CGCATGC.AAGCC-...............................................................................................................................    8943
H1K.CM.96.MP535           ....................................................................................................................................................    8604
H101_AE.TH.90.CM240       ---.--.---A-........................................................................................................................................    9203
H102_AG.NG.x.IBNG         ---.--.-----.-----..----------.--A--.---.-------..---.--.-..........................................................................................    9201
H1N.CM.95.YBF30           ....................................................................................................................................................    9182
H1O.BE.87.ANT70           ---.---.AG..........................................................................................................................................    9754
H1O.CM.91.MVP5180         ---.---.AGAAGC.AGT---..--CT-A..--TG..--CAACC.CTGG-G-...CTA..G-.......--TC...........................................................................    9793
CPZ.CD.90.ANT             ....................................................................................................................................................    9068
CPZ.CM.01.SIVcpzCAM13     ---..--CT-CA........................................................................................................................................    9284
CPZ.CM.05.SIVcpzEK505     ---..--CT-AA........................................................................................................................................    9132
CPZ.CM.05.SIVcpzMB66      ---..--A---A-.......................................................................................................................................    9281
CPZ.CM.98.CAM3            ---..--A---G..A-C----..------TA--CA..--C..A-.-A----..---..--.-------TCC...............................CTCAGAA.-TA--A--A..GTA-GT-----------.....    9170
CPZ.GA.88.GAB2            ---..--CTC-A..A-----..-----T-A--TG..--C..A-.-A----..---..GT.-------TCC...............................CTCAGAC.TAAA.CATAA.G-A----------------..    9146
CPZ.CM.98.CAM5            ....................................................................................................................................................    9261
CPZ.CM.05.SIVcpzMT145     ---..--CTGCA........................................................................................................................................    9208
CPZ.TZ.00.TAN1            ---..A-.TG-T.-......................................................................................................................................    9326
CPZ.CM.05.SIVcpzLB7       ---..--CTGCA........................................................................................................................................    9170
CPZ.US.85.CPZUS           ---..--CT--A..A------..------TA--.C..--C..A-.--------..---..--.-------TCC...............................CTCAGAA.-T---AA-A..GTA-GT--------------..    9781
CPZ.GA.88.GAB1            ---.TGTAT--G..A-C----..------TA--TA..--C..A-.AC----..---..--.-------TCC...............................CTCAGAT.TAAA--ATA..GTCAA---------------..    9811
MON.CM.99.L1              ....................................................................................................................................................    9448
MON.NG.x.NG1              A-AAAAAAAAAA--AAA-AAAAAAAAAAAAAAAAAAA...............................................................................................................    8260
GSN.CM.99.CN166           ....................................................................................................................................................    9414
GSN.CM.99.CN71            ....................................................................................................................................................    9410
MUS_1.CM.01.1085          ....................................................................................................................................................    9420
MUS_1.CM.01.CM1239        ....................................................................................................................................................    9508
MUS_2.CM.01.CM1246        ....................................................................................................................................................    9484
MUS_2.CM.01.CM2500        ....................................................................................................................................................    9430
RCM.NG.x.NG411            TT-AACCAGAATC-.AGT---TG..CCCATTT-A.....TCTCTCA-TGAAACGACTCTGGGGT-G-GA..............TC--TCAG.......ATATTTGTGGCAGAAGT.AC-.G-AGGCTAAGAA------C-A-C--G.    9557
RCM.GA.x.GAB1             AGTTTAC.TG-A-GC.A----TGTGCCTGTTT-A....CCTCTCAGCAGT-AACGACTCTGGGGTAG-GA..............TC--TCAG.......ATTCTTGTGGCAGAAGAG-C--G..GGCTAAGAA--.T--C---C--GT    9465
DEN.CD.x.CD1              ....................................................................................................................................................    9679
                  3' LTR R repeat end 3' LTR U5 start                                                                     3' LTR U5 end
MAC.US.x.239              AGA..A-TAAG-T-..GT---TGT-CCCAT-TCT.CCTAGCC-CC.GCCTGG-...CAAC.TCGG-..-CTCAATA....ATAAGAAGA-CCTGGTCTGTTAGGACCCTTTCTGCTT.TGGGAA-CCG-AGCA--------C------..   10279
H2A.GW.x.ALI              AGA..A-CAGGTT-.AAG---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CA-..TCGG-GTTC-CCTGA....GTAACAAGA-CCTGGTCTGTTAGGACCCTTTCTGCTT.TGGGAA-CGG--GCA--------C------..   10353
H2A.DE.x.BEN              AGA..A-CAAGTT-.AGT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CA-..TCGG-GTTC-TCTGA....GTAACAAGA-CCTGGTCTGTTAGGACCCTTCTCGCTT.TGGGAATCCA--GCA--------C------..   10359
H2A.SN.x.ST               AGA..A-CA...........................................................................................................................................    9672
H2B.GH.86.D205            AGA..A-CAAGT-..AGT---TGT-CCCAT-TCT.CCTAGTC-C..GCCTGG-...CAC..TCGG-..GCTCCAC.....TGA.TAGA-CCTGGTCTGTTAGGACCCTTCTTGCTT.TGGGAA-CCG-AGCG--------C------..   10269
H2B.CI.x.EHO              AGA..A-CAAGAC-.AGT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CA-..TCGG-..GCTCCAC......TA..GAAA-CCTGGTCTGTTAGGACCCCTTCTGCTT.TGGGAA-CCA--GCA--------C------..   10242
H2G.CI.92.ABT96           AGA..A-TAAG---GTGT---TGT............................................................................................................................    9599
H2U.FR.96.12034           ....................................................................................................................................................    9607
MAC.US.x.EMBL_3           AGA..A-TA...........................................................................................................................................    9264
SMM.US.x.H9               AGA..A-TA...........................................................................................................................................    9638
SMM.US.x.PGM53            AGA..A-TAAG---.GT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CA-C.TCGG-..-CTCGAAT..CATAAGAAGA-CCTGGTCTGTTAGGACCCTTTCTGC..................................   10184
SMM.US.x.SIVsmH635F_L3    AGA..A-TAAG---.GT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CA-C.TCGG-..-CTCGACA..CATAAAGAAGA-CCTGGTCTGTTAGGACCCTTTCTGCTT.TGGGAA-CCG--GCA--------C------..   10315
SMM.SL.92.SL92B           AGA..A-CAAGTC.......................................................................................................................................    9613
STM.US.x.STM              AGA..A-TAAG---.GT-G-TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CAAC.TCGG-..-CTCAA............GA-CCTGGTCTGTTAGGACCCTTTCTGC..................................    9892
MNE.US.x.MNE027           AGA..A-TAAG-C-..GT---TGC-CCCAT-TCT.CCTAGTC-CC.......................................................................................................    9664
DRL.x.x.FAO               T--ATG-ATGAGT-.AGT--ACTC-TGT-TTT-G........T-CT-AGTGAGACTCTGGTT-CTGGAGA..............TC--TCAGACT-GGTGAGAGCTATCCTGAGCTTTAG-..GAGAGTAAA-A--------C----..    9731
MND_2.x.x.5440            ATAGCAAGTAAG-G.TACTC-TCCAT-TTGTC-TCAATAGAA...ACTCTGG-TACC-GAG-........................-----.T----TCTGTGGCAGAGATTTGATAT-TAG-G--AGTAGAGAA--------C----..    9647
MND_2.CM.98.CM16          ....................................................................................................................................................    9518
MND_2.GA.x.M14            ....................................................................................................................................................    9529
OLC.CI.97.97CI12          ....................................................................................................................................................    8678
GRV.ET.x.GRI_677          AGT..C-CTA-ATTG.G---CAAG--CT-ATTG-TGCGCCGA-CCT.CTAGAGG.....TG-A-CTCTCTTACTGGGTTCTCCTG-A--CAGGT.................................G-G..-G--A---C----..    9623
VER.KE.x.TYO1_patent      AGA..G.CT-ATCTGAGTCAAGTG-CCT-ATTGACG.CC-CAC-CTCTTGAACG.....GG-ATCTTCCTTACTGGGTTCTCTCTCT.......................GACCCAGGCGAG--..........--A---C----..    9778
VER.KE.x.9063             AGA..G.CT-ATCTGAGTCAAGTGCCCT-ATTGACG.CC-CAC-C..A-GCAGG....GG--C-GTTCCTTACTGGGTTCTCTCTCTGACCCAG.................................GCG---G--A---C----..    9815
VER.KE.x.AGM155           AGA..G.CT-ATCTGAGTCAAGTGCCCT-ATTAATG.CC-CAC-C..TTGAACG....GG-GAAGTTCCTTACTGGGTTCTCTCTCAAACCCAG.................................GCGA--G--A---C----..    9794
VER.DE.x.AGM3             AG-.CA-AGACTT..--GAGAGA....CATT-C-TG-TGCATCCAGCACAAG..................TGAGGGAAGACCCTGATGGAAT--ACCATGGAGAAGTCTTGGTCTGGAAG.........-T--ATCCCATGT-----..    9325
TAN.UG.x.TAN1             A--..A-T--ATCTGAG.CAAGTGCC-CATT-CGCGCCC-TCC-CAG-............................GAACCCTCGTTACTGGG-...........TTTCTTATC-AGGC-G--..........-------C----..    9784
SAB.SN.x.SAB1C            AGTTTACAAG-A--GCA----TGTGCCCATTTATTCCTCAG........................GACAA-CCTGGTTACTAAGGATCCC.........TGA.........-AGAA------CT-----..   10036
MND_1.GA.x.MNDGB1          -CA.................................................................................................................................................    9215
DEB.CM.99.CM40            ....................................................................................................................................................    9227
DEB.CM.99.CM5             ....................................................................................................................................................    9158
TAL.CM.01.8023            ....................................................................................................................................................    8845
TAL.CM.00.266             ....................................................................................................................................................    9338
SUN.GA.98.L14             AG...-TTTACTGCAAGCAA-TGT-CCCGTGT--TCTC--TCTCC..............TTCTAA-CCCTGTTCGACTGCTCTCAG-TAGGGAGC-A...........TTACTGG--GCC-AGTGATCC-G--CTG-CGG------..   10006
WRC.CI.97.97CI14          ....................................................................................................................................................    9484
WRC.CI.98.98CI04          ....................................................................................................................................................    9461
WRC.GM.05.Pbt_05GM_X02    A--AAGCGTG-GGTTAGTT-TTGT---AAGAAGTAGAACCCT-G-AG-CT-.................................................................................................    8709
LST.CD.88.524             ....................................................................................................................................................    8594
LST.CD.88.447             ....................................................................................................................................................    8588
LST.KE.x.lho7             T--AAA-TACAAGCTAGT--CTCA-AGTAGTC--T-TC-CCTTC-CCC........TGGTTCAG-GATC.................T--CA-TAGAGAGATTGGAGCCTTGTGAT--GGGAC--GGCT-CCA-...............    9957
SYK.KE.x.SYK173           .....ACGTC-GCT.C-CA-CTGT-C---TG-GT.CACC-CCCGA.A-GTG.................................................................................................    9597
SYK.KE.x.KE51             ....................................................................................................................................................    9068
COL.CM.x.CGU1             ....................................................................................................................................................    8728
```

V

HIV-1/SIVcpz Proteins

Contents

V-1	Introduction		**301**
V-2	Annotated features		**302**
V-3	Sequences		**304**
V-4	Alignments		**310**
	V-4.1	Gag	310
	V-4.2	Pol	318
	V-4.3	Vif	332
	V-4.4	Vpr	336
	V-4.5	Tat	338
	V-4.6	Rev	340
	V-4.7	Vpu	342
	V-4.8	Env	346
	V-4.9	Nef	360

V-1 Introduction

The HIV-1/SIVcpz protein alignments contain the same sequences as the complete nucleotide alignment, but they are annotated in more detail.

V-2 Annotated features

Features of HIV-1 annotated in the alignment that follows.

Feature	Protein	Location	Page
Gag start, p17 start	Gag	1	310
membrane binding	Gag	1-30	310
nuclear localization	Gag	11-35	310
nuclear localization	Gag	106-116	310
phosphorylation site	Gag	112	310
p17 end	Gag	132	310
p24 start	Gag	133	310
CyPA binding	Gag	205-241	312
major homology region	Gag	285-304	312
p24 end	Gag	364	314
p2 start	Gag	365	314
p2 end	Gag	377	314
p7 start	Gag	378	314
Zn motif	Gag	392-404	314
Zn motif	Gag	413-425	314
p7 end	Gag	432	314
p1 start	Gag	433	314
p1 end	Gag	448	314
p6 start	Gag	449	314
Vpr binding	Gag	455-460	314
Vpr binding	Gag	489-494	316
p6 end, Gag end	Gag	501	316
Gag-Pol TF start	Pol	1	318
Gag-Pol TF end	Pol	56	318
protease start	Pol	57	318
protease end	Pol	155	320
p66, p51 RT start	Pol	156	320
M41L	Pol	196	320
D67N	Pol	222	320
K70R	Pol	225	320
D110 catalytic site	Pol	265	320
polymerase motif	Pol	337-342	322
T215Y	Pol	370	322
K219Q	Pol	374	322
p51 RT end	Pol	595	324
p15 RNase H start	Pol	596	324
p66 RT, p15 Rnase H end	Pol	715	326
p31 Integrase start	Pol	716	326
p31 Integrase end	Pol	1003	330
Pol end	Pol	1003	330
Vif start	Vif	1	332
Vif end	Vif	192	334
Vpr start	Vpr	1	336
oligomerization	Vpr	2-43	336
amphipathic α-helix	Vpr	16-34	336
H(S/N)RIG motifs	Vpr	73-84	336
frameshift in HXB2	Vpr	72	336
Vpr end in HXB2	Vpr	79	336

Feature	Protein	Location	Page
Vpr end	Vpr	97	336
Tat start	Tat	1	338
disulfide bonding	Tat	22-37	338
NLS	Tat	49-57	338
exon 1 end	Tat	72	338
exon 2 start	Tat	73	338
Tat end	Tat	101	338
Rev start	Rev	1	340
exon 1 end	Rev	25	340
exon 2 start	Rev	26	340
NLS	Rev	34-49	340
Leu-rich effector domain	Rev	75-83	340
Rev end	Rev	116	340
Vpu start	Vpu	1	342
transmembrane domain	Vpu	1-27	342
cytoplasmic domain	Vpu	28-82	342
α-helix	Vpu	31-49	342
phos	Vpu	53	342
phos	Vpu	57	342
α-helix	Vpu	59-72	342
Vpu end	Vpu	83	342
Env start	Env	1	346
signal peptide	Env	30	346
gp120 start	Env	31	346
glycosylation NVT	Env	88-90	346
V1	Env	131-156	346
glycosylation NDT	Env	136-138	346
glycosylation NSS	Env	141-143	348
glycosylation NCS	Env	156-158	348
V2	Env	156-196	348
glycosylation NIS	Env	160-162	348
glycosylation NDT	Env	186-188	348
glycosylation NTS	Env	197-199	348
glycosylation NKT	Env	230-232	348
glycosylation NGT	Env	234-236	348
glycosylation NVS	Env	241-243	348
glycosylation NGS	Env	262-264	348
glycosylation NFT	Env	276-278	350
glycosylation NTS	Env	289-291	350
glycosylation NCT	Env	295-297	350
V3	Env	296-331	350
glycosylation NNT	Env	301-303	350
V3 tip	Env	312-315	350
glycosylation NNT	Env	339-341	350
glycosylation NKT	Env	356-358	350
CD4	Env	368	350
CD4	Env	370	350
V4	Env	385-418	350
glycosylation NST	Env	386-388	350
glycosylation NST	Env	392-394	350
glycosylation NST	Env	397-399	350
glycosylation NNT	Env	406-408	352
CD4	Env	427	352

Feature	Protein	Location	Page
CD4	Env	429	352
CD4	Env	438	352
glycosylation NIT	Env	448-450	352
V5	Env	460-471	352
glycosylation NES	Env	463-465	352
fusion peptide	Env	510-527	352
gp120 end	Env	511	352
gp41 start	Env	512	352
immunodominant region	Env	588-607	354
glycosylation NAS	Env	611-613	354
glycosylation NKS	Env	616-618	354
glycosylation NHT	Env	624-626	354
glycosylation NYT	Env	637-639	354
glycosylation NGS	Env	750-752	356
glycosylation NAT	Env	816-818	356
gp41 end	Env	857	358
Env end	Env	857	358
Nef start	Nef	1	360
myristoylation	Nef	2-7	360
acidic cluster	Nef	62-65	360
poly-P helix	Nef	69-78	360
phos	Nef	77	360
phos	Nef	81	360
HXB2 premature Nef end	Nef	124	360
normal Nef end	Nef	207	362

V-3 Sequences

Sequences included in the HIV-1 protein alignments.

Name	Accession	Proteins	Author	Reference
B.FR.83.HXB2	K03455	All	van Beveren, C.P.	*(in) Weiss,RL, Teich,N, Varmus,H and Coffin,J(Eds); RNA TUMOR VIRUSES, SECOND EDITION,* **2**, Vol 2: 1102-1123; Cold Spring Harbor Laboratory, Cold Spring Harbor (1985)
A1.GE.99.99GEMZ011	DQ207944	All	Zarandia, M.	*ARHR* **22**(5):470-476 (2006)
A1.KE.00.KER2008	AF457052	All	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
A1.KE.00.KNH1144	AF457066	All	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
A1.KE.00.KSM4024	AF457077	All	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
A1.KE.00.MSA4069	AF457080	All	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
A1.KE.00.NKU3005	AF457089	All	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
A1.RU.00.RU00051	EF545108	All	Thomson, M.M.	*ARHR* **23**(12):1599-1604 (2007)
A1.RW.93.93RW_024	AY713406	All	Brown, B.K.	*J Virol* **79**(10):6089-6101 (2005)
A1.SE.95.SE8891	AF069673	All	Carr, J.K.	*AIDS* **13**(14):1819-1826 (1999)
A1.SE.95.UGSE8131	AF107771	All	Carr, J.K.	*AIDS* **13**(14):1819-1826 (1999)
A1.TZ.01.A173	AY253305	All	Arroyo, M.A.	*ARHR* **20**(8):895-901 (2004)
A1.UA.01.01UADN139	DQ823357	All	Saad, M.D.	*ARHR* **22**(8):709-714 (2006)
A1.UG.92.92UG037	AB253429	All	Sakamoto	Unpublished
A1.UG.99.99UGA07072	AF484478	All	Harris, M.E.	*ARHR* **18**(17):1281-1290 (2002)

Name	Accession	Proteins	Author	Reference
A2.CD.97.97CDKTB48	AF286238	All	Gao, F.	*ARHR* **17**(8):675-688 (2001)
A2.CY.94.94CY017_41	AF286237	All	Gao, F.	*ARHR* **17**(8):675-688 (2001)
A.SN.01.DDI579	AY521629	All	Meloni, S.T.	*J Virol* **78**(22):12438-12445 (2004)
A.SN.01.DDJ369	AY521631	All	Meloni, S.T.	*J Virol* **78**(22):12438-12445 (2004)
A.SN.96.DDJ360	AY521630	All	Meloni, S.T.	*J Virol* **78**(22):12438-12445 (2004)
A.CD.02.02CD_KTB035	AM000055	All	Vidal, N.	*ARHR* **22**(2):182-187 (2006)
A.CD.97.97CD_KCC2	AM000053	All	Vidal, N.	*ARHR* **22**(2):182-187 (2006)
A.CD.97.97CD_KTB13	AM000054	All	Vidal, N.	*ARHR* **22**(2):182-187 (2006)
B.AR.04.04AR151516	DQ383752	All	Pando, M.A.	*Retrovirology* **3** , 59 (2006)
B.AU.87.MBC925	AF042101	All	Oelrichs, R.B.	*J Biomed Sci* **7**(2):128-135 (2000)
B.BO.99.BOL0122	AY037270	All	Carr, J.K.	*AIDS* **15**(15):F41-F47 (2001)
B.BR.03.BREPM2012	EF637046	All	Sa-Filho, D.	*ARHR* **23**(9):1087-1094 (2007)
B.CA.97.CANB3FULL	AY779553	All	Cali, L.	*ARHR* **21**(8):728-733 (2005)
B.CN.05.05CNHB_hp3	DQ990880	All	Tan, J.-x.	*Chin Med J(Engl)* **120**(9); 831-3 (2007)
B.CO.01.PCM001	AY561236	All	Sanchez, G.I.	*Am J Trop Med Hyg* **74**(4); 674-7 (2006)
B.GB.83.CAM1	D10112	All	Saurya, S.	*ARHR* **19**(1):73-76 (2003)
B.GB.86.GB8	AJ271445	All	Farrar, G.H.	*J Med Virol* **34**(2):104-113 (1991)
B.GE.03.03GEMZ010	DQ207942	All	Zarandia, M.	*ARHR* **22**(5):470-476 (2006)
B.IT.05.SG1	DQ672623	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Calugi, G.	*J Virol* **80**(23):11892-11896 (2006)
B.JP.05.DR6538	AB287363	All	Sakamoto	Unpublished
B.KR.05.05CSR3	DQ837381	All	Cho	Unpublished
B.NL.00.671_00T36	AY423387	All	Geels, M.J.	*J Virol* **77**(23):12430-12440 (2003)
B.RU.04.04RU128005	AY682547	All	Galkin, A.N.	*ARHR* **22**(11):1192-1197 (2006)
B.TH.00.00TH_C3198	AY945710	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Watanaveeradej, V.	*ARHR* **22**(8):801-7 (2006)
B.UA.01.01UAKV167	DQ823362	All	Saad, M.D.	*ARHR* **22**(8):709-714 (2006)
B.US.04.ES10_53	EF363127	All	Blankson, J.N.	*J Virol* **81**(5):2508-2518 (2007)
B.US.99.PRB959_03	AY331296	All	Bernardin, F.	*J Virol* **79**(17):11523-11528 (2005)
C.AR.01.ARG4006	AY563170	All	Carrion, G.	*ARHR* **20**(9):1022-1025 (2004)
C.BR.04.04BR013	AY727522	All	Sanabani, S.	*ARHR* **22**(2):171-176 (2006)
C.BW.00.00BW07621	AF443088	All	Novitsky, V.	*J Virol* **76**(11):5435-5451 (2002)
C.CN.98.YNRL9840	AY967806	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Qiu, Z.	*ARHR* **21**(12):1051-1056 (2005)
C.ET.02.02ET_288	AY713417	All	Brown, B.K.	*J Virol* **79**(10):6089-6101 (2005)
C.GE.03.03GEMZ033	DQ207941	All	Zarandia, M.	*ARHR* **22**(5):470-476 (2006)
C.IL.99.99ET7	AY255824	All	Harris, M.E.	*ARHR* **19**(12):1125-33 (2003)
C.IN.99.01IN565_10	AY049708	All	Khurana	Unpublished
C.KE.00.KER2010	AF457054	All	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
C.MM.99.mIDU101_3	AB097871	All	Takebe, Y.	*AIDS* **17**(14):2077-87 (2003)
C.MW.93.93MW_965	AY713413	All	Brown, B.K.	*J Virol* **79**(10):6089-6101 (2005)
C.SN.90.90SE_364	AY713416	All	Brown, B.K.	*J Virol* **79**(10):6089-6101 (2005)
C.SO.89.89SM_145	AY713415	All	Brown, B.K.	*J Virol* **79**(10):6089-6101 (2005)
C.TZ.02.CO178	AY734556	All	Arroyo, M.A.	*AIDS* **19**(14):1517-1524 (2005)
C.UY.01.TRA3011	AY563169	All	Carrion, G.	*ARHR* **20**(9):1022-1025 (2004)

Name	Accession	Proteins	Author	Reference
C.YE.02.02YE511	AY795906	All	Saad, M.D.	*ARHR* **21**(7):644-648 (2005)
C.ZA.04.04ZASK164B1	DQ056405	All	Rousseau, C.M.	*J Virol Methods* **136**(1-2):118-125 (2006)
C.ZA.05.05ZAPSK240B1	DQ369991	All	Rousseau, C.M.	*J Virol Methods* **136**(1-2):118-125 (2006)
C.ZM.02.02ZM108	AB254141	All	Tatsumi	Unpublished
D.CD.83.ELI	K03454	All	Alizon, M.	*Cell* **46**(1):63-74 (1986)
D.CM.01.01CM_0009BBY	AY371155	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
D.KE.01.NKU3006	AF457090	All	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
D.KR.04.04KBH8	DQ054367	All	Cho	Unpublished
D.TD.99.MN011	AJ488926	All	Vidal, N.	*JAIDS* **33**(2):239-246 (2003)
D.TZ.01.A280	AY253311	All	Arroyo, M.A.	*ARHR* **20**(8):895-901 (2004)
D.UG.99.99UGK09259	AF484498	All	Harris, M.E.	*ARHR* **18**(17):1281-1290 (2002)
D.YE.01.01YE386	AY795903	All	Saad, M.D.	*ARHR* **21**(7):644-648 (2005)
D.YE.02.02YE516	AY795907	All	Saad, M.D.	*ARHR* **21**(7):644-648 (2005)
D.ZA.90.R1	EF633445	All	Jacobs, G.B.	*ARHR* **23**(12):1575-8 (2007)
F1.AR.02.ARE933	DQ189088	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Aulicino, P.C.	*ARHR* **21**(2):158-164 (2005)
F1.BE.93.VI850	AF077336	All	Laukkanen, T.	*Virology* **269**(1):95-104 (2000)
F1.BR.01.01BR125	DQ358802	All	Sanabani, S.	*Infect Genet Evol* **6**(5):368-377 (2006)
F1.ES.x.P1146	DQ979023	All	Sierra	Unpublished
F1.FI.93.FIN9363	AF075703	All	Laukkanen, T.	*Virology* **269**(1):95-104 (2000)
F2.CM.02.02CM_0016BBY	AY371158	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
F2.CM.95.MP257	AJ249237	All	Triques, K.	*ARHR* **16**(2):139-151 (2000)
F2.CM.97.CM53657	AF377956	All	Carr, J.K.	*Virology* **286**(1):168-181 (2001)
G.BE.96.DRCBL	AF084936	All	Oelrichs, R.B.	*ARHR* **15**(6):585-589 (1999)
G.CM.01.01CM_4049HAN	AY371121	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
G.CU.99.Cu74	AY586547	All	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
G.ES.99.X138	AF450098	All	Delgado, E.	*JAIDS* **29**(5):536-543 (2002)
G.GH.03.03GH175G	AB287004	All	Takekawa	Unpublished
G.KE.93.HH8793_12_1	AF061641	All	Salminen, M.	*ARHR* **8**(9):1733-1742 (1992)
G.NG.01.01NGPL0669	DQ168576	All	Carr	Unpublished
G.PT.x.PT2695	AY612637	All	Esteves	Unpublished
G.SE.93.SE6165	AF061642	All	Carr, J.K.	*Virology* **247**(1):22-31 (1998)
H.BE.93.VI991	AF190127	All	Janssens, W.	*AIDS* **14**(11):1533-1543 (2000)
H.BE.93.VI997	AF190128	All	Janssens, W.	*AIDS* **14**(11):1533-1543 (2000)
H.CF.90.056	AF005496	All	Gao, F.	*J Virol* **72**(7):5680-5698 (1998)
J.CD.97.J_97DC_KTB147	EF614151	All	Abecasis, A.B.	*J Virol* **81**(16):8543-8551 (2007)
J.SE.93.SE7887	AF082394	All	Laukkanen, T.	*ARHR* **15**(3):293-297 (1999)
K.CD.97.EQTB11C	AJ249235	All	Triques, K.	*ARHR* **16**(2):139-151 (2000)
K.CM.96.MP535	AJ249239	All	Triques, K.	*ARHR* **16**(2):139-151 (2000)

Name	Accession	Proteins	Author	Reference
01_AE.CF.90.90CF11697	AF197340	All	Anderson, J.P.	*J Virol* **74**(22):10752-10765 (2000)
01_AE.CN.05.FJ051	DQ859178	All	Huang, H.L.	*Chin Med J* **119**(19):1622-1628 (2006)
01_AE.CN.06.FJ054	DQ859180	All	Huang, H.L.	*Chin Med J* **119**(19):1622-1628 (2006)
01_AE.HK.x.HK001	DQ234790	All	Tsui	Unpublished
01_AE.TH.01.01TH_R2184	AY945730	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Watanaveeradej, V.	*ARHR* **22**(8):801-7 (2006)
01_AE.TH.02.OUR769I	AY358062	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Tovanabutra, S.	*ARHR* **20**(5):465-75 (2004)
01_AE.TH.90.CM240	U54771	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Carr, J.K.	*J Virol* **70**(9):5935-5943 (1996)
01_AE.US.00.00US_MSC1164	AY444804	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Tovanabutra, S.	*ARHR* **21**(5):424-429 (2005)
02_AG.CM.02.02CM_4082STN	AY371141	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
02_AG.EC.x.ECU41	AY151001	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Carrion, G.	*ARHR* **19**(4):329-32 (2003)
02_AG.GH.03.GHNJ196	AB231898	All	Tatsumi	Unpublished
02_AG.NG.01.PL0710	DQ168577	All	Carr	Unpublished
02_AG.NG.x.IBNG	L39106	All	Howard, T.M.	*ARHR* **10**(12):1755-1757 (1994)
02_AG.SE.94.SE7812	AF107770	All	Laukkanen	Unpublished
02_AG.SN.98.MP1211	AJ251056	All	Toure-Kane, C.	*ARHR* **16**(6):603-609 (2000)
03_AB.RU.97.KAL153_2	AF193276	All	Liitsola, K.	*ARHR* **16**(11):1047-1053 (2000)
04_cpx.CY.94.CY032	AF049337	All	Gao, F.	*J Virol* **72**(12):10234-10241 (1998)
05_DF.BE.x.VI1310	AF193253	All	Laukkanen, T.	*Virology* **269**(1):95-104 (2000)
06_cpx.AU.96.BFP90	AF064699	All	Oelrichs, R.B.	*ARHR* **14**(16):1495-1500 (1998)
08_BC.CN.97.97CNGX_6F	AY008715	All	Piyasirisilp, S.	*J Virol* **74**(23):11286-11295 (2000)
09_cpx.GH.96.96GH2911	AY093605	All	McCutchan, F.E.	*ARHR* **20**(8):819-826 (2004)
10_CD.TZ.96.96TZ_BF061	AF289548	All	Koulinska, I.N.	*ARHR* **17**(5):423-431 (2001)
11_cpx.GR.x.GR17	AF179368	All	Paraskevis, D.	*ARHR* **16**(9):845-855 (2000)
12_BF.AR.99.ARMA159	AF385936	All	Carr, J.K.	*AIDS* **15**(15):F41-F47 (2001)
13_cpx.CM.96.1849	AF460972	All	Wilbe, K.	*ARHR* **18**(12):849-56 (2002)
14_BG.DE.01.9196_01	AY882421	All	Harris, B.	*ARHR* **21**(7):654-660 (2005)
15_01B.TH.99.99TH_MU2079	AF516184	All	Viputtijul, K.	*ARHR* **18**(16):1235-1237 (2002)
16_A2D.KR.97.97KR004	AF286239	All	Gao, F.	*ARHR* **17**(8):675-688 (2001)
17_BF.PE.02.PE02_PCR0155	EU581828	All	Carr	Unpublished
18_cpx.CU.99.CU76	AY586540	All	Thomson, M.M.	*AIDS* **19**(11):1155-63 (2005)
19_cpx.CU.99.CU7	AY894994	All	Casado, G.	*JAIDS* **40**(5):532-537 (2005)
20_BG.CU.99.Cu103	AY586545	All	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
21_A2D.KE.91.KNH1254	AY945737	All	Visawapoka, U.	*ARHR* **22**(7):695-702 (2006)

Name	Accession	Proteins	Author	Reference
22_01A1.CM.01.01CM_0001BBY	AY371159	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
23_BG.CU.03.CB118	AY900571	All	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
24_BG.CU.03.CB378	AY900574	All	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
24_BG.CU.03.CB471	AY900575	All	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
25_cpx.CM.01.101BA	DQ826726	All	Carr	Unpublished
27_cpx.CD.97.97CDKTB49	AJ404325	All	Vidal, N.	*ARHR* **16**(18):2059-2064 (2000)
28_BF.BR.99.BREPM12609	DQ085873	All	Sa Filho, D.J.	*ARHR* **22**(1):1-13 (2006)
29_BF.BR.02.BREPM119	AY771590	All	Sa Filho, D.J.	*ARHR* **21**(2):145-151 (2005)
31_BC.BR.02.110PA	EF091932	All	Santos, A.F.	*AIDS* **20**(16):2011-2019 (2006)
32_06A1.EE.01.EE0369	AY535660	All	Adojaan, M.	*JAIDS* **39**(5):598-605 (2005)
33_01B.MY.05.05MYKL007_1	DQ366659	All	Tee, K.K.	*JAIDS* **43**(5):523-529 (2006)
34_01B.TH.99.OUR2478P	EF165541	All	Tovanabutra, S.	*ARHR* **23**(6):829-833 (2007)
35_AD.AF.05.05AF095	EF158041	All	Sanders-Buell, E.	*ARHR* **23**(6):834-839 (2007)
36_cpx.CM.00.00CMNYU830	EF087994	All	Powell, R.L.	*ARHR* **23**(8):1008-1019 (2007)
37_cpx.CM.00.00CMNYU926	EF116594	All	Powell, R.L.	*ARHR* **23**(7):923-933 (2007)
38_BF1.UY.05.UY05_4752	FJ213780	All	Ruchansky	Unpublished
39_BF.BR.03.03BRRJ103	EU735534	All	Guimaraes, M.L.	*AIDS* **22**(3):433-435 (2008)
40_BF.BR.04.04BRRJ115	EU735538	All	Guimaraes, M.L.	*AIDS* **22**(3):433-435 (2008)
42_BF.LU.03.luBF_05_03	EU170155	All	Struck	Unpublished
43_02G.SA.03.J11223	EU697904	All	Badreddine, S.	*ARHR* **23**(5):667-674 (2007)
U.CD.83.83CD003	AF286236	All	Gao, F.	*ARHR* **17**(12):1217-1222 (2001)
U.CD.90.90CD121E12	AF457101	All	Mokili, J.L.	*ARHR* **18**(11):817-823 (2002)
U.GR.99.GR303	AY046058	All	Paraskevis, D.	*J Gen Virol* **82**(Pt 10):2509-2514 (2001)
U.NL.95.U_NL_95_H10986_D1	EF029066	All	van der Hoek, L.	*ARHR* **23**(3):466-470 (2007)
CPZ.CD.90.ANT	U42720	All	Vanden Haesevelde, M.M.	*Virology* **221**(2):346-350 (1996)
CPZ.CM.01.SIVcpzCAM13	AY169968	All	Nerrienet, E.	*J Virol* **79**(2):1312-9 (2005)
CPZ.CM.05.SIVcpzEK505	DQ373065	All	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.CM.05.SIVcpzLB7	DQ373064	All	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.CM.05.SIVcpzMB66	DQ373063	All	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.CM.05.SIVcpzMT145	DQ373066	All	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.GA.88.GAB1	X52154	All	Huet, T.	*Nature* **345**(6273):356-359 (1990)
CPZ.TZ.00.TAN1	AF447763	All	Santiago, M.L.	*J Virol* **77**(3):2233-2242 (2003)
CPZ.US.85.CPZUS	AF103818	All	Gao, F.	*Nature* **397**(6718):436-441 (1999)
CPZ.CM.x.SIVcpzDP943	EF535993	All	Heuverswyn, F.V.	*Virology* **368**(1):155-171 (2007)
CPZ.CM.x.SIVcpzMB897	EF535994	All	Heuverswyn, F.V.	*Virology* **368**(1):155-171 (2007)
N.CM.02.DJO0131	AY532635	All	Bodelle, P.	*ARHR* **20**(8):902-908 (2004)
N.CM.04.04CM_1015_04	DQ017382	All	Yamaguchi, J.	*ARHR* **22**(1):83-92 (2006)
N.CM.04.04CM_1131_03	DQ017383	All	Yamaguchi, J.	*ARHR* **22**(1):83-92 (2006)
N.CM.95.YBF30	AJ006022	All	Simon, F.	*Nat Med* **4**(9):1032-1037 (1998)
N.CM.97.YBF106	AJ271370	All	Roques, P.	*AIDS* **18**(10):1371-1381 (2004)
O.BE.87.ANT70	L20587	All	Vanden Haesevelde, M.	*J Virol* **68**(3):1586-1596 (1994)
O.CM.91.MVP5180	L20571	All	Gurtler, L.G.	*J Virol* **68**(3):1581-1585 (1994)
O.CM.96.96CMABB637	AY169810	All	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.98.98CMA104	AY169802	All	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.99.99CMU4122	AY169815	All	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.FR.92.VAU	AF407418	All	Vartanian, J.P.	*J Gen Virol* **83**(Pt 4):801-805 (2002)

Name	Accession	Proteins	Author	Reference
O.SN.99.SEMP1299	AJ302646	All	Toure-Kane, C.	*ARHR* **17**(12):1211-1216 (2001)
O.US.99.99USTWLA	AY169814	All	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)

```
                        membrane binding                                                                          phosphorylation site
                        Gag start, p17 start nuclear localization                                                 nuclear localization                          p17 end  p24 start
B.FR.83.HXB2            MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEEEQNKSKKKAQQAAA..........DTGHSNQ.............VSQNYPIVQNIQGQMVHQAI  147
A1.GE.99.99GEMZ011      ..---------Q--K--S-----------R---L-----------L--S----T---Q--ME---AFR--T-D-K-----I-V--------DV-------X--X-IX--X-QXX---..............T-XT...............G---------A----I--SM  140
A1.KE.00.KER2008        ..---------K--A--------------R---L--------K--L-----------Q--MN-I--A----T------F-A------------V---------V--I-----Q-T-----..............---NNSK..............--H-------A----I--SL  145
A1.KE.00.KNH1144        ..---------K-----------------RM--L-----------L--S----A---Q--IE-I--A-K--T------F--I---------K-DV------------L-----Q-T-----..............---SGSKVS............----------A----I--TL  147
A1.KE.00.KSM4024        ..---------K--A-------------------L-------D---L--S----T---Q--ME---A-K--T---K--F----------------V------------L.H---Q-T-----..............---N-SK...............----------A----I--SL  144
A1.KE.00.MSA4069        ..----I----K--A--------------R---L------M----L-------T---Q--MK------T--T-----------V--------DV------------M----QQ-T-----..............---N-ST...............----------A----I--PL  145
A1.KE.00.NKU3005        ..---------K--A---------R---R----------------L-------T---Q---E----A----T--------------------DV------------IKKR--Q-T-----..............---S-S-...............----------A----I--NL  145
A1.RU.00.RU00051        ------I----K--A---------R-Q-R---L-----------L--S--------Q---E----A-R--X---K-----X-------L--DV------------I-----Q-T--T-T..........GA-N--K...............----------A-----X-SX  147
A1.RW.93.93RW_024       ..---------K--A--------------RM--L-----------L--S---ST---Q--IE----A-K--T--IK-------------R--DV---------V--I-----Q-T-----..............---N-SRS..............----------A-------PL  146
A1.SE.95.SE8891         ..-----K---EKK-A---M-------------M---R----KS-L--S----T---Q--MDK---A-R--T------------------K-AV------------I---N-Q-T-----..............---N-SN...............--R-------A-------SL  145
A1.SE.95.UGSE8131       -----------K--A---------N---R---L--------K--L--------A---Q--IE---SAPK--T--------------------DV------------I-----Q-KE----..............---N-SK...............-------M--A-------VM  147
A1.TZ.01.A173           ..----I----K--A--------------R---L-------D---L--S----T---Q--MN----AVK--T--IK--F--------------DV------------I-----Q-T-----..............---D-SK...............-------I--A----I--NL  145
A1.UA.01.01UADN139      ..---------K--A--------------RI--L-----------L--S----A---Q---E----A-K------K------V-----------R----G------I-----Q-T--S-T..........G--S-SK...............----------A----T--SM  145
A1.UG.92.92UG037        -----------K--A--------------R---L-----------L--S----T---Q--ME---SA-R--T--------------------V------------I-K---Q-T-----..............---S-SK...............----------A----I--SL  147
A1.UG.99.99UGA07072     ..---------K--A--------------R---L-------D---L-------A---Q--MA---AIK--T--IK--F--------------D------------M-D---Q-T----V..............---S-SK...............----------A----I--PV  145
A2.CD.97.97CDKTB48      -----------K-EA--------------R---L--------K-SI--S----ET---R-F-----A-E--T----------I-V--F--K--V---------------C-Q-T-----..............---S-SSQNYRGS.......S---------A--------V  153
A2.CY.94.94CY017_41     ------I----K--A--------------R---L--------K-SI-------P------IR----A----T---K------VV--W----VDV---------------Q-TQHAA-..............---N-S................----------A---------  143
A.SN.01.DDI579          .#---------K-NK-------------Q-RI--L-----------L--S----T---Q--MQ----A-K--T--------------------DV------------I---N-Q-T----T..............---N-R-...............----------A-------SL  145
A.SN.01.DDJ369          .#---------K--A--------------RM--LI----------L--S----A---Q--IE---SA-K--T---K--F--------------V------------I-KRH-Q-T-----A..............---S-R-...............----------S-------SL  146
A.SN.96.DDJ360          .#---------K--K-------------Q-RM--LI----------L--S----T---Q--MD----A-K--T------F--------------DV-----V------I-KG--Q-T-----..............---N-R-...............----------A-------SL  145
A.CD.02.02CD_KTB035     -----------K--A--------------R---L-------D---L--S--------Q--IEK---V-S--T--IK----L-------------VR---------AL----QRT-----A..........AGDS-K................----------A-------S-  147
A.CD.97.97CD_KCC2       -----------K--S--------------RM--L--------K--L-------A---Q--MR---SA----T---K-----------------K----------------L----QQ-T----S..........AGDS-R................----------A--------L  147
A.CD.97.97CD_KTB13      -----------K--S--------------R---L-----------L--S----G---Q--M-----A-N--T--IK-----------------D---------T---M-------QAA-..............NA-DNSK..............----------A--------L  145
B.AR.04.04AR151516      .#-----I---K-----------H---R----------------------------E----A-K----F--------------KMDV------------------S---------T..............---NNS-...............----------L--------L  145
B.AU.87.MBC925          -----------------R-----R----Q-------------S---------------R----A------DF----------------V-------ERV-------E-----------..............---N-SKQ..............----------M---------  148
B.BO.99.BOL0122         ..--------A----K--------------R--------------------------G---------------------------V-F----K--V--------------------------..............---N--GNSNQ..........--H-F-----L-------SL  149
B.BR.03.BREPM2012       -----------Q--K--R-----------R----------------------S-D------E------K-------------V-------VDV-------E--------------------QQAAA.....N--NNS-...............----------L-------P-  152
B.CA.97.CANB3FULL       -----------D-----R-----------R------------------------G------E---------------------I-V------K--V-------------------------..............G-EN-S-..............----------L--------L  147
B.CN.05.05CNHB_hp3      -----------Q-----R------------------------------------A-------Q--------------K--F-A--V-----------V-------E-------------------A..............---NNS-...............--H-------L-------P-  148
B.CO.01.PCM001          ..---------------F----------------------------------S--------------------------K--F--I-V----------V-------R----------------..............---N-S-...............----------L--------L  145
B.GB.83.CAM1            ---------------K--------------------------------------------------------------------K-DV-------E--------------------..............G--N-S-...............----------L---------  147
B.GB.86.GB8             ----------------------------R---V---------------------------E-----------R---F--------------DV---------------S------------..............-K-N--SNQ............----------L---------  149
B.GE.03.03GEMZ010       ..---------K--Q-------------Q---------G-----Y-------------K--I------P------K-------V--------GV---Q---E------------------..............--ENK--..............----------L---------  145
B.IT.05.SG1             ------LXC--K----------------Q-----------------------------G------------------K-------V------G----------E---------------V--A..............---NNS-...............G-----------------L  148
B.JP.05.DR6538          ---------------K--------------------------------------A---K---------A---------K--F--I-V------K-A-Q------E-----------------..............---N-S-...............----------L---------  147
B.KR.05.05CSR3          -----------Q--------------------L-----------------S----A------I---------------K--H-AI-------R--DV-------E-VA---S----T---------..............--ASNS-...............--------------------  147
B.NL.00.671_00T36       ------------------------R----------------------------------------A--------K--F----------A---V-------E-----------R----QQAEA........-A-KN-P...............----------L---------  150
B.RU.04.04RU128005      X-----I--A----K--------------R-----X---------------X-----------E------------X-------I-V------K--V-------E-----------------..............S--N-S-...............----------L---X-----  147
B.TH.00.00TH_C3198      ..---------Q-----R-----------R---L----------------------------------------F--------------DV-------E-V-------Q----T--AA........G--N-S-...............-----------------P-  147
B.UA.01.01UAKV167       ..---------A--K-------------------------------------------------------K-------V------K-------------------------..............-AANKS-...............----------L--------L  145
B.US.04.ES10_53         -----------K--K-------------T-----------------------------A------M-H-H---------------A-----------DVR-------RV-------C---------NTET......N--N-S-...............----------L---------  151
B.US.99.PRB959_03       -----------K--Q--R-------N-R-------------------------------A-------------------K-----I-----------DV--------------------------..............A--NNS-SSQ.............----------L-------PL  150
C.AR.01.ARG4006         ..----I-R--K--T--R-K-K-----T-MM--L-----------D------T------IR----A-----K--I-------------AK---R---------K---E--QQ-T---..............EAADKGK..............----------L-------PL  143
C.BR.04.04BR013         --------R-EK--T--R-----------MM--L-----------LA------A------IK----A-K--TK--I-------------K---R---------Q------QL-I---..............EAADKEK.............I---------H---W---P-  145
C.BW.00.00BW07621       ------I-R-EK--K----------R-H-MI----------------L-----------K--IK----A----T-----------------KK-DVR-------Q--K---Q-QQ-T---..............E-ADKK..............------------------SL  144
C.CN.98.YNRL9840        ------I-R--K--K------------H-M---L-----------L--------D--K--IK----A----T------F-----------EG--VR----------------IQQ-T---..............KEDDGK...............----------L---------  144
C.ET.02.02ET_288        ..------R-EK--K------------H-M---L-----------L--------T--K--IK----A-K--T------F-----------EG-D-R-------R-------VQQ-T---G..........EADKGK...............----------L-------SL  143
C.GE.03.03GEMZ033       ..----I-R--K--K------------H-M---L-----------L-------AD--K--IK----A----T------------------KG--VR---D---------KECQR-T---..............EAADKK...............----------L---------  142
C.IL.99.99ET7           ..----I-R--Q---------------H-M---L-----------VL-------A---K--MR--H-A----T---K-------------KN-DV------------------QQ-T---..............TATDET...............----------L-----Y-NL  142
C.IN.99.01IN565_10      --------R--K--K------------R-MI--L-----------L-------A---K--IK--H-A----T------------------AD--VR---D-----------CQQ-T---KGA..........-GKV-QKGADGK.........--L-------L---------  153
C.KE.00.KER2010         ..----I-R--K--A------------H-R---L-----------L-----------K--IK----A----T------------------NK--VR-----------------QQ-T--G...........KAADGK...............----------L--------L  142
C.MM.99.mIDU101_3       ------I-R--K--K------------H-M---L-----------L-----------K--IK----A----T------F-----------SG--VR----------------I-Q-#---..............KEADGK...............----------L---------  143
C.MW.93.93MW_965        ..----I-R--K--T----K-------H-M---L-----------L--A----A---K--IK----A----T------H-----------AN---R-----------------QQ-T-K-..............EAADKGK..............----------H---------  143
C.SN.90.90SE_364        ..----I-R--K--S--R---------H-M---L-----------L-------------MK----A----T--------------------D-KVQ-----------------QQ-T---..............EAAGGK...............----------L-------PL  142
C.SO.89.89SM_145        ..----I-R--K--A--R---------H-MI--L-----------L-----------K--MK----A----T------F-----------EK--VR-----------------QQ-T---..............KAADGK...............----------L---------  142
C.TZ.02.CO178           ..----I-R--K--K--R---------H--I--L-----------L-----------K--IK--H-A----T------F-A-SV------G---R----------------CQQ-T---..............KAADGK...............----------L-------P-  142
C.UY.01.TRA3011         ..----I-R--K--T--R---K-----H-MM--L-----------LD----------K--IK----A----TK--I-------------EK--VR--------K-----KQQAE...............AADKGK...............----------V-------PL  139
C.YE.02.02YE511         ..----I-R--K--T--------------M---L---------------------K--IR----A----TA--K-------------KD-AV-----------------QQ-T---..............TAEDKK...............------------------SL  142
C.ZA.04.04ZASK164B1     ------I-R--K--T----K-------H-M---L-----------L--S----E--K--VK---------T------F----V-----KG-KVL----------------CQQ-TK--E...........KAADGK...............----------L---------  145
C.ZA.05.05ZAPSK240B1    --------R-EK--K------------H-M---------------L------SA---K--MQ----A----T------------------R--DVR-----------------QQ-T--.............AKADGK...............----------L-------P-  143
C.ZM.02.02ZM108         ------I-R-EK--A------------Q-M---L-----------L-----------K--MK---------T-----------V-----E---VR---------------K-CQQ-I-AQQ-...........EAADKK...............----------L---------  146
D.CD.83.ELI             -----------K--K--------------R------------Y-L-----------K--I-----AI---T------------------KG-DV-------E-M----------------..............---NNS-...............----------L---------  147
D.CM.01.01CM_0009BBY    ..---------K--A---------R-R-R----------------L-----------K--IS------K--------F----------N---V-------------------.-KAQQ-AA...........NK-NNS-...............----------L---------  141
D.KE.01.NKU3006         ..---------K--E--------VN-R------------------I-------P-------E----AI-----GIK--------------K--QVA-------------T--------TT-...........--KS-S-...............----------L--------L  145
D.KR.04.04KBH8          -----------K--T*-------R--M-------------K---L--------K-----I------I-----K-K--F-----------E--NV---------------------------T-...........--KNNS-...............----------L-------PL  146
D.TD.99.MN011           -----------K--D--R-Q-----N--------------------L--------D--K---------AIK----------F-----------EK-QV------------------------VL-...........---N--...............--H-------L--------L  147
D.TZ.01.A280            ..---------Q--A--------------Q----------------L-----------K-#....--AI-------K--F-----------RK--V-------E-L---T----------T-...........---S-S-...............----------L---------  141
D.UG.99.99UGK09259      ..------R-EK--D--G-----------R---L-----------L-------A---K---E-----I------VK--F---V-L---A--KVA-------------A---R-----T-...........---N-S-...............--K-----R---------P-  145
D.YE.01.01YE386         ..---------K--D----------N-R-----------------L-----------K--M-K---TI-----GVK--------------EG-K-----------L----T---E------..............--RN-P...............----------L---------  144
D.YE.02.02YE516         ..---------K--E----------S---R---------------L--------D--K---------A-K---------F-----------E--GV------------------I----P-----..............---N-S-...............----------L---------  145
D.ZA.90.R1              -----------R--A--R---------Q-R---------------L-----------K--MA-----I--------------------E--DV---Q-----------------------..............---N-S-...............----------L-------P-  147
F1.AR.02.ARE933         -----------K--T--------------R---L-----------ID------A----K-I---H--I-------K-------V--S------V-------E-L--D----QQ-R-----..............--....................-----V---P-----Y-SL  143
F1.BE.93.VI850          ------I----K--E----Q-------R--M--LI----------LD---------QK-IR-------------K--F----V--Y----AGVT--------L-------QQ-T-----..............-K-...................----------L-------SL  143
F1.BR.01.01BR125        -----------K--A--------------R---L-----------L------------K-I--------------K--F----V--Y----V-V-------E-L--------SQQKTQQ-AA........EK-...................----------L-------SL  145
F1.ES.x.P1146           -----------K--A--------------R---L-----------D-------P---XK-IA----------------------I-V--Y----V-V----D--E-L--------Q-T-----..............-K-...................----------L--------L  143
F1.FI.93.FIN9363        -----------K--A-------------Q-RI--L-----------ID---------QK-IA-I---I-------------I-V--F---K--V----------L-------QQ-T-----AA.........-K-...................----------L---------  145
F2.CM.02.02CM_0016BBY   ..---------K--D--------------R--------K------L-------T---K--I----S--------IK--------V--Y---K-Q-R---------LQ---D-YQQ-T-P---..............-K-...................----------L---------  141
F2.CM.95.MP257          -----------K--A--------------R-----------K---L-------T---KK-I-------------K--F--IVV--Y---K--VR---------LQ---D-HQQ-T---T-DKGV......SK-...................-------L--L-------SL  147
F2.CM.97.CM53657        ..----L----K--DL-------------R-----------------L--------K---K--I-------------K--F--IVV--Y-----K-G---------LQ---D--QQ-T-P---..............-K-...................----------L-------SL  141
G.BE.96.DRCBL           -----------K--A-------------R-RM--L-------D---L-------A---QK-MA----A----T--IK--F------------K--VR------EEV-KI-K--QQ-..........................ENS-S-...............----------A---------  140
G.CM.01.01CM_4049HAN    ..---------K--S--------------Q----------M----L--D----A---Q-AM--I--A----T--I---L--L-V------K--V------EEV-KI-KR-QQ-I----R.........-E-D---..............S---------A----I----  145
G.CU.99.Cu74            -----------R--A--------------RM--L--------K--L--------G---Q--MN----A---AT---K----------------K--V------EEV-KA-K--QQ-T----R.........-E-NNS-...............----------A---P-----  147
G.ES.99.X138            -----------K--A--------------RX--L-----------L--D----A---Q--X-----A----T------F-A-----------GV-------EEV-KA-K--Q--Q----M.........-E-NNS-..............A---------A---------  147
G.GH.03.03GH175G        -----------K--A--------------RM--L------M----L-S-----A---Q--MM----A----T------F---------------V-------E-V-KL-KN-QQ-Q-A-K..........AE-NN--...............----------A-------P-  146
G.KE.93.HH8793_12_1     ----X-X----K--A----X-T-------R---L------M----L-------A---Q--MS----AIX--T--IK--F-----------PK--V-------EEV-KI-K--QQ-I----R.........-E-N-S-...............----------A---------  147
G.NG.01.01NGPL0669      ..---------K--A--------------RI--L-----------L-------A---Q--MN----A----T--IK----A---------R---VR-------EV-KL-KN-Q-ET-K--T..........S--NN--...............----------A---------  145
G.PT.x.PT2695           -----------K--A--------------RM--L--------K--L--D----A---Q--MR----A----T------F---------------V-------EEV-KT-K--Q-QQ-A-M.........-E-N-S-...............----------A---------  146
G.SE.93.SE6165          ---------T--K--A---------R-S--I--L-----------L--D----A---Q--MR---------T--IK----A-------------V-------EEV-KI-K--QE-I----M.........-K-N---...............----------A---------  147
H.BE.93.VI991           -----------K--A---------R---R---L-----------L--D----AD--Q---------A-K--T-D-Q-----I-V--------DV-------G----I---N-QRT---P-AA.........-KEKDSK..............I---------A---------  149
H.BE.93.VI997           -----------R--TL-------------R-----------------L------SA---L--IE--R--IK--T---X--F---------L-----V-------G----I---RQQ-T---T-..........NKERD-K..............----------A-------P-  147
H.CF.90.056             -----------K--A--------------R---L-----------L-------P---L--IE-I--AIK--T---K--F-L--V-----RK-DV-------------I----QQ-T-----..............-KEKD-K..............----------A---------  147
J.CD.97.J_97DC_KTB147   ....#-ESI---SEA----W--RR-E---RR----------D-Y-L-----**A--GE---V-I--D-*--A--IK--F--E-------*--D-R-----------L*----E--KKE-..........VKKYNS-..............--H----L--M*-EL----L  135
J.SE.93.SE7887          ------I----K--D--------------RI--L-------D---L-----SAK--Q---V----A----TQ-IK---------------------M---E----I---N-QQ--K-ET.........-KKDNS-...............----------L---P----L  147
K.CD.97.EQTB11C         -----------K--K----Q---------R---L-----------L--N----V------IR----------------F-------W---S-QVR---------L-----RTQQ-T-..............QGKADKG..............----------L--------L  143
K.CM.96.MP535           -----------K--A-------------------L-----------L-------T------IT-I---I------IK-----I-V--F---K--V---------L-------QR-T-..............QEAADKG..............----------L--------L  143
```

```
                                membrane binding
                          Gag start, p17 start  nuclear localization                                                          phosphorylation site
                                                                                                                              nuclear localization                                   p17 end  p24 start
B.FR.83.HXB2                MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEEEQNKSKKKAQQAAA.........DTGHSNQ.............VSQNYPIVQNIQGQMVHQAI  147
01_AE.CF.90.90CF11697       -----------K--A--------------RM--L-----------L-------A--Q-LIE---SN-K-----IK--F--I---W------DV-----------V---N-Q-T-----.........G--SNSK.............----------A-------PL  147
01_AE.CN.05.FJ051           -----------K--A--------------RM--L-----------I-------A--Q--IE---ST-K------K--F--I-V-W------D-------------A-R--QQ-T-----.........G--S-SK.............----------S-------PL  147
01_AE.CN.06.FJ054           --------T--K--A--------------RM--L-----------L-------A--Q-LIE---ST-K------K------V--W------DV---------M--V-K---Q-T-----.........---S-SK.............----------A-------SL  147
01_AE.HK.x.HK001            ------I-T-EK--A--------------MI--L-----------L-------A--Q--IE---ST-K------K--F------W-------V---------L--I---NQ--T-----.........G--SNSK.............----------A----I--SL  147
01_AE.TH.01.01TH_R2184      ..---------K--A--------------R---L-----------L-------A--Q--IE---ST-K------K--F------W---L---V-------------V-K-RQQ-T--T--.........---S-SK.............----------A-------P-  145
01_AE.TH.02.OUR769I         ..---------K--A----------R----M--LI----------L-------A--Q--ME------K--------F---V--W----K--V-------------V----QQ-I-----.........G--S-ST.............----------A-------PV  145
01_AE.TH.90.CM240           -----------K--A----------R---R---L-----------L--SF---A--Q--IE---ST-K--L---K--F------W-------V-------------V---QR-T-----.........G--S-SK.............----------A----A--PL  147
01_AE.US.00.00US_MSC1164    ..------T--K--K----------R-------L-----------L-------A--Q-LIE---T-K------K--F--L---W-------V---------M--V---QQ-T-----.........G--S-GK.............----------A-------PV  145
02_AG.CM.02.02CM_4082STN    ..---------K--A--------------R---L-----------L-------T--Q-LMN---T-KA---VK-----L---W-----------------M--I-----Q-T--E--.........A--S...............S---------A-------SM  142
02_AG.EC.x.ECU41            .................................#-----------L-------A--QHL-E---SH-KI-----K--F--I---W-----------------L--V-----QQK-----.........A--S...............S---------A-------S-  110
02_AG.GH.03.GHNJ196         -----------K--A------------Q-----------------L-------A--Q-L-E---ST-R------K-----I---W------D-R-----------A-----Q-T-----.........A--S---VRV.........S---F----A----I--SM  150
02_AG.NG.01.PL0710          ..---------K--S------------R--I--------------L--D---A--Q-LME---ST-R------K--F--I---W------DV-------------I---N-T-QAA-.........A--G...............S---------A-------SM  140
02_AG.NG.x.IBNG             -----------K--A--------------R---L-----------L-------A--Q-LME---SA-R------K--F--I---W------D-----------M--I-----Q-V--T--.........A--S...............S---------AK---T--SM  144
02_AG.SE.94.SE7812          -----------K--A--------------R---L-----------L-------A--Q-LME---ST-R------K--F--I---W------D-----------V--V-K---Q-T-----.........AA-S...............S---------A----T--S-  144
02_AG.SN.98.MP1211          #----------K--A--R-----------R---L-----------L-------A--Q--IE---ST-K------K--F-----------------V-----------L--I-E---Q-T-----.........A--S...............S---F----A-------S-  143
03_AB.RU.97.KAL153_2        -----------K--A-----------E--RI--L-----------L--S-------Q---E----T-K------K-------------------------------I-----Q-T----T.........G--S-SK.............----------A----T--SM  147
04_cpx.CY.94.CY032          -----------K--A--R-----------R---L-----------L-------A--Q-LME---ST-K-----------IT--W------DVQ-----------I-S---Q-T-----.........AA-G-SN.............----------A-------S-  147
05_DF.BE.x.VI1310           -----------K--A--------------R---L-----------L-----------K--IS----AI---T---K-------------E---V-------E----------R-KA-Q-EA.........GA-N-S-.............A-----------------L  148
06_cpx.AU.96.BFP90          -----------K--E---------------M--L-----------I-------A--Q--IE---ST-K------K--F-------------KVT-----I-----I-K---Q--H---.........A--N-SN.............L---------A--------M  146
08_BC.CN.97.97CNGX_6F       .#----I-R--K--K------------H-M---L-----------L-----------K--IK----A----T------F----------EE--VR---------------IQQ-T--..........AKKAEEK.............----------P-------PL  142
09_cpx.GH.96.96GH2911       -----------K--A------K-------RM--L------------D-S---TG--Q---E------R------K-------------K--VR-----------I-G---Q-T---T-.........N--NNSK.............----------A---------  147
10_CD.TZ.96.96TZ_BF061      -----------K--E------------Q-----L-------D---I----------K--I----AI-----IK------------E--KVA-------------T--------T-.........---N-S-.............----------L-------PL  147
11_cpx.GR.x.GR17            -----------K--A--------------R---L-----------L--S----G---K--M-----A-G--T----------A-----L---V---------V--I------QQAA-...........-S-N--K...........----------A--------M  145
12_BF.AR.99.ARMA159         -----------------------------R---------------------------K-I-----------------F----V------K--V-------E-L-------QQ-T-----.........-K-.................----------L-------PL  143
13_cpx.CM.96.1849           -----------K--K------------R-RM--L-----------L-------T--Q---R------------K-----------------V-------EEV-KV--S-QQRT----K.........NE-N---.............----------A-------S-  147
14_BG.DE.01.9196_01         -----------K--A--------------RM--L-----------L--D----A--Q--M-------K--T--I---F-----------K--V-------EEV-KA-K-XQE-Q-A-M..........-E-NNS-.............A---------A--------L  146
15_01B.TH.99.99TH_MU2079    -----------K--A-------------------L-----------I--S---SA--Q---E---ST-K------K---------W--QN---V-------------V---NQQ-T-----.........G--S-ST.............--H-------A-------PV  147
16_A2D.KR.97.97KR004        -----------K--A--------------R---L-----------I----------K--I---E--IK-----IK--F-A-V-----------VR--------L------H-Q-T-P-..........AADTGSS.............G---------A----TY-NL  145
17_BF.PE.02.PE02_PCR0155    ..---------------------------------------------------------K-IA--------------------V------KV-V-------E-L----X-CQQ-------.........NK-.................----------L-------SL  141
18_cpx.CU.99.CU76           -----------K--K--------------R---L-----------L-------A--Q--IE---S-IK-----IK-------------G-DV--------------I-----Q-T-----.........AA-S-S-.............----------A-------S-  147
19_cpx.CU.99.CU7            ------L----K--A------S-----Q-RM--L-----------I--S---SA--Q--IE---ST-K----------F------W----Q--V----------L--V-----Q-T--ET-..........--RN-SH.............--------HKL--------L  147
20_BG.CU.99.Cu103           -----------K--E--------------RM--L-----------L-------V--Q--IN----A-K--T---K--F-------------GV---------EV-KL-K--QQ-M----T.........GE-N-S-.............----------A---PI----  147
21_A2D.KE.91.KNH1254        ..---------K--A--------------R---L-----------I----------K--IV-----I-------K--F----------X--DVR------E----------N-KV-Q-TA..........--KN-S-.............----------A---W-----  146
22_01A1.CM.01.01CM_0001BBY  ..------R--K--A------------Q-RM--L-------D---LD------A--Q--IE------K-----IK-----I------------DV-------------I-----Q-T--G--.........G--S-SN.............----------A----I---L  145
23_BG.CU.03.CB118           -----------K--A--------------RM--L-------D---L-------A--Q--MS---A-K--T--IK--F-A-------------V--------EV-KL-K--QQ-M----T.........GE-N-G-.............----------A---P-----  147
24_BG.CU.03.CB378           -----------K--S------------Q-RM--L-----------L-------A--Q---N---A-K--T--IK--F-------------GV---N---EEV-KL-K--QQ-M----T.........GE-N-S-.............----------A---PI----  147
24_BG.CU.03.CB471           -----------K--A--------------RM--L-----------L-------AG--Q---N---A----T---K--F-----------K--V-------EEV-KL-K--QQ-M----T.........GE-N-S-.............----------A---PM---L  147
25_cpx.CM.01.101BA          ..---------K--A--------------R---L-----------L-------P--Q--MN---SA-K---Q--K----AI-V-----L--D-------------I-----Q-T--E--.........---S-SSQ.............----------A-------SL  146
27_cpx.CD.97.97CDKTB49      ------E----Q--A----H-------Q-RI--L-----------L-------A---LD-------A-K------Q----L--------R---V-------EEV-KI-RN--Q-T--E-K..........-K-INS-.............----------A---------  147
28_BF.BR.99.BREPM12609      -----------------------------R-----------------------------------I-------------------V------K---------------------------.........-A-N-S-.............----------L--------L  147
29_BF.BR.02.BREPM119        ..........#------------------Q---V-----------------------K--E------------K--F----V------K-DV-------E----------S-Q-A--.........G--N-S-.............-----------------P-  135
31_BC.BR.02.110PA           ------I-R--K--T----K-K-----Q-MM--L-----------LD-------A--K--MN-I--A----TK--I-------------EN--VRN---------K-----IQQ-T---..........EAADKGK.............----------L---------  145
32_06A1.EE.01.EE0369        -----------K--E----------R-------------------L-------A--Q--VE-V--A-R------K--F-------------KVA-----------I-----Q------T.........A--N...............S---------A----T--S-  144
33_01B.MY.05.05MYKL007_1    -----------K--A--------------R-----------D---L-------A--Q-LIE---T-K------K--F--I---W---------------I-----V-T--QQ-K-----.........G--SNSK.............----------A-------SL  147
34_01B.TH.99.OUR2478P       ..---------K---------------S-QM--L-----------L-------A--Q--IE---ST-K------K--F--I---W----G-DV-----------L--V---QQ-T-----.........G--SGXK.............----------A-------PV  145
35_AD.AF.05.05AF095         ..----I----K--A--------------R---L-----------L--S----T---Q--ME----A-K--T------F-----------K-A-------------I-----Q-------.........---S-K..............----------A-------SL  144
36_cpx.CM.00.00CMNYU830     ..---------K--A--------------RI--L-------------D-----SA--Q--IE----T-K------K-----------------DV-----------KL-SN--Q-T-----.........---S-ST.............--R-------A--------L  145
37_cpx.CM.00.00CMNYU926     ..------T--K--A------------Q-RM--L-----------L-------A--Q--IE---ST-K------K--F------W------DVR---------L--V-.NKQ-TQ-A--.........G--S-CK.............I---------A----I--SL  143
38_BF1.UY.05.UY05_4752      ------I-T--K--A--------------R---L-------------D------P----K-IV-------------K-----IVV--F--RKV-V---------L-------QQ-T-----.........GQ-...............----------L-------S-  143
39_BF.BR.03.03BRRJ103       -----------------------------R---L----------------------------I---------------------V------K-DV----------------------------..........---NNS-.............----F----L--------L  147
40_BF.BR.04.04BRRJ115       --------------K--------------Q--------------------------------K------------K--F-------------D-------------------M------AA..........---N-SK.............----------L--------L  149
42_BF.LU.03.luBF_05_03      #---------------------------------------------------------------------A--------K-------------K-DVR-----------------X------..........---N---.............----------L---------  146
43_02G.SA.03.J11223         -----------K--A------X-------RM--X-----------I-----SA--Q--ME-XXST-R------K--F--IX--WX-----D-----------V--M-----Q-T-----.........A--S...............S---------A-------SL  144
U.CD.83.83CD003             -----------K--A----------R---R---L-------K--I-------A--Q--I-----A--------K-I--L-------------V------------V---G-Q-------.........-K-SNS-.............----------M---------  147
U.CD.90.90CD121E12          ...K-------K--A----------R---R---L-----------L-------A--Q--IV----AF-------K-------------------------E----A---G-Q-------.........-K-DNSK.............----------M-------P-  144
U.GR.99.GR303               -----------K--A------------Q-R---L-----------L-----------Q--ME----A-K------K-----------------DV-------------L-KR--Q-T-----.........---GTSK.............----------A----I--PM  147
U.NL.95.U_NL_95_H10986_D1   --------R--Q--T-------------------L-------------------A--Q--IM----A----T----------------CK---------------I---N---TEA..........---KTS-.............--------R-L--------L  144
CPZ.CD.90.ANT               ---G----R-EK--T--S-----------MI--L---RS--Q---LSSS-------EKAIH--S--IEIR-P-II--F--ICV-W---KGEK----EQ-VKTVKMKVMQTQAETGSSQTASRGMLLRLLLLNKQWCQRHLSG.......EGR----IVDAG-IAR--PL  163
CPZ.CM.01.SIVcpzCAM13       --------T--K------V------R---MM--L-----------C---M-SA--LKL-Q--E-A-K---G---F--L-V-W---SN--VE--QK--EQLKKHHGEQVR-NENTENNESRQG....ASAS-G.............T-G-------A------TP-  152
CPZ.CM.05.SIVcpzEK505       --------T--K--Q----Y---------MM--L-----------C---MD-A---A-L-R--E-A-K---G---F--L-V-----NN-KVQN-Q---E-LR-KMKAEQ-EPEPEQ-A..........GAAAAPESS.............I-R---L---A-------PL  150
CPZ.CM.05.SIVcpzLB7         ------I-T--K--D---V------R-R-MM--L-----------C---M--T---E-LIT--E-A-K---N-K--F--IVV-W------A---------TLTAVRQ-QQ-EKEA---AARD......NDS................G-R-----T-A---P----L  147
CPZ.CM.05.SIVcpzMB66        --------T--K-------Y---------MM--L-----------C--S-M--T---K-L-Q--E-A-K---G---F--IVV-W----G-PV---------QLQ-A-Q-G-QEVAA-T-..........-GTST.................--R-F---A-A-------P-  145
CPZ.CM.05.SIVcpzMT145       ------I-T--R--A--------------MM--L-----------C---M-NA---EKL-K--E-A-K---G----F--L-V-W---K---VE--QQ--Q-LKAAV-PEQGETPS...............................A-A-F-V-R-A---LT-NP-  137
CPZ.GA.88.GAB1              --------T--K------V------R-R-MM--L-----------CD---M-SK---TKL-Q--E-A-K---G----F--L-V-W-I-SD-TVE--QK--EQLKRHHGEQQS-TESNSGSREGGA....SQ-A-ASAG............I-G---L---A---------  155
CPZ.TZ.00.TAN1              --------R-DK--T--S---KSR-R---LI--L---GS--Q---M----M-NV---WK-IL-----VDI--P-II--F--ICV-----AGERVQ--E--VKIVKMKLTVQ-NNSTATSSGQRQNAGEKEE-VPPSGNTGNTGRATETPSG-RL--VITDA--VAR--P-  170
CPZ.US.85.CPZUS             --------T--R--A--------------MM--L-----------C---M--AD--L-L-K--E-A-K---G---F--L-V-W---S-VTVE--QQ--V-LK-VVQLQ-QQEEKEQQQQEA......SGSNI...............G-S---VI--A--------M  149
CPZ.CM.x.SIVcpzDP943        ------I-T--R--A--------------MM--L-----------C---M--G---E-L-K--E-A-K---G---F--L-V-W---K--PVE--QQ--T-LK-TAAXEQEV--P.............QPQQDSN.............--R-F-V---A---LM--PM  144
CPZ.CM.x.SIVcpzMB897        --------T--K--D----Y---------MM--L-----------C---MD-T---E-LIN---A-T---G---F--L-V-W------RV---------L--V-Q---V-VK-E............AAAADPSGGE...........-R-F---T-A-------SL  147
N.CM.02.DJO0131             --------T--K--Q--A-Y---------RM--L-----------C---MD-AS--A-L-N--E-A-A---G-----AL-V-----SN-QVHN-Q------K-IKE-KERHKSEPKNPEA.......GAAAAADNN...........I-R---L---A----I--PL  152
N.CM.04.04CM_1015_04        --------T--K--K--A-Y----------M--L-----------C---MD-AH--N-L-N--E-A-N---G-----X-I-V-----ST-PVLN-Q------K-KKEQH-SE-KNPE-..........GAAAAADNN...........I-R---L---A------HPL  149
N.CM.04.04CM_1131_03        --------T--K--K--A-Y---------RM--L-----------C---MD-AK--A-L-N--E-A-N---G-------L-V-----ST-PVLN-Q------K-KKEQH-PE-KNPE-..........GAAAAADNN...........I-R---L---A----I--PL  149
N.CM.95.YBF30               --------T--K--Q--S-Y---------RM--L-----------C---MD-AD--AKL-N--E-A-K-----------AL-V-----S--Q-HN-Q------K-K-EQH-PEPKNPE-..........GAAAATDSN...........I-R---L--TA-------PL  149
N.CM.97.YBF106              --------T--K--Q--A-Y---------R---L-----------C---MD-AN--A-LIN--E-A-K---G----X--L-V-----SN-PVHN-Q------K-K-EQH-SEPKKPE-..........G-AAAADSS...........I-R---L---A-------PL  149
O.BE.87.ANT70               ---S----T-SK--A--Q---K--S----R---L-----------C--E----A--NEKL-Q--E-A-K---DS-Q--W-AIVV-W---N-YK-G--QQ-IQ-LK-VMGSR-SADAAK............EDTSAR-.............AG------S-A---------  145
O.CM.91.MVP5180             --------T-SK--A--R------S--A-R---L---------Y-C-------A--TE-L-Q--E-A-K----D-K--W-AI-V-W---N-FD-R--QQ-IQ-LK-VMASR-SAEAAK............EETSPR-.............T-------T-A---------  145
O.CM.96.96CMABB637          --------T-SK--A--Q---K--S----R---L-------D---C-XE----A--NETL-Q--E-A-K---XX-K--W--I-V-W---N-FKVE--QQ-IQ-LK-VMGNR-SADTAK.............ENN-K-.............T-----V-A-M-------PL  144
O.CM.98.98CMA104            --------T-SK--A--Q-K-K--C----R---L-----------C--E----A--NEKL-Q--E-A-K----S-Q--W--I-V-W---N-FKVE--QQ-IQ-LR--KGGR-SAGAAK............EDTSAR-.............TG----V---A---------L  145
O.CM.99.99CMU4122           ---S----K-SK--A--Q---K--S----R---L-----------C--E----A--NEKL-Q--E-A-K---DS-Q--W--IVV-W--XN-YR-E--QQ-IK-LR-VMXSX-SASDAK............EDTXAR-.............TG-----IP-A-------P-  145
O.FR.92.VAU                 --------T-TK--A--Q---K--S--------LI----G-D---C--E---A--NE-L-Q--E-A-K---S*--W-AI-V-W---N-YKVE--QQ-VR-LK--MK-R-PVSSTK............ENTS-K-.............TG----V-A-A--------L  144
O.SN.99.SEMP1299            --------T-SK--A--Q---K--C----R---L-------D-S-C--E---A--NEKL-Q--E-A-K----S-Q--W--I-V-W---N--KVE--QQ-IQ-LK-VMGSR-SAGTAK............EDTSAR-.............TG------T-A-------SL  145
O.US.99.99USTWLA            I-------T-KK--T*-Q----S-S----Q---L----K-----C--D---A--NE-L-Q--E-A-K---DS-Q--*-AIVV-R---K-YRV---QQ-IQ-LX-VXXSX-SADNX............EDTXTK-.............T-------T-A-X-A---PL  142
```

```
                                                                           CyPA binding                                          major homology region
B.FR.83.HXB2            SPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMT...NNPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVK 314
A1.GE.99.99GEMZ011      -----------I--------------------------M---I------------D-----------L--AQ---FP-------------------------...S-------D------------------S--------------------F--------T-D-- 307
A1.KE.00.KER2008        -----------I--RG----------------------M---I------------------------L--AQ----P---I-D------------P----T---...--------D-------------------V-----K--------------F-V------T---- 312
A1.KE.00.KNH1144        -----------I--------------------------M---I------------D-----------T-------P------------------P---------...S-------D-------------------V-----K--------------F--------T---- 314
A1.KE.00.KSM4024        -----------I--------------------------M---I------------D-------------------P------------------I---------...S---V---D-------------------V-----K--------------F--------T---- 311
A1.KE.00.MSA4069        -----------I---S----------------------M---I------------D-----------L--AQ----P---I--------------P----I---...A-------D-------------------V-----K--------------F-V------T---- 312
A1.KE.00.NKU3005        -----------I--R-------------D----H---M---I--------E---D-----------L--I-------------------------P----A---...G-------D-------------------V-----K--------------F-A------T---- 312
A1.RU.00.RU00051        ---X----X--I-------------T-----------XM-X-I------------D-----------L--AQ---FP---L---------------------...S-------D-------------------V------------------------F--------T---- 314
A1.RW.93.93RW_024       -----------I--------------------------M---V------------D-----------L-------P------------------P---------...G-------D-------------------V-----K--------------F--------T---- 313
A1.SE.95.SE8891         -----------I---G------------------G---M---I------------D-----------L-------P------------------P---------...S-------D-------------------V-----K--------------F--------T---- 312
A1.SE.95.UGSE8131       -----------I--R-------------------H---M---I--------E---D-----------T--I----V------------------------A---...S---V---D-------------------V-----K--------------F-A------T---- 314
A1.TZ.01.A173           -----------I--------------------------V---I------------D-----------L---Q---P---I--------------P---LQ---...G-------N-------------------V-----K--------------F-A------T-D-- 312
A1.UA.01.01UADN139      -----------I--------------------------M---I------------D-----------L--AQ---FP-------------------------...S-------D-------------------V------------------------F--------T-D-- 312
A1.UG.92.92UG037        -----------I--------------------------M---I------------D-----------L------V-------------------P----A---...G-------D-------------------V-----K--------------F--------T---- 314
A1.UG.99.99UGA07072     --------------RG----------------------M---I------------D-------------------P-----------------NI---------...S-------D-------------------V----VK--------E-----F-I------T---- 312
A2.CD.97.97CDKTB48      -------------------------T-----------------------------D-----------L---Q---P--------------A--N----------...S-----------------------------V-----K--------------F--------T---- 320
A2.CY.94.94CY017_41     -------------------------T-----------------------------D-------------------P----------------------------...SD----------------------------V------------------------F--------T---- 310
A.SN.01.DDI579          -----------I-------------------T------M---I------------D-----------M-------P----------------------------...S---V------R-----------------V------------------------F-C------T-D-- 312
A.SN.01.DDJ369          -----------I-D-G------------------H---M---I------------D-------------------P-----------------N----------...S---V---DM-R-----------------V------------------------F-C------T---- 313
A.SN.96.DDJ360          -----------I--------------------------M---I------------D-----------I-------P----------------------------...S---V------R-----------------V------------------------F-C------T---- 312
A.CD.02.02CD_KTB035     -------------------------A-----C-----------------------D-----------L------LP---T------------------------...S-----------------------------V-----K--------------F--------T---- 314
A.CD.97.97CD_KCC2       -------------------------------C------M----------------D-----------M-------P----------------------------...S-----------------------------V-----K--------------F--------T-D-- 314
A.CD.97.97CD_KTB13      --------------R----------------C-----------------------D-----------L-------P----------------------------...S-----------------------------V------------------------F-V--------D-- 312
B.AR.04.04AR151516      -----------I-------------------------------------------------------L--I----V---------------------------...-------------------X---------V---------------------------------------- 312
B.AU.87.MBC925          --------------------------------------------------------------------------------------------------------...---------------------------------------------------------------D-- 315
B.BO.99.BOL0122         -----------I-------------------------------------------------------L--AQ--------I-------------------T---...-----------------------------V---------------------------------------- 316
B.BR.03.BREPM2012       -------------------------------------------------------------------L------------------------------------...---------------------------------------------------------------D-- 319
B.CA.97.CANB3FULL       -------------------------------------------------------------------L------V-----------------------------...---------------M-------------------------------------------------- 314
B.CN.05.05CNHB_hp3      -----------------N-------T----------------------------------------------Q---V----L-D--------------------...A---V-------G----------------G-------------------------------D-- 315
B.CO.01.PCM001          -------------------------------------------------------------------L------------------------------------...H------------------------------------------------------------------ 312
B.GB.83.CAM1            -------------------------------------------------------------------L------------------------------------...-------------------------------------------------------------------- 314
B.GB.86.GB8             ----------------------------A---------------------------------------------V-----------------------------...-----------------------------V---------------------------------D-- 316
B.GE.03.03GEMZ010       -------------------------T-----------------------------D--------------------------I------------------VT---...S------------------------------------------------------------D-- 312
B.IT.05.SG1             ----X---------------X----------------------------------------------L------V-------------------------A---...----------------------------XS-----K-----X-----------I---------- 315
B.JP.05.DR6538          -------------------------------------------------------------------L------------------------------------...-------------------------------------------------------------------- 314
B.KR.05.05CSR3          -------------------------------------------------------------------L------------------------------------...---------------M-------------------------------------------------- 314
B.NL.00.671_00T36       -------------------------------------------------------------------L------------------------------------...----------------------------------K-------------------------------- 317
B.RU.04.04RU128005      -------X---I-------------------------------------------------------L-------------I---X---------------A---...---------------XX----------V------------X--------------T---- 314
B.TH.00.00TH_C3198      -------------------------T-----------------------------------------L------------------------------------...S--A------------------------------------------------------------D-- 314
B.UA.01.01UAKV167       -------------------------------------------------------------------L---------------------------------A---...-----------------------------V---------------------------------D-- 312
B.US.04.ES10_53         -----------I----------------------A--------------------------------L--AQ---V-----I-------------------T---...--------D----------------------------------------------------------- 318
B.US.99.PRB959_03       -------------------------T-----------------------------------------L--AQ--------I-------------------T---...-------------------------------------------------------------------- 317
C.AR.01.ARG4006         -----------I-------------T-----------------------------D-----------L--X----V-------------------N-----A---...S-------D-------------------V-----K--------------F--------T---- 310
C.BR.04.04BR013         -A-----------------------T-----------------------------D-----------L---------------------------------A---...G-Q-----------------------------V-----K--------------F--------T-D-- 312
C.BW.00.00BW07621       -----------I-------------T-----------------------------D-----------L---------------------------------A---...S---V-------------------------V-----K--------------F--------T---- 311
C.CN.98.YNRL9840        -----------I-------------T-----------------------------D-----------L---------------------------S-----A---...G--SV---D-------------------V------------------------F--------T-D-- 311
C.ET.02.02ET_288        -----------I--------I----T-----------------------------D-----------I--------Q------------------------A---...S-------D-------------------V------------------------F--------T-D-- 310
C.GE.03.03GEMZ033       -----------I---G-N-------T-----------------------------D-----------L------V--------------------------A---...S-------D-------M-----------V-----K--------------F--------T-D-- 309
C.IL.99.99ET7           -----------I-------------T-----------------------------D-----------L------V-------------------P---LA---...S---V---D-------------------V-----K--------------F--------T-D-- 309
C.IN.99.01IN565_10      -----------I-------------T-------------------R---------D-----------L---P---------------------------VA---...G---V-----------------------V-----K--------------FR-------T---- 320
C.KE.00.KER2010         -----------I-------------T-----------------------------D-----------L------NP-------------------------A---...S-----------------------------V------------------------F--------T---- 309
C.MM.99.mIDU101_3       -----------I-------------T-----------------------------D-----------L------V--------------------S-----A---...G---V---D-------------------V-----K--------------F--------T-D-- 310
C.MW.93.93MW_965        -----------I-------------T-----------------------------D-----------L---------------------------------A---...--------D-------------------V-----K--------------F-------ST---- 310
C.SN.90.90SE_364        -----------I---G-N-------T-----------------------------D-----------L------V--------------------N----VA---...G-------D-------------------V-----K--------------F--------T-D-- 309
C.SO.89.89SM_145        -----------I--------I----T-----------------------------D-----------L---Q---V-----------------------------...S-------D-------------------V-----K--------------F--------T-D-- 309
C.TZ.02.CO178           ---------------G-N-------T-----------------------------D-----------L---------------------------------A---...S-------D-------------------V-----K--------------F--------T-D-- 309
C.UY.01.TRA3011         -A-----------------------------------------------------D-----------L--------------------------------MA---...----V-------------------------V------------------------F--------T---- 306
C.YE.02.02YE511         -----------I-------------T-----------------------------D-----------L------V--------------------S-----A---...G---V---D-------M-----------V-----K--------------F--------T-D-- 309
C.ZA.04.04ZASK164B1     -----------I---G-N-------T-----------------------------D------------------------------------------------...--------D-------------------V-----K--------------F--------T-D-- 312
C.ZA.05.05ZAPSK240B1    -----------I--------I----T-----------------------------D-----------L---------------------------------A---...----V---D-------------------V------------------------F--------A-D-- 310
C.ZM.02.02ZM108         ----------------------------------V--------------------D-----------L--PQ---V----I-------------------T---...S-------D-------------------V-----K--------------F-------ST---- 313
D.CD.83.ELI             -----------I-------------------------------------------------------L---------------------------------A---...S---------------V-------------V---------------------------------D-- 314
D.CM.01.01CM_0009BBY    -----------I-----N----------A--------------------------------------L------V-------------------------MA-I-...G---V-------------------------V---------------------------------------- 308
D.KE.01.NKU3006         -----------I-------------T-----------------------------------------L--I----N-------------------------A---...S-----------------------------V-----K---------------------------D-- 312
D.KR.04.04KBH8          -----------------N-------T------------S---A-------------------S----L-------------L----------------D------...S-----------------------------VG----K-------------------------------- 313
D.TD.99.MN011           -----------I-----N-------------------------------------------------L------V-----------------------------...S-----------------------------V---------------------------------D-- 314
D.TZ.01.A280            -----------I-------------------------------------------------------L---------------------------------A---...----V-------------------------V---------------------------------D-- 308
D.UG.99.99UGK09259      -----------I-----N-------------------------------------------------L---------------------------------D---...G-----------------------------V---------------------------------D-- 312
D.YE.01.01YE386         -----------I-------------------------------------------------------L------V--------------------------A---...---A-------------------------V---------------------------------D-- 311
D.YE.02.02YE516         -----------I-----N--------------------------------------------G----L--AQ--------I--------V------E---R---...S-----------------------------V---------------------------------------- 312
D.ZA.90.R1              -----------I-------------------------------------------------------L---------------------------------A---...--------D-------------------V---------------------------------D-- 314
F1.AR.02.ARE933         -------------------------------------------------------D-----------L--A-----P------------------P----Q---...S---V-------------------------V-----K--------------F-A------T-D-- 310
F1.BE.93.VI850          -----------I----------------------T--------------------D-----------L-----AP--------------------------Q---...G---V---D-------------------V-----K--------------F-V--------D-- 310
F1.BR.01.01BR125        -----------I-------------------------------------------D-----------L-------P-------------------N-----Q---...S---V-------------------------V------------------------F--------T---- 312
F1.ES.x.P1146           -------------------------------------------------------D-----------L-------P-------------------------Q---...S---V-------------------------V------------------------F--------T---- 310
F1.FI.93.FIN9363        -----------I-------------------------------------------D-----------L-------P-------------------------Q---...S---V---D-------------------V------------------------F-A------T---- 312
F2.CM.02.02CM_0016BBY   -----------I-------------------------------------------D-----------L---Q---P---I---------------------A---...S---V-------------------------V------------------------F--------T---- 308
F2.CM.95.MP257          -----------I-------------------------------------------D-----------L-------P-------------------------A---...S---V-------------------------V-----K--------------F--------T---- 314
F2.CM.97.CM53657        -----------I-------------------------------------------D-----------L---Q---P---I---------------N-----A---...S---V-------------------------V-----K--------------F-------------- 308
G.BE.96.DRCBL           -------------------------T----------------------------------D------L--QQ--------I-D-T------A--------R---...S---V-------------------------V------------------------F--------T---- 307
G.CM.01.01CM_4049HAN    T---------------------------------S--------------------D-----------L--PQ----P---I-------------------R---...S-------D-------------------V-----K--------------F--------T---- 312
G.CU.99.Cu74            -------------------------------------------------------D-----------L--PQ----P---I------------------------...-----------------------------V-----K--------------F--------T---- 314
G.ES.99.X138            -------------X-----------------------------------------D-----------M--XX----P---I-------------------R---...S-----------------------------V-----------------X-F--------T-X-- 314
G.GH.03.03GH175G        -------------------------------------------------------D-----------L--QQ--------I-D-T---------------R---...S-----------------------------V------------------------F-A------T---- 313
G.KE.93.HH8793_12_1     -------------------------------------------------------D-----------I--PQ----P---I--------------N-----R---...S-----------------------------V------------------------F--------T---- 314
G.NG.01.01NGPL0669      -----------I-------------T-----------------------------D---D-------L--QQ----P---I-------------------T---...-----------------------------V-----K--------------F--------T-D-- 312
G.PT.x.PT2695           -------------------------------------------------------D-----------I--QQ----P---I-------------------R---...S-----------------------------V------------------------F--------T---- 313
G.SE.93.SE6165          T-------------------------------------L----------------D-----------M--QQ---FP---I--------------S-----T---...G-----------------------------V------------------------F-C--------D-- 314
H.BE.93.VI991           --------------------------------------A----------------D-----------L-------P-------------------------VA---...G-------D-------------------V-----K--------------FRV------T-D-- 316
H.BE.93.VI997           -X------------------------------------A----------------D-----------L-------P-------------------------A---...G--S----D-------------------V-----K--------------F--------T---- 314
H.CF.90.056             --------------------------------------A----------------D-------------------P-------------------------A---...G--A----D-------------------V-----K--------------F--------T-D-- 314
J.CD.97.J_97DC_KTB147   -----------I---G-N-------------------------------------D-----------L------V--------------------N-----A-I-GNGG-Q-------------------------V------------------------F--------T---- 305
J.SE.93.SE7887          -----------I-------------------------------I-----------D-------------------------V----------------------...G-----------------------------V------------------------F-A------T-D-- 314
K.CD.97.EQTB11C         -----------I-------------------------------------------D-----------M---Q---P---I-------------------T---...S-----------------------------V------------------------FRV------T---- 310
K.CM.96.MP535           -----------I-------------T-----------------------------D---D-------L-------P-------------------------A---...S---V-------------------------V------------------------F--------T---- 310
```

```
                                                                                     CyPA binding                                                           major homology region
B.FR.83.HXB2                SPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMT...NNPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVK 314
01_AE.CF.90.90CF11697       ---------------G-N--------------------M---I------------D-----------L--------P------------X----N---------...S-------D-------------------V--------------------F--------T---- 314
01_AE.CN.05.FJ051           ---------------G-N--------------------M---I---------------------------------P---------------------------...--------D-------------------V-----------------------------T---- 314
01_AE.CN.06.FJ054           ---------------G-N-------T------------M---I---------------------------IQ----P---I-------------------A-V-...----V---D-------------------VG--------------------------------- 314
01_AE.HK.x.HK001            ---------------G-N--------------A-----M---I---------------------------------P---------------------------...S--X----D-------------------V-----------------------------T---- 314
01_AE.TH.01.01TH_R2184      ---------------G-N--------------------M---I---------------------------------P---------------------------...--------D-------------------V-----------------------------T---- 312
01_AE.TH.02.OUR769I         ---------------G-N--------------------M---I----------------------------Q----P---I-------------N-----R---...--------D-------------------VG--------------------------------- 312
01_AE.TH.90.CM240           ---------------G-N--------------------M---I------------------P--------------P---------------------------...--------D-------------------V-----------------------------T---- 314
01_AE.US.00.00US_MSC1164    ---------------G-N--------------------M---I----------------------------Q----P---L-------------N---------...--------D-------------------V-----------------------------T---- 312
02_AG.CM.02.02CM_4082STN    -----------I-------------T------------M---I------------D--------------------P---------------------------...S---V----------V------------V--------------------F--------T-D-- 309
02_AG.EC.x.ECU41            -----------I--------------------------M---I------------D-----------A---Q----P---------------------------...S--------------V---------------------------------F--------T---- 277
02_AG.GH.03.GHNJ196         -----------I---G----------------A-----M---I------------D--------------------P---------------------------...S--------------V------------V--------------------F--------T-D-- 317
02_AG.NG.01.PL0710          -----------I-------------T------------M---I------------D-------------------AP---------------------------...S---V----------V------------V--------------------F--------T---- 307
02_AG.NG.x.IBNG             -----------I---G----------------------M---I------------D--------------------P---------------------------...S--------------V------------V--------------------F--------T---- 311
02_AG.SE.94.SE7812          --------------------------------------M---I------------D--------------------P---------------------------...S--------------V------------V--------------------F--------T-D-- 311
02_AG.SN.98.MP1211          -----------I--------------------------M---I------------D-------D---T--------P---------------------------...S-------D-------------------V--------------------F--------T---- 310
03_AB.RU.97.KAL153_2        -----------I--------------------------M---I------------D-----------L--AQ---FP--------------SS-----------...S-------D-------------------V--------------------F--------T-D-- 314
04_cpx.CY.94.CY032          -----------I--------------------------M---I------------D-------D---T--------P---------------------------...S---V-------------------T---I--------------------F-C------T---- 314
05_DF.BE.x.VI1310           -----------I-------------------------------------------------------L---Q---V------D-----------------A---...----------------------------V-----------------------------T---- 315
06_cpx.AU.96.BFP90          -----------I-D-----------T------------M---I------------D--------------------P---I-----------------------...S---------------------------VG----K--------------F--------T---- 313
08_BC.CN.97.97CNGX_6F       -------------------------T-----------------------------D-----------L-------V----------------------------...----------------------------------K--------------F--------T-D-- 309
09_cpx.GH.96.96GH2911       ----X------I-------------T------------M---I------------D-------------------NP------------------V----A---...G--------------V------------V--------------------F--------T-D-- 314
10_CD.TZ.96.96TZ_BF061      T------------------------------------------------------------------L---Q---V----I-------------------R---...S---------------------------V-------------------------------D-- 314
11_cpx.GR.x.GR17            -----------I--------------------------M---I------------D-------------------AP----------------GN---------...G----------R----------------V--------------------F------------- 312
12_BF.AR.99.ARMA159         -------------------------------------------------------D-------D---L--------P-----------------------Q---...S---V--------------------------------------------F--------T---- 310
13_cpx.CM.96.1849           -------------------------------------------------------D-----------M--QR---FP---I-------------------T---...----------------------------V--------------------F--------T-D-- 314
14_BG.DE.01.9196_01         -------------------------------------------------------D-----------M--QQ----P---I----------S--N-----R---...S---------------------------V-----K--------------F--------T---- 313
15_01B.TH.99.99TH_MU2079    ---------------G-N--------------------M---I--------------------K------------P----------------------V----...S--A----D-------------------V-----------------------------T---- 314
16_A2D.KR.97.97KR004        -------------------------------------------------------D-----------L--------P---------------------------...S---V-----------------------V--------------------F--------T---- 312
17_BF.PE.02.PE02_PCR0155    -----------I---G-N-------T-----------------------------D-----------L---Q----P-----------------N-----Q---...S---V-----------------------V-----K--------------F--------T---- 308
18_cpx.CU.99.CU76           ----------AI--------------------------M---I------------D-----------I---Q--------I-D-----------------R---...--------D-------------------V--------------------F--------T---- 314
19_cpx.CU.99.CU7            -----------I-----N-------------------------------------------------L--------------------------------A---...S---------------------------V-------------------------------D-- 314
20_BG.CU.99.Cu103           -------------------------------------------------------D-----------T--QQ----P---I----------A--------T---...----------------------------V-----K--------------F--------T---- 314
21_A2D.KE.91.KNH1254        -----------------N-------------------------------------------------L--I---------I-------------------A---...G---------------------------V---------------------------------- 313
22_01A1.CM.01.01CM_0001BBY  ---------------G-N-------------T--T---M---II-----------D-----------M---Q---YP-----------------------A---...S---V-----------------------V--------------------F--------T-D-- 312
23_BG.CU.03.CB118           -------------------------------------------------------D-----------L--QQ--------I-------------------R---...----------------------------V-----K--------------F--------T---- 314
24_BG.CU.03.CB378           -------------------------------------------------------D--------------PQ----P---I-------------------R---...S---------------------------V-----K--------------F--------T---- 314
24_BG.CU.03.CB471           -------------------------------------------------------D-----------L--PQ----P---I-------------N-----R---...S---------------------------A-----K--------------F--------T---- 314
25_cpx.CM.01.101BA          -------------------------------------------------------D-----------I---Q----P---I-------------N-----R---...S---------------------------V--------------------F--------T---- 313
27_cpx.CD.97.97CDKTB49      T------------------------------------------------------D-----------T--PQ---VP---I-D-----------------A---...----------------------------V--------------------F------------- 314
28_BF.BR.99.BREPM12609      -----------I-------------------------------------------D-----------L--------P-----------------N-----A---...S---V-----------------------V--------------------F-------CT---- 314
29_BF.BR.02.BREPM119        A-----------------------------------------S------------D--------------------P-----------------------Q---...S---V-----------------------V--------------------F--------T---- 302
31_BC.BR.02.110PA           -----------I-------------T-----------------------------D---------------Q----------------------------T---...G---V-----------------------V-----K--------------FR-------T-D-- 312
32_06A1.EE.01.EE0369        -----------I--------------------------M---I------------D-----------LY-TQ---FP---------------------------...S-------D-------------------V--------------------F--------THD-- 311
33_01B.MY.05.05MYKL007_1    ---------------G-N--------------------M---I--------------------D--KT--------P-----------------N---------...--------D-------------------V-----------------------------T---- 314
34_01B.TH.99.OUR2478P       ---------------G-N--------------------M---I--------------------D-------Q----P---------------------------...S-------D-------------------V-----------------------------T---- 312
35_AD.AF.05.05AF095         -----------I--------------------------M---I------------D-----------L---------------------------P----A---...G-------D-------------------VG----K--------------F-V------T---- 311
36_cpx.CM.00.00CMNYU830     ---------------G-N--------------------M---I------------D-------------------HP---I---------E---------A---..N-P-IS-----------------------V--------------------F--------T-D-- 313
37_cpx.CM.00.00CMNYU926     -----------I-------------T---D----H---A---V------------D-------D-----------HP---L------------------VA---...S-------D-------------------V--------------------F-V-----ST---- 310
38_BF1.UY.05.UY05_4752      -----------I-------------------------------------------D-----------I---Q----P---I-------------N-----Q---...S---V-----------------------V--------------------F-A------T---- 310
39_BF.BR.03.03BRRJ103       -----------I-------------------------------------------D-----------L------------------------------------...----------------------------V-------------------------------D-- 314
40_BF.BR.04.04BRRJ115       -----------I-------------------------------------------D-----------L--------P---I----------S--------Q-I-...G---V-----------------------V--------------------F------------- 316
42_BF.LU.03.luBF_05_03      ----------------------------A------------X-------------D-----------L--------P-----------------N-----Q---...S---V-----------------------V---------X----------F--------T---- 313
43_02G.SA.03.J11223         -----------I-------------T---------X-------------------D-----------L--PQ----P---I------------------XT---...S--------L------------------V-----K--------------F--------T-D-- 311
U.CD.83.83CD003             -----------I-------------------------------------------D-------D---L--------P---------------------------...S-------D-------------------V-----K--------------F-V-----S----- 314
U.CD.90.90CD121E12          -----------I-------------------------------------------D-------D---L--------P---------------------------...S-------D-------------------V--------------------F-------S----- 311
U.GR.99.GR303               -----------I-------------T---X--------M---II-----------D-----------L--TQ---FP---------------------------...S--------..-----------------V--------------------F----XK--T---- 312
U.NL.95.U_NL_95_H10986_D1   -------------------------------------------------------D--------------------P-----------------------A---...S---V-----------------------V-----K--------------FI---------D-- 311
CPZ.CD.90.ANT               T---------C----N-N----------------H-------A--D--G---V---V----------L--T----VQA--L---T----------V---MQ--STPQQ-GGV---D-------M----V--X---V---E-K-----------------I-------P-- 333
CPZ.CM.01.SIVcpzCAM13       -----------------N--------------L-L-V-----A-----G---V-R-V----------M------------L---------------------A-...A-------D--R--V-----RV-K--C-V-----K-------K---------M---------- 319
CPZ.CM.05.SIVcpzEK505       -----------------N-------M------------------------------V----------G----M---P---V---------------A--VA---...A---V---D--R---V------------A-----K-----T-----------------T---- 317
CPZ.CM.05.SIVcpzLB7         -------------------------A--A-----------------------I---V-----Q----L--I-M----------------------V--------...S---------------M----V------V-----K-----------------------T---- 314
CPZ.CM.05.SIVcpzMB66        -------------------------M------------------------------V----------L-------VP------------------I---V----...S-------D--------------K--C-V-----K-----S-----------------T---- 312
CPZ.CM.05.SIVcpzMT145       ----------A------N-------M-----------------I--------I---V--D-------T--T-----------------------S-E--V----...A-------D--R---V-------K----V-----K-------------------------D-- 304
CPZ.GA.88.GAB1              --------------------------------L---V-----A-----G---V---V----------L--T---------L---------------------T-...A-------DV-R--V------V----C-V---------------------------------- 322
CPZ.TZ.00.TAN1              ---------R-I---G-N----------------Y---S---A--E----------V----------A--A----QQA-ML---T-A------------VL---TPQAQGGV---D------------L------V-----K-----------------I-------P-- 340
CPZ.US.85.CPZUS             ----------A------N-------M----------V-----AI----G---V---A------D*--T--L----------------------------V----...A-------D--R--VV-----V-K--C-VG----K------------------------P--- 315
CPZ.CM.x.SIVcpzDP943        ---------------N-N-------M----------V-----AI----G---V---V-----L----T------------L--A------------A---A---...A---V------R--VV-----V------V-----K-----------------------T---- 311
CPZ.CM.x.SIVcpzMB897        -------------------------M-----------------I--------I---V----------T--I-----P------------------I---V----...S---T---D------V------------V-----K-----------------------T-D-- 314
N.CM.02.DJO0131             T----------I-----N-------M--------A---------------------V----------T--XPX--LP---L---------------X--VT---...S---V-------K--V----R-------V---E-K-------------------------D-- 319
N.CM.04.04CM_1015_04        T----------I-----N-------M--------S---S-----------------V----------T---PV--LP---L----X----------P---T---...S-------D--R---V----R-------V---E-K-----------------------T-D-- 316
N.CM.04.04CM_1131_03        T----------I-----N-------M--------S---S-----------------V----------T---PV--LP---L---------------P---T---...S-------D--R---V----R-------V---E-E----------------V------T-D-- 316
N.CM.95.YBF30               T----------I-------------M--------S---------------------V------D---T---PV--LP---L-D-------------A--VA---...A---V---D--R---V----R-------V---E-K-----------------------T---- 316
N.CM.97.YBF106              T----------I-----N--I----M--------S---S-----------------V----------T--APV--LP-----D-------------A--VA---...S-------D--R---V----R-------V---E-K-----------------------T-D-- 316
O.BE.87.ANT70               ----------A------N--I----M------ISY-I-----AI----G-L-V---V-----V----T--PPV--LP---I---T----------Q----H-T-..RP-Q-----D--RK--V-----M-K----V-----K-----------------------T---- 313
O.CM.91.MVP5180             ----------A------N--I----M------V-Y-I-----AI----G-L-V---V----------T--PAM--LP---I---T----------Q----I-T-..RGANS----D--RK--V-----M-K----V-----------------------------T---- 313
O.CM.96.96CMABB637          ----------A------N--I----M------I-Y-I-----AI----G-L-V---V------D---T--PPI--LP---I---T---------NQ---VH-I-..RA-A-----D--RK--V-----M-K----V-----K-----------------------T--X- 312
O.CM.98.98CMA104            ----------A------N--I----M---X---SY-X-----AI----G-L-V---V----------T--IAV--LP---I---T--------------VH-T-..RA-N-----D--RK--V-----M-K----V-----K---------------------------- 313
O.CM.99.99CMU4122           ----------A------N--I----M------I-Y-I-----AI----G-L-V---V-----S----T--PAM--LP-------T----------Q---VH-T-..RA-N-----X--RK--V--X--M-K----V---------------------------------- 313
O.FR.92.VAU                 ----------A--K---N--I----M--------Y-V-----AI----G-L-V-E-V--D---D--KT--PVV--LP---I---T---------------H-T-..RA-Q-----D--RK--VI----M-K----V-----K-R-E--------Q----------T---- 312
O.SN.99.SEMP1299            ----------A------N--I----M------I-Y-T-----AI----G-L-V---V----------T--PA---LPV--I---T----------Q---VH-I-..RP-Q-----D--RK--V-----V-K----V-----------------------------T---- 313
O.US.99.99USTWLA            ----------A------N--I----M------ISY-I-V---AI----G-L-V---V------D---T--PPI--LP---I---T----------Q----H-TS..RPGN----XD--RK--V-----M-K----V------X--------------------------- 310
```

```
                                p24 end p2 start      p2 end p7 start     Zn motif          Zn motif        p7 end p1 start   p1 end p6 start   Vpr binding
B.FR.83.HXB2             NWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNS.....ATIMMQRGNFR.NQRKIVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCT....ERQANFLGKIWPSY....KGRPGNFLQSR...PEPT........... 456
A1.GE.99.99GEMZ011       ----D-------------------SG-----------------------------LQ-A......N----KS---.GP-R.I----------L--------------------------....--------R----S....-------P---...---S........... 447
A1.KE.00.KER2008         G----------------S--R---TG----------------S------------AQQA......NV-------K.G-KR.I----------L-------------------------I....-------------S....-------P---...----........... 452
A1.KE.00.KNH1144         G----------------S--R---TG------------------------------QHT......N---------.G-KR.I----------L-----------------D--------....-------------S....-------P---...----........... 454
A1.KE.00.KSM4024         G----------------S--R---AG------------------------------QHT......N---------.G-KR.I----------L--------------------------....-------------S....-------P-N-...L---........... 451
A1.KE.00.MSA4069         G---D------------S--R----G-S--------------S------------AQ-T......KV---K----.G--R.I----------L--------------------------....-------------S....-------PLNS...L---........... 452
A1.KE.00.NKU3005         G-------I--------S--R---SG-----------------------------MQQT......N---------.G-KR.I----------L--------------------------....-------------S....-------P---...----........... 452
A1.RU.00.RU00051         --------------------R---AG----------------------------H-Q-......TN----K----GPPKR.I-X--------L--X-----------------------....-------------S....X---X--P---...----........... 455
A1.RW.93.93RW_024        G----------------S--R----G-S--------------S------------AQQT......N---------.G-K-.I----------L-------------------------S....--------R----S....-------P---...----........... 453
A1.SE.95.SE8891          G----------------S--R---SG------------------------------QQA......NV-------K.G-KR.I----------L--------------------------....-------------N....R------P---...----........... 452
A1.SE.95.UGSE8131        G---D------------S--RG--AG----------------S------------AQQT......N---------.G-K-.I----------L--------------------------....-------------H....-------P---...---S........... 454
A1.TZ.01.A173            G----------------S------SG-----------------------------AQQT......N---------.G-KR.I----------L--------------------------....-------------S....-------P---...----........... 452
A1.UA.01.01UADN139       --------------------R----G-----------------------------AQGA......N----KS---.GPKR.I----------L--------------------------....--------R----S....-------P---...---S........... 452
A1.UG.92.92UG037         G-------I--------S--R---AG------------------------------QHT......N--------K.G-KR.I----------L-K--------------R---------....-------------S....-------P---...----........... 454
A1.UG.99.99UGA07072      G---D------------S------SG----------------S-------------QQA......N--V------.G-KR.I----------L--------------------------....-------------S....-------P---...----........... 452
A2.CD.97.97CDKTB48       ----D------------S--R----G------------------------------Q-T......N--I-----K.G-KR.I----------L--------------------------....-------------N....-------P---...T---........... 460
A2.CY.94.94CY017_41      ----D-----------RS--R----G-S--------------S-----------H-QST....NTN---------.G-KR.I----------L--------------------------....-------------N....-------P---...T---........... 452
A.SN.01.DDI579           -----------------S--R---TG------------------------------QHT......NV--------.G-KR.I----------L--------------------------....-------------N....-------P---...----........... 452
A.SN.01.DDJ369           -----------------S--R---SG---------------------------T--QHA......N----K---K.G-KR.I----------L--------------------------....-------------S....-------P---...----........... 453
A.SN.96.DDJ360           -----------------S--R---TG------------------------------NHA......N---------.G-KR.I----------L--------------------------....-------------S....-------P---...----........... 452
A.CD.02.02CD_KTB035      ----D------------S--RG--AG------------------------------Q-T......N----K----.GP-R.I----------L--------------------------...N---------V---N....R------P-K-...L---........... 455
A.CD.97.97CD_KCC2        ----D---------------R---AG----------------AIS.---V-E-RRMM-T.....TIM-LKG----.G-KR.I----------L--------------------------....-------------N....-------P-KG...L---........... 454
A.CD.97.97CD_KTB13       ----D------------A------TG----------------S------------MQ-P......N----K----.GP--.I----------L--------------------------....---------V---N....-------P-N-...L---........... 452
B.AR.04.04AR151516       -------------------------G----------------S----I----------P.....-N---------.----T-----------I----K---------------------....--------RL---H....-----------...----.........AP 456
B.AU.87.MBC925           -----------------------------------------------------------.....-----------.----T-----------I-K------------------------....-------------H....-----------...----........... 457
B.BO.99.BOL0122          -----------S-------------G-------------------------------G-.....--V-------K.YTK-T-----------I-----------------------E--....-------------H....-----------...----........... 458
B.BR.03.BREPM2012        ------------------------------------------S----I-------M---.....-N---------.P---------------I-K------------------------...F-------------H....---------N-...---R..........P 463
B.CA.97.CANB3FULL        ----------------------------------------------------------P.....NAV--------.----P-----------I-K------------------------....--------R----H....-----------...----........... 456
B.CN.05.05CNHB_hp3       -----------S----R--------G----------------S----I-----------.....-N---------.----------------I-K-----------R---------E--....-------------H....-----------...---S........... 457
B.CO.01.PCM001           -------------------------G---------------------------------.....TA-------YK.----P-----------I-K-----------------------I....-------------N....-------P-N-...L---........... 454
B.GB.83.CAM1             -----------------------------------------------------------.....-----------.----T--------V--I-K-----------------------N....-------------H....-----------...----........... 456
B.GB.86.GB8              --------------------------------------------------------NS.......------K----.S---T-----------I-----------------------E--....-------------H....-----------...---I........APP 460
B.GE.03.03GEMZ010        -------------------------------------------------------ISQP.....P----------.----T-----------I-K------R-----------------....-------------H....-----------...----........... 454
B.IT.05.SG1              ----------Y--------X-----G----------------S-----------L---P.....T---------X.---X------------I--------------------------....-------------S....----X------...----........... 457
B.JP.05.DR6538           -----------------------------------------------------------.....-S--V------.-----I----------ISK------------------------....-------------H....-----------...----........... 456
B.KR.05.05CSR3           -----------------------------------------------------------.....----------K.G---M-----------I-K------------------------....-------------H....-------P---...L---........... 456
B.NL.00.671_00T36        -----------S---------------------------------------------SA.....PA-------H-.---RT-----------I-------X-------X-X--------....X---X--------H....-----------...----....APSQSRP 466
B.RU.04.04RU128005       ----------------------------------------------------------Q.....-X--------K.X-KRS-------R-X-I---------X--X--X----------....K------X-----H....-----------...----........... 456
B.TH.00.00TH_C3198       --------------------R--------------------------------------....A--V--------.----T-----------I-K-------R----------------....------------AH....-----------...----........... 457
B.UA.01.01UAKV167        -----------------------------------------------I--------AQP.....-----------.----T-----------I-K--------------R---------....-------------H....-----------...-G--.......APRT 458
B.US.04.ES10_53          ------------------------------------------------------H-NS......-A----K----.----T-----------L----------------R---------....-------------H....---------N-...T--S........... 459
B.US.99.PRB959_03        ---------------------------S-------------------------------.....N----------.----------------I-----------------------E--....-------------H....-----------...----........... 459
C.AR.01.ARG4006          --------------------R----G-S----------------------------N-T......N--------K.GPKRTI------R---L--------------------------....-------------N....------X----...----........... 451
C.BR.04.04BR013          ----D---------------R----G-S----------------------------N-T......N--------K.GPKRPI----------L-K------------------------....---V---------H...KG----------...----.......APPL 458
C.BW.00.00BW07621        ----D---------------R----G-S--------------S------------AN-T......N--L-KN--K.GP-R------------I--------------------------....-------------H....-----------...----........... 452
C.CN.98.YNRL9840         ----D---------------R----G-S--------------S------------TNST........----S--K.GPKR------------I--------------------------....-------------H....-----------...----........... 451
C.ET.02.02ET_288         --------------------R----G-S---------------------------ANQA......N----K---K.GP-RTI----------L--------------------------....---------L---H....---------N-...-A--........APL 454
C.GE.03.03GEMZ033        ----D---------------R----G------------------------------NST......N--------K.SPKR------------I-----------------------E--....-------------H....-----------...----........... 450
C.IL.99.99ET7            ----D---------------R----G---------------------------T-AN-T......H------N-K.GPKR------------L--------------------------....---------F---H....-----------...----........... 450
C.IN.99.01IN565_10       ----D---A-----------R----G-S--------------S------------ANS-.......-L---S--K.GPKR-I----------I-K---------------------E--....-------------H....---------N-...----........... 460
C.KE.00.KER2010          ----D---------------R----G--------S-------S------------AN-A......N--I--S--K.SPKR------------I--------------------------....-------------H....---------N-...----........... 450
C.MM.99.mIDU101_3        --------------------R--------------R------S------------TNSA.......-L---S--K.GPKRT-----------I----------------Q---------....D--------L---H....-----------...----........... 450
C.MW.93.93MW_965         ----D------------N--R---SG----------------S-----------HAG-A......N-----S--K.GP-RT-----------L--------------------------....-------------.....-----------...----........... 450
C.SN.90.90SE_364         ----D---------------R---AG----------------S------------AN-P......H-L---N--K.GP-R------------I-K------------------------....------------AH....-----------...----........... 450
C.SO.89.89SM_145         ----D---------------R----G-S---------------------------ANAN....NIN-L------K.GPKR------------I--------------------------....-------------H....---------N-...----APPAESFRVRP 463
C.TZ.02.CO178            ----D---------------R----G-----------------------------ANSK......N-----S--K.GPKRT-----------I--------------------------....-------------H....-----------...----........... 450
C.UY.01.TRA3011          ----D---------------R----GS-----------------------------N-T......N-----S--K.GPKRTI----------L--------------------------....-------------H....-------P---...L---........... 447
C.YE.02.02YE511          ----D---------------R----G-S----------------------------N-P......N-----N--K.GPKR------------I-K-----------------------N....-------------H....-----------...----........... 450
C.ZA.04.04ZASK164B1      ----D---------------R----G----------------S------------AN-T......N-----S--K.GP---I----------I-K------------------------....-------------H....-----------...---S........... 453
C.ZA.05.05ZAPSK240B1     ----D---------------R----G------------------------------N-G......N----KS--K.GPKR.I----------L----------------Q---------....-------------H....---------N-...----........... 450
C.ZM.02.02ZM108          ----D------------S--R----G-S--------------S------------AQ-T......N-----N--K.G--R------------I--------------------------...N-------------H....-----------...----........... 455
D.CD.83.ELI              -------------------------Q----------------S------------A-------V T-A-------K.GP---I----------I-K-----------R-------L----....--------R----H....-----------...----........... 457
D.CM.01.01CM_0009BBY     -------------------------G----------------S------------A-GA....S-A----K---K.GP-RT-----------I-K------------------------....-------------H....-----------...----........... 451
D.KE.01.NKU3006          ------------------------------------------S------------A--T....ANA------R-K.GT---I----------I-K-------R----------------....-------------H....---------N-...----........... 455
D.KR.04.04KBH8           ------------------------AG-S--------------S------------A-S-....N-A--------K.GP--------------I-K-------R----------------....-------------H....N------P---...T---........... 456
D.TD.99.MN011            -------------------------G----------------S----I-------A-S-....A-A--------K.-PKR------------I-K---------------------E--....-------------S....-----------...----........... 457
D.TZ.01.A280             ------------------------------------------S------------A--V....N-A--------K.GP---I----------I-K------------------------....-------------H....-----------...----........... 451
D.UG.99.99UGK09259       ---------------------------------------------------------TG....N-GV-------K.G--R-I----------L----K-----------R---------....-------------H....-----------...----........... 455
D.YE.01.01YE386          ------------------------------------------S------------A-KG.....NAV-------K.GP--TI------------K-------R----------------....-------------H....-----------...----........... 453
D.YE.02.02YE516          -------------------------G-S--------------S-----------IA-SN....N-A--------K.-P--------------I-K------------------------....--------R----S....-----------...----........... 455
D.ZA.90.R1               -------------------------Q-----------------------------A--L....ATAV-------K.G----I----------I-K--------------R---------....-------------H....-----------...----........... 457
F1.AR.02.ARE933          ----D----------------G--SG----------------S------------A-ST......-V---KS--K.G-KR---------V--I-K--------------R---------....-------------N....-----------...----........... 451
F1.BE.93.VI850           G---D-------------------TG----------------S-------------ANS......A----KS--K.G--RV-----------I----------------R---------....-------------N....-----------...----........... 450
F1.BR.01.01BR125         G---D--------------------G-----------------------------A-HA......-V---KS--K.G--RSI----------L----K-----------R---------....-------------H....---------N-...----........... 453
F1.ES.x.P1146            G---D---I----------------G-----------------------------KA-G.....TA----KS--K.G--R------------I-K--------------R---------....-------------Q....---------N-...----........... 451
F1.FI.93.FIN9363         G---D------------I---G--IG----------R----------I-------ANT.......----KS---.G--R------------I----------------Q---------....-------------N....-----------...----........... 452
F2.CM.02.02CM_0016BBY    G---D--------------------G---------------------I---------AT......SVL--KS--K.G-KR------------I-K-------R----------------....-------------N....-------I---...----........... 449
F2.CM.95.MP257           G------------------------G----------------S----I------KA-GA......A----KS--K.G--R------------I---------R------Q---------....---------M---N....---------N-...----........... 455
F2.CM.97.CM53657         G---D------------I---G--TG---------------------I---------ST......S-L--KS--K.G--RN-----------I-K-------R----------------....-------------N....---------N-...----........... 449
G.BE.96.DRCBL            S---D---I--------I---G--QG----------------S------------ASGA....A-A----KS--K.GP-RTI----------L-----------------------E--....-------------N....---------N-...----........... 450
G.CM.01.01CM_4049HAN     G---D--------------------G----------------S------------ASAT.....TA----KS--K.GP-RTI----------L--------------------------....--------R----N....---------NS...----........... 454
G.CU.99.Cu74             G---D------------A--R----G----------------S------------ASGA....G-AVL---S--K.GP-RNI----------L--------------------------....---------L---N....---------N-...---S........... 457
G.ES.99.X138             G-------------------R-X--G----------------S------------A-GA......-X---KS--K.GP-RTI----------L-K------------------------....-------------N....---------N-...----........... 455
G.GH.03.03GH175G         H-------I------------S--TG------------------------------SGA.....TS----NS---.GPKRNI----------L-----------------------E--....-------------N....---------N-...---S........... 455
G.KE.93.HH8793_12_1      G---D---------------R----G-----------------------------A-GA....A-A----KS--K.GPKRNI----------L--------------------------....-----#-------N....---------N-...----........... 456
G.NG.01.01NGPL0669       ----D---------------R----G------------------------------SGA....A-A----KS--K.GP-RNI----------L----------------R---------....-------------N....---------N-...L---........... 455
G.PT.x.PT2695            G---D---------------R----G-S--------------S------------ASG.......------KS--K.GPKRMI----------L--------------------------....-------------N....---------N-...----........... 454
G.SE.93.SE6165           G---D---------------R---QG-S--------------S------------ASGA....A-A-----S--K.GP-RTI----------L--------------------------....-------------N....---------N-...T---........... 457
H.BE.93.VI991            ----D-----------R-------RG-SI-------------S---------------A....S-A----K---K.GP-RT---S-------I----------------Q---------....G------------S....-------P-K-...L---........... 459
H.BE.93.VI997            ----D---------------R---QG-SI-------------S---------------A....N-A----KS--K.GP--------------I----------------R---------....-------------S....-----------...----........... 457
H.CF.90.056              --------------------R---QG-SI-------------S---------------T....NTA----K---K.G---F-----------I----------------R---------....-------------S....-----------...----........... 457
J.CD.97.J_97DC_KTB147    ----D---------------R---AG-------------RVSS-R---FG--I---N-......TN-I-------.G-KRS-----------I-K----------*--------I-N--....---------V---S....-----------...----........... 445
J.SE.93.SE7887           ----D-------------------SG--------------------------------T......N---------.DHKR---------Q--I-K------------------------....-------------S....-----------...----........... 455
K.CD.97.EQTB11C          ----------------R-------SG---------------------------------......AV-------K.G--R-I----------L-------------------------S....---------F--LN....-E-------N-...----........... 451
K.CM.96.MP535            ----D--------------------G-S--------------S----I----------P......VV---K---K.GH--------------I--------------------------....-------------H....-----------...----........... 451
```

```
                                                                       p24 end p2 start      p2 end p7 start    Zn motif          Zn motif          p7 end p1 start     p1 end p6 start  Vpr binding
B.FR.83.HXB2                NWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNS.....ATIMMQRGNFR.NQRKIVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCT....ERQANFLGKIWPSY....KGRPGNFLQSR...PEPT........... 456
01_AE.CF.90.90CF11697       -----------------S------TG-------------------------------QHA......N--------K.G-KR.I----------L---------R------Q---------....------------LN....-------P---...L---........... 454
01_AE.CN.05.FJ051           -----------------ST-----IG----------------S------------AQHT......N--------K.G-KR.I----------L--------------------------....-------------N....-------P---...----........... 454
01_AE.CN.06.FJ054           -----------------S------TG-S--------------S-----------S-QHT......N--------K.G-KR.I----------L---------E-------------------N-------------N....-------P---...----........... 455
01_AE.HK.x.HK001            -----------------S------AG----------------S----------N-AQQT......-V------XK.G-KR.I------R---L--------------------------....------------PN....-------P---...---S........... 454
01_AE.TH.01.01TH_R2184      -----------------S------A-----------------S------------AQHA......N-L------K.GRKT.I------R---L--------------------------....---------L---N....-------P-N-...----........... 452
01_AE.TH.02.OUR769I         -----------------S------TG----------------S-----------HAQQA......N--------K.G-KR.I----------L--------------------------....-------------N....-------P---...----........... 452
01_AE.TH.90.CM240           -----------------S------TG----------------S-----------HAQHA......---------K.G-KR.I------R---L---------Q----------------....-------------N....-------P---...----........... 454
01_AE.US.00.00US_MSC1164    -----------------S------TG-S--------------S-----------H-QHA......N--------K.G-KR.I------R---L--------------------------..ERQG-----------H....-------P---...----........... 454
02_AG.CM.02.02CM_4082STN    --------------------RG---G-----------------------------AQQA......N---------.G--T.I----------L----K----R----------------....-------------N....-------P---...----........... 449
02_AG.EC.x.ECU41            --------------------R----G-----------------------------AQQA......N---------.GP-T.I----------L-K--K----R----------------....-------------S....-------P---...----........... 417
02_AG.GH.03.GHNJ196         --------------------R----G-S--------------S------------AQQ-......NV--------.G--T.I----------L----K----R----------------....-------------N....-------P---...---S........... 457
02_AG.NG.01.PL0710          G----------------S--R----G------------------------------QQT......NV--------.S--T.I----------L----K---------------------....-------------S....-------P---...----........... 447
02_AG.NG.x.IBNG             -----------------S--R---TG------------------------------QQA......NV--------.G--T.I----------L----K----R----------------....-------------S....-------P---...----........... 451
02_AG.SE.94.SE7812          -----------------S--R----G----------------S-------------QQP......N---------.G--P.I----------L----K---------------------....-------------S....-------P---...----........... 451
02_AG.SN.98.MP1211          -----------S----R--------G-------S---------------------AQQ-......N---------.G--T.I----------L----K----R----------------....-------------S....-------P---...----........... 450
03_AB.RU.97.KAL153_2        --------------------R---SG--------V--------------------Q-A......N----KS---.GPKR.I----------L--------------------------....--------R----S....-------P---...---S........... 454
04_cpx.CY.94.CY032          -----------------S------TG----------------S------------AS-AAA....A----KSK-K.G--RTI----------L--------------------------....--------RM---S....---------N-...----........... 457
05_DF.BE.x.VI1310           ------------------------QQ----------------S------------A---AA....-A-------K.GP---I----------I-K-----G--------R---------....---------V---H....-----------...---S........... 458
06_cpx.AU.96.BFP90          ----D---------------R----G-----------------------------ASVG......A----KS--K.GPKRSI----------L--------------------------....-------------N....---------N-...----........... 454
08_BC.CN.97.97CNGX_6F       ----D---------------R---SG-S--------------S------------TN-T.......-L---S--K.GSKR------------I-K------------------------....---------F---H....-----------...----........... 449
09_cpx.GH.96.96GH2911       -*------------------R----G------------------------X------Q-T......N-L------K.G-KR.I---------XX--------------------------....-------------N....-------X---...---S........... 453
10_CD.TZ.96.96TZ_BF061      ------------------------------------------S------------A-GG.....N---------K.GPK-SI------------K------RR------R----L---P...ER------------N....-----------...----........... 457
11_cpx.GR.x.GR17            -----------------S--R---VG----------------S------------AQHT......N--------K.G-KR.I----------L------------*-------------....-------------S....-----------...----........... 451
12_BF.AR.99.ARMA159         ----D-------------------SG-----------------------------A-ST......AVL--KSS-N.G--R---------T--I-K--K--------------------P....-------------N....---------N-...----........... 451
13_cpx.CM.96.1849           ----D---------------R----G-----------------------------ASSAAA...-A----KS--K.GPKR-I------R---L-------------------------M....-------------N....-----------...-G--....APPESRP 465
14_BG.DE.01.9196_01         S---D---------------R----G----------------S-X----------ASGA......---L-KS--K.GP-RQI----------L--------------------------....-------------N....---------N-...----........... 454
15_01B.TH.99.99TH_MU2079    --------A--------S------SG-S--------------S-----------H-QQT......NV-------K.G-KR.I----------L----------------E---------....--------RL---#....f------P---...---S........... 452
16_A2D.KR.97.97KR004        -----------------S--R----G-S--------------S-------------Q-T...NSNIM---K----.G-K#.I----------L--------------------------....-------------H....S------P---...T---........... 454
17_BF.PE.02.PE02_PCR0155    G-------------------R---SG-------------------------------GT......-V---KS--K.G--R--------R---I-K--------------R---------....---------V---N....R--------N-...T---........... 449
18_cpx.CU.99.CU76           -----------------N--R----G----------------S-----------H-Q-T......N--------K.G-KR.I----------I--------------------------....-------------N....---------K-...----........... 454
19_cpx.CU.99.CU7            -------------------------G----------------S------------A---....H------K---K.GL--V-----------I--------------------------....-------------H....-----------...----........... 457
20_BG.CU.99.Cu103           G---D---------------R----G------------------------------SGA....G-S-L--KS--K.GPKRNI----------L--------------------------....---------L---H....-----------...----........... 457
21_A2D.KE.91.KNH1254        ----D------------S-------G----------------S------------A-S-....A-AV-------K.GP---I----------I-K--------------R---------....---------V---H....-----------...----........... 456
22_01A1.CM.01.01CM_0001BBY  -----------------S--R---TG------------------------------QHT......G--------K.G-KR.I----------LT------R------------------L.TL-------------S....-------P-G-...----........... 455
23_BG.CU.03.CB118           G---D---------------R----G-S---------------------------IPGA....G-S-----S--K.GPKRNI-------D--L-----------------D--------...EG--------L---H....-------P-N-...L---........... 458
24_BG.CU.03.CB378           G---D---I-----------R----G------------------------------SGT....G-S----KS---.GP-RNI----------L--------------------------....---------L---H....-------P-N-...L---........... 457
24_BG.CU.03.CB471           G---D---------------R----G------------------------------SGT....G-S----KS--K.GTKRNI----------L----------------R---------....---------F---H....-------P-N-...L---........... 457
25_cpx.CM.01.101BA          G-------------------R----G-S----------------------------SGT....ATA--------K.GPKRNI----------L--------------------------....-------------S....---------N-...----........... 456
27_cpx.CD.97.97CDKTB49      ----D-------------------SG----------------S------------ASGAA....-A----KS--K.GP--TI----------L---------R----------------....-------------N....-----------...---S........... 457
28_BF.BR.99.BREPM12609      G---D--------------------G----------------S------------A--T......-V-------K.G-----------I----------------R---------....-------------H....---------N-...----........... 455
29_BF.BR.02.BREPM119        ----D--------------------G----------------S-----------HA--T......A-L--KS--K.G--RV-----------I-K--------------R---------....-------------H....---------N-...----........... 443
31_BC.BR.02.110PA           ----D---------------R----G-S----------------------------N-T......N-----S--K.GPKRPI----------I-----------------------E--....-------------H....----X------...----........... 453
32_06A1.EE.01.EE0369        --------------------R----G-----------------------------AQ-A......N----KS---.GPKR.I----------L----------------Q---------....-------------S....---------N-...----........... 451
33_01B.MY.05.05MYKL007_1    -----------------S------TG----------------S-----------HAQHT......G--------K.G-KR.I----------L--------------------------....---------L---H....-----------...----........... 454
34_01B.TH.99.OUR2478P       -----------------S------TG----------------S------------TQQA.....NMM-------K.G-KR.I----------L--------------------------....-K-I-----F---N....-------P---...----........... 453
35_AD.AF.05.05AF095         G----------------S--R---TG----------------S------------AQH......TN----K---K.GHKR.I-------D--L----------------R---------....-------------H....-------P---...----........... 451
36_cpx.CM.00.00CMNYU830     -----------------S------TG------------------------------QH......-G--I--S---.G--T.I----------L--------------------------...N-------------S....-------P---...----........... 454
37_cpx.CM.00.00CMNYU926     -----------------S------TG-----------------------------T-AQQ......TS-----S--K.GPKRNI----------L--------------------------....--------R----S....-------P---...----........... 451
38_BF1.UY.05.UY05_4752      ----D----------------G--IG---------------------------------P......IV---KN--K.G--RSI----------L----------------R---------....-------------N....---------N-...----........... 451
39_BF.BR.03.03BRRJ103       ------------------------------------------S------------A--Q.....TNV--------.----T-----------V-----------------------E--....-------------H....---------T-...----........... 456
40_BF.BR.04.04BRRJ115       ----D--------------------G-----------------------------I--TA...NTAV---KS--K.G--R-------------I----------------R---------....-------------H....---------N-...L---........... 460
42_BF.LU.03.luBF_05_03      G---D--------------------G-----------------------------A-G......TAVL---S--K.GP--------------I-K--------------R---------....--------------....---------N-...---R..........P 455
43_02G.SA.03.J11223         ----D------------N--R---SG----------------S------------ASGA....A-A----K---K.GP-RNI----------L--------------------------....-------------H....-----------...----........... 454
U.CD.83.83CD003             ----D---I---------------TG----------------S------------TN.......TA--------K.GP-R-------------I-K--------------R---------....-------------N....---------N-...----........... 454
U.CD.90.90CD121E12          ----D---T---------------TG-----------------------------M--......T-V-------K.G-----------I-K--------------R---------....-------------N....---------N-...T---........... 452
U.GR.99.GR303               ------------------------SG------------------------------QS......PN----K----.G-KR.I----------L---------X---------------P....--------R----S....-------P---...----........... 452
U.NL.95.U_NL_95_H10986_D1   ----D----------------------S-------------------I-------A-K......TVV-------K.GTKR-I----------I--------------------------....--------RV---N....-----------...----........... 452
CPZ.CD.90.ANT               A-------I--------H------TG-S----L---------A----------ASAN-AQ...GTAVFL----GNRGGKRPL---------------K---R----R--Q----L-N-PA.TNTGKV-----PT-TW...WGC-----V-KE.EVV---........... 484
CPZ.CM.01.SIVcpzCAM13       T---D------------Q-------G-S--------------A----------TM-QSQART...D-FF-K-PGAT.P-RKI--Y-------L----K---R-------Q--------S....G--V-----G---R....S------V-N-...T---........... 463
CPZ.CM.05.SIVcpzEK505       -----------------N--R----G-S--------------A----------T-AQTA.....TSVF------K.GI--TI----------L----K-----------Q--------R..SG---------V--LS....-------P-TT.TRK---........... 463
CPZ.CM.05.SIVcpzLB7         ----D------------S--R----G-S--------------A------------ANAA.....TV--------K.GPKRVI----------I----K---R-------Q---R-Q--N....MGKV-----L---S....Q------P---...----........... 456
CPZ.CM.05.SIVcpzMB66        ----D------------S--R----G----GDEP-F------S------------AQH-........NDAKRQ-K.GPKR------------I----K---R-------Q-----RN--...N---------L---S....---------K-...----........... 452
CPZ.CM.05.SIVcpzMT145       ----D---I--------N--R----G----------------S-------------NKTE......-F----GQNRPPN-KI----------L--------R-------Q--------N..QNNGRI-----N--PW....G-----YV-G-...K---........... 448
CPZ.GA.88.GAB1              ----D------------Q-------G----------------S----------.SMVQNQ..GR-DVFF-K-QGA.GPKRKI----------L----K---R----R--Q-------------G--V-----G---R....S------V-N-...T---........... 466
CPZ.TZ.00.TAN1              T----------------H------QG------L---------S---KI-----ASA-AGG.....VN-L-G-KRPPLKKGQLQ------V----------------R--Q---------TRNNSTGV-----RT-LW....GC-----V-N....................... 482
CPZ.US.85.CPZUS             T-------------------R----G------V---------A----------C-MK-P.....SSVFL-K--AG.KPGRKI----------L--------R-------Q---------A..GN--------H-SPSWSGGSK------EN-...K---........... 463
CPZ.CM.x.SIVcpzDP943        T---D--X---------N--R----G----------------A-------------Q--.....TSVFL----NG.RPTRKI----------L--------R-------Q----I----A..GD--------G-SPW.SGGSK-------G-...K---........... 458
CPZ.CM.x.SIVcpzMB897        -----------------Q--R----G----------------S------------ANT.......-V------VK.I-KR------------I--------R-------Q----Q-N-P...DMQR---------PR....---------K-.PE-PEP........... 457
N.CM.02.DJO0131             -----------------Q------------------------A-------------QTA.....TSVFA-GR--K.GI-RTIR---------L----K---RG------Q--------K..NEG--------G-SPF....-------P-TTTTRK---........... 466
N.CM.04.04CM_1015_04        ----------------XQ-------S-S-------------------------A-AQAA.....TSVFV-----K.GT-RTI----------L----K---RX------Q--------K..NEG--------G-SPF....-------P-TA.TRK---........... 462
N.CM.04.04CM_1131_03        -----------------Q-------GSS------------------------TA-AQAA.....TSVFV-----K.GT-RTI----------L----K---RG------Q--------K..NEG--------G-SPF....-------P-TT.TRK---........... 462
N.CM.95.YBF30               -----------------QL------G----------------A-------------QQP....TTSVFA-----K.GI--PI----------L----K---RG------Q--------K..NEG--------S-SPF....-------P-TT.TRK---........... 463
N.CM.97.YBF106              -----------------Q-------G----------------A----------A-AQTA.....TSVFV-----K.GI--TI----------L----K---RR------Q--------K..NEGX-------G-SPF....-------P-TT.TRR---........... 462
O.BE.87.ANT70               -----------------Q---S---G-------V--------T----------ATAQQDLKGGYTAVF----QNP.IRKGTI----------I----------------Q--------R...NGK-------Y--PG....GT----YV-RP...AH-S........... 461
O.CM.91.MVP5180             -----------S-----Q-------E-------V--------T---KI-----ASAQQDLKGGYTAVF----QNP.-RKGPI----------I-K-------R------Q--------K...NG--------Y--PG....GT----YV-KQ...VS-S........... 461
O.CM.96.96CMABB637          ---------------X-Q---S---G-------V--------T----I--X--AS-NQDLKGGYTAVF----QNP.TRQRPI----------L----------------Q--------R...NGK------RY--PG....GT----YV-KQ...VP-S........... 460
O.CM.98.98CMA104            -----------------Q-------G-------V--------T----------ASAQQDLKGGYTAVF----QNP.IR-GP-----------I-------------N--QD-------R...NGK-------Y--PG....GT-----M-KQ...VA-S........... 461
O.CM.99.99CMU4122           -----------------Q---S---G-------V--------T----------ATAQQDLKGGYTAVF----QNP.GRKGPI----------L----------------Q--------R...NGK-------Y--PG....GT----YA-RQ...VS-S........... 461
O.FR.92.VAU                 -----------------Q-------G-S-----V--------T----------ASAQQDLKGGYTSVF----QKP.IRKGPI----------VTK--------------Q--------KN...R--T---K-Y--PR....GT----YV-EQ...VS-S........... 460
O.SN.99.SEMP1299            -----------------Q---S---G-------I--------T----------.AAAQDLKGGYTAVF----QNP.SRKGPI----------L----------------Q--------K...NG--------Y--PG....GT----YA-RQ...VS-S........... 460
O.US.99.99USTWLA            -----------S-----Q-------G-------V--X-----T---KI-----ASAQQDLKGGXTAVF----QNL.GRKGP-----------V----------------Q--------R...NG--------Y--PG....GE-----M-KQMPP-A-S........... 461
```


HIV-1/SIVcpz Proteins

HIV-1/SIVcpz Proteins — Gag

```
                                Vpr binding                                  Vpr binding
                                                                             p6 end, Gag end
B.FR.83.HXB2                    ...APPEESFRS.....GVETTTPPQKQEPIDKE......LYPLTSLRSLFGNDPSSQ*   500
A1.GE.99.99GEMZ011              ...---A-H-GM.....E-.I--SL---QK-R-.....QHP-SI--K-------L---   492
A1.KE.00.KER2008                ...---A-I-GM.....E-.I-S-P---QKER-.....QTP-FV--K-------L---   497
A1.KE.00.KNH1144                ...---A-L-GM.....E-.IAS-L---QKG--.....QTP-SI--K-------L---   499
A1.KE.00.KSM4024                ...---A-T-GM.....E-.MAS-L-K-QK-RD.....QAT--V--K-------L---   496
A1.KE.00.MSA4069                ...---A-IYGM.....EG.IAS-----QR--.....QEA-SI--K------LL---   497
A1.KE.00.NKU3005                ...---A-LLGI.....KE-.IAS-P---QN-RG.....QNP-SV--K------LL---   497
A1.RU.00.RU00051                ...---A-N-GM.....E-.IA-SL---QK-R-.....QPP-SI--K-------L---   500
A1.RW.93.93RW_024               ...---A-LLGM.....E-.IA--P---QK-R-.....QGP--V--K-------L---   498
A1.SE.95.SE8891                 ...---A-I-GM.....E-.I-S-P---KK-R-....SEAP--I--K-------L---   498
A1.SE.95.UGSE8131               ...---A-..MM.....E-.IAS-P---QNNPP........SV--K------LL---   493
A1.TZ.01.A173                   ...---A-L-GM.....EG.IASLP---QK-R-.....QVP--V--K-------L---   497
A1.UA.01.01UADN139              ...---A-N-GM.....E-.I--SL---QK-R-.....QQP-SI--K-------L---   497
A1.UG.92.92UG037                ..AP-AA-I-GM.....RE-.IVS-P---QN-RD.....QNP-SV--K------LL---   500
A1.UG.99.99UGA07072             ...---A-L-GM.....E-.M-S-P---QK-R-.....QVP--V--K-------L---   497
A2.CD.97.97CDKTB48              ...---.....M.....EE-.I-SSL----..NR-.....PST-AI--K------LL---   498
A2.CY.94.94CY017_41             ...---A-NL-M.....E-.I-SSL---LETR-.....PYN-AI--K-------LL--   497
A.SN.01.DDI579                  ...---ADLLGM.....E-.I-SSP---QK-R-......PP--V----------L---   496
A.SN.01.DDJ369                  ...---A-LLGR.....E-.I--SP---QK-R-......QP--I--K-------L---   497
A.SN.96.DDJ360                  ...---ADLLGM.....E-.IPSSP---QK-R-......PP-SV----------L---   496
A.CD.02.02CD_KTB035             ...---A--LGM.....E-.I-F-----MG-R-.....QSS--V--K------LL---   500
A.CD.97.97CD_KCC2               ...---A--LGM.....G-.I-SSL---IK-RD.....QQS--I--K-------*---   498
A.CD.97.97CD_KTB13              ...---A--LGM.....GK.--SSL---LR-RD.....QQA--V--K-------*---   496
B.AR.04.04AR151516              EPT--------F.....E--P--S---GT----------------K-------*-T-   503
B.AU.87.MBC925                  ...---------F.....E-----S------E----------A--K----------   502
B.BO.99.BOL0122                 ...---------F.....E-----F----QK-RG.....LYP--D--K-------*-K-   503
B.BR.03.BREPM2012               EPT---A----F.....E----------V-----------A--------------   511
B.CA.97.CAN3FULL                APPVQ-L----F.....E--SP-S----QT-Q----------A--K-----------   504
B.CN.05.05CNHB_hp3              ...---------F.....E----------------------------------   502
B.CO.01.PCM001                  ...---------F.....E-V---L-----T--------------K----------   499
B.GB.83.CAM1                    ...---------F.....E-K---S----------------A--------------   501
B.GB.86.GB8                     EPT--------F.....G-----S------N-------.P---A--------------   508
B.GE.03.03GEMZ010               ...---------F.....E--A--SP---T---D----------A--K----S------   499
B.IT.05.SG1                     ...---------F.....E-X--------------------M-----------------   502
B.JP.05.DR6538                  ...---------F.....E-I---S----------------M---A--------------   501
B.KR.05.05CSR3                  ...---------F.....E--A--------T----------Q---A--------------   501
B.NL.00.671_00T36               EPT--------F.....E-----S-R---T----------A--K----S--*---   513
B.RU.04.04RU128005              ...---------F.....E-K---S-----VE----------A--X----------   501
B.TH.00.00TH_C3198              ...---------F.....E--A--S--------D----------A--K----------   502
B.UA.01.01UAKV167               EPT--------F.....E----------------------A--K----------   506
B.US.04.ES10_53                 ...------I-GF.....E-AAA----------------K-----------------   504
B.US.99.PRB959_03               ...---------F.....E-APA-S----T----------SA--------------   504
C.AR.01.ARG4006                 ...---A----F.....E-.---SP----K-R-----------------K----S--L---   492
C.BR.04.04BR013                 EPT---A----F.....A-.---SP----K-R-----------------S--L---   502
C.BW.00.00BW07621               ...---A---KF.....E-.-N-TP----K-R-----------------K----S--L---   493
C.CN.98.YNRL9840                ...---A----F.....E-.---AP---SK-R-----------------L---   492
C.ET.02.02ET_288                EPT---A----F.....E-.A---P---LK-R.........A----K-------L---   498
C.GE.03.03GEMZ033               ...---A----Y.....E-.---AP----K-R----------I--K----S--L---   491
C.IL.99.99ET7                   ...---M----F.....E-.---AP---SK-R-----------------K----S--L---   491
C.IN.99.01IN565_10              ...---A----F.....E-.---TP----K-R-----------------K----S--L---   501
C.KE.00.KER20I0                 ...---A----F.....E-.-ALAP---LKER-----------------K----S--L---   491
C.MM.99.mIDU101_3               ...---A---GF.....G-.---AP----K-R-----------------L---   491
C.MW.93.93MW_965                ...---A----F.....E-.---AP----KEDR......EP-----K-------L---   493
C.SN.90.90SE_364                ...---A----F.....E-.---TP...QK-R-----------.-K----S--*L--   486
C.SO.89.89SM_145                EPT---A----F.....E.#---AP---LKE..........H----K----S--L---   504
C.TZ.02.CO178                   ...---A----F.....E-.---AP----KGG----------I--K----S--L---   491
C.UY.01.TRA3011                 ...---A----F.....E-.A--SPR--QK-R-----------------K----S--L---   488
C.YE.02.02YE511                 ...---L----F.....E-.---AP---AK..-----------------K----S--LLP-   489
C.ZA.04.04ZASK164B1             ...---A----F.....E-.---AP----K-R-----------------K----S-LL---   494
C.ZA.05.05ZAPSK240B1            ...---A----F.....E-.---TP---LK-G-.......SH----K----S--L---   491
C.ZM.02.02ZM108                 ...---L-----------.---AP----K-R----------I--K----S------   491
D.CD.83.ELI                     ...---A---GF.....E-.I--S----QK----------------K-------L---   501
D.CM.01.01CM_0009BBY            ...---I---GF.....E-.V--S----QK--D....KE----A--K-------L---   497
D.KE.01.NKU3006                 ...---A---GF.....E-.I--S----QK---------M--SA--KL------L---   499
D.KR.04.04KBH8                  ...---A---GF.....E-.I--S--R-QK-N-----------D--K-------L-K-   500
D.TD.99.MN011                   ...---A---GF.....E-.I--S----QK--D....KE----A--K-------L---   503
D.TZ.01.A280                    ...---A-I-GF.....E-.IK-S----QK--D....KE---SA--K-------L---   497
D.UG.99.99UGK09259              ...---A---GF.....E-.I-------QQ---------KH--A--K----S---L--   499
D.YE.01.01YE386                 ...---A---GF.....E-.V--S----QK-Q-.....----A--K----S--*L--   496
D.YE.02.02YE516                 ...---A---GF.....E-.I--S----QK--D....KD----------------L---   501
D.ZA.90.R1                      ...---A---GF.....E-.I--S----QK----------------K----S--L---   501
F1.AR.02.ARE933                 ...---A---GF.....RE-.I--SP---QK-EG.....LYP--A--K-------*L--   495
F1.BE.93.VI850                  ...---A---GF.....RE-.I--SP---QK-G-.....LYP--A--K-------*---   494
F1.BR.01.01BR125                ...---A---GF.....E-.IN-SP---QKEEG.....HNP--A--K-------*---   497
F1.ES.x.P1146                   ...-A-A---GF.....-..VEVTLH--QKGEG.....LYP--A--K-------*---   493
F1.FI.93.FIN9363                ...---A--LGI.....RE-.V--SPR--QKEEG.....QYP--A--K-------*---   496
F2.CM.02.02CM_0016BBY           ...---A-G-GF.....E-.I--SP---QK--G.....LYP----K-------*---   493
F2.CM.95.MP257                  ...---A---GF.....E-.IA-SP---QK--------.QVP--I--K----S-Q*---   499
F2.CM.97.CM53657                ...---A---GF.....E-.I--SP---QK--------.MYP-----K-------*---   493
G.BE.96.DRCBL                   ...---A-N-GF.....E-.IA-SP---QKE-----------S--K------Q*-K-   493
G.CM.01.01CM_4049HAN            ...---A---GV.....E-.IA-SP---RE----------------K----S--*---   497
G.CU.99.Cu74                    ...---A---GF.....Q-.IA-SP---QKE-D..........-I--K----S--*---   497
G.ES.99.X138                    ...---A----F.....E-.IA-SP---QKE----------A--K----S--*---   497
G.GH.03.03GH175G                ...---A---GF.....R-.-A-SP---RE-------.E-------K-------*---   499
G.KE.93.HH8793_12_1             ...---A---GF.....E-.IA-SP-P--KE--------IH--A--K----S--*---   499
G.NG.01.01NGPL0669              ...---A---GF.....E-.IA-S-----RE--------.P--A--K------Q*---   498
G.PT.x.PT2695                   ...---A---GF.....E-.IA-SP----K----------------K----S--*---   497
G.SE.93.SE6165                  ...---A--LGF.....E-.IA-SP---MKE---------.LYP--K----S--*---   498
H.BE.93.VI991                   ...---A---GF.....E-.I--SPR--LKEQ--------.P------------Q*---   501
H.BE.93.VI997                   ...---A---GF.....E-.M-SSP---LK----------.P-FA--K-------L---   500
H.CF.90.056                     ...---A---GF.....E-.M--SP---QLKDK......EP--A-------S--LL--   501
J.CD.97.J_97DC_KTB147           ...---A---GF.....-EK.I--S-----RKEL........--SA--K----S------   488
J.SE.93.SE7887                  ...---A--LGL.....E-..I-SP----K----------------K----S--L---   498
K.CD.97.EQTB11C                 ...---A---GF.....EK.I--SLR--MK-Q......QGP-----K----S--L---   496
K.CM.96.MP535                   ...---A---GF.....E-.I--SPR--TK---------.QSP-----K-------L---   496
```

```
                                    Vpr binding
                       Vpr binding               p6 end, Gag end
B.FR.83.HXB2                  ...APPEESFRS.....GVETTTPPQKQEPIDKE......LYPLTSLRSLFGNDPSSQ*    500
01_AE.CF.90.90CF11697         ...---A--LGM.....-E-.I-SFP---QK-------.....HPS--V--K-------L---    499
01_AE.CN.05.FJ051             ...---A-NWEM.....RE-.I--SM---QK--D.....HPP--V--K-------L---    499
01_AE.CN.06.FJ054             ...---A-NWGM...GD-A-.--SSL---QK-RV.....HRP--V--K-------L---    502
01_AE.HK.x.HK001              ...---A-NW-.........-.I-SLL---QK-------.....HPS-SI--K-------L---    496
01_AE.TH.01.01TH_R2184        ...---A-NWGM.....-E-.I-SSL---QK-------.....HPP--V--K-------L---    497
01_AE.TH.02.OUR769I           ...---A-NWGM.....-E-.--S-L---QK-------.....QPP--V--K-------L---    497
01_AE.TH.90.CM240             ...---A-NWGMGEEIT-E-.I-SLP---QK-------.....HPP--V--K-------L---    504
01_AE.US.00.00US_MSC1164      ...---A-NWGM.....-E-.I-SLL---QK-------.....HPP-SV---------L---    499
02_AG.CM.02.02CM_4082STN      ...---A---GM.....-E-.--S-L---QQEPR..EKGLYP-----K-------*---    496
02_AG.EC.x.ECU41              ...---A--V-I.....-E-.I-SSPN--TR--G.....LYP-V---K-------*---    461
02_AG.GH.03.GHNJ196           ...---A--LGM.....-E-.VASTP----G--G.....IYP-----K-------*---    501
02_AG.NG.01.PL0710            ...---A--LGM.....-E-.I-SSPQ---R--G.....LYP-----K-------*L--    491
02_AG.NG.x.IBNG               ...---A---GM.....-E-.IP-SPQ---R--G.....LYP-----K-------*L--    495
02_AG.SE.94.SE7812            ...---A--LGI.....-E-.I-SS-----G--G.....LYP--A--K-------*---    495
02_AG.SN.98.MP1211            ...---A---GM.....-E-.I-SSP----G--G.....LYP-----K-------Y---    495
03_AB.RU.97.KAL153_2          ...---A-N-GM.....-E-.I--SL---QK-R-.....QHP-SI--K----D--L---    499
04_cpx.CY.94.CY032            ...---A-CLER.....KE-.--SSL----R--------.......-------K----S--L---    501
05_DF.BE.x.VI1310             ...---A----F.....-E-.IASSP---QK-EG.....LYP--A--K-------L---    503
06_cpx.AU.96.BFP90            ...---I---GF.....-E-.IA-SP---SKE--...EKG----A--K----S--*---    500
08_BC.CN.97.97CNGX_6F         ...---A----F......E-.---AP----K-R-.........----------S--L---    490
09_cpx.GH.96.96GH2911         ...-X-A---GM.....RE-.X-SSXX-XKDGEX......PP--X--K-------L---    497
10_CD.TZ.96.96TZ_BF061        ...---A---GF.....-E-.I--S----QK--------......PH--A--K-------L---    501
11_cpx.GR.x.GR17              ...---A---GF.....-E-.---SP----KE--.......---I---K----S--L---    495
12_BF.AR.99.ARMA159           ...---A---GF.....-E-.I-SSP-R-QK-EG.....QYP--A--K-------*-*-    494
13_cpx.CM.96.1849             EPT---A---GF.....-E-.IA-S-----K..-.......T---AA-K----S--*-*-    508
14_BG.DE.01.9196_01           ...---A---GF.....-E-.IA-SP----KE--......S------K----S--*---    497
15_01B.TH.99.99TH_MU2079      ...---A-NWGM.....-E-.--SLL---QK-R-.....HTP--V--K-------LL--    497
16_A2D.KR.97.97KR004          ...---A-D-GM.....-E-.---L----LKNR-.....QHT-AI--K-------LL--    499
17_BF.PE.02.PE02_PCR0155      ...---L---GF.....-EK.I--DP---QK-EG.....QYP--A--K-------*---    493
18_cpx.CU.99.CU76             ...---A--LGF.....-E-.IA-SP---SKEEK.......--FS--K----S------    497
19_cpx.CU.99.CU7              ...-------GF.....-E-.I--S---DQK--------......----A--K-------L---    501
20_BG.CU.99.Cu103             ...---A---GF.....-E-.IA-SP--GQKE-D.........-I--K----S--*---    497
21_A2D.KE.91.KNH1254          ...---A---GF.....-E-.I--S--R-QKEEL.......------K-------L---    499
22_01A1.CM.01.01CM_0001BBY....---A--VGM.....-E-.I-SS----QRE--......QPP--V--K-------L---    500
23_BG.CU.03.CB118             ...---A---GF.....-E-.IA-SP--G-KEQD.........-I--K----S--*---    498
24_BG.CU.03.CB378             ...---A-E..........-.IA-SP---QKE-D.........-I--K----S--*---    492
24_BG.CU.03.CB471             ...---A---GF.....-E-.IA-SP---QKE-D.........-I--K----S--*---    497
25_cpx.CM.01.101BA            ...---A---GF.....-E-.I--SP----T---.......---MA--K----S--*---    499
27_cpx.CD.97.97CDKTB49        ...---A----F.....-E-.V--SP----K---........V------------L---    500
28_BF.BR.99.BREPM12609        ...---A----F.....-E-.---SP---QKEEG.....LYP--A--K-------*---    499
29_BF.BR.02.BREPM119          ...---A----F.....-E-.I--SP---EKNEG.....LYP--A--K-------*-K-    487
31_BC.BR.02.110PA             ...---A----F......E-.---SP---QK-R-.........-----K-------L---    494
32_06A1.EE.01.EE0369          ...---A----F.....EE--AA-SS----KK--.......---SA----------*---    495
33_01B.MY.05.05MYKL007_1      ...--------F.....-E----V-----------.......---V---K----------    499
34_01B.TH.99.OUR2478P         ...--------F.....-EG---S-----------.......---SA--K-------LL--    498
35_AD.AF.05.05AF095           ...---A---GF.....-E-.I--S----QK-G-......----A--K----------    495
36_cpx.CM.00.00CMNYU830       ...---A-N-GI.....-E-.MSSS----SR--G.....LYP-----K-------*L--    498
37_cpx.CM.00.00CMNYU926       ...---A---EV.....KE-.VASSL--#.fG--.....LYP--A--K----S--*P--    492
38_BF1.UY.05.UY05_4752        ...---A---GF.....-E-.IA-SP---QKNEG.....LYP--A--K-------*---    495
39_BF.BR.03.03BRRJ103         APP--S-----F.....-EG----SR-T-LT-RD....KE-----------S------    506
40_BF.BR.04.04BRRJ115         ...---A---GF.....-E-.I-S-P-R-QKEEG.....LYP--A--K------Q*--L    504
42_BF.LU.03.luBF_05_03        EPT---A---GF.....-E-.I--SL---QK-EG.....PYP----------------    503
43_02G.SA.03.J11223           ...---A---GF.....-E-.IA-S-----RE--......SP--A--K-------W-K-    498
U.CD.83.83CD003               ...---A---GF.....-E-.---SP----R------....SLYP-----K----S---L--    500
U.CD.90.90CD121E12            ...---A---GF.....-E-V---SP---LR--------......SP--A--K----S--LL--    497
U.GR.99.GR303                 ...---A---GF......G-EIV-SP--G-K--------......----A--K----S--*---    495
U.NL.95.U_NL_95_H10986_D1     ...---A---GF.....-E-.I--SP---MK-------.....QDP-----K----S--W---    497
CPZ.CD.90.ANT                 ...---I-IYQ......EEHKR-QKGLKG..EE-.......-P-SY--K----K-Q*WKF    525
CPZ.CM.01.SIVcpzCAM13         ...---A--YGY......Q-KEKNQEE--EGN......RLYP-----K----S------    507
CPZ.CM.05.SIVcpzEK505         ...---I--YGY......Q-EK-TQGTEREEKEKTE.SSLYP-----K----S---L--    512
CPZ.CM.05.SIVcpzLB7           ...---I---GF......QPK--QK-EDREE--TE..NKLYP-----K-------L---    504
CPZ.CM.05.SIVcpzMB66          ...---I---GF......R-EEPAQKEEVKSEETE..NKLYPQ----K-------L---    500
CPZ.CM.05.SIVcpzMT145         ...---M---GF......Q-...EKKVK-EVRE-...SKPPS-----K----------    492
CPZ.GA.88.GAB1                ...---I--YGY......Q-EEKSQE-K-GESSL........--P---K----S------    509
CPZ.TZ.00.TAN1                ...T-EKGKAQE......Q--AQT-VVPTAPPL-....MTMKGGF--K-I--S-Q**--    526
CPZ.US.85.CPZUS               ...---I-D-GY......Q-E-VTQE--GKEKEP.......FQ----K----S----E-    506
CPZ.CM.x.SIVcpzDP943          ...---I---GY......Q---VNK-GEKKEMEVEKE.PLYPQ----K----S----E-    507
CPZ.CM.x.SIVcpzMB897          ..S---I---GF......QDQKIDSL--XVRNEXQTENRLYP--D--K-------L--Q    508
N.CM.02.DJO0131               ...---L-NYG-......Q-ERSTQG-ERQENT-NSL...YP-----------L----    513
N.CM.04.04CM_1015_04          ...---L--YG-......Q-EGSTQG-EMQENQ-KTENSLYP-----------L----    512
N.CM.04.04CM_1131_03          ...---L--YG-......Q-ERSTQR-EMQENQ-KTENTLYP-----------L----    512
N.CM.95.YBF30                 ...---L--YGF......Q-EKSTQG-EMQENQ-RTENSLYP----------------    513
N.CM.97.YBF106                ...---L--YG-......Q-EKSTQG-EMQENQ-KTETSLYP-----------LX---    512
O.BE.87.ANT70                 ...---M-EEVK.....-Q-..NQE--GG-NE.........---FA--K----T-Q*L-Q    501
O.CM.91.MVP5180               ...---M-EAVK.....EQ-..NQS--..GDQE-.......---FA--K----T-Q*--Q    501
O.CM.96.96CMABB637            ...---M-EVVK.....-Q-..NQN-N-GGDQD-.......---FA-------T-Q*L*Q    501
O.CM.98.98CMA104              ...---M-EVVK.....-Q-..NQE--ED..SN-.......---FA--K----T-Q*--Q    501
O.CM.99.99CMU4122             ...---M-EE-K.....-P-..SQE--GD..PN-.......---FA-------T-Q*L-Q    501
O.FR.92.VAU                   ...---M-EMMK.....-Q-..EDQNQK.GEQN-.......---FA-------T-Q*--Q    501
O.SN.99.SEMP1299              ...---MTEEMK.....-Q-..NQE--ED..QN-.......---FA-------T-Q*L-Q    500
O.US.99.99USTWLA              ...---M-EVVX.....-XX..NQN-GED..SN-.......---FA--K---XT-Q*--Q    501
```

```
                              Gag-Pol TF start                                                            Gag-Pol TF end  protease start
B.FR.83.HXB2                  FFREDLAFLQ....GKAREFSSEQT...RANS..............PTRRELQVWG.....RDNNSPSEAGADRQGT......VSFNFPQVTLWQRPLVTIKIGGQLKEALLDTGADDTVLEE.MSLPGRWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTP  137
A1.GE.99.99GEMZ011            ----N---Q-....RE--K------...--I-..............--S-T-WDG-.....---PL-.---TE----A.....S------I---------VR--------------------D.ID---K-------------------E-------K------------  137
A1.KE.00.KER2008              ----N---Q-....RE---------...G---..............--S-D-WDG-.....---.-L-----E-E--N.....PTLC---I------V--V-V--EMR--------------D.IN---K---------------K----VS-----K------------  137
A1.KE.00.KNH1144              ----N---Q-....-E--K------...GT--..............--S-A-WDG-.....--.SL--K---E----D.....STL-L--I---------V--A---R--------------D.ID---K---------------K--E-V------K------------  137
A1.KE.00.KSM4024              ----N---Q-....-E--K------...GT--..............S-S-D-WDG-.....---.GL---E--E--SS.....YTLS---I---------VR--------------------D.IN---K---------------K-----------K------------  137
A1.KE.00.MSA4069              ----N---Q-....-E--K-----P...GT--..............--S-D-WDG-.....---.SL-----E----G.....GT-----I------V--VR-E------------------D.ID---K---------------------------K------------  137
A1.KE.00.NKU3005              ----N---Q-....-E--K------...G---..............--S-A-RDK-.....---.SL-----E---SE.....PT-S---I---------V---------------------D.IN---K---------------------------K------------  137
A1.RU.00.RU00051              ----N---Q-....XE--X------...----..............--S---WDG-.....--SPL.--T--E----A.....S------I---------V---------------------D.IN---K---------------------------K----A-------  137
A1.RW.93.93RW_024             ----N---Q-....-E--K------...----..............--S-APWDG-.....---.SP-----E----G.....PT-S---I-------------------------------D.IN---K---------------------------K------------  137
A1.SE.95.SE8891               ----N---Q-....-E--K------...G---..............--S-D-WDG-.....---.-L----EE---VG....GTTL----I-F-------VR--MQ----------------D.IN---K---------------K-----------K------------  138
A1.SE.95.UGSE8131             ----N---P-....-E-GK------...G-I-..............--S-D..DG-.....---.SL-----KQPT.........-S---I------I--VR--------------------D.IN---K---------------------------K------------  131
A1.TZ.01.A173                 ----N---Q-....RE--K--P---...G---..............--S-A-WDG-.....---.SL---T--EG---G....PTLS---I---------VR--------------------D.IN---K---------------K-----P-----KR-----------  137
A1.UA.01.01UADN139            ----N---Q-....RE--K------...--I-..............--S---WDG-.....---.-PLP-T-TE----A....A------I---------VR--------------------D.ID---K---------------------------K------------  137
A1.UG.92.92UG037              ----N---Q-....RE--K------...-T--.............PTSS-D-WDE-.....---.SL-----E---PE.....PT-S---I---------V---------------------D.IN---K---------------K-----------K------------  138
A1.UG.99.99UGA07072           ----N---Q-....-E--K------...G---..............--S-A-WDG-.....---.DL-----E----G.....PTLS---I---------VR--------------------D.IN---K---R-----------K-----------K------------  137
A2.CD.97.97CDKTB48            ----N---Q-....RE--------D...----..............--.....NG-.....---LLA.---.E.--AV.....HPC----I---------V--E--R---------------D.IN--WK---------------------V-----KR-----------  130
A2.CY.94.94CY017_41           ----N---Q-....RE--K-----N...----..............--S---ENG-.....----.LLP---TGD---I....Q-C----I---------V--E--------------------.IN---K---------------------A-----KR-----------  137
A.SN.01.DDI579                ----N---Q-....-E--K------...----..............--S-SPGDGR.....---LL-.---EG-----.....A--S---I------I--VR--------------------D.IN---K---------------------------K------------  136
A.SN.01.DDJ369                ----N---Q-....-E---------...----..............--S-APWEG-.....---.-PL----EG-----....T------I---------V-V-------------------D.IN---K---------------------------K------------  136
A.SN.96.DDJ360                ----N---Q-....RE---L-----...----..............--S-SPGDGR.....--TLL-.---EG-----.....A--S---I------I--VR--------------------D.IN---K---------------------------K------------  136
A.CD.02.02CD_KTB035           ----S---Q-....R----L-P-K-...G---..............-AS--PRDG-.....----LP-.--RDGG---E....L-ISL--I---------VRV-------------------D.ID---K---------------------------K------------  137
A.CD.97.97CD_KCC2             ----N---Q-....RE---L-P-R-...----..............--S--PGDG-.....---LL-.---DKG--PA.....I-LDL--I------V--VRVE------------------D.IN---K---------------K-----------K------------  137
A.CD.97.97CD_KTB13            ----S---Q-....R----L-PK--...----..............--S--PGNG-.....K--LL-.--R-EG--PA.....G-LSL--I------I--V---------------------D.IN---K---------------------------K------------  137
B.AR.04.04AR151516            ----T---P-....-E--K-P----...----PRA.........NS----------.....--T--L---NN-----......--LDL--I--------N------I------------X--D.-N---K---------------------S--------V---------  142
B.AU.87.MBC925                --------P-....-----------...----..............----------.....G----L-------E--......---S---I---------V-V--------------------.-N--------------------------------------------  137
B.BO.99.BOL0122               ----N---P-....RE---------...----..............---G------.....----PL----TE----V.....S--G---I--------K--V-------------------D.-N---K---------------K-----P------------------  138
B.BR.03.BREPM2012             --------P-....-E----P---A...--QTRA..........NS--S-------.....G----------G----......---S---I-----------------------------V--.-N---K---------------------P------------------  141
B.CA.97.CANB3FULL             ----N---P-....-----------...----PT...........S--T-------.....--IT-L-------P--......---S---I--------------------------------.-N-------------------------S--------V----I----  140
B.CN.05.05CNHB_hp3            ----N---P-....-E---------...--I-..............---G----R-.....G----T----------......I---L--I--------P------I---------------D.-N-------------------------P--------V----I----  137
B.CO.01.PCM001                ----N---Q-....-E---------...----..............----------.....GGD-----T--N----......-------I--------------------------------.-----K-R-----------------H--V-----------------  137
B.GB.83.CAM1                  ----N---P-....-E---------...----..............----------.....-E---L----------......---S---I--------------------------------.-N-------------------------P------------------  137
B.GB.86.GB8                   --------P-....------P----...----PTR........ADS--------R-.....-----LT-T---K---......---S---I------I------------------------D.-N-------------------K------V-----------------  143
B.GE.03.03GEMZ010             ----N---P-....-----------...----..............---G------.....G-S--L---D----S-......--LS---I---------VRV--------------------.-N-------------------------P------------------  137
B.IT.05.SG1                   ----------....----X-----A...----..............----------.....--X------------N......-------IP----------------------------*--.-N------------V------------P-----X------*-R---  135
B.JP.05.DR6538                --------P-....R-----P----...----..............----------.....-----L---------N......---S---I--------------------------------.-N---K---------------------P-------------I----  137
B.KR.05.05CSR3                ----N---P-....-E---L-----...----..............----------.....G-SD-S----------......---S---I-------------------------N------.---------------------------M------------------  137
B.NL.00.671_00T36             --------P-....-E---------...----PFSEQ....TRANS----------.....-----L----N-----......---S---I------I-------------------------.-D-----------------------------S---T----------  147
B.RU.04.04RU128005            --X-----P-....------P----...----..............----------.....-E---L-----G-K--......I-LS--XI------V-----E-------------------.-----K------X--------------T-D----------------  137
B.TH.00.00TH_C3198            ----N--CP-....-E---------...----..............---G------.....--S--I--------RP......I--S---I-------------------------------D.-N-------------------------P--------V----I----  137
B.UA.01.01UAKV167             ----N---P-....------P----...----PTNR.......ANS---G------.....----------T-----......---S---I--------------------------------.-N--------------------------V-----------------  144
B.US.04.ES10_53               ----N---P-....---G--PP--N...--I-..............---GD-R---.....GGSS-----------K......--LD---I--------------------------------.IN-------R-----------------P------------------  137
B.US.99.PRB959_03             ----N---P-....------P----...----..............----------.....-GTS-L---D------......---S---I--------------------------------.-N-------------------------P------R------I----  137
C.AR.01.ARG4006               ----N---Q-....----K-XP---...----..............--S-----R-.....G-.-PL-----E----......--L----I-----------V--------------------.-K---N--------------------VM-----K------------  133
C.BR.04.04BR013               ----N---P-...G-E----PT---...----PTTR.......ANS--S-----R-.....G-.-PL---TEG----......-------I------I---------R---I-----------.IE---K----I-V------R------VP-----K----A-------  141
C.BW.00.00BW07621             ----N---P-....-E----PP---...----..............--S-----R-.....G-.KPH-----E----......--L----I--------S--V---I---------------D.IN---K---------------------------KR-----------  133
C.CN.98.YNRL9840              ----N---P-....-E----PP--A...----..............--S-----R-.....G-.-PS----IE----......--L----I--------S-RV---T----------------.IN---K---------------------P-----K------------  133
C.ET.02.02ET_288              ----T---P-....-E----P----...S--PTR.........ANG--S-----R-.....GG.-PS-----E---S......--L----I----------R---------------------.IN---K---------------------H-----K------------  139
C.GE.03.03GEMZ033             ----N---P-....-E---LPP--A...--D-..............--S-----R-.....G-.-PR-----E----......--L----I--------S--V---I----------------.I----K---------------------V-----K----S-------  133
C.IL.99.99ET7                 ----I---P-....-E----P----...----..............--NG----R-.....-N.-PR-KT-VE----......--L----I--------S--V---IR---------------.IN---K---------------------------K------------  133
C.IN.99.01IN565_10            ----N---P-....------P----...----..............--S-----R-.....G-.DPH-----E----......--L----I----------RV---I----------------.LA---K---------------------T-----K------------  133
C.KE.00.KER2010               ----N---P-....-E----P---A...----..............--S-----R-.....G-.SPC-----E-E--......--L----I---------V-----IR---------------.IN---K--------------------VI-----K-T----------  133
C.MM.99.mIDU101_3             ---KT---P-....-E----PP---...----..............--SG--R-R-.....G-.-PS----TE----......--L----I--------SV-V---II---------------.IN---K---------------------SM----K------------  133
C.MW.93.93MW_965              ----N---.-....-E----P----...----..............--S-----R-.....G-.-PC-----EGRQGT.....-TL----I-------IS------I----------------.IN---K------------------E-VP-----K--L---------  134
C.SN.90.90SE_364              ----N--CP-....-Q----P----...----..............--S-----R-.....G-.-PH....-E----......--L-.S-I---------------I----------------.IN---K---------------------P-----K------------  129
C.SO.89.89SM_145              ----N---P-....-E----P---A...----PTSRELQGQARANS--S-----R-.....G#.-PR-----E-T--......--L----I--------S--V---I---------------D.IN---K---------------------P-----K----S-------  144
C.TZ.02.CO178                 ----N---P-....-E----P---A...----..............--S-----R-.....--.-PR----TEGR--......--L----I--------S-RV---V----------------.IN---K---------------------S-----K------------  133
C.UY.01.TRA3011               --------P-....-E--K-PP---...----..............--S-----R-.....-G.-PL-----E----......-------I--------------------------------.-K---N---------------------V-----K----S--I----  133
C.YE.02.02YE511               ----N---P-....------P----...----..............--I-----R-.....GN.-PR-K--GE..--......--L----I------V----V---I----------------.IN---K---------------------------KR-----------  131
C.ZA.04.04ZASK164B1           ----N---P-....-E---LP----...--F-..............S-S-----R-.....G-.-PC-----E----......--L----I------V-S--VE--I----------------.I----K---------------------S-----K------------  133
C.ZA.05.05ZAPSK240B1          ----N---P-....-E---L-P--A...----..............-AS-----R-.....G.-PH------EGR-I......-TL----I--------S--V---V--------------D.IN---K---R-----------------P-----K------------  133
C.ZM.02.02ZM108               ----N---P-....-E----PP---...----..............--T............G-.-PS-----E----......-------I--------S--V---I----------------.IN---K---------------------P-----K------------  128
D.CD.83.ELI                   ----N---P-....---G-L-PK--...----..............--S---R---.....--.-PL-KT--E----......-------I--------A-----------------------.-N---K---------------------P-----Q------------  136
D.CM.01.01CM_0009BBY          ----N---P-....---G-L-----...----..............--N---R---.....-G.-PL--T--EG-RQ....GT--LS---I-----------V-------------------D.IN---K--------------------VP-----Q------------  138
D.KE.01.NKU3006               ----N---P-....---G-LP----...----..............--S---R---.....G-.-TLP-T--E---D......---S---I---------V------R---------------.ID---K----------------------VD----R-----------  136
D.KR.04.04KBH8                --------P-....W---KLP---N...---R..............--S---R---.....G-.-PL--T--E--R-......--LG---I--------KV-----I---------------D.IE---K----------------------------------------  136
D.TD.99.MN011                 ----N---Q-....---GK------...----..............--S---R---.....--.-PL--T--EG-RQ....GT---S---I-----------V---Q---------------D.IN---K------------------E--PV-----T-V---------  138
D.TZ.01.A280                  ----N---P-....R----LP----...----..............--S-D-R---.....G-.KTL--T--E---Q....GT---S---I--------------------------------.-----K---------------------------Q------------  138
D.UG.99.99UGK09259            ----N---P-....-E----P----...----..............--S---R---.....G-.-TS--T-TT---E......A-LS---I---------V---E------------------.ID---K----------------------V-----R-----------  136
D.YE.01.01YE386               ----N---P-....-----LP----...----..............-AS---R---.....GG.-TL--T--E-P-A......---S---I------V--VM---------------------.-N---K----------------------V-----------------  136
D.YE.02.02YE516               ----N---Q-....-E-GKL-----...---C..............--S---R---.....--.-PL--T--EG-RQ....GS--S----I----------N----Q---------------D.ID---K------------------E-V-V-----------------  138
D.ZA.90.R1                    ----N---P-....---G-------...----..............--S---R---.....G-.-PL--T-TE---A......---DL--I---E----------------------------.-N---K---------------------PL------------I----  136
F1.AR.02.ARE933               ----N---Q-....-E--K-P----...----..............-AS---R-QR.....G-.-PL-----E-R--.....VP-LS---I--------A-N--------------------D.IN---K---------------K---N----------TC--------  137
F1.BE.93.VI850                ----N---Q-....-E--K-P----...----..............--S---R-QR.....G-.-PL-----E-R--.....VP-LS---I---------------I---------------D.IN---K---------------K---N--------------------  137
F1.BR.01.01BR125              ----N---P-....-E----P---A...-T--..............--S---R-RR.....G-.KPL-----EGG--.....QP-LS---I----------RV-------------------D.IN---K---------------K---N--------------------  137
F1.ES.x.P1146                 ----N---PE....-E-GK-P----...--Y-..............C-S---WIRS........G-N-AS-TEGR--.....VS-LSL--I-----------V----R--------------D.IN---K---------------K-----T-D------T---------  135
F1.FI.93.FIN9363              ----N---Q-....-E--K-P--*-...----..............-AS--PRDQR.....-G.--L-----E-R--.....VP-LS---I-------------------------------D.IN---K---------------K---H--------------------  136
F2.CM.02.02CM_0016BBY         ----N---Q-....RE-W--H----...---G..............-AS-G-R-RR.......--LP-----E----.....VS-LD---I------V----VE---R--------------D.IN-S-K---R----------------VP-----Q------------  137
F2.CM.95.MP257                ----NV--Q-....-E--K------...----..............-AS---R-R-.....G-.S-LP----E----.....GS-LD---I------V----V----R--------------D.IN---K--------------------VS-----Q------------  137
F2.CM.97.CM53657              ----N---Q-....-E-WK------...----..............-AS-K-R-RR.......--LP-----E---N.....VP-LD---I-----------VE---R--------------D.IN-S-K---R-----------------P-----Q------------  137
G.BE.96.DRCBL                 ----N---Q-....-E----P---A...----..............------R-R-.....G-.SPLP----EGK--......I-SI---I------I-KVR-----I---------------.ID---K-------------------------S-KR-----------  136
G.CM.01.01CM_4049HAN          ----N---Q-....RE---L-P--P...----..............--S---RSRR.....G-.SPLP----KGE-A......I-LDL--I---------V------I--------------D.IN---K-------------------------S-K------------  136
G.CU.99.Cu74                  ----T---Q-....RE-G----K-A...--I-..............--C---R-RT.....G-.SPLP--R-EGE--......--L-L--I---------VR-----I--------------D.IN---K---------------------------K------------  133
G.ES.99.X138                  ----N---Q-....-E-----P--A...G---..............--------RR.....G-.SPLP--R-EGK-I......-PLSL--I---------VR----PI--------------D.IN---K----L-----------------V----K------------  135
G.GH.03.03GH175G              ----N---Q-....-E--------A...--I-..............S-S---R-RE.....G-.SPLP----EGK-G.....TI-L----I---------VR-----I---------------.IN---K---------------------------K------------  137
G.KE.93.HH8793_12_1           -#--N---Q-....-E--------A...----..............------R-RR.....GN.SPLP--R-EGK-D......T-LS---I---------V------I--------------D.IK---K-------------------------S-K------------  135
G.NG.01.01NGPL0669            ----N---Q-....-E--K-----A...--D-..............------R-RR.....G-.SPLP----EGE-I......T-LSL--I------I--VR-----I--------------Q.IN---K---------------K-----P---E-K------------  136
G.PT.x.PT2695                 ----N---Q-....-E-----P--A...----..............------R-RR.....G-.SPLP--R-EG--V......IPL-L--I---------VR-----I--------------D.IN---K---------------------------K------------  136
G.SE.93.SE6165                ----N---Q-....-E--------D...-T--..............--C-KPR-RR.....G-.SPLP---DEGK-A........ISL--I---------V------I---------------.IN------------------------VP---S-K-----I------  134
H.BE.93.VI991                 ----N---Q-....------PP-EA...----..............--S---R-RR.....G-.HPL-----E-T--......-------I------I--V--E------------------D.IN---K------------------E-VA---F-K------------  135
H.BE.93.VI997                 ----N---Q-....RE--K--P--A...----..............--S---R-R-.....G-.DLLP----EG---........-LC--I---------V--E--R----------------.IN-L----------------------VA-----K------------  135
H.CF.90.056                   ----N---Q-....RE--K--P--A...-T--..............--S---R-RR.....G-.DPL-----AEGQG......T-LS---I---------V--E--R----------------.IN---K------------------E-VA-----K------------  136
J.CD.97.J_97DC_KTB147         ----S---Q-....-E---L-P--A...-T--..............--S-K-R-R-.....E-.--LP-T-TEEGTI......---S---I---------VRV--------------------.ID---K------------------ND-----E-K------------  135
J.SE.93.SE7887                --------Q-....RE---L-P---...----..............--S--PRAR......-G.DPLP-T--EG---......--S----I----------R-----R--------------D.ID--RK------------------NEVP---E-K-------I----  135
K.CD.97.EQTB11C               ----V--SQ-....RE--K------...----..............--S---W-R-.....E-.-PL--T-NE-S--.....GS------I------V--V-V----R---------------.IN---K--------------------VCM----Q------------  137
K.CM.96.MP535                 ----N---P-....-E---------...----..............--S---R-R-.....G-.-PL----DQ----.....EP------I------I----V----R---------------.IN---K---------------------V-----Q------------  137
```

Gag-Pol TF start — Gag-Pol TF end, protease start

B.FR.83.HXB2	FFREDLAFLQ....GKAREFSSEQT...RANS..............PTRRELQVWG.....RDNNSPSEAGADRQGT......VSFNFPQVTLWQRPLVTIKIGGQLKEALLDTGADDTVLEE.MSLPGRWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTP	137
01_AE.CF.90.90CF11697	----N--SQ-.....-E---------..----..............--S---GDG-.......--.-LLP----E---------.PF--S---I-F-------V--E-----------------D.IN------------------------------KR-----------	137
01_AE.CN.05.FJ051	----N---Q-.......---GK-------..----..............S-S-K-GNE-.......--.-PLN-T--E---P.....SS-LS---I---------V-------------------D.IN---K--------------------------K------------	137
01_AE.CN.06.FJ054	----N---Q-.......---G----K---..----..............-SS-K-GDGRR...G--.-LLT----E---------.SST-S---I---------V--E---R-------------D.IN---K------------------V-----K------------	139
01_AE.HK.x.HK001	----N--SQ-.....-E-----------..--F-..............--SGK-.E.......G-.-LLT----EG--------.PF---L--I---------V-------------------D.IN---K--------------------------K------------	134
01_AE.TH.01.01TH_R2184	----A---Q-.......---GK-------..----..............--S-K-GDG-.......--.-LLT----E---------.PS--S---I------V--V-------------------D.IN---K-----------------E--P-----K------------	137
01_AE.TH.02.OUR769I	----N---Q-.......---G---------..-T--..............-AS-K-GDG-.......--.-L-T----E---------.AS--S---I---------V-------------------D.IN---K--------------------------K------------	137
01_AE.TH.90.CM240	----N---Q-.......---G---------..----..............--S-K-GDG-RDNGG--.-LLT----E---------.SS--S---I---------V-------------------D.IN---K-------------K-----------K------------	142
01_AE.US.00.00US_MSC1164	----N---P-.......---GK-------..----..............--S---GDG-.......--.-LLT----E---------.SS-LS---I---------V-------------------D.IN---K--------------------------K------------	137
02_AG.CM.02.02CM_4082STN	---KN---Q-.....-E--K--------..-T--..............--S---WNG-.......--.-L-----T..A--EGKGTIS-L----I---------VR-----I---------------.IN---K--------------------P-----K------------	140
02_AG.EC.x.ECU41	----N---Q-.....-E--K--P----..GT--..............S-S--R-NR-.....G-.-LL--S-DEG---.....IS-C----I---------V------I---------------.I----K--------------------P-----K------------	137
02_AG.GH.03.GHNJ196	----N---Q-.....-E--K--------..GTI-..............--S---GDG-.....-G.SLNP---GG--N.....IS-L----I---------VR--C--I--------------KVE---K--------------------------KR-----------	138
02_AG.NG.01.PL0710	----N---Q-.......----K--------..G---..............--S-KPGDG-.......--.-LL-T--TEG--A....IS------I--------AVR-E---I--------------D.IN---K--------------------A-----K------------	137
02_AG.NG.x.IBNG	----N---Q-.....-E--K--------..GT--..............S-S---WDG-.......--.T-L-T--TEG--A....IS------I---------VR-E---I--------------D.IN---K--------------------------K------------	137
02_AG.SE.94.SE7812	----N---Q-.....-E--K--------..GT--..............-AS--PWDR-.......--.-LL---TGG--.......IS-LS---I---------V------I---------------.IN---K--------------------------K------------	137
02_AG.SN.98.MP1211	----N---Q-.....-E--KL-------..GT--..............--S---WDG-.......--.-LLP---TGG--.......IP------I-------F-VR-E---IK-------N-----K.IN---K-------------L-----------K--M-S-------	137
03_AB.RU.97.KAL153_2	----N---Q-....RE--K---------..--I-..............--S-K-WDG-.......--.-PLP-T-TEG--.......AS------I---R-----VR--------------------D.IN---K--------------------------K------------	137
04_cpx.CY.94.CY032	----NV--Q-....RE--K-----A..----..............-A-GM-REER.....G-.-LL----TEG--.......I------I-----------L---IR-------------.IN---K--------------------P-----K------------	136
05_DF.BE.x.VI1310	----S---P-.....-E---LPP----..G-L-..............-AS-----------.....G-.SLL-----EGR--......VP-LS---I------V---R--------------------D.IN---K---------------------V----Q------------	137
06_cpx.AU.96.BFP90	----N---Q-.....-E--------A..----..............--H---RFRR.....G-.SPLP-T-VEGE-G...KGAI-LSL--I---------VR-----I--------------D.IN---K--------------K-----------K------------	139
08_BC.CN.97.97CNGX_6F	----I---P-.....-E----PP----..----..............--S-----R.....G-.-PS----TE---------......L----I--------S--V---I---------------.VN---K-----------------E--P-----K------------	133
09_cpx.GH.96.96GH2911	----N---Q-....RE---L-X----..G-IG..............-XS---WDE-.......--.XLLXXXXEG-GXA......P--X---I------I--V--------------------.IN---K--------------------------Q------------	136
10_CD.TZ.96.96TZ_BF061	----N---Q-....R----LP-------..----..............S-S---R--------.....G-.-TL--T--E---A......A-LS---I-----------------------------------.-N---K-----M-------------------Y------------	136
11_cpx.GR.x.GR17	----N---Q-.....-E----PT--A..----..............--S---R-R-.....G-.-PLP-T--QGE--......I-Y----I------V-PV-VA---------------------.------------------------K--ED-T---E-K-------I----	136
12_BF.AR.99.ARMA159	----N---Q-.....-E--K-P---A..----..............-AS---W-RR.....G-.-LL-----E-R--......VP-LS---I--------I--V--------------------D.IN---K--------------K---NV-------------I----	137
13_cpx.CM.96.1849	----N---Q-....RE-----------..-T--PTREQ....TRANS--S---R-RR.....G-.SPLP---TE..-D......I-LSC--I------I-IV---E--I-----------K-.IN---------------------NMS---E-K-------I----	144
14_BG.DE.01.9196_01	----N---Q-.....-E-----P--A..----..............-------R-RR.....G-.SPLP--R-EGK-V......IPL-L--I---------VR-----I--------------D.IN---K-----------------------G-K------------	136
15_01B.TH.99.99TH_MU2079	----T---Q..................--I-..............--S---GDG-.......--.-LLT----E---------.YS--S---I---------------------P----P---D.INF--K-N--------------------P-----K------------	136
16_A2D.KR.97.97KR004	----N---P-....RE---------N...-T--..............--S-G-WNG-.....G-.-PLA----EK---------.TH-C----I---------V--E---R--------------D.IN---K--------------------T----EKR----------	137
17_BF.PE.02.PE02_PCR0155	----S---Q-.....-E--K-P---D..----..............-AI--FR-R-.....E-.-PR-----E-R--......VP-LS---I---------------------------------D.IN---K-R-------------K---N-P---------V------X--	137
18_cpx.CU.99.CU76	----N---Q-.....-E-----P-KA..----..............--SG--G-RR.....G-.SPLP--RVEGGE.......I--L---I---------V------I--------------D.IN---K--------------------------KR-----------	135
19_cpx.CU.99.CU7	----N---P-.....-E--KLP------..----..............-------R--------.....GN.-PL--T-SE---------......--LSL--I-----------------------------------D.IN---K-R--------------------P-----Q------------	136
20_BG.CU.99.Cu103	----T---P-.........------P--A..----..............--C---R-RR.....G-.S-LP--R-EGE-.........L-L--I---------VR-----I--------------D.IN---K--------------------------K------------	133
21_A2D.KE.91.KNH1254	----G---P-.....-E-G-L----A..----..............--S---R------.......--.-PL--T--E-GTV.......-L----I---------V-V--------------------.-N---K---	135
22_01A1.CM.01.01CM_0001BBY	----N---Q-.....-E---L--G--..----..............--S--RGDG-.......--.-LLP----E-E--......AS--S---I---------------------------------D.ID---K--------------------------K------------	137
23_BG.CU.03.CB118	----T---P-........-----L----A..----..............--C---RIRR.....G-.SPLP--R-EGA-.........L-L--I---------VR-----I--------------D.IN---K--------------------P-----K------------	133
24_BG.CU.03.CB378	----T---P-.......---GKL----A..----..............--C-G..........G-.SPLP--R-EGE-.........L-L--I---------VR-----I---------------.IN---K--------------------------KR-----------	128
24_BG.CU.03.CB471	----I---P-.....-E--------A...--D-..............--C---R-RR.....G-.SPLP--R-EGE-.........L-L--I---------VR-----I--------------D.IN---K--------------------------K------------	133
25_cpx.CM.01.101BA	--------Q-....RE--KL----A...--D-..............--C---R--R.....G-.-PLP--------I......I-YG---I---------V-----II--------------D.IN---K---------------------M--S-K------------	136
27_cpx.CD.97.97CDKTB49	----N---Q-.....-E-----P--A...--I-..............--S-----R-.....GG.DPL--T--EG--S.......-L----I---------V--E------------------D.IN---K--------------K--E-VVV----K------------	135
28_BF.BR.99.BREPM12609	----N---P-.....-E--K-P---A..----..............-AS-----RR.....G-.-PL-----E-G--......VP-LSL--I-----------V----R--------------D.IN---K--------------K---D---D------T---------	137
29_BF.BR.02.BREPM119	----N---P-.....-E--K-P---A..----..............--S------R.....G-.-PL----GKER--......VP-LS---I--------E--V--------------------D.IN---K--------------K--EN-------------------	137
31_BC.BR.02.110PA	----N---P-.....-E--X-PP---..----..............--SG----R.....G-.-PL-----EG---.........L----I---------V-----R---------------.IE---N-----------------N--P-----K------------	133
32_06A1.EE.01.EE0369	----N---Q-.....-E---L----A..----..............--C-----R-.....G-SSPLL----EEK-A......I-LS---I------I--VR--E--I--------------D.IN---K--------------------H-----K------------	137
33_01B.MY.05.05MYKL007_1	----A---P-.......---G-LP-KP-..----..............---G-----------.....------S--------------.I-C----I--------------------------------D.IN------------------------------KR-----------	137
34_01B.TH.99.OUR2478P	----I---Q-.......---G--P-------..----..............---G-----------.....-----I--T-----------.I--S---I---------V-------------------D.IN---K--------------------P-----------I------	137
35_AD.AF.05.05AF095	----N---P-.....-E--K--------..----..............--S---R-------.....G-.-ALP-T--E-R-VV.......-S---I---------V-----QR--------------D.IN---K--------------K-----------K------------	136
36_cpx.CM.00.00CMNYU830	----N---Q-.....-E----------..-T--..............--S---WDR-.......--.VLF----IEG---.....IS-L----I---------VR-----I--------------D.IN---K--------------------P-----K------------	137
37_cpx.CM.00.00CMNYU926	----N---Q-....RE--K--------..G---..............--S---RGER.....GS.SLL-----..G---.....IP--S---I------IT-V------I--------------D.IN---K--------------------P-----K------------	135
38_BF1.UY.05.UY05_4752	----N---Q-.....-E--K-P---A..----..............--S---W-RR.....G-.SPL-----EER--......VP-LS---I-----------V----M--------------D.IN---K--------------K---N--V----Y--T---------	137
39_BF.BR.03.03BRRJ103	----N---P-.....-E------D--..----PT..........S-I-----------........-----L--D------Q....GT--L----I------I--V---------------------.-N---K----I--------R-K--ED-P------RV-----I----	142
40_BF.BR.04.04BRRJ115	----N---P-.....-E----P---A...--D-..............-AS---W-RR.....G-.-L------EGR--......VP-LS---I------V--VRV----R---------N----D.IN---K--------------K---D-P-------------------	137
42_BF.LU.03.luBF_05_03	----N---------.....-E--K-PP--A...--KARA..........NS-AS---W-RR.....G-.-PL-----E-R-SV.....P-L----I-----------V-E-I------------------.--------------------K---N-S-D---------------	141
43_02G.SA.03.J11223	----N---P-.....-E---LP---A..----..............--C---R-RR.....G-.S-LP----EGE-I......T-ISL--I--------KV---E--I--------------Q.ID------------------K-----P---E-K------------	136
U.CD.83.83CD003	----N---Q-.....-E---------..----..............--S---R-RR.....G-.-PLA----EG--.....VPLP------I---------V-------------------V--.-T---K--------------------------K------------	138
U.CD.90.90CD121E12	----N---Q-...-----------N...----..............S-S---R-RR.....GS-HPLA----EG--......VS--S---I---------V-------------------V--.-T---K--------------------------KR-----------	137
U.GR.99.GR303	--------Q-...-----------...--D-..............--S---R-R.......-GDSPLP----KG--A.....I--S---I---------VR-----I--------------G.I----K-----R--------------------K---------S--	136
U.NL.95.U_NL_95_H10986_D1	----S---Q-....RE----P-------..----..............--SG--R-RR.....G-.-PL----DE---------.GS------I-----M--VN----V--------------D.IN---K-R-R---------------E-VC-----Q------------	137
CPZ.CD.90.ANT	----TDPHVV...GVQT--LCA-GG.SSG--..............SS-H-D-SGGA.....QE...D--GSQGGG-.......T-LV--EIP-----MMEVL-Q--KCQ-----------V-G.IH-Q-N----T---------S-Q--NKVP-Q-GDRTVLA---L--N-	137
CPZ.CM.01.SIVcpzCAM13	----G---P-....RE--QLCT--D...----..............--S---W-PR...EGEEPRGTGRGEQ.........TIP-S----I-------IPVR-E---C-----------IQD.IQ-Q-I---------------K---NVS---E-K-VQ---------	136
CPZ.CM.05.SIVcpzEK505	----G--S--....RE--K-PPDNN.KE----..............-SN---W-S-...GEDHTGDREGRK.GEDREL...SVPTL----I------IL-V----EI------------I--.IQ-E-K---------------K---NVI---Q-K--V---------	143
CPZ.CM.05.SIVcpzLB7	----T----P.........-----KL--------..--H-..............-AN---W-SA....KNHPETGGQG.RGQDREQ...AVP-L----IS-----I----V--Y--G------------D.LE-G---------------K---N-P---G-K--L---------	140
CPZ.CM.05.SIVcpzMB66	----T---Q-...........-----P--E-...-T--..............S-N---R-Q-...GGTCPEGG--E.RG.DREQ...AVS-A----IS-----V--VR-E-------------------.IE-G--------------------NVT-D---KR-V---------	140
CPZ.CM.05.SIVcpzMT145	---KK--S-G....-E---LCAR-E...GT--..............S-YGI-RIS-.........GEESQ-GGEG-EQ...AA-------I-------IPV--E---I-----------I-D.LN---K-S--------------K-FERVN---E-K-ITS---------	136
CPZ.GA.88.GAB1	----R---P-....RE--QLCA--N...-T-G..............--D---W-P-...GREEPGEERGR...E-SIS.......T-L--I-------IPV-VE---C-----------I-R.IQ-Q-L---------------K-F-NVH---E-R-VV---------	135
CPZ.TZ.00.TAN1	----THPLVG....VQT--LCA-H.......................-RE--GSGA......G-STDT-G-NCPTT-DDD....ERRVL---N-----MM-V-VQ--VCQ---------S-FCN.IK-K-Q-T--T-------VP-SE-YN-PVQ-GNKEVRA-----E--	133
CPZ.US.85.CPZUS	----T-VPIVERGIKET--LPGK-E...G-H-..............S-N-GFR-SR...RDS-TGETGKG...KGALS.......A----I-------LR--VA--IV------------DN.IQIE-T-R---M---------K---HVN---E-R--Q-S--------	139
CPZ.CM.x.SIVcpzDP943	----R-VSVE.RGIKET--LPAG-E...GTH-..............S-H--FW-S-...NNSEQTGRK-RDGGGER....APVP-I-L--I-------LGV--E---I-----------IDN.VQ-T-----------------K---N-A---E-R--T---------	144
CPZ.CM.x.SIVcpzMB897	----N--SP-....RE----PP-E-.RTTRTK.............C--N--FW-S-...SENRFSETXGQERGXDREQ...TVP--G---IS-----V-AAR---VI--V------------.IE-E-K-R--------------------P-----KR-----------	145
N.CM.02.DJO0131	---KG-VS--....RET-QLPPDNNNKE--H-..............-AT---W-S-...GEEHTGEGEAGEHRELS......IPTL----I------VI-V---KEVR------------I--.IQ-E-K-------------------NVT-D-Q-R--V---------	142
N.CM.04.04CM_1015_04	---KG-VS--....RET--LPPDSN.KE--H-..............-AT---W-S-...GGEHTGKGDAGEPGEDRELS...VPTL----I------VISV---NEVR---X---------I--.IQ-E--------------------N-T-D-Q-R--V---------	144
N.CM.04.04CM_1131_03	----G-VS--....RET--L-PDNN.KE--H-..............-AT---W-S-...GEEHTEKGDAGEPGEDREHS...VPTL----I------VISV---NEVR------------I--.IQ-E--------------------N-T-D-Q-R--V---------	144
N.CM.95.YBF30	----E-VS--....RET-KLPPDNN.KE--H-..............-AT---W-S-...GEEHTGEGDAGEPGEDRE...LSVPT-----I------VI-V---KEVR------------I--.LQ-E-K-------------------N-TVD-Q-R--V---------	144
N.CM.97.YBF106	---KG-VS--....RET-KLPPDNN.KE--H-..............-AT---W-S-...GEEHTGKGDAGEPGEDRD...LSVPTL----I------VXAV---KEIR------------I--.IQ-E-K-------------------N-T-D-Q-R--V---------	144
O.BE.87.ANT70	---QI--SGG....HE--QLCA-TS...TPI-..............--DGGGSEGT.....GE.SGTERGPE.-A.........L-VCL--IP--D--I--ARV--H-C-V-----------NN.IQ-E-K--------------KE--NVTV--E-REVQ---------	133
O.CM.91.MVP5180	----V--SGG....HE--QLCA-TS...VPI-..............--NGGGSEGT.....-E.SESEGG...SGRA.......-PICL--IP--D--I--A-V--H-C-------------NN.IQ-E---T-------------KE-NNVTV-VQ-KEVQ---------	133
O.CM.96.96CMABB637	---QV--SGR....HE--QLCT-AS...API-..............--NGGGSEGA.....-E.SESEPR-G.PGRA......L-VCL--IP--D--V-VA-V--H-C-V-----------NN.TQ-E---T-------X-----KE-GNVTV--X-K-VQ---------	135
O.CM.98.98CMA104	---QI--SGG....HE--Q-YA-TS...GPI-..............-SDGGSSEGT.....GE.PGTERG...LKR-......L-VCL--IP--D--V--A-V--H-C-V-----------ND.IQ-E-K-----------C-R-KEFNNVTV--E-REVQ---------	133
O.CM.99.99CMU4122	----I--SGG....HE--QLCT-TS...VPI-..............SSDGGGKEGT.....-E.SGTEGG...PER-......L-VCL--IP--D--V--ARV--H-C-V-----------NN.IQ-E-K-T-------------KE-YXVKV--E-REVQ---------	133
O.FR.92.VAU	--K-I--SER....HE--QLCAGTS...VPI-..............--NGRDDEGA.....-G.G-E--G--E.........RAL-VCL--IP--D--V--ARV----C-V-----------NN.IQ-E-K-A-------------KE-YNVAV-LE-REVR---------	134
O.SN.99.SEMP1299	----I--SGG....HE--QLCA-TS...VPI-..............--DDGGNEGT.....-E.SGTEGG...PER-......L-VCL--IP--D--V--ARV--H-C-V-----------TN.IQ-E-K-T-------------KE-N-VPV--E-REVL---------	133
O.US.99.99USTWLA	----V--SGG....RE--Q-YA-ANAPSSPI-..............--NGGGSEXA.....XX.SESGGG...LER-......L-VCL--IP-XD--I--A-V--H-C-V-----------SN.IQ-E-K-T-------------KE-XDVTV--E-R-MQR---------	136

```
                           protease end p66, p51 RT start      M41L  D67N K70R          D110 catalytic site
B.FR.83.HXB2               VNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICT..EMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLP 305
A1.GE.99.99GEMZ011         -------M------------------T-----------------------------TD--N..D---------------------------------------------------------------R------------------S-----------T---------------- 305
A1.KE.00.KER2008           -------M------------------T--------R------------------T---Q..------------------I-----------------------------------------R-------------------KN----------L---------------- 305
A1.KE.00.KNH1144           -------M---------------------------------------------T----..---------Q---------I------------------------------------------------------------H-------------T------K-------- 305
A1.KE.00.KSM4024           -------M---------------------------------------------T----..-------------------I--------------------------------------------------------------N----------T---------------- 305
A1.KE.00.MSA4069           -------M---------------------------------------------T---M..-------------------I-------------------------------------------------------------KS----------T---------C------ 305
A1.KE.00.NKU3005           -------M---------------D-----------------------------T---A..-------------------I--------------------------------------------------------------S----------T---R--T-------- 305
A1.RU.00.RU00051           -------M---LX-------------T--------------------R--TD--K..---------------------------------------------------------------------------------N-----------------V-------- 305
A1.RW.93.93RW_024          -------M---------------------------------------------T----..-------------------I-----------------------------------------------------------------------------T---------------- 305
A1.SE.95.SE8891            -------M------------------T--------RI------------T---R..-------------------------------------------------..#------*---------------------KN---------------------------- 300
A1.SE.95.UGSE8131          -------M---------------------------------------------T----..-------------------I------N-DR-------------------------------------------------S----------T---------------- 299
A1.TZ.01.A173              -------M---------------------------------------R--T---M..-------------------I--------------------------------------------------------------S----------T------V-------- 305
A1.UA.01.01UADN139         -------M---------------------------------------------TD--K..---------V----------V------S---------------------------------------------------N-------------------------- 305
A1.UG.92.92UG037           -------M---------------S-------------RI------------T---A..D--R----------------I--------------------------------------------------------------S----------T---------------- 306
A1.UG.99.99UGA07072        -------M------------------T--------RI------------TK--K..------------------------------------------------------R-----------------------------N-------------------------- 305
A2.CD.97.97CDKTB48         -------M-V-L-----------------------------------------T---K..----------------------------D---------------------------------------------------H---------------------------- 298
A2.CY.94.94CY017_41        -------M-V-L-----------------------------------------T---K..--------------------------------------------------------------------A---------------H-------------T------V-------- 305
A.SN.01.DDI579             -------M---------------K----T------------------------TA--K..--E------------------------------------------------------------------------------S----------T------V-------- 304
A.SN.01.DDJ369             -------M---------------K----T------------------------TA--K..--E----------------------------------------------------------Q-------------------G----------T------V-------- 304
A.SN.96.DDJ360             -------M---------------K----T------------------------TA--K..--E-----------------------------------------------------------------------------IN----------T------V-------- 304
A.CD.02.02CD_KTB035        -------M-------------------------------I-----S-------T---R..-------------------I--------------------------------------------------------------N----------T---------------- 305
A.CD.97.97CD_KCC2          -------M-------------S-------------------------------T---N..-------------------I---R---------------------------------------------------------AN----------T---------------- 305
A.CD.97.97CD_KTB13         -------M---------------------------------------------T----..--------R----------I--------------------------------------------------------------G-------------------A-------- 305
B.AR.04.04AR151516         ---------------------------------------------------------..-----------------------------------------------------------------------------------Y-------------T---------------- 310
B.AU.87.MBC925             --------------------------------------------------------IVM----------------------------G-----------------------------------------------------Y-------------T---------------- 307
B.BO.99.BOL0122            ---------------------------------------------------------..-------------------------------------------------------------R---------------------------------T---------------- 306
B.BR.03.BREPM2012          -------------------------------------R-------------------..---------------------------------------------------------------------------------------------------T---------------- 309
B.CA.97.CANB3FULL          ---------------------------------------------------------..-L--D----------------------G-------------------------------------------------I-----L-------------------------- 308
B.CN.05.05CNHB_hp3         I----------L-----------------------------------------M----..-----------------------------------------------------------------------------------------------R-----------D--- 305
B.CO.01.PCM001             ---------------------------------------------------------..-----------------------------------------------------------------------------------------------------V-------- 305
B.GB.83.CAM1               ---------------------------------------------------------..------------------------------------------------------------------------------------K------------T---------------- 305
B.GB.86.GB8                -----------L---------------------------------------------..----------------------------------------------------------S-----------------------KE-------------------------- 311
B.GE.03.03GEMZ010          -----------------------D---------------------------------..------------------------------------------------------------------------------------KE-----------T---------------- 305
B.IT.05.SG1                ---XE---*-XMV---------------------------X------*-----X----..-------X--X---X-X----X-X----------X------X---------------------------------X-------------------X---------------- 301
B.JP.05.DR6538             -----------L---------------------------------------------..------------------------------------------------------------------------------------K------------V---------------- 305
B.KR.05.05CSR3             -------------------------I-------------------------------..------------------------------------------------------------------------------------K---------------------------- 305
B.NL.00.671_00T36          -----------L---------------------------------------------..-L------------------I--------------------------------------------------------------K------------T---------------- 315
B.RU.04.04RU128005         -------------------------------S-------------------------..----------------------------------------------------------SX----------------------K---------------------------- 305
B.TH.00.00TH_C3198         -----------L---------------------------------------------..------------------------------------------------------------------------------------K------------V---------------- 305
B.UA.01.01UAKV167          ---------------------------------------------------------..--------R---D---------------------------I-----------------------------------------K---------------------------- 312
B.US.04.ES10_53            ---------------------------------------------------------..--R--------------------------------------------C-------------------------------------------------------------- 305
B.US.99.PRB959_03          -------M---L---------------------------------------------..------------------------------------------------------------------------------------P---------------------------- 305
C.AR.01.ARG4006            -------M---L-----------------------------------------T---E..-------T-------------------------------------------------------------------------------------------------------- 301
C.BR.04.04BR013            A------M---L-----------------------------------------T---D..-------------------I------N-NR---------------------------------Q-------------------S----------K---------------- 309
C.BW.00.00BW07621          -------M---L----------------------------R------------TA--D..-------T-------------------------------------------------------------------------N-------------------------- 301
C.CN.98.YNRL9840           -------M---L-----------------------------------------TA--D..-------T-----------I-------------------------------------------------------------Y-G-------------------------- 301
C.ET.02.02ET_288           -----------L-----------------------------------------T---E..----------------------------------------------------------------------------------S----------T---------------- 307
C.GE.03.03GEMZ033          -------M---L-----------------------------------------TA--E..-------T-------------------------------------------------------------------------------------------------------- 301
C.IL.99.99ET7              -------M---L-----------------------------------------TA--E..-------T-------------------------------------------------------------------------S-------------------------- 301
C.IN.99.01IN565_10         I------M---L-----------------------------------------TA--D..-------T-----------I-----E--------------D---------------------------------------Y---------------------------- 301
C.KE.00.KER2010            -------I---L-----------------------------------------T---E..-------E------------------------------------------------------------------------KN----------R---------------- 301
C.MM.99.mIDU101_3          -------M---L-------------------------R---------------TA--E..-------T-----------I------------------------------L-----------------------------Y---------------------------- 301
C.MW.93.93MW_965           -------M---L-------------I---------------------------TA--E..-------T-------------------------------------------------------------------------N-------------------------- 302
C.SN.90.90SE_364           -------M---L-----------------------------------------TA--E..-------T-------------------------------------------------------------------------G-------------------------- 297
C.SO.89.89SM_145           -------M---L-----------------------------------------TA--E..-------T-------------------------------------------------------------------------G----------T---------------- 312
C.TZ.02.CO178              -----------L-----------------------------------------MA--E..-------T-------------------------------------------------------------------------G----------T---------------- 301
C.UY.01.TRA3011            -------M---L-----------------------------------------TA--D..-------T-----------------------------------------------------------------------------------T---------------- 301
C.YE.02.02YE511            -------M---L-------------------------R---------------T---E..D------T------------------------------------------------------------------------KE-------------------------- 299
C.ZA.04.04ZASK164B1        -----------L---------------------------------S-------T---Q..D------T-------------------------------------------------------------------------S-------------------------- 301
C.ZA.05.05ZAPSK240B1       -------M---L-----------------------------------------TA--D..--R----T-----------I-------------------------------------------------------------G----------V---------------- 301
C.ZM.02.02ZM108            -----------L-----------------------------------------TA--E..-------T------------------------------------------------------Q---------------------------------------------- 296
D.CD.83.ELI                -----------------------------------------------------T----..D--------R----------I---------------------------------------------------------------------------S-------------- 304
D.CM.01.01CM_0009BBY       ---------------------------------------I-------------T----..--------R---------------------------------------------------------------------K------------T---------------- 306
D.KE.01.NKU3006            -----------------------------------------------------T----..-------------------I-----------------------------------------------------------------------------T---------------- 304
D.KR.04.04KBH8             --------------------------------V----R---------------T----..--------R------------------------------------------------------------------------------------------------------- 304
D.TD.99.MN011              -----------------------------------------------------T----..--------R----------I--------------------------------------------------I----------A------------------------T-------- 306
D.TZ.01.A280               -----------------------A-----------------------------T----..--------R----------I---------------------------------------------------------------K------------T---------------- 306
D.UG.99.99UGK09259         -----------------------------------------------------T----..-------TR----------I------------------------------I-------------------------------H-------------T---------------- 304
D.YE.01.01YE386            ----------------------------R------------------------T---A..D--------R----------I---------------------------------------------------------------H-E-----------T------T-------- 304
D.YE.02.02YE516            -----------------------------------------------------I----..--------R----------I--------------------------------------------------------------S----------T---------------- 306
D.ZA.90.R1                 -----------L-----------------------------------------T----..-------------------I--------------------------------------------------------------C---------------------------- 304
F1.AR.02.ARE933            ---F-I-M--------H------K------------------------L---S---I..-----------------X---------------------------------------------------------------------------------------------- 305
F1.BE.93.VI850             -------M-----------V---------------------------------T---L..----------------------------S--------K---------------------------------------------K--K----------V---------------- 305
F1.BR.01.01BR125           -------M---------------N-------------R---------------T----..--------R---------------------------------------------------------------------------------------V---------------- 305
F1.ES.x.P1146              -------M---------------------------------------------T---A..D--------------------------------------------------------------------------------KE---F-------L---------------- 303
F1.FI.93.FIN9363           ---V---M---------------------------------------------T----..D--------R-----------------------------------------------------------------------K------------V---------------- 304
F2.CM.02.02CM_0016BBY      -------M---------------------------------------------T----..----------------------------------------------------------------------------------KE-------------------------- 305
F2.CM.95.MP257             -------M---------------------------------------------T----..---------------------------------------------------------------R------------------KE-------------------------- 305
F2.CM.97.CM53657           -------M-------------S-------------------------------T----..----------------------------------------------------------------------------------KE-------------------------- 305
G.BE.96.DRCBL              I------M-----------------------------R---------------T---N..-------------------I---------R--------------------P------G---Q-R------------------N----------T-----------..#- 301
G.CM.01.01CM_4049HAN       I------M---------------K-------------R---------------T---N..D------------------I-------------------------------------------R---------------------------------------------- 304
G.CU.99.Cu74               I------M-------------------------------I-------------TD--A..D------------------I--------------------------------------------------------------K------------T---------------- 301
G.ES.99.X138               I------M--K------------------------------------------T----..--R----------------I------------------------------------S-----------------------S----------T---------------- 303
G.GH.03.03GH175G           I------M-----------------------------RI--------------T---K..D--------------------------------------------------------------R------------------S-------------------------- 305
G.KE.93.HH8793_12_1        I------M---------------------------------------------T----..--R----------------I---R----------------------------------------------------------S----------T---------------- 303
G.NG.01.01NGPL0669         I------M---------------------------------------------T----..--------R----------I--------------------------------------------------------------S-------------------------- 304
G.PT.x.PT2695              I------M---------------------------------------------T----..--R----------------I------------------------------------S-----------------------S----------T---------------- 304
G.SE.93.SE6165             I------M-----------------------------R---------------T---K..--E----------------I-----------------------------------------------------------------------------V-------- 302
H.BE.93.VI991              -------I---M-----L-----------T-----------------------T---L..-------------------I------N--R------------------------------------S-----G-------H-------------T---------------- 303
H.BE.93.VI997              -------I-----------------------------R---------------T---M..-------------------I-------------------------------------------------S--------------K---------------------------- 303
H.CF.90.056                -------I---------------------------------------------T----..--------R------S--I-------------------------------------------------S--------------KE-------------------------- 304
J.CD.97.J_97DC_KTB147      -------M---L---------------------------I------R---E--TQ--E..---------------------------------#-----------------------------------------------Y---------------------------- 302
J.SE.93.SE7887             -------M---L---------------------------I-------------TQ--A..--E-----RV-------------------------------------------------------------------------Y---------------------------- 303
K.CD.97.EQTB11C            -------M-------------------------------------------------..-------------------------------I------------P------------------------------------K---------------------------- 305
K.CM.96.MP535              -----------------------------------------------------T----..-----------------------------------------------------------------------------------K-------------------V-------- 305
```

```
                              protease end p66, p51 RT start                    M41L           K70R
                                                                                               D67N                 D110 catalytic site
B.FR.83.HXB2                  VNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICT..EMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLP 305
01_AE.CF.90.90CF11697         -------M---------------D----T--------------*-----*--T---K..---E---------------------E----------------------------V--------------------------Y-G--------------------------- 303
01_AE.CN.05.FJ051             -------M---------------D----T-----------------------A---K..---E-------E-----------------------------------------------------------------------S------------------V-------- 305
01_AE.CN.06.FJ054             -------M---------------D------------S----------R----T---K..---E-------------------------------------------------------------------------------S-----------L--------------- 307
01_AE.HK.x.HK001              -------M---------------D----------------------------T---K..---E----------------------------------------------I-------------------------------KX--------------------------- 302
01_AE.TH.01.01TH_R2184        -------M---------------D----T--------R--------------T---K..-----------------------------------------------------------------------I-----------S-----------T--------------- 305
01_AE.TH.02.OUR769I           I------M---------------D-------------R--------------T---K..---E-------------------------------------------------------------------------------S--------------------------- 305
01_AE.TH.90.CM240             -------M---------------D----T-----------------------T---K..---E-------------------------------------------------------------------------------S--------------------------- 310
01_AE.US.00.00US_MSC1164      I------M---------------D----T-R---------------------T---K..---E-------------------------------------------------------------------------------N--------------------------- 305
02_AG.CM.02.02CM_4082STN      -------M--------------------------------------------KD--S..--------------------I-------------------------------------------------------------K------------V--------------- 308
02_AG.EC.x.ECU41              -------M---------L----------------------------------T----..D---------------------------------------------------------------------------------K------------V--------------- 305
02_AG.GH.03.GHNJ196           -------M--------------------------------------------TD---..------Q-------------------------------------P------------------------------------H-------------V--------------- 306
02_AG.NG.01.PL0710            -------M------------------------------I-------------T---M..----------------------------------------------------------------------------------K------------V--------------- 305
02_AG.NG.x.IBNG               -------M--------------------------------------------TD---..----------------------------------------------------------------------------------K------------V--------------- 305
02_AG.SE.94.SE7812            -------M--------------------------------------------TD---..---------R----------I------N--R-------------------I-------------------------------K---------------------------- 305
02_AG.SN.98.MP1211            -----K-I-------------------L--------------------L-H-TD--A..-------------------------R---------------------------------------------------------------------V--------------- 305
03_AB.RU.97.KAL153_2          -------M---L----------------T-----------------------TD--K..------------------------------------G---------------------------------------------Q------------T--------------- 305
04_cpx.CY.94.CY032            -------M---L----------------------------------------T----..D-------------------I-------------------------------------------------------------PE-----------T--------------- 304
05_DF.BE.x.VI1310             ----------------------------------------------------T----..D--------R----------I---------R-----------------------------------I---------------KE--------------------------- 305
06_cpx.AU.96.BFP90            I------M--------------------------------------------T---K..------------------------------------------------------------------------------I----K-----------T--------------- 307
08_BC.CN.97.97CNGX_6F         -------M---L----------------------------------------TA--D..--------T----D------I---R----S----------------------------------------------------K------------V--------------- 301
09_cpx.GH.96.96GH2911         -------M------------------------------I-------------T---A..----------------------------------------------------------------------------------K------------V--------------- 304
10_CD.TZ.96.96TZ_BF061        ----------------------------------------------------T---K..---------R----------I------------------------------------------------------------Y----------------------------- 304
11_cpx.GR.x.GR17              -------M--------------VD----------------------------T----..--------------------------------------------------I----------------------------L---S-----------L--------------- 304
12_BF.AR.99.ARMA159           -----------L----------------------------------------T----..-----------------------------------------------------------------------------------A----------CV------T-------- 305
13_cpx.CM.96.1849             -------M--------------------------------------------T---K..----------------------------------------------------------------------------------K------------V--------------- 312
14_BG.DE.01.9196_01           I------M--------------------------------------------T----..D--R----------------I-------------------------------------S-----------------------KS-----------T--------------- 304
15_01B.TH.99.99TH_MU2079      -------M---L----------------------------------------T---K..---E----------------I-------------------------------------------------V------------S--------------------------- 304
16_A2D.KR.97.97KR004          -------M-V-L----------------------------------------T---K..--------------------------------------------------------------M---------R--------H-------------T------V-------- 305
17_BF.PE.02.PE02_PCR0155      ----------M-----------------------------------------T----..----------------------------------------------------------------------------------KE--------------------------- 305
18_cpx.CU.99.CU76             -------M--------------------------------------------T----..--------------------I-------------------------------------------------------------PE--------------------------- 303
19_cpx.CU.99.CU7              ----------------------------------------------------T----..D-------------------I--------------------------------------------------------------S--------------------------- 304
20_BG.CU.99.Cu103             I------M---------------D---------------------------------..-----------------------------------------------------------------------------------------------------W--------- 301
21_A2D.KE.91.KNH1254          ----------------------------------------------------T----..---------R----------I------------------------------------------------------------S-------------T--------------- 303
22_01A1.CM.01.01CM_0001BBY    -------M--------------------A---------I-------------T---K..--------------------I------------------------------------------------------------Y-G-----------T--------------- 305
23_BG.CU.03.CB118             T------M-------------------------------------------------..-L-------------------------N-------------------------------------------------------E--------------------------- 301
24_BG.CU.03.CB378             I------M-------------------------------------------------..----------------------------------------------------------------------------------KE-----------V--------------- 296
24_BG.CU.03.CB471             I------M-------------------------------------------------..--------------------------------------------------------------------------------------------------------------- 301
25_cpx.CM.01.101BA            I------M--------------------------------------------T---N..D--------R----------I-------------------------------------------R-----------------KE-----------T---A----------- 304
27_cpx.CD.97.97CDKTB49        ----------------------------------------------------T----..---------R-------------------------------------------------------------------------N-----------V--------------- 303
28_BF.BR.99.BREPM12609        -------M-------------------------------------------------..----------------------------------------------------------------------------------KE--------------------------- 305
29_BF.BR.02.BREPM119          -------M-------------------------------------------------..-------------------------------------------------------------Q--------------------K-------------------T-------- 305
31_BC.BR.02.110PA             -------M---L----------------------------------------TA--D..--------E-------------------------------------------------------------------------KS--------------------------- 301
32_06A1.EE.01.EE0369          -------M---------L-------------------R-----------R--T----..--------------------I-------------------------------------------------------------K---R------------------------ 305
33_01B.MY.05.05MYKL007_1      -------M--------------------------------------------I----..----------------------------G---------------------------------------------------------------------------------- 305
34_01B.TH.99.OUR2478P         -----------L----------------------------------------A---A..--------------------I------------------------------------------R------------------K---------------------------- 305
35_AD.AF.05.05AF095           -------M---L-----------N----R-----------------------T----..--------------------I--------------------------------------------------------------------------T--------------- 304
36_cpx.CM.00.00CMNYU830       I--V---M--------------------A-----------------------T---A..------------------------R---------------------------------------R-----------------K---------------------------- 305
37_cpx.CM.00.00CMNYU926       -------M---------------N----S---------------S-------T---N..---------RV---------I-------------------------------------------R------------------S-----------T--------------- 303
38_BF1.UY.05.UY05_4752        -----------L----------------------------------------T----..-----------------------------------------------------------------------------------------------V--------------- 305
39_BF.BR.03.03BRRJ103         ---V--D-M--M--------------------------------------------A..FL-E-----------------------GG-----VA------------------------I--N--------------I----E-----------V--------------- 310
40_BF.BR.04.04BRRJ115         ------D----L---------------------------R------------T----..-L-------------------------N--R---IT----------------------G---N-------------------KE-----------L--------------- 305
42_BF.LU.03.luBF_05_03        -----------L----------------------------------------K----..------------------------------------------X---------------------------------------K------------T--------------- 309
43_02G.SA.03.J11223           I------M-------------------------------------A------T---L..--------------------I-----------------X----------------------X---------------------S-----------T---------C----- 304
U.CD.83.83CD003               -------M--------------------------------------------T---L..-----------------------------------------------------------------------------------N--------------------------- 306
U.CD.90.90CD121E12            -------M--------------------------------------------T----..-----------------------------------------------------------------------------------N--------------------------- 305
U.GR.99.GR303                 -------V--------------------A---------I----XX-X-----T----..------------------------------R---------------------------------------------------K------------T--------------- 304
U.NL.95.U_NL_95_H10986_D1     -------M-----------------------------------------R--TD---..----------------------------------------------------------------------G-----------PG--------------------------- 305
CPZ.CD.90.ANT                 -------V-CLL----------KV-------E-----R------SK---E--K---D..KL-A-N---R---D------I-------TS--------------------I-----------Q---------------I---Q------------V---------C----- 305
CPZ.CM.01.SIVcpzCAM13         -------I--------V----S---------------RI-----SA------T---Q..--------------------I--------S--------------------------------E---------------C---K------------V--------------- 304
CPZ.CM.05.SIVcpzEK505         -------F-----------------I-----------R--------------T----..---------R----------I---------------------------------------------------------C----N-----------V--------------- 311
CPZ.CM.05.SIVcpzLB7           -------I--------------------Q-----------------------T----..---Q-----R------------------------------------------------------R-------------C----N--------------------------- 308
CPZ.CM.05.SIVcpzMB66          -------I--------------------S--------R-----------R--T----..---------RV---------I-----------------------------------------Q---------------C----N-----------V--------------- 308
CPZ.CM.05.SIVcpzMT145         -------V--KL-----------D--K-F---------------SK---E--T---K..--------------------I-------------------------E-------------------------------C---P----F----------------------- 304
CPZ.GA.88.GAB1                -------I---L----V----S----------------------SA------T---Q..--------------------I---------------------------------------------------------C---K-------------------V-------- 303
CPZ.TZ.00.TAN1                I------I-K-L------------V-K-Q--E------------SK---E--T---K..TL-------AV---------I-------TS--------------------L----------R-RNM------------I---P------------L--N---K-F------ 301
CPZ.US.85.CPZUS               -------I-------------S-K----R--------R-------A------T---Q..--------TRV---------I-------------------------------------------R-------------C---KE-----------V--------------- 307
CPZ.CM.x.SIVcpzDP943          -------I---------------G-------------R--------------T---Q..--------------------I-------GE------------R-----------------------------------C----S-----------V--------------- 312
CPZ.CM.x.SIVcpzMB897          -------I-------------T---------------R--------------T----..---------R----------I---------------------------------------------------------C----E-----------V--------------- 313
N.CM.02.DJO0131               -------F-----------------------------R-------A---E--RQ--A..---Q-----R----------I----------------------------------------Q----------------C----------------T--------------- 310
N.CM.04.04CM_1015_04          -------F---------------K---------------------A---E--R----..---------R----------I---R-------------------------------------Q---------------C----------------V--------------- 312
N.CM.04.04CM_1131_03          -------F-----------L---K-------------E-------A---E--R----..---------R----------I-----------------------------------------Q---------------C----------V-----V--------------- 312
N.CM.95.YBF30                 ---------------------------------------------T---E--R----..---------R----------I-----------------------------------------Q---------------C---K---------------------------- 312
N.CM.97.YBF106                -------F-----------------------------R-------A---E--R----..---------R----------I-------------------------E---------------Q------X--------C---K---------------------------- 312
O.BE.87.ANT70                 -------I--GL-----------AP-------------------SK---E--TA--Q..---Q-----R----------I-------G-----------------E-----------G---Q-Q-------------C---P------------V--------------- 301
O.CM.91.MVP5180               ---L---I--GL-----------AP-------------------SR---E--TA--Q..---Q-----R----------I-------------------------------------G---QRQ-------------C---P------------V------V-------- 301
O.CM.96.96CMABB637            -------I--GL-----------AP-------------L-----SR---E--TA--Q..---Q-----R-------X--I-------------------------------------X---QRQ-------------C----------------V------V-------- 303
O.CM.98.98CMA104              -------I--GL---------S-AP---R---------------SK---E--TA--Q..---Q-----RV---------I----------------------K--------------G-----Q-------------C---P------------V------V-------- 301
O.CM.99.99CMU4122             -------I--GL-----------AP---X---------I-----SK---E--TA--Q..---Q-----RV---------I--------S----------------------------G---Q-Q-------------C---P------------V--------------- 301
O.FR.92.VAU                   ---L---I--GL-----------TP-------------------SR---E--T---K..-L-Q-----R----------I-------G-----------------------------G---Q-Q-------K-----C---PE-----------V------V-------- 302
O.SN.99.SEMP1299              -------I--GL-----------AP-------------I-----SK---E--TA--Q..---Q-----R----------I-------------------------------------G---Q-Q-------------C---P------------V--------------- 301
O.US.99.99USTWLA              ---L---I--GL----X------TP-------------------SR---E--TA--Q..-XQ-----R---X------I-----X----X-XX-------X-----XXX-XX----XX--QRQ-------------C---P---------X--V------V-------- 304
```

```
                                                      K219Q
                              polymerase motif    T215Y
B.FR.83.HXB2               QGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAELELAENREILKEPVHGVYYD 475
A1.GE.99.99GEMZ011         -----------C----------SK--E-----------------------------A--MS--F------------------------------M-LD----------------------------K--------A----DIVT----------------R--------- 475
A1.KE.00.KER2008           -----------C----------SK--E-I--------------------A------A---S--F------------------------------E----E--------------------------K--------A----DIVT-----------------D-------- 475
A1.KE.00.KNH1144           ----------------------SK--E-I--------------------A------A---S--F------------------------------E--N-E---------------------A----KH--R----A----D-VT------------K------------- 475
A1.KE.00.KSM4024           ----------------------AK----I--------------------A------A---N--F-----------------------R---------Q----------------------------K--------A----D-VT-----------------D-------- 475
A1.KE.00.MSA4069           ----------------------SK--E-I--------------------A-V----A---S--------------------------------------E---------------------A----K--------A----DIV------------------D-------- 475
A1.KE.00.NKU3005           ----------------------LK--EMI--------------DL----A-V----N---S--F------------------------------E----E---------------------A-------------A----DTVT-----------------D-------- 475
A1.RU.00.RU00051           -------S---C---R------SK--E------------------------V----A---S--FAX--X-------------------------M--D-------------------------------------A----DIVT-------------------------- 475
A1.RW.93.93RW_024          ----------------------SR----I--------------------A------A---S--F-----------------------------------E---------------------A----K--------A-----IVT-------------------------- 475
A1.SE.95.SE8891            --------------I-------V---E-I--------------------A-V----A---S--F------------------------------Q--N-EN--------------------V----K--------A----DIVT-----------------D-------- 470
A1.SE.95.UGSE8131          ----------------------SK--E-I----------------E----------A---S--FF---Q-------------------------K----E---------------------A----K--------A----DIVT-----------------D-------- 469
A1.TZ.01.A173              ----------------------SK--E-I---------------------------N---K--F---------------------------------D-E---------------------A----K--------A----DIVTM------------------------- 475
A1.UA.01.01UADN139         -------S--------------LK--E-----------------------------A---S--F------------------------------M--D-------------------------------------A----DIVT--------------L----------- 475
A1.UG.92.92UG037           -----------A----------SK--------------------------------E---K--F------------------------------E----E---------------------A----K--------A----DIVT-----------------D----A--- 476
A1.UG.99.99UGA07072        ----------------------SK--E-I----------------E-----V----A---S--F------------------------------Q--N-E--------------------------K--------A-T--DIV------------------D-------- 475
A2.CD.97.97CDKTB48         ------------------D---AR--EM-----------------D---A------A------F------------------------------K--------------------------A------------------DIV---R------E-------D-------- 468
A2.CY.94.94CY017_41        ----------------------SK-TELI----------------S---V------A---K--FY-----------------------------K--------------------------A----K--------A----DIVT--K------E-------T-------- 475
A.SN.01.DDI579             -----------Y----------S---E-I------------N--------------A---K--F------------------------------Q--------------------------A-------------A----DIVT-----------------D----A--- 474
A.SN.01.DDJ369             -----------Y----------S---E-I---------------------------A---K---------------------------------Q--D----------------------------K--------A----D-VV-----------------D-------- 474
A.SN.96.DDJ360             -----------Y----------L---E-I---------------------------A---K--F------------------------------Q--------------------------A----K---R----A----DIVT-----------------D----A--- 474
A.CD.02.02CD_KTB035        -----------C----------S---------------------------------N---K--F------------------------------Q--D----------------------------K--------A----DIVT----------------------E--- 475
A.CD.97.97CD_KCC2          ------------T---------SK--E----------------------A------A---K--F------------------------------Q-------------------------------K---R-I--A----DIVT-------------------------- 475
A.CD.97.97CD_KTB13         ----------------------AK---M---------------------A------D---K--F------------------------------Q-------------------------------K--------A------VE-------------------------- 475
B.AR.04.04AR151516         ---------------------------------------------------V----K---K--F---------------------------------------------------------G----K--------A-S----V---K----------------------- 480
B.AU.87.MBC925             -----------C--------------E------------------------------------F---------------------------------------------------------A----K-----I--------IV---K----------------------- 477
B.BO.99.BOL0122            -----------C------------------------------------------------K------------------------------------H----------------------------K------------------------------------------- 476
B.BR.03.BREPM2012          ------------------D-----------------------------------------K--F------------------------------M--------------------------A----K---------------V--------------------------- 479
B.CA.97.CANB3FULL          -------S---------------------------------------------------WK--FY--------------------------------------------------------A----K------------------------------------------- 478
B.CN.05.05CNHB_hp3         -----------C-------------------------------------A-V----E------F---------------------------------------------------------A----KE----------------------------------T------- 475
B.CO.01.PCM001             -----------C---------------------------------------V----E---K--FA-----------------------------I-------------------------------K---------------V--SA----------------------- 475
B.GB.83.CAM1               ----------------------------------------------------------------------------------------------M--------------------------A----K---------------V--------------------------- 475
B.GB.86.GB8                --------------------------E------------------------------------FS----------------------------------------------------------------------A-------T-------------------------- 481
B.GE.03.03GEMZ010          --------------------------------------------------------K---S--F-----------------------------V---------------------------A--------------------VS----------------R--------- 475
B.IT.05.SG1                ---------C-----R--------------------------------SI-------X-----F-----X---X--------------------K---X---------------X-----------K--------------I--------------------S---A--- 471
B.JP.05.DR6538             -----------C--------------E---------------------------------G--F---------------------------------------------------------A----K--------A-----I-V----------------------A--- 475
B.KR.05.05CSR3             --------------------------N----------------------A---D------K--F--------------------------------------------------D----------AK--------A----------K----------------------- 475
B.NL.00.671_00T36          --------------------------------------------M-----R---------K--FD-----------------------------------X-X----------X---------------------------IV---A----------------------- 485
B.RU.04.04RU128005         -------S-------------------------------------------------------F----------------------XX--X---K--D-------------------------------------A----------QG---------------------- 475
B.TH.00.00TH_C3198         ----------------------------I-------------------------------K--F------------------------------I--V-----------------------G--R-KE--------------V--------------------------- 475
B.UA.01.01UAKV167          ---------------R-----------------------------------------------F------------------------------K--------------------------A----K---------------V--------------------------- 482
B.US.04.ES10_53            ---------------------------L-----------------E-------K---------F------------------------------E--------------------------A----K---R-----------V--------------------------- 475
B.US.99.PRB959_03          ----------------------E---E-------------------H---------E---K--F------------------------------E----E--------------------------K----------S----------------------R--------- 475
C.AR.01.ARG4006            -------S----X---------A---E----------------------A---K--E---K---------------------------------Q-------------------------------K-------------DIV---A----------------------- 471
C.BR.04.04BR013            ----------------------AK--E----------------------A------E---K--FI---Q-------------------------Q---------------------------------------------DIV--------------------------- 479
C.BW.00.00BW07621          ---------------R------T---E----------------------V------R---K--F------------------------------Q----ED----------------------------------A----DIV--------------------------- 471
C.CN.98.YNRL9840           ---------------R------A---E----------------------A------E------F------------------------------Q----------------------------------------A----DIV--------------------------- 471
C.ET.02.02ET_288           ----------------------AK--E-----------------------------E---K--F------------------------------Q----E--------------------------K--------A----DIV--------------------------- 477
C.GE.03.03GEMZ033          ----------------------TR---M---------------------A------E---K--F------------------------------Q---EE-----------------------------------A----DIV--------------------------- 471
C.IL.99.99ET7              ---------------R------A---E----------------------A------G------F------------------------------Q----------------------------------------A----DIV--------------------------- 471
C.IN.99.01IN565_10         -----------A----------A---E----------------------A------A---K--F------------------------------Q-------------------------------K---R---------DIV--------------------------- 471
C.KE.00.KER2010            ----------------------AK--E----------------------A------E------F------------------------------Q--D--------------------------R--H------------DIV--------------A------------ 471
C.MM.99.mIDU101_3          ---------------R------A---E-----------------------------E---K--F------------------------------Q-------------R--------------------------A----DIV-----------------------A--- 471
C.MW.93.93MW_965           ---------------R------A--------------------------A------E-------------------------------------Q----------------------------------------A----DIV--------------------------- 472
C.SN.90.90SE_364           -------S--------------A---E------------------M---A-V-K--E---K--F------------------------------Q--D-E---------I-------------------------A----D-V--------------------------- 467
C.SO.89.89SM_145           -----------C---R------AK--E----------------------A------E---K--F------------------------------Q----------------------------------------A----DIV--------------------------- 482
C.TZ.02.CO178              ----------------------A---E----------------------A------E------F------------------------------Q---------------------------------------------DIV-----------------R--------- 471
C.UY.01.TRA3011            ----------------------A---EV---------------------A-V----E---K---------------------------------Q----------------------------------------A----DIV-----------------------A--- 471
C.YE.02.02YE511            ----------------------A---E------------------K---A-V----E---K--F------------------------------Q--D-------------------------------------A----DIV--------------------------- 469
C.ZA.04.04ZASK164B1        ----------------------A-----------------------L--E---Q--D---K--F---G--------------------------Q--T------------------------------------------DIV--------------------------- 471
C.ZA.05.05ZAPSK240B1       -----------C------D---AK--E------------------E---A------T---K--F------------------------------Q----------------------------------------A----DIV--------------------------- 471
C.ZM.02.02ZM108            -----------C----------TK--E----------------------A------E---K--F-----------------------------------E---------------------A-------------A-S--DIV-----------------------A--- 466
D.CD.83.ELI                --------------------------EM-------------------------K--E------F-R--------------------------S-K----E--------N--ER--------------------------------------------------------- 474
D.CM.01.01CM_0009BBY       -----------C--------------ELI--S-----------------S------E---K------------D--------------------K----E---------------------S----K--------A----------K-------------R--------- 476
D.KE.01.NKU3006            --------------------------EV---------------------A------E---K--F------------------------------K----E-------------------------------C---A----------K----------------------- 474
D.KR.04.04KBH8             --------------------------E----------------------A---K--E---K---------------------------------T----E---------------------A----K--------A----------K----------------------- 474
D.TD.99.MN011              -----------C------D-------EL---------------------A---K--E---K---------------------------------T----E--------------------------K---R----A-T----V---K-------------R--------- 476
D.TZ.01.A280               --------------------------E-----------------------------D---K--F------------------------------K----E---------------------S----K--------A----------K-------------R-T---A--- 476
D.UG.99.99UGK09259         --------------------------EV----------------------T-----K---K---------------------------------K----E---------------------A---------CI--A------V---A----------------------- 474
D.YE.01.01YE386            --------------------------EV---------------------I------E---K--F------------S-----------------K----E-------------------------------C---A------V--------------------------- 474
D.YE.02.02YE516            -----------C--------------EL-------------------------K--E-------------------------------------T--Q-EN-------------------------K---R----A----D-----R-------------R--------- 476
D.ZA.90.R1                 -----------C-------------------------------------A------E------F------------------------------K----E------------------------------------------V--------------------------- 474
F1.AR.02.ARE933            -----------C----------TK---M----------------------------K---T--F------------------------------Q-------------------K-----------K--------A----DIV---A----------------------- 475
F1.BE.93.VI850             -----------C----------MK--------------------------------E------F-----------------H------------Q--N-----------------------------P-------A----DIV---A-------K-----R--------- 475
F1.BR.01.01BR125           -----------C------D---AK--G----------------------A------E------F--------R---------------------Q--N--------------------------R-KN-------A----DIV---A----------------------- 475
F1.ES.x.P1146              -----------C------D---AK--E-----------------L----A------E---K--F------------------------------Q--N-------------------------------------A----DIV---T-------------------A--- 473
F1.FI.93.FIN9363           -----------C----------TR--------------------------------E---K--F------------------------------Q--D------D-------------X----------------A----DMV---A--N-------------------- 474
F2.CM.02.02CM_0016BBY      ----------------------AR--E--V-------------------A------E------F----------------------------A-Q--N-S------------------------R-KH-------A----D-V---A----------------------- 475
F2.CM.95.MP257             --------------I-------E---E----------------------A------E------F----------------------------A-Q--D-S------------------------R-KH------------D-V---A----------------------- 475
F2.CM.97.CM53657           ----------------------VK--E-----------------------------E------F----------------------------A-Q--D-S------------------------R-KH-------A----D-V---V-----M----------------- 475
G.BE.96.DRCBL              ----------------------T---E----------------------A------E------F------------------------------Q--N-EN-------------------------K-----I--A----DIVSM-A---M------------------- 471
G.CM.01.01CM_4049HAN       ----------------------AK--E----------------------A------K---E--F------------------------------Q--D-E--------------------------KH--R----A----DIV---A---M-----K------------- 474
G.CU.99.Cu74               --------------------------E-E--------------------A----------S--F------------------------------Q--D-E--------------------------K--------A----DIV--------------------------- 471
G.ES.99.X138               ----------------------IK--E----------------------A------K---S--F------------------------------Q--D-E--------------------------K--------A----DIV---A----------------------E 473
G.GH.03.03GH175G           --------------I-------TK-------------------------A-V----E---K--F------------------------------Q--D-E--------------------------KH--R----A-----IVS--A---M------------------- 475
G.KE.93.HH8793_12_1        -------P--------------IK--EM---------------------A--K---E------F------------------------------Q----E-----------------------------------A----DIV---A----------#------------ 472
G.NG.01.01NGPL0669         ----------------------T---EM---------------------A------N------F------------------------------Q--D-EN--------------------A----K-------------DIV-----------------R--------- 474
G.PT.x.PT2695              ----------------------IK--E----------------------A------K---N--F------------------------------Q--D-E-------------------------IK--------A----DIV---A----------------------E 474
G.SE.93.SE6165             ---------------R------AN--EM---------------------A------E---K--F------------------------------Q--D-E--------------------------TH-------A----DIVS--A---M---------R--------- 472
H.BE.93.VI991              --------------------------EVI--------------------E------A------F----Q------------------------VK-------------------------------K---X----A-----IV---K-------------------A--- 473
H.BE.93.VI997              --------------------------E-I--------------------A------A------F-----------------------------VK----------------------------------------A----D-V---K-------------R--------- 473
H.CF.90.056                ------------------A---E---EM---------------------A------A---K--F----------------------------TVK---------------------------N---K--------A----DI----K-------------R--I------ 474
J.CD.97.J_97DC_KTB147      -----------C----------AK--EV-----------------E---K------E------F-----------------------------VQ----E--------------------------K-------------DIVA--A----------------------- 472
J.SE.93.SE7887             -----------C------K---ER--E------------------E---R--K---E---K--F------------------------------Q----ED-------------------------K------K-A----DIV---R---------K------------- 473
K.CD.97.EQTB11C            -----------C----------RK---M-L-------------------A------E------F------------------------------Q--D----------------------F-----K--------V----DIV---A----------------------- 475
K.CM.96.MP535              -----------H----------IK--EM-------------------P--------E---K--F------------------------------Q--D----------------------------K--------V----DIV---A----------------------- 475
```

```
                                                
```

```
                                                                                                                                                  p51 RT end p15 RNase H start
B.FR.83.HXB2              PSKDLIAEIQKQGQGQWTYQIYQEPF.KNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGK.TPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEK.EPIVGAETFYVDGAANRETKLGKAGYVTNRGRQKVVTLTDTTNQKTELQAIYLALQ  642
A1.GE.99.99GEMZ011        -----V--------------------.---------KKGS----------AV---VA---------.----R---------A--M----------------------------.--------------------I-------D-------P--E--------H--H----  642
A1.KE.00.KER2008          -----T--------D-----------.---------K-S-------R--A-V---VVM--------.-----------------MD---------------------------.D--A-----------H------------D-------S--E--------H-------  642
A1.KE.00.KNH1144          ---E----V-----D-----------.---------Q-S----------A-V----AM-G------.----R------------RD---------------------------.D---------------------------D-------S--E--------H-------  642
A1.KE.00.KSM4024          --------------D-----------.---------R-S----------ADV---VVM--------.-----------------MD---------------------------.D--A----------------I-----I-D-------S--E--------H-------  642
A1.KE.00.MSA4069          --------------D-----------.---------K-S------------V---VVM--------.-----------------MD---------------------------.D-----------------N---------D-------S--E--------H-------  642
A1.KE.00.NKU3005          --------------D-----------.---------K-S-------R--A-V--RVVM--------.-----------------MD---------------------------.D--A------------------------D------ISX-E--------H--H----  642
A1.RU.00.RU00051          -----V--------D-----------.---------KKGS----------AV---VA---------.----R----R----A--M----------------------------.----------------------------DK------S--E--------H--H----  642
A1.RW.93.93RW_024         --------------D-----------.---------K-S-------R--A-V---VVM--------.-----------------MD---------------------------.D--E----Y-------------------D------IS-AE--------H--H----  642
A1.SE.95.SE8891           --------------D-----------.---------K-S-------R--V-V---VVM--------.-----------------MD---------------------------.D---------------------------D---------AE--------H-------  637
A1.SE.95.UGSE8131         --------------D-----------.---------KKS-------R--A-V---VVM--------.----------D--D---MD---------------------------.D--A-V----------------------D-------S--E--------H--H----  636
A1.TZ.01.A173             ------V-------D-----------.---------K-S----------A-V---VV---------.-----------------MD---------------------------.D--I------------------R-----D-------S--E--------H--Q----  642
A1.UA.01.01UADN139        -----V--------D-----------.---------KKGST---------AV---VA---------.----R---------A--M----------------------------.--------------------I-------D---K---P--E--------H--H----  642
A1.UG.92.92UG037          --------------D-----------.---------K-S----------A-V---VVM--------.-----------------MD---------------------------.D--A------------------------D-------S--E--------Y--H----  643
A1.UG.99.99UGA07072       --------------D-----------.---------KKS----------A-V---VVL--------.-----------------MD--------D------------------.D--S-V----------------------D-------S-DE-----A--H--L----  642
A2.CD.97.97CDKTB48        -----------------S--------.---------KRKST---------------A---------.I---R--------------W-------------------------T.---A------------------------D-----I-P--E--------H-------  635
A2.CY.94.94CY017_41       --------------D-----------.---------KR-ST----I-----------M--------.-----------------A----------------------------.---A------------------------D-----I-S--E--------H-------  642
A.SN.01.DDI579            --------------N-----------.---------K-SS------R--V-V---VV---------.----R------------M------Y---------------------.D--I------------------------D-------S-ME--------H--H----  641
A.SN.01.DDJ369            -----V--------D-----------.---------K-S-------R--V-V----V---------.-------------D----------C---------------------.D---------------------------D-------S--E--------H--H----  641
A.SN.96.DDJ360            --------------D-----------.---------KKSS------R--V-V---VV---------.-------------D---M------C---------------------.D---------------------------D-------S--E--------H--H----  641
A.CD.02.02CD_KTB035       -----V-------HD-----------.---------KK-S------R----V---VAV--------.----R---------A--MD---------------------------.D---------------------------D-------S--E--------H-------  642
A.CD.97.97CD_KCC2         --------------------------.---------KK-S------R--S-V---VA---------.----R---------A--MD---------------------------.D---------------------------D-------S--E--------H-------  642
A.CD.97.97CD_KTB13        ---E----------D-----------.---------KK-S------R--------VAL--------.----R---------A--MD---------------------------.D--A-V----------------------D-------S--E--------H-------  642

B.AR.04.04AR151516        ------------D-------------.---------X----------------N--A--G------.I---R--------DA--M----------------------------.---A------------------R-----D-------S----------X---H----  647
B.AU.87.MBC925            --------------------------.----------T----------------------------.I---------------------------------------------.---I-------------------------K------S-------------------  644
B.BO.99.BOL0122           --------L----E-------F-D--.----------AK------I----------A---------.----------------------------------------------.D--E-V------------------------------S-------------------  643
B.BR.03.BREPM2012         ---E----------------------.---------------------------------------.I---R---------A--I----------------------------.----------------------------D-------P--------------H----  646
B.CA.97.CANB3FULL         --------------R-----------.--------------------------KQ-----------.--------------A-------------------------------.----------------------------D-------SIP------------N----  645
B.CN.05.05CNHB_hp3        -------------L----------QY.-----------------------------A---------.--------------A-------------------------------.---E-------------------------K---R----N---------H--H----  642
B.CO.01.PCM001            --------L--------S--------.-------------T---------------A---------.--------------A-------------------------------.-----------------D--S-------D-------S--------------H----  642
B.GB.83.CAM1              --------L-----------------.---------K---T---------------A---------.-------------DA--ID---------------------------.-------------------R--------D-------P-------------------  642
B.GB.86.GB8               -----V-------L------------.---------K-------------------A---------.--------------A--M----------------------------.---I------------------------D-------P--------------H----  648
B.GE.03.03GEMZ010         --------V---E-------------.-----------------------------A---------.I---R------------A---------------------------T.----------------------------D-------S--------------H----  642
B.IT.05.SG1               ----X--------------X-F--X-.---X--------------I----------SX--------.----R---------A-------------------------------.---E---------------R--I-----T---R---S-P---I--------H----  638
B.JP.05.DR6538            -----V--V-----------------.-----------K-----------------A---------.--------------A--M----------------------------.---------------S------------D-------H-A------------H----  642
B.KR.05.05CSR3            -----V-------L------------.----------V------------------AQ-------R.I-------------I--M----------------------------.---------------S----A-----I-D---------------------------  642
B.NL.00.671_00T36         -T------------------------.---------KV------------------A---------.--------------A-------------------------------.----------------------------D------IPIK------------H----  652
B.RU.04.04RU128005        -X------------------------.---------K---T---------------A---------.----R---------A-------------------------------.---E--------------------------------S--------------H----  642
B.TH.00.00TH_C3198        -------------L-----------Y.-----------------------------A---------.--------------A----------V--------------------.---E-------------------------K---------------------H----  642
B.UA.01.01UAKV167         --------V---E-------------.---------K-------------------AI--------.----R------------M---------------------------T.----------------------------D-------S--------------H----  649
B.US.04.ES10_53           --------------------------.-------------------R---------A---------.I-------------A-------------------------------.-----------------------------K---------------------H----  642
B.US.99.PRB959_03         --------L---EY------------.------------------I----------A---------.----R-----------------------------------------.---L------------------------D-------S-P------------R----  642

C.AR.01.ARG4006           ------------E-------------.---------K--T------R---------AQ---I----.----R---------S---D---------------------------.---T----------------M-------D-----I-A--------------Q----  638
C.BR.04.04BR013           -------------ED-----------.---------K--T------R---------AQ--------.----R---------A---D---------------------------.---P------------------------------I-S--E-----------K----  646
C.BW.00.00BW07621         -------D-----ND-----------.---------K--T----------------SR--------.----R-------------D---------------------------.D-L--V-----------D--I-------DK--R--I---E-----------Q----  638
C.CN.98.YNRL9840          S-------------D---------S-.---------K--T----------------AM--------.----R-------------D---------------------------.D--A-V--------------I-------D---K-I-S--E------------I---  638
C.ET.02.02ET_288          -------------DD-----------.---------KR-T-------L--------AL--------.I---R-------------D---------------------------.---A----------------I-------D-----I-S--E-----------Q----  644
C.GE.03.03GEMZ033         -------------HD-----------.---------K--A----------------AM--------.----R-------------D---------------------------.---A-V---------S------------D-----II-I-E-----A-----Q----  638
C.IL.99.99ET7             -------------DD-----------.---------K--T---------VG-----AM-G------.I---R-------------D--------D------------------.-----V--------------I-------D-----I-S--E-----A-----QI---  638
C.IN.99.01IN565_10        --------L----HD-----------.---------K--T---------A------AL--------.----R-------------D---------------------------.-----V-----------D--I-------D-----I-S--E----------------  638
C.KE.00.KER2010           -------------HD-----------.---------K--T----------------A--------R.----R-------------D---------------------------.D--A----------------V-------D-------SISE-----------Q----  638
C.MM.99.mIDU101_3         --------------D-----------.---------K--T----------------AM--------.----R-------------D---------------------------.D--A-V--------------I-------D---K---S--E------------I---  638
C.MW.93.93MW_965          ---E---------HD-----------.---------K-KS-----------T----AM--------.----R-------------D---------------------------.---A----------S-----I-------DK------S--K-----A-----Q----  639
C.SN.90.90SE_364          ---E---------HD-----------.---------K--T----------------A---------.----R-------------D---------------------------.D--I----------------I-------D-------SI-E-----------Q----  634
C.SO.89.89SM_145          -------------HD-----------.---------K--T----------------AQ--------.----R-----------------------------------------.---A----------------M-----I-D-----I-S--E-----------QI---  649
C.TZ.02.CO178             -------------HD-----------.---------K--T----------------AM--------.----R---------A---D---------------------------.---A------------------------DK----IIS--E-----------Q----  638
C.UY.01.TRA3011           ------------D-------------.--------TK--T------R---------AL---I----.----R---------A---D---------------------------.---A----------------M-------D-----I-S--E-----------Q----  638
C.YE.02.02YE511           -------------YD-----------.---------KI-T----------------AM--------.I---R-------------D---------------------------.--M--V--------------M-------D-----I-S-NE-----A-----H----  636
C.ZA.04.04ZASK164B1       -------------HD-----------.---------K--T----------------AL-------R.----R-------------D---------------------------.D-MI-V----------------------D------IS--E-----------L----  638
C.ZA.05.05ZAPSK240B1      -------------E------------.---------K--T----------------AI--------.----R-----------------------------------------.---A-----------TKD--I-------D-----IIS--E-----------Q----  638
C.ZM.02.02ZM108           -------------HD-----------.---------K--T----------------AL--------.V---R-------------D---------------------------.---S----------------I-------D-------SI-E-----------Q----  633

D.CD.83.ELI               -------------H------------.----------------------A----R-S--------R.----R------------A----------------------------.---I------------------------D-------P--------------N----  641
D.CM.01.01CM_0009BBY      --------V----ES--SW-------.---------K-KS----------------A----I----.I---R---------A--M----------------------------.---I------------------------DK------PF-------------N----  643
D.KE.01.NKU3006           --------------D---------QY.---------KL------------------AQ-C------.----R-----------------------------------------.---I----------S-------------D--------I-------------N----  641
D.KR.04.04KBH8            --------L-----------------.---------K---T---------------AQ-------R.----R---------A------------------------------R.---I---------------R--R-----D------IP--------------N----  641
D.TD.99.MN011             --------V---------W-------.----------IK-----------------SL--------.----R---------A--M----------------------------.---M------------------------D-------PF-------------N----  643
D.TZ.01.A280              ------------E-------------.---------K---T---------------AL-G------.I---R-----------------------------------------.---I---------------R--------DK------S--------------N----  643
D.UG.99.99UGK09259        --------------D---------QY.---------KR------------------AQ-C------.----R------------A----------------------------.----------------------------D-------S--------------N----  641
D.YE.01.01YE386           --------------D---------QY.---------K-------------------AQ-C------.----R-----------------------------------------.---A-----------S----I-------D-------S--------------N----  641
D.YE.02.02YE516           --------V---------W-------.------------------------V----AQ--------.----R---------A--M----------------------------.---I------------------------D-------PF-------------N----  643
D.ZA.90.R1                -------------D------------.----------T-----------A----NLS---------.----R-----------------------------------------.--------------------V-------D-------PF-------------N----  641

F1.AR.02.ARE933           -------------------------L.------------S----------------AQ--------.--------L----D---ISI-------------------------T.---A----------S-----K-------D-----A-P--E-----A--K-VQ----  642
F1.BE.93.VI850            -------------D---------N--.---------KV-S----------------AL--------RS-------L----D----D--------------------------T.---A--D-------S-----K-------DK-K----S--E-----A----------  643
F1.BR.01.01BR125          --------------------------.---------K--S----------------AQ--------.----R--------D-------------------------------T.---A----------S---S-K-------DK----A-S--E-----A-----Q----  642
F1.ES.x.P1146             ------V-----------------T-.---------K--ST---------------AQ-C------.--------------A---D--------------------------T.---M----------S-----Q-------D-------S--R-----A-----H----  640
F1.FI.93.FIN9363          ------PKL-------------R---.---------K--S----------------AL--------.--------L----D-------------------------------T.---A----------S-----K-------D-------S--E-----A-----H----  641
F2.CM.02.02CM_0016BBY     -------------HD----------H.R--------KR-S-----------V----A--G------.V---R----------------------------------------T.---I-----------S----I-------D-------S--E-----------H----  642
F2.CM.95.MP257            -------------HD----------H.----------RKS-----------V---VA--G------.V---R---------I------------------------------T.---I----------------------I-D-------S--E-----------Q----  642
F2.CM.97.CM53657          --------V----HD----------H.----------R-S-----------V----A--G------.V---R----------------------------------------T.---L----------------I-------D-------P--E-----------H----  642

G.BE.96.DRCBL             --------V----L------V----Y.---------KGGS-----------V----A--G-I----.I-------K-----V------------------------------T.---P-V--Y-------------------DK---III---E-----A--H--Q----  638
G.CM.01.01CM_4049HAN      ---E----V----LD----------Y.---------KKGS-----------V----A---------.I-------R-----V------------------------------T.---P----Y-------------------DK-K--IIA--E-----A--H--H----  641
G.CU.99.Cu74              ---EI---V----ND----------Y.---------KRGS---------V-V----S---------.--------R-----V---D--------D--------------R--T.---L----Y-------------------DK-K---I---------------Q----  638
G.ES.99.X138              ---E----V----LD----------Y.---------KRGS-----------V----A----I----.--------R-----V------------D--------------R--T.---A----Y-------------------DK-K--II---E-----A-----QI---  640
G.GH.03.03GH175G          ---E----V----LD----------Y.---------QRGST----------V----A--G------.V-------R-----V------------------------------T.---P----Y-------------------DK-K--II---E-----A-----Q----  642
G.KE.93.HH8793_12_1       ---E----V---E------------Y.---------KRGT-----------V----A---------.--------R-----V------------D--------------R--T.---P----Y-------------------DK-K--II---E-----------Q----  639
G.NG.01.01NGPL0669        --------V----PD----------Y.---------KRGS------R---------A---------.V-------R-----V---D--------D--------------R--T.---P------------------------DK-K---I-IQE--------H--Q----  641
G.PT.x.PT2695             ---E----V----P-----------Y.---------KRGS-----------V----A---------.--------R-----V------------D--------------R--T.---I----Y-------------------DK-K---I---E-----A-----Q----  641
G.SE.93.SE6165            ---E----V----LD----------Y.---------KRGS-----------V----A---------.--------R-----I---D-----------------------R--T.---P----Y-------------------DK-K--II---E-----A-----Q----  639

H.BE.93.VI991             ---E---------PD-----------.---------K--S-----------V----A---------.I---R--------------H---------------H---------T.---E----Y-----------M-------D--K--I-S--E----------------  640
H.BE.93.VI997             -------------PD-----------.---------K--N---S------------A---------.I------------------H--------*------H---------A.---A----Y-----------I-------DK-K----A--E----------------  639
H.CF.90.056               ---------R----------------.---------K--T-----I----------S---------.I---R------------------------------H---------T.---A----Y-I-----------------D--K----S--E----------------  641

J.CD.97.J_97DC_KTB147     SA-E----V----LD-----------.---------KR-T------R----V----SC--------.I---R-------------D------------II-------------.---M----------------I-------DK--------N-----------------  639
J.SE.93.SE7887            SA-E----V----LD-----------.---------KR-S---------A-V----AL-A------.----R----R--------D---------------------------.---M----------S-----T-------DK------------------H------R  640

K.CD.97.EQTB11C           -------------H-----------Y.----------I-S-----------V---VAM--------.----R--------G-------------------------------T.--------------------Q-------DK-----ISI-E-----------H----  642
K.CM.96.MP535             -------------ND----------H.------------S----------------A--G------.----R----------------------------------------T.---------------H----K-R-----D-------SI-E-----A-----C----  642
```

```
                                                                                                                     p51 RT end p15 RNase H start
B.FR.83.HXB2               PSKDLIAEIQKQGQGQWTYQIYQEPF.KNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGK.TPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEK.EPIVGAETFYVDGAANRETKLGKAGYVTNRGRQKVVTLTDTTNQKTELQAIYLALQ  642
01_AE.CF.90.90CF11697      ---E-V--------D-----------.----------K-S------R--A-V---VA---------.----R----R-------A--------------------------.D--A----L------S--N---------D-------S--E--------H--H----  640
01_AE.CN.05.FJ051          -----V--V-----D-----------.-------D--K-S------R--A-V---A---------.---------------M--------------------------.D--L----------S-----------D-------S--E--------H--H----  642
01_AE.CN.06.FJ054          --------V-----D-----------.----------K-S------R---V---A---------.----R---R--------------------------------.D-------------S-----------------------S--E--------H-------  644
01_AE.HK.x.HK001           -----V--V-----D-----------.----------K-S------R--A-V---V---------.A-------R------M--------------------------.D-------------S-----------D-------S--E--------H-------  639
01_AE.TH.01.01TH_R2184     -----V--V-----D-----------.----------K-S------R---V---A---------.I---R---R------M-------------------------#D-------------------------K------S--E--------H--H----  642
01_AE.TH.02.OUR769I        -----V--V-----D-----------.---R------K-S------R---V---A-----L---.----R-----------M--------------------------.---------------S-----------DK------S--E--------H--H----  642
01_AE.TH.90.CM240          -----V--V-----D-----------.----------RGS------R---V---VA---------.----R---R------M--------------------------.D-------------S-----------D-------S--E--------H--H----  647
01_AE.US.00.00US_MSC1164   --------V-----D-----------.----------KGS------R---V---A--G-------.----R-----------M--------------------------.D-----------------------------------S--E--------H--H----  642
02_AG.CM.02.02CM_4082STN   -T------------D-----------.---------KK-S------R--A-----VA---------.----R---R---A-------------------------------.D--A--------------------I-DK------S--E--------H-------  645
02_AG.EC.x.ECU41           -T---V--------D-----------.---------KK-S---------A-V---VAI--------.----R---R---A-------------------------------.D----V----------------------D----R--S--E--------H-------  642
02_AG.GH.03.GHNJ196        -A---V--------D-----------.---------KK-S-------------V---VA----I---.----R---------A--MD--------------------------.D-------------------------D-------S--E--------H--H----  643
02_AG.NG.01.PL0710         -T---V--------D-----------.---------KK-S-------------V---V-V-------.----R--------------V--------------------------.D--P-I-----------------------D-----------E--------H--H----  642
02_AG.NG.x.IBNG            -T---V--L-----D-----------.---------KK-S-------------V---VAM--------.----R---R-------M--------------------------.D------------------I-------D-------S--E--------H--H----  642
02_AG.SE.94.SE7812         -T---V--------D-----------.---------KK-S-------------V---VA---------.----R---R---A--M--------------------------.D------------------I-------D-------S-AE-------H--H----  642
02_AG.SN.98.MP1211         -T------------D-----------.---------KK-S-------------V---VA----V----.----R---R---A--M--------------------------.D--I------------------R-----D----N--S--E--------H--L----  642
03_AB.RU.97.KAL153_2       -----V--------------------.----------L-----------------A---------.--------------------------------------------.-------------------S-------D-------S---------------H----  642
04_cpx.CY.94.CY032         ------------------------H.----------T-S------R---------AM-C------.----R--------D--------------------------T.D--A-----------------Q-------D-------S-SE---------------  641
05_DF.BE.x.VI1310          --------V----L------------.----------A--V---------A------A---------.----R--------I------------------------------.---L---------------------I-DK----A-S--E-----A-----H----  642
06_cpx.AU.96.BFP90         --------V----Y-----------H.----------IKS----------------AL---------.----R-----------------------------------T.-------------------Q-------D-----------E------------N----  644
08_BC.CN.97.97CNGX_6F      ---E----------D-----------.---------K--T----------------AM--------.I---R------------D--------------------------.D--A-V--------------I-------D---K-I-S--------------I---  638
09_cpx.GH.96.96GH2911      -----V--------D-----------.----------K-S------R---V---V-----------.-----------------A--------------------------.D--I----------------------------I-S--E--------H--H----  641
10_CD.TZ.96.96TZ_BF061     --------------D----------H.---------KR-T----------------AQ-C------.--------------I---D-------D----------------.-------------------------D----RAISI-------------N----  641
11_cpx.GR.x.GR17           --------V----LD-----------.---------KR-T------R--A-V----AL-G------.I---R------------D--------------------------.---I-------------------------K-----I-PIE--------E--H---K  641
12_BF.AR.99.ARMA159        --------V-----------F-----.-------------------------------A---------.---------L----D--------------------------T.---AE---------S-K---K-------DK------S--E-----A--H-------  642
13_cpx.CM.96.1849          -E-E--V-------E----------Y.---------KRGS-----------V---VA---------.L-------R-----V---D-------D--------------R--T.---L----Y--------------------D--K--II---E-----------Q----  649
14_BG.DE.01.9196_01        ---E----V----L-----------Y.---------KRGS-----------V----A---------.---------R-----V-----------D-----------X--R--T.---A----Y--------------------DK----II---E-----A-----Q----  641
15_01B.TH.99.99TH_MU2079   -----V--V-----D-----------.----------K-S------R---V----A---------.----R-------------M--------------------------.D-------------S-----------D-------S--E-----------Q---R  641
16_A2D.KR.97.97KR004       --------V----PD-----------.---------KR-ST----------------VM--------.I---R-----------------------------------T.---T-----------------------DK----I-S------------H-------  642
17_BF.PE.02.PE02_PCR0155   -------------------V------.--------------I---------------------------.-------------D--------------------------T.---AE---------------K----X--D-----A-S--E-----A--H--Q----  642
18_cpx.CU.99.CU76          -----V-------YD-----------.---------KR-S----------V-V---VA---------.-------------------------------S--F----------T.---A-----------------I-------DK-----ISI-E--------E--H----  640
19_cpx.CU.99.CU7           --------V-----D-----------.---------K--------------------SL--------.I---R----------A--R--------------------------.---M-------------------------D-------S-------------N----  641
20_BG.CU.99.Cu103          -------------------------.----------KS----------A------AQ---------.-----------------M--------------------------.---------------------------D--K----S--------------H----  638
21_A2D.KE.91.KNH1254       ---E----------D-----------.---------KN-S----------------AM-A------.I----------K---M---------D------------------.---A-----------------------D------I-S--E-----------------  640
22_01A1.CM.01.01CM_0001BBY -------------D-----------.----------K-S-----------V---VA--G-------.----R----------A--M--------------------------.D--I----------S-----------D-------S--E-I------H--H----  642
23_BG.CU.03.CB118          -------------L-----------.----------KST----------V----AQ---------.----R----------A--M--------------------------.---------------------------D--K----S--------------H----  638
24_BG.CU.03.CB378          -------------------------.----------KS------------------AQ---------.---------------A--M--------------------------.---------------------------D--K----P--------------H----  633
24_BG.CU.03.CB471          ---E----------------------.----------KN------------------AQ---------.---------------A--M--------------------------.---E-----------------------D--K--I-S--------------H----  638
25_cpx.CM.01.101BA         ------------------------Y.---------KRGS----------------C---------.----R------------D-------D------------------.D--I----------------------D-------S--E-------H-------  641
27_cpx.CD.97.97CDKTB49     ---E----V----LD-----------.---------KR-A------R---V----AS-C------.----R-----------------------------------Q.-------------------K-------DK-----I---E-----------Q----  640
28_BF.BR.99.BREPM12609     ------V-------D-----------.----------T----------------VA---------.I---T----------I------------------------------.-------------------Q-------D-------P--------------H----  642
29_BF.BR.02.BREPM119       -------------------------.----------T-----------------A--G-------.I---R--------D--------------------------T.---A---------S-----K-------D-----A-P--E-----A--H--Q----  642
31_BC.BR.02.110PA          -------------D-----------.---------K--T------R--------AL---I----.----R---R---A---D--------------------------.---I-----------------------D-----S-S--E-----------Q----  638
32_06A1.EE.01.EE0369       --------V----ND----------H.---------KIKS----------------AM---V----.----R---------------------D-----------------T.-----V-------------K-------D---------SE------------N----  642
33_01B.MY.05.05MYKL007_1   -----V--V-----D-----------.----------R-S------R---V----A---------.----R---R--R---M--------------------------.D--I----------S-----------DK------S--E--------H--H----  642
34_01B.TH.99.OUR2478P      -----V--V-----D-----------.----------K-S------R---V----A---------.----R-------------M--------------------------.D-------------S----R-----I---------S--E--------H--H----  642
35_AD.AF.05.05AF095        --------------D-----------.---------KK-S-----------V---V-L--------.I---R---------A--MD--------------------------.D-------------------------D-------S--E--------H--H----  641
36_cpx.CM.00.00CMNYU830    -------------ED-----------.----------K-S-----I----------A---------.I-------------I--M--------------------------.D-------------------------D-------S--E--------H--H----  642
37_cpx.CM.00.00CMNYU926    -E-E---#L-----D------F--S-f---------KK-S------R--A-----VA----I----.----R---R---A--RD--------------------------.-------------G-----Q-------DK-----I-SF-E--------H--H----  639
38_BF1.UY.05.UY05_4752     -----V-------ND-----------.---------KV---------------------------.I------L-----------------------------------T.---AE---------S-----K-------D-----A-S--E-----A-----LI---  642
39_BF.BR.03.03BRRJ103      ---E----------S-----------.-I-------K-K------R---------AL---------.----R-------D--------------------------T.---I-----------H----A-------D--------SFM----------------  647
40_BF.BR.04.04BRRJ115      -------------------------.---------K-------------------A---------.I---R--------D---S--------------------------T.---A-V----------------------D-----A-S------------------  642
42_BF.LU.03.luBF_05_03     -------------------------.--------G-------------------A--G-----R.I-X-------------I------------------------------.D--------------------X------D-----A-S--E-----A-----X----  646
43_02G.SA.03.J11223        -T-E----------D-----------.---------KRGS-------E--QV---VA---------.----R------------------------------R--------.D----V-----------S-Q-------D-------S--E--------H-------  641
U.CD.83.83CD003            -----V--------D--S--------.---------KR-S----------------SM---------.I-------------S---D--------------------------.D--M-----------------K-------D------IS--E-----A----L-----  643
U.CD.90.90CD121E12         --------------D--S--------.---------KR-S----------------SM---------.------------------AD--------------------------.---A----------S----K-------DQ----TIS--E-----A----L-----  642
U.GR.99.GR303              -------------H-----------Y.----------V-S-----I-----V----------------.----R-----G-------D----P-------.------------D-.G------------------R-------D----X---A--E-----------H----  640
U.NL.95.U_NL_95_H10986_D1  --------------D----------Y.----------RKS----------------AM---------.----R-----------------------------------.D--E------------------M-------D-------S--E-----------Q----  642
CPZ.CD.90.ANT              -GLP-K-T-----S---------NEG.-L--A-----PT-T---E-R--AGV----GL-------E.V---Q---TR---DA--SD----------------IR---N-LA.D--PE-----------NSQ---------D---SR-KH-QK----QA-----LM--E  642
CPZ.CM.01.SIVcpzCAM13      -D-E----V----E-------F--QH.R---------Q-ST----IR--A-VI---AV---------.--R-R--V---S-----A-----------I---------S--T.D--PT-D-Y-----------K-----I-DK-K--IIS-EN----QA--G-L-----  641
CPZ.CM.05.SIVcpzEK505      -D-E-------------------DLH.---------K--ST----IR----V---VAL---------.----R--V---V---------------D----------------T.---S----Y---------------F--D------TSISE----QA----VLM---  648
CPZ.CM.05.SIVcpzLB7        -A------V-----S---------Q-.----------QGS---------A-VM---AI---------.I---R------A--A---D----------Y----SH--F-----Q.---P-----------------------D--K--IIP--E-----A-----Q----  645
CPZ.CM.05.SIVcpzMB66       ---E-----------------F--QY.-------------S----------V---VAL---------.V-R-R---------A---D----------Y---------------Q.D--P-----------------------DK----IIS--E-----A-----Q----  645
CPZ.CM.05.SIVcpzMT145      -E-E-----------------F---H.---------GKQ-A-----IR--A-T----A---------.----R--V---M--A--AD----------I-----------T.---E----Y-----------R------------IHS-E------A-----VLM---  641
CPZ.GA.88.GAB1             -D-E---------NC------F---H.----------Q-S-----IR--A-------A---------.----R--V---S-A--A-----------I----------S--T.---PTTD-Y-----------T--------DK-K--IIS-EN----QA--K-LL----  640
CPZ.TZ.00.TAN1             -KLP-Q-A-----------------EG.---------KSP-T---EIR--AGLI---GN---I---I.V---L--VS----SQ---D---V--V------I-----IR---N-LS.D--PE------------DS-K-R----------YRSKD-EN----QA--W-VD---K  638
CPZ.US.85.CPZUS            -E-E-V--V----RS------F--RH.----------Q-S-----IR--V-V----A---------.V-----LV---V-----S---------D-----------------K--T.-A-E------------Q-----K-------D------II--EN-----A--T-V----K  644
CPZ.CM.x.SIVcpzDP943       -E-E----V----NS------F--QH.----------Q-S-----IR--A-V---VA---------.----R--V---V-----S---------D-----------------X--T.---P-------------D--K-------D------I-N-E------A--T-V-----  649
CPZ.CM.x.SIVcpzMB897       ---E--------------------QH.---------K--A--S---------V---VAL---------.----R----------A---D----------Y---------------T.---------------------------D------IIS--E----RA----------  650
N.CM.02.DJO0131            -G-E----------D----------Y.---------KR-S-----I---V-V---VA--G------.I-------V---V--A---DN--------------------------T.---T----Y-------K---------F--D--K----SIE------------L----  647
N.CM.04.04CM_1015_04       -E-E-----------S---------SX.---------K-KS-----I-----V---VA---------.--------XV---V--A----N--------------------------T.---S----Y-------K---------F--D-------S-EN------------LM---  649
N.CM.04.04CM_1131_03       -E-E-----------S---------SH.---------K--S-----I-----V----A---------.---------V---V--A----N----------------P---------T.---S----Y-----------------F--D-------SIK---------H--LM---  649
N.CM.95.YBF30              -G-E-V-------------------LH.---------K--S-----I---V-V-R-VA---------.----R--V---V--A---DH--------------------------T.---S-----------------------F--D-------SIA------A-----LM---  649
N.CM.97.YBF106             -G-E----------------------Y.---------KX-S-----I-E-AAV---VA---------.---------V---V-------H--------------------------T.---S----Y-------K---------F--D-------SIEN-----A-----L----  649
O.BE.87.ANT70              -D---WVN-----GE---------EH.--------T-QKAS----IR--A-VI--VSQ---I-----.L------VTR-------AD-------------S----I----R--S.---M----Y-------------------EQ-K--IIK-DE-----A--M--L----  638
O.CM.91.MVP5180            -D---WVS---H-E------V--DEH.----------QKAS----IR--A-V---VSQ-A-------.L---R--VTR-------A--------------S----I-------T.-----------------N------------EQ-K-NIIK-EE-----A--M-VLI---  638
O.CM.96.96CMABB637         -N---WVN---H-K----------DEH.----------QKAS---EIR--A-V---VAQ--------.L------VTR-------AD-------------S----I-------S.--------Y-----------R--------EQ-K--IIN-EE-----A--M-VLI---  640
O.CM.98.98CMA104           -D---WVSX----E----------DEH.--------T-QKAS----IR--A-VL--VSQ--------.L------VTR-------AD-----------X--S----I----R--S.---M--D-----------D----------EQ-K--IIK-DE-----A--M-VL----  638
O.CM.99.99CMU4122          -D---WVN-----E------V---EH.--------T-Q-AS----IRI-A-VI--VSQ-A-------.L--------TR-------AD-------------S----I----R--S.---M----Y-------N-----------EQ-K---IK-EE-----A--M--X---K  638
O.FR.92.VAU                -D---WVN-----E----------DEH.----------QKAS----IR--ANVI--VSQ-A-I----.L------VTR-V-----AD-------------S----I----K--N.---I-------------D----------I-EQ-K--IIK-DE-----A--M-VL----  639
O.SN.99.SEMP1299           -D---WVN------W---------DEH.---------T-QKAS----IR--A-VL-RVSQ-A-I----.L--------TR-------AD-------------S----I-------S.---M----Y--------D-----------EK----IIK-EE----RA--M-VL----  638
O.US.99.99USTWLA           -D---WVNX----E---S--V--DEH.---------T-QKAS----IR--A-V-X-ASQ-A---X--.L------VTR-------AD-------------S----I-------T.---------Y--------D----------EQ-K--IIQ-EE-----A--M-XL----  641
```

p66 RT, p15 Rnase H end p31 Integrase start

```
B.FR.83.HXB2            DSGLEVNIVTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQ.DEHEKYHSNWRAMASDFNLPPVVAKEIVASCDKC.QLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQ 810
A1.GE.99.99GEMZ011      ---S--------------V-----R----I--K---K--E--R---S-------------------N---R----E------.E---------K----------I------------.------------------------------I--------------------- 810
A1.KE.00.KER2008        ---S--------------------R----I------K--G-D----S-------------------S---R-----------.E---R--N----------I--II-----------.----------------M----------------------------------- 810
A1.KE.00.KNH1144        ---S--------------------K-----------K--G-D----S------------K------N---------------.E---R----------------I-------C----.------I---------------------I----------------------- 810
A1.KE.00.KSM4024        ---S--------------------R-----------K--G-D----S-------------------S---------------.E---R------T---------I---------N--.----------------------------I-I--------------------- 810
A1.KE.00.MSA4069        ---S--------------------R-----------E----D----S-------------------N---------------.E---R------T----------I------C----.------I--------------------------------------------- 810
A1.KE.00.NKU3005        ---S--------------------R----I------K--G-D----S-------------------S---------------.E---R------T---------I------------.---------------------------------------------------- 810
A1.RU.00.RU00051        ---SK-------------------R--A-I--K---KX-E------S-------------------X---------------.ED--R-----K---N------I------------.------I--------------------------------------------- 810
A1.RW.93.93RW_024       ---S--------------------R---G-------K--E-D----S-------------------S---------------.E---R------T----------------------.---------------------------------------------------- 810
A1.SE.95.SE8891         ---S--------------------R-----------K--G-D----S---------------------V-RI----------.E---R------T---------I------------.----------------M----------------------------------- 805
A1.SE.95.UGSE8131       ---S--------------------K-----------K--E-D----S-------------------S---------------.E---R------T---------I------------.---------------------------------------------------- 804
A1.TZ.01.A173           ---P--------------------R-D--I------K--E-D----S-------------------S---------------.ED-------------------I------------.------V---------M----------------------------------P 810
A1.UA.01.01UADN139      ---S-------------------------I--K---K--E--R---S-------------------D---R----E------.E---------K----------I------------.------------------------------I--------------------- 810
A1.UG.92.92UG037        ---S--------------------R-----------K--E-D----S-------------------S---------------.E---R----------------I------------.---------------------------------------------------- 811
A1.UG.99.99UGA07072     ---S--------------------R-----------K--G-D----S-------------------S----I----------.E---R------T---------I------------.----------------M----------------------------------- 810
A2.CD.97.97CDKTB48      ---S--------------------S-----------K--E--R---S-------------------C---------------.E-------------H------I-----------W#R---------------------------I-V--------------------- 803
A2.CY.94.94CY017_41     -----------------------ER----I------K--E--R---S-------------------S---------------.E---R---------H-------------------.-----------------------------------------------T---- 810
A.SN.01.DDI579          ---S----------------------D--I------K--E------S-------------------S---------------.E---R------T---------I------------.------------------------------I--------------------- 809
A.SN.01.DDJ369          ---S--------------------R----I------K--E------S-------------------N---------------.E---R------T---------I------------.------K-----------------------I--------------------- 809
A.SN.96.DDJ360          ---SK-------------------K----I------K--E------S-------------------S---------------.E---R------T---------I------------.------------------------------I--------------------- 809
A.CD.02.02CD_KTB035     ------------------------R----I------K--E---I--S------------K------S------------E--.A------N-------------I------------.----------------------------------------M----------- 810
A.CD.97.97CD_KCC2       ---S--------------------I----I------K--E------S-------------------S---------------.E---R------T----------------------.---------------------------------------------------- 810
A.CD.97.97CD_KTB13      ---S----------V---------R---G-------K--E------S-------------------S---------------.E---R------T----------------------.---------------------------------------------------- 810
B.AR.04.04AR151516      ------------------------K----I-S----------R-----------------------S---------------.ED------------N-------------------.------I--------------------------------------------- 815
B.AU.87.MBC925          -------------------------------S--------------------------------------------------.E------N----------I----------C----.---------------------------------------------------- 812
B.BO.99.BOL0122         ---P--------------------R----I-S------------------------------------V-------------.------------I------------Q---N----.-V--------------------------I----------------------- 811
B.BR.03.BREPM2012       ------------------------K------S----E-----------------------------S---------------.E------N-------------I------------.---------------------------------------------------- 814
B.CA.97.CANB3FULL       ---S--------------------K------S-------Q--R-----------------------S---------------.E---------------------------------.----------------------------I----------------------- 813
B.CN.05.05CNHB_hp3      ---A--------------------K------S-----------I----------------I---------------------.E------N----------I----------C----.----------------------------I---------------I--T---- 810
B.CO.01.PCM001          ------------------------K------S--------------------------------------------------.ED-------------------I------------.----------------V-----------I----------------------P 810
B.GB.83.CAM1            ------------------------K------S----E---------------------------------------------.E---------------------------------.---------------------------------------------------- 810
B.GB.86.GB8             ------------------------R----V-S----------------------------------S----I----------.ED------------N-------------------.----------------------------I----------------------- 816
B.GE.03.03GEMZ010       --------------------G---K------S-------N-------------------------------I----------.E-----------L--------------I-C----.-I--------------------------I----------------------- 810
B.IT.05.SG1             -XV---------------L-----KRX----S----------------------------------V-------*-----X-.X-P-N----*-PX---LXR--I-P-KL-P-G-N-.------LL----X---L---NG--*K----X-P-----XF------------ 803
B.JP.05.DR6538          ------------------------K------S--------------------------------------------------.E---------------------------------.-I--------------------------I----------------------- 810
B.KR.05.05CSR3          ------------------------K-----------------R---T-------------------N---------------.ED-----T-----V--------------------.----------------------------I-----------V----------- 810
B.NL.00.671_00T36       ------------------------K------S----------------------------------T---------------.E------------------------------N--.----------------------------I----------------------- 820
B.RU.04.04RU128005      ---------------------K--K------S----------------------------------S------X--------.X---------------------------------.----------------------------I----------------------- 810
B.TH.00.00TH_C3198      ------------------------K------S--------------------------D-I---------------------.E---------------------------------.----------------------------I------------------T---- 810
B.UA.01.01UAKV167       ------------------------K------S--------------------------------------------------.----------------------------------.----------------------------I----------------------- 817
B.US.04.ES10_53         ------------------------K--------------N------------------------------------------.E----------------G------R---------.----------------------------I----------------------- 810
B.US.99.PRB959_03       --E---------------------K-----------------------------------------S---------------.E------N--------------------------.------------------------------------S--------------- 810
C.AR.01.ARG4006         ------------------------K--------------Q------S-------------------S---R-----------.ED------------NE-----I------------.------T---------------------I----------------------- 806
C.BR.04.04BR013         ---S--------------------K-----------------R---S------------K------S---------------.E--------------E-----I------------.----------------------------I----------------------- 814
C.BW.00.00BW07621       ---P--------------------K---------------------S-------------I-----S---------------.--------------NE-----I------------.------I---------------------T-----------M----------- 806
C.CN.98.YNRL9840        ---S--------------------K-----------------R---S-------------------S---------------.E--------------------I--------G-Q-.----------------------------I----------------------- 806
C.ET.02.02ET_288        ---S--------------------K----V------E-----R---S-------------------S---------------.E-------------NE-----I------------.----------------------------I----------------------- 812
C.GE.03.03GEMZ033       ---S--------------------K--------------N--RI--S-------------------S---------------.E------N-------------I------------.----------------------------I--XX------------------- 806
C.IL.99.99ET7           ---S--------------------K-----------------RI--T-------------------N---------------.E-------------NE------------------.------I--------------------------------------------- 806
C.IN.99.01IN565_10      ---S--------------------K-----------------R---S-------------------S---R-----------.ED--R----------------I----------Q-.----------------------------I----------------------- 806
C.KE.00.KER2010         ---S--------------------K--------------R------S-------------------S---------------.E--------------E-----I------------.------T---------------------I----------------------- 806
C.MM.99.mIDU101_3       ---S--------------------K-----------------R---S-------------------S---------------.E--------------------I----------Q-.----------------------------I-----------M----------- 806
C.MW.93.93MW_965        ---S--------------------K----I---------N---I--S-------------------S---------------.E--------------E-----I------------.------I---------------------I----------------------- 807
C.SN.90.90SE_364        ---S--------------------K-----------------R---S-------------------S---------------.E--------------E-----I------------.------I---------------------I----------------------- 802
C.SO.89.89SM_145        ---S--------------------K-----------------R---S-------------------S---------------.E--------------E--------R---------.----------------------------I----------------------- 817
C.TZ.02.CO178           ---S--------------------K-----------------R---S-------------------S---------------.E--------------E-----I------------.-Q----I---------------------I----------------------- 806
C.UY.01.TRA3011         ------------------------K-----------------R---S-------------------N------------R--.ED-------------E------I-----------.------T---------------------I----------------------- 806
C.YE.02.02YE511         ---S--------------------K---------------------S-------------------S---------------.E------N-------------I------------.------I---------------------I----------------------- 804
C.ZA.04.04ZASK164B1     ---S--------------------K--------------N------S-------------------T---------------.E--------------------I------------.------T--------------------------------------------- 806
C.ZA.05.05ZAPSK240B1    ---S--------------------K--------------N--R---S-------------------R---------------.E--------------E-----I------------.----------------------------------------M----------- 806
C.ZM.02.02ZM108         ---S--------------------K----------------DR---S-------------------N---------------.E--------------E--------R---------.----------------------------------------M----------- 801
D.CD.83.ELI             ------------------------K-----------------------------------------Q---------------.E------N--------------------------.---------------------------------------------------- 809
D.CM.01.01CM_0009BBY    ---S--------------------K------S----E-----------------------------S---------------.E---------------------------------.---------------------------------------------------- 811
D.KE.01.NKU3006         ---S--------------------K------S-------T------S-------------------N----I----------.E---R--N--------------------------.------L--------------------------------------------- 809
D.KR.04.04KBH8          ------------------------K------S----------------------------------N------------E--.ED-------------E-------------C----.-Q-------------------------------------------------- 809
D.TD.99.MN011           ---S--------------------K----I-S----------------------------------S-----------SQ--.---D------------------------------.-I-------------------------QI----------------------- 811
D.TZ.01.A280            --------I---------------K------S----------------------------------S---------------.ED--R---------------S--------C----.---------------------------------------------------- 811
D.UG.99.99UGK09259      ------------------------R--------------Q------S-------------------N----I----------.E------N----------I---------------.------L---------------------I----------------------- 809
D.YE.01.01YE386         -------------------H----K---------------------S-------------------N-V--I----------.E------N---S-VGE--I---------------.------L---------------------I----------------------- 809
D.YE.02.02YE516         ---S--------------------K----V-S----------------------------------S-----------NQ--.ED-D--------------I---------------.-Q--------------------------I----------------------- 811
D.ZA.90.R1              ------------------------K------S----------R-----------------------Q---------------.E------N--------------------------.------------------------------I--------------------- 809
F1.AR.02.ARE933         ---A--------------------K----------------------------------------------I----------.ED-----N--------------------------.----------------------------------------L----------- 810
F1.BE.93.VI850          ---S--------------------K----I---------Q--R---S---------------------V--I----------.E------N-------------I------------.--------------------------------------E------------- 811
F1.BR.01.01BR125        ---S--------------------K---------------------S------------------------I----------.E------N--------------------------.----------------------------------------L----------- 810
F1.ES.x.P1146           ---S-----------------------------------Q--RI--S------------------------I----------.ED--------K-----------------------.----------------------------I----------------------- 808
F1.FI.93.FIN9363        ---S--------------------K--------------Q---I--S------------------------I----------.E------N--------------------------.----------------------------I----------------------- 809
F2.CM.02.02CM_0016BBY   ---S-----------------H--K--------------Q------S-------------------T---------------.E------N-------------I------------.------I---------------------I----------------------- 810
F2.CM.95.MP257          ---S-----------------H--K----I---------Q--R---S-------------------T---------------.E---------------------------------.----------------------------I----------------------- 810
F2.CM.97.CM53657        ---S-----------------H--K--------------Q--R---S-----------------------------------.E---------------------------------.----------------------------I----------------------- 810
G.BE.96.DRCBL           --QS--------------------R----I----------------S-------------------S---------------.E---R-----K----------I------------.----------------------------I----------------------- 806
G.CM.01.01CM_4049HAN    --RP--------------------R---------------------S-------------I-----S----I----------.ED--R------T----------------------.----------------------------I----------------------- 809
G.CU.99.Cu74            ---S--------------------R---------------------S-------------------T---------------.E---R----------------II-----------.----------------------------I----------------------- 806
G.ES.99.X138            ---S--------------------R---------------------S-------------------S---------------.ED--R-----------------I-----------.----------------------------I----------------------- 808
G.GH.03.03GH175G        --KS--------------------R---------------------S-------------------N---------------.E---R----------------I------------.------L---------------------I-I--------------------- 810
G.KE.93.HH8793_12_1     ---S--------------------R---------------------S-------------------S---------------.ED------------------HI---------H--.----------------------------I----------------------- 807
G.NG.01.01NGPL0669      ---S--------------------R-----------------R---T-------------------N---------------.E---R---------N-------------------.----------------------------I----------------------- 809
G.PT.x.PT2695           --RP-------------------NK---------------------S-------------------S---------------.ED--------K-----------------------.----------------------------I----------------------- 809
G.SE.93.SE6165          --RS--------------------R--A------------------S-------------------S---------------.E---R--N-------------I------------.----------------------------I-I--------------------- 807
H.BE.93.VI991           E--P--------------------K-----------E------F--S-------------------S---------------.VQ-------------------I------------.----------------------------I----------------------- 808
H.BE.93.VI997           ------------------------K-----------E---------S-------------------S---------------.EA--R--N-------E-----I------------.-----------------------------------------XP--------- 807
H.CF.90.056             ------------------------K-----------E---------S-------------------S-V-------------.E---R--N----V--------I------------.---------------------------Q-----------------------K 809
J.CD.97.J_97DC_KTB147   ---Q--------------------K-----------E-----R---S-------------------S------------R--.E---R----------------I------------.----------------M-----------I------I---------------- 807
J.SE.93.SE7887          ------------------------K-----------E---------S-------------------S---------------.ED--------------------------------.---------------------------------------------------- 808
K.CD.97.EQTB11C         ---S--------------------K----------------DR---S-------------------S---------------#K------N--------------------------.------I---------------------I----------------------- 810
K.CM.96.MP535           ---S--------------------K---D-------------RI--S-----------------------------------.E------N-------------I------------.------I---------------------I----------------------- 810
```

```
                              p66 RT, p15 Rnase H end p31 Integrase start
B.FR.83.HXB2                DSGLEVNIVTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQ.DEHEKYHSNWRAMASDFNLPPVVAKEIVASCDKC.QLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQ  810
01_AE.CF.90.90CF11697       ---SD--------H----------R----I------K--E--R---S-------------------S---------------.E---R------T---------I-------N----.----------------------------------------------------  808
01_AE.CN.05.FJ051           --ES--------------------R----V------E---------S-------------------S---------------.E---R------T---------I-------N----.----------------------------------------------------  810
01_AE.CN.06.FJ054           ---S--------------------RN---I------EI------H-S-------------I-----S---------------.E---R------T-----------------N----.----------------------------I-----------------------  812
01_AE.HK.x.HK001            ---S--------------------R----I-S----E---------S-----------NK------S--------N------.E---R------T---------I-------N----.----------------------------------------------------  807
01_AE.TH.01.01TH_R2184      ---S--------------------R----V------E---------S-------------------S---------------.E---R------T----------I-----------.------I---------------------------------------------  810
01_AE.TH.02.OUR769I         ---S--------------------R----V------E---------S-------------------S---------------.E---R------T---------I-------N----.----------------------------------------------------  810
01_AE.TH.90.CM240           ---S--------------------R----V------E---------S-------------------S---------------.E---R------T---------I------TN----.----------------------------------------------------  815
01_AE.US.00.00US_MSC1164    ---S--------------------R----V------E---------S-------------------S---------------.E---R------T---------I-------N----.----------------------------I-----------------------  810
02_AG.CM.02.02CM_4082STN    ---S--------------------R-D---------K--E------S-------------------S---------------.E---R----------------I------------.---------------------------R------------------------  813
02_AG.EC.x.ECU41            ---S--------------------R--------------G-D----S-------------------N---R-----------.ED--R------T----------------------.------I------G--------------I-----------------------  810
02_AG.GH.03.GHNJ196         ---S--------------------K-D--I--L---K--E-D----S-------------------N---R-----------.E---R---------N------I------------.----------------------------I-----------------------  811
02_AG.NG.01.PL0710          ---SQ-------------------R----I------K--E-D----S-------------------N-V-------------.ED--R----------------I--R---------.------I------G--------------I------------------S----  810
02_AG.NG.x.IBNG             ---S--------------------R-----------K----D----S-------------------N---------------.E---R-----K----------I------------.-M-----------G--------------I-----------------------  810
02_AG.SE.94.SE7812          ---S--------------------R------T----E----DR---S-------------------S---------------.E---R-----------------------------.----------------------------------------------------  810
02_AG.SN.98.MP1211          E--S--------------------R-----------K--E-D----S-------------------K---------------.E---R----------E-----I------------.-------------------I--------I-I---------------------  810
03_AB.RU.97.KAL153_2        ------------------------K------S--------------------------------------E-----------.E------G--------------------------.----------------------------I-----------------------  810
04_cpx.CY.94.CY032          ---S-----------I--------R---D----------R-D----S-------------------N---------------.E------N------------S----------N--.------------------------------M---------------------  809
05_DF.BE.x.VI1310           ------------------------K--------------E-K----S-----------------------------------.----------------------------------.----------------------------------------------------  810
06_cpx.AU.96.BFP90          ---P--------------------K------------------I--S-------------------N---------------.ED--R---------N------I------------.------I---------------------I-----------------------  812
08_BC.CN.97.97CNGX_6F       ---S--------------------K-----------------R---S-------------------N---------------.E--------------------I----------Q-.----------------------------I-----------------------  806
09_cpx.GH.96.96GH2911       ---S--------------------R-----------K--E------S-------------------S---------------.E---R-----------------------------.----------------------------I-----------------------  809
10_CD.TZ.96.96TZ_BF061      ---S--------------------K---------------------P-------------------R---------------.E------N--------------------------.------L---------------------------------------------  809
11_cpx.GR.x.GR17            ------------------------K------S--------------S-------------------S--------------P.EG--R-----------------I------N----.---------------R------------I-----------------------  809
12_BF.AR.99.ARMA159         ---S--------------------K------S--------------S---------------------V--I----------.E------N--------------------------.----------------------------T-----------L-----------  810
13_cpx.CM.96.1849           ---S--------------------R---------------------S-------------------N---------------.ED--R------T----------------------.------T---------------------I-----------------------  817
14_BG.DE.01.9196_01         ---S-I------------------R---------------------S-------------------S---------------.ED--------------------------------.----------------------------I-X---------------------  809
15_01B.TH.99.99TH_MU2079    ---S--------------------R----V------E---------S-------------------S---------------.E---R------T---------I------------.----------------------------------------V-----------  809
16_A2D.KR.97.97KR004        ------------------------R----I--L---K--E--R---S---------E---------S---------------.------R-------H--------------C----.-V--------------------------I-----------------------  810
17_BF.PE.02.PE02_PCR0155    ---S--------------------K--------------E------S---------------R--------I----------.E---R--N--------------------------.----------I-----------------------------L-----------  810
18_cpx.CU.99.CU76           ---Q--------------------K----I------E-----------------------------S---R-------E---.E--D---N--------------------------.---------------------R------I-I---------------------  808
19_cpx.CU.99.CU7            ------------------------K------S----------R-----------------------S---------------.E--------------------I------------.----------------------------------------------------  809
20_BG.CU.99.Cu103           ------------------------K------S----------------------------------S---------------.E---R---------N------I------------.----------------------------I------------------S----  806
21_A2D.KE.91.KNH1254        --E---------------------R--------------G--R---S-------------------S---------------.E-------------H------II------C----.----------------------------------------------------  808
22_01A1.CM.01.01CM_0001BBY  ---S--------------------R-------L---K--E------S-------------------S----I----------.ED--R------T-V-------I---------N--.------I---------------------------------------------  810
23_BG.CU.03.CB118           ---S--------------------K------S----------------------------------S---------------.E---------K-----------------------.----------------------------I-----------------------  806
24_BG.CU.03.CB378           ------------------------K------S----------------------------------S---------------.E---R----------------I------------.----------------------------I-----------------------  801
24_BG.CU.03.CB471           ------------------------K------S----------------------------------N---------------.E---R----------------I------------.-------------N--------------I-----------------------  806
25_cpx.CM.01.101BA          ---S--------------------NX---I----------------S------X------------S---------------.E---R----------------I------------.------L---------------------I-----------------------  809
27_cpx.CD.97.97CDKTB49      ---S------------------S-K---------------------S-------------------S----I----------.ED--R--T-------------I------------.----------------------------I------------------T----  808
28_BF.BR.99.BREPM12609      ------------------------K------S---------------------------------------I----------.-----------S-D-------I------------.----------------------------------------------------  810
29_BF.BR.02.BREPM119        ---S--------------------K-----------H----T----S---------------------V--I----------.ED-----N---S---------I------------.----K-------------------V-----I---------L-----------  810
31_BC.BR.02.110PA           ---S--------------------K-----------------R---S-------------------K---------------.E---R----------------I------------.------T---------------------I-----------------------  806
32_06A1.EE.01.EE0369        ---S--------------------K---------------------S-------------------S---------------.E---R---------N------I------------.-------------------I--------I-I---------------------  810
33_01B.MY.05.05MYKL007_1    ---S--------------------R----V------E---------S------------K------S---------------.E---R------T---------I-----A-N----.------------------RV--------------------------------  810
34_01B.TH.99.OUR2478P       ---S--------------------K----V------E---------S-------------------S---------------.E---R------T---------I-------N----.------------------------------------------T---------  810
35_AD.AF.05.05AF095         ---S--------------------K-----------K-VE-D----S-------------------T-----------E---.E---R----------------I---Q---N----.----------------M-------------I---------------------  809
36_cpx.CM.00.00CMNYU830     ---SK-------------------R-----------K--E-D----S-------------------K----I----------.ED--R----------------I------------.----------------------------I-I---------------S-----  810
37_cpx.CM.00.00CMNYU926     --ES--------------------R-D--I------KI-E--R---S-------------I-----S----GV---------.ED-KR---------N------I------------.----------------------------------------------------  807
38_BF1.UY.05.UY05_4752      ---P--------------------K--------------N------S------------------------I----------.----------------------------------.-------------------I--------I-I---------L-----------  810
39_BF.BR.03.03BRRJ103       ---S--------------------K------SK---E-----------------------------S---------------.E------N--------------------------.----------------------------------------------------  815
40_BF.BR.04.04BRRJ115       ---------S-----------L--K--------------N--R-----------------------S---------------.E------N--------------------------.----------------------------------------------------  810
42_BF.LU.03.luBF_05_03      ------------------------K----I---------X------S-------------------T----I----------.ED--------K-------I----X---I-C----.----------------------------X-----------L-----------  814
43_02G.SA.03.J11223         ---S--------------------R-----------K--E-D----S-------------I-----S---------------.E-X-R------------D----------------.-Q----------------------------M---------------------  809
U.CD.83.83CD003             -----A------------------S---------------------S-------------A---------------------.E---R----------------I------------.----------------------------------------L-----------  811
U.CD.90.90CD121E12          -----A------------------S---------------------S-----------------------------------.E--------------------I------------.------I---------------------I-----------------------  810
U.GR.99.GR303               ---S--------------------T-----------K--G--R---S-------------I-----N---------------.ED-XR---------N-------I--------AGI.R-----------------------------P------N--------------  808
U.NL.95.U_NL_95_H10986_D1   ---SA-------------------N---D----------Q-DQ---S-------------------N----I----------.ED-----T--------------------------.----------------M-----------I----A------------------  810
CPZ.CD.90.ANT               --TGP------------VL-GT------P--EE--QK---R-QI--S-------------------Q---Q----E------.ED-D-------SL-DEY----I-----I-Q----.HV----R------------V----------I-----S--F-----MAD---K  810
CPZ.CM.01.SIVcpzCAM13       --EQQ---------V-----S---H-----------E----D----S-------------------S---------------.E---R-----K-L-GE-----I-------N----.-V----L---------T--M--------I-I----I----------------  809
CPZ.CM.05.SIVcpzEK505       -A-Q----------V----HS---K-----------E-----RI--S-------------I-----T---------------.E---R-----K----------I------------.------I---IN----V-----------I-----------L-----------  816
CPZ.CM.05.SIVcpzLB7         ---P--------------L-------D---------N---------T-------------------S------------R--.E------N-------------I------------.----------------------------------------M-----------  813
CPZ.CM.05.SIVcpzMB66        --EV----------------G---T----I------E---------S-------------I-----S---------------.E------N-------------I-------N----.------I---------------------I-----------M-----------  813
CPZ.CM.05.SIVcpzMT145       ---Q---V------V---L-SH--H---DI------E------I--S-----R-------------S-----------E---.E---R-----K---Q------I-------Q----.-------------------I--------I-I---------------------  809
CPZ.GA.88.GAB1              --DQQ---------V-----S---H-----------E------I--S--------------------------------R--.E---R-----K----------I-------H----.-V-----------------V----------I---------------------  808
CPZ.TZ.00.TAN1              ---AQ---------VM-VL-GL----D-PI-E---QK-TQ-TAI---------------E------KN---I------NE--.ED-D------K-L-DEY----------I-Q-P--.HI----I-----Y--E---I----------I--------F-------E---R  806
CPZ.US.85.CPZUS             --ENT--V--G---V----HS---------------E------S-IS-------------I-----S---------------.ED-D-------T----------------------.-P----I---------------------I-I---------L-----------  812
CPZ.CM.x.SIVcpzDP943        -AEHI---------V----HS-----------L---E------S--S-------------------S---R-----------.E------N--------------------------.----------------M-----------------------L-----------  817
CPZ.CM.x.SIVcpzMB897        --EV----------------S---I---DI--------V----I--S-------------I---------------------.E------N--K---DE-----I------------.----------------------------I-----------------------  818
N.CM.02.DJO0131             ---Q---------------HSH--K------S----E-----R---S-------------------S-----------E---.ED-XR---------------S-------------.------I----N----V-----------I-----------L-----------  815
N.CM.04.04CM_1015_04        -X-Q-----------M---HS---K------S----E-----R---S-X-----------------S-----------E---.E---R-----K----------I------------.-----------N----V-----------I-----------L----------X  817
N.CM.04.04CM_1131_03        E--Q-----------M---HS---K------S----E-----R---S-------------------S-----------E---.E---R-------------X--I------------.-----------N----V-----------I-----------L-----------  817
N.CM.95.YBF30               E--RD----------M---HS---K------S----E-----R---S-------------------S----I------E---.ED-DR-----K----------I------------.-----------N----V-----------I-----------L-----------  817
N.CM.97.YBF106              E--Q-A---------M---HS---K---D--G----E-----R---S---------------X---S---X-------E---.E---R-----K----------I------------.----------IN----V-----------I-----------L-----------  817
O.BE.87.ANT70               --KET------------V-SS--T----PI-Q----E-T---Q---T------------KI-----KD--R----E---Q--.ED--------K-L--E-G---------I---P--.HI----I--------EV--I----M---I-I--------F------------  806
O.CM.91.MVP5180             --KEQ---------V----SS--T--D-PI-Q----E-T---R---T------------KI-----KD--R----E---Q--.ED----------L----G---I-----I---P--.HI----T-----Y--E---M--------I-I-------DF------------  806
O.CM.96.96CMABB637          --KET---------V----SS--T--D-PI-Q----E-T---RT-IT---------X--KI-----KD--R----E---X--.ED----------L----G---------I---P--.HI----I-----Y--E---M----------I--------F------------  808
O.CM.98.98CMA104            --KEK---------V----SS--T----PI-Q----E-T---Q---T-------------I-----KD--R----E---Q--.ED------X---L----G---------I-N-P--.HI---XI--------EV--M----T-----I--------F------------  806
O.CM.99.99CMU4122           --KET--------F-----SS--T--D-PI-Q----E-T---Q----------------KI-----KD--R----EX--Q--.ED----------L----G---------I-N-P--.HI----I--------EV--I----T-----I--------F------------  806
O.FR.92.VAU                 --KEK---------V----SS--T--D-PI-Q----E-T-R-Q---T------------KI-----KD--R----E---Q--.ED----------L----G---------I---P--.HIRR-----------E---M--------I-I--------F------------  807
O.SN.99.SEMP1299            --KET----Q----V----SS--T----S--Q----E-T---Q---T------------KI-----KD--R----E---Q--.ED----------L----G---------I-N-PQ-.HI----I--------EV--M--------I-I--------F------------  806
O.US.99.99USTWLA            --KER---------VX---SS--T--D-PI-Q----E-T-X-QX--TX-----------KI----XKD--R----E---Q--.ED-------X--L----G---------I---P--.HIRE-P--X------E---M--------I-I--------F------------  809
```

```
B.FR.83.HXB2                ETAYFLLKLA.GRWPVKTIHTDNGSNFTGATVRAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRNPLWKGPAKLLWKGEGAVVIQDNSDIKVVPRRK 979
A1.GE.99.99GEMZ011          ----------.------VV-----P---SSA-K------D-Q------------------------------E---------------------------------I----------------I-------------D-I--------------------N--------- 979
A1.KE.00.KER2008            ----------.------VV---------S-AFK------NVQ--Y-------------------------------------------------------------I----S-------------------------D-------------------------------- 979
A1.KE.00.KNH1144            -----I----.------V----------S-A-K------D-Q------------------------------E---------------------------------I--------------H---------------D-V------------------------------ 979
A1.KE.00.KSM4024            -A--------.------VV---------S-AFK------N-Q--------------------Q--Q------E---------------------------------I--------------H----------F----D-V------------------------------ 979
A1.KE.00.MSA4069            -----I----.------VV---------S-A-K------N-Q----------------------------------------------------------------I------------------------------D-------------------------------- 979
A1.KE.00.NKU3005            -A--------.------VV---------S-AFK------NV-------------------------------E---------------------------------I---------R----------K---------D-V------------------------------ 979
A1.RU.00.RU00051            ----------.------V----------SSA-K------N-Q------------------------------E---------------------------------I---------R--------------------D-V------------------------------ 979
A1.RW.93.93RW_024           ----------.------VV---------S-A-K------N-Q------------------------------E---------------------------------I------------------------------D-I------------------------------ 979
A1.SE.95.SE8891             ----------.------VV---------S-AFK------NVQ------------------------------E---------------------------------I------------------------------D-----------------L-------------- 974
A1.SE.95.UGSE8131           -A--------.------VV---------S-AFK------NVQ------------------------------E-----R---------------------------I------------------#-----------D-I-------------------Q---------- 972
A1.TZ.01.A173               ----------.------VV---------S-AFK------SVQ--Y---------------------------E---------------------------------I------------------------------D-I------------------------------ 979
A1.UA.01.01UADN139          ----------.------V------P---SSA-K------N-Q------------------------------E---------------------------------I------------------------------D-I--------------------N--------- 979
A1.UG.92.92UG037            ----------.------VV---------S-A-K------NV-------------------------------E---------------------------------I----S-L-------------K---------D-I------------------------------ 980
A1.UG.99.99UGA07072         ----------.------VV-----P---S-AFK------NVQ------------------------------E---------------------------------I---------R----H---M------F----D----------------------G--------- 979
A2.CD.97.97CDKTB48          ----------.------V------P---S-A-K------DV-----------------------------------------------------------------I----------------I-------------D-I--------------------G--------- 972
A2.CY.94.94CY017_41         D----I----.------V------P--IS---K--------Q----------------------------------------------------------------I-------------R----------------D-I------------------------------ 979
A.SN.01.DDI579              -A--------.------V----------S-AMK------N--------------------------------E---------------------------------I------------------------------D-I-------------------------I---- 978
A.SN.01.DDJ369              ----------.------V----------S-A-K------N-Q------------------------------E---------------------------------I---------Q------I-------------D-I------------------------------ 978
A.SN.96.DDJ360              ----------.------V----------S-AMK------D--------------------------------E---------------------------------I------------------------------D-I-------------------------I---- 978
A.CD.02.02CD_KTB035         -A---I----.------V----------STA-KT------VQ----------------------------------------------------------------I---------------VL-------------D-I---------------------E-------- 979
A.CD.97.97CD_KCC2           -----I----.------V----------STA-K------S-Q------------------------------E----------------------------------------L-----------------------D-I------------------------------ 979
A.CD.97.97CD_KTB13          -A---I----.------V----------S-A-K------N-Q-----------------------------------------------------------------------------------------------D-I------------------------------ 979
B.AR.04.04AR151516          -----I----.-----------------ST--K--------N-----------------------------------------------------------------------------------------------D----R--------------------------- 984
B.AU.87.MBC925              ----------.-------------P---SN--K--------------------------I----------------------------------------------I------------------------------D-------------------------------- 981
B.BO.99.BOL0122             -----I----.-----------------SNA-K-------V--------------------N---------------------------------------------------------------------------D-------------------------------- 980
B.BR.03.BREPM2012           ----------.-----------------S-A-K-------V-----------------------------------------------------------------I------------------------------D-------------------------------- 983
B.CA.97.CANB3FULL           ----------.-----------------S---K----------------------------------------------------------------------------L---------------------------D-------------------------------- 982
B.CN.05.05CNHB_hp3          -----I----.-------------R---STS-K----------------------------N------------------------------------------------------R--------------------D-------------------------------- 979
B.CO.01.PCM001              -----I----.-----------------ST--K-------V--------------------N------------------------------------R--------------------------------------D-------------------------------- 979
B.GB.83.CAM1                ----------.-------------G--IST--K----------------------------N--------------------------------------------I------------------------------D-------------------------------- 979
B.GB.86.GB8                 D----I----.-------------R---ST--K--------------------------------------------------------------------------------------------------------D-------------------------------- 985
B.GE.03.03GEMZ010           ----------.------V----------SNALK-----------------------------------A--------------------------------------------------------L------F----D-------------------------------- 979
B.IT.05.SG1                 --------I-.-----------------SS--K-------------------------------------------------------------------------I-------------N--------------N-D-F------------------------------ 972
B.JP.05.DR6538              ----------.-------V---------ST--K--------------------------------------------------------------------------------------------------------E-------------------------------- 979
B.KR.05.05CSR3              -----I----.-----------------ST--K-------V---------------------------A--------------------------------------------------------------------D-I---------------------E-------- 979
B.NL.00.671_00T36           ----------.-------------P--ISTA-K---------------------------------------E-------------------------E-------I------------------------------D-------------------------------- 989
B.RU.04.04RU128005          ----------.-----------------SN--K--------------------------------------------------------------------------------------------------------D----------------X-----TE-------- 979
B.TH.00.00TH_C3198          ----------.-------------R---SNS-K---------------------------------------------------------Y-------------------------R--------------------D-------------------------------- 979
B.UA.01.01UAKV167           ----------.------IV---------SNA-K---------------------------------------------------------------------------M----------------------------D-------------------------------- 986
B.US.04.ES10_53             -----V----.-----T-----------STA-K---------------------------------------E--------------------------------------------------------------N-D-------------------------------- 979
B.US.99.PRB959_03           -----I----.-----R-V-----P---S-A-K--------------------------------------------------------------------------------------------------------D-------------------------------- 979
C.AR.01.ARG4006             -----I----.------V----------SNA-K--------Q----------------------------------------------------------------I----------------I-------------D-I----------------L------------- 975
C.BR.04.04BR013             -----I----.------V----------SN--K--------Q----------------------------------------------------------------I----------------I-------------D-I----------------L------------- 983
C.BW.00.00BW07621           ----YI----.------I----------S-A-K--------Q---------------------------------------------------R------------I---------R----R-IQ------------D-I-------------------------I---- 975
C.CN.98.YNRL9840            -----I----.------V----------S-A-K--------Q----------------------------------------------------------------I----S-----------I-------------D-I----------------------V------- 975
C.ET.02.02ET_288            ----YI----.------V----------SSA-K--------T--------------------------V------------------------R----D-C-----I-------------N--L-------------D-I------------------------------ 981
C.GE.03.03GEMZ033           ----YI----.------V----------SSA-K--------Q----------------------------------------------------------------I----------------I-------------D-I------------------------------ 975
C.IL.99.99ET7               ----YI----.------V----------S-A-K--------Q------------------------------E---------------------------------I----------------I-------------D-I------------------------------ 975
C.IN.99.01IN565_10          ----YI----.------V----------S-A-K--------Q----------------------------------------------------------------I----------------I-------------D-I--------------------NE-------- 975
C.KE.00.KER2010             ----YI----.------V----------S-A-K--------Q----------------------------------------------------------------I---------R------IQ------------D-I------------------------------ 975
C.MM.99.mIDU101_3           -----I----.------V----------S-AAK--------H-----------------------R----------------------------------------I----------------I-------------D-I------------------------------ 975
C.MW.93.93MW_965            ----YI----.------V----------S-A-K--------Q--------------------------------------------L-------------------I---------R------I-------------D-I------------------------------ 976
C.SN.90.90SE_364            -----I----.------V----------S-A-K--------Q------------------------------E--------------------R------------I----------------I-------------D-I------------------------------ 971
C.SO.89.89SM_145            ----YI----.------V----------SNA-K--------H---------------------------------------------------R------------I----------------I-------------D-I------------------------------ 986
C.TZ.02.CO178               ----YI----.------V----------STA-K--------Q------------------------------E--------------------R------------I------------------------------D-I------------------------------ 975
C.UY.01.TRA3011             -----I----.-----QV------P---SN--K--------Q----------------------------------------------------------------I---------#------I-------------D-I----------------L------------- 974
C.YE.02.02YE511             ----YI----.------V----------S-A----------Q------------------------------E---------------------------------I----------------I-------------D-I------------------------------ 973
C.ZA.04.04ZASK164B1         ----YI----.------V----------STA-K--------Q------------------------------E---------------------------------I------------------------------D-I------------------------------ 975
C.ZA.05.05ZAPSK240B1        ----YI----.------VL---------S-A-K-------------------------------------------------------------------------I----S-----------I-------------D-I------------------------------ 975
C.ZM.02.02ZM108             -----I----.------V------T---S---K--------Q----------------------------------------------------------------I----------------I-----------N-D-I----------------LR------------ 970
D.CD.83.ELI                 ----------.------VV---------S-A-K------------------------------------------------------------RR-----------I----------------I-------------D-I-------------------K---------- 978
D.CM.01.01CM_0009BBY        ----Y-----.------VV---------S-A-K-------------------------------------------------------------------------I---------R----H--N----------N-D-I------------------------------ 980
D.KE.01.NKU3006             ----------.------VV---------S-A-K-------------------------------------------------------------------------I----------------I-------------D-I---------------------E-------- 978
D.KR.04.04KBH8              ----------.------V----------S-V-K-----------------------------D---------E---------------------------------I---------R--------------------D-I-------------------S---------- 978
D.TD.99.MN011               ----------.-------V-----R---S---K-----------------------------------E-------------------------------------I---------R--------------------D-I------------------E----------- 980
D.TZ.01.A280                -----I----.------V----------S---K----------------------------R--------------------------------------------I----S-L--T--------------------D-I-------------------H---------- 980
D.UG.99.99UGK09259          -A--------.-----RV----------SNALK-----------------------------------K-------------------------------------I------------------------------D-I------------------------------ 978
D.YE.01.01YE386             ----------.------V------G--IS-A-K-------------------------------------I-----------------------------X-----I------------------------------D-I---------------------E-------- 978
D.YE.02.02YE516             ----------.------VV---------SGA-K-------------------------------------------------------------------------I------------------------------D-I------------------------------ 980
D.ZA.90.R1                  ----------.------VV---------S-A-K------------------------------------------------L------------------------I------------------------------D-I------------------------------ 978
F1.AR.02.ARE933             -----I----.----A--------P---S-A-K-S------Q--------------------------------------------------------------------S-----R------I-------------D-V---------------------E-------- 979
F1.BE.93.VI850              -----I----.------I----------S-A-K-S------Q-----------------I----------------------------------------------I---S-----R----------------------V---------------------E--I----- 980
F1.BR.01.01BR125            D----I----.------I------P---S-A-K--------Q--------------------------K-------------------------------------I---S-----R------I-------------D-V---------------------E-------- 979
F1.ES.x.P1146               -----I----.------R----------SSA-K------D-Q----V---------------------------------------------------E-------I----------------I-------------D-V---------------------E-------- 977
F1.FI.93.FIN9363            D----I----.------M----------S-A-K--------Q----------------------------------------------------------------I---------------V--------------D-V---------------------E-------- 978
F2.CM.02.02CM_0016BBY       -----I----.------I----------STV-K--------Q----------------------------------------------------------------I----------------S--------F----D-I--------------------NE-------- 979
F2.CM.95.MP257              -A---I----.------I----------S-V-K--------Q----------------------------------------------------------------I----------------S--------F----D-V--------------------NE---I---- 979
F2.CM.97.CM53657            ----------.------I----------STV-K--------Q----------------------------------------------------------------I-------------------------F----D-I--------------------NE-------- 979
G.BE.96.DRCBL               -----I----.------I----------S-A-K------S-T----------------------------------------------------------------I----S-------------------------D-I--------------------NE-------- 975
G.CM.01.01CM_4049HAN        --G--I----.------V----------S-A-K------N-T--------------------------A-------------------------------------I----S-------------------------D-I--------------------NE-------- 978
G.CU.99.Cu74                -----I----.------V----------S-A-K------N-T----------------------------------------------------------------I----S--------------------F----D-I--------------------NE-------- 975
G.ES.99.X138                -----I----.------I----------STA-K------N-T----------X-----------------------------------------------------I----S-----A--------------F----D-I--------------------NE-------- 977
G.GH.03.03GH175G            -----I----.------VL---------S-A-K------N-T----------------------------------------------------------------I----S--------------------F----D-I--------------------NE-------- 979
G.KE.93.HH8793_12_1         -----I----.------V----------S-A-K---R--D-T-----------------------------------------------------------S----I----S-----------XQ-------F----D-V--------------------NE-------- 976
G.NG.01.01NGPL0669          -----I----.------V----------SSA-K------S-T----------------------------------------------------------------I----S-------------------------D-I---------------------E-------- 978
G.PT.x.PT2695               -----I----.-----RI----------S-A-K------D-T----------------------------------------------------------------I----S--------------------F----D-I--------------------NE-------- 978
G.SE.93.SE6165              -----I----.-----TV----------S-A-K------N-T----------------------------------------------------------------I----S-------------------------D-V--------------------NE-------- 976
H.BE.93.VI991               -----I----.------M----------S-A-K------D-H------------------------------------R-----------------------R---I-------P--------SQ--K---------D-I---------------------E-------- 977
H.BE.93.VI997               -----I----.------M------T---STA-K------D-Q-D----------------------------------R---------------------------I----------------SN--K---------D-I---------------------E-------- 976
H.CF.90.056                 ----------.S-----V----------S-A-K------D-Q----------------------------------------------------------------I----------------SN--K---------D-I---------------------E-------E 978
J.CD.97.J_97DC_KTB147       ---F-I----.-----RV----------S-A-K-------V-------------------------------E---------------------------------I----S-----------I-------------D-I---------------------E-------- 976
J.SE.93.SE7887              -A-F-I----.------V----------SGA-K------D--------------------------------E---------------------------------I---------R--------------------D-I-------P-------------E-------- 977
K.CD.97.EQTB11C             -----I----.-----RV----------S-V-K------D--------------------------------E---------------------------------I------------------------------E-I------------------.#-E-------- 977
K.CM.96.MP535               -----I----.------V------T---STV-K-------V----------------------------------------------------------------------------------LN--K---------E-I---------------------E-------- 979
```

```
B.FR.83.HXB2                 ETAYFLLKLA.GRWPVKTIHTDNGSNFTGATVRAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRNPLWKGPAKLLWKGEGAVVIQDNSDIKVVPRRK  979
01_AE.CF.90.90CF11697        ----------.--*---V----------S-AMK------NVQ------------------------V-----E---------------------------------I-----E-----X------------------D-I------------------------------  976
01_AE.CN.05.FJ051            ----------.------V----------S-A-K------DVR------------------------------E----P----------------------------I------------------------------D-I------------------------------  979
01_AE.CN.06.FJ054            ----------.----A-V----------S-A-K------N-R------------------------------E--------L------------------------I------------------------------D-I-------------------D----------  981
01_AE.HK.x.HK001             ----------.------V----------S-A-K------NVR--------------------D---------E---------------------------------I------------------------------D-I------------------------------  976
01_AE.TH.01.01TH_R2184       ----------.------V----------S-A-K------NVR------------------------------E---------------------------------I---------R--------------------D-I------------------------------  979
01_AE.TH.02.OUR769I          ----------.------V----------S-A-K------NV-------------------------------E---------------------------------I------------------------------D-I------------------------------  979
01_AE.TH.90.CM240            ----------.------V----------S-A-K------NVQ------------------------------E---------------------------------I------------------------------D-I------------------------------  984
01_AE.US.00.00US_MSC1164     ----------.------V----------S-A-K------NV---------------------------E---E---------------------------------I------------------------------D-I--------------------N---------  979
02_AG.CM.02.02CM_4082STN     -----I----.-----QV---------IS-A-K------D------------------------------------------------------------------I----S-------------------------D-I------------------------------  982
02_AG.EC.x.ECU41             -----I----.------I----------S-A-K------NVT----------------------------------------------------------------I----S-----------I-------------D-I----------------V-------------  979
02_AG.GH.03.GHNJ196          -----I----.-----RV----------S-A-K------DV-----------------------------------------------------------------I----S-----------I-------------D-I------------------------------  980
02_AG.NG.01.PL0710           -----I----.-----RV----------S-A-K------N-T----------------------------------------------------------------I----S-------------------------D-I---------------------E--------  979
02_AG.NG.x.IBNG              -----I----.------V----------S-A-K------NVT----------------------------------------------------------------I----S-----------I-------------D-I------------------------------  979
02_AG.SE.94.SE7812           -----I----.------V----------S-A-K------NVT--------------------------------------------L------R------------I----S-------------------------D-I------------------------------  979
02_AG.SN.98.MP1211           -----I----#----f-VV---------S-A-K------KVTPGI-NS---------------------------------------------R------------I----S------------N--K---------D-I------------------------------  978
03_AB.RU.97.KAL153_2         -----V----.------I----------STA-K-----------------------------Q--QT---------------------------------------I----------------I-------------D-I--------------------N---------  979
04_cpx.CY.94.CY032           -----I----.------M--A---P---S-A-K------D-N----------------------------------------------------------------I----S-------------------------E-I------------------------------  978
05_DF.BE.x.VI1310            D----I----.------M------P---SGA-K--------Q------------------------------------------T---------------G-----I---T------------I-------------D-V---------------------E--------  979
06_cpx.AU.96.BFP90           -----I----.------V----------S-A-K------N-T------------------------------E---------------------------------I----S-----------I-------------D-I---------------------E--------  981
08_BC.CN.97.97CNGX_6F        -----I----.------V----------S-A-K--------Q------------------------L-------------------------------------------------R------I-------------D-I------------------------------  975
09_cpx.GH.96.96GH2911        -----I----.------V----------S-A-K------NVT----------------------------------------------------R-----------I----S-------------------------D-I--------------------N---------  978
10_CD.TZ.96.96TZ_BF061       ----------.------VV---------S-A-K-------------------------------------------Q-----------------------------I----------------I-------------G-I------------------E-----------  978
11_cpx.GR.x.GR17             -----I----.------V----------SNA-K--------Q--------------------G--E------E-------------------G-------------I------L-------------K---------D-I--------------I---------------  978
12_BF.AR.99.ARMA159          -----I----.-------------P--IS-M-K--------Q--------------------------S-------------------------------------I---S-----R------I-------------D-V-------------------D-E--------  979
13_cpx.CM.96.1849            -----I----.------V----------S-A-K------N-T------------------------------E---------------------------------I------L---------I-------------D-I---------------------E--------  986
14_BG.DE.01.9196_01          -A---I----.------I---------ISN--K--------------------------------------X----------------------------T----------S--------------------F----D-I--------------------NE--------  978
15_01B.TH.99.99TH_MU2079     ----------.------V----------S-AMK------NVR---WD------------I----------I-E---------P-----------------------I-L-----P----PN-.S-------------D-I------------------------------  977
16_A2D.KR.97.97KR004         -----I----.------V------P--IS-P-K-------VQ----------------------------I-------------------S---------------I-------------R----------------D-I-----------------------------R  979
17_BF.PE.02.PE02_PCR0155     -----I----.-------------T---S-A-K--------Q----------------------------------------------------------------I-M-S-----R------I-------------D-V---------------------E--------  979
18_cpx.CU.99.CU76            ----L-----.------V------P---S-A-K------N-T----------------------------I----------------------------------------SE-----F------------------D-I---------------------E--------  977
19_cpx.CU.99.CU7             ----------.A-----V------P---S-A-K------NVQ----------------------------------------------------------------I----S--------------S----------D-I------------------------------  978
20_BG.CU.99.Cu103            -----I----.------V------P--IS-A-K------N-T----------------------------------------------------------------I----S-------------------------D-I--------------------NE--------  975
21_A2D.KE.91.KNH1254         ----WI----.------V------P--ISTP-K-----------------------------------------------------------------D-------I-------------R----------------D-I------------------E-----------  977
22_01A1.CM.01.01CM_0001BBY   -----I----.------V----------S-A-K------E-Q----------------------------I-E---------------------------------I------------------------------D-I------------------------------  979
23_BG.CU.03.CB118            -----I----.------V---X--P--IS-A-K------D-T----------------------------------------------------------------I------------------------------D-I-------------------SNE--------  975
24_BG.CU.03.CB378            -----I----.------V------P--IS-A-K------D-T----------------------------------------------------------------I----S--------------------F----D-I--------------------NE--------  970
24_BG.CU.03.CB471            -----I----.------V------P---S-A-K------S-T----------------------------------------------------------------I---SS-------------------------D-I--------------------NE--------  975
25_cpx.CM.01.101BA           ----------.------V------P--IS-A-K---------------------------------------E---------------------------------I----S-----------I-------------D-I---------------------E--------  978
27_cpx.CD.97.97CDKTB49       -----I----.------V------R---STA-K------NVT----------------------------------------------------------------I----S-------------------------D-I---------------------E--------  977
28_BF.BR.99.BREPM12609       ----------.-------V---------SNA-K------E----------------I-------------------------------------------------I---S-----R--------------------D-V------------------------------  979
29_BF.BR.02.BREPM119         --G--I----.-----R-------P--ISTA-K-----------------------------------E-I-----------------------------------I---S-----R------------------N-D-I---------------------E--------  979
31_BC.BR.02.110PA            -----I----.-----RV----------SN--K--------Q----------------------------------------------------------------I----------------I-------------D-I----------------L-------------  975
32_06A1.EE.01.EE0369         -----I----.------V----------S-A-K------N-T----------------------------------------------------------------I----S-------------------------D-I---------------------E--------  979
33_01B.MY.05.05MYKL007_1     -I--------.------V----------S-A-K------NVQ------------------------------E---------------------------------I------L-----------------------D-I------------------------------  979
34_01B.TH.99.OUR2478P        ----------.------V----------S-A-K------NVR----------------------------------------------------------------I------------------------------D-I------------------------------  979
35_AD.AF.05.05AF095          ----------.------VV---------S-AFK------S-Q------------------------------E---------------------------------I------------------------------D--------------------------------  978
36_cpx.CM.00.00CMNYU830      ----------.------V----------S-A-K------D-Q----------------------------------------------------------------I----S-------------------------D-I------------------------------  979
37_cpx.CM.00.00CMNYU926      ----------.-----RV----------S-A-K------NV-----------------------------------------------------------------I----S-------------------------D-I------------------------------  976
38_BF1.UY.05.UY05_4752       -----I----.-----R-------P---S-A-K--------Q--------------------------Q-------------------------------------I---S-----R------I----X--------D-V---------------------E--I-----  979
39_BF.BR.03.03BRRJ103        -----I----.-----------------ST--K----------------------------N------------------------------------------------------R--------------------D--------------------------------  984
40_BF.BR.04.04BRRJ115        ----------.------IV-----P---SN--K--------Q----------------------------------------------------------------I---S------------S-------------D-V---------------------E--------  979
42_BF.LU.03.luBF_05_03       -----I----.-------------X--XSTV-K-----------------------------------E-------------------------------------I------------------------------D-V------------------------------  983
43_02G.SA.03.J11223          D----I----.------V----------S-A--------SVT--------------------X-----------------------------------------X-I----S-L-----------------------D-I---------------------E--------  978
U.CD.83.83CD003              -----V----.-----TV----------S-A-K--------Q----------------------------------------------------------------I------------------------------D-I---------------------E--------  980
U.CD.90.90CD121E12           D----V----.-----T-----------S-A-K--------Q------------------------------E---------------------------------I----------------IQ------------D-I---------------------E--------  979
U.GR.99.GR303                ----------.------V----------S-AMK------NVT--------S-------------------------------------------------------I---------R------I-------------D-I---------------------E--------  977
U.NL.95.U_NL_95_H10986_D1    ----YI----.------V------P---S-A-K-------------------------------------------------------------------------I-M-------T--------------------D-I---------------------E--------  979
CPZ.CD.90.ANT                S---------.S------------A---S-A-K------N-Q--------------------Q--Q----I-----Q-----V---H-------------TP-Q--L--L------TQ--N--L---Q---H-----D-V-----Q-----------K-QEE--------  979
CPZ.CM.01.SIVcpzCAM13        -----I----.S-----VL-----T---SNA-K------N---K------------------------E------------------S------------T-----I----------------L---K---------D-I-----T-------------QGE--------  978
CPZ.CM.05.SIVcpzEK505        -----I----.------V----------S---K------N-Q-------------A--------------I-----------------------------T-----I---------TK--T--L-V-----------E-T--------------------G---------  985
CPZ.CM.05.SIVcpzLB7          -----I----.S-----I----------STA-K------Q-Q----------------------------------------------------------------I---------T------S-V-K-------R-D-I----------------L--QEE--------  982
CPZ.CM.05.SIVcpzMB66         -----I----.------V----------SS--K--------Q----------------------------I-----------------------------------I--L------TK--Q--SN--K---------D-I----------------L--QEE--------  982
CPZ.CM.05.SIVcpzMT145        -----V----.-------------P---SNA-K--------R--------------------------E-------------------------------------I----------------L---Q--------KD-I-------------------QDE--------  978
CPZ.GA.88.GAB1               ----------.-------------P---S-A-K------D-------------------L----------------------------------------T-----I---------S------L-V-K---------D-I-----T-------------QGEL-------  977
CPZ.TZ.00.TAN1               -----I----.------K------P---STA-K------Q-QH-------------------Q--Q--E-------Q-R---I---Y-----------E-T----LL--LT-N----Q-----L-V---------A-D-I-----R-----------KEGE---------  975
CPZ.US.85.CPZUS              -----I----.------V----------SS-----------Q----------------------------I-----------------------------T-----I----SEL--DL-----L-V-----------D-I-----------------KE-EEV-------  981
CPZ.CM.x.SIVcpzDP943         -----I----.------V----------SN--K--------Q----------------------------I-E---------------------------T-----I---------TN-----L-V-----------D-I-----R-----------K-KEEV-------  986
CPZ.CM.x.SIVcpzMB897         -----V----.------V----------S-A-K------N-Q----------------------------------------------------------------I---------R------S-V-K---------E-I----------------L--KE---------  987
N.CM.02.DJO0131              -----I----.------V------P---S---K------N-T----------------------------I---------------------------E-T-----I---------TK--T--L-------------D-I--------------------G---------  984
N.CM.04.04CM_1015_04         -----I----.------V------P---S---K------S---------------A--------------I-----------------------------T-----I---------TK--T--L-V-----------D-I--------------------G---------  986
N.CM.04.04CM_1131_03         -----I----.------V------P---S---K------S---------------A--------------I-----------------------------T-----I---------TK--T--L-V-----------D-I--------------------G---------  986
N.CM.95.YBF30                -----I----.------V----------S---K------N---------------A--------------I-----------------------------T-----I---------TN--T--L-V-----------D-I--------------------G---------  986
N.CM.97.YBF106               -----I----.------V------P--IS---K----------------------A--------------I----------------------------XT-----I---------TK--T--L-V---X-------D-I--------------------G---------  986
O.BE.87.ANT70                ----------.A-----V------P---ST-MK------N-QH---------------A------S--Q---------R--------V------------T-----I--L-SQ---T------L-XHK---------D-I-----Q-------------KG---------  975
O.CM.91.MVP5180              ----------.A-----V------P---S-AMK-----T--QH---------------A------S--Q------------------V------------T----LI--L-SQ---T------L--N----------D-I-----Q-------------KG---------  975
O.CM.96.96CMABB637           ----------.A---IRVL-----P---S--MK------N-RH---------------A------S--Q------------------V-----------CT----LI-ML-SQ---T----H-L---K-Q-------D-I-----Q-------------KG---------  977
O.CM.98.98CMA104             ----------.A-----V------P---S--MK------N-XH---------------A------S--E---------R---------------------T----LI--L-SQ---T------L-------------D-I-----Q-------------KG---------  975
O.CM.99.99CMU4122            ----------.A-----V------P---S--MK-----TN--H---------------A------S--Q---------R--------V------------T-----I--L-SQ---T------F---K-Q-------D-T-----Q-------------KG---------  975
O.FR.92.VAU                  ----------.A---I-IL-----P---S-AMK------N-QH---------------------ES-MQ------------------V------------T----LI--L-SQ---T------L-------------D-I-----Q-------------KG---------  976
O.SN.99.SEMP1299             ----------.A-----V------P---S-AMK------N--H---------------A------S--Q------------------V--Y---------T-----I--L-SQ---T------F---K-Q-------D-I-----Q-------------KG---------  975
O.US.99.99USTWLA             ----------.A-----I-----RP---S--MK-----T--QH--X--------X---A------S--Q------------------V------------T----LI--L-SQ---T------L-------------D-I-----Q-------------KGE--------  978
```

```
                                p31 Integrase end
                                      Pol end
B.FR.83.HXB2            AKIIRDYGKQMAGDDCVASRQDED*  1003
A1.GE.99.99GEMZ011      ------------------------   1004
A1.KE.00.KER2008        ------------------G-------  1004
A1.KE.00.KNH1144        ------------------G-------  1004
A1.KE.00.KSM4024        ------------------G-------  1004
A1.KE.00.MSA4069        -----E------------G-------  1004
A1.KE.00.NKU3005        ----L-------------G-------  1004
A1.RU.00.RU00051        ------------------------   1004
A1.RW.93.93RW_024       V-----------------G-------  1004
A1.SE.95.SE8891         --------------G---G-------  999
A1.SE.95.UGSE8131       ------------------------   997
A1.TZ.01.A173           ------------------G-------  1004
A1.UA.01.01UADN139      ------------------G-------  1004
A1.UG.92.92UG037        V---K-------------G-------  1005
A1.UG.99.99UGA07072     ------------------G-------  1004
A2.CD.97.97CDKTB48      ------------------G-------  997
A2.CY.94.94CY017_41     ------------------G-------  1004
A.SN.01.DDI579          ------------------G-------  1003
A.SN.01.DDJ369          ------------------G-------  1003
A.SN.96.DDJ360          ------------------G-------  1003
A.CD.02.02CD_KTB035     --------------G---G-------  1004
A.CD.97.97CD_KCC2       ------------------G-------  1004
A.CD.97.97CD_KTB13      ------------------G-------  1004
B.AR.04.04AR151516      ------------------------   1009
B.AU.87.MBC925          ------------------------   1006
B.BO.99.BOL0122         ------------------------   1005
B.BR.03.BREPM2012       V---K-------------------   1008
B.CA.97.CANB3FULL       ------------------------   1007
B.CN.05.05CNHB_hp3      ------------------------   1004
B.CO.01.PCM001          ------------------------   1004
B.GB.83.CAM1            ------------------------   1004
B.GB.86.GB8             ------------------------   1010
B.GE.03.03GEMZ010       ------------------------   1004
B.IT.05.SG1             ------------------------   997
B.JP.05.DR6538          V-----------------------   1004
B.KR.05.05CSR3          ---F--------------------   1004
B.NL.00.671_00T36       ------------------------   1014
B.RU.04.04RU128005      ------------------------   1004
B.TH.00.00TH_C3198      ------------------------   1004
B.UA.01.01UAKV167       ------------------------   1011
B.US.04.ES10_53         ------------------------   1004
B.US.99.PRB959_03       ------------------------   1004
C.AR.01.ARG4006         V---K--------A-----------  1000
C.BR.04.04BR013         V---K--------A-------N--Q  1008
C.BW.00.00BW07621       ----K--------A----G------  1000
C.CN.98.YNRL9840        ----K--------A--E-G------  1000
C.ET.02.02ET_288        ------------A----G----N-   1006
C.GE.03.03GEMZ033       V-----------A----G------  1000
C.IL.99.99ET7           ----K--------A----G------  1000
C.IN.99.01IN565_10      ----K--------A----G------  1000
C.KE.00.KER20I0         ----K--------A----G------  1000
C.MM.99.mIDU101_3       ----K--------A----G------  1000
C.MW.93.93MW_965        V---K--------A-----G-----  1001
C.SN.90.90SE_364        V-----------A----G------  996
C.SO.89.89SM_145        ------------A----G------  1011
C.TZ.02.CO178           V---K--------A-----------  1000
C.UY.01.TRA3011         V---K--------A--M-N----N-  999
C.YE.02.02YE511         V---K--------A----G------  998
C.ZA.04.04ZASK164B1     ----K--------A----G-----Q  1000
C.ZA.05.05ZAPSK240B1    ----K--------A----G------  1000
C.ZM.02.02ZM108         V-----------A----G-----Q  995
D.CD.83.ELI             V-----------------------   1003
D.CM.01.01CM_0009BBY    ------------------------   1005
D.KE.01.NKU3006         -----------------------Q  1003
D.KR.04.04KBH8          ------------------------   1003
D.TD.99.MN011           --------------G---------  1005
D.TZ.01.A280            ----------V-------------  1005
D.UG.99.99UGK09259      ------------------------   1003
D.YE.01.01YE386         ------------------------   1003
D.YE.02.02YE516         V-----------------G------  1005
D.ZA.90.R1              V-----------------------   1003
F1.AR.02.ARE933         ---------------M-G------  1004
F1.BE.93.VI850          ------------V-----G------  1005
F1.BR.01.01BR125        ------------------G------  1004
F1.ES.x.P1146           --------------G---G--N---  1002
F1.FI.93.FIN9363        ------------------G------  1003
F2.CM.02.02CM_0016BBY   V-----------------G------  1004
F2.CM.95.MP257          ------------------G------  1004
F2.CM.97.CM53657        ------------------G------  1004
G.BE.96.DRCBL           ------------------G------  1000
G.CM.01.01CM_4049HAN    ----K--------G----G------  1003
G.CU.99.Cu74            --------------A---G------  1000
G.ES.99.X138            ---------------G------  1002
G.GH.03.03GH175G        ----K-------------G------  1004
G.KE.93.HH8793_12_1     ------------------G------  1001
G.NG.01.01NGPL0669      ------------------G------  1003
G.PT.x.PT2695           ------------------G------  1003
G.SE.93.SE6165          ------------------G------  1001
H.BE.93.VI991           ------------------G------  1002
H.BE.93.VI997           ------------------G------  1001
H.CF.90.056             ------------------------   1003
J.CD.97.J_97DC_KTB147   ------------------G------  1001
J.SE.93.SE7887          ------------------G------  1002
K.CD.97.EQTB11C         ------------------G------  1002
K.CM.96.MP535           ------------------G------  1004
```

p31 Integrase end
Pol end

B.FR.83.HXB2	AKIIRDYGKQMAGDDCVASRQDED*	1003
01_AE.CF.90.90CF11697	----------G----N-	1001
01_AE.CN.05.FJ051	----------G------	1004
01_AE.CN.06.FJ054	----------G------	1006
01_AE.HK.x.HK001	----------G------	1001
01_AE.TH.01.01TH_R2184	----------G------	1004
01_AE.TH.02.OUR769I	----------G------	1004
01_AE.TH.90.CM240	----------G------	1009
01_AE.US.00.00US_MSC1164	----------G------	1004
02_AG.CM.02.02CM_4082STN	----------G------	1007
02_AG.EC.x.ECU41	V----------G----N-	1004
02_AG.GH.03.GHNJ196	----K----------G------	1005
02_AG.NG.01.PL0710	----------G------	1004
02_AG.NG.x.IBNG	----------G------	1004
02_AG.SE.94.SE7812	V--V----------G------	1004
02_AG.SN.98.MP1211	T--L----------GG-N---	1003
03_AB.RU.97.KAL153_2	----------	1004
04_cpx.CY.94.CY032	----------N----G------	1003
05_DF.BE.x.VI1310	----------G------	1004
06_cpx.AU.96.BFP90	----------G------	1006
08_BC.CN.97.97CNGX_6F	----K--------A----G------	1000
09_cpx.GH.96.96GH29I1	-----H----------G------	1003
10_CD.TZ.96.96TZ_BF061	V---K--------A----------Q	1003
11_cpx.GR.x.GR17	----------GS-----	1003
12_BF.AR.99.ARMA159	----------G------	1004
13_cpx.CM.96.1849	----------G------	1011
14_BG.DE.01.9196_01	----------G------	1003
15_01B.TH.99.99TH_MU2079	----------G------	1002
16_A2D.KR.97.97KR004	----------G------	1004
17_BF.PE.02.PE02_PCR0155	----------G------	1004
18_cpx.CU.99.CU76	----------G------	1002
19_cpx.CU.99.CU7	----------G------	1003
20_BG.CU.99.Cu103	----------G------	1000
21_A2D.KE.91.KNH1254	----K--------G-------N---	1002
22_01A1.CM.01.01CM_0001BBY	-----E----------G----N-	1004
23_BG.CU.03.CB118	----------G------	1000
24_BG.CU.03.CB378	----------G----N-	995
24_BG.CU.03.CB471	----------G------	1000
25_cpx.CM.01.101BA	----------M-G------	1003
27_cpx.CD.97.97CDKTB49	----K----------G------	1002
28_BF.BR.99.BREPM12609	----------	1004
29_BF.BR.02.BREPM119	----------	1004
31_BC.BR.02.110PA	V---K--------A--M--------	1000
32_06A1.EE.01.EE0369	-----E--------E---G------	1004
33_01B.MY.05.05MYKL007_1	----------G----------	1004
34_01B.TH.99.OUR2478P	----------G------	1004
35_AD.AF.05.05AF095	----------G------	1003
36_cpx.CM.00.00CMNYU830	-----N----------G------	1004
37_cpx.CM.00.00CMNYU926	----------M-G------	1001
38_BF1.UY.05.UY05_4752	----------G------	1004
39_BF.BR.03.03BRRJ103	----------G------	1009
40_BF.BR.04.04BRRJ115	----------	1004
42_BF.LU.03.luBF_05_03	----------	1008
43_02G.SA.03.J11223	----------N----G------	1003
U.CD.83.83CD003	----------N-	1005
U.CD.90.90CD121E12	----------N-	1004
U.GR.99.GR303	----------G------	1002
U.NL.95.U_NL_95_H10986_D1	----------V----------	1004
CPZ.CD.90.ANT	----KE-R-KIEDR-DL-G--N---	1004
CPZ.CM.01.SIVcpzCAM13	----------	1003
CPZ.CM.05.SIVcpzEK505	----------G----G---SQ	1010
CPZ.CM.05.SIVcpzLB7	----------E-AN-L-D----N-	1007
CPZ.CM.05.SIVcpzMB66	----------NQ	1007
CPZ.CM.05.SIVcpzMT145	----------A-S-------S-	1003
CPZ.GA.88.GAB1	----------N---	1002
CPZ.TZ.00.TAN1	----KE-------AGGMDD--N-T-	1000
CPZ.US.85.CPZUS	----K--------A-SM-G----S-	1006
CPZ.CM.x.SIVcpzDP943	----------GSM--G---SQ	1011
CPZ.CM.x.SIVcpzMB897	----------VE-N--L-D----S-	1012
N.CM.02.DJO0131	----------G----G-A-NQ	1009
N.CM.04.04CM_1015_04	----------G----G---NQ	1011
N.CM.04.04CM_1131_03	----------G----G---NQ	1011
N.CM.95.YBF30	----------G----G---NQ	1011
N.CM.97.YBF106	----------G----GX--NQ	1011
O.BE.87.ANT70	-----E-------T-SM--G-T-SE	1000
O.CM.91.MVP5180	----------T-SM-N--T-SE	1000
O.CM.96.96CMABB637	---L---------A-SM--G-TXSE	1002
O.CM.98.98CMA104	-----H-------A-SM--G-T-SE	1000
O.CM.99.99CMU4122	-----N-------T-SM--G-X-SE	1000
O.FR.92.VAU	----------A-S---G-T-GE	1001
O.SN.99.SEMP1299	-----H-------T-SM--G-T-SE	1000
O.US.99.99USTWLA	----------T--M--G-T-SE	1003

HIV-1/SIVcpz Proteins

```
                      Vif start
B.FR.83.HXB2          MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHYESPHPRISSEVHIPLGD..ARLVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPELADQLIHLYYFDCFSDSAIRKALLGHIVSPRCEYQAGHN.KVGSLQYLALAALITPK.....KIKPP  162
A1.GE.99.99GEMZ011    --------------------N----------K---E---------R--K-----------..----VK------A--K-----H-------Q------I--D--------H------E------IV------K-N-PP--S.----------K--V--T.....RER--  162
A1.KE.00.KER2008      --------------------N---------I-K--K----------R--KV----------..-K--VR------P--K-----H------L------------------H--N---E------I--QV---S--------.----------K--V---.....-T---  162
A1.KE.00.KNH1144      ---K----------------N----------K--KN-V----F--RN--V-----------..-K--VR------P--N-----H------L-S------D---------------G------I--EL--------T--S.----------K--V--Q.....-T---  162
A1.KE.00.KSM4024      --------------------N----------K--KN------F--R--KV-----------..-K--VI---------K--------------RG-----I--D--------H--A--T-------I--QV-R-----PT---.----------K--V--T.....-TR--  162
A1.KE.00.MSA4069      --------------------N----------K--SK-W--------R--KA-------V-E.------R-----N---KA-Q--H-------LRK-------DQ---------------E------I--QV-----D--T---.----------K--V---.....-T---  162
A1.KE.00.NKU3005      --------------------N--------H--K--K-----------R---V-----------..--I--R---------K--Q--H-------L-----LT-DM-------H--N----------I--QV--------T---.----------K--V--P.....R-R--  162
A1.RU.00.RU00051      --------------------N-----------------L-------R---V-----------..----VR---------K--Q--H-------Q--X---I--D-------------------------E--R-----------.-----X----K--V--T.....RGR--  162
A1.RW.93.93RW_024     --------------------N----------K-TKN---------R---V-----------..---IVR---------K-----H-------L------I--D--------H------E---R-I--QV---S--HT---.Q---------K--V--T.....-TR--  162
A1.SE.95.SE8891       --------V-----------N---------I-K--K--L---F--R--KV---------E.----VR-----Q---K-----H-------LR----I--D--------H--K----------I--E-------------.----------K--V---.....--R--  162
A1.SE.95.UGSE8131     --------------------N---------I-K--A----------R--KV-----------..---IVR------P-----Q--H-------LR----I--D--------H--N----------I--RV---S----T---.----------K--V---.....RT---  162
A1.TZ.01.A173         --------------------N----------R--KD---------R--KV-----------..----VR-----Q---K-----H-------L------I--D--------H--------------------I--QV-R---D-------.----------K--V---.....-----  162
A1.UA.01.01UADN139    -----------------X--N----------K---K-V-K-----IN-KV-----------..----VK---------K-----H-------Q------I--D----------------E------IV-Q-------N-P-----.---P------K--V--I.....RER--  162
A1.UG.92.92UG037      --------------------N---------I-RR-K----------R--KV-------I--..--I-VR-----Q---K-----H-------L------I--D--------H--N----------I--QV-----D--T---.----------K--V--S.....R----  162
A1.UG.99.99UGA07072   --------------------N----------K--K-------F--R--KV-----------..----V------Q---K--Q--H-------L------I--D--------H--N--------R-I--QV---S---------.----------R--V--T.....RTR--  162
A2.CD.97.97CDKTB48    --------------------N----------K---E---------R---V--------KE..---IVR------P--K-----H-------QG-----I--D--H-----------E----R-I--E--R----------.----------R--VAST.....RT---  162
A2.CY.94.94CY017_41   --------------------N---------I-R--K--V-K-----RN-----------E..--IIVR------I--K-----H-------QN--H--I--D--H-----------E------II-E-------------.----------K-VVAST.....RT---  162
A.SN.01.DDI579        --------------------N---------I-K--KD-------DCR---A-----------..-M--VR------P--Q-----H-------Q------I--D--------H----------R-I--Q--R------T---.----------R--VAS-.....-TR--  162
A.SN.01.DDJ369        -----------------K--N----------R--KK-V-K-----R--KV---Y----E..-I--VR------P-----Q--H-------Q------I--D--------H--N--------------Q--R-S---------.----------K--V--E.....RT---  162
A.SN.96.DDJ360        --------------------N---------I-K--KN---------R---V-----------..-I--VR------P--Q-----H-------Q------I--D--------H--N----------I--Q--R-----PT---.----------R--V---.....-----  162
A.CD.02.02CD_KTB035   --------V--------K--N---------I-K--K--V----F--R--KA-----------..---IVR-----Q---------H-----QL------I--D--------H--E----------I--N------D-------.----------K--VA--.....GT---  162
A.CD.97.97CD_KCC2     ------------------A-N--------H--K--I----------R--K------------..----VK----------G---H-----KLR-----I--D--------H--T---E------I---R---K---------.----------K--V--T.....RA---  162
A.CD.97.97CD_KTB13    --------------------N---------I-K--K----------RP-K------------..----VR-----QP-----Y--H-----KMR-----I--D--------H--------------------I------------------.----------K--VQ--.....RR---  162
B.AR.04.04AR151516    -----------------S-----------T----KR-I-------TN-------------..---I-----------E-------------EK------DI-------------------------------------R-S--------.-------------VK--.....RKR--  162
B.AU.87.MBC925        ----------------------------------K--C-------TN-----------E..-K---V-----------------------------G-D---------------E-----------R-----------.-----------------.....-----  162
B.BO.99.BOL0122       -----------------K----------I-K--K--C------NL-----G------P..-K------------------------E--A---T-DI-------------A------I--R----S-DFR----.-------------V--.....-T---  162
B.BR.03.BREPM2012     ------------------------------I---KE---------T------------..------------------------R-------D-----------N---E----N-I--R----K---------.R-------------R.....RV---  162
B.CA.97.CANB3FULL     -----------------K---------K-IT---KE-I-K----N-----------------..-KV-V-----------------------E-------D-----------------E----N----R----S---------.--------------------.....Q----  162
B.CN.05.05CNHB_hp3    -A---------------------------I-R--K--C-K---D-T--K------------..-K----------------------------RD-----AE-S-----------------E------I--RV--SS-N-PE--S.-------------V---.....RR---  162
B.CO.01.PCM001        -----------------------------I-K--K------------T--K----------E..N---V----------------------------R-------D--------------E----N-I--NL-R-----------.--E-------T------.....-----  162
B.GB.83.CAM1          -----------------------------I----KK-S------T-------------E..----V---------------------------T-G-N-----D-----------------E------IV-RL-------------.----------T---A--.....-----  162
B.GB.86.GB8           -----------------------------I----KK-V-K----NT-------------E..-------------------------------R--R-----G------Q-------E----N----RT---S-K-------.----------T------.....-----  162
B.GE.03.03GEMZ010     -----------------K--H----Y---I-K---D-R-------T--K-------------..-E--VI------S-------------------------DS---------------E----R-II-RL-------------.----------T--VK--.....-K---  162
B.IT.05.SG1           --------------------H---I--K-M-K--KN-V------N-----G----------.GK-------------Q--Q--H-------E-------D--R-----------E----N-I--RR---L---------.-----------------R.....RR---  162
B.JP.05.DR6538        -----------------------Y-----R-TER-V-------T--K-------------.DK-----------------------------R-------N----------E---EA---HTI--R------------T.--------------K--.....-----  162
B.KR.05.05CSR3        -----------------------------I-R--K--V------NI---V------------..-K--V--------------------------G--N-----T------R-----------------I------S----------.----------T------.....-T---  162
B.NL.00.671_00T36     -----------------------Y---R----KK---------T------------E..-K----------------------------ER--T-----D------V------E----N-I--RL-------------.----------V------.....-----  162
B.RU.04.04RU128005    -----------------K--N-I------I-R--KK-V------TN-K-------------..-K---K---------Q-----H-------RG-N-----D----------------EXX--N-II----------PT---.-------------IT--N.....-R---  162
B.TH.00.00TH_C3198    -----------------K-----------I-R--K--I-K---D-S--K------------..-T-----------------------------N-N-----G-----------------E----N-IV-RV---S-D-------.-------------V------.....-R---  162
B.UA.01.01UAKV167     --------------------N--------I-----K-C-K-----TN--T-----------..-K--VI----------E--------------K---I--D-----------------G----S-I--RL---S----T--S.-------------VK--.....RQ---  162
B.US.04.ES10_53       -----------------S------Y----K---------------K----------E..----------------------------R-------------------------E----N-I--Q--------------.--------------------.....RR---  162
B.US.99.PRB959_03     ------------------A----------K--K----------R--------------..-K--A-----------------------------R-------N------M------E----H-I---R---S---------.----------T--V---.....-T---  162
C.AR.01.ARG4006       -------L-----------A-N--------I-K-----Y-------R--KV-----------..------I-----Q---------H-------LXK-------G------MH----T-----------R-------------.----------T--LK--.....-R---  162
C.BR.04.04BR013       -------L--------K----N--------I-KR-K--Y-------R--K-----------E..-K--------XQ----N----H-------LR-------D-------TH----A-----------R-----D-------.----------T---K--.....--R--  162
C.BW.00.00BW07621     -------L--------K----N----------KR-N-----------R--KL-----------..------K-----------------H-----KLR--------G------MH----A----------QV-I---D-------.----------T---K--.....-R---  162
C.CN.98.YNRL9840      -------L--R-----K----N--------I-RR-N--L-----D-R--KV---------E.------K-----Q---------H-------LR--N--IE-G---------------A-------I-----I---D------N----------T---K--.....-----  163
C.ET.02.02ET_288      -------L--------K--A-N--------I-RR-N--L-K-----RN--T---------E.---IVK-----Q---------H-------LRG-T-----C------MH----AE------I--QR-----D-------.----------T--LK--.....-R---  162
C.GE.03.03GEMZ033     -------L-----------K-N----------KR-S-------F--R--KT------------..------K-----Q----E----H-------LRK-------G------MH----A-------I-----I---D-------.----------T---K--.....QR---  162
C.IL.99.99ET7         ------GL--------K-K--N----------KR-S----------R--KV-A----------..------K------------------H-------LR--------G------MH----AK------I-----I---D-------.----------T---K--.....-R---  162
C.IN.99.01IN565_10    -------L--------K----N--------I-RR-K--------D-G--KV------------..------K-----Q---------H-------LRS-----E-G------MH----A-------I-----I---D-------.----------T---K--.....-----  162
C.KE.00.KER2010       -------L-------------N-------HI-RR-S-------F--R--KV---------E.------K-----Q---------H-----KL--N-----G------MH----A-------I-----I-----------.----------T---K-N.....-----  162
C.MM.99.mIDU101_3     -------L--------K----N--------I-RR-K--------D-R--KV---------E.---I-K-----Q----E----H-------LR------E-G---------------A-------I-----I---D------N----------T---K--.....-----  163
C.MW.93.93MW_965      -------L-------------N----------RR-D----------R--KV------------..------K-----Q---------H-------LR--------G------MH----A-------I---------D-------.----------T---K--.....-----  162
C.SN.90.90SE_364      ------GL--------K----N--------T-RR-K----------R--KV------------..------K------------------H-------LR--------G------MH----T-------I-----I-----------.----------T---K--.....-----  162
C.SO.89.89SM_145      -------L-------------N--------I-KR-S----------R--KV---------E.------K-----Q---------H-------LR--------G------MH----A-------I-----I---D-------.----------T---K--.....-----  162
C.TZ.02.CO178         -------L-------------N----------RR-S----------R--KV---------G.------------Q---------H-------LR--------G------MH----A-------I-----I---D-P-----.----------T---K--.....-----  162
C.UY.01.TRA3011       -------L---I---K----N----------KR-S--Y-------R--K-----------E..-K---------Q----E----H-------LR----I--G------MH----A----E-I---R-----D-------.Q---------T---R--.....-----  162
C.YE.02.02YE511       ------GL-------------N--------I-RR-S----------R--KV---------E.------K------------------H-------LR--------G------MH----T-------I-----I---D----T.----------T---K--.....-----  162
C.ZA.04.04ZASK164B1   ------AL-------------N--------T-KR-K--Y-------R--KV---------E.------------Q----E----H-------LR--------G------M-----A-------I-----I---D-------.----------TT--K--.....R----  162
C.ZA.05.05ZAPSK240B1  -------L--------K----N-------HI-KR-K----------RN-KV------------..------K-----Q---------H-------LR--------G------MH----A-----Q-I---------D-------.----------T---K--.....RT---  162
C.ZM.02.02ZM108       -------L--------K--A-N--------I-KR-N----K-----RN-K-------------..-------------------------H-------LR--------G------MH----A-------I---------D-------.--------T-T---K--.....-R---  162
D.CD.83.ELI           -----------------K-------------K--NR-------------K----------E..------K---------E-------------R--------G------M------E------I--D--------------.----------T---A--.....Q----  162
D.CM.01.01CM_0009BBY  -----------------N-------Y--H--K--K-----Q--------K-------------..-K--V-----------------------------------G------M------A-------I---R----D-------.----------T---E--.....-----  162
D.KE.01.NKU3006       --------------------N----Y----K--Y--L------C-N-K----------E.-----VK---------E-------------R--------G------I-----AE-----------T---N-------.----------T---K--.....R----  162
D.KR.04.04KBH8        -----------------N-------Y---I-K-----V----D-N----------------..-K--V-----N---K-----------L--R--I--G------MH-------------------I---R----D-------.----------T------.....R----  162
D.TD.99.MN011         --------V--------N-------Y-----KE-K--------D----K----------E.-----VK----N------------------R--------G------TH-------------------I--QR-----D-------.----------T---K-R.....-R---  162
D.TZ.01.A280          --------------------------------R--Q--L-----D-H--K------------..----V----------E-------------R--------G------M-----T-------I---------D-------.----------I---K--.....-----  162
D.UG.99.99UGK09259    ------------------A-------IHI-K--K--L-----DC---K----------E..----VR---------E-------------R--------G---------------AE------I--Q--T---N-------.----------T---AQ-.....-----  162
D.YE.01.01YE386       --------------------N----------K--Q--L-----DCL--K------------..----VK-----------------------------R--------G------I-----AE------I---V-T---N-------.----------T---K--.....GT---  162
D.YE.02.02YE516       -----------------N-----------I-XE-K--------D---K----------E..-K--VI----N--------------QR--------S------T------E------I--LR-----D-------.----------T---A-R.....RR---  162
D.ZA.90.R1            -----------------N-------Y--H--K--K--------D---K----------E..E--VK---------K---P---------R--I----G------I------E------I---R-----------N---------K---AS-.....-T---  163
F1.AR.02.ARE933       -----------------N-------Y--H--K--KK------F--R---V-------EE..-K---I-----------------------XGN-R--I--G------I-S-----E----G-I---RI----N-------.RL---------T---A-Q.....-T---  162
F1.BE.93.VI850        ------L----------N-------Y-----K--K--S----FQ-R---V-------EE..VK---------P---E------------QGK-R--I--G------I------E------I---RI----N-------.----------T---A-E.....-T---  162
F1.BR.01.01BR125      -----------------N-------Y--H--K--KR-L----F--R---V-------EE..-K--------------------------QG--R--I--G------I------E------I---KI----N-------.Q---------T---A--.....-----  162
F1.ES.x.P1146         --------V--------S-------Y--HI-K--KK-C----F--RN--V-------KE..-KI-V----------E-----------QGK----I--D------I------E------I--YR-----N-------.--K-------T---A--.....-T---  162
F1.FI.93.FIN9363      -------T---------N-------Y--H--K--KR------F--R--K--------EE..-E---------N------------------QG--R--I--G------I------E------I---RI----D-------.----------T--VS--.....-A---  162
F2.CM.02.02CM_0016BBY --------------------N----------RR-Q----------R---V--------EN.DSK---V---------K-----H-------Q---R-----G------K------E------I---R-----N-------.----------T--V---.....-V---  163
F2.CM.95.MP257        ----------------K----N--------I-KR-A----------R---V--------EE.DSK---I--------------------------Q---R-----G--------H------------------I--QR-----N-------.----------T------.....-----  163
F2.CM.97.CM53657      --------------------N----------RR-K----------R---V--------K-.DSK---VS-------------------------Q---R-----G-----------------E------I--QR-----N-------.----------T------.....-----  163
G.BE.96.DRCBL         --------------------N---------I-K--K----------R--KV-------------..-K--VR-------DKE---H-------QG-------D--H----H--N---E------I--QT-R-S---P-----.----------KV-VA-T.....RRR--  162
G.CM.01.01CM_4049HAN  -------V------------N------KHI-K-T-H---------R--K-------------..-M-IVR---------K-----H-------QR-----I--D--H-----------E------I--E--------------.----------K--VP-T.....RKR--  162
G.CU.99.Cu74          --------L-----------N----------K--KD-W-------R--K--------------..---IVR---------K-----H-----KQRK----I--D--------H--K--------Q-I--EV-R---------K.----------K--V--T.....-----  162
G.ES.99.X138          ----------------K----N----------K--K-------F--R--KV-------------..-T--VR---------K-----H-------QR-----I--D--------H--------------------I--R------------T.----------K-----T.....-A---  162
G.GH.03.03GH175G      --------------------N---------I-K--K-----R----R--KV-------------..-T--VK---------QA----N-T-----QR--N--I--D--H----H------E------I--E--------------.----------K--VK-A.....RK---  162
G.KE.93.HH8793_12_1   --------------------N-------L---K----------F--R--KV-------------..-K--VS-----N---K--Q--H-------QR-----I--D--------H------E----N-I--R--R-----------.----------K--V--T.....-TR--  162
G.NG.01.01NGPL0669    ----------------K----N---------I-K-------G-----R---T-------I--..-KI-VI---------KA-Q--H-------QR-----I--D--------H-----T-------I--EL-R-----------.----------K--VK--.....-----  162
G.PT.x.PT2695         ------------------A-N----------K--K--S----F--M--KV--D--------..-K--VR---------K-----H-------QR-----I--D--------H--------------------G-----------T.----------K--A--T.....-A---  162
G.SE.93.SE6165        --------------------H----------K--------P--A-R---V------------..-T--V----------K--Q--H-------QR--R---E-D--H----H--------------------I--Q--------------.Q---------KV-V-S-.....RSR--  162
H.BE.93.VI991         -----------------K--N-----------K--KK-V-------TN-KT-------V--..-----------------------H-------QE-----I--D--------H--------------------I---R---I-D-----R.----------T---S--.....RT---  162
H.BE.93.VI997         -----------------NK------Y---I-K--NR-V-----D-----------------..------------------------------Q--------G------TH-----E----G-I--RV-------------.Q---------T--VA--.....-T---  162
H.CF.90.056           -----------------N-------Y--HI-R----------F--T------------E..------------N----E-----------L-------E-G------MH-----E------I--RV-R---N-P---K.Q--T------T--VA--.....-----  162
J.CD.97.J_97DC_KTB147 -----------------N-------Y--HN-K---K-L-----D-N--KV---------C..-E--V------------------------L---R-----G------MHH-----EA-----I----------------.----------T---K--.....-R---  162
J.SE.93.SE7887        -----------------N-------Y--N--K---K-L-----D-N--K----------E..-I---------Q-------------------QR--R-----G------MC-------------------I--Q-----D-------.----------T---R--.....RR---  162
K.CD.97.EQTB11C       -----------------N-------Y-----K--NR-R-----D-N--K-------------..-E--V----------E-----------L-K-R-----G------I------E----------R-----------T.Q---------T---A--.....-T---  162
K.CM.96.MP535         -----------------N-------Y--HI-K--NR-Y-------R--K-------------..-E--V------L-----------------L---R---E-D------I------E--V---I---R------C-----.----------T--VA-R.....RP---  162
```

```
                              Vif start
B.FR.83.HXB2                  MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHYESPHPRISSEVHIPLGD..ARLVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPELADQLIHLYYFDCFSDSAIRKALLGHIVSPRCEYQAGHN.KVGSLQYLALAALITPK.....KIKPP  162
01_AE.CF.90.90CF11697         --------------------Y--------I-K--KN---------Q--KV-------I--.-----R---------K-----H-------QRK-------D---RQ---Q-----------M--QV-R-----PT---.----------K--A---.....-TR--  162
01_AE.CN.05.FJ051             --------------------N--------I-K--KK-S-------Q--KV----Y----E..GK---R-----Q---K-----H-----KQRQ-N--I--D-------QQ----------R-I--QV-RS--D-PS---.----------K--T---.....--R--  162
01_AE.CN.06.FJ054             --------------------N----------KN-K----------Q--KV---------.-----R-----Q---K--Q--H-------QRK----I--D-------Q----------R-I--QV-RH----PS---.----------K--T---.....--R--  162
01_AE.HK.x.HK001              --------------------N----------K--KK-Y-------Q--K---------E.-----K-----Q---K--Q--H-------QRN----I---V--R----Q-----A-----R-I--QV-RYK---PS---.----------R--TG--.....GSR--  162
01_AE.TH.01.01TH_R2184        --------------------N----------K--KK---------QN-KV---------E..-K---R-----Q---K-----H-------QRE----I--D---R----H------------I--QVDRRK---PS--S.Q---------K--T---.....R-R--  162
01_AE.TH.02.OUR769I           --------------------N--------I-K--KK-S-------Q--KV---------E..-K---R-----Q---Q-----H-------QRN----I--D-------IQ----------S-I--QV-RS----PS---.----------K--T---.....R-R--  162
01_AE.TH.90.CM240             --------------------N--------I-K--KK---------Q--KV---------E.------R-----Q---K-----H-------QRK----I--D---K----Q--G----------I--QV-RR----PS---.----------K--T---.....R-R--  162
01_AE.US.00.00US_MSC1164      --------------------N----------K-VNK-A-------Q--KV---I-----E..-I---R-----Q---K-----H-------QRN-N--I--D--------Q------------I--QV-RR----PS---.----------K--T---.....R-R--  162
02_AG.CM.02.02CM_4082STN      --------------------N----Y---R-R--KD------F-CK--KA----------.-----R-----N---------H-F----KQ-K----I--D---H----H------E------I--QV---K---------.Q---------K--V--R.....-T---  162
02_AG.EC.x.ECU41              --------------------N----Y-----K--K-----------R--KV---------.----VR----------------H------KQR-----I--D---------H------E------I---V-R-K--------.----------K--VA-T.....RT---  162
02_AG.GH.03.GHNJ196           --------T-------K----N----------K--KD------F--R--KA---------.----VR-----N----------H-------QR-----I--D------------------E-----V---QV-R-S-------S.----------K--VA-T.....RR---  162
02_AG.NG.01.PL0710            --------------------N--------I-K--Q--X-------R---V----------.----VR---------A------H-------Q------I--D---------H-------------IV-EL-R----------.----------K---E--.....-T---  162
02_AG.NG.x.IBNG               --------------------N----------K--K-------F--R--KVC---------.----VR----------------H------KQ------I--D-----------N----------I--E--R----------.----------N--VA-T.....-T---  162
02_AG.SE.94.SE7812            --------------------N----Y---K-R--KD---------S---V----------.----VR----------------H------KQR-----I--D---------H------------I--Q--R---------T.----------K--V--T.....RT---  162
02_AG.SN.98.MP1211            --------------------N----------K--HR---------R--KV---------..-T-IVR-----QP---------H----V-QQ------I--D---------H------E---------QV-R-K--------.----------K--V--T.....RR---  162
03_AB.RU.97.KAL153_2          --------------------N------I-I-K-----V-K-----RN------------..-K---K-----------------A------E--------N---------------E----N-I---R---S---R----.-------------R---.........---  162
04_cpx.CY.94.CY032            -A-------------K----N----------K--K-*--------R--KV---------E.----VR-----QP--Q-----H-------LR--------D-------MH------E------I---R-------------.--------------S--.....-T---  161
05_DF.BE.x.VI1310             -----------------N-------Y--H--K--NR-C-----F--RN-------------..-K--V-------------------------R--------S-------V-------E----N-I--R-------F-----.----------T-------.....-T---  162
06_cpx.AU.96.BFP90            -----------------N-------Y--HI-K--KR-N------D-N--K----------S..-E---------N----K-------------LR--R-----SM------IH------E-----------R-----D------.----------T---K-E.....-R---  162
08_BC.CN.97.97CNGX_6F         -------L--------K----N----------RR-N---------D-R--KV---------E.------K-----Q----------H-A-----LRS-N--IE-G-------M------A--------I---V-I---D------N----------T---K--.....R----  163
09_cpx.GH.96.96GH2911         -----------------N-------Y--HT-K--KR---------D-N--K----------A..-E--V------------K-------------L---R-----S-------IH------E------II-DR----------.----------T---N-R.....-T---  162
10_CD.TZ.96.96TZ_BF061        -------L-------------N------------R--N--S-------R--K---------E..-S-I-K---------E----H-------L------I--G-------MH-----AE------I--QV-T---N-----K.Q---------T--------.....-T---  162
11_cpx.GR.x.GR17              -------------A-------N----------K--K--L-------R---V--------EE..DM--V-----------E-------------R---R--------------MH-----AE------I---R-------P--R-.----------K--V---.....RT---  162
12_BF.AR.99.ARMA159           -----------------N-------Y--H--K--KR------F--R---V--------EK..-E---------QP----------------QG--R--I--G---R---I------E------I---KI----N------.----------T---A--.....-T---  162
13_cpx.CM.96.1849             -----------------N-------Y--H--K--GR-------D-N--K-----------..-K--V-----------K-------------L---R-----G-------TH------------I---R-----D------.----------T---H--..........  162
14_BG.DE.01.9196_01           --------------------N----------R--K-------F--G--KV----------..-K--VR---------K-----H-A-----QRS-----H-D---------------ET---N-I-----R-S---R----.--------------T--T.....-A---  162
15_01B.TH.99.99TH_MU2079      --------------------N----------K--KK---------Q--KV--Q-----EE.------K-----Q---K-----H-------QRG----I-------R-----Q------E----R-I--QV-RYS-D-PS---.----------K--T---.....R-R--  162
16_A2D.KR.97.97KR004          --------------------N----------K--KD-C--------IN--V---------G..-KIIVK------P--K--------------QE-----I--D--GH----H------------I--K--R----------.----------K--VGET.....RT---  162
17_BF.PE.02.PE02_PCR0155      -----------------N-------Y-IH--K-TKN-S-----F--RN-------------..-K---A-----------------A-----QGG-R--I--G-------T------AE----N-I---RI----N-----K.----------T---A--.....-T---  162
18_cpx.CU.99.CU76             --------------------N-I--------K--KN---------R--KV----------..-T--V-----------KE----H-------LR--R-----D-----------------E----N-I---R----------.-------------------.....-T---  162
19_cpx.CU.99.CU7              --------------------N--------I-K--NK-V-------RN-KV----------..-K--VK---------K-----H-------QR-----I--D---------H--------------QV-R-----------.----------K--V--T.....RSR--  162
20_BG.CU.99.Cu103             --------------------H----------K--K--W-----F--R--KV---------..-K-IVR---------K----XH-----KQRE----I--D---HM------N------------I--E--R----FH----.----------K--V--T..........  162
21_A2D.KE.91.KNH1254          -------V-------------N--------I-R--KD-V-----Q-R---V--------EE..-K-IVR------S--K-----H-------Q-G----I--D---H-----------E------I--RL---K-------S.----------K--VAA-.....RT---  162
22_01A1.CM.01.01CM_0001BBY    --------------------N----------K--------------R--K----------.----VR----------KE----H-------Q--------D-------M------------------I--Q----S----T---.----------R--VA--..........  162
23_BG.CU.03.CB118             --------------------H-----XK-I-K--K--C-K-----TN------------E..-K-I-K------------------------------G-------Q-------E----S-I---R---S---------.----------K--V--T..........  162
24_BG.CU.03.CB378             --------------------N--I-----I-K--KD-C-K--F--RN-KV---------E..-K-IVR---------K-----H------KQRE----I--D-------QH--K-------------EK-R-S---H---S.----------K--V--T..........  162
24_BG.CU.03.CB471             -----------------K--N--------I-K--K--G-----F--RN--V---------N..-K-IVK---------K--Q--H------KQRE----I--D-------QH--K------------I--EQIR-S---------.----------K--VI-T..........  162
25_cpx.CM.01.101BA            -----------------NK------Y--H--K--KK---------N--------------..-E-AV-----------E-------------L---R-----G---R---MH-----T-------I---R----S-------K.----------K--VA--..........  162
27_cpx.CD.97.97CDKTB49        -----------------N-------Y---I-K---R-------D-N--------------..-V--V-----------E------------LR--R-----G---R---TH------E------I--YR---K-----------------------R.....-T---  162
28_BF.BR.99.BREPM12609        --------------------N--------I-R--KK-V-K-----T---V----------..-K--VI--------------------------E--------D---------------E----N-I--R--R---D--T---.----------T--TK--.....-T---  162
29_BF.BR.02.BREPM119          -----------------N-------Y-IHI-K---K-L------NR--K---------EE..-K---I-----Q-----------------QG--------G-------SH--N---E----N-I--R--RS--------T.----------T--L--R.....-T---  162
31_BC.BR.02.110PA             -------L--------K--X-N---------X-RR-S--X-------R--KX---------E.------K-----Q----------H-------LRS-------D-------MH-----A-------I---R-----D------.------X---T---K--.....-R---  162
32_06A1.EE.01.EE0369          --------N-L-------N-------Y--HI-K--KN-------D-H--K----------S..-E---------N----E-------------L---K--I--S-------MH------E------I-----RH---------.----------K---K--.....-R---  162
33_01B.MY.05.05MYKL007_1      --------V------------N------K-I-K--K--L-------Q--KV-------I-E..-KIA-R-----Q---K-----H-------Q-K----I--D---------N-----------R-I--QV-RY----PS---.----------KV-T-Q-.....R-R--  162
34_01B.TH.99.OUR2478P         --------------------N-I--------K--KK---------Q--KV---------E.------R-----Q---K--Q--H-------QRN----I--D---R----Q------E----R-I--QV-RR----PS---.----------K--T---.....R-R--  162
35_AD.AF.05.05AF095           --------------------N--------I-K-TK--V-K-----R--KV----------..-KIIVR---------K-----H-------L------I--DV-------H--N---------VI--QV---S-Q------.----------K--V--T..........  162
36_cpx.CM.00.00CMNYU830       --------------------N-------HI-KT-KD---------R--KV----------.------R---------------H------KQRK----I--D---------H------------I--Q------D-P----.----------K--V--T.....-R---  162
37_cpx.CM.00.00CMNYU926       --------------------N-------HI-K--K----------RD--V---------E..-W--VR-----QP---------H-------LRK-R--I--D---------H----------T-I--R--R-----P---S.----------K--VP-T.....RKR--  162
38_BF1.UY.05.UY05_4752        -----------------N-------Y-IH--K--K--C-K--F--RN--V--------KE..-T-A--------------------------LGG-R--I--G-------I-------E----N-I---R-----N-P----.----------T---A--.....-T---  162
39_BF.BR.03.03BRRJ103         -----------------N-------Y-KH--K--KE-V-------NN------------E..-K--V-----------E-------------R--------D-----------------E----H-I--R--R---N------.----------T--------.....R----  162
40_BF.BR.04.04BRRJ115         -----------------K----------HI-RR-GR---------T--K----------E..-E--V--------P-----------------R--------D---------H-----E----N----R----S-N------.----------T---K--..........  162
42_BF.LU.03.luBF_05_03        --------------------N----------K--K--Y-------T--K-----------..X----K------------------------------N---------------E----NTI---R------------.Q---------T--------.....-X---  162
43_02G.SA.03.J11223           --------------------N-X--------K--K----------RN-KV-------XX-..-T-XXX----XQ----X-Q--H------XQ-K-------D---------H------E------I--Q--R----------.----------K--VK-T.....RK---  162
U.CD.83.83CD003               -----------------K--N--------I-K--K--V-K-----TN------------N.---I-V-A-----------------------QG-----I--D---------H------------I------------T---.----------TT-VA-T.....-R---  162
U.CD.90.90CD121E12            --------------------N----------K--NK-V-K-----IN-----------EN.---I---A-----------------------QG-----I--D-------MH------------I---------D--T---.----------T--VP-I.....-R---  162
U.GR.99.GR303                 -----------------N-------Y--H--K--KR-V-----D-N------------RE..-K--V-----------E-------------M-S-R-----S-------MH------E------I--Q------D------.----------T---N--.....RT---  162
U.NL.95.U_NL_95_H10986_D1     -----------------S-------Y--H--K--K----------R---V--------K-..-K------R------G--Y---------GQRK-------G-------MH--------------------R------------.----------TV-V---.....R----  162
CPZ.CD.90.ANT                 I-K-*KIEMI--------NI--------IWETKVLKP-K------ND--KKGE------PTLDKK--V-VF---QC---P----H-------CGK-I-------T---M--Q---P----Q-V-Q-I--ERILTY-H-KK--S.Q--T------FCKILEFRGY.PKGPRRQ  167
CPZ.CM.01.SIVcpzCAM13         ---K---------------KA-N----Y---KTK--KD--------H-N--V-------FK-.NSK-I-----A-T----A----H----Q--MGS-V------YV--K---SQ-----A-----R-I-----E---A-RE--R.Q--T--F---K--VKE-.....RLR--  163
CPZ.CM.05.SIVcpzEK505         --------V--------V---N----------K-------------TDN-K----I-----E..-K---V-----MP---P----H-------QGI-R--I-------K----------TA----Q-V--RP-I-K---P---K.Q----------I-WVGVQ.....-RR--  162
CPZ.CM.05.SIVcpzLB7           -----K-L----I------KA-N----Y-I---R--------------EN-KV---I-----E..-K---V-----TP----------H-----KQ-K-----I-------HM------N--AE------I--RV------FP---K.---T------IS--LQRQ.....PTR--  162
CPZ.CM.05.SIVcpzMB66          ------------------K--N----------R---N---------NN-K-----------E..-K---------T----A----H-------QRS-R--I--D-----------------AE------I--Q--------PT---.----------ST--KS-.....PRR--  162
CPZ.CM.05.SIVcpzMT145         -------Q---------VK--N----Y---R-K--KN-------DGTN--MA-QIY--INQ..-K-IV------MP---------H--A----ERG---------T--R----H-----TAT---N-I-EKA-L-K--FPE--S.--------W--KQ-VKT-.....-NR--  162
CPZ.GA.88.GAB1                ------------------K--N----Y-I-R-K-----------DH-N-KVA--I---FR-.YSK-IV----A-SP---A----H----Q--LGS-V-----FT--R---SQ-----AET---R-I--QL-A------KE--R.Q-----F---K---SER.....RHR--  163
CPZ.TZ.00.TAN1               -------QV--MI----L---T------IFTTKCCKD-K------TDT-KRAG-I----TE.RSK--VLH----AC---P----H-IGL---QGK-----I---T-------TR--T--AAG-V-Q-I--ERILTF-HF-S--R.Q--T--F--FRKVVESQDKQPKGPRR-  168
CPZ.US.85.CPZUS               -------L---------VKA-N----Y---R-K--K----------TN--VN------IDK..-KI--I-----K--QA---H-------QGT-K--I---M--K----------TA------V--RP-I-K---PT---.------F---K--VGQS.....-RR--  162
CPZ.CM.x.SIVcpzDP943          --------V--------VK--N----Y---RTK---K---------SN--V---I-----EK..-K--V------MP--HE----H-------YGS-R--I---M--R-----H-----AA------V--TQ-V-K---P----Q.Q----------K-WVGVE.....-KR--  162
CPZ.CM.x.SIVcpzMB897          -----K-------I----VKA-N---------K-------------EN-KV---I---V--..-K---V-----MP----------H-------Q-------I-------H-----H-----AE------I--Q-------FP---T.---T------IS--LKR-.....PK---  162
N.CM.02.DJO0131               --------V------R--K--H----------R--E--Y------TN--K--------V-L..-----------T----S---H-----KLGK-K------M--K-------N--TA----Q-V--RP-F-----P----.Q--T--H---T-WVGV-.....-R---  162
N.CM.04.04CM_1015_04          --------V-----------I-N-------H--K--K--Y-------R--KT--------V-Q..-T---V-----T---QP---H-------LRK-K------M--K-----------T-----Q-V--RQ-F-----P----.Q-DT------T-WVGI-.....-R---  162
N.CM.04.04CM_1131_03          --------V--------K--I-N-------H--K--K--Y------TR--K---------V-Q..-----V-----T---QP---H-------LRK-K------M--K----------TA----Q-I--RQ-F-----P----.Q--T------T-WVGI-.....-R---  162
N.CM.95.YBF30                 --------V--------K--K-N----------K--K--Y------TH--K---------V-Q..----TV-----T---QS---H-------LRK-K------M--K----H-----TA----Q-V--RP-L-----P----K.Q--T------T-WVGA-.....-R---  162
N.CM.97.YBF106                --------V------X-K----N----------K--K--Y------TK--KT---X---V-X..-X---V-----T---QP---HX-------QGK-K------M--K---C---N--TA----Q-V--RP-L----X-P---X.Q--T-----XT-XVGV-.....-RR--  162
O.BE.87.ANT70                 -------L-------QKVKA-N----Y-K-R-R-TEN-W-------RN--V--S-Y--V-V..-HV-V------MP----E----H--------Y-K-K--I----T--RM---H--T--TA--V----I--QR-LTK---PT--S.Q--T--L---R-VVKAR.....SR---  162
O.CM.91.MVP5180               -------L----I--QKVKA-N----Y-K-M-K--AN-R-------RN-KV--A-Y--VAE..-DI-V------MP---EE---H------QY-E-K--I----T--RM---H--T--TE-------I--QR-LTK---L---S.Q--T--F---K-VVKV-.....RN---  162
O.CM.96.96CMABB637            -------L-------QXVKA-N--X-Y-K-X-K--QK-Y------TRN-----G-Y--V-T..-NI-V------MP----E----H------QY-K----I----T--RI---H--T--TG----R-I--QR-LT----P---S.Q--T--L---R-VVKD-.....RS---  162
O.CM.98.98CMA104              -------L-------QKVKV-N----Y-K-R-R--KD-Y-----X-RN--V--G-Y--V-Q..-CI-V------MP---EDQ--H------QY-K-R--I----T--RM------T--TE-------T--QR-LTK-D-PE--S.---T--L---RVVVKT-.....RN---  162
O.CM.99.99CMU4122             -------L------XQKVK--N----Y-KHR-R-TKE-H-K--F-FRN--A--G-Y--V-M..-CV-V------MP----E----H------QY-N-T--I----T--RI--QH--T--TE------II-QRIXTK---S---S.Q--T--L---RVVVKE-.....RN---  162
O.FR.92.VAU                   -------L-------QKVKA-N----Y-K-R-R--TD-Y-K-----IN--V--R-Y--V-P..-CIMV------MP---EEQ--H------QCRK-I--I----T--K-----H--T--TE----R-T--QR-L-Q---PE--S.Q--T--L---R-VVKA-.....RNR--  162
O.SN.99.SEMP1299              -------L--------Q-VK--N----Y-K-R---T-D-Y----F--RN--V--G-Y--V-G..PWI-V------MP----E----H--------Y-K-K--I----T--RM--I---A--TE-------I--QR-LT----P---S.Q--T--L---RVVVKE-.....RN---  162
O.US.99.99USTWLA              -------L-------QKVK--NX---Y-K-R-K--KD-Y------TRN-----G-Y--V-G..-MI-V------MP----E----H------EY-N-K--I----T--RM---H--T--T-------C--QK-LT---FHK--S.Q--T--L---RVVVGE-.....TSR--  162
```

```
                                Vif end
B.FR.83.HXB2          LPSVTKLTEDRWNKPQKTKGHRGSHTMNGH*  192
A1.GE.99.99GEMZ011    ----R------------R-R--N-----C-  193
A1.KE.00.KER2008      ----K--A---------IR---E--------  193
A1.KE.00.KNH1144      ----RI-A------H---R-P-E------C-  193
A1.KE.00.KSM4024      ----RI-A---------R-LK--------C-  193
A1.KE.00.MSA4069      ----RI--------R----P-E------C-  193
A1.KE.00.NKU3005      ----R--A----------R-P--------C-  193
A1.RU.00.RU00051      ----R------------R-R--N-----C-  193
A1.RW.93.93RW_024     ----R--A------R---D---------C-  193
A1.SE.95.SE889[       ----R--A-----------------------  193
A1.SE.95.UGSE8131     ----R-----------------------C-  193
A1.TZ.01.A173         ----R--A---------IR-P---------  193
A1.UA.01.01UADN139    ----R---------S---R-R--N-----C-  193
A1.UG.92.92UG037      ----K--A----------R---E------C-  193
A1.UG.99.99UGA07072   ----R--A----------R------------  193
A2.CD.97.97CDKTB48    ----R--V----------R----------C-  193
A2.CY.94.94CY017_41   ----R--V----------------------C-  193
A.SN.01.DDI579        ----R------------R---------C-  193
A.SN.01.DDI369        ----R------------R--------S-C-  193
A.SN.96.DDI360        ----R--A----------R----------C-  193
A.CD.02.02CD_KTB035   ----R------------R----------C-  193
A.CD.97.97CD_KCC2     ----R------------R------------  193
A.CD.97.97CD_KTB13    ----R--A--------------N-----C-  193
B.AR.04.04AR151516    ----K-----------------N---S---  193
B.AU.87.MBC925        ----------------I-------I-----  193
B.BO.99.BOL0122       ---LQI-----------------E-------  193
B.BR.03.BREPM2012     ----R----------H-------N-S-----  193
B.CA.97.CANB3FULL     ----A----------------E-------  193
B.CN.05.05CNHB_hp3    ----A-----------RI-----N---S---  193
B.CO.01.PCM001        ----K----------R-I-------------  193
B.GB.83.CAM1          ----R--------------------------  193
B.GB.86.GB8           ------------------------------  193
B.GE.03.03GEMZ010     ---IR-----------R--------P-----  193
B.IT.05.SG1           ----Q---------------------ID---  193
B.JP.05.DR6538        ----A--------------------------  193
B.KR.05.05CSR3        ----A----------------K---------  193
B.NL.00.671_00T36     ----R----------R------E---H---  193
B.RU.04.04RU128005    ----A--A-X--------R---EN-------  193
B.TH.00.00TH_C3198    ----A-----------R----------S---  193
B.UA.01.01UAKV167     ----R----------------------H---  193
B.US.04.ES10_53       --N-A----------R-------N-------  193
B.US.99.PRB959_03     -------------E----------------  193
C.AR.01.ARG4006       ----R--V----------RDR--N-I-----  193
C.BR.04.04BR013       ----K--V----------RDR----------  193
C.BW.00.00BW07621     ----R--V----------R-R--N-------  193
C.CN.98.YNRL9840      ----K--V-----N---IR-R--N-------  194
C.ET.02.02ET_288      ----S--V----------R-R--N--V----  193
C.GE.03.03GEMZ033     ----R--V----------R-R--N-------  193
C.IL.99.99ET7         ----R--V----------R-R--N-------  193
C.IN.99.01IN565_10    ---IR--V----------R-R--N-------  193
C.KE.00.KER2010       ----Q--V-----N----R-R--N-------  193
C.MM.99.mIDU101_3     ---IK--V-----N---IR-R--N-------  194
C.MW.93.93MW_965      ----K--V----------R-R--N-------  193
C.SN.90.90SE_364      ----R--V----------R-R--N-------  193
C.SO.89.89SM_145      ---IR--V-----N--RIR-R--N-------  193
C.TZ.02.CO178         ---IR--V-----N----R-R--N---S---  193
C.UY.01.TRA3011       ----K--V--------R-RDR--N-------  193
C.YE.02.02YE511       ----R--V----------M-R--N-------  193
C.ZA.04.04ZASK164B1   ---TR--V----------RDR--N-------  193
C.ZA.05.05ZAPSK240B1  ----R--V----------R-R--N-------  193
C.ZM.02.02ZM108       ----R--V----------R-R--N-------  193
D.CD.83.ELI           ----R-----------Q-R------------  193
D.CM.01.01CM_0009BBY  ----R------------R----N-------  193
D.KE.01.NKU3006       ----R-----------R----------H---  193
D.KR.04.04KBH8        ----R-----------E---R---P-----W  193
D.TD.99.MN011         ----R-----------Q--------------  193
D.TZ.01.A280          ----K-----------R-------P------  193
D.UG.99.99UGK09259    ----R-----------R--------------  193
D.YE.01.01YE386       ----R-----------R--------------  193
D.YE.02.02YE516       ----K-----------Q--------------  193
D.ZA.90.R1            ----R------------R------------  194
F1.AR.02.ARE933       ----Q--V----------R------------  193
F1.BE.93.VI850        ----Q--V--------E-R------------  193
F1.BR.01.01BR125      ----Q--V----------RD-----------  193
F1.ES.x.P1146         ----K--V-----E-R-IR-R--N-------  193
F1.FI.93.FIN9363      ----K--V--------EIR------------  193
F2.CM.02.02CM_0016BBY ----K--V--K--N-----------------  194
F2.CM.95.MP257        ----R--V-----N----R------------  194
F2.CM.97.CM53657      ----R--V-----N----R------------  194
G.BE.96.DRCBL         ----R--------E----R---ENP------  193
G.CM.01.01CM_4049HAN  ----S---------S---R---ENP------  193
G.CU.99.Cu74          ----K-----------R-R---ENP------  193
G.ES.99.X138          ----R-----------R-R----NP------  193
G.GH.03.03GH175G      ----R----N------------ENP-----W  193
G.KE.93.HH8793_12_1   ----R---------S---R---E-P------  193
G.NG.01.01NGPL0669    ----R----------R------ENP--S--W  193
G.PT.x.PT2695         ----R-----------R-R----NP------  193
G.SE.93.SE6165        -----E-A----------R---ENP------  193
H.BE.93.VI991         ----R--V----------R------------  193
H.BE.93.VI997         ----K--V--G-------R---------R--  193
H.CF.90.056           ----R--V----------R------------  193
J.CD.97.J_97DC_KTB147 ----Q--V--------------E--------  193
J.SE.93.SE7887        ----Q--V----------T---E--------  193
K.CD.97.EQTB11C       V---Q--V----------R--------S-Q-  193
K.CM.96.MP535         V---K--V----------R-----Q------  193
```

Vif end

```
B.FR.83.HXB2                LPSVTKLTEDRWNKPQKTKGHRGSHTMNGH*   192
01_AE.CF.90.90CF11697       ----R------------R---ENP------   193
01_AE.CN.05.FJ051           ----K------------IR---E-P--H---   193
01_AE.CN.06.FJ054           ----K------------IR---ECPP-----   193
01_AE.HK.x.HK001            ----K-------SE----R---ENP------   193
01_AE.TH.01.01TH_R2184      ----K------------IG---ENP------   193
01_AE.TH.02.OUR769I         ----K------------IR---ENP-T----   193
01_AE.TH.90.CM240           ----EI.----------.R---ENP------   191
01_AE.US.00.00US_MSC1164    ----R------------IR---ENP------   193
02_AG.CM.02.02CM_4082STN    ----R--A-----E----R-----RP-----   193
02_AG.EC.x.ECU41            ----K--A--------RIR----NRS-----   193
02_AG.GH.03.GHNJ196         ----K------------R----NRPI----   193
02_AG.NG.01.PL0710          ----R-------KE----R-----RS-----   193
02_AG.NG.x.IBNG             ----R--A----KE----R-----RP-----   193
02_AG.SE.94.SE7812          ----K--A-----E----R-----RS-----   193
02_AG.SN.98.MP1211          ----K--A-----E----R----NRS-----   193
03_AB.RU.97.KAL153_2        ---------------R--D-------S---   193
04_cpx.CY.94.CY032          ----K--V----------R-R-ENQI-----   192
05_DF.BE.x.VI1310           ----R--------------R--N-----Y-   193
06_cpx.AU.96.BFP90          ----Q--V----------R---E--------   193
08_BC.CN.97.97CNGX_6F       ---IK--V-----N---IR-R--N-------   194
09_cpx.GH.96.96GH29I1       ----K--V--------R-R---ENQ------   193
10_CD.TZ.96.96TZ_BF061      ----R----------R--------------   193
11_cpx.GR.x.GR17            ----R------------R-R--N--V--C-   193
12_BF.AR.99.ARMA159         ----K--V-----N----R------------   193
13_cpx.CM.96.1849           ----K--V----------RD--ENQ------   193
14_BG.DE.01.9196_01         ----R----------R-R----NP------   193
15_01B.TH.99.99TH_MU2079    --R-K-------D----IR---ENP--D---   193
16_A2D.KR.97.97KR004        ----R-------------------------   193
17_BF.PE.02.PE02_PCR0155    ----K--V--------E-R---E--------   193
18_cpx.CU.99.CU76           ----R--A----------R----NQ------   193
19_cpx.CU.99.CU7            ----R------------R---ENP------   193
20_BG.CU.99.Cu103           ----R----------R-R---ENP------   193
21_A2D.KE.91.KNH1254        ----R-------------------L-----   193
22_01A1.CM.01.01CM_0001BBY---R------------R---EYP-----W   193
23_BG.CU.03.CB118           ----R----------R-R---DN-------   193
24_BG.CU.03.CB378           ----R----------R-R---EN-------   193
24_BG.CU.03.CB471           ----R----------R-RD--EN-------   193
25_cpx.CM.01.101BA          ----R------------R---ENQ------   193
27_cpx.CD.97.97CDKTB49      ----K--V----------R-L-ENQ------   193
28_BF.BR.99.BREPM12609      ----R------------I-----N-------   193
29_BF.BR.02.BREPM119        ----K-------------D-----------   193
31_BC.BR.02.110PA           ----R--V-----A----RDR--N-------   193
32_06A1.EE.01.EE0369        ----Q--V-----E----R---E------R-   193
33_01B.MY.05.05MYKL007_1    ----K------------IR---ENPATS---   193
34_01B.TH.99.OUR2478P       ----K----------QIR---ENP-T----   193
35_AD.AF.05.05AF095         ---IR--A-----E----R-P-E--------   193
36_cpx.CM.00.00CMNYU830     ----K------------R-----RST----   193
37_cpx.CM.00.00CMNYU926     ----K------------R---ENP------   193
38_BF1.UY.05.UY05_4752      ----K--V-------PE-R------------   193
39_BF.BR.03.03BRRJ103       ---IK----------R------E---I----   193
40_BF.BR.04.04BRRJ115       ---IK----------R--------------   193
42_BF.LU.03.luBF_05_03      ----R-------------------------   193
43_02G.SA.03.J11223         ----A----X----S---R-----RX-----   193
U.CD.83.83CD003             ----R--V------------K-----H---   193
U.CD.90.90CD121E12          ----R--V---------I---K---------   193
U.GR.99.GR303               ----K--V-------------E-Q-----   193
U.NL.95.U_NL_95_H10986_D1   ----R--V----------R---E---I----   193
CPZ.CD.90.ANT               F--LSI---------RRMR---ENQ*-S-C-   197
CPZ.CM.01.SIVcpzCAM13       ----A--------T--R----PEN---S---   194
CPZ.CM.05.SIVcpzEK505       -------------RH--I-D------T--L-   193
CPZ.CM.05.SIVcpzLB7         ----S--A------S----------I-S---   193
CPZ.CM.05.SIVcpzMB66        ----R--A-----NS---RD-K-N-I-----   193
CPZ.CM.05.SIVcpzMT145       ----A--A--------R-RD-QE-PMSS-C-   193
CPZ.GA.88.GAB1              ----A---------H-R--V-QENL-R----   194
CPZ.TZ.00.TAN1              --------------HRT-T-R-EN--LS-C-   199
CPZ.US.85.CPZUS             --------------C-----QPEN---S---   193
CPZ.CM.x.SIVcpzDP943        ----A--A------H--IRDQQE--------   193
CPZ.CM.x.SIVcpzMB897        ----S--A-------R-----KE--------   193
N.CM.02.DJO0131             --------------H---Q----NPI-----   193
N.CM.04.04CM_1015_04        ----A--------RH---Q----NPI-----   193
N.CM.04.04CM_1131_03        ----A---------H---Q----NP------   193
N.CM.95.YBF30               -------------EH--MQ----NPI-----   193
N.CM.97.YBF106              -------------ER---Q----NPI---P-   193
O.BE.87.ANT70               ----Q---------HLRIRDQLK-PS-----   193
O.CM.91.MVP5180             ----QR--------W-IRDQL---S-----   193
O.CM.96.96CMABB637          ----Q---G----RHLRIRDQQE--S--X--   193
O.CM.98.98CMA104            ----Q--------RHL-IRDQLE--S-----   193
O.CM.99.99CMU4122           ----Q---------HLRIMDQL-N-S-----   193
O.FR.92.VAU                 ----Q--------RHLRI-DQP--QL-H---   193
O.SN.99.SEMP1299            ----Q-------SRHLRIRDQLE--S-----   193
O.US.99.99USTWLA            V---Q---G----RC--I-DQLE--S---QY   193
```

Vpr start; amphipathic α-helix; oligomerization; frameshift in HXB2; Vpr end in HXB2; H(S/N)RIG motifs; Vpr end

```
B.FR.83.HXB2            MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIW.LHGLGQHIYETYGDTWAGVEAIIRILQQLLFIH#FRIGCRHSRIGVTRQ....RRARNGASRS*  96
A1.GE.99.99GEMZ011      ------------------M-----D--H--------P-.---------N------E-------T------V-.-----Q-----IEQR....--G---S----  97
A1.KE.00.KER2008        --------------Y--------------------GP-.------Y--H------E----ML--------V-.-----T-----IIPR....----D-SG---  97
A1.KE.00.KNH1144        --------------Y---M-----D--H--------P-.---------D------E-------T------V-.-----Q-----II-G....--G---S----  97
A1.KE.00.KSM4024        --------------Y---M--------H--------P-.---------N------E--------------V-.-----Q-----II-G....--GI--SG---  97
A1.KE.00.MSA4069        ------------------M-----D--H--------P-.---------D------E-------L------T-.-----Q-----IIPGRRG.--V--------  100
A1.KE.00.NKU3005        --------------Y---M--------H--------P-.----------------E--------------V-.-----Q-----II-G....--V-D-SG---  97
A1.RU.00.RU00051        --------------Y---M--------H--------P-.---------N------E--------------V-.-----Q-----IN-R....--V---S----  97
A1.RW.93.93RW_024       --------------Y---M-----D--H--------P-.---------------QE-----V--------V-.-----Q-----IIPG....--G---S----  97
A1.SE.95.SE8891         --------------Y---A---V----H--------P-.---------N------E-------T------V-.-----Q-----II-R....--V-D-S----  97
A1.SE.95.UGSE8131       --------------Y---M-----D--H--------P-.------Y--Y------E-----V-T------V-.----......#II-G....----D-S----  90
A1.TZ.01.A173           --------------Y------------H--------P-.-------V--------E--------------V-.-----Q-----II-GG...--V---S----  98
A1.UA.01.01UADN139      ---V--------------M--------H--------P-.---------N------E--------------V-.-----Q-----IVQR....--G---S----  97
A1.UG.92.92UG037        --------------Y---M-D---D--H--------P-.---------H------E--------------V-.-----Q-----INIRG...--V-D-SG---  98
A1.UG.99.99UGA07072     --------------Y------------H--------P-.---------N------E-------T------V-.-----S-----I--R....----D-S----  97
A2.CD.97.97CDKTB48      ----------P-------V--------Q--------P-.---------N-----RE-----V--------V-.-----Q-----IIPR....--V-D----P-  97
A2.CY.94.94CY017_41     --------------Y---M--------Q--------H-.------Y--N------E---V---Y------V-.-----Q-----II-R....--V-D----P-  97
A.SN.01.DDI579          --------------Y---M--------Q--------P-.---------N------E-------M------V-.-----Q-----II-RG...--V-D----P-  98
A.SN.01.DDJ369          --------------Y---M-----D--H--------P-.------YV-N------E------Q-------V-.-----Q-----I--RG...--G-D------  98
A.SN.96.DDJ360          --------------Y---M------I-Q--------P-.------Y--N------E-------M------V-.-----Q-----II-RG...--V-D----P-  98
A.CD.02.02CD_KTB035     --------------Y---M--------H--------P-.------Y--N------E-----T---H------.Y----Q-----II-S....--R-D-TG---  97
A.CD.97.97CD_KCC2       --------------Y---------------------V-.---------N------E----------------.Y----Q-----I-PS....-GV--------  97
A.CD.97.97CD_KTB13      ------N-------Y---V--------H--------M-.---------N------E-----V-M--------.Y----Q-----IIPS....--I--------  97
B.AR.04.04AR151516      ------N-------Y------------S---X----X-.-QS---Y--------------------------.-----------I-L-....-----------  97
B.AU.87.MBC925          --------------Y---------------------P-.---------------------------------.-----Q-----I---....---------P-  97
B.BO.99.BOL0122         ---------------------------S--------T-.-Q----------------------------L--.-----------I-Y-....K----------  97
B.BR.03.BREPM2012       -----Q--------F------------S----------.--N------------------LL-S--------.-----Q-----II--....-----------  97
B.CA.97.CANB3FULL       --------------Y------------S--------V-.-----------------------------M---.-----Q-----I---R...-AR--------  98
B.CN.05.05CNHB_hp3      --------------Y----------I-R--------R-.--E---Y---------V--------------T-.-----Q-----I---....-----------  97
B.CO.01.PCM001          --------------Y------------RV------SS-.-------------N------VL--T--------.------L----I.Q-....-G---------  96
B.GB.83.CAM1            ---T--------Q-Y------------S--------P-.-LA-------------P----L--S--------.-----------IIQ-....--S--------  97
B.GB.86.GB8             --------------Y------------S-------S--.------Y--------------------------.-----------I---....-----------  97
B.GE.03.03GEMZ010       ---T----------S---A--------S--------P-.--N---Y---------T-------T--------.-----------I---....------S----  97
B.IT.05.SG1             --------------Y-R-------------------T-.------Y-----R--*------T--M--------------H-----M--H....-----------  97
B.JP.05.DR6538          ---------T----Y------------R--------P-.--S------------------------------.-----Q-----IIP-R...-AR--------  98
B.KR.05.05CSR3          --------------Y---S--------R--------P-.--S---Y---------T----L--M--------.-----Q-----II--....-----------  97
B.NL.00.671_00T36       --------------Y-A----------S--------G-.---------------------L-----------.-----Q-----I--R....G-V------P-  97
B.RU.04.04RU128005      ------N-------Y------I--------------V-.------Y---------X--------------V-.-----------I---....-----------  97
B.TH.00.00TH_C3198      ---------------------------S----------.---------D-----------LL-T--------.-----------I-Q-....-----------  97
B.UA.01.01UAKV167       ----------P---Y-A-A--------S--------P-.--S---Y--------------------------.-----Q-----II--....-----------  97
B.US.04.ES10_53         --------------Y--------------------TV-.---------------------------------.-----Q-----I---....-----------  97
B.US.99.PRB959_03       --R---N-------Y---------------------P-.--S---Y------------------------V-.-----Q-----II--....-----------  97
C.AR.01.ARG4006         --------------Y------------Q--------P-.--S---Y---------M--------------V-.-----------II--....-----------  97
C.BR.04.04BR013         --------------Y------------Q--------P-.--S---Y---------T--------------V-.-----N-----IM--....-----------  97
C.BW.00.00BW07621       --------------Y------------Q--------P-.--S------N------T----------------.-----Q-----IM--....--T--------  97
C.CN.98.YNRL9840        ---P-----------------------Q--------P-.------YV--------T---T---L--------.-----Q-----IW--....-----------  97
C.ET.02.02ET_288        --------------YS--------D--Q--------P-.--S---Y--N------R----L-TF---S----.-----H-----I---....-----------  97
C.GE.03.03GEMZ033       --------------Y------I-----Q--------P-.--SI--Y---------T---V---V--------.-----Q-----IL--....-----------  97
C.IL.99.99ET7           ---------------------------Q--------P-.--S---Y---------T----L--T------T-.-----Q-----IMP-....-------R---  97
C.IN.99.01IN565_10      --------------Y------------Q--------P-.------YV--------T-------L--------.-----Q------L--....-----------  97
C.KE.00.KER2010         ---P----------Y------I-----Q--------P-.--SI--Y---------T----------------.-----H-----IL--....-----------  97
C.MM.99.mIDU101_3       ---P-----------------------Q--------P-.------YV--------T---I------------.L----Q-----IV--....-----------  97
C.MW.93.93MW_965        --------------Y---A--------Q--------P-.--N-------------T---V------------.-----H-----IL--....--T---PN---  97
C.SN.90.90SE_364        --------------Y------------Q--------P-.--S-------------T----------------.-----Q-----IIQ-....-----------  97
C.SO.89.89SM_145        ---S----------Y------I-----Q--------P-.--N-------------T--------------T-.-----Q-----IL--....-----------  97
C.TZ.02.CO178           ---P----------Y---A--I--D--Q--------P-.--S------A-H----N----L--M------V-.-----Q-----ILQ-....-----------  97
C.UY.01.TRA3011         --------------Y------------Q--------P-.--S-------------M--------------V-.-----Q-----IL--....-----------  97
C.YE.02.02YE511         -------H------Y------------Q--------P-.--S---Y---------T-----L--------V-.-----Q-----TI-H....-------N---  97
C.ZA.04.04ZASK164B1     --------------Y------I-----Q--------P-.--S---Y--------------LL-----M----.-----H-----ILQH....-----------  97
C.ZA.05.05ZAPSK240B1    ---------------------------Q--------P-.--N---Y---------T--QV------------.Y----Q-----IW--....-------N---  97
C.ZM.02.02ZM108         --------------Y------------Q--------P-.------Y---------T-----------M----.-----S-----II--....-----------  97
D.CD.83.ELI             -----A--------Y---A--------S----------.--S-------------V----------------.-----Q-----II--....------S----  97
D.CM.01.01CM_0009BBY    --------------Y------I-----S--------V-.----------------------T-V--------.-----Q-----I---....------S----  97
D.KE.01.NKU3006         --------------YDA----------S--------M-.--S------D------T----L-----------.-----L-----II--....--TG-------  97
D.KR.04.04KBH8          -----G--------Y---A--------S--------P-.--D------A------V----LL-V------V-.-----Q-----II--....------S----  97
D.TD.99.MN011           -----A--------Y------------S----------.---------------------L-------M---.-----------IQ-T....A-----P----  97
D.TZ.01.A280            --------------Y------------S--------L-.--N---F--------------------------.-----Q-----I-L-....--T---S--P-  97
D.UG.99.99UGK09259      --------------Y------------S--------M-.----------------T----LL-F------V-.-----Q-----I---....--T--------  97
D.YE.01.01YE386         --------------Y------------S--------P-.--S---Y--D------T----L--T--------.-----Q-----IL--....--I--------  97
D.YE.02.02YE516         -----A--------Y------------S--------A-.------------------------T-----L--.-----Q-----ISP-....------S--P-  97
D.ZA.90.R1              --------------Y------------S-------IT-.--S------------------------------.-----Q-----II--....------S----  97
F1.AR.02.ARE933         --------------Y---A-D---------------P-.--S------N-----------LL----------.-----H-----I---....-----------  97
F1.BE.93.VI850          -----G--------Y---A--I--------------P-.---------N------E----------------.-----------IVP-....--V--------  97
F1.BR.01.01BR125        --------------Y---------------------P-.--S------N------E-------M--------.-----H-----IM--....--I--------  97
F1.ES.x.P1146           --R-----------Y---------------------P-.------Y--S------E----------H-----.-----Q-----I---....--V--------  97
F1.FI.93.FIN9363        -----G--------Y---A-----------------PF.---I--Y--N------E------K---------.-----H-----II--....--V--------  97
F2.CM.02.02CM_0016BBY   ---P----------Y------------H--------E-.---------N------E-------T--------.-----H-----II--....--I---NG---  97
F2.CM.95.MP257          ---P----------Y------------H--------E-.---------N------E----------------.-----H-----II--....--I---SG---  97
F2.CM.97.CM53657        ---P-----------------------H--------E-.---------N------E--Q----T--------.-----H-----II--....--L---NG---  97
G.BE.96.DRCBL           --R---------------A--------H--------P-.------L--N------E----------------.-----Q-----IIP-R...--V-D-SG---  98
G.CM.01.01CM_4049HAN    ---V----------Y---------------------P-.------Y--N------E--------------V-.-----Q-----I-P-....--V---PG---  97
G.CU.99.Cu74            ------------------------------------P-.------YV-N------E----------------.-----------ISLR....--V-D-PG---  97
G.ES.99.X138            --------------Y---A-----------------P-.------Y--N------E----------------.-----Q-----IVP-....--V-D-PG---  97
G.GH.03.03GH175G        ------N-------Y---A-----------------P-.------Y--N------E-------M--------.Y----Q-----IIP-....--V-D-P--P-  97
G.KE.93.HH8793_12_1     ---V----------Y---A-----G-----------P-.---------N------E----L-----------.-----Q-----I-LR....--V---P----  97
G.NG.01.01NGPL0669      ------N-------Y------------V--------P-.---------N------E----------------.-----Q-----I-PR....--V-D-SG---  97
G.PT.x.PT2695           --------------Y---A-----------------P-.------Y--N------E----------------.-----Q-----I-PR....--V-D-PG---  97
G.SE.93.SE6165          --------------Y---A-----------------L-.---------N------E----------------.-----Q-----I-PR....--V-D-PG---  97
H.BE.93.VI991           ---------------------------S-----Y--P-.--D------N------E-------T------A-.-----Q-----I---....--V---P----  97
H.BE.93.VI997           -----------------*---------S--------D-.--Q------N------G----L--T--------.-----H-----II--....--V---P----  96
H.CF.90.056             -------------------------I----------V-.--Q------N------V----L--T--------.-----Q-----I---....--V---P----  97
J.CD.97.J_97DC_KTB147   --------------Y------------R--------P-.--N------N------E----------------.-----H-----II--....--G---S----  97
J.SE.93.SE7887          -------H------YH--------------------P-.--S---Y--S------E----------------.-----H-----IIP-....--G--------  97
K.CD.97.EQTB11C         --------------Y----I-I-----R--------P-.--D------NN-----E----------------.-----H-----IL--....--GG-------  97
K.CM.96.MP535           --------------N------I-----R--------P-.--N------T------EL---------------.-----H-----IIP-....--G---S----  97
```

```
                                                                                    Vpr end in HXB2
                                                                       frameshift in HXB2
                                       amphipathic α-helix
                           Vpr start   oligomerization                               H(S/N)RIG motifs      Vpr end
B.FR.83.HXB2               MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIW.LHGLGQHIYETYGDTWAGVEAIIRILQQLLFIH#FRIGCRHSRIGVTRQ....RRARNGASRS*   96
01_AE.CF.90.90CF11697      --------------Y------------I--------P-.------Y--N------E-------T------V-.-----Q-----II-G....--G----G---   97
01_AE.CN.05.FJ051          --------------Y-A-A-----D-----------P-.---------D------E-------T------V-.-----Q-----ILPG....--G----G---   97
01_AE.CN.06.FJ054          ------------V-S---------------------P-.------Y--NN-----E-------T------V-.-----Q-----IIPG....--G----G---   97
01_AE.HK.x.HK001           --R-----------Y---------------X----XP-.--S---Y--D------E-----T-T------V-.-----Q-----IIPG....--G--------   97
01_AE.TH.01.01TH_R2184     -------R------Y------------T---K----P-.---------N------E--------------V-.-----Q-----IMPG....--G----G---   97
01_AE.TH.02.OUR769I        --------------Y---------------------P-.------Y--NN-----E--------------V-.-----Q-----IIPG....--G--------   97
01_AE.TH.90.CM240          ------#f------Y---------------------P-.-Q----Y--NN-----E-------M------V-.-----Q-----IMPG....--G---TG---   95
01_AE.US.00.00US_MSC1164   --------------Y---------D-----------P-.------Y--YN-----E----LVIF------V-.-----Q-----IIPG....--G----G---   97
02_AG.CM.02.02CM_4082STN   --R-----------S---------------------P-.---------D------E----------H---V-.-----Q-----II-G....--G----G---   97
02_AG.EC.x.ECU41           --------------F------------H--------Q-.------Y--N------E--V-----------V-.-----H------I-G....--G----G---   97
02_AG.GH.03.GHNJ196        --------------S-Q----------Q--------P-.------Y--N------E--------------V-.-----Q-----II-G....--G---SG---   97
02_AG.NG.01.PL0710         --R-----------F------------H--------P-.----------------E--R---T---------.-----Q-----IIPG....--G----G---   97
02_AG.NG.x.IBNG            --R-----------S------------H--------P-.----------------E--K-----------V-.-----Q-----IIQG....--G----G---   97
02_AG.SE.94.SE7812         --R-----------F------------H---T----P-.----------------E--G-----------V-.-----Q-----II-G....--G----G---   97
02_AG.SN.98.MP1211         --R-----------F------------H--------P-.----------------E--G-----------V-.-----Q-----IIPG....--G----G---   97
03_AB.RU.97.KAL153_2       --------------Y------------S--------V-.--S---Y--------------------------.-----Q-----IIQR....-----------   97
04_cpx.CY.94.CY032         --------------N---------------------P-.---------N------E----------------.-----Q-----I-P-R#..EE-GD------   98
05_DF.BE.x.VI1310          --------------Y---I--------H--------P-.--S---Y--N------E-----V----------.-----------IL--....--I---S----   97
06_cpx.AU.96.BFP90         --------------Y------------S--------P-.--NI--Y--S------E----LL-T--L---V-.-----H-----IIP-....--G--------   97
08_BC.CN.97.97CNGX_6F      ---P-----------------------Q--------P-.------YV--------T---T-L----------.-----Q-----IV--....-----------   97
09_cpx.GH.96.96GH2911      --------------N---------------------P-.---------D------E--------------V-.-----H-----II-GR...--G-D--G---   98
10_CD.TZ.96.96TZ_BF061     --------------Y-----------------------.--N-------------T----------------.-----P-----IV--....------T-T--   97
11_cpx.GR.x.GR17           --------------YS--M--------H--------P-.---------N------E--------------V-.-----Q-----II-G....--G----R-Y-   97
12_BF.AR.99.ARMA159        ---P----------Y---------------------P-.---------N------E----------------.-----Q-----IM--....--V--------   97
13_cpx.CM.96.1849          --------------N------------H---K----P-.---------N------E-----T-M--------.-----Q------V--....--G---S----   97
14_BG.DE.01.9196_01        --------------Y---------------------P-.------Y--N------E----------------.-----Q-----IIP-....--V---PG---   97
15_01B.TH.99.99TH_MU2079   -G------------Y-G-P------------KFS.-P-.------S--TN-----E-----V---H----V-.-----H----AIIPGR...-AR---PG---   97
16_A2D.KR.97.97KR004       --------------Y---------D--R--------P*.--S---Y--A------T-----------M----.-----Q-----IMP.....-G----S----   95
17_BF.PE.02.PE02_PCR0155   -----G--------Y---------------------P-.--S------N------E----------------.Y----H-----I---....--V--------   97
18_cpx.CU.99.CU76          --------------N---------------------P-.------Y--N------E----------------.-----Q-----IV--....--G--------   97
19_cpx.CU.99.CU7           ------N-------Y---A-----D-----------P-.------YV-T------E----------------.-----Q-----I-PR....--V-D-PG---   97
20_BG.CU.99.Cu103          ------------------------------------P-.-----V---N------E----L-----------.-----Q-----ISLR....--V---P----   97
21_A2D.KE.91.KNH1254       --------------S------------S--------T-.--S-------------------T----R-----.-----Q------V--....------S----   97
22_01A1.CM.01.01CM_0001BBY ------------V-Y------V-------------GP-.-----E---N------E--------------V-.-----Q-----II-G....--G----G---   97
23_BG.CU.03.CB118          ------------Q-----------------------P-.-----V---N------E--------------T-.-----Q-----ISPR....--G---PG---   97
24_BG.CU.03.CB378          ------------------------------------P-.---------N------E----------------.-----Q-----ISLRRR..L-G-D-SG---   99
24_BG.CU.03.CB471          --------------Y---------------------P-.---------N------E------------M---.-----------ISL-....--G-D-SG---   97
25_cpx.CM.01.101BA         ---T----------N------------H--------P-.------X--N------E--V-------------.-----Q-----I-P-....--G---P----   97
27_cpx.CD.97.97CDKTB49     --------------N------------T--------P-.--D----V-N------E----------------.-----H-----II--....--G---S----   97
28_BF.BR.99.BREPM12609     --------------Y------------S----------.---------------------L-------M---.-----------IN--....-----------   97
29_BF.BR.02.BREPM119       ---------S----Y---A--------S--------V-.--S---Y---------T----L-----------.-----Q-----II--....--P--------   97
31_BC.BR.02.110PA          --R-----------Y------------Q--------P-.-------------------------------V-.-----Q-----IL--....-----------   97
32_06A1.EE.01.EE0369       --R------------------------S---K----P-.--S---Y--N------E----------------.-----N-----IIP-....--G---T----   97
33_01B.MY.05.05MYKL007_1   --------------R---A------------------Q-.---------H------E----M--T------V-.-----Q-----IIPG....--G----G---   97
34_01B.TH.99.OUR2478P      -----A------------------------------P-.------Y--DN-----Q--------------V-.-----Q-----ILPG....--G--------   97
35_AD.AF.05.05AF095        --R-----------Y------------H--------P-.------Y--N-----------------------.-----------II-G....--V-D-SG-P-   97
36_cpx.CM.00.00CMNYU830    --------------F------------H--------P-#------Y---------E-----M--------V-.-----Q-----II-G....--G----G---   97
37_cpx.CM.00.00CMNYU926    --------------Y---A-----------------P-.------Y---------Q------K---------.-----Q-----I-PR....--V---P----   98
38_BF1.UY.05.UY05_4752     -----G--------Y------------H--------P-.--S---Y--N------E------------M---.-----H-----I---....--V---N----   97
39_BF.BR.03.03BRRJ103      ----S---------Y-Q----------G----------.-------------------K-------------.-----------INP-....-----------   97
40_BF.BR.04.04BRRJ115      ------N-------Y------------S--------P-.--N---Y---------T-----------M----.-----H-----IM--....--S---S----   97
42_BF.LU.03.luBF_05_03     --------------Y---------------------P-.----------------E---V------------.-----------INIR....-------N---   97
43_02G.SA.03.J11223        ---V----------X---------X--H--------P-.---X-----N------E-------X--X---X-.-----Q-----II-G....--G-H--G---   97
U.CD.83.83CD003            --------------Y-A----------H--------P-.--N---Y--N------Q-------S--------.-----Q-----I---....----H-P----   97
U.CD.90.90CD121E12         --------------Y------------Q--------P-.--N------N------Q----L-----------.-----Q-----II--....----H-P--P-   97
U.GR.99.GR303              --------------N------------H--------PC.---------N------E-----L----------.-----Q-----IV--....--G---S----   97
U.NL.95.U_NL_95_H10986_D1  ----------P---Y-Q-A--I-----R--------P-.---------N------E--------------T-.-----Q-----ILQ-....--G----N---   97
CPZ.CD.90.ANT              -----Q-E------M---L--T---I---------QPT.-QH--NWV-AN---S-R--QVL-IL--KA--T-.Y-H--A-----............H-PR-R-   89
CPZ.CM.01.SIVcpzCAM13      --H-------P---Y---A--I--D------K-Y--S-.--Q------T------E--------------V-.--V--H-----IIP-....--RA---N-P-   97
CPZ.CM.05.SIVcpzEK505      --------------Y---A-----D-----L-----P-.------YF-N------E-------T--------.Y----Q-----I-P.....Q-R--------   96
CPZ.CM.05.SIVcpzLB7        ---V----------Y---A--V-----Q--------P-.--Q---Y--D------E--Q-------------.Y----Q-----I-P-....--R-D--N---   97
CPZ.CM.05.SIVcpzMB66       ---L----------Y------I-----R--------D-.--Q------T-----------------------.Y----H----ININQQ...--R----N---   99
CPZ.CM.05.SIVcpzMT145      ----------S---YEQ-L-DT---I-Q--------P-.--A--E--------S----Q-----------V-.-----H-----INP-....--R-D-TN-A-   97
CPZ.GA.88.GAB1             ----------P---YQ--A--T--------------P-.--Q---F--D------V----------H-----.--L--Q-----ILP-....--RS--SN---   97
CPZ.TZ.00.TAN1             -----N-N------YT--L-DI---I-Q---K----PI.-Q-V-NWVFTI---S-E--QEL-K---RA--T-.Y-H--I-----#.......fESH--S   88
CPZ.US.85.CPZUS            ---V------A---Y---A-DV-----E--L-----P-.-Q----Y--S------E----------------.Y----N-----ISLARRTPQGR---S----  101
CPZ.CM.x.SIVcpzDP943       ----------A---Y---A-----------------A-.---------D-----------------------.Y----Q-----INP.....G-R--------   96
CPZ.CM.x.SIVcpzMB897       --------------Y--------D---Q--------Q-.--D------N-----------------------.Y----Q-----ILP-....G-R---SN---   97
N.CM.02.DJO0131            -------A------Y---------------------T-.---------N------E----------------.Y----Q-----I-P.....-GR--------   96
N.CM.04.04CM_1015_04       -------A------Y---A-----D-------------.---------D------E-------MX-------.Y----Q-----I-P.....Q-R--------   96
N.CM.04.04CM_1131_03       -------A------Y---A-G---D-------------.---------D------E-------M--------.Y----Q-----I-P.....Q-R--------   96
N.CM.95.YBF30              --R----A------Y---A-------------------.---------N------E----------------.Y----Q-----I-P.....Q-R---T----   96
N.CM.97.YBF106             --R----A------Y---A-------------------.---------N------E----------------.Y----Q-----I-P.....Q-R--------   96
O.BE.87.ANT70              ------N---AK--F---A--------A--------P-.--A---Y---------V--M-----------T-.Y----Q-----INPRG...-GR---S----   98
O.CM.91.MVP5180            ----L-N---A---F------------E--------P-.-QAC--Y---------E--M-----------T-.Y----Q-----ILPSNTRG-GR---S----  101
O.CM.96.96CMABB637         ----------A---F--X---------A--------A-.-Q----Y--D------E--M---------I-A-.-----Q-----I-PSNARG-GR---S----  101
O.CM.98.98CMA104           ---T------A---F---A-----D--T--------P-.-XA---Y---------L--M--T-----I--A-.-----Q-----INPSNTRG-GR---S----  101
O.CM.99.99CMU4122          -------H--A---F---A--T--D--A--------P-.--A---Y---------V--I--M---------T-.-----Q-----INPSNTRG-GR---S----  101
O.FR.92.VAU                ----------A---V-A-A--------A--------N-.-QA---Y---------E--I--T-----II-T-.Y----Q-----IIPSNARR-GR-D-SH---  101
O.SN.99.SEMP1299           ----------A---F---S--------A--------P-.-QA---Y--D------V--M-----------A-.-----Q-----I-PSNTRG-GR---S----  101
O.US.99.99USTWLA           V--VS-----A---F----IQ------T--------P-.------Y---------E--M--------I--T-.Y----Q-----IIPSNTRG-GR-D-S--P-  101
```

```
                          Tat start        disulfide bonding         NLS        exon 1 end  exon 2 start          Tat end
B.FR.83.HXB2              MEPVDP.RLEPWKHPGSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRQ..RRRAHQNSQTHQASLSKQPTSQ..PRG.DPTGPKE*KKKVERETETDPFD*   100
A1.GE.99.99GEMZ011        -D----.N----N----------S--------W---L--LK-G-----------H..-CGTPHS-KD--IPIP---LPH..TQR.-Q---E-S-----SKA---R---   102
A1.KE.00.KER2008          -D----.N-A--Q------T-P-NK-F-----Y--P--LLN-G------------..--GTP-SNKD--NPIP---LP-..TQ-.NS---E-S-----SK--A-R-A-   102
A1.KE.00.KNH1144          -D----.N-D--N------T-P-NK-------Y---Y--LN-G------------..--GTPHS-KD--NP-P--SMP-..TQ-.VS---E-S-----SQA-A-R---   102
A1.KE.00.KSM4024          -D----.N----N------T---SK--------------LN-G------------..--G-P-SNKN--NPIP---IP-..AQ-.VS-----S--M-SKA-A-R---   102
A1.KE.00.MSA4069          ------.D----N------T-P-SK------------S--LN-G------------..--GTL-S-KD--NPVP---IP-..TQ-.NS-----S--E-SKA---R---   102
A1.KE.00.NKU3005          -D----.N----N------T-P-SK-------Y--L---Q--G-----------K-..--GTP-S-KD--NPIP----P-..A--.-------ST----GKA---R---   102
A1.RU.00.RU00051          -D----.N----N----------SK-------W------XX-G----------KR..--GTPHS-KD--TPIP---LP-..TQR.XQ---E-S-TE--XKAK--R---   102
A1.RW.93.93RW_024         -D----.N----N------T---NK-------Y--PS--LN-G------------..--GTP-S-KD--NPIQ--SIP-..TQ-.I----E-S-----SK----R---   102
A1.SE.95.SE8891           -D----.N----N------A---SA-------Y--P---LN-G------------..---TP-S-KD--NPVP--SIP-..AQ-.-S---E-S-----SK----R-A-   102
A1.SE.95.UGSE8131         -D----.N----N------T-P-NKYF--R--Y--L---QH-G------------..--GTP-S-KD--DPIP--SIPP.AQ-.I----E-S-----SKA--R-A-   102
A1.TZ.01.A173             -D----.N----N------A---SP-------Y------LN-G-----------K-..---PP-S-KD--NPIP---IP-..TQ-.IS---E-S-----SKA---R---   102
A1.UA.01.01UADN139        -D----.NI---N----------S-----R--W---I--LK-G-----------P..--G-PHS-KD--NTI----LPH..TQR.-Q---E-S-----SKA---RL--   102
A1.UG.92.92UG037          -D----.S----N--------P-NK---V---Y---C--LN-G----------KP..--GTP-SNKD--NPIP---IPR..TQ-.-S---E-S-----SK--A-RYA-   102
A1.UG.99.99UGA07072       -D----.N----N------A---ST-H-----Y--P---LN-G------------..----P-S-KD--NPVP---IP-..TQ-.VS--QE-S-----SK----R---   102
A2.CD.97.97CDKTB48        ------.N----N----------N--------Y---R--LN-G-----------P..--GPD-GNTD--NPVP--SLP-..TQR.-S---E-S-----SKA---R---   102
A2.CY.94.94CY017_41       ------.K----N-----------K----R--Y---L---N-G-----------P..--KPSPSNKD--NPIP--SLP-..AQR.V----E-P--E--SKA---R---   102
A.SN.01.DDI579            ------.N----H----------S-----A--W------LN-G-----------K-..--GTP-SNKD--NPVP---LP-..TP-.-------S-----SK----R-A-   102
A.SN.01.DDJ369            ------.N----Q----------S-----A--W---A--LN-G------------..--GTP-SNKD--NPVR---IP-..AQ-.-S---E-S-----SK----Q---   102
A.SN.96.DDJ360            ------.N----H----------NK----A--W------LN-G-----------K-..--GTS--NKD--NPVP---IP-..TQ-.-S---E-S-----SK----R---   102
A.CD.02.02CD_KTB035       --L---.N-------------P-SK-------Y--YS--LK-G----------TH..--GTSHS-KD--NHIP--SLP-..TQW.ES---E-S--EM-SK----R---   102
A.CD.97.97CD_KCC2         ------.N----E--------P-NR-------Y--P---VN-G----------KL..---TPRSNKD--NPIP---L--..AQR.-----E-S--E--SKA-P-R---   102
A.CD.97.97CD_KTB13        ------.N---------------NR----V--Y--YM--LK-G---F------KL..--GTPYS-KD--NPIP---LP-..AQR.-S---EKS--E--SK--P-R-A-   102
B.AR.04.04AR151516        ------.--------------P--K--------------V--G------------..---PP-D-----VP---------.L--.-T---EXS-----------DA-   102
B.AU.87.MBC925            ------.------------R----P----------------------------..----P------V------A--.---.--A---S------------R--   102
B.BO.99.BOL0122           ------.----------------------------------K-G------------..---SPPD-----VP-P---A--.---.-----Q--S----------H-GR-   102
B.BR.03.BREPM2012         ------.--------------P--K-------L-------K-G------------..----P-D-EN--V-----SA--.---.-----E-S-----S------K--   102
B.CA.97.CANB3FULL         ------.------------R---N-----------------S-G------------..----P-D-E---V------S--.L--.-------S------------DA-   102
B.CN.05.05CNHB_hp3        ------.----------------N-------S-------TK-G------------..---TP-D-----V--P--------.---.--E---S-----S--A--R--   102
B.CO.01.PCM001            ------.--------------------------------TK-G------------..---S-PD-KI---P--T-------.---.---------S-E---------RV--   102
B.GB.83.CAM1              ------.------------R---------Q------------G------------..---TP-S-K----------A--.FQ-.-------S-----G----H-G--   102
B.GB.86.GB8               ------.-I----------R--------------------LR-G------------..---LPED--I--V--P-------.-Q-.-------S-----SK-----S-W   102
B.GE.03.03GEMZ010         -D----.-----------------------------------G--------W--..----P-D-K------P---A--.---.--------Q-----SK-----DH-   102
B.IT.05.SG1               ------.N----------------T-------Y---Y--TS-G------------..----P-D-P-D-V------I--.-H-.-------S---------M--VA-   102
B.JP.05.DR6538            ------.-----E------R---N---------------TK-G------------..--GPP-G-----V--P---L--.S--DH----T-S-------A----EY-   103
B.KR.05.05CSR3            ------.-----N----K-E-P-N--------L------TR-G------------..----P-D-----V----------.-P-.-------S-----------VNQ   102
B.NL.00.671_00T36         ------.----------------N---------------TR-G------------..---TP-D-----V----------.L--.-------S-----------GH-   102
B.RU.04.04RU128005        ------.-----E-----------A----------------K-G------------..---SP-S-EA--VP--N-----.-G-.-------Q----------N-DA-   102
B.TH.00.00TH_C3198        ------.----------------------------------K-G---F--------.------D----------------.---.--A---S-----G---A--R--   102
B.UA.01.01UAKV167         ------.------------R----K---------------M--G------------..---PP-D-EA--VP-P---A--.---.-------S-E---SK-----DH-   102
B.US.04.ES10_53           ------.----------------NS-------Y--P------G------------..----TND--N--IP---SAP-.-G-.-------S-E----T-----DP-   102
B.US.99.PRB959_03         ------.------------R---------Q---------LR-G------------..---TPPEN----V------S--.L--.-------S-E---SK--------   102
C.AR.01.ARG4006           ------.K----N--------P--K------SY------L--G------------..--S-PPSGED--NLI-E--L---.T--.-QA-SE-S-----SK-K------   102
C.BR.04.04BR013           ------.N----N--------P-NS----S-SY------L--G------------..--S-SSS-ED--NLI----LP-.----.-Q--SE-S-----SK--------   102
C.BW.00.00BW07621         --Q---.N----N--------P-NK-F--S-SY--L---Q--G------------..--H-PPG-KD--NPI----LP-.T--.----SE-S-----SK--------   102
C.CN.98.YNRL9840          ------.N---------------N-----Y-SY--L---Q--G-----------R.--S-P-S-ED--NLI----LPR.TQ-.----SE-S-----SK-K------   102
C.ET.02.02ET_288          ------.N----N--------P---------SY--P---L--G------------..--S-PPS-ES--DPI----LP-.T--.N---SE-S-----SK------A-   102
C.GE.03.03GEMZ033         ------.N----N-----------K-F-----Y--PA--L--G------------..--S-PPS-ED--NPI----LPR.TQ-.NS--SE-S-----SK--------   102
C.IL.99.99ET7             ------.N----N------S---NT----A-SY--L---QK-G----------RP.G-S-PPS-EDN-NLI----L---.T--.-S--SE-S--E--SK-----Y--   103
C.IN.99.01IN565_10        ------.N----N----------N-----N-SY--L---QK-G------------..--S-PPS-ED--NLI----LPR.TP-.----SE-S-----SK-K--L---   102
C.KE.00.KER2010           --L---.K----N----------NT------SY--L---QK-G--------W--..----PPS-ED--DLI----LP-.T--.----SE-S-----SK-K------   102
C.MM.99.mIDU101_3         ------.N----N------E---N-----R-SY--L---QK-G------------..--S-P-S-ED--NPI----LPR.TQR.----SE-S-----SK-KA-----   102
C.MW.93.93MW_965          -D-I--.N----N----------NR----R--Y--L--LQ--G------------..--S-PPS-ED--NP-----LP-.T--.-S--SE-S-----SK-------W   102
C.SN.90.90SE_364          ------.N----N----------NK----R-SY--L---Q--G------------..--STPPS-ED--NPI----L---.T--.----SE-S-----SK--------   102
C.SO.89.89SM_145          ------.E----N---C------N-----L-SY--L---Q--G------------..--S-P-S-ED--NPI----LP-.T--.----SE-S-----SK-K------   102
C.TZ.02.CO178             ------.N----N--------P-NQ----H-SY--L---Q--G------------..-----PS-ED--DHI-E-SLPR.TQ-.----SE-S-----SK--------   102
C.UY.01.TRA3011           ------.N----N----------N--F----SY--L---Q--G------------..--S-PPS-ED--IPI----LP-.A--.-Q--SEKS-----SK--------   102
C.YE.02.02YE511           ---I--.K----N------R---NK----A--Y--L---QK-G------------..--S-PPD-ED--NLV----LP-.T--.-Q--SE-S-----SK------NW   102
C.ZA.04.04ZASK164B1       ------.N----N--------P-NP------SY--L---Q--G---Y-------Q.--S-PPS-KD--NPI---SLPR.TQ-.-S--SE-S--E--SK-KR-Q---   103
C.ZA.05.05ZAPSK240B1      ---I--.N---------------NG------SY--L---QK-G------------..--S-PPS--D--NP-P---L--.PQ--.----SE-S-----SK-K---S--   103
C.ZM.02.02ZM108           ------.N----N------Q-P--K-F--Y-SY--L---Q--G----------S-..--S-PPSNKD--DLV----L---.T--.----S--S-----SK--A--S--   102
D.CD.83.ELI               -D----.N----N------R-P-NK-H-----Y--P---LN-G------------..--GPP-GG-A--VPIP---S--.---.-------Q-----S-A----*-C   101
D.CM.01.01CM_0009BBY      -D----.N-----------S---NS----Q--Y---L---K-G----------G-..--GPP-G-----VPVP---S--.---.-------------SK--AP-Y-W   102
D.KE.01.NKU3006           ------.N----R------R-P-NK-------Y--Y------G------------..---SP-GD-A--DPIP---S--.---.-------------SKA-A----W   102
D.KR.04.04KBH8            -D----.N-----------R---N--------L------LK-G------------..---PP-GD-A--VP-----P--.---.-------Q--E--SK--ANQ--W   102
D.TD.99.MN011             -D----.N----N------S-S-SS----I--Y---L-----G----------K..--GLP-DNK---VAVQE--S--.H--.-------------SK--A----W   102
D.TZ.01.A280              -D----.N----N------R-P-N--------Y---N-----G------------..---TP-GD-A--DPIP---S--.---.-Q-----------SKA-A-R--W   102
D.UG.99.99UGK09259        ------.N----N------E-P-NK-------Y---H--V--G----------L.--S-SPGG-A--DPIP---S--.L--.-Q-----------SQA-AT-Y-W   102
D.YE.01.01YE386           ------.S----N------R-P--S----W--Y---L--M--G------------..---TP-DN-A--V-IP---S--.---.-----------------A----W   102
D.YE.02.02YE516           -D----.N---------R-S-P-SP-----------------G----------G-..---SP-SN----VPVP---S--.---.-------------SQAKAP---W   102
D.ZA.90.R1                -D----.N----N--------P-NK-H-----Y---L---S-G------------..--KPPHSD-D--VPIPE--S--.---.-------------S-A--NQ--W   102
F1.AR.02.ARE933           --L---.N-D--N------T-P-------R-----YR-LAR-G------------..-H-TPHGT-I--DLVP---L---.A--.N------S--E--SKAK---C--   102
F1.BE.93.VI850            ------.S-D--N------T-P--K----R------W--T--G------------..-H-TP-S--V--N--P---L---.A--.-------S--E--SKAK---CA-   102
F1.BR.01.01BR125          ------.N-D--N------T-P-NK----R-----YW--A--G------------..---TP-SD-I--DLIPE--V---.AQR.N------S-EE---KAK-H-*--   101
F1.ES.x.P1146             ------.N-D--N------T-P-NK----N-----YH--A--G-----------RH-H-TP-D-KV--NPVPE--L---.A--.-------S--E--SK-K---*--   103
F1.FI.93.FIN9363          --L---.N-D--N------P-P-NK----R-----YW--A--G------------..-H-TP-S--I--DPVP---L---.--R.N------S--E--SKAK---C--   102
F2.CM.02.02CM_0016BBY     --M---.K-D--N-------NP-NK----R------L--TS-G------------..--S-SHS-KD--DPV----L---.A--.-----E-S-----SK-K---*--   101
F2.CM.95.MP257            --V---.N-D---------E-P-NK-----------L--TR-G------------..---TP-SGEV--DPV----L---.T--.--K--E-S-----SK-K---S--   102
F2.CM.97.CM53657          --M---.K----N------E-P-NK-----------L--TR-G------------..----S-S-EI--NPVP---LP-.AG-.N-N--E-S-----SK-KP--S--   102
G.BE.96.DRCBL             -D----.K-------------P-N-----SVAAL.----LN-G----------KH..--GTPHS-KD--TPVP---F-T.T--.N----Q-S--E--SK--------   101
G.CM.01.01CM_4049HAN      -D----.N----N------I-P-N--------W---Y--LN-G----------SH..---TSPG-TD--NPVP---LPT.T--.N--DSE-S-----SK-K---Y--   102
G.CU.99.Cu74              -D----.N----N----------N-----R--W---S--LH-G----------KH..--GTP-G-KD--NPVP---LPT.T--.N------S--E-ASK--------   102
G.ES.99.X138              -D----.N----N------S---------R--W------LN-G----------KH..--GPP-SNKD--NPVP---LPT.T--.N----E-S---M-SK-K-N----   102
G.GH.03.03GH175G          -D----.N----N------T-P-NK----V--W------LN-G----------K-..--GTP-G-KD--NPVP---LPI.S--.NQ--*--P--E-ASK---G-CA-   101
G.KE.93.HH8793_12_1       -D----.N----N----------NK----I--W------LN-G----------KH..---PP-DNKD--NPVP---LPT.T--.N------S-----SK-K------   102
G.NG.01.01NGPL0669        -D----.N----N------R---N--------W------LN-G-----------H..--STPAG-KD--NPVP---LPT.T--.N------S--E--SKA-------   102
G.PT.x.PT2695             -D----.N----N-----------K----V--W------LN-G----------KH..--GTP-S-TD--NPVR---LPT.T--.N--D-E-S--EM-SK---N----   102
G.SE.93.SE6165            -D----.N----N--------P-NK-F--V--W------LN-G----------KH..--GTP-S-KG--DPVP---LPT.T--.N------S--E-ASKA-A-QC--   102
H.BE.93.VI991             -D----.NQ---N------R---N--------Y---L--LK-G---Y--------..--GTPKSL-D--TLIP---L-R.TS-.-----EK-----ASK-------W   102
H.BE.93.VI997             -D----.NQ---N---N----P-N-----Q-S----L--LK-G----------SR..--ATPASV-D--NHIP---L-R.T--.------------ASK--A--C--   102
H.CF.90.056               -D----.K----N------Q---N--------Y---M--LK-G----------S-..-H-TPASL-D--N-I----L-R.TH-.-------Q--E-ASK-----*--   101
J.CD.97.J_97DC_KTB147     -D----.NI---NQ---------NQ----R--Y--PL--LK-G----H-------..--A-PSG-K---DPIPG--L---.TQR.ES--QEKS--E--SKA-P-R---   102
J.SE.93.SE7887            ------.NR---N-------------------Y------LQ-G------------..--S-PPG-K---DLIP---L---.TQR.K----E-S--E--SKA-P-R---   102
K.CD.97.EQTB11C           ------.NI---NQ---------NQ-------Y---L--LQ-G---C---E----..-TTTPYA-KN-KDPIP---LP-.A--.-------S--E--SK-K---*--   101
K.CM.96.MP535             -D----.NI---NQ---------NQ----R--Y---I--LK-G----N------P.--TTPY--EN--DP-R---L--.---.EQ-D---S-----SK-K--Q---   102
```

```
                              Tat start            disulfide bonding           NLS          exon 1 end exon 2 start              Tat end
B.FR.83.HXB2                  MEPVDP.RLEPWKHPGSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRQ..RRRAHQNSQTHQASLSKQPTSQ..PRG.DPTGPKE*KKKVERETETDPFD*  100
01_AE.CF.90.90CF11697         ------.S----N------T---SK-------W---L--LK-G----------KH..--GPPPG-KD--NPIP---LPT..T--.N------S--E-AKKAK----A-  102
01_AE.CN.05.FJ051             ------.N----N------T---N--------W---L--LK-G----H-------..--GTP-S-KD--NPVP---L-I..I--.N------S--E-ASKA----*--  101
01_AE.CN.06.FJ054             ------.N----N----K-T---SK----I--W---L--LQ-G----------KH..--GTP-S-KD--NLIRE-SLPI..I--.NQ----KS----ASK-----C--  102
01_AE.HK.x.HK001              ------.N----N----K-T----K----I-SW---L--LK-G----------KH..--GTPRS-EG--NPVP--SL-T..S--.NQ-D-E-S----AGK-----*--  101
01_AE.TH.01.01TH_R2184        ------.N----N------T---SK----I--W---L--LK-G----------KH..--GTP-S-KD--NP-PE--LPI..I--.N--D---S--E-ASK----R*--  101
01_AE.TH.02.OUR769I           ------.N----N----K-T---SK----I--W---L--LK-G----H-----KH..----P-S-ED--DPIP--SLPI..I--.N------S--E-ASK-----C--  102
01_AE.TH.90.CM240             --L---.N----N------T---SK-------W---L--LK-G----------KH..--GTP-S-KD--NPIP---LPI..I-R.N--D---S--E-ASKA---QC--  102
01_AE.US.00.00US_MSC1164      ------.N----N------T---S--------W---I--LK-G----------KH..--GTP-S-KS--NPIRE--FPI..I--.N------S--E--SKA----C--  102
02_AG.CM.02.02CM_4082STN      --L---.N----N------T---ST-------W---L--LN-G-----------R..--GTP-SH-D--SPVP----TT..T--.N-A--E-ST---ASK---G-*--  101
02_AG.EC.x.ECU41              ------.S-D--N------A---NK----M--W---I--LK-G-----------R..--GTP-SRPD--NPVP---LPT..T--.N------S--E-ASK---N-C--  102
02_AG.GH.03.GHNJ196           -DL---.N-D--N------T---SK----I--W---LG-LN-G-----------P..--GTP--R-DY-NPAP---LPT..T--.IQADQ--S----ASKP-P--C--  102
02_AG.NG.01.PL0710            --L---.S----N------T---SK----I--W---L--LK-G-----------R..--GTP-SH-D--NPV----LPT..T--.N---A--S--E-ASK----RCA-  102
02_AG.NG.x.IBNG               --L---.S----N------T---SK----M--#---L--LN-G-----------R..--GTP-SR-D--NPVP---LPT..T--.N--D---S--E--SK-K---C--  101
02_AG.SE.94.SE7812            --L---.S----N------T---SK----I--W---L--LN-G----------KR..--GTP-SR-DN-DPVP---LPT..T--.N-A----S--E-AGK-----C--  102
02_AG.SN.98.MP1211            --L---.K----N------T---SK----L--W---L--LN-G-----------R..--GTP-SR-D--NPVP---LPT..T--.EQ-----S--E-ASK---G-C--  102
03_AB.RU.97.KAL153_2          ------.---------------------L------MK-G-------------..----P-DN--D-V--P---A--..---.-----------M------H---  102
04_cpx.CY.94.CY032            ------.D----N------T-D-NK-F-----W------LK-G----------KH..--GSL-G-KG--NLIP---L--.Q-N-.-S---E-Q----ASK--A---A-  103
05_DF.BE.x.VI1310             -D----.N----N------R---NQ----M--Y---N-----G-----------P..---PP-G--A--DPVPE--P--..---.------KQ--E--SK--A-Q--W  102
06_cpx.AU.96.BFP90            ------.KI---NQ---R------K-------Y--P---LN-G-------------..--Q-PPG-KN--DPV----L--..TQR.EQ---EKS--E--SKA-P-R---  102
08_BC.CN.97.97CNGX_6F         ------.N----N------E---N-----R-SY--L---QK-G-------------..--S-P-S-ED--NLI---SLPR..TQ-.-QA-SE-S-----SK-KPA----  102
09_cpx.GH.96.96GH2911         --L---.N----N------A---NK----R--Y---H--LN-G-------------..--A-SCG-KSN-DPIP---LP-..T--.N------S--E--SK-KS-Q---  102
10_CD.TZ.96.96TZ_BF061        ----H-.N----N----------NK-H--A-SY--L---Q--S-------------..--S-PPS-ED--NLI----S--..---.----------------A-A----W  102
11_cpx.GR.x.GR17              ---G-I.NID--NQ---------NQ----R--Y---H--LK-G---------W--..--T-SRS-KN--DPIPE--LP-..ASR.N----E-P--E--SKA-PA----  102
12_BF.AR.99.ARMA159           --L---.N----N----K-S-P--K----R-----YW--V--G-------------..----P-G---N--------I--..A--.N------S--E--SKAKS--C--  102
13_cpx.CM.96.1849             -D----.KI---NQ-------P--K---------------G-------------..--S-SH---N--DLIP---LPI..---.NQ-----S--E--SK-------W  102
14_BG.DE.01.9196_01           -D----.N----N-----------------W------LN-G----------KR..--GPP-S-ED--NPVP---LPT..T--.N----E-S-----SK-K-N----  102
15_01B.TH.99.99TH_MU2079      -D---L.N----N----R-T---N-----I--W---I--LK-G----------KH..--GTP-S-KD--NPAP---L-I..S--.K--D--KS--E-ASKA----*--  101
16_A2D.KR.97.97KR004          -D----.N----N------R---NK----R--Y--P---LN-G-----------P..---PP-GGTG--NPIPE-SLPR..-QR.I---TE-S-TE--SKA----C--  102
17_BF.PE.02.PE02_PCR0155      --L---.N-D--N------TXP--R----R-----YW-----G-------------..---SP-G-K---S--P---IP-..A--.N------S--E--SKA----*--  101
18_cpx.CU.99.CU76             ------.NI---NQ---------N-----R--Y---L--LK-G-------------..--TTPYD-KH--DLIPE--L-R..TSW.-------Q-E--ASKA-A-QC--  102
19_cpx.CU.99.CU7              -D----.N----N----------NT----------L--TK-G-------------..--GPP-SH-D--NPVP---LPD..TE-..---T--S--E--SK-K------  101
20_BG.CU.99.Cu103             -D----.N----N----------NK----V--W------LN-G----------KH..--GPPPG-KD--NPVP---IPT..T--.N------S-EE--SK-T---CA-  102
21_A2D.KE.91.KNH1254          -D----.N----N------R-P-NK-H-----Y--T---L--G----------S-..---PP-SG--P-VHIP---S--..---.-------------S---------  102
22_01A1.CM.01.01CM_0001BBY    --L---.S----N------T---N--F--V--W------LK-G----------KH..--GTP-G-KD--NPIP---LP-..TKR.-S---E-S--E--SK---NR-A-  102
23_BG.CU.03.CB118             -D----.N-D--N----------NK----M--W------LK-G----------KH..--GTP-D-KD--IPVP---LPT..T--.N------S--E--SK--A-----  102
24_BG.CU.03.CB378             -D----.N----N----------NK----V--W---L--LK-G----------KH..--GTPLSNKD--NPVP---LPT..A--.N-----KS--E--SK-KA-----  102
24_BG.CU.03.CB471             -D----.N----N----------N-----I--W------LN-G-----------H..--GTPPS-KD--NPVP---IPT..T--.N------S----ASK--P-----  102
25_cpx.CM.01.101BA            -D----.NI---NQ-----T---N-----I--W---L--LN-G-------------..---TPPG-KD--YPVP---IPT..T--.NQ-----S----ASK-A---S--  102
27_cpx.CD.97.97CDKTB49        -D----.NI---NQ---------NQ----S--Y---L--LK-G---FH-------..--ATPYG-KN--DPIP--SIP-..TQR.VS---E-SE----SKA---RL-W  102
28_BF.BR.99.BREPM12609        ------.-----Q------R---NK-------L------MK-G-------------..---P--DN-NN-V------A--..S--.N----E-S--------T----A-  102
29_BF.BR.02.BREPM119          ------.--------------P-N-----R------K-----G-------------..---SPRH-PA---P-----S--..---.-------S--------KA-Q---  102
31_BC.BR.02.110PA             ------.N--X-N----------N-------SY------L--G-------------..--S-PPX-ED--NLI----LP-..X--.-Q--SE-S------K-K------  102
32_06A1.EE.01.EE0369          -----H.KI---NQ---R---P--P-------Y--YM--LN-G---Y--------..--T-PPG-KN--DPE----V--..TQR.E----EK---E--SKATPNR---  102
33_01B.MY.05.05MYKL007_1      --L---.N----N----K-T---S-----R--W---L--LK-G----------KHKH--GTP-S-ED--NP-P---L-I..S--.N--DS--S----ASKA--GLCA-  104
34_01B.TH.99.OUR2478P         ------.N----N------T---SK----N--W---L--LK-G----------KH..--GTP-S-KD--NPIPE--LPI..I--.NQ-----S--E-ASKA----*--  101
35_AD.AF.05.05AF095           -DL---.D-D--N------T-P-NK-F--V-SY------LN-G-------------..---TP-S--D--NP-P----P-..AQ-.IS---E-S-----SKA-P-R---  102
36_cpx.CM.00.00CMNYU830       --L---.NI---N------R---N--F-----W------LN-G-----------R..--G-P-SR-D--DPV----LPA..I--KN------S--E-ASK---N-C--  103
37_cpx.CM.00.00CMNYU926       -D----.N---------------N---------------TK-G-------------..--T-A-SR-DN-H-VP---LPT..T--.N------S--E-ASKA----CA-  102
38_BF1.UY.05.UY05_4752        --I---.N-D--N------A-P-N-----R-----YY--EK-G-------------..--KTSSS-PI--DPVP--S-P-..A--.N------S--E--SKAK---*--  101
39_BF.BR.03.03BRRJ103         ------.----------------NK----L--Y----------G-------------..-P-TP-SN-L--DLVPE--L--..AQ-.N------S--E--SKAK--LS--  102
40_BF.BR.04.04BRRJ115         -D----.-----------------K-F-----------M--G-------------..---SP-GR--C-V------A--..H--.-------Q-E----------E--  102
42_BF.LU.03.luBF_05_03        ---I--.K----E----------------R----------R-G-----------P..--NPPPD-E----------AT-..---.-------X-----------DX-  102
43_02G.SA.03.J11223           --L---.S----N------T---SK-------W---L--LN-G-----------R..--GTP-XR-D--NPV----LPI..TSE.N------S--E--SK--------  102
U.CD.83.83CD003               -D----.-----N----------NS----R--L------M--G----------GK..---TP-SGPN--NIV----S--..---.----QE-P-----KK-T------  102
U.CD.90.90CD121E12            -D----.K----N----------N--------L------VK-G----------GK..---TP-SDPD--NIVP---L--..---.-----E-P-----KK-IP-----  102
U.GR.99.GR303                 -D----.NI---NQ--------------------------R-G-------------..--A-PHDNKN--DLIP--SL--..A--.-S-----S-----SK-K------  102
U.NL.95.U_NL_95_H10986_D1     ---I--.KI---NQ---------NQ-------Y---L--LK-G----N------P..--TTPYD-KN--DPIPE--F--..A--.-------P--E--SK--------  102
CPZ.CD.90.ANT                 -D---A.ETP--L--PAT-A-P-N-----C--Y--PL--TK-G--------R-ARRN--TTAES-ENN-DPV--*SLPK..TSR.IQSSQ-K-E----EK-GSGGGPC  103
CPZ.CM.01.SIVcpzCAM13         ---I--.E---------------N-----A-----MY--TK-G----------AR..--T-FAS-KNN-DPVR---L--..---.NQERSEKQAE---SQAAA--*--  101
CPZ.CM.05.SIVcpzEK505         ------.E----N----------N--H--V--Y--VY--TK-G----------S-..---TP-SNKS--DP-P---L--..RL-.-Q--Q--Q--TL-S*-TA--CA-  101
CPZ.CM.05.SIVcpzLB7           ---I--.S-D--N--------P-N--F-----Y--PF--TK-G---------W-S..---TSKS-PHY-DPI----FPP..S--.-SLS-E-Q-----SKAQ--Q*-C  101
CPZ.CM.05.SIVcpzMB66          ---I--.N----N----------N-----Q--Y---L--TK-G-------R--K-..---TSES--N--DPVP---L--..-G-.IE--Q-KS--E--SQ-TS-Q-A-  102
CPZ.CM.05.SIVcpzMT145         ---IE-.D----N----K-R-E-NK-F-----Y--IL--TK-G----------GR..-KP-SS-QD.NKDP-R---L--..-IR.-Q--SEKQ--E---TS-A-QS-W  101
CPZ.GA.88.GAB1                -D-I--.D-----------R-V-N-----A--Y--IY--TK-G----------TTR.--T-PAG-KNN-D-IP---L--..S--.NKE-SEKST-E-ASK--A-Q*--  102
CPZ.TZ.00.TAN1                ....#-.QVA--E---AA-E-P------------P---TK----------R-G...-K.SAVH-TNN-DPVRQ-SLPK..RSR.IQNSQE-SQEE--A--TSGGRPR   95
CPZ.US.85.CPZUS               -D----.NI----Q---------N-----------V---TK-G----------GRR.--TTAKS-ADN-DP-R---F--..QL-.NQ--Q--Q--AL-KQ-TS-QR--  103
CPZ.CM.x.SIVcpzDP943          ------.D----Q------R---N-----C-----M---EK-G-----------GR.-HPPATS-THN-DP-RQ--L-R..QQ-.-Q--QE-Q-----SQA-AH-RA-  103
CPZ.CM.x.SIVcpzMB897          -D-I--.N----N--------P-N-----M--Y---F--TR-G---Y------K...--G-TES-KD--DTIP---L--..SG-.TQ--Q--S--E--KT-TS-QCA-  101
N.CM.02.DJO0131               ------.-----N----------NG----Y--Y--MC--TK-G----------S-..---TPKS-KN--DPIPE--L--QQQP-.-Q--Q-KQ--AL-DKA----*--  103
N.CM.04.04CM_1015_04          ------.-----N----------NK----N--Y--V---TK-G----------S-..--KTPNS-KDY-DPIPE--L--..QP-.-Q--Q--Q--AL-SK--A-LC--  102
N.CM.04.04CM_1131_03          ------.-----N----------N--------Y--V---TK-G----------S-..---TPNS-KNY-DPIPE--L--..QP-.-Q--Q-XQ--AL-XK--A--C--  102
N.CM.95.YBF30                 ------.-----N----------N-----R--Y--LY--TK-G----------S-..---TP-S-KS--DLIPE--L--..QQ-.-Q--Q-KQ-EAL-SK--A--C--  102
N.CM.97.YBF106                ------.-----N----------NK-------Y--MC--TK-G----------S-..---PPKS-KD--DPIPE--L-R.QQP-.-Q--Q-KQ--AL-GK--A--C--  103
O.BE.87.ANT70                 -D----.EVP--H------QIP-N-----R--Y--Y---VR-G----------G...-P.-AASHPD-KDPVP--SPTI..TKR.KQERQE-QEEE--KKAGPGGYPR  100
O.CM.91.MVP5180               -D----.EMP--H----K-Q-P-N-----R--Y--Y---TK-G----H------...-PA-AASYPDNKDPVPE-SL-H..TGR.KQKRQE-QE----K--GPSGQPC  101
O.CM.96.96CMABB637            -D----.EVP--H------P-P-NA-------Y--YL--TK-G----H------...-PA-AASCSNNKDLVPE--STS..TNR.KQKRQE-QEE---AK-G-GGDPC  101
O.CM.98.98CMA104              -D----.EMP--H------Q-P-N-----S--Y--Y--LVR-G-----------...-P.-TASHPDNKDPVPE-SP-I..TNR.K*ECQE-QE-E--K--GPG*YHC   98
O.CM.99.99CMU4122             -D----.EMP--H------Q-P-NK-------Y--Y---TS-G-----------...-P.-AAHHPDNKDLVPE-SPTI..TNR.KQERQE-QE----K--GP-*YPC   99
O.FR.92.VAU                   -D-I--.EMP--H------Q-P-NA----Q--Y--YL--TK-G----H------...-PA-AASYPDNKDPVPE-SP-N..NKR.K*ERQE-QE-E--TK-GAGGLHC  100
O.SN.99.SEMP1299              -D----.EMP--H------Q-P-NK----A--Y--Y---AS-G-----------...-PA-AARHPDN-DIVPE-L-YI..TNR.KQKRQE-QE-E--N-ACPR*YPG  100
O.US.99.99USTWLA              -D---H.EMP--H----S-Q-P-NT----S--Y--I---TK-G-----------...-PTTAERHTDNKDPVPE-SPTI..TSR.K*ERQ--QE-E--KKAGAGG*PR   99
```

```
                          Rev start            exon 1 end exon 2 start   NLS                              Leu-rich effector domain                                          Rev end
B.FR.83.HXB2              MAGRSGD..SDEELIRTVRLIKLLYQSNPPP..NPE.GTRQARRNRRRRWRERQRQIHSISERILGTYLGRSAEPVPLQLPPLERLTLDCN.......EDCGTSGTQ.......GVGSPQILV.ESPTVLESGTKE*   116
A1.GE.99.99GEMZ011        -------..T-A--LTA--I--F------C-..T-R.------K-Q-----A-----D-------S-C--------------I---H---S.......---------QSQGTET---G---SG.K-SGI-G------   124
A1.KE.00.KER2008          -------.------LKAI-I--I------Y-..K-K.------K-------A-----D-L-----S-C---P--------------H---S.......---------.......---R--VS-.---DI-G----N-   117
A1.KE.00.KNH1144          -------.------LTA--I--I------C-..K-R.-S----K-------A-----D-------S-C----E--------S--G-R---C.......---------QSQGVEA---G--VS-.---GI------N-   124
A1.KE.00.KSM4024          -------.----G-L-AI-I--I------F-..K-K.-S----K-------A-----D-------SDC----T-------------H---H.......---------QSQGVET---R--VS-.---AI-G----N-   124
A1.KE.00.MSA4069          -------.------FKA--I--I------Y-..K-K.------K-------A-----D-------S-C---P----------I---H---H.......K--------QSQGAET---G--VSG.---DI-G----N-   124
A1.KE.00.NKU3005          ------N.------LKA--T--I------H-..K-G.------K-------A-----D---Q---S-C-------------------H---C.......---------QSQGVET---RT-VSG.---GI------N-   124
A1.RU.00.RU00051          ------N..A--G-LTA--I---------Y-..K-R.-X----K-------X-----D---X---S-CX---X-----P-------H----.......---------QSQGTET---R---S-.---GI-X------   124
A1.RW.93.93RW_024         -------.------LKA--I--I--K---Y-..K-R.-S----K-------A--K--D---K---SAC---------FL-----G-H---S.......---------QSQGVET---R--VSW.---V--G---EK-   124
A1.SE.95.SE8891           -------.------LKA--I--I------Y-..K-R.------K-------A--K--D-L-----S-C---PQ-----L---I---H---S.......---------QSQGAET---G--VSE.--SII-G----T-   124
A1.SE.95.UGSE8131         -------.------LKA--T--I------Y-..Q-K.-S----K-------A-----D-L-Q---SAC----E--------------HI--H.......--------#.........fT-VSG.--SV--D----D-   112
A1.TZ.01.A173             ------N.----D-LKA--I--I------Y-..K-K.-S----K-------A-----D-------S-C----P-------------H---C.......---------QSQGVET---G--VSG.---G--GT---K-   124
A1.UA.01.01UADN139        -------..P-A--LTA--I--I------Y-..T-R.------K-------A-----D-------S-C---------SF----I---HIN-S.......---------QSQGTET---E---SG.--SGI-G----G-   124
A1.UG.92.92UG037          ------N..P----L-AI-I--I------Y-..E-K.------K-------A-----DTL---V-S-C---P-----------I---H---S.......---------QSQGVET---RT-VSG.---V--G----N-   124
A1.UG.99.99UGA07072       -------.------LKA--I--I------Y-..K-K.-S----K-------A-----D---K---S-C---------F----I---H---S.......---------QSQGTET---R--VSG.--SNI-G----N-   124
A2.CD.97.97CDKTB48        -------..P--D---AI-I--I------Y-..K-R.------K-------A-----D-------S-C---PT-------------H---S.......-----C---QSQGAET---R--TS-.--SGI-G--IED-   124
A2.CY.94.94CY017_41       T----D-..P--S-LQAI-T--I------Y-..K-R.-S---Q--------A-----D-----V-R-C---PT-------------HI--S.......--------LQSQGTET---RS-ES-.--SVI-G---E--   124
A.SN.01.DDI579            ------N.------L-AI-I--I------Y-..K-Q.------K-------A-----D-L-----S------E-------------H---S.......---------.........GS*-S-.-CSAI-GE---KQ   112
A.SN.01.DDJ369            -------.------L-AI-I-RI--E---F-..K-R.-----Q--------A-----N-------S-C-E-PT----F------G-HI--S.......---------QSQGTET---G-S-SG.--SVI-G*---K-   123
A.SN.96.DDJ360            ------N.------L--I-I--I------Y-..K-R.------K-------A-----D-----L-S-C---PE-------------H---S.......---------.........GS*-S-.--SAI-GE---KQ   112
A.CD.02.02CD_KTB035       ------H..T----LTA--I--II-----Y-..K-S.-S----K-------A-----D-------S-C----E----F----I---H-N-S.......----A----QSQGTET---E---S-.---G--G-AAEK-   124
A.CD.97.97CD_KCC2         ------S.---A--LAAI-I--I------Y-..K-R.------K-------A--N--D-------R-C----E----FC-------H---S.......---------QSQGTET---G---S-.--SG------EN-   124
A.CD.97.97CD_KTB13        ------S.------LTA--I--I------Y-..K-R.------K-------A--N--D-L-----S-C---PE---F--------HI--S.......---------QSQGTET---R---P-.--SDI-V----N-   124

B.AR.04.04AR151516        -------.----D-LK------F---------.SS-.------X------------QTL-HW--S-H---P------------------D.......---------.......-X--S--P-.--X----------   117
B.AU.87.MBC925            -------.------LK------F------L-..S--.---------------------RE--G---S-F--GPT------------------.......K--------.......------V--.-------------   117
B.BO.99.BOL0122           -------.----D-LQ-----F---------..---.------------------H-QAV-AW-IS-H----T------------------.......---------.......---D-----.------G------   117
B.BR.03.BREPM2012         -------.------L----I-F---------..---.--------------A-----RK--DW--S-H---P-----------------S.......---------.......----T--S-.--S---D------   117
B.CA.97.CANB3FULL         -------.------LK-----F---------.SS-.----------------------RTL-----S-----T------------------.......---------.......----------.-------------   117
B.CN.05.05CNHB_hp3        ----RE-..N----LK--K---F------L-..---.---K---------A-----RE-----D----G-T----------------S.......---------.......R-------G.---A--D------   117
B.CO.01.PCM001            -------.----D--Q---F-R----L----..SR-.-------------------G---AW--RSH----------------N---S.......---A-----.......------V--.---A------E--   117
B.GB.83.CAM1              -------.------LKA----------L-..SSK.------------------H-RA--NW--S-H---P-----------------S.......K--------.......----------.---A---------   117
B.GB.86.GB8               -------.----DFLK---F--F---------..--K.--------------A------Q-G-W--SAF---P----------I-------D.......---------.......------V--.---A--DP-----   117
B.GE.03.03GEMZ010         -------.------L-----------------..S--.--------------A-----RT--DW--R-H-D--K--------S-------S.......---R-----.......----------.----I-G------   117
B.IT.05.SG1               -------.------LK--Q---Y------FR..S-T.-------------K--W-Q-L-GW--SNH---PM-------------I-TG.......-------S-.......---T--V--.----------E--   117
B.JP.05.DR6538            -------.----D-LKA-----F------Y-..SL-GT-----------------R-------C---E-P--------------S---S.......---------.......----------.---AI--------   118
B.KR.05.05CSR3            -------.---D--LK------F---------..S-Q.------------------Q---AWV-RNH----------P-----------S.......---------.......---D--V--.--S-I--------   117
B.NL.00.671_00T36         -------.------L-------Y---------.SS-.-----------------R-RA---W--SNH---P-----------------S.......---------.......----------.---AI-D----G-   117
B.RU.04.04RU128005        -------.----D-L-A----F---I-----.SQ-.------------------RTL-GW--SD----PX------------------.......---------.......---N-----.-P-P---------   117
B.TH.00.00TH_C3198        -----E-.-------K-----R-------L-..S--.--------------A-----RE------D----G-T------------------.......---------.......--------L.------DP-A-K-   117
B.UA.01.01UAKV167         -------.----D-LK------Y---------..---.--------------A-----RT--DW--SNH---P----------------S.......---------.......---G-----.---A---P-----   117
B.US.04.ES10_53           --------.NV---LT---I--F---------.SQ-.------------------RTL-GW--SN-------------I-----------.......---------.......------V--.---A--D------   117
B.US.99.PRB959_03         -------.------LQKI---RF---------.SS-.--------------A--------GW-VS-L---P-----------------S.......---------.......-E-N-----.-PS------D---   117

C.AR.01.ARG4006           -------.----A-LLA--I--T------Y-..K--.---R---------A------------I-C---PT---------I---NIN-S.......-SG------.......R--N-*-FR.KPCA--G-RA-KE   116
C.BR.04.04BR013           -------.----A-LQA--I--I------Y-..---.--------------A------------T-C----T----F----I---NIN-S.......-GS------QPQGTPE---D-*VFR.KPCAI-G-RA-KE   123
C.BW.00.00BW07621         --------.N-DT-LQA--I--I------Y-..K--.--------------A------------S-V---P-----F----I---HIG-S.......-SG------QPQGTPE-M-N-*-SG.N-CA--G-RA-KE   123
C.CN.98.YNRL9840          --------.G--A-L-A--I--I------Y-..E-R.------K-------A--KR---------S-C---P----HF----I---HIGDS.......-SG-A----QSQGTTE--A--*-SG.K-CA--G--A-KE   123
C.ET.02.02ET_288          -------.---AA-LQA--A--I------Y-..K--.-----Q--------A-------L-----SDC---HP--------------N---S.......-SG------QPQGTTE---N-*VSG.QYCA--G-RA-KE   123
C.GE.03.03GEMZ033         -------.----A-LQA--I--I------Y-..E-K.-----Q--------A------------S-C---------------I---HI--S......QSS-A----QSQGTTE-----*-FG.K-GA--G----KE   123
C.IL.99.99ET7             --------.DQ-AA-LQA-KI--I------Y-..K--.------K-------A--K---T------S-C------------------HI--S.......-SS------QPQGNTT-----*VSG.K-CA--G-RA-KE   124
C.IN.99.01IN565_10        -------.----A-LQA--I--I------Y-..E-Q.-----Q--------A-----C-------S-C---P-----F----I---HISDS.......-RS------PSQGTTE---N-*-SG.K-C---GF-A-KE   123
C.KE.00.KER2010           --------.N-D--L-V--I--I------Y-..K--.------K-------A---R---------SAC---P-----F----I---HI--S.......-GG------QSQGTTE-----*VSG.KLCA--G--A-KE   123
C.MM.99.mIDU101_3         -------.----A-L-A--I--I------Y-..E-R.-----QK-------A--K--------V-S-C---P------L---ID--HISDS.......-SG-A----QSQGTTG-----*-SG.R-CA--G--A-KE   123
C.MW.93.93MW_965          -------.----A-LQA--I--I------Y-..K--.--------------A--------G---A-C---P-----------I---HI--S.......-SS------QSQGTAE-----*VSG.K-CA--GARA-KE   123
C.SN.90.90SE_364          -------.----A-LQA--I--I------Y-..K--.------K-------A--------------DC---P-----------I---HI--S.......-SS--PD-E.........---N-*VSG.KPXA--GF-A-KE   116
C.SO.89.89SM_145          -------.----A-L-A--I--I------Y-..K--.------K-------A-------------S-C---PT---------I---HIGGS.......-SG------QSQGTTER----*VSG.K-WA--G--A-KE   123
C.TZ.02.CO178             -------.---DA--QV-KI--II-----Y-..E-K.--------------A-------------SAC-E-P--------------HI--S.......-SG------QSQGTTE---N-*VPG.RLRA--G---EKE   123
C.UY.01.TRA3011           -------.----A-L-A--I--F------Y-..K--.--------------A------------T-C----P-S-------I---NIN-S.......-SG------QPQGTTE---N-*VFR.-PCAI-GPRA-KE   123
C.YE.02.02YE511           --------.N-AA-LQ--KI--I------Y-..KL-.--------------A--K-----G---SCC-E-P-----------I---HI--S.......-SS--T---QSQGNTE----S*VSR.K-CA--GARA-KE   123
C.ZA.04.04ZASK164B1       --------.S---A-LQA--I--I------F-..E-K.------K-------A---E-N-------S-C---P-----F----I---HIGDS.......-NL------.........---N-*VSG.KPY---G-RA-KE   117
C.ZA.05.05ZAPSK240B1      -------.----A-LQA--I--I------Y-.NPN-.--------------A------Q------V-----PT---------I---HI--S.......-SS------QSQGTTE---N-*-SG.K-R-I-G-RA-KE   124
C.ZM.02.02ZM108           ------V.----A-LQAT-I-RI------Y-..K--.------K-------A------Q------S-C----------FR---I---HI--D.......-SS------QSQGTTE---N-*VSG.K-CA--G-RA-KE   123

D.CD.83.ELI               -------.----D-LKA-----F---------..S--.--------------A-----RE-A----------P--------------N-N-S.......---R-----.......---H---S-.------------E-Q   117
D.CM.01.01CM_0009BBY      ----RE-.----D-LKV-----F---N-----..S--.--------------A---HL-T-G---S--------------------N-N-S.......---------.......-------S-.------G---E--   117
D.KE.01.NKU3006           --------.N--D-L-AI----I---------..S--.--------------A--------G--IS------E---F---------N-N--.......---------.......--H-V--.---AI-D------   117
D.KR.04.04KBH8            -------.----D-LKAI----F------L-..S--.--------------A-----N--G--IS----------------I----IN--.......---------.......--S-VSG.--SAI-DT-----   117
D.TD.99.MN011             ----REE..N--DFLK-I----W--K------.ST-.--------------A--------G---S---------S-----I---N-N-S.......---EP----.......---S--S-.------------E--   117
D.TZ.01.A280              -------.------L-AI----I---------..SL-.--------------A-----D--G---S--------------------N-N-H......T--------.......----H-VS-.-P-AI-D---E--   117
D.UG.99.99UGK09259        --------.C-GA-LQA-K---I---------..SS-.--------------A----L-T-G---S-----LE-----------------S.......---------.......----H-VF-.---A--D------   117
D.YE.01.01YE386           -------.------LK-I----F---------..S--.-----------------------G---S-----E--------------N-N--.......---A-----.......----H-VP-.-PSA--D------   117
D.YE.02.02YE516           ----RE-.----D-LKAI----F---N-----..S--.--------------A---HL---G---S-L--G----------I---N-N-S.......---P-----.......-------S-.-PS----P--ED-   117
D.ZA.90.R1                -------.----N-LTAI-I--F---------..S--.--------------A-----N--G--V-F--------------------------S.......---------.......-------S-.---A---------   117

F1.AR.02.ARE933           -------.---T--LTAL-Y--I------Y-..K--.--------------A-----RA--D---SSC---PE-------------HIN-S.......---------......PGTEE-L-G---AG.--HA-----I-N-   117
F1.BE.93.VI850            -------.---T--LKA-KC--I------Y-..K--.--------------A-----RAL-D---SSC----E-------------HIN-S.......---------......QGPEE----S--SG.--HA---------   117
F1.BR.01.01BR125          -------.------LKAI-Y--I------Y-..K-R.-------------NA---H-RE-----SSC---PE-------------HIN-S.......---E-----......QGTEE---N--TSG.-PRA-----D---   117
F1.ES.x.P1146             -------SDT-T--LK---C--I------Y-..K--.--------------A------K-----RSC--GPV-------------HIN-S.......---------......QGTEK-------S-.--HA--G---Q--   119
F1.FI.93.FIN9363          -------.---T--LKA-KY--I------Y-..S-D.--------------A-----RA-----SSC---LE-------------HIN-S.......---------......QGTEE-------SG.-HH----------   117
F2.CM.02.02CM_0016BBY     -------.----A-LTA--T--I------Y-..K--.--------------A-----RE-----S-C------------------HI--S.......--S------......QGTER-------S-.--RS--G---EN-   117
F2.CM.95.MP257            --------.R----LKA--Y--I------Y-..KL-.---K----------A------Q-----S-C---LQ---R----L--K-HIN-S.......---------......QGTEK-------S-.--RA--G------   117
F2.CM.97.CM53657          -------.------LKA--Y--I------Y-..K--.---T----------A--N--RQ-----S-C---PT-------------HIN-S.......---E-----......QGTNK---N---SG.--RA--GP-N---   117

G.BE.96.DRCBL             ------S..T----LTA--I---------S-..P--.--------------A------------S-C---PE-------------H---S.......--G------QSQGTEI-------F-.--SV--G------   124
G.CM.01.01CM_4049HAN      ------A..T-G--LQA-QI--I------Y-..P--.----T-K-------A--K---T------S-C---PT----F----I---C---S.......--S------QPQGTET---R--V-W.---AA-G------   124
G.CU.99.Cu74              ------S..T----L-A--I--I------Y-..P--.------K-------A--K--------V-S-C----X--------------S---S......K-S------QSQGTET---R---S-.---G--G------   124
G.ES.99.X138              ------S..T--D-L-AI-I--I------Y-..P--.------K-------A------------SACV--PE---S------------------.......--SX-X---EXEGTET---W---XG.-CSIE-G------   124
G.GH.03.03GH175G          ------S..N----L-A--I--I------Y-..SA-.------K-------A---VRAL---F-SAC---PT------------C---S.......--S------QSQGTET---R---S-.---V--G------   124
G.KE.93.HH8793_12_1       ------S..T--D-L--I-I--I------Y-..P--.------K-------A------------SSC-----------------H--SS.......--G------QSQGTET---G---F-.--SV--G--A---   124
G.NG.01.01NGPL0669        --------.T--A-LQA--I--I------Y-..P--.----------K---A---------K---SDC---R-----F--------H---S.......--SRIT---QSPGTET---G--TS-.---A--G----N-   124
G.PT.x.PT2695             ------S..T----L-AA-I--I--E---Y-..P--.----T-K-------A-----------V-S-C----T-------------H---S.......--S--L---QPQGTEP-L-G---S-.--SV--G------   124
G.SE.93.SE6165            ------S..T----L-A-KA--I------Y-..P--.--------------A-----SA-----TA----P-------------H---S.......--S------QPQGTET---RS*-F-.--SV--GQ-----   123

H.BE.93.VI991             --------.N--G-L-AC-I-R-------Y-..E-A.-----Q--------A--------G--V-A-C--GP-------------------S.......---------......EK------TSG.---A--GT-A---   117
H.BE.93.VI997             ------A..G--Q-PQVCKI--II-----Y-..E-A.--------------A-----RA-----TDC---PP------------------.......K--------......EK-K-G---P-.--S---GT-----   117
H.CF.90.056               ------A..--T--LQVCKI--I------C-..E-T.--------------A-----RE-----TSC---PP---T----------N-S.......---------......EK-E-----SL.--S-I-GT-----   117

J.CD.97.J_97DC_KTB147     -------.----Q-LLA-----I---G--Y-..K-N.-S------------A--N--D-------SAC---P-----F----I---R---S.......----D----......TQ---D---SG.--GVL-GA-----   117
J.SE.93.SE7887            -------.---DQ-LLA-----I------Y-..K-N.-S------------A--N--D------PSSC---P-----------I---R---S.......----N----......TQ---D---SG.-PCM--GA-----   117

K.CD.97.EQTB11C           ----R--..-EQQ-LTP--I--I------Y-..K--.--------------A-----RE--Q-V-SSC----T------------S-N-D.......--S------......QGTEGEL-----P-.-PD---G--D---   117
K.CM.96.MP535             ----R--..P--Q-LT---T--I--E---Y-..-L-.-S--T---------A--K--S-----L-SAC-------------I-K-N-N-D.......--P------......KGTEG-L-----S-.-PC----------   117
```

```
                              Rev start          exon 1 end exon 2 start   NLS                                Leu-rich effector domain                                    Rev end
B.FR.83.HXB2                  MAGRSGD..SDEELIRTVRLIKLLYQSNPPP..NPE.GTRQARRNRRRRWRERQRQIHSISERILGTYLGRSAEPVPLQLPPLERLTLDCN.......EDCGTSGTQ.......GVGSPQILV.ESPTVLESGTKE*   116
01_AE.CF.90.90CF11697         ------N..T--D-LQA--I--I------Y-..P--.------K-------R-------L-----VACV---T------------HI--S.......---------QSQGTET---G---SG.--SVI-G----N-   124
01_AE.CN.05.FJ051             --------..N----L-AA-T--I------Y-..SS-.---------------A-----RE-----SSCV--CT------------H---S.......---------QSQGTET---R---FG.K-SV--G----N-   124
01_AE.CN.06.FJ054             ------S..T----L-A--I--I--E---F-..SS-.------K-------A-----RA-----S-CM----------------HIE--.......-G----A-QSQGSET---R---SG.-PSDI--P---N-   124
01_AE.HK.x.HK001              ------S..T----LGA--A--I------S-..PA-.----T-K-------A-----RK-----SSVV---T--------S---H---.......---------QSQGTET---R---SG.--SVI-GA---N-   124
01_AE.TH.01.01TH_R2184        ------S..I----L-A--I--I------Y-..SS-.----T-K-------A-----GE-----S-C----T------------H---S.......--S------QSQGTET---R---SG.--SVI-GQ---N-   124
01_AE.TH.02.OUR769I           ------S..T----L-A-KI-RI------F-..SS-.------K-------A-----RA-----SAC----T------------H---S.......---------QSQGTET---K---SG.--SGI-GP---N-   124
01_AE.TH.90.CM240             ------S..T----L-A--I--I------F-..SS-.----T---------A-----SA-----S-C----T---------V---N---S.......--G------QSQGTET---R---SG.--SVI-GP---N-   124
01_AE.US.00.00US_MSC1164      ------S..T----L-A--A-RI--E---F-..SS-.---------------A-----RA-----SAC----T------------H---S.......---------QSQGTET---R---SG.--SVI-GP--EN-   124
02_AG.CM.02.02CM_4082STN      --------..A----L-AI-I--V--------..P-G.------K-------A----VRE------SSC---P----------I---C---S.......--R------QSQGTET-------SG.--SII-G------   124
02_AG.EC.x.ECU41              --------..A----L-A--I--I------Y-..P--.------K-------A-----RA-----S-C---P-----L---I---C---S.......-N---P---QSQGTET-------SG.-CSII-G------   124
02_AG.GH.03.GHNJ196           --------..P----L----I--I-H----Y-..P--.-S--T-K-------AS-S--RA----L-SSC---PK---------I-------.......---------QPQEPAT-------S-.--SNI-G------   117
02_AG.NG.01.PL0710            --------..A----L-AI-I--I------Y-..P-G.------K-------A-----GAL-----S-C---P----SF----I---S---S.......---------QSQGTKT------VS-.-H-VI-G------   124
02_AG.NG.x.IBNG               --------..A----L-A--I--I------Y-..P--.----T-K-------A-----RA-----S-C---------------I---N---S.......---------LSQGTET-------S-.--YII-G------   124
02_AG.SE.94.SE7812            ------N..A----L-A--T--I------Y-..P--.------K-------A-----RA-----S-C---------------I---R---S.......---------.......-------S-.--SI--G------   117
02_AG.SN.98.MP1211            --------..A--G-L-A--I-RI------Y-..P--.-S-----------A----VRA-----S-C---P-----P---I---C---S.......--S------QSQGTET----S-TS-.--SVI-G------   124
03_AB.RU.97.KAL153_2          --------.------LK-I----F---------..S--.-----------------H-----Q--S-----PE---L-H----------S.......---------.......---------.------D------   117
04_cpx.CY.94.CY032            ------N..I--D-FKAA-A--I------Y-.N--T.-------------A--K---L----A-----P-----------K---N-S.......---------.......DK------VS-.-L-A--GT-A---   118
05_DF.BE.x.VI1310             --------..R--D-LKA-----I------L-..S--.---------------A-----N--G--L-S------E------------N-N-S.......---------.......-------S-.-P-AI--------   117
06_cpx.AU.96.BFP90            --------.--DR-LLA--I--I------Y-..K-N.-S------------A--N--D-----V-S-C-------------I---R---S.......----N----.......---N---SG.-PDML-GT--T--   117
08_BC.CN.97.97CNGX_6F         --------.---A-L-A--I--I------Y-..E-R.------K-------A--N-L---------SAC---PT------------HITDS.......-RG------QPQGTTE-----*-SG.K-CA--G--A-KE   123
09_cpx.GH.96.96GH2911         --------..N--Q-LAA--V--I------Y-..K--.---------------A--S--N-------RAC---PE------------H-N-S.......----N----.......--AD---SG.--CAI-G------   117
10_CD.TZ.96.96TZ_BF061        --------.---A-LQA--I--I---------..S--.-----------------------GG---DA----PE------------N-N--.......---------H.......---GH--S-.----I-D------   117
11_cpx.GR.x.GR17              --------.--AQ-LAAA-I--I----S-Y-..K-A.--------Q-----A--N-L----Q---S-C----E-----P---I---H---S.......----N----.......---ES--PG.--S---GPR-E--   117
12_BF.AR.99.ARMA159           --------.-----LKAA-----------Y-..K--.---------------A--S--RA-----ID-----PE----F---------G--.......---------.......---------.-------------   117
13_cpx.CM.96.1849             --------.---V-LT---I--I---N--Y-..SQ-.------K-------A--------GQ---S-C----------------H---.......---------QSQGTET---G---P-.---GI-G----N-   124
14_BG.DE.01.9196_01           ------S..A--D-LKA--I--I------Y-..P--.------K-------A-------------SACV--XE------------H---S.......--S--L--RQSQGTET---G---S-.--SV--G------   124
15_01B.TH.99.99TH_MU2079      ------S..T----L-A--I--I-H----Y-..SA-.-S--T---------A--KR-RE------SPC-----------------S---.......---------QSQGTET---R---SG.--SV--GP---N-   124
16_A2D.KR.97.97KR004          --------..P--D-L-A--A-RI------S-..D-R.-S----K-------A-----RA------SAC---PT------------H---S.......---------QPQGTET---R---S-.-PSV--G---E--   124
17_BF.PE.02.PE02_PCR0155      --------.---D-LKA-----A------S-..K--.---------------A-----RE-----S-V---P-----L---I-------.......---------.......---N-----.----I--------   117
18_cpx.CU.99.CU76             --------.---Q-LT--SI--I------Y-..E-A.---------------A-----SA------N-C--GP-----------------S.......-----F---QSQGTET---GS-VP-.--SG--GT-D---   124
19_cpx.CU.99.CU7              --------.---D-L-VI-I--I------Y-..TQR.#f----K-------A-------------S-C----E----F--------H-N-S.......---------.......R--G---S-.--SV--G--A-N-   115
20_BG.CU.99.Cu103             ------S..T--D-L-VA-I--I------Y-..P--.---------------A-----RAL-----T-C---PP----------------S.......--S------QSQGTET---R--VS-.---G--G------   124
21_A2D.KE.91.KNH1254          ------A..---D-LKV---L-SI---------..S--.---------------A-------------S-----PE------------N-NSS.......---------.......-T-G---P-.---A------E--   117
22_01A1.CM.01.01CM_0001BBY    ------S..T----L-A--I--I------Y-..K-R.------K-------A--K--D-L-----SAC-----------------H---S.......---------QSQGTET---R---S-.---DI-GA---K-   124
23_BG.CU.03.CB118             ------S..T----L----I--F------Y-..P--.---------------A---------Q---ADC---PP----------------S.......--I------QSQGTETR--G---S-.---A--G------   124
24_BG.CU.03.CB378             ------S..T----LLAI-I--I------Y-..P--.---------------A-----------V-T-C---P-----L-----------S.......-SS------QSQGTET---GS--S-.--SV--G------   124
24_BG.CU.03.CB471             ------G..T----LQA--I-RI------Y-..P--.---------------A--S--R------T-C---PT----F------------S.......-GS------HSQGTET---E---S-.--SV--G------   124
25_cpx.CM.01.101BA            --------..N----L-A--T-NI--R---Y-..P--.---------------A-----RQ-----TSC---LE------------HI--S.......--S------QSQGTET---R---SG.--SV--G------   124
27_cpx.CD.97.97CDKTB49        -----EG..N--R-LTA--I--I------Y-..K-R.-SQ---K-------A-----D--G----S-C---PT----F----I---C--.......---------QSQGTET---G--TS-.-H-AI-GT-A---   117
28_BF.BR.99.BREPM12609        --------.---D--K-I-I--F---------..S--.----------------------L-RW--SNH---PE----F----I-------.......---------.......------V--.---A--G------   117
29_BF.BR.02.BREPM119          --------.---D-LG----------------..SR-.-------------------N---TW--S-R---PE---H--------------.......-------P-.......---------.-P----------   117
31_BC.BR.02.110PA             ----R--.---A-LQX--I--I------Y-..X--.--------------G-----------S-C----T----F----I---NIN-S.......-SG------.......---N-*VFE.KPSA--G-RA-KE   116
32_06A1.EE.01.EE0369          --------..N--Q-LLA--I--I-N----Y-..K-N.-S------------A--H--D---H---S-C---P----------I---R---S.......----N----.......---D---PE.KPGVL-GT-----   117
33_01B.MY.05.05MYKL007_1      ------STST----LKA--I--I------Y-..SA-.----T-K-------A----VCAL-D---SAC----T-----H---I---N---S.......--------Q....GTET---R---SG.-PSVI-GT---N-   122
34_01B.TH.99.OUR2478P         ------S..T----L-A--T--I------Y-..SS-.------K-------A--K--RE--D---ISC----T-----------------S.......K--------.......------V--.---A--D------   117
35_AD.AF.05.05AF095           --------.-----L-A--I--I------H-..K-R.-S----K-------A--N--D---Q----AC---P-------------H---R.......---------QSQGVET---R--VS-.-P-V--G----N-   124
36_cpx.CM.00.00CMNYU830       --------..A----L-V--T--I------Y-..PS-GR-----K-Q-----A-----RA------S-C---FE------------H---S.......---------QPQGTEA---G--V--.---GI-GT---N-   125
37_cpx.CM.00.00CMNYU926       --------.---Q-L-A--I-NI------Y-..P--.------K-------A-----RAL-----S-C------------------H-N-S.......----RD----QPQGTET-L-G---S-.-P-G--G----N-   124
38_BF1.UY.05.UY05_4752        --------.---K-LQAAQY--I------R-..KL-.---------------A-----RE------DSC---P-----P-------HIN-S.......---E.......QRTEQ---N----G.---AI--------   117
39_BF.BR.03.03BRRJ103         --------.--Q--LKAISY--I------Y-..K-K.---------------A-----CQ------RSC---LE------------HIN-S.......-N--.......QGADER-----TSG.-H-AI-G------   117
40_BF.BR.04.04BRRJ115         --------.---D-LKA---V-Y---------..ST-.-------------------RK--DW--S-H---PT----F------------S.......--S------.......---T--VF-.-P-A------E--   117
42_BF.LU.03.luBF_05_03        --------..R--T-LQ------R---------..S--.-------X-----------QTX-GW--N-----PE-----H-----------S.......---------.......------V--.---A---T-----   117
43_02G.SA.03.J11223           --------..A----L-X--I--I------Y-..S-A.R-------------A-------------SAC---R-------------H---S.......--SR-P---QSPGTET---G---P-.---V---------   124
U.CD.83.83CD003               ------E..-----L-A--I--I---------..---.------K-------R--Q--R------S-C---P----H---------N---S.......K..........GTAT----T--PG.--CA--G------   113
U.CD.90.90CD121E12            ------E..-----L-AI-I--I------Y-..---.---------------K--Y--R------SAC---PT------------H---S.......-..........GTET----T--SG.--CA-------K-   113
U.GR.99.GR303                 --------.---Q-LT-I-I--I------Y-..K--.---------------A-------------L-C----E--------SI--H---S.......----N--D......#R--AX---SGf-PFCYMG---A-E   117
U.NL.95.U_NL_95_H10986_D1     ----R--..Q--Q-LT---I--I------F-..KL-.-----------K---A-------------S-C-----------------S.......--SR......EGTEGEL-----P-.-PCA---------   117
CPZ.CD.90.ANT                 ----EELEGT--Q-LKA-KI--I------Y-..K-A.-S-A----K----KK--D-VEGLAA-V-R-LV-GPD-HN-VD--D-QN-S-EPL.......GSPP-PLQNYFQTVQPSGEQS-HPL.DNN.......QN-   119
CPZ.CM.01.SIVcpzCAM13         -----EH..D-AR-LQA-KI--I--E---Y-..S--.---N---------A-----RE--N-V-AAH---PPQ-GH-E--E-DK-S-H-L.......-TV-KV--TVNQHPQTDT-ETDSTE.GDNPI-GKRI-N-   124
CPZ.CM.05.SIVcpzEK505         ------V..N----L-AI-V--I------Y-..-DS.------K-------A-----RAL-----SSCV-GLQ--ST-P----D--S-NPE.......--L---E-EHPQGTATT**TQN-F-.GNNSL-GK-A-N-   122
CPZ.CM.05.SIVcpzLB7           --------..R----L-A--I--I------F-..H--.---S---------A-H---SE-AG---S-C-E-PEG----P---ID--S-NSE.......--I-DC--IHSEGTET-----T-SA.GNS---GK-A-N-   124
CPZ.CM.05.SIVcpzMB66          ---G--N..-----L-A--I--I------Y-..S--.-S------------A-----N-L-----S-C-E-PTG--S-P---I------SA.......--I--G--DPPQGTET-T---NTPE.GHS-I-GT-A-N-   124
CPZ.CM.05.SIVcpzMT145         -----EG...--N-L-AI-I-RI--E---Y-..SQS.---R-----K----------NQ-G--V-ASC----S--PQ-P---I---CISDL.......-TT-DT--IQPQGSTTNRY--TPTE.RD-A--DNRNQ--   123
CPZ.GA.88.GAB1                -----EP.QD-AR-LQA-KI--I------Y-..S--.---K---------A--K--SE--G-V-A-----PPK-GD-E--E-DK-S-Q-V.......-TTQDV--SNTSQPQTAT-ETVPAG.GNYSI-GK-A-N-   125
CPZ.TZ.00.TAN1                ----EE....-AN-LY---I--I--D---Y-..SGA.-S-T---------Q--H-VDALAS---QYR--GPQ--PH-DI-D-SK-H--PLDQPASTETGDNQL---PSNSA*NYHSRS*TDH.*VEQ*YLC*L-SF   124
CPZ.US.85.CPZUS               ------G.DA--Q-L-A--I--I--A---F-..-NS.------------KN-----NE--Q---S-C-----S--PD-P---IGG--INSE.......--STNP--EHQQGTATVE*TQST-L.GGNRL-GK-A-S-   124
CPZ.CM.x.SIVcpzDP943          -----EG.DA-TR-LQV--I--I--D---Y-..DNK.---------------A---H--AL------SC----P--PT-P---ID----NPE.......--I-N---EHQQRTAVAE-HQ-TP-.G-DSI-GK-ATN-   125
CPZ.CM.x.SIVcpzMB897          ------S...-----L-A--I--I------F-..S--.-P------------K-----SAL-----S-C---LE----------------SE.......--A-DC--EVPXRTET---N-N--E.-HSP--GTRAEN-   123
N.CM.02.DJO0131               ------V..N----L-A--I--I------Y-NN-NQ.-------------T-----RE--Q---D-C--GLQ---D-P-----D-----PA.......--S-AP--EPHQGTATTE*TQNT--.GNNSI-GKRV-G-   125
N.CM.04.04CM_1015_04          ------V..N--K-LTA--I--I------Y-..-NQ.---------------A-----CA------S-C--GLQ--GDPP----------NT.......--S-DP--EPQQGTATIE*TQKT--.GNTSI-GKRV-R-   123
N.CM.04.04CM_1131_03          ------V..N----LTA--I--I------Y-..-NQ.------X------XA-----RA--X---S-C--GXQ--GDPP-X--X-----XT.......--S-AP--ESQQGTATTE*TQXTX-.GXTSI-GKRV-R-   123
N.CM.95.YBF30                 ------V..N----L-A--V--I------Y-..-SK.---------------A-----RA------SSC--GPP---D-P----D-----TE.......--S--P--ESQQGTATTE*TQNT--.GNTCI-GKRV-G-   123
N.CM.97.YBF106                ------V..N--D-LKA-KI--I------Y-.DSSQ.---------------A-----RA-------A---GPQ---D-P----G-----HK.......--S-DP--ESQQGTATTE*TQNT--.GNNSI-GKRV-R-   124
O.BE.87.ANT70                 -----E....-DQ-LQAIQI--I------Q-..S-R.-S-N--K-------R--A-VDTLAA-V-A-VVHGPQNNNIVD-----Q-SIRDP.......-GDQL-EAWTVDPRAEDN*CL-N-C.SCN-I-ATRIA--   121
O.CM.91.MVP5180               -----E...E-QQ-LQAIQI--I------C-..T-A.-S-N--K-------R--A-VD-LAT---A-VVHG-QDNNLVD-----Q-NIRDP.......-ADRLP--GTVDPGTKDN*-L*T-W.SCNAI-ATRIEK-   121
O.CM.96.96CMABB637            ------...D-QQ-LQA--I--I------Q-..A-T.-S-N----------R--A-VETLAA-V-A-VVH--QDDNLVE-----Q-SIRDP.......-RDQLP--WTGDPRAEDN*-L-TNR.SCX-I-ATRIA-K   122
O.CM.98.98CMA104              -----D....-DQ-LQAI-I--I------H-..S-I.-S-S----------R--A-VDT-AA---A-VVHGPQNNNLVD-----Q-SIRDP.......-SDQP--NWTVDPRAEDN*WL-T-*.SCN-I-ATRIA--   120
O.CM.99.99CMU4122             -----D....-VP-LH-I-I--I------Q-..S-I.-S-S----------R--A--DTLAA-F-ASVVHGPQDNNLVE-----Q-SIRDP.......DGDQP---GTVDPRTEDN*CL-N-C.SCS-I-ATRIA--   121
O.FR.92.VAU                   -----E...E-QQ-LQAIQI--I------H-..TTS.-S-N----------R--A-VD--AA-V-ASVVHGPQDNNLVD-----Q-S-RDP.......-VDQ-P--WTVDHRAEDN*-L-T-*.SYN-I-ATRIAN-   121
O.SN.99.SEMP1299              -----D...G-QP-L-AIQI--I-----SHT..S-T.-S-S----------T-HAHVDTLAA---A-VVHGPQDNNLVE-----Q-SIRDP.......DGDQP---WTVDSGTEDN*CL-T-H.SCN-I-ATRVA--   122
O.US.99.99USTWLA              ------...E-QQ-LNAIQI--I------Q-..S-A.-S-N--K-------R--A-VD-LAT---A-IVHGPPDNN-VE--D--Q-NIRDP.......-VDQPA--WTVHSMAKDN*-L-T-C.SCN-I-ATRIEN-   122
```

phos phos
transmembrane domain — cytoplasmic domain
Vpu start — α-helix — α-helix — Vpu end

```
B.FR.83.HXB2          MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHYESPHPRISSEVHIPLGD..ARLVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPELADQLIHLYYFDCFSDSAIRKALLGHIVSPRCEYQAGHN.KVGSLQYLALAALITPK.....KIKPP 162
A1.GE.99.99GEMZ011    ---------------------N----------K---E---------R--K-----------..----VK------A--K-----H-------Q------I--D--------H------E------IV------K-N-PP--S.----------K--V--T.....RER-- 162
A1.KE.00.KER2008      ---------------------N--------I-K--K----------R--KV----------..-K--VR------P--K-----H-------L------------------H--N---E------I--QV---S--------.----------K--V---.....-T--- 162
A1.KE.00.KNH1144      ---K-----------------N----------K--KN-V----F--RN--V----------..-K--VR------P--N-----H-------L-S-------D---------------G------I--EL--------T--S.----------K--V--Q.....-T--- 162
A1.KE.00.KSM4024      ---------------------N----------K--KN------F--R--KV----------..-K--VI---------K-------------RG-----I--D--------H--A--T-------I--QV-R-----PT---.----------K--V--T.....-TR-- 162
A1.KE.00.MSA4069      ---------------------N----------K--SK-W-------R--KA-------V-E..-----R-----N---KA-Q--H-------LRK-------DQ--------------E------I--QV-----D--T---.----------K--V---.....-T--- 162
A1.KE.00.NKU3005      ---------------------N-------H--K--K----------R---V----------..--I--R---------K--Q--H-------L------LT-DM-------H--N----------I--QV--------T---.----------K--V--P.....R-R-- 162
A1.RU.00.RU00051      ---------------------N----------------L-------R---V----------..----VR---------K--Q--H-------Q--X---I--D-------------------------E--R----------.------X---K--V--T.....RGR-- 162
A1.RW.93.93RW_024     ---------------------N----------K-TKN---------R---V----------..---IVR---------K-----H-------L------I--D--------H------E----R-I--QV---S---HT---.Q---------K--V--T.....-TR-- 162
A1.SE.95.SE8891       --------V------------N--------I-K--K--L----F--R--KV---------E..----VR-----Q---K-----H-------LR-----I--D--------H--K----------I--E-------------.----------K--V---.....--R-- 162
A1.SE.95.UGSE8131     ---------------------N--------I-K--A----------R--KV----------..---IVR------P-----Q--H-------LR-----I--D--------H--N----------I--RV---S----T---.----------K--V---.....RT--- 162
A1.TZ.01.A173         ---------------------N----------R--KD---------R--KV----------..----VR-----Q---K-----H-------L------I--D--------H-------------I--QV-R---D------.----------K--V---.....----- 162
A1.UA.01.01UADN139    ------------------X--N----------K---K-V-K-----IN-KV----------..----VK---------K-----H-------Q------I--D---------------E------IV-Q------N-P----.---P------K--V--I.....RER-- 162
A1.UG.92.92UG037      ---------------------N--------I-RR-K----------R--KV-------I--..--I-VR-----Q---K-----H-------L------I--D--------H--N----------I--QV-----D--T---.----------K--V--S.....R---- 162
A1.UG.99.99UGA07072   ---------------------N----------K--K-------F--R--KV----------..----V------Q---K--Q--H-------L------I--D--------H--N--------R-I--QV---S--------.----------R--V--T.....RTR-- 162
A2.CD.97.97CDKTB48    ---------------------N----------K---E---------R---V--------KE..---IVR------P--K-----H-------QG-----I--D---H-----------E----R-I--E--R----------.----------R--VAST.....RT--- 162
A2.CY.94.94CY017_41   ---------------------N--------I-R--K--V-K-----RN------------E..--IIVR------I--K-----H-------QN--H--I--D---H-----------E------II-E-------------.----------K-VVAST.....RT--- 162
A.SN.01.DDI579        ---------------------N--------I-K--KD-------DCR---A----------..-M--VR------P--Q-----H-------Q------I--D--------H-----------R-I--Q--R------T---.----------R--VAS-.....-TR-- 162
A.SN.01.DDJ369        ------------------K--N----------R--KK-V-K-----R--KV----Y----E..-I--VR------P-----Q--H-------Q------I--D--------H--N-------------Q--R-S--------.----------K--V--E.....RT--- 162
A.SN.96.DDJ360        ---------------------N--------I-K--KN---------R---V----------..-I--VR------P--Q-----H-------Q------I--D--------H--N----------I--Q--R-----PT---.----------R--V---.....----- 162
A.CD.02.02CD_KTB035   --------V---------K--N--------I-K--K--V----F--R--KA----------..---IVR-----Q---------H------QL------I--D--------H--E----------I--N------D------.----------K--VA--.....GT--- 162
A.CD.97.97CD_KCC2     --------------------AN-------H--K--I----------R--K-----------..----VK----------G----H------KLR-----I--D--------H--T---E------I---R---K--------.----------K--V--T.....RA--- 162
A.CD.97.97CD_KTB13    ---------------------N--------I-K--K----------RP-K-----------..----VR-----QP-----Y--H------KMR-----I--D--------H-------------I----------------.----------K--VQ--.....RR--- 162
B.AR.04.04AR151516    ------------------S-----------T----KR-I-------TN-------------..---I------------E-------------EK-------DI---------------------------R-S--------.-------------VK--.....RKR-- 162
B.AU.87.MBC925        -----------------------------------K--C-------TN------------E..-K---V-------------------------------G-D---------------E----------R------------.-----------------.....----- 162
B.BO.99.BOL0122       ------------------K-----------I-K--K--C------NL-----G-------P..-K----------------------------E--A---T-DI-------------A-------I--R----S-DFR----.-------------V---.....-T--- 162
B.BR.03.BREPM2012     ------------------------------I----KE---------T--------------..------------------------------R--------D-----------N---E----N-I--R----K--------.R---------------R.....RV--- 162
B.CA.97.CANB3FULL     ------------------K---------K-IT---KE-I-K----N---------------..-KV-V-------------------------E--------D---------------E----N----R----S--------.-----------------.....Q---- 162
B.CN.05.05CNHB_hp3    -A----------------------------I-R--K--C-K---D-T--K-----------..-K---------------------------RD-----AE-S---------------E------I--RV--SS-N-PE--S.-------------V---.....RR--- 162
B.CO.01.PCM001        ------------------------------I-K--K----------T--K----------E..N---V-------------------------R--------D---------------E----N-I--NL-R----------.--E-------T------.....----- 162
B.GB.83.CAM1          ------------------------------I----KK-S-------T-------------E..----V------------------------T-G-N-----D---------------E------IV-RL------------.----------T---A--.....----- 162
B.GB.86.GB8           ------------------------------I----KK-V-K----NT-------------E..------------------------------R--R-----G-------Q-------E----N----RT---S-K------.----------T------.....----- 162
B.GE.03.03GEMZ010     ------------------K--H----Y---I-K---D-R-------T--K-----------..-E--VI------S--------------------------DS--------------E----R-II-RL------------.----------T--VK--.....-K--- 162
B.IT.05.SG1           ---------------------H---I--K-M-K--KN-V------N-----G---------..GK-------------Q--Q--H--------E--------D---R-----------E----N-I--RR---L--------.----------------R.....RR--- 162
B.JP.05.DR6538        --------------------------Y-----R-TER-V-------T--K-----------..DK----------------------------R--------N-----------E---EA---HTI--R------------T.--------------K--.....----- 162
B.KR.05.05CSR3        ------------------------------I-R--K--V------NI---V----------..-K--V-------------------------G--N-----T-------R--------------I-------S--------.----------T------.....-T--- 162
B.NL.00.671_00T36     --------------------------Y---R----KK---------T-------------E..-K---------------------------ER--T-----D-------V-------E----N-I--RL------------.-----------V-----.....----- 162
B.RU.04.04RU128005    ------------------K--N-I------I-R--KK-V-------TN-K-----------..-K---K---------Q-----H--------RG-N-----D---------------EXX--N-II----------PT---.------------IT--N.....-R--- 162
B.TH.00.00TH_C3198    ------------------K-----------I-R--K--I-K---D-S--K-----------..-T-----------------------------N-N-----G---------------E----N-IV-RV---S-D------.-----------V-----.....-R--- 162
B.UA.01.01UAKV167     ---------------------N--------I-----K-C-K-----TN--T----------..-K--VI----------E--------------K----I--D---------------G----S-I--RL---S----T--S.-------------VK--.....RQ--- 162
B.US.04.ES10_53       ------------------S-------Y-----K----------------K----------E..------------------------------R------------------------E----N-I--Q-------------.-----------------.....RR--- 162
B.US.99.PRB959_03     -------------------A------------K--K----------R--------------..-K---A------------------------R--------N-------M-------E----H-I---R---S--------.----------T--V---.....-T--- 162
C.AR.01.ARG4006       -------L-----------A-N--------I-K-----Y-------R--KV----------..-----I-----Q---------H-------LXK-------G-------MH-----T-----------R------------.----------T--LK--.....-R--- 162
C.BR.04.04BR013       -------L--------K----N--------I-KR-K--Y-------R--K----------E..-K--------XQ----N----H-------LR--------D-------TH-----A-----------R-----D------.----------T---K--.....--R-- 162
C.BW.00.00BW07621     -------L--------K----N----------KR-N----------R--KL----------..-----K---------------H------KLR--------G-------MH-----A----------QV-I---D------.----------T---K--.....-R--- 162
C.CN.98.YNRL9840      -------L--R-----K----N--------I-RR-N--L-----D-R--KV---------E..-----K-----Q---------H-------LR--N--IE-G--------------A-------I-----I---D------N----------T---K--.....----- 163
C.ET.02.02ET_288      -------L--------K--A-N--------I-RR-N--L-K-----RN--T---------E..---IVK-----Q---------H-------LRG-T-----C-------MH-----AE------I--QR-----D------.----------T--LK--.....-R--- 162
C.GE.03.03GEMZ033     -------L-----------K-N----------KR-S-------F--R--KT----------..-----K-----Q----E----H-------LRK-------G-------MH-----A-------I-----I---D------.----------T---K--.....QR--- 162
C.IL.99.99ET7         ------GL--------K-K--N----------KR-S----------R--KV-A--------..-----K---------------H-------LR--------G-------MH-----AK------I-----I---D------.----------T---K--.....-R--- 162
C.IN.99.01IN565_10    -------L--------K----N--------I-RR-K--------D-G--KV----------..-----K-----Q---------H-------LRS-----E-G-------MH-----A-------I-----I---D------.----------T---K--.....----- 162
C.KE.00.KER2010       -------L-------------N-------HI-RR-S-------F--R--KV---------E..-----K-----Q---------H------KL---N-----G-------MH-----A-------I-----I----------.----------T---K-N.....----- 162
C.MM.99.mIDU101_3     -------L--------K----N--------I-RR-K--------D-R--KV---------E..---I-K-----Q----E----H-------LR------E-G--------------A-------I-----I---D------N----------T---K--.....----- 163
C.MW.93.93MW_965      -------L-------------N----------RR-D----------R--KV----------..-----K-----Q---------H-------LR--------G-------MH-----A-------I---------D------.----------T---K--.....----- 162
C.SN.90.90SE_364      ------GL--------K----N--------T-RR-K----------R--KV----------..-----K---------------H-------LR--------G-------MH-----T-------I-----I----------.----------T---K--.....----- 162
C.SO.89.89SM_145      -------L-------------N--------I-KR-S----------R--KV---------E..-----K-----Q---------H-------LR--------G-------MH-----A-------I-----I---D------.----------T---K--.....----- 162
C.TZ.02.CO178         -------L-------------N----------RR-S----------R--KV---------G..-----------Q---------H-------LR--------G-------MH-----A-------I-----I---D-P----.----------T---K--.....----- 162
C.UY.01.TRA3011       -------L----I---K----N----------KR-S--Y-------R--K----------E..-K---------Q----E----H-------LR-----I--G-------MH-----A-----E-I---R-----D------.Q---------T---R--.....----- 162
C.YE.02.02YE511       ------GL-------------N--------I-RR-S----------R--KV---------E..-----K---------------H-------LR--------G-------MH-----T-------I-----I---D-----T.----------T---K--.....----- 162
C.ZA.04.04ZASK164B1   ------AL-------------N--------T-KR-K--Y-------R--KV---------E..-----------Q----E----H-------LR--------G-------M------A-------I-----I---D------.----------TT--K--.....R---- 162
C.ZA.05.05ZAPSK240B1  -------L--------K----N-------HI-KR-K----------RN-KV----------..-----K-----Q---------H-------LR--------G-------MH-----A-----Q-I---------D------.----------T---K--.....RT--- 162
C.ZM.02.02ZM108       -------L--------K--A-N--------I-KR-N----K-----RN-K-----------..---------------------H-------LR--------G-------MH-----A-------I---------D------.--------T-T---K--.....-R--- 162
D.CD.83.ELI           ------------------K-------------K--NR------------K----------E..-----K----------E-------------R--------G-------M-------E------I--D-------------.----------T---A--.....Q---- 162
D.CM.01.01CM_0009BBY  ------------------N-------Y--H--K--K-----Q-------K-----------..-K--V----------------------------------G-------M------A-------I---R-----D------.----------T---E--.....----- 162
D.KE.01.NKU3006       ---------------------N----Y-----K--Y--L------C-N-K----------E..----VK----------E-------------R--------G-------I------AE-------------T--N------.----------T---K--.....R---- 162
D.KR.04.04KBH8        ------------------N-------Y---I-K-----V-----D-N--------------..-K--V------N----K------------L---R--I--G-------MH-------------I---R-----D------.----------T------.....R---- 162
D.TD.99.MN011         --------V---------N-------Y-----KE-K--------D----K----------E..----VK-----N------------------R--------G-------TH-------------I--QR-----D------.----------T---K-R.....-R--- 162
D.TZ.01.A280          --------------------------------R--Q--L-----D-H--K-----------..----V-----------E-------------R--------G-------M------T-------I---------D------.----------I---K--.....----- 162
D.UG.99.99UGK09259    -------------------A--------IHI-K--K--L-----DC---K----------E..----VR----------E-------------R--------G--------------AE------I--Q--T---N------.----------T---AQ-.....----- 162
D.YE.01.01YE386       ---------------------N----------K--Q--L-----DCL--K-----------..----VK------------------------R--------G-------I------AE------I---V-T---N------.----------T---K--.....GT--- 162
D.YE.02.02YE516       ------------------N-----------I-XE-K--------D----K----------E..-K--VI-----N-----------------QR--------S-------T-------E------I--LR-----D------.----------T---A-R.....RR--- 162
D.ZA.90.R1            ------------------N-------Y--H--K--K--------D----K----------E..E---VK----------K----P--------R--I-----G-------I-------E------I---R------------N----------K---AS-.....-T--- 163
F1.AR.02.ARE933       ------------------N-------Y--H--K--KK------F--R---V--------EE..-K---I-----------------XGN-R--I--G-------I-S-----E-----GI---RI----N------.RL--------T---A-Q.....-T--- 162
F1.BE.93.VI850        ------L-----------N-------Y-----K--K--S----FQ-R---V--------EE..VK----------P---E------------QGK-R--I--G-------I-------E------I---RI----N------.----------T---A-E.....-T--- 162
F1.BR.01.01BR125      ------------------N-------Y--H--K--KR-L----F--R---V--------EE..-K---------------------------QG--R--I--G-------I-------E------I---KI----N------.Q---------T---A--.....----- 162
F1.ES.x.P1146         --------V---------S-------Y--HI-K--KK-C----F--RN--V--------KE..-KI-V-----------E------------QGK----I--D-------I-------E------I--YR-----N------.--K-------T---A--.....-T--- 162
F1.FI.93.FIN9363      -------T----------N-------Y--H--K--KR------F--R--K---------EE..-E---------N-----------------QG--R--I--G-------I-------E------I---RI----D------.----------T--VS--.....-A--- 162
F2.CM.02.02CM_0016BBY ---------------------N----------RR-Q----------R---V--------EN.DSK---V---------K-----H-------Q---R-----G-------K-------E------I---R-----N------.----------T--V---.....-V--- 163
F2.CM.95.MP257        ----------------K----N--------I-KR-A----------R---V--------EE.DSK---I-----------------------Q---R-----G--------H-------------I--QR-----N------.----------T------.....----- 163
F2.CM.97.CM53657      ---------------------N----------RR-K----------R---V--------K-.DSK---VS----------------------Q---R-----G---------------E------I--QR-----N------.----------T------.....----- 163
G.BE.96.DRCBL         ---------------------N--------I-K--K----------R--KV----------..-K--VR--------DKE----H-------QG--------D---H----H--N---E------I--QT-R-S---P----.----------KV-VA-T.....RRR-- 162
G.CM.01.01CM_4049HAN  -------V-------------N------KHI-K-T-H---------R--K-----------..-M-IVR---------K-----H-------QR-----I--D---H-----------E------I--E-------------.----------K--VP-T.....RKR-- 162
G.CU.99.Cu74          --------L------------N----------K--KD-W-------R--K-----------..---IVR---------K-----H------KQRK----I--D--------H--K--------Q-I--EV-R---------K.----------K--V--T.....----- 162
G.ES.99.X138          ----------------K----N----------K--K-------F--R--KV----------..-T--VR---------K-----H-------QR-----I--D--------H-------------I--R------------T.----------K-----T.....-A--- 162
G.GH.03.03GH175G      ---------------------N--------I-K--K-----R----R--KV----------..-T--VK---------QA----N-T-----QR--N--I--D---H----H------E------I--E-------------.----------K--VK-A.....RK--- 162
G.KE.93.HH8793_12_1   ---------------------N------L---K----------F--R--KV----------..-K--VS-----N---K--Q--H-------QR-----I--D--------H------E----N-I--R--R----------.----------K--V--T.....-TR-- 162
G.NG.01.01NGPL0669    ----------------K----N--------I-K-------G-----R---T-------I--..-KI-VI---------KA-Q--H-------QR-----I--D--------H-----T-------I--EL-R----------.----------K--VK--.....----- 162
G.PT.x.PT2695         -------------------A-N----------K--K--S----F--M--KV--D-------..-K--VR---------K-----H-------QR-----I--D--------H----------------G------------T.----------K--A--T.....-A--- 162
G.SE.93.SE6165        ---------------------H----------K--------P--A-R---V----------..-T--V----------K--Q--H-------QR--R---E-D---H----H-------------I--Q-------------.Q---------KV-V-S-.....RSR-- 162
H.BE.93.VI991         ------------------K--N----------K--KK-V-------TN-KT-------V--..---------------------H-------QE-----I--D--------H-------------I---R---I-D-----R.----------T---S--.....RT--- 162
H.BE.93.VI997         ------------------NK------Y---I-K--NR-V-----D------------------..-----------------------------Q---------G-------TH------E-----GI--RV------------.Q---------T--VA--.....-T--- 162
H.CF.90.056           ------------------N-------Y--HI-R----------F--T-------------E..-----------N----E------------L-------E-G-------MH------E------I--RV-R---N-P---K.Q--T------T--VA--.....----- 162
J.CD.97.J_97DC_KTB147 ------------------N-------Y--HN-K---K-L-----D-N--KV---------C..-E--V------------------------L---R-----G-------MHH-----EA-----I----------------.----------T---K--.....-R--- 162
J.SE.93.SE7887        ------------------N-------Y--N--K---K-L-----D-N--K----------E..-I---------Q-----------------QR--R-----G-------MC-------------I--Q------D------.----------T---R--.....RR--- 162
K.CD.97.EQTB11C       ------------------N-------Y-----K--NR-R-----D-N--K-----------..-E--V-----------E------------L-K-R-----G-------I-------E----------R-----------T.Q---------T---A--.....-T--- 162
K.CM.96.MP535         ------------------N-------Y--HI-K--NR-Y-------R--K-----------..-E--V------L-----------------L---R---E-D-------I-------E--V---I---R------C-----.----------T--VA-R.....RP--- 162
```

```
                                                                 phos phos
                    transmembrane domain                     cytoplasmic domain
                   Vpu start                     α-helix                   α-helix          Vpu end
B.FR.83.HXB2                    MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHYESPHPRISSEVHIPLGD..ARLVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPELADQLIHLYYFDCFSDSAIRKALLGHIVSPRCEYQAGHN.KVGSLQYLALAALITPK.....KIKPP  162
01_AE.CF.90.90CF11697           --------------------Y--------I-K--KN---------Q--KV-------I--..-----R--------K-----H-------QRK-------D---RQ---Q------------M--QV-R-----PT---.----------K--A---.....-TR--  162
01_AE.CN.05.FJ051               --------------------N--------I-K--KK-S-------Q--KV----Y----E..GK---R-----Q---K-----H------KQRQ-N--I--D-------QQ----------R-I--QV-RS--D-PS---.----------K--T---.....--R--  162
01_AE.CN.06.FJ054               --------------------N----------KN-K----------Q--KV----------.------R-----Q---K--Q--H-------QRK----I--D---------Q----------R-I--QV-RH----PS---.----------K--T---.....--R--  162
01_AE.HK.x.HK001                --------------------N----------K--KK-Y-------Q--K----------E.------K-----Q---K--Q--H-------QRN----I---V--R----Q-----A-----R-I--QV-RYK---PS---.----------R--TG--.....GSR--  162
01_AE.TH.01.01TH_R2184          --------------------N----------K--KK----------QN-KV----------E..-K---R-----Q---K-----H-------QRE----I--D---R----H------------I--QVDRRK---PS--S.Q---------K--T---.....R-R--  162
01_AE.TH.02.OUR769I             --------------------N--------I-K--KK-S-------Q--KV----------E..-K---R-----Q---Q-----H-------QRN----I--D-------IQ----------S-I--QV-RS----PS---.----------K--T---.....R-R--  162
01_AE.TH.90.CM240               --------------------N--------I-K--KK---------Q--KV----------E.------R-----Q---K-----H-------QRK----I--D---K----Q--G---------I--QV-RR----PS---.----------K--T---.....R-R--  162
01_AE.US.00.00US_MSC1164        --------------------N----------K-VNK-A-------Q--KV---I-----E..-I---R-----Q---K-----H-------QRN-N--I--D--------Q------------I--QV-RR----PS---.----------K--T---.....R-R--  162
02_AG.CM.02.02CM_4082STN        --------------------N----Y---R-R--KD------F-CK--KA-----------.------R-----N---------H-F----KQ-K----I--D---H----H------E------I--QV---K---------.Q---------K--V--R.....-T---  162
02_AG.EC.x.ECU41                --------------------N----Y-----K--K----------R--KV-----------.-----VR-----------------H------KQR-----I--D---------H------E------I---V-R-K----------.----------K--VA-T.....RT---  162
02_AG.GH.03.GHNJ196             --------T-------K----N----------K--KD------F--R--KA-----------.-----VR-----N-----------H-------QR-----I--D----------------E-----V---QV-R-S-------S.----------K--VA-T.....RR---  162
02_AG.NG.01.PL0710              --------------------N--------I-K--Q--X-------R---V-----------.-----VR-----------A----H-------Q------I--D---------H------------IV-EL-R-----------.----------K---E--.....-T---  162
02_AG.NG.x.IBNG                 --------------------N----------K--K-------F--R--KVC----------.-----VR-----------------H------KQ------I--D-----------N----------I--E--R-----------.----------N--VA-T.....-T---  162
02_AG.SE.94.SE7812              --------------------N----Y---K-R--KD---------S---V-----------.-----VR-----------------H------KQR-----I--D---------H------------I--Q--R---------T.----------K--V--T.....RT---  162
02_AG.SN.98.MP1211              --------------------N----------K--HR---------R--KV------------..-T-IVR-----QP---------H----V-QQ------I--D---------H------E---------QV-R-K---------.----------K--V--T.....RR---  162
03_AB.RU.97.KAL153_2            --------------------N------I-I-K-----V-K-----RN--------------..-K---K-----------------A------E--------N----------------E----N-I---R---S---R----.-------------R---.....-----  162
04_cpx.CY.94.CY032              -A-------------K----N----------K--K-*--------R--KV----------E.-----VR-----QP--Q-----H-------LR--------D-------MH------E-------I---R-------------.--------------S--.....-T---  161
05_DF.BE.x.VI1310               -----------------N-------Y--H--K--NR-C----F--RN--------------..-K--V-------------------------R--------S-------V-------E----N-I--R-------F-----.----------T----------.....-T---  162
06_cpx.AU.96.BFP90              -----------------N-------Y--HI-K--KR-N-----D-N--K-----------S..-E---------N----K------------LR--R-----SM------IH------E-----------R-----D------.----------T---K-E.....-R---  162
08_BC.CN.97.97CNGX_6F           -------L--------K----N----------RR-N---------D-R--KV----------E.------K-----Q---------H-A-----LRS-N--IE-G-------M------A-------I---V-I---D------N---------T---K--.....R----  163
09_cpx.GH.96.96GH2911           -----------------N-------Y--HT-K--KR--------D-N--K-----------A..-E--V-----------K------------L---R-----S-------IH------E------II-DR------------.----------T---N-R.....-T---  162
10_CD.TZ.96.96TZ_BF061          -------L-------------N----------R--N--S-------R--K-----------E..-S-I-K----------E----H-------L------I--G-------MH-----AE------I--QV-T---N-----K.Q---------T---------.....-T---  162
11_cpx.GR.x.GR17                -------------A-------N----------K--K--L-------R---V---------EE..DM--V-----------E--------------R---R--------------MH-----AE------I---R-------P--R-.----------K--V----.....RT---  162
12_BF.AR.99.ARMA159             -----------------N-------Y--H--K--KR------F--R---V---------EK..-E---------QP------------------QG--R--I--G---R---I-------E------I---KI----N------.----------T---A---.....-T---  162
13_cpx.CM.96.1849               -----------------N-------Y--H--K--GR-------D-N--K------------..-K--V----------K------------L---R-----G-------TH-------------I---R-----D------.----------T---H---.....-----  162
14_BG.DE.01.9196_01             --------------------N----------R--K-------F--G--KV-----------..-K--VR----------K-----H-A-----QRS-----H-D----------------ET---N-I-----R-S---R----.-------------T--T.....-A---  162
15_01B.TH.99.99TH_MU2079        --------------------N----------K--KK---------Q--KV--Q-----EE.------K-----Q---K-----H-------QRG----I------R----Q------E----R-I--QV-RYS-D-PS---.----------K--T---.....R-R--  162
16_A2D.KR.97.97KR004            --------------------N----------K--KD-C-------IN--V----------G..-KIIVK------P--K-------------QE-----I--D--GH----H--------------I--K--R-----------.----------K--VGET.....RT---  162
17_BF.PE.02.PE02_PCR0155        -----------------N-------Y-IH--K-TKN-S----F--RN--------------..-K---A-----------------A-----QGG-R--I--G-------T------AE----N-I---RI----N-----K.----------T---A---.....-T---  162
18_cpx.CU.99.CU76               --------------------N-I---------K--KN---------R--KV-----------..-T--V----------KE----H-------LR--R-----D----------------E----N-I---R-------------.------------------.....-T---  162
19_cpx.CU.99.CU7                --------------------N--------I-K--NK-V-------RN-KV-----------..-K--VK----------K-----H-------QR-----I--D---------H-----------------QV-R----------.----------K--V--T.....RSR--  162
20_BG.CU.99.Cu103               --------------------H----------K--K--W----F--R--KV-----------..-K-IVR----------K----XH------KQRE----I--D---HM------N-----------I--E--R----FH----.----------K--V--T.....-----  162
21_A2D.KE.91.KNH1254            -------V------------N--------I-R--KD-V-----Q-R---V---------EE..-K-IVR------S--K-----H-------Q-G----I--D---H----------E------I--RL---K-------S.----------K--VAA-.....RT---  162
22_01A1.CM.01.01CM_0001BBY      --------------------N----------K-------------R--K------------.-----VR----------KE----H-------Q-----------D-------M--------------I--Q----S----T---.----------R--VA--.....-----  162
23_BG.CU.03.CB118               --------------------H-----XK-I-K--K--C-K-----TN-------------E..-K-I-K-----------------------------------G-------Q------E----S-I---R---S---------.----------K--V--T.....-----  162
24_BG.CU.03.CB378               --------------------N--I-----I-K--KD-C-K--F--RN-KV----------E..-K-IVR----------K-----H------KQRE----I--D-------QH--K-------------EK-R-S---H---S.----------K--V--T.....-----  162
24_BG.CU.03.CB471               -----------------K--N--------I-K--K--G----F--RN--V----------N..-K-IVK----------K--Q--H------KQRE----I--D-------QH--K-----------I--EQIR-S---------.----------K--VI-T.....-----  162
25_cpx.CM.01.101BA              -----------------NK------Y--H--K--KK---------N---------------..-E-AV----------E------------L---R-----G---R---MH-----T-------I--R----S-------K.----------K--VA--.....-----  162
27_cpx.CD.97.97CDKTB49          -----------------N-------Y---I-K---R-------D-N---------------..-V--V----------E------------LR--R-----G---R---TH------E------I--YR---K---------.-----------------R.....-T---  162
28_BF.BR.99.BREPM12609          --------------------N--------I-R--KK-V-K-----T---V-----------..-K--VI-----------------------E--------D----------------E----N-I--R--R---D--T---.----------T--TK--.....-T---  162
29_BF.BR.02.BREPM119            -----------------N-------Y-IHI-K---K-L------NR--K-----------EE..-K---I----Q----------------QG--------G-------SH--N---E----N-I--R--RS--------T.----------T--L--R.....-T---  162
31_BC.BR.02.110PA               -------L--------K--X-N--------X-RR-S--X-------R--KX----------E.------K-----Q---------H-------LRS-------D-------MH-----A--------I---R-----D------.------X---T---K--.....-R---  162
32_06A1.EE.01.EE0369            --------N-L-------N-------Y--HI-K--KN-------D-H--K-----------S..-E---------N----E------------L---K--I--S-------MH------E------I-----RH---------.----------K---K--.....-R---  162
33_01B.MY.05.05MYKL007_1        --------V-----------N------K-I-K--K--L-------Q--KV-------I-E..-KIA-R-----Q---K-----H-------Q-K----I--D---------N----------R-I--QV-RY----PS---.----------KV-T-Q-.....R-R--  162
34_01B.TH.99.OUR2478P           --------------------N-I---------K--KK---------Q--KV----------E.------R-----Q---K--Q--H-------QRN----I--D---R----Q------E----R-I--QV-RR----PS---.----------K--T---.....R-R--  162
35_AD.AF.05.05AF095             --------------------N--------I-K-TK--V-K-----R--KV-----------..-KIIVR----------K-----H-------L------I--DV--------H--N---------VI--QV---S-Q------.----------K--V--T.....-----  162
36_cpx.CM.00.00CMNYU830         --------------------N-------HI-KT-KD---------R--KV-----------.------R-----------------H------KQRK----I--D---------H------------I--Q------D-P----.----------K--V--T.....-R---  162
37_cpx.CM.00.00CMNYU926         --------------------N-------HI-K--K----------RD--V----------E..-W--VR-----QP---------H-------LRK-R--I--D---------H-----------T-I--R--R-----P---S.----------K--VP-T.....RKR--  162
38_BF1.UY.05.UY05_4752          -----------------N-------Y-IH--K--K--C-K--F--RN--V---------KE..-T-A-------------------------LGG-R--I--G-------I-------E----N-I---R-----N-P----.----------T---A---.....-T---  162
39_BF.BR.03.03BRRJ103           -----------------N-------Y-KH--K--KE-V-------NN-------------E..-K--V-----------E--------------R--------D----------------E----H-I--R--R---N------.----------T---------.....R----  162
40_BF.BR.04.04BRRJ115           -----------------K----------HI-RR-GR---------T--K-----------E..-E--V-------P--------------------R--------D---------H-----E----N----R----S-N------.----------T---K---.....-----  162
42_BF.LU.03.luBF_05_03          --------------------N----------K--K--Y-------T--K------------..X----K---------------------------------------N---------------E----NTI---R-------------.Q---------T---------.....-X---  162
43_02G.SA.03.J11223             --------------------N-X---------K--K---------RN-KV-------XX-..-T-XXX----XQ----X-Q--H------XQ-K-------D---------H------E------I--Q--R-----------.----------K--VK-T.....RK---  162
U.CD.83.83CD003                 -----------------K--N--------I-K--K--V-K-----TN------------N..--I-V-A---------------------QG----I--D---------H------------I------------T---.----------TT-VA-T.....-R---  162
U.CD.90.90CD121E12              --------------------N----------K--NK-V-K-----IN-----------EN..--I---A---------------------QG----I--D-------MH-----------------I---------D--T---.----------T--VP-I.....-R---  162
U.GR.99.GR303                   -----------------N-------Y--H--K--KR-V-----D-N-----------RE..-K--V----------E------------M-S-R-----S-------MH------E------I--Q------D------.----------T---N---.....RT---  162
U.NL.95.U_NL_95_H10986_D1       -----------------S-------Y--H--K--K----------R---V---------K-..-K------R------G--Y----------GQRK-------G-------MH-------------------R-------------.----------TV-V---.....R----  162
CPZ.CD.90.ANT                   I-K-*KIEMI--------NI--------IWETKVLKP-K------ND--KKGE------PTLDKK--V-VF---QC---P----H-------CGK-I------T---M--Q---P----Q-V-Q-I--ERILTY-H-KK--S.Q--T-----FCKILEFRGY.PKGPRRQ  167
CPZ.CM.01.SIVcpzCAM13           ---K--------------KA-N----Y---KTK--KD--------H-N--V-------FK-.NSK-I-----A-T----A---H----Q--MGS-V-----YV--K---SQ-----A-----R-I-----E---A-RE--R.Q--T--F---K--VKE-.....RLR--  163
CPZ.CM.05.SIVcpzEK505           --------V--------V---N----------K-------------TDN-K----I-----E..-K---V-----MP---P----H-------QGI-R--I------K----------TA----Q-V--RP-I-K---P---K.Q----------I-WVGVQ.....-RR--  162
CPZ.CM.05.SIVcpzLB7             -----K-L----I-----KA-N----Y-I---R-------------EN-KV---I-----E..-K---V-----TP---------H------KQ-K----I------HM------N--AE------I--RV------FP---K.---T-----IS--LQRQ.....PTR--  162
CPZ.CM.05.SIVcpzMB66            -----------------K--N-----------R---N---------NN-K-----------E..-K---------T----A---H-------QRS-R--I--D-----------------AE------I--Q--------PT---.----------ST--KS-.....PRR--  162
CPZ.CM.05.SIVcpzMT145           -------Q---------VK--N----Y---R-K--KN-------DGTN--MA-QIY--INQ..-K-IV-----MP---------H--A----ERG---------T--R----H----TAT---N-I-EKA-L-K--FPE--S.--------W--KQ-VKT-.....-NR--  162
CPZ.GA.88.GAB1                  -----------------K--N----Y-I-R-K------------DH-N-KVA--I---FR-.YSK-IV----A-SP---A----H----Q--LGS-V-----FT--R---SQ-----AET---R-I--QL-A-----KE--R.Q-----F---K---SER.....RHR--  163
CPZ.TZ.00.TAN1                  -------QV--MI-----L---T------IFTTKCCKD-K------TDT-KRAG-I----TE.RSK--VLH----AC---P----H-IGL---QGK-----I---T-------TR--T--AAG-V-Q-I--ERILTF-HF-S--R.Q--T--F--FRKVVESQDKQPKGPRR-  168
CPZ.US.85.CPZUS                 -------L---------VKA-N----Y---R-K--K----------TN--VN------IDK..-KI--I-----K---QA----H-------QGT-K--I---M--K-----------TA------V--RP-I-K---PT---.------F---K--VGQS.....-RR--  162
CPZ.CM.x.SIVcpzDP943            --------V--------VK--N----Y---RTK---K---------SN--V---I----EK..-K--V------MP--HE----H-------YGS-R--I---M--R----H-----AA------V--TQ-V-K---P---Q.Q----------K-WVGVE.....-KR--  162
CPZ.CM.x.SIVcpzMB897            -----K------I----VKA-N----------K-------------EN-KV---I---V---..-K---V-----MP---------H-------Q------I------H-----H-----AE------I--Q--------FP---T.---T-----IS--LKR-.....PK---  162
N.CM.02.DJO0131                 --------V------R--K--H----------R--E--Y-------TN--K---------V-L..-----------T----S----H------KLGK-K-------M--K-------N--TA----Q-V--RP-F-----P-----.Q--T--H---T-WVGV-.....-R---  162
N.CM.04.04CM_1015_04            --------V----------I-N-------H--K--K--Y-------R--KT--------V-Q..-T---V-----T---QP----H-------LRK-K-------M--K-----------T-----Q-V--RQ-F-----P-----.Q-DT------T-WVGI-.....-R---  162
N.CM.04.04CM_1131_03            --------V--------K--I-N-------H--K--K--Y------TR--K---------V-Q..------V-----T---QP----H-------LRK-K-------M--K----------TA----Q-I--RQ-F-----P-----.Q--T------T-WVGI-.....-R---  162
N.CM.95.YBF30                   --------V--------K--K-N----------K--K--Y------TH--K---------V-Q..-----TV-----T---QS----H-------LRK-K-------M--K----H-----TA----Q-V--RP-L-----P---K.Q--T------T-WVGA-.....-R---  162
N.CM.97.YBF106                  --------V-----X-K-----N----------K--K--Y------TK--KT---X----V-X..-X---V-----T---QP----HX------QGK-K-------M--K---C---N--TA----Q-V--RP-L---X-P---X.Q--T-----XT-XVGV-.....-RR--  162
O.BE.87.ANT70                   -------L-------QKVKA-N----Y-K-R-R-TEN-W-------RN--V--S-Y--V-V..-HV-V------MP----E---H-------Y-K-K--I---T--RM---H--T--TA--V---T--TA--V---I--QR-LTK---PT--S.Q--T--L---R-VVKAR.....SR---  162
O.CM.91.MVP5180                 -------L----I--QKVKA-N----Y-K-M-K--AN-R-------RN-KV--A-Y--VAE..-DI-V------MP---EE---H------QY-E-K--I---T--RM---H--T--TE------I--QR-LTK---L---S.Q--T--F---K-VVKV-.....RN---  162
O.CM.96.96CMABB637              -------L-------QXVKA-N--X-Y-K-X-K--QK-Y-------TRN-----G-Y--V-T..-NI-V------MP----E---H------QY-K----I---T--RI---H--T--TG----R-I--QR-LT----P---S.Q--T--L---R-VVKD-.....RS---  162
O.CM.98.98CMA104                -------L-------QKVKV-N----Y-K-R-R--KD-Y-----X-RN--V--G-Y--V-Q..-CI-V------MP---EDQ--H------QY-K-R--I---T--RM------T--TE------T--QR-LTK-D-PE--S.---T--L---RVVVKT-.....RN---  162
O.CM.99.99CMU4122               -------L------XQKVK--N----Y-KHR-R-TKE-H-K--F-FRN--A--G-Y--V-M..-CV-V------MP----E---H------QY-N-T--I---T--RI--QH--T--TE------II-QRIXTK---S---S.Q--T--L---RVVVKE-.....RN---  162
O.FR.92.VAU                     -------L-------QKVKA-N----Y-K-R-R--TD-Y-K-----IN--V--R-Y--V-P..-CIMV------MP---EEQ--H------QCRK-I--I---T--K----H--T--TE----R-T--QR-L-Q---PE--S.Q--T--L---R-VVKA-.....RNR--  162
O.SN.99.SEMP1299                -------L-------Q-VK--N----Y-K-R---T-D-Y----F--RN--V--G-Y--V-G..PWI-V------MP----E---H-------Y-K-K--I---T--RM--I---A--TE------I--QR-LT----P---S.Q--T--L---RVVVKE-.....RN---  162
O.US.99.99USTWLA                -------L-------QKVK--NX---Y-K-R-K--KD-Y-------TRN-----G-Y--V-G..-MI-V------MP----E---H------EY-N-K--I---T--RM---H--T--T-------C--QK-LT---FHK--S.Q--T--L---RVVVGE-.....TSR--  162
```


HIV-1/SIVcpz Proteins

```
B.FR.83.HXB2          LPSVTKLTEDRWNKPQKTKGHRGSHTMNGH*   192
A1.GE.99.99GEMZ011    ----R------------R-R--N-----C-   193
A1.KE.00.KER2008      ----K--A---------IR---E-------   193
A1.KE.00.KNH1144      ----RI-A------H---R-P-E------C-   193
A1.KE.00.KSM4024      ----RI-A---------R-LK-------C-   193
A1.KE.00.MSA4069      ----RI--------R----P-E------C-   193
A1.KE.00.NKU3005      ----R--A---------R-P--------C-   193
A1.RU.00.RU00051      ----R------------R-R--N-----C-   193
A1.RW.93.93RW_024     ----R--A------R---D---------C-   193
A1.SE.95.SE8891       ----R--A----------------------   193
A1.SE.95.UGSE8131     ----R-----------------------C-   193
A1.TZ.01.A173         ----R--A---------IR-P---------   193
A1.UA.01.01UADN139    ----R---------S---R-R--N-----C-   193
A1.UG.92.92UG037      ----K--A---------R---E------C-   193
A1.UG.99.99UGA07072   ----R--A---------R-----------   193
A2.CD.97.97CDKTB48    ----R--V---------R---------C-   193
A2.CY.94.94CY017_41   ----R--V----------------------C-   193
A.SN.01.DDI579        ----R------------R---------C-   193
A.SN.01.DDI369        ----R------------R--------S-C-   193
A.SN.96.DDI360        ----R--A---------R---------C-   193
A.CD.02.02CD_KTB035   ----R------------R---------C-   193
A.CD.97.97CD_KCC2     ----R------------R-----------   193
A.CD.97.97CD_KTB13    ----R--A--------------N-----C-   193
B.AR.04.04AR151516    ----K-----------------N---S---   193
B.AU.87.MBC925        ----------------I------I-----   193
B.BO.99.BOL0122       ---LQI---------------E-------   193
B.BR.03.BREPM2012     ----R---------H-------N-S-----   193
B.CA.97.CANB3FULL     ----A----------------E-------   193
B.CN.05.05CNHB_hp3    ----A----------RI-----N---S---   193
B.CO.01.PCM001        ----K---------R-I------------   193
B.GB.83.CAM1          ----R-------------------------   193
B.GB.86.GB8           ------------------------------   193
B.GE.03.03GEMZ010     ---IR----------R--------P-----   193
B.IT.05.SG1           ----Q--------------------ID---   193
B.JP.05.DR6538        ----A-------------------------   193
B.KR.05.05CSR3        ----A---------------K--------   193
B.NL.00.671_00T36     ----R---------R------E---H---   193
B.RU.04.04RU128005    ----A--A-X-------R---EN-------   193
B.TH.00.00TH_C3198    ----A----------R---------S---   193
B.UA.01.01UAKV167     ----R----------------------H---   193
B.US.04.ES10_53       --N-A----------R--------N-----   193
B.US.99.PRB959_03     -------------E----------------   193
C.AR.01.ARG4006       ----R--V---------RDR--N-I-----   193
C.BR.04.04BR013       ----K--V---------RDR----------   193
C.BW.00.00BW07621     ----R--V---------R-R--N-------   193
C.CN.98.YNRL9840      ----K--V-----N---IR-R--N-------   194
C.ET.02.02ET_288      ----S--V---------R-R--N--V----   193
C.GE.03.03GEMZ033     ----R--V---------R-R--N-------   193
C.IL.99.99ET7         ----R--V---------R-R--N-------   193
C.IN.99.01IN565_10    ---IR--V---------R-R--N-------   193
C.KE.00.KER2010       ----Q--V-----N----R-R--N-------   193
C.MM.99.mIDU101_3     ---IK--V-----N---IR-R--N-------   194
C.MW.93.93MW_965      ----K--V---------R-R--N-------   193
C.SN.90.90SE_364      ----R--V---------R-R--N-------   193
C.SO.89.89SM_145      ---IR--V-----N--RIR-R--N-------   193
C.TZ.02.CO178         ---IR--V-----N----R-R--N---S---   193
C.UY.01.TRA3011       ----K--V--------R-RDR--N-------   193
C.YE.02.02YE511       ----R--V---------M-R--N-------   193
C.ZA.04.04ZASK164B1   ---TR--V---------RDR--N-------   193
C.ZA.05.05ZAPSK240B1  ----R--V---------R-R--N-------   193
C.ZM.02.02ZM108       ----R--V---------R-R--N-------   193
D.CD.83.ELI           ----R----------Q-R------------   193
D.CM.01.01CM_0009BBY  ----R------------R----N-------   193
D.KE.01.NKU3006       ----R----------R---------H---   193
D.KR.04.04KBH8        ----R----------E---R---P-----W   193
D.TD.99.MN011         ----R----------Q--------------   193
D.TZ.01.A280          ----K----------R-------P------   193
D.UG.99.99UGK09259    ----R----------R--------------   193
D.YE.01.01YE386       ----R----------R--------------   193
D.YE.02.02YE516       ----K----------Q--------------   193
D.ZA.90.R1            ----R------------R------------   194
F1.AR.02.ARE933       ----Q--V---------R------------   193
F1.BE.93.VI850        ----Q--V--------E-R------------   193
F1.BR.01.01BR125      ----Q--V---------RD-----------   193
F1.ES.x.P1146         ----K--V-----E-R-IR-R--N-------   193
F1.FI.93.FIN9363      ----K--V--------EIR------------   193
F2.CM.02.02CM_0016BBY ----K--V--K--N-----------------   194
F2.CM.95.MP257        ----R--V-----N----R------------   194
F2.CM.97.CM53657      ----R--V-----N----R------------   194
G.BE.96.DRCBL         ----R--------E----R---ENP------   193
G.CM.01.01CM_4049HAN  ----S---------S---R---ENP------   193
G.CU.99.Cu74          ----K----------R-R---ENP------   193
G.ES.99.X138          ----R----------R-R----NP------   193
G.GH.03.03GH175G      ----R----N------------ENP-----W   193
G.KE.93.HH8793_12_1   ----R---------S---R---E-P------   193
G.NG.01.01NGPL0669    ----R----------R------ENP--S--W   193
G.PT.x.PT2695         ----R----------R-R----NP------   193
G.SE.93.SE6165        -----E-A----------R---ENP------   193
H.BE.93.VI991         ----R--V---------R------------   193
H.BE.93.VI997         ----K--V--G-------R---------R--   193
H.CF.90.056           ----R--V---------R------------   193
J.CD.97.J_97DC_KTB147 ----Q--V--------------E--------   193
J.SE.93.SE7887        ----Q--V----------T---E--------   193
K.CD.97.EQTB11C       V---Q--V---------R--------S-Q-   193
K.CM.96.MP535         V---K--V---------R-----Q------   193
```

```
B.FR.83.HXB2                LPSVTKLTEDRWNKPQKTKGHRGSHTMNGH*   192
01_AE.CF.90.90CF11697       ----R-------------R---ENP------   193
01_AE.CN.05.FJ051           ----K------------IR---E-P--H---   193
01_AE.CN.06.FJ054           ----K------------IR---ECPP-----   193
01_AE.HK.x.HK001            ----K-------SE----R---ENP------   193
01_AE.TH.01.01TH_R2184      ----K------------IG---ENP------   193
01_AE.TH.02.OUR769I         ----K------------IR---ENP-T----   193
01_AE.TH.90.CM240           ----EI.-----------.R---ENP------   191
01_AE.US.00.00US_MSC1164    ----R------------IR---ENP------   193
02_AG.CM.02.02CM_4082STN    ----R--A-----E----R-----RP-----   193
02_AG.EC.x.ECU41            ----K--A--------RIR----NRS-----   193
02_AG.GH.03.GHNJ196         ----K-------------R----NRPI----   193
02_AG.NG.01.PL0710          ----R-------KE----R-----RS-----   193
02_AG.NG.x.IBNG             ----R--A----KE----R-----RP-----   193
02_AG.SE.94.SE7812          ----K--A-----E----R-----RS-----   193
02_AG.SN.98.MP1211          ----K--A-----E----R----NRS-----   193
03_AB.RU.97.KAL153_2        ---------------R--D-------S---   193
04_cpx.CY.94.CY032          ----K--V----------R-R-ENQI-----   192
05_DF.BE.x.VI1310           ----R---------------R--N-----Y-   193
06_cpx.AU.96.BFP90          ----Q--V----------R---E--------   193
08_BC.CN.97.97CNGX_6F       ---IK--V-----N---IR-R--N-------   194
09_cpx.GH.96.96GH291I       ----K--V--------R-R---ENQ------   193
10_CD.TZ.96.96TZ_BF061      ----R-----------R--------------   193
11_cpx.GR.x.GR17            ----R-------------R-R--N--V--C-   193
12_BF.AR.99.ARMA159         ----K--V-----N----R------------   193
13_cpx.CM.96.1849           ----K--V----------RD--ENQ------   193
14_BG.DE.01.9196_01         ----R-----------R-R----NP------   193
15_01B.TH.99.99TH_MU2079    --R-K-------D----IR---ENP--D---   193
16_A2D.KR.97.97KR004        ----R--------------------------   193
17_BF.PE.02.PE02_PCR0155    ----K--V--------E-R---E--------   193
18_cpx.CU.99.CU76           ----R--A----------R----NQ------   193
19_cpx.CU.99.CU7            ----R-------------R---ENP------   193
20_BG.CU.99.Cu103           ----R-----------R-R---ENP------   193
21_A2D.KE.91.KNH1254        ----R--------------------L-----   193
22_01A1.CM.01.01CM_0001BBY-----R-------------R---EYP-----W   193
23_BG.CU.03.CB118           ----R-----------R-R---DN-------   193
24_BG.CU.03.CB378           ----R-----------R-R---EN-------   193
24_BG.CU.03.CB471           ----R-----------R-RD--EN-------   193
25_cpx.CM.01.101BA          ----R-------------R---ENQ------   193
27_cpx.CD.97.97CDKTB49      ----K--V----------R-L-ENQ------   193
28_BF.BR.99.BREPM12609      ----R------------I-----N-------   193
29_BF.BR.02.BREPM119        ----K--------------D-----------   193
31_BC.BR.02.110PA           ----R--V-----A----RDR--N-------   193
32_06A1.EE.01.EE0369        ----Q--V-----E----R---E------R-   193
33_01B.MY.05.05MYKL007_1    ----K------------IR---ENPATS---   193
34_01B.TH.99.OUR2478P       ----K-----------QIR---ENP-T----   193
35_AD.AF.05.05AF095         ---IR--A-----E----R-P-E--------   193
36_cpx.CM.00.00CMNYU830     ----K-------------R-----RST----   193
37_cpx.CM.00.00CMNYU926     ----K-------------R---ENP------   193
38_BF1.UY.05.UY05_4752      ----K--V-------PE-R------------   193
39_BF.BR.03.03BRRJ103       ---IK----------R------E---I----   193
40_BF.BR.04.04BRRJ115       ---IK-----------R--------------   193
42_BF.LU.03.luBF_05_03      ----R--------------------------   193
43_02G.SA.03.J11223         ----A---X----S---R-----RX-----   193
U.CD.83.83CD003             ----R--V------------K-----H---   193
U.CD.90.90CD121E12          ----R--V--------I---K---------   193
U.GR.99.GR303               ----K--V--------------E-Q------   193
U.NL.95.U_NL_95_H10986_D1   ----R--V----------R---E---I----   193
CPZ.CD.90.ANT               F--LSI---------RRMR---ENQ*-S-C-   197
CPZ.CM.01.SIVcpzCAM13       ----A--------T--R----PEN---S---   194
CPZ.CM.05.SIVcpzEK505       -------------RH--I-D------T--L-   193
CPZ.CM.05.SIVcpzLB7         ----S--A-----S----------I-S---   193
CPZ.CM.05.SIVcpzMB66        ----R--A-----NS---RD-K-N-I-----   193
CPZ.CM.05.SIVcpzMT145       ----A--A--------R-RD-QE-PMSS-C-   193
CPZ.GA.88.GAB1              ----A---------H-R--V-QENL-R----   194
CPZ.TZ.00.TAN1              --------------HRT-T-R-EN--LS-C-   199
CPZ.US.85.CPZUS             --------------C-----QPEN---S---   193
CPZ.CM.x.SIVcpzDP943        ----A--A------H--IRDQQE--------   193
CPZ.CM.x.SIVcpzMB897        ----S--A-------R-----KE--------   193
N.CM.02.DJO0131             --------------H---Q----NPI-----   193
N.CM.04.04CM_1015_04        ----A--------RH---Q----NPI-----   193
N.CM.04.04CM_1131_03        ----A---------H---Q----NP------   193
N.CM.95.YBF30               -------------EH--MQ----NPI-----   193
N.CM.97.YBF106              -------------ER---Q----NPI---P-   193
O.BE.87.ANT70               ----Q---------HLRIRDQLK-PS-----   193
O.CM.91.MVP5180             ----QR---------W-IRDQL---S-----   193
O.CM.96.96CMABB637          ----Q---G----RHLRIRDQQE--S--X--   193
O.CM.98.98CMA104            ----Q--------RHL-IRDQLE--S-----   193
O.CM.99.99CMU4122           ----Q---------HLRIMDQL-N-S-----   193
O.FR.92.VAU                 ----Q--------RHLRI-DQP--QL-H---   193
O.SN.99.SEMP1299            ----Q-------SRHLRIRDQLE--S-----   193
O.US.99.99USTWLA            V---Q---G----RC--I-DQLE--S---QY   193
```


HIV-1/SIVcpz Proteins

```
                                                
```

Env start — signal peptide — gp120 start — glycosylation NVT — glycosylation NDT — V1

B.FR.83.HXB2	MRVKE...KYQ.HLWRWGW....RW.....GTMLLGMLMI.CSATEK..LWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLVNVTENFNM.WKNDMVEQMHEDIISLWDQSLKPCVKLT.PLCVSLKCTDLKND..............	137
01_AE.CF.90.90CF11697	-G--GTQMNWP.---K-.............--LI--LVI-.---SDT..------------RD-D---------N-QE----------V----------IH-E------D-.---N-----Q--V---------------.----T-H--KA-LN..............	136
01_AE.CN.05.FJ051	-----TQMNWP.N--K-.............--LI--LVI-.-N-SDN..-------------D-D-----------HK---------------------MH-E--------.---N-----Q--V---------------.----T-N--NAAVN..............	136
01_AE.CN.06.FJ054	-----TQMNWP.N--K-.............--LI--LVI-.---SNN..-------------D-D---------R-HE-----I-V-------------IL-E--------.---N-----Q--V---------------.----T-N--TANLP..............	136
01_AE.HK.x.HK001	-----TQMNWP.N----.............--LI--LVI-.---SDN..-------------D-D-----------HE---------------------IK-GD-------.-X-K-A---Q--V---------------.----T-N--QANWK..............	136
01_AE.TH.01.01TH_R2184	-----TQMTWP.N--K-.............--LI--LVI-.---SDN..------------RD-D-----------HE--A------------------IH-E--------.---N-----Q--V---------------.----T-N---A-WT..............	136
01_AE.TH.02.OUR769I	----GTQMSWP.---TG.............-ILII-LVI-.---S-N..------------RD-D-----------HE---------------------IH-E--------.---N-----Q--V---------------.----T-N--NANLT..............	136
01_AE.TH.90.CM240	-----TQMNWP.N--K-.............--LI--LVI-.---SDN..------------RD-D-----------HE---------------------IH-E--------.---N-----Q--V---------------.----T-N--NANLT..............	136
01_AE.US.00.00US_MSC1164	-----TQMNWP.N--K-.............--LI--LVI-.---SD-..-------------D-E---------R-HE---------------------IY-E--------.---K-A------V---------------.----T-N--TANLT..............	136
02_AG.CM.02.02CM_4082STN	---MGIQKN-P.L----.............-MIIFCIMI-.-N-KD...-------------D-E------------------------------S---MP-E----D---.---N------V--------G-----R--.----T-H--NFNRS..............	135
02_AG.EC.x.ECU41	---MGIQKN-P.LF---.............-VIIFWTM--.-N-AN...------------RD-E------------Q-----------------S---IN-G--------.---N---------------------E--.----T-N-QNVT-I..............	135
02_AG.GH.03.GHNJ196	---MGTQKN-P.C----.............-IIIFWIML-.-KGED...------------RD-D----------S-------------------S---IY-E--------.---N---------------------E--.----T-E-HSVT-S..............	135
02_AG.NG.01.PL0710	-G-M-IQKSCP.P--K-.............-MIIFWTMI-.-N-D-...-------------D-E---------------A------------------IH-E--------.---K------------------------.----T-N--RL-N..............	135
02_AG.NG.x.IBNG	---MGIQKN-P.L----.............--NIFWIMI-.-N-EQ...-------------T-E----------------------------------IH-E----K---.---N------------------------.----T-D-HNFN-S..............	135
02_AG.SE.94.SE7812	---MGIQKN-P.L----.............-MIIFWIMT-.---GN...------------RD-E----------------------------------IH-G----D---.---S---------------------Q--.----T-H-Q-NLTS..............	135
02_AG.SN.98.MP1211	---MGILKSCP.PF---.............MI---WI-I-.-N-EN...-----------RRD-E------------E-----------------S---MH-E--------.---N------V-----------------.----T-E-HNYNYT..............	136
03_AB.RU.97.KAL153_2	-----IRK.H..----..............--LF------.-----N..----------------------------SK--------Y-------S---IP-K--------.G--N------------------------.----T-N-----KE..............	133
04_cpx.CY.94.CY032	---MGMQRN-P.---E-.............--LI--LVI-.---SNN..------------RD-E-------E----EK----I----------------A-I--------.-----------------NEG----A---.S---TFT-INATTT..............	136
05_DF.BE.x.VI1310	---RGMQRNWP.--GK-.............-LLF--I-I-.----D-..F-----------R--K----------G--K-----------------------G--------.-----------T-----------------.----T-N---F-AN..............	136
06_cpx.AU.96.BFP90	----GIQTSW-.---K-.............--LI--LVI-.---SKN..M---------A-ED-D-I----------SPDK------------------IS-K--------.---N--D-----------E---------.----T-T--NATLG..............	136
08_BC.CN.97.97CNGX_6F	---RGTRRN--.QW-I-.............-VLGFW----.-NVEGN..---------------K------------E-----------------S---I-ME--------.-N----N-----VT--------------.----T-E-RNVSSN..............	136
09_cpx.GH.96.96GH2911	-----IQRNW-.N----.............-L-I------.-N-A-N..---------------E---------Q--SK-K------------------IG-E--------.---N------K-----------------.----T-D-HRVN-S..............	136
10_CD.TZ.96.96TZ_BF061	-T-TGIQRNCR.QW-I-.............-ILGFW----.-N-R--..--------------T-------------KA-A--I-------------R-I--E--------.---G--D------------G----I---.----T-N-NGINAT..............	136
11_cpx.GR.x.GR17	-----TQRNWH.N----.............-L-IF-----.-N-EK...M-----------RD-D-I-------RT-S--K-------S--------R-LS-E--------.---N--------V-----E---------.----A-N---A.................	132
12_BF.AR.99.ARMA159	---RGMQRNW-.--GN-.............-FLF--I-I-.-N---N..--------------------------S-ER-----------------------E------D-.---N------T-----------------.----T-N-SNANAT..............	136
13_cpx.CM.96.1849	--ARGIQMTW-.N--K-.............--LI--LVI-.---SNN..------------RD-E------------S--K------------------IN-G--------.-E-N------T----I---------S--.----T-N-NNITS-..............	136
14_BG.DE.01.9196_01	-KA-GIQRNC-.R--K-.............-VT-I-----.-N-V-Q..--------------T----------R--------------------S-L----E--------.---N------------------------.----T-D--N-NET..............	136
15_01B.TH.99.99TH_MU2079	-----TQMNWP.N--K-.............--LI--LVI-.---A-N..-------------D-S---------------A--G-----------S---EI-G--------.-E-------Q------------------.------N----R--..............	136
16_A2D.KR.97.97KR004	---RGIQRN--.---K-.............-ILI------.SK---D..------------RD-E---------------A------------------IN-E--------.---G------------------------.----T-N-SRV--T..............	136
17_BF.PE.02.PE02_PCR0155	---RGMQRN--.---K-.............-I----I---.-N--KD..------------R--X----------G-E------------------------E------D-.---N------T-----------------.----T-N--NANAT..............	136
18_cpx.CU.99.CU76	---RGTQRN-P.S--K-.............--II----LM.----DH..-------------D-V----------S-S--S--------------D---IT-G--------.---N--------------E---------.----T-N--NVXXS..............	136
19_cpx.CU.99.CU7	---MGTRRS-P.P--N-.............--LI--LVIM.---FDN..------------RD-E------------------------------S---IN-E----E---.---N------A-----------------.----T-D.....................	130
20_BG.CU.99.Cu103	----GIQTNW-.C--T-.............--LII-LGI-.---SNN..------------ED-N-P----------S--S------------------IP-E--------.---N-----N--------E---------.----T-N--N-SVN..............	136
21_A2D.KE.91.KNH1254	---R-MKRNC-.N--K-.............-I--F-----.----DN..------------R--NAS----------TK-A--I---------------IE-E--------.---N--------------E---------.----T-N-AGSTTS..............	136
22_01A1.CM.01.01CM_0001BBY	---RGTQMTW-.N----.............--LI--LVI-.---QN...------------R--E------------P-----------------S---IE-A--------.---K-----Q------------------.----T---E-ARFN..............	135
23_BG.CU.03.CB118	----GIQKNW-.R--T-.............--LII-LGI-.---SDN..------------ED-Y-P----------S--S------------------IP-X--------.---H--------------E---------.----T-N--KVDCE..............	136
24_BG.CU.03.CB378	----GIQMNW-.---T-.............--LII-LGI-.---SNN..-------------D-D-P--------S-S--S------------------IP-E--------.---N--------------E---------.----T-N--T-NSN..............	136
24_BG.CU.03.CB471	---RGIPMNW-.L--T-.............--LII-LGI-.---SSN..------------ED-Y-P----------VA-S----------------H-IP-Q--------.---N-----N--------E---------.----T-N--SVDCN..............	136
25_cpx.CM.01.101BA	---R-IQRN-P.L----.............-ILII-LGI-.---SDN..-------------D-E-X----------SP-A--------------S---IE-T--------.---N--------------E------Q--.----T-N-VT-NCS............NR	138
27_cpx.CD.97.97CDKTB49	----GTQRNW-.G----.............-L-I---A--.---K-D..------------RD-D-----------HT---------------------II-E--------.---N--------------E---------.----T-N--NANAT..............	136
28_BF.BR.99.BREPM12609	---RGMQRNWK.--GT-.............-LLF--I-I-.-N--NQ..S------------------------N-VK--A--I-----------I------K--------.---N--------------E-I-------.----T-N--NNVTV..............	136
29_BF.BR.02.BREPM119	--ARGTRKNC-.---E-.............-IL-------.---EQ...-----------------------N-------A--------------D---IP-K--------.-----------Q----------Q------.----T-N-S--G-N..............	135
31_BC.BR.02.110PA	----GIQRNW-.QW-I-.............-ILGFW----.YNVRGN..------------R--K------------------------------S---LD-G--------.-E----D---Q-----*-----------.----T-N-S-AISN..............	135
32_06A1.EE.01.EE0369	---RGTQMNWS.---K-.............--LI--LVI-.-N-SKD..----------A-ED-D-I---------HS--R------------------IH-K--I-----.---H---------K----E---------.----T-H-KNATFK..............	136
33_01B.MY.05.05MYKL007_1	-----TRMNWP.N--K-.............--LI--LVIM.---SNN..------------RD-D-----------HE---------------------IH-K-I------.---K-A---Q--V------------Q--.----I-H--NANLT..............	136
34_01B.TH.99.OUR2478P	-----TQMNWL.N--K-.............--LI--LVI-.---SDN..------------RD-D-----------HE---------Y-----------IH-E--------.---N-----Q--V---------------.----T-N--NAALTNVTI..........	140
35_AD.AF.05.05AF095	T--MGTQMNW-.N----.............---I---II-.---V-N..-------------D-E------------E--M------------------IP-G----D---.---K------T-----------------.----T----NATTNI.............	137
36_cpx.CM.00.00CMNYU830	--ARGTQRNCL.---K-.............--LII-LVI-.---SDL..------------RD-N----------------------------------ID----------.---N-----Q------------------.----T-E-QNFNGSNNH...........	139
37_cpx.CM.00.00CMNYU926	---MGIQRN-P.P--N-.............--LI--LVI-.---NNL..-------------D-D---------T--VP--------LPGTHRPH.--KINVG---KD---WKN-MVEQMQIDI-SLWGPKL.---E---PS-RH#-NSSHKV-PQDQRE..........	140
38_BF1.UY.05.UY05_4752	---RGMQRNW-.P-GK-.............SLLF--I-I-.-N-ED...----I-------R---------------------------------D---I--G------D-.-R-P------T-----------------.----T-H--N-TTK..............	135
39_BF.BR.03.03BRRJ103	---RGMQRNW-.--GK-.............-LLF--I-I-.-N-A-N..------------RD--------------E----------------------E-K------D-.---N------T-----------------.----T-N-SNVNVT..............	136
40_BF.BR.04.04BRRJ115	----GIKKNC-.R--K-.............-M----L---.-N-NDQ..W-------------------------G-E---------------------I--G--------.----------------------------.----T---AAFNGT..............	136
42_BF.LU.03.luBF_05_03	---TGIMKN--.-X-K-.............-I----I---.S-----..---------------X------------------I---------------------------.---N------------------------.----T-N--NXXXXXXXXX.........	141
43_02G.SA.03.J11223	-----IQRNW-.Y--TM.............WXLI--LXI-.---SNN..------------ED-X-P-----N----S--K--X---------------MT-E----P---.---N------D-------E------M--.----T-T--NATCN..............	136
U.CD.83.83CD003	-----IQRN--.---K-.............SLII---I--.-K-I--..S------------D-E------------EK-S--------------S---L--G--------.---K------------------------.F---T-N-I-V--S..............	136
U.CD.90.90CD121E12	--A--IQRN-H.Y--K-.............SLII---II-.-KGE-Q..S------------D-D----------G-EK-A--------------S---L-----------.---N-----------------------S.S---T-N---VGTN..............	136
U.GR.99.GR303	--L-GIQRNC-.N----.............AL-IF-----.-K--DN..-------------D-E-----------HSA-K---L----------S---MI-G--------.---N------T-----------------.----T-N-KNPSCV..............	136
U.NL.95.U_NL_95_H10986_D1	---RGMQRNWR.--G--.............-ILF--I---.---AD-..----I-------QD-N----------S-SN-R------------------LP----------.---N----V---------E---------.----T-N---VT-R..............	136
CPZ.CD.90.ANT	--KPIHII........................WGLA-L-QFIEKGTNEDY---F------RN--P-----TN-SMTS---------TS---I--D-IV-R.L-TSVW--A.Y--Y---S-T--MXQ-FQ--H--------.-M-IKMN--GYNGT..............	127
CPZ.CM.01.SIVcpzCAM13	-K-M-ERKRSF.NGFYMT............MTICMA-L-P-LTSD...E-----------R-VN-V---------HS--A--I------------------H----D---.-N-Q-----Q---S--------------.----TM--SNVTRQ..............	136
CPZ.CM.05.SIVcpzEK505	-K-T-MQKNW...................LICC--IG-I...KIIGS.E------------RD-E-V---------HS--A--I---Q-----------LIP-----R-D-.---N--D--Q-------E----------.----T-S-SSWRSV..............	130
CPZ.CM.05.SIVcpzLB7	----GTWTSCM..........PYYVVGILSLCS-IIS...STKN-N..---------------EA----------QKA-A--I--SQ-------D-KAIE-S--------.-T-A--N--QQ-V-----------I---.----T-H-SIP-F-..............	137
CPZ.CM.05.SIVcpzMB66	-KAM-TQRNCR.T-SLKEI..........ILCT-VLGI-.GIIKCEDNM-----------R--D-----------QNP-------SQ---S-N---E-IE-T----Y--A.-E-N---------VN-----V------I.----T-N-SLF-CI...........KEN	143
CPZ.CM.05.SIVcpzMT145	-K-MGMKYYWW.RNSCRSI..........SIK-ILIGW-.A-CFGEENW-----------R--K----------S-S--A--I---Q-------T---L-P-----E---.-E-Y--D--Q-------E----------.----T-T-NNPT-T..............	140
CPZ.GA.88.GAB1	-K-M-.KK-RD.WNSLSII..........TIITIIL-TP.-LTS-...------------HD-DPV---------HS--A--I---Q-------S----F-P--I-S---.---N--D---------------------.----T-Q-SKANFS..............	136
CPZ.TZ.00.TAN1	*KKMK.NLIG...ITLI..............LIITI-G-.GFS-....YYT--F---------QP--------DITSRDK--I----N---L---Y--T-A--SIR----.EE-Y--QE-K---L--FQ--F-------.-F-IKMT--MTNTT..............	128
CPZ.US.85.CPZUS	-K-M-.KK-R...--LS............YCL-SSLIIP.GLSS....--A---------RDVE------------KQ-A--I---Q-------------H-P---K-D-.-E-N-A---Q--------------I---.----TMT-LNPDSN..............	131
CPZ.CM.x.SIVcpzDP943	---M-RIRGS.....................LW-MG-I-P-LTGSD.Q--A----------D-E------A--S-LNK-A--I--SQ------------HVP-I--T-D-.-E-N-----QT-----------------.----TMH-IPYNIS..............	130
CPZ.CM.x.SIVcpzMB897	-K-MGTQRSWK.QFLSLR..........IVMCLW-LGV...IRSENN..---------------K-----------QIA-A--I--SQ---------E-IT-E----E-DA.-N-N--D--Q--L---------------.----T-N-STNTTT..............	137
N.CM.02.DJO0131	V-MMGMQIGW...................PFFC-MIS-T-.G-KK....Y-A---------RDVE-V---------HS--A--I---Q-------------P-D---P---.-E-K-A--VQ-------E----------.----TMN-SNSNGN..............	129
N.CM.04.04CM_1015_04	---MGMQSGW...................PFFC--ISSI-.G-EQ....H-----------RDTE-V---------HS--A--I---Q------------P-I----H-D-.---N-A---Q-------E-------X--.----TMH-XNSNGX..............	129
N.CM.04.04CM_1131_03	K--MGMQSGWVTRRVLGMMSGW.......PFFC--ISSI-.G-EQ....H-----------R-TD-V---------HS--A--I---Q------------H-I----K-D-.---N-A---Q-------E----------.----TMH-N-SNGF..............	141
N.CM.95.YBF30	-K-MGMQSGW..........MGMKSGWLLFYLLVSLIKV-.G-EQ....H-----------R--E-----------HS--A--I---Q-------------L-P---K---.-E-K-AD--Q-------E----------.----TML-N-SYGE..............	138
N.CM.97.YBF106	R--MGMQSGW...................PFFC--IS-T-.G-DP....H-----------RD-E-V---------HS--A--I---Q-------------L-T---Y---.-E-K-A---Q-------E----------.----TML-NNSNGN..............	129
O.BE.87.ANT70	-KAM-KRN-K...--TL..............YLAMALITP.-LSLRQ..-YA---A-----ED--PV-------NLTS--K--I--SQ-------T-Y-YP-H---DD--I.---Y-----Q---------------QM-.F---QME--NIAGT..............	133
O.CM.91.MVP5180	-K-MKKNNRK...S-SL..............YIAMAL-IP.-LSYS-.Q-YA---S-----E--APV-------NLTS--Q--I--SQ---------H-FP-G---D--DI.---Y--D----------E------E-M-.F---QMN-V--QTN..............	134
O.CM.96.96CMABB637	-K-M-MKN-K...-CIL..............CI-MAL-IP.-LNSDQ..-YA--FS-----ED--P--------NLTS--K--I--SQ-------S-H-YP-I---DK--I.-E-Y--D--Q-------E-------QM-.F---QMN---VNAT..............	133
O.CM.98.98CMA104	-KAMKKKN-R...--IL..............YIVM-LITP.-LSSSQ..FYA---A-----EN-NPV-------NLTSP-Q--I--AQ---------H-YP-H---DR-DI.-E-Y-----Q---TD--E--------M-.FM--QME-N-INTT..............	133
O.CM.99.99CMU4122	--AM-KRNRQ...-GIL..............YIVMALIIP.-LSYDQ..-YA---S-----E--KPX------VNLTS--K--I--SQ-------T-H-YP-K---DK--I.-E-Y-----Q-------E-------QM-.F---QMN--NVNDT..............	133
O.FR.92.VAU	-KAMGKRNRK...-GI-..............CLI-ALIIP.-LSCNQ..-YA---S-----ED-KP--------NLTS--Q--I---Q-------S-N-YE-K---DK--I.-E-Y--D--Q----D----------QM-.F---QMN---I--S..............	133
O.SN.99.SEMP1299	-K-M-QRNRK...-GIL..............CIVMALITP.-LSYNQ..HYA---A-----E---PV-------NLTS--Q--I--SQ-------S-Y-YP-TK--D---I.---Y-----Q-------E-------QM-.F---QMN--NYVQG..............	133
O.US.99.99USTWLA	-KAM-XKNRK...S-IL..............CXIMALIIP.-LNS-Q..-YA---S-----ED-SP-------VNLTS--Q--I--SQ-------S-N-YD-K---DY--I.---Y--D------------------QM-.F---QMN---V-DN..............	133

HIV-1/SIVcpz Proteins

glycosylation NDT, glycosylation NCS, glycosylation NIS, glycosylation NSS (V1); glycosylation NDT, glycosylation NTS (V2); glycosylation NKT, glycosylation NGT, glycosylation NVS, glycosylation NGS

```
B.FR.83.HXB2            ...........TNTNSSSGRMIMEKGEIKNCSFNISTSIRGKVQKEYAFFYKLDIIPIDND.............................TTSYKLTSCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAE  267
A1.GE.99.99GEMZ011      ...........SSN---VN-NDSMVE-MR-----MT-EL-D-RKTV-SL--R---VATGIN............................DSKQ-R-IN----AM------IT-------------------RDPN-------K-----------K---------------  267
A1.KE.00.KER2008        ......NTKVN-TSAPTTAATVNNDSDM----Y-MT-EL-D-K--M-SL--R--VEL-NKNN..........................SNS-E-V-IN-K--TL-------------------------M-KD-E-----L-K-----------K---------------  274
A1.KE.00.KNH1144        ........TDVSG-RGNITI-KEME---------MT-E--D-K--V-SL--R--VV--NQGNSS.......................SKNSSE-R-I-----A----------------------------RD-E-----E-K-----------K---------------  276
A1.KE.00.KSM4024        .............--VNIT-EEE-IKNL------MT-EL-D-R--VNSL--R---VK-NESS...........................KSSE-R-IN----A----------------------------KDEK-------K-----------K--------------G  266
A1.KE.00.MSA4069        .............NATGRVDNSTMRE-V---T--VT-E--D-RK--HSL--RV-LVQ--EMQG........................NNSDSK-R-IN----A-----------------------------DMNYK-Q---Q-----------K---------------  269
A1.KE.00.NKU3005        .......TNANNC-HAN-SITNDMA---------MT-EL-D-K--V-SL------VQ-NGNNGS.......................EYNGSE-R-IN----A-----------------------------DEE-------K-----------K--------------K  276
A1.RU.00.RU00051        .................-MSNSSRY--M------MT-EL-DRKR-VHSL------VSTSSN............................ESGP-R-IN----AM-------V--------------------Q-D-------R-----------K---------------  266
A1.RW.93.93RW_024       ....................FTNI-E--T-----VT-EL-D-TK-VHSL--R---V-LNEYSSSNKS.................SNNNNSSQE-V-IN----A--------T-------------------KDTDY------R-----------K---------------  269
A1.SE.95.SE8891         ...............-VTNITSDMA---------MT-E--D-R--VH-L--R--VV-M--N.............................NSL-R-IN----A-----------------------------D-E-------K---S-----------------------  262
A1.SE.95.UGSE8131       ...........NTNST-ANLTDSV---MR------T-EL-D-KK-V-SL------VK-NKNKSFRGK................NSSGNSSSDR-R-IN----A-----------------------------EDE-----T-R--------------------------K  279
A1.TZ.01.A173           ......TDSDNSDNATN-TITREMI----------T-E--D-K--V-SL----VVE-S-S.............................NSSQ-R-IN----A--------T-------------------K----------N---S-----------------------  272
A1.UA.01.01UADN139      ........TLNCTASNCTISNSSLYE-MR-----MT-EL-D-EK-VHSL------VSTGSN............................DSGE-R-IN----AM-------T--------------------DT--------R-----------K---------------  270
A1.UG.92.92UG037        ...........NI---ITNSSVNMRE--------MT-EL-D-NR-V-SL----VVQ-N-GN...........................NSSNL-R-IN----AL-------T-------------Y------D-E-----L-K---------------------------  269
A1.UG.99.99UGA07072     .........NL--S-TTVNITEDM----------MT-EL-D-R--MHSL--R---VQ-NGNRS.........................NSSEQ-R-IN----A----------------------------KDRD-------K-----------K--A------------  272
A2.CD.97.97CDKTB48      ....................NTNSTE------Y-MP-ELKD-T--V-SL--E--VVLLNRSKN.........................SSYST-R-I----------------------------Y-----KD-E---K-S-S---S---A------A------------  260
A2.CY.94.94CY017_41     ............THSN--STQSPINE------Y-TT-IL-D-T--V-SL--R--VVQL-ESENKN.....................TSGSN-L-R-IN----T--------T-------------------KDPR-----S-K---S-------K--A------------  272
A.SN.01.DDI579          ..............FTNTENVTKAYE-M-------T-EL-D-K--I-SL--R--VV--E-GSNS.......................TENS-E-R-IN----A----------------------------KE-K-------R-----------K---------------  269
A.SN.01.DDJ369          ............-S--N-SVKNEMP-------Y-VT-A--D-R-NV-SL--RV-VV--NENSNS.......................SENSSE-R-IN----A--------T-------------------KD-D-----T-R-----------K---T-----------  271
A.SN.96.DDJ360          ..................RYDTKD-E-M-------T-EL-D-R-SV-SL----VV--N-NSS.........................NNSSNE-R-IN----A----------------------------KD-K-------R-----------K---------------  264
A.CD.02.02CD_KTB035     ......NVDISSHI-N-YVNVSV-EHAMT-----VT-A--D-Q--VHSL---V-LVQLGE-KGSS....................NSSNSSIL-R-IN----A----------------------------KD-E-------K-----------K---------------  280
A.CD.97.97CD_KCC2       .................NNTKVEVP--MT-----MT-ELSD-K--VRSL--RI-LVQ-G-NTND......................SSNRSLQ-R-IN----T------------V---------------KDQE-------K-----------K---------------  267
A.CD.97.97CD_KTB13      .............-AHKVSVSPEMQEKVA-----VT-EL-DRK--V-SL--RE-LV---E-KNNSNSASISTSSNSSNCSISDSNSTCINIST-R-IN----T------------------V----------E-E-------K---S-------K--------------K  293
B.AR.04.04AR151516      .........ATNS-RINNSSLGTME-Q-------ISR-EN--R--X-L-NS--VVX--K...............................EQ--M-IN---------------------F-T----------D-K-----------------------------------  266
B.AU.87.MBC925          ....ATNTTSSSE-VMEK-EKV-------------T-N---------L-----VV-V-K...............................N-----I---------------------Y------------KDEK---K-L-----------------------------  273
B.BO.99.BOL0122         .....................NNST-DM------TAG-KD--K----L--N--VV---D-N............................GS-T-R-IN----I---T----T---------T----------D-N-------K---------------------------  259
B.BR.03.BREPM2012       ..............LTNDS-WEDM-E---------T-G-KD--K--H-L-NR--VV----AEN.........................GSR---M-RN----M-------------------------------------S---K---------------------------  277
B.CA.97.CANB3FULL       ............---TN--RE-R--E-------K-T-NR-D------L-----VV----...............................S---R-I------------------------------L----D-K-----------------------------------  266
B.CN.05.05CNHB_hp3      .................NNSSTT------------T--TGT--.RDS-L-N---VL----SSTSYN..................SSTSNRYH--R-IH----R------------------T-----L----------I-------------------T----------D  273
B.CO.01.PCM001          ..............S-GRVMPT-KPE--------VT---GS--.Q-T-L--R--VE---KN.............................DSR-R-IH------K----------------T-------Q--D-K-----------------------------------  262
B.GB.83.CAM1            .........RTNSSDWDRREGEKM----------VT----N--R----L-----VV---KA.............................N---T-IH-----------------------T----------D-K---K-------------------------------  268
B.GB.86.GB8             ...............TNTDNS---G--M-------T----D-M-----LL-----VS-GS-............................N----I--K-DA---------I--------F----------------D-K-------------------------------  262
B.GE.03.03GEMZ010       ...............TTNSSNSWKEV-M------VT---GN-MH----LL----LV--EEN............................NN-I-T-I-------------I-----------------V---DEK-------K-------------T------------K  263
B.IT.05.SG1             ..........................-M------VTNN-GKTA.N---LL--Y--VQ--DY.............................N---R-I----T--------I---------------------D-R---S-------------------------------  248
B.JP.05.DR6538          .......STETN-RTNT-TTSVPDREPM-----DVT--L-D-IT----L--RP-VV---QTNNT......................YNNTYST-R-IH------------M----------------------NS-------------------K---------------  277
B.KR.05.05CSR3          .........TT--GT-KWNGDNSTI----K-Y--TT-GMID-MKE---L--N---V-L--SLAKNS..................SHNKTSYAN-R-I-------K-----I----------T----------DTK-------K---------------------------  279
B.NL.00.671_00T36       .......SSNNITGSNNNSNLEQMAR--S------T-T-KN-R-R-F-LLS---V---................................SY--M-IN---------------Q-------T----------D-K-------K---------------------------  270
B.RU.04.04RU128005      ............-TD-TTNNST--G-E-------TT-N--N-M----L-----V---G-...............................N---M---X----------------------------KENK-------K---------------------------  266
B.TH.00.00TH_C3198      .............NS-TTNSSMIGG--M------VT-N-GA--.T---L---T--VQ-A--.............................N---R--K------------I-------R--------------AR-----------------------------------  263
B.UA.01.01UAKV167       .VTNTNGTITSNS--NGTSWEY--G--M------VT---GNEMKED--L------KS-K--.............................N---I-IN--N----------T--------------------D-K---K-------------------------------  276
B.US.04.ES10_53         ....................LANMT------T---T-TM-D----A--L--ST-VVK-N-ND...........................NNS--R-I----------------------------------RD-K-----------------------------------  265
B.US.99.PRB959_03       .............--TT-NWEK--P---------VT----D-------L-----VVQ--DID...........................NG-N-R-I-----------------------------------------S---N---------------------------  265
C.AR.01.ARG4006         ..........ISX-I-NDTILDDM----------MT-EL-D-QK-VH-L--R--XV-LN-EN...........................DSSD-R-IN----A--------T-D-----------Y-----------------K----------K---------------  269
C.BR.04.04BR013         ........IIGY-NQTIKYNET--V-Q-------MT-EL-D-KR-VH-L--R---V-LEDGSN.........................NTYDN-R-IN----T--------T-D-----------Y----------------N-----------K---------------  273
C.BW.00.00BW07621       ....................TNVTNADY-------T-EL-D-RK----L--R---VQLGENN..........................ANSSE-R-IN----T--------N-D-----------Y------T---D-----------------K-------------S-  260
C.CN.98.YNRL9840        ...........SSNGTYNIIHNETYK-M------AT-VVEDRK--VH-L--R---V-L-ENNSSEK...................SSENSSEY-R-IN----A----------D-------T---Y------DN--------H-----------K---------------  276
C.ET.02.02ET_288        ...SNSNVTSNP-VT---TDASKVI-QM-------T-EV-D-NK-A--L--R---V-LT--N...........................SSDE-R-IN----T----------D---L-------Y------D---------K-----------K------------I--  275
C.GE.03.03GEMZ033       .............STNAKINSTSVGE-MR-----VT-EL-D-EK-Q--L--------L-KGNSNAN...................QGNDTSGE-R-IN----T----------D-----------Y----------------N-----------K---------------  273
C.IL.99.99ET7           .............VT-NVTINNDTIN-L-------T-E--D-KKS-F-L--R---V-LNEQ............................NSSE-R-IN----AV---------D-----------Y----------------N-----------K---------------  265
C.IN.99.01IN565_10      ..............-V--NSDNNTR---R--T--MT-ELKD-KKQVS-L--R---V-L-GKNESS.....................NPNTSSE-R-IN----A----------D-----------Y-------E--------S-----------K---------------  271
C.KE.00.KER2010         ......ATSNS--GTGNDTYENDMTR-M------TT-E--D-Q--V-SL------VSLE-GN...........................NSST-R-IN----A-------I--------------Y----------------Q-----------K---------------  273
C.MM.99.mIDU101_3       ....NSTGRNVSSSSG-N-TY-GNYT--------TT-VLKDRK-SV--L--R---V-LNEKKSS.......................ENSSEY-R-INW---A--------T-D-----------Y-----KD-K------RR-----------K---------------  279
C.MW.93.93MW_965        .............-N-GTVNVNDTMY--------MT-EL-D-KKQV--L------VSLNENS..........................NNSSE-R-IN-------------T-D-----------Y----------T-I---K-----------K---------------  267
C.SN.90.90SE_364        ...............NNATSNDNPN----------T-EL-D-RRN---L--R---V-LSGSKN........................SSNSSE-R-IN----A----------D-----------Y----------------N-----------K---------------  267
C.SO.89.89SM_145        ............I-SVGNVTSNDN-E--------AT-ELSD-K--V--L--R---V-LNEGE..........................NNSRK-R-IN----A----------D-----------Y------E---------R-----------K---------------  268
C.TZ.02.CO178           ......ATIREYVALEHNASLGNYSD--------VT-ELKD-TKRVH-L--L---E-LNEKEDK.......................DANF-T-R-IN----A----------D-----------Y----------------K-----------K-------------T-  277
C.UY.01.TRA3011         ..............RTKNNHTDNM-------T--MT-VL-D-KE-VH-L------V-L-GN............................TSSN-R-IN----A----------D-----------Y------E---------N-----------K--------------A  264
C.YE.02.02YE511         .............-VTNDVTFNDTMR---------T-E--D-RKQ---L--R---V-L-GKE..........................ENPSE-R-IN----AV---------D-----------Y------R---------N-----------K---------------  268
C.ZA.04.04ZASK164B1     .VTKNNTSNGT-TPSTTTPTPN-DT-GM-------T-EL-D-EK-VN-L--R---V-LKTNE..........................PNS-D-R-IN----T-A-----I--D-----------Y-----S----------K-----------K---------------  279
C.ZA.05.05ZAPSK240B1    ....................NTDMEK-MR-----AT-E--D-Q--A--L---P-LV-LEEN.............................SS--R-IN----T--------T-D-----------Y----------------L-----------K---------------  257
C.ZM.02.02ZM108         .....TSSVKENT--NAIFNITEI-K-MS-----VT-EL-D-KKQ---L--R-----LNR-ENTSD...................NSTETPKE-R-IN--S-T----------D-----------Y----------------N-----------K---------------  282
D.CD.83.ELI             .............NGTMGNNVTT-EKGM------VT-VLKD-K-QV--L--R---V-----SS.........................TNS-N-R-IN----A----------------------------RD-K-----------------------------------  268
D.CM.01.01CM_0009BBY    ..............STETELNNSST--M-------T-EL-D-KKQ---L------VE--KKDNST.....................DNSTDGT-R-IN----A-----------------------------D-R-------------------K---------------  272
D.KE.01.NKU3006         .ETMKNETMKNE-MKNLNDTIKNLNETMT----RVT-EL-D-EKQVH-L--R--VVKT-GG.............................NN--R-IN------------IT------------------------------K-----------K---------------  283
D.KR.04.04KBH8          ..................TINNNST--M------VT-VV-D-KKQ---L--R---V-M---.............................N---R-IN----TV------I-------------------------------------------K---------------  259
D.TD.99.MN011           .................VTVSNSSL--M-------T-VV-D-KKQ---L--R---VEMGDNE...........................TSGT-R-IN----A-----------------------G-----D-K-------------------K------F--------  263
D.TZ.01.A280            .........NSG-GTNTTDPRLI----M-------T-E--D-RKQVQ-L-----VV---KK............................NNN--T-MH----A-K--------------------------KD-K------KK--------------------------G  269
D.UG.99.99UGK09259      ........DATNKNATNNTSNVTNEIGM-I----VT-EV-DRKKQVS-L-----VVQ--DT.............................N---R-IN----A--------T-------------Y----------------Q-----------K---------------  269
D.YE.01.01YE386         .........NT-SG-NTTNNITN-DIGM------VT-EL-D-KKQVH-L-----VV----NN...........................SAD--R-IN----A--------T-------------Y------D-R-----L-R---------------------------  271
D.YE.02.02YE516         .................TVDNNSTRE-M------VT-VV-D-QKQ---L------VAM--...............................ST-R-IN----T-----------------------G-----E-NYT-K---N-----------K------F--------  259
D.ZA.90.R1              ...............-G-TTTEKPTL-VQ-----VA-VV-D-RKEAS-L--RP--V---Q-NK........................NGNK-T-R-IN----S----------------------------RD-K-------------------K---------------  273
F1.AR.02.ARE933         ...............TTNNTLAD-P-A-Q-----VT-EV-D-KL-VH-I--R---V--SGSN..........................DNSST-R-IN----T---------WD-------T---Y-----XD-R-------R-----------T------S--------  265
F1.BE.93.VI850          ....................SQEKP-AMQ-----MT-EV-D-KL-LS-L--R---V--G-N............................NSSE-R-IN----T---------WD-----------Y------D-R-------K-----------K---------------  258
F1.BR.01.01BR125        ..........QNS-LKEGDALKE-P-A-Q-----MT-EL-D-QM-VH-L--R---V--S-ISN.........................NSNGE-R-IN----T---------WD-----------Y-----KDEK-------K-----------K---------------  270
F1.ES.x.P1146           ...............EQNSTLKE-E-A-------MT-VVND-KL-VQSL--R---V-LN-NNKNISNND.............TTSNNTSSYEE-R-IN----T-K-------WD-----------Y------EEE-------R-----------K--------------G  285
F1.FI.93.FIN9363        ..........NDTLSDQ-STLKE-P-A-Q-----MT-EVED-K--VH-L--R---E--S-NN...........................SREE-R-IT----T---------WD-----------Y-----KD-R-------R-------------------------S-  268
F2.CM.02.02CM_0016BBY   ..........AN---QPNNITLE-Q----------T-E-KD-RKN-F-T---H--V--EKN.............................----R-------TV---------D----Y------Y------D-R---K-L-------------K---------------  267
F2.CM.95.MP257          .........VTSSN-TTLAPNVTISE-M-------T-E--D-QK----L----VVQ-N-S...............................N--R-IN----TL---------D--------------------------L-R-----------K---------------  268
F2.CM.97.CM53657        .........ITNGNSTLDNITLE-Q----------T-E-NDIKK--S-I--R--VV--N-S.............................-SE-R-L-----TV---------D------------------D-E-----L-R-----------K---------------  268
G.BE.96.DRCBL           .................-T-N-T-EYRMT-----MT-EL-D-KKA---L--RT-VV--NEMNNEN.....................NGTNS-W-R--N--V-T-K------T-------------------VD-K-----T-N-----------K---------------  267
G.CM.01.01CM_4049HAN    .........TNGTIDN-NNITVDG-E--T------T-E--D-IRQ---L-----VV--N--.............................NR--M-IN--V-S-K------D-------------------RD-N-----S-R-----------K---------------  268
G.CU.99.Cu74            .............NSTGRNCSVEKGE---------T-E-KD-KKQ---L--R---V--NDNRDNST...................VSNSSPEN-R-IN--V-A-K--------D-----------------RD-E-------N-----------K---------------  274
G.ES.99.X138            ............S--KTDNSSTVD---L-------T-E--D-KKQ---L--R---V--TDNGN.........................SSADN-R-IN--V-T-K--------D-----------------RD-X-----S-K-----------K--I------------  269
G.GH.03.03GH175G        ...NCTRKHPNGTVDYN-TVDNSSCEI--------T-EL-D-SK----L--RP-VV---GNDN.........................STYSD-R-IN--V-Y-K------N-D-----------------GD-K-----L-N-----------K------------VT-  278
G.KE.93.HH8793_12_1     ...............TVANESVSAQ----------T-E--DRKR----L--R--V---ND-SSN......................STGNYSN-R-IN--V-T-K------D-D--------G--------KE-E-------Q-----------K---------------  268
G.NG.01.01NGPL0669      ........CTKN-NTG-TNVTCNKEV-M-------T-E--D-QK----L----MVS-NDGDSNA......................TNYNA-N-R-IN--V-T-K---------------------------D-N-----T-K---S-------K------------V--  277
G.PT.x.PT2695           GSTTSTSTVGS-AST-TV-STAGYRE-L-------T-E-KDRRKQ---L-----V--NDGRN..........................SSANN-R-IN--V-T-K--------D-----------------RD-E-----L-K-----------K---------------  296
G.SE.93.SE6165          ...GNKRNNSTD-STETNNSTVDNP---------VT-E--D-KK-------R--VV--N-..............................A-N-R--H--V-T-K------T-D---------T-------RD-E-------K-----------K---------------  272
H.BE.93.VI991           .........NV-KS-N-TDIN-G-IQ-QR-----VT-A--D-N--VH-L--RA--VQ--EGER........................NKSDNH-R-IN------K---------------------------G-K-----------------------------------  273
H.BE.93.VI997           .........DTNSSSTVNATSSPSAN-LT-----VT-V--D-Q-RVH-L--R--VV---ETSNN......................NNSNS-K-R-IN---------------D----------------------------------------K---------------  272
H.CF.90.056             ..............T-N-TSSMEAG--LT-----VT-VL-D-Q--VH-L--R--VV----N.............................S-Q-R-IN------------------------------------------L-----------------------------  263
J.CD.97.J_97DC_KTB147   ..............TT-PSPT-TNPLGM-------T-E-GDRRK----L---Q-VVS---S.............................NS--I-IN------K---------------------------DNK-------K-----------K------------I--  270
J.SE.93.SE7887          .............S-TT-NSSVSSPDIMT------T-E--N-RKQ---L--RQ-VV---SN.............................NKN-I-IN------K--------Q------------------D-N-----S-K-----------K------------I--  264
K.CD.97.EQTB11C         .......RTNANKNDTNINATVTSTD---------T-ELKD-KKRVS-L------VQ-KQSEI........................NQSESEDR-IN----TV-----------------------------N-C------------------K---------------  275
K.CM.96.MP535           ............-STNNATSTVVSPA---------T-E-KD-KK--S-L--R--VL-LNGEGN.........................NSS-E-R-IN----T---T----T-------------------KD-R-------K-----------K---------------  269
```

```
                     glycosylation NIS
  glycosylation NDT  glycosylation NCS                                          glycosylation NVS
                                                                                glycosylation NGT
        V1                          V2                                          
  glycosylation NSS            glycosylation NDT    glycosylation NTS    glycosylation NKT    glycosylation NGS

B.FR.83.HXB2                 ...........TNTNSSSGRMIMEKGEIKNCSFNISTSIRGKVQKEYAFFYKLDIIPIDND............................TTSYKLTSCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAE 267
01_AE.CF.90.90CF11697        ...........DTY-GTAKLNDTIGD-VR-----VT-EL-D-K-EVH-L--VP--VR-GEKNKN......................SSGNSSE-I-IH------K------T-D-------T---Y------E-N-------K---S-------K--------------- 273
01_AE.CN.05.FJ051            .........NSKSITV-NNTVGNITDDV---T--MT-EL-D-K--A--L--R--LV----.............................N-E-R-IN------K-----I--D-------T---YV-----D-N-------K---S---A------A------------ 267
01_AE.CN.06.FJ054            .........KNNQ-GEYIEIIGNMTE-VR-----MT-GL-D-Q--VH-L--R---A--NENNS........................NENN-E-R-IN------K-P---I--D-------T---Y--F---D-N-------K---S-------K-------------S- 273
01_AE.HK.x.HK001             .......SNNT-QNIN-XVTIGNMTD-VR-----MT-ELTD-Q--V--L------VE-N-...............................ST-R-IN------K--------D-------T---YV-----D-K-S-----N---S-------K--------------- 268
01_AE.TH.01.01TH_R2184       .........KVS-PTDVPSIVGNVTD--------MT-EL-D-K-SVH-L--R---V-MKANS..........................SNS-E-R-IN------K-----I--D-------T---Y---R--E-N-T-----K---S-------K--------------- 271
01_AE.TH.02.OUR769I          ..........NVKSTNEAPSIGNITD-VR-----MT-EL-D-KK-VH-L--R---VQ-E-N............................SSE-R-IN------K-----I--D-------T---Y-L----D-K-----L-K---S-------K--------------- 267
01_AE.TH.90.CM240            .........NGSSKTNV-NIIGNITD-VR--T--MT-ELTD-K--VH-L------VQ-EDKK...........................TSSE-R-IN------K-----I--D-------T---Y------D-N-------K---S-------K--------------- 270
01_AE.US.00.00US_MSC1164     ..........NV-ITKD-NKIGNLTD-VR-----MT-EL-D-KH-VH-L--R--LV--E-N............................NSSE-R-IN------K-----I--D-------T---Y------D-N-------N---S-------K--------------- 268
02_AG.CM.02.02CM_4082STN     ................NYNISDDMRE---------T-EL-D-KK----L--RQ-VVQ--SE.............................VGQ-R-IN----A----------Q-----------------RD-S-----S-K-----------K------------T-- 260
02_AG.EC.x.ECU41             ....................TEEMRE--------MT-EL-D-K--VS-L--R--VVQ-N-SQ...........................DNSH-R-IN----A--------T--------------------------K---K-----------K--I----------S- 258
02_AG.GH.03.GHNJ196          ............SE-KIGNIS-EMQ---------MT-EL-D-NR-MH-L--RQ--V-MNESLVSI....................NTTNS-DQ-R-IN----TV----------------------------D-N-----L-R-----------K--------------- 273
02_AG.NG.01.PL0710           ................N-TFNNSMV--------XMT-VL-D-K-NVH-L-----VVQ-NSNS...........................NSSE-R-IN----A-------IT--------------------D-E-S-----K-----------K--------------- 262
02_AG.NG.x.IBNG              .............YSN--NLTSDMN----------T-EV-D-KK-MH-L--R--VVQ-NEN............................NGSQ-R-IN----A----------------------------KD-G-------K-----------K--------------- 264
02_AG.SE.94.SE7812           ................-GNISENMQ---------MT-EL-D-K--V--L--RY-VVQ-NETG...........................DNIQ-R-IN----A-----------------------------DEK-------K-----------K--------------- 262
02_AG.SN.98.MP1211           ................RNNSK-N-VQ-M------MT-VLKD-KK-MA-L---I--V---KN.............................A-Y-R-IN----A-----------------------------E-D-S---S-K-----------K--------------- 261
03_AB.RU.97.KAL153_2         ..............VT-TNTSSIKMM-M-------T-DL-D--K----L-----VVQ---...............................D--R-I------V------I---------------------D-K-------------------K--------------- 258
04_cpx.CY.94.CY032           ................N-TNGTVI-EG------D-T-E--D-KK----L--RI--V--NARVPING..................SNRNNS-EE-M-IN--A-T-K---------------------------E-N-T-L-------S-R-----K--------------T 272
05_DF.BE.x.VI1310            .........STA-STTN-TTLKE-T-AVQ-----MT-EVND-KL-VH-L--R---V--SS-D..........................SSNS--R-IN----T-K-------WD-----------Y------E-K-S-----K-----------K--------------K 271
06_cpx.AU.96.BFP90           .NKTLGNNSTNSTLGNN-TIVDDISK---------T-E--D-TK----L--RP--V--GD-S...........................NNSD-R-IN--V-T-K--------------------------RD-D-------K-----------K--------------- 278
08_BC.CN.97.97CNGX_6F        ............GNGTYNETYNESVK--------AT-LL-DRKKTV--L--R---V-LNDENSG.......................ENSSEY-R-IN----A--------T-D-------T---Y------D-I-----Q-H-----------K--------------- 271
09_cpx.GH.96.96GH2911        ...............KPNTTSDGSEVRMQ------T-EL-D-H--V-SL--RS--VS-NAS.............................NNE-R-IN----A--------T-------------------RD-E-----I-K---S----------------------- 262
10_CD.TZ.96.96TZ_BF061       ...................NDTVGDPRMT-----TT-EVKD-KKQ---L-----VVQ-GDS............................N---R-IN----A--------T----------------I---DRK-------K-I---------K--------------- 258
11_cpx.GR.x.GR17             .....................RDNATD----T---T-ELED-KKN-R-L--R--VV--NDSS..........................SNIGQ-R-IN--V-AVK---S--T-------------------RD-E-------K-----------K--------------- 255
12_BF.AR.99.ARMA159          .......ANNA-ANGTQNNTLEE-P-A-Q-----MT-E--D-QL-VH-L--R---V--G-G.............................SET-R-IN----T--*------WD-------T---Y-L----D-K-------Q-----------K--------------- 269
13_cpx.CM.96.1849            ..............S-INSST-QGRE------Y-MT-GVGDRR--V-SL--RV--E--EGNSNS......................GNSSSSE-R-I-------K------T--------------------DEE-------R---S-------K--------------- 270
14_BG.DE.01.9196_01          .............ESGN-TTNSSSW--M------VTSIT-D----Q--Y--N---VQ--SES...........................NS-K-R-I-----I--------T--------------------D-K----------------------------------- 266
15_01B.TH.99.99TH_MU2079     ................TNINRTI-G---------T----T--.-D--L--S--VEA---...............................---R-RK----I------------------T------------K----------------------------------- 258
16_A2D.KR.97.97KR004         ................I-STQSPDSNNTM----ETT-EL-D-K--VQ-L--E---VQLNSSDSN.......................DTLNRQ-R-IH-D--T----------------------------KDPE-------K---S----------A------------ 267
17_BF.PE.02.PE02_PCR0155     .KSASAKNVST-ASXANDS#TLE-P-A-L-----MT-EV-D-QL-V-SL--X---V--S-N.............................GSE-R-IN----T---------WD------------------D-K-------KD----------K----LIVV.------ 274
18_cpx.CU.99.CU76            ...........G-NTXA-NSS-#ATE-M-----F-T-EV-X-QKQ---L--X---VS-VGXVf.........................SNYSD-R-IN--V-T-K--------D-------------------QP-------R-----------K--------------T 268
19_cpx.CU.99.CU7             ..........................HM------MT-EL-D-K-QI-SL------V--GD..................................-IIN----A--------T--------------------D-E-----K-R-----------K--------------- 240
20_BG.CU.99.Cu103            ...................STVTSSED-NI---Y-T-E-KDRKKQ---L--R--VV--NDKNSN......................NSNNYSD-R-IN--V-A-K--------D-----------------RD-E-------K-----------K--------------- 265
21_A2D.KE.91.KNH1254         ......SVNSTNSSV-TNSTSEL-AAGMQ-----VT-VV-D-QKQ---L--RP--V----N............................DNS--R-IN----A-------I--------------------RD-Q-------K-----------K--------------- 272
22_01A1.CM.01.01CM_0001BBY   ...............ITNSSVEEL-T-VT-----MT-EL-D-RREV-SL--R---VG-NG..............................NN-FR-IN----A--------T-------------------KE-E-------N---S-------K--------------- 260
23_BG.CU.03.CB118            ...........-RNCTTTNGTVNNSAD--------T-E--D-SK----L------M---DKNNG......................SNKNYSD-R-IN--V-T-K--------D-----------------RD-E-------R-----------K--------------- 273
24_BG.CU.03.CB378            .........QTK-CTEE-NSTVINCK---------T-ELQD-KKQ---L---I--V--N-SSS........................NNNYSD-R-IN--V-A-K--------D-----------Y-----RDEE---K---R-----------K--------------D 273
24_BG.CU.03.CB471            ...........-RNCTEGINRTTNYED-------AT-E-KDRQKQ---L-----VVQ-NDKDNSS.....................NNNNYSD-R-IN--V-T-K--------D-----------------RDAE-------K-----------K--------------- 274
25_cpx.CM.01.101BA           INNSNCSHSINNS-EKNNNTIANGCQR--------T-ELKD-KK----L-----VV--EDGN..........................NSND--R-IN----A-K------T-------------------KD-K-------K-----------K--------------- 282
27_cpx.CD.97.97CDKTB49       ..............-NHTTINDNM---L------VT-A-KD-Q--VH-L-----VVQ-G--TN.........................TTGAE-R-IN------K------T----------------------------------S-------K-------------S- 267
28_BF.BR.99.BREPM12609       .......TNMNATNTTGNSSVEAME--M------VT---GD-M-----L-----VV--A--.............................NA--T-IN------------------------------L----------------------------------------- 270
29_BF.BR.02.BREPM119         ............TNTN--DWGR-----M-------T-DL-D-K-----V-----VT--G--.............................N---R-I------L------T---------------------D-K---S------------------------------- 264
31_BC.BR.02.110PA            .ANKTNANQTNA-QT-NNSST-DM----------MT-EL-D-IR-VH-L--R---V-LNENN..........................SSSGD-R-IN----A----------D-----------Y-----K--N--R----N-----------K--I------------ 278
32_06A1.EE.01.EE0369         ......NATINGTSEGTKNITVNN-E-MT-----VT-E-ED-KKT---L-----V---EGKNN........................NGNSSA-I-IN--ATN-E--------------------------RD-N-------K-----------K-------------G 276
33_01B.MY.05.05MYKL007_1     .........NV-RQTNAPSILENVTE--S-----TT-EL-D-K--V--L------VS-GGNK..........................SNSSE-R-IN------K--------D-------T---Y------D-N-------K---S-------K--------------- 271
34_01B.TH.99.OUR2478P        ..................TNGPNITE--R-----MT-E--D-K--V--L-----LVQ-NGS.............................GGE-R-IN------K-----I--D-------T---Y------D-N-------K---S-------K--------------- 263
35_AD.AF.05.05AF095          ....................TENMN----------T-EL-D-R--V-SL-----VVQ-ED-............................NSS--R-IN----A--------T--------------------ENK-----S-K---S-------K--------------D 259
36_cpx.CM.00.00CMNYU830      ..................NVSKDMQ---------TT-VL-D-I--VHSL--R---V---ASN..........................NDSSR-R-IN--A-A--------T-------------------KD-Q-----L-K---S-------K--------------K 265
37_cpx.CM.00.00CMNYU926      ................#KSLDGG-EI------Y-MT-EL-D-K--I-SL-----VV--GDN.............................NET-R-IN-------------T------Y-------------D-EY------R--------------------------- 264
38_BF1.UY.05.UY05_4752       .........SQNG-SN--STLKEDI-V-Q-----MT-EV-D-QL-V-SL--R---V--S-GNNGS.....................RSNSSGE-R-IN--S-T---------WD-----------Y-------N--------------------K-M-----------S- 275
39_BF.BR.03.03BRRJ103        .............ND-QNSTLKE-A-ALQ-----MT-E--D-Q--VH-L--R---V--SDNSNG.......................SNGSGK-R-IH----T---------WD------------------D-K-------K-----------K--------------- 270
40_BF.BR.04.04BRRJ115        .........VHN--QGIDSTLKE-PEA-Q-----MT-EL-D-Q--VH-L------V-LNS-GN........................NSSEKY-R-IN----T---------WD-----------Y----------------------------K--------------- 273
42_BF.LU.03.luBF_05_03       ................XXXXXXXXXX--------VTLG-KS-M-----L--RS-VV----NS...........................NS-N-M-IX----XL------I---------------------D-K-------K------------------------X-- 268
43_02G.SA.03.J11223          ................KTNVKCEDV-NL------VT-EL-D-KK----I-----VV---GNSN.........................NESGD-R-IN--V-T--------T-D------------X----KE-N-----Q-K---S-------K--------------- 265
U.CD.83.83CD003              ..................TNNNT-EAT-T----KVP-ELKD-TETVHTL-----VV-LNVTNN........................SSISST-R-IN----T-----------------------------D-K-------K-------------------------S- 264
U.CD.90.90CD121E12           .....................NTAAN-LT----AVT-ELKD-KKNVH-L-----VV-LNETN..........................SSYGT-R-IN----T-----------------------------D-E-------R-------------------------S- 259
U.GR.99.GR303                ...........ALNC-NVNATDEMR--M----Y-MT-EM-D-K--V-SL--R--VVQMGS-NS........................KSNYSE-R-IN-S--A--------T---------T-------R--EEK-------K---S-------K--------------K 271
U.NL.95.U_NL_95_H10986_D1    .............---GNNATVE----L-------T-VLQD-KR----L--RF-VV-LNS-NN........................SSNYSE-R-IN--I-T-K------D-------------------KD-R-------K-----------K--------------- 269
CPZ.CD.90.ANT                ...PTTPSTTTSTVTPKTTTP-VDGMKLQE-N--Q--GFKD-K--MK-I---G-LMKCQDNN...........................E-NC-Y-WH---TT---S-E-ST-------------Y---R-EDED-T-V-M-K---V-H-----S-M-A-W-----TYQT 267
CPZ.CM.01.SIVcpzCAM13        .NSTSNSNNNP-LK-NT--KNETGVLQM---T--TT-EL-D-KKQV-SL--VD-LQSLGSG............................TGDT-TMIN---TA-----------------------------DVN-S-K-K-R-----H-----K---T---II-----T 277
CPZ.CM.05.SIVcpzEK505        ................NNSVNQTNHVQMQ-----VT-EL-D-KKQV-SL--MG----L-TNNS........................SGNNSQ-R-IN---TAV------I-------Y----P----I---DQD-----E-N-----------K--I----I------T 260
CPZ.CM.05.SIVcpzLB7          .................NSSINSSNETM------TT-EL-D-KKQT-SL--VD-LTQ-NKT.............................-E----IN---TA-------I-----------------L-R-----Y--S--N-----H-----K--I----I------- 261
CPZ.CM.05.SIVcpzMB66         GNTTNCTVQIS-GND-TANNITVGTIDMY-----AT-EL-DRKKQV-SL--RQ-LE-LEGNKPP......................EGDKNAL-R-YN---TAM----S--L----------------L---D-N-T-I-Q-K-----H------------F----T-E- 291
CPZ.CM.05.SIVcpzMT145        ...............-CTNSTDDRL-DMR-----VT-EL-D-KRQV-SL--VE--TA-G-..............................NST-R-IN---TA--------T----------------L---DIDYK-NET-K-----H-----K--AT---I----T-D 265
CPZ.GA.88.GAB1               ............QAKNLTNQTSSPPL-M------VT-EL-D-KKQV-SL--VE-VVNLG-E.............................NNT-RIIN---TA--------T--------------------D-D-S-K-K-------H-----K---T----I------ 265
CPZ.TZ.00.TAN1               ...NKTLNSAT-TLTPTVNLSSIPNY-VY-----QT-EF-D-KKQI-SL--RE--VKE-GN.............................NN--Y-HN--------E-D-ST------R---------L--RDQN-T-K-Q-S---V-H-----Y-MIA-A-H-----E- 266
CPZ.US.85.CPZUS              ..................SSAVNTTDIMR------T-EL-D-KKQV-SL--VD-LAH-N-...............................NT-R-IN---TA--------T-----------P--------E-D-K-K-E-K-----------K---T---II-----T 252
CPZ.CM.x.SIVcpzDP943         .............STGNDT-NSSTSQMV-K-F--MT-EL-D-QKQV-SL--VD--VX-NG..............................XX--R-IN---TA--------T-----------P----M---EPD-----K-R-I---------K---T-X-I------- 257
CPZ.CM.x.SIVcpzMB897         ......................EASQVQY-----VT-EL-D-KKQV-SL--RE--TSL-SNKT.........................VKNGT-R-IN---TA--------T------Y---------L---DQN-K-K-T-R-----H-----K------F-------- 260
N.CM.02.DJO0131              ........RTTDEKEKPGNGTDL-ARHM-------T-E-HD-KKQA-SL--VE-VV-LNDG............................NNST-R-IN---TAV------TT-----------P----M---EAN-----E-K-----------K--I----I-----DK 263
N.CM.04.04CM_1015_04         ............-XTVDN-ELDIGYKQM------VT-ERKD-KKLA-SL--AE-VVQLNESD..........................STNQT-R-I--K-TSV------TT-----------P----MX--EGN-S-K-E-K-----------K-TI----IX----DT 261
N.CM.04.04CM_1131_03         ............-ETVGN-TLDIGYKQM------VT-EREDRKK-A-SL--TE-VVQLNDSD..........................SDSNT-T-IN-K-T-VR-----TT-----------P----MR--EGN-S-K-E-K-----------K--I----IM----D- 273
N.CM.95.YBF30                .........ERN---MTTREPDIGYKQM------AT-ELTD-KKQV-SL--VE-VV--NAY.............................NKT-R-IN---TAV------T------------P----M---EGN-S-N-S-------------K--I----I-----NT 270
N.CM.97.YBF106               ..........SAGNSTTNRTEDL-DRQM-------T-E--DRKKQV-SL--VE-VV--KDGT...........................DNNT-R-IN---TAV------TT-----------P----M---EGN-S-N-S-------------K--I----I-----DT 262
O.BE.87.ANT70                ......................TNENLM-K-E--VT-V-KD-KE-KQ-L--VS-LMELNETSST......................NKTNSKM-T--N--STT----------------------Y--F---STE-----T-R-ITV-T-------T-----I---T-SK 259
O.CM.91.MVP5180              ................KT-LLNETIN-MR-----VT-VLTD-KEQKQ-L--VS-LSKVNDSNA.........................VNG-T-M--N--STI-K------------------T-Y--F---DTD-----L-H-I-V-T-----K-T-----I---T-SR 263
O.CM.96.96CMABB637           ......TSNNTPR--TATSTPGI-EMKV-K-E--VT-ILKD-KE-KQ-L--VS-LTKVN-SETG.......................NNTEPI-T-IN--STT-K------N-------------Y--F---DTA-----L-K-I-V-T-----K-T-----I---TIS- 274
O.CM.98.98CMA104             ...............NTNTTTHNPSPDM-Q-N--VT-V-KD-KE-KQ-L--VS-L-RFNETEN........................NKTREI-T-IN--STTT--------------N------Y--F----IE-----T-N-I-V-T-----K-TI----I---T-SK 264
O.CM.99.99CMU4122            ..................TNATAASSAV-K-E--VT-VLKD-KE-QQ-L--RS-LNK-E-ST...........................TS-M-T-IN--STT-K------T-----------G-Y--F---DTE-----Q-R-ITV-T-----K-T-----V---T-S- 258
O.FR.92.VAU                  ............I-TTN-PLNSNNTK-V-Q-D--VT-VLKD-QE-KQ-L--VT-LVK-NATS...........................NE-M-R-IN--STT-R--------------------Y--F---ETG-----L-K--TV-T-----K-T-----I---A-SK 264
O.SN.99.SEMP1299             .........NY--NS-INNDTSSPENLV-Q-E--VT-VVKD-KE-KQ-L--VS-LMK-NEAN..........................DTKDM-T-IN--STT-K--------------------Y--F---STE-------N-ITA-T-----K-T-----I---T-S- 268
O.US.99.99USTWLA             .........TTNT-TAKTTNPEE-TNPV-K-D--VT-VVKD-QE-KQ-L--VS-LLK-G-G.............................-NI-T-IN--S-T-K------T-------------Y--F---ETG-------K-I-V-T-----K-T---H-IF--TISK 265
```

```
                     glycosylation NNT
                 glycosylation NCT   V3 tip                                                                                glycosylation NST      glycosylation NST
   glycosylation NFT                          V3                                                        CD4                  V4
              glycosylation NTS                         glycosylation NNT     glycosylation NKT     CD4        glycosylation NST
                       *                                  *                                            *     *
B.FR.83.HXB2          EE.VVIRSVNFTDNAKTIIVQLNTSVEINCTRPNNNTRKRIRIQRGPGRAF....VTIG.KIGNM.....RQAHCNISRAKWNNTLKQIASKLREQFGNNK.........TIIFKQSSGGDPEIVTHSFNCG.GEFFYCNSTQLFNSTW............FNSTWSTEG 404
A1.GE.99.99GEMZ011    KK.-M---K-I---G-I------KP-N-T-I--G--I-TS-..RI---QT-....YGT-DV--DI.....-K-Y--VN-TE-----QK-SIQ--NY-N-K..........----RN-----L-VT-------.-------T-G------................DGNDTV 399
A1.KE.00.KER2008      KR.-M---A-I---T-N-----KDP------------RGV..HI-L--R-....Y-T.QVV-DI.....---Y--V-KS--DD---KVVYQ--KY-N.K..........----NS-----V--T--M-T--.-------TSG------.................PVNQE 403
A1.KE.00.KNH1144      -K.-K--TE-I-N---------VEP-R----------ESV..RI---Q--....FAT-DI--DI.....------V--SQ--K--Q-V-EQ---H-K-K..........----NS-----L--T-------.-------TSG------................NT-MS- 408
A1.KE.00.KSM4024      .-.-K---E-I-N---N-----DKP-K-----------SV..Y...........A-DI--DI.....------VNASD-SEA-RKVVLT-Q.....................#--L-VT-------.-------TSG------...............-WND-AN 374
A1.KE.00.MSA4069      Q-.IA---E-I-----------VNP-R----------RD-..GI---QT-....YAA-AI--DI.....---Y--V---A--K--QEV-KQ--QE-N-K..........-R-ANH----L--T-------.-------TSG-----Y.............NGTDIA 401
A1.KE.00.NKU3005      NN.-T---A-IS----I--I----P-T-----------SV..RI---Q--....Y-TD.I--DI.....------V-KSE--K--Q--VIQ--KHWN.K..........-N-TEP----L--T-------.-------TSD------...............PVNASME 407
A1.RU.00.RU00051      GN.-T---E-I---T-------SKP-N-T-I--G----TS-..RI---QT-....YAT-DI--DT.....-K----V--TE-G-ATEE-RKQ-QTY-P-K..........-S--N-----L--T-------.-------T-G---NASNN.................. 394
A1.RW.93.93RW_024     RN.-T---E-I-N-----L---VKP-R------Y---FR-V..RI---QS-....HASDNIV-DI.....-K----V--SE--E--Q-VVTQ--GY---K..........-V-TKP----L--I------R.-------TSS------.............-N-A-MQK 403
A1.SE.95.SE8891       T-.-M--AE-I-N-I-N----F-K----I-I--------S-..RI---Q--....YAT-DI--DI.....---Y-DVN-TE--EA-QKVVNQ-KTH-K-K..........----NS-----L--T-------.-------TSG------.................NG-DS 393
A1.SE.95.UGSE8131     --.-R---E-IS---------FTKP---I---------S-..RI---Q--....YGM-DI--DI.....-K----V--S---E---KV-IQ--KYWN.T..........----TN-----L--T-------.-------TSG----S-.............NEND-KV 411
A1.TZ.01.A173         G-.IQ---E-I-N---N----FTEP-Q-I-I---------V.HI---QT-....YAT.DI---I.....---Y-TVN-TE--T--QKVGKQ--DLYF------..........----TS-----L--T-------.-------TSG---GLFNGTW............NGNH 406
A1.UA.01.01UADN139    N-.-M---E-I---G-N-----TEP-N-T-I--G----TS-..RI---QT-....YAT-EI---P.....-K----V--VA-----QK-SIQ--KY-N-K..........-T--N-----L-VT-------.-------T-E------................GENGTI 402
A1.UG.92.92UG037      GK.-M---E-I-N-V-N------E--T-----------RSV..RI---QT-....YAT-DI--DI.....-------V-GSQ--K--H-VVEQ--KYWN--..........----NS-----L--T-----A.-------TSG------..................VN- 398
A1.UG.99.99UGA07072   G-.-M---E-I-N---N------KP-N-T-I-------QSV..RI---Q--....YAT-NI--DI.....---Y-DVN-TE--K--QEV-YQ-GKH---K..........----AN-T---L--T-------.-------T-N---NI-...............SNASM 403
A2.CD.97.97CDKTB48    GK.-M---E-I-----N----F-KP-P-----------S-..RF---Q--....Y-NNNI--DI.....---------ITE--A---KVVEQ---H-P-K..........----NS-----L--T-------.-------T-G------...............ENG-NK 392
A2.CY.94.94CY017_41   GGKIM---E-I-N---N----FTKP-L-T-I-------S-..RF---Q--....Y.TNEI--DI.....--------NKTL--D--QKV-EQ---K-PKK..........----TN--------T-L----A.-------T-G---G--...........W-NGTWNGP 407
A.SN.01.DDI579        --.-M---E-L-N---------VNP-Q-S--------TSV..RI---Q--....YAT-DI--DI.....-------V--EN-TE--Q-V-KQ-KRH-N.K..........----TG-----L--T-------.-------TSG------..............-D-AINDT 402
A.SN.01.DDJ369        GK.-M---E-L-N---------DEP-Y-----------SV..HI---Q--....YAT-DI--DI.....-------V--ED--K--YKVVEQ--KH-N-R..........-N-TAP----L--T-------.-------TSG-----Y..........NSTWNN-DS 407
A.SN.96.DDJ360        --.-M---E-L-N---------V-P-Q----------TSV..RI---Q--....YAT-DI--KI.....-----TVR--Q--T--Q-V-DQ--KH-N.K..........-N-TRP----L--T-------.-------TSG------...............NSTTIKQ 396
A.CD.02.02CD_KTB035   GK.-M---E-I-N---------ET-N-T-L-F------SV..RI---QV-....YAT-GI--DI.....-K-F-IVNATE-DT-ISRVREQ--AH-P-K..........-V-AN-T---L--T--T---.-------TSG-N-NS-..................GT 408
A.CD.97.97CD_KCC2     GG.IK---A-ISY---N-----DIP-K---S------TSV..RI---QT-....YAT-DI---I.....---------L--TA--D--YNVSKA---H-P-K..........----NK-----L-VTQ-M-I--.-------TSG----S-...........APNG-WIN 401
A.CD.97.97CD_KTB13    -T.-I--TE-IS--G-N----FVDP-R---I------TSV..RI---QT-....Y.TSDI--DI.....---Y-EVNKTQ--M--QKV-DQ--KF-P-K..........-N-N-T---L-VT--I---R.-------TSS----S-.............PIN--TS 425

B.AR.04.04AR151516    -D.I----E---N---I------ET---------S-----SV..HI--------....YAT-EI--DI..........-------E------R-L-R------NTT..........-V--NH-----------T----.-------T-------...................REDNS 397
B.AU.87.MBC925        G-.-I---E-I-N---------KDP-------------S-..HI---------....YAT-DI---I.....---Y-TLN--R--D------E--G---K-K..........-V-N---------M------.--------------G--.............-D-ERSWS 407
B.BO.99.BOL0122       --.I--K-E-------------K-IS------------S-..NI-----L....YAT-EI---I.....---------ATE-----E--VT--G----V--..........----N---------M------.-------T-E-----N..........FTSTWNM-S- 397
B.BR.03.BREPM2012     --.I----K-------I---------K---------S-..HM-W-----....YAT-EI---I.....---------GTN-T---R-V-Q--E-FN.-K..........-V-SP--------AM------.-------T-P---NSKWVNG...............- 407
B.CA.97.CANB3FULL     --.-----A-----T-------ET--------------S-..HIA-------....YAT-EIT-DI.....-K-----------A-E--VK--G---K-K..........-V-N---------M------.-------T-------...............VN--WN 398
B.CN.05.05CNHB_hp3    --.-----K-------I-----KQ--------------..TM----VL....Y-T-QIV-DI.....-K----L--PQ-DD--R-VVG--------K..........----N--A-----V-M-IS---.-------TSR------YNNGT.............WN 406
B.CO.01.PCM001        --.I----E--------------V-------------G-..HL---K--....YAS-DIV--I.....------TLN-T--EK--G-V-N-----YK-K..........----N------L-T-M------.-------T-------..............NENSTWNGT 397
B.GB.83.CAM1          K-.-----E---N---------KEP-------L------S-..AI----TV....YATDRI--DI.....------L-ST---------VT--K-----K..........----N---------M------.-------T-----T--...........LFNGTWNDT 403
B.GB.86.GB8           -K.-----D--M--T-------KEA-------------G-..YI----R-....Y-T-RI--DI.....--------KE------H--VIE--K--R-K..........-V-N---------M------.------KTA-------.................NSTGN- 394
B.GE.03.03GEMZ010     --.-I---E-----V--------E--Q-----------S-..HI---------....YAT-EI--DI.....------L-SDQ-S----H--T------K-K..........-A-N---------TMFN----.-------T-----N--..............LNNDT 394
B.IT.05.SG1           --.-----A--S--T--------KT-I------G-----S-..HM-----W....Y-T-EIV--V.....------L--TA--D---LVVE--K---PKK..........-V-NR-A-------M------.-------L-------.................. 374
B.JP.05.DR6538        K-.-----E---N---------KDP---------GY--I--VS-GP-R---I....-ATKKIE-DT.....KK-Y-TLNKTQ--D--R-V-E--------K..........-V-N---------M---T--.------------------.............HGNISNV 412
B.KR.05.05CSR3        -D.I----E-I-N---N-----EL-K------------S-..HI---K--....YAA-DI---I.....---------KS-TD---K--D--S-----K..........-V--P---------M---T--.-------T-----GN-...............SNG-WN 411
B.NL.00.671_00T36     -D.-----K-----T-------KE--------------S-..HI---------....YAT-EI--DI.....------L------D--N--VG----LYK-K..........-V-NS---------M-----R.-------T-------.................DVNAT- 402
B.RU.04.04RU128005    --.--V--R--S---N-----KDP-Q------S-----S-..SI---X--....YAT-DI--DI.....-X---L-G-D-TK--E--VK--X--YN.K..........-V-----------XM------.-------T-K-----...............ENN--NS 398
B.TH.00.00TH_C3198    --.-M---S--S--T-V------E--Q----------I-G-..HI---Q-W....Y-T-QI--DI.....-K---L-STN--------VK------R-K..........-V-----------TML-----.-------TSK------................NTNGT 394
B.UA.01.01UAKV167     GD.I----E-----V--------E-IT-----------G-..HI---GT-....FAT-AI---I.....------L---Q--K--E--VI--GK--NSTT.........I-FDQ-PP------LM-T----.-------TSK------..............NNGTWT-NID 413
B.US.04.ES10_53       --.-I---A-I---T-----H--N--G-T----G----QS-..HIA---------....YAT-AIT--I.....-K------IVS--D----VVE--Q-H-P-K..........-A--N---------M-R----.-------T-D--TWNGTWN.................. 394
B.US.99.PRB959_03     --.-----E---N---------E--I------------S-..NI---------....YAT-DI--DI.....---------EKA--K-----VT-----YE-K..........-V-N--T-------M-T----.-------T-------...................NGTE 395

C.AR.01.ARG4006       --.II---K-I---T-----H--E-------------ES-..RI---QT-....YAT-DI--DI.....-E-----TTA--K--HELSK--KKL-P-K..........-Q-TNH----L--T------R.-------T-------Y..............X-NG--I 402
C.BR.04.04BR013       --.I----K-IS--V-----H-KEP------------ES-..RI---QT-....YAT-DI--DI.....---------K-A--K--QEVST--A-H-P-K..........-K-TK-----I--T------M.-------TSG---G-Y...............NGTESN 405
C.BW.00.00BW07621     -G.I----E-I-N-------H--E----V----G---RSV..RI---QT-....YAT-DI---I.....---------KT---K--SR-GE--K-H-P-K..........-R-NS-A---L--T------R.-------T-R---G-Y..................NST- 390
C.CN.98.YNRL9840      G-.II---E-L-N-V-----H--Q----V--------TS-..RI---QT-....YAT-EI--DI.....---------E----E--QRVGE--A-H-P-K..........-N-TS-----L--T------R.-------TSG---G-Y..................MPN 405
C.ET.02.02ET_288      D-.TI---E-L-N---I------EP-----------QS-..RI---QT-....YAT-EI--DI.....---------GDR--K--ENVKER-QK--P-K..........-L-AP-----L--T------R.-------TSK---T-G..................... 401
C.GE.03.03GEMZ033     --.II---EDINN---I---H--V---------G---QS-..RI---QT-....YAR-DI--DI.....--------NEN---T--QKV-E--KKY-P-K..........-A-NS-----L--T-------.-------TSK--ENRM..............FND-YR 405
C.IL.99.99ET7         K-.I-V--E-L-N---I---H-KNH---V----------SM..RI---QT-....YAT-SI--DI.....---Y-T-NGM---E--QGVGE--QKH-P-R..........-R-EP-----L-VT------R.-------TSN-----Y..................RSN 394
C.IN.99.01IN565_10    K-.II---E-L-N-------H--Q----V----------S-..RI---QM-....YAT-DI--DI.....---Y---.SSN--E--YNVSK--ADH-P-K..........-N-TS-----L--T------R.-------TSK-----Y...............KYTNGT 402
C.KE.00.KER2010       --.IM---E-I-N-V-----H--K-I--V-I--G----SV..RI---QT-....YAT-DI--DI.....-K-Y---KE------Q-VRK--E-H-P--..........-E--PHP-----VA-------.-------TSG---E-L...................TRLN 403
C.MM.99.mIDU101_3     --.II---E-L--HF--------QP---E---------S-..RI---QT-....YAT-EI--DI.....---------KQ---E--QRVGR--A-L-P-R..........-V-NS-P---L--A------R.-------TSG---G-Y.................ND-YR 410
C.MW.93.93MW_965      --.IIV--E-L---V-----H--E----V---------SV..RI---QT-....YAT-AI--DI.....---------TI---K--QGVEK--K-H-P-K..........-E--P-----L--T------R.-------TSN--..-S...................NLF 394
C.SN.90.90SE_364      G-.II---E-L-N-------H--E-I--V-A-------SM..RI---QT-....YAT-DI--DI.....--------.SGN--A--EKVKG--Q-H-PGK.........N-S-EP-----L--T------R.------DTSX---G-T...................... 392
C.SO.89.89SM_145      G-.IM---E-L-N-------H--Q----R---YA-----SV..RI---QT-....Y.TNDI--DI.....---------GD---R--Q-VGK--A-H-P-K..........-K-EP-----L--T------R.-------TSG-----Y..................NPN 396
C.TZ.02.CO178         --.II-S-E-I--S--I---H--E--P-V--------TS-..RI---Q--....YAT-DI--DI.....---------EST--K--QWVGI----YEHF--..........----NS-----L--T-------.-------TSK---G-F.............NGTFNSS 411
C.UY.01.TRA3011       K-.II---K-I---T-I---H--E-------------TS-..RI---QT-....YAT-EI--DI.....---------KKA-KE--QKVRN--A-H-P-K..........-N-TKH----L--T------R.-------TSS-----YN...........TTG-E-SN- 400
C.YE.02.02YE511       --.I----E-L---V-----H-KNP-D-K---------SM..RI---QT-....YAT-SI--DI.....---------EK---D--QE-GK--K-H-P-K.........S-R-EPH----L--TM-----R.-------TSK---R-Y............MFNSTYMPN 403
C.ZA.04.04ZASK164B1   -G.II---E-L-N-------HF-E--A-----------S-..RI---QT-....YAT-EI--DI.....---------KKA--EI-QKVKE--K-H-SDK.........E-N-NH-----L-VTM-T---.-------TS----TSY.................VFNNT 410
C.ZA.05.05ZAPSK240B1  --.II---E-L---G-----H--ET#K-V----A-----S-..RI---Q--....Y.AHDI--DI.....----------TE--K--ERVKK--G-H-P-K..........-K--PH----L-VT------R.-------T-N---...............ENIT-VNI 386
C.ZM.02.02ZM108       --.I---TE-L---------H-T-P-Q-V---H------SV..RI---Q--....YAT-EV--DI.....-A-----EGL--Q--HNVSE-F--Y-P-KT.........MNFTQPAA---L--T------R.-------TSK---RAF..............NGTYYWN 416

D.CD.83.ELI           --.-I---E-L-N---N--AH--E--K-T-A--YQ---Q-T..PI-L-QSL....Y-TRS.RSII.....G---------Q-SK--Q-V-R--GTLLNKT..........I-K--P---------T-------.-------TSG------...............NI-A-NNIT 402
D.CM.01.01CM_0009BBY  -D.IM---E-----V-N----F--TIN-T-I--G-----S-..HF---Q-L....YAND.I--DI.....-K------GKD-----QEV-R--GNLYN.K..........--S-RP---------T-------.-------TSK---R--S..........IGTWED--GN 407
D.KE.01.NKU3006       --.TI---E-L-N----------E--T-----S---QS-..RI---Q--....Y-TE.V---I.....-L-F-E-NGTE--K--Q-V-K--GNL-P-K..........------P-----L--T------R.-------TS-------...............D-NK--DTT 417
D.KR.04.04KBH8        K-.II---E-L-N---N------V--P-K-E-------RG-..HIA--Q-L....FATTNV--DI.....-R------ST---E--Q-L-K--GDL-NKT.........E-N--P-----L--T----V--.-------TSG------...................MVNSTV 391
D.TD.99.MN011         --.I----E-L---T-N------ETIN-T-I-------RSV..HI---Q-L....Y-TN.V--DI.....---------G-------QEV-R--GAALNTT.........Q----P---------T-------.-------TSK--DREF..............NITGNN 394
D.TZ.01.A280          --.II---E-L-N-V--------ET-K------------G-..RI---QT-....F-AE.VT-DI.....-K-Y----G-E-DK--Q-V-T--GDLLNKT.........I-N-SP---------T-------.-------TSL---T--...............IKGTQ 399
D.UG.99.99UGK09259    --.II---E-L-N---I------E--V------Y----G-..HV---Q--....Y.TNNIV-DI.....--------------K--E-V-K--SSL-NQT..........----RP---------T------R.-------TAR---YE-............SNNTLAST 402
D.YE.01.01YE386       --.II---E-LAN-V-N------E--T------Y---QGV..HM-----L....Y-TN.I---I.....-K---TV--DE--K--QKVTNR-EKLIS.K..........-VK-QP---------T-------.-------TSE---QSAW..................... 396
D.YE.02.02YE516       --.II---E-I---I--------ETIN-T-I-------QG-..HIA--Q-L....Y-TN.V--DI.....---------GTQ-----QK--K--GKLYNKT..........--V--P---------T-------.-------TST--TR-I............NTTNDN-VN 393
D.ZA.90.R1            --.II---E-L-N---N------A---------Y---QNT..-I---QT-....YASRMI--DI.....---------EE--KR--Q-V-I--GNLLNKT..........--N-RPP-------TS------.-------TSG------NRTGN..............N 405

F1.AR.02.ARE933       -D.II---H-IS-------H--E--Q-----------S-..HL---Q--....YAT-DI---I.....-K----V-GMQ--T--ERVKA--KPL-P-A..........-K-NS-----L--TM-----R.-------TSG---...................... 388
F1.BE.93.VI850        -G.I----Q-ISN-------H--E--Q-----------G-..HL---QT-....YAT-AI--DI.....-K------GTQ-----EYVKAE-KSH-P--T.........A-K-N------L--TM-----R.------DTSG---D-G..................... 385
F1.BR.01.01BR125      KD.II---Q-IS-------H--E--Q-----L------S-..RI---Q--....YAT-DI--DI.....-K----V-G-Q--K--GRVRA--KSH-PDA.........A-K-NS-----L--TM-----R.------DTST---..................... 393
F1.ES.x.P1146         -D.II--FQ-IS--T-----H--E--Q-----------S-..RI---Q--....YAT-DI--DI.....-K----V-G-Q--K-VER-SA--KSY-P-K..........-Q-NS-----L--TM-----R.-------TSK--TNNI..................... 411
F1.FI.93.FIN9363      GG.II---Q-LS-------H--E--Q-----------S-..RI---QS-....YAT-EI--DI.....-K-----GEQ--K--DRVKAE-KLH-N.K..........-Q-NS-----L--TM-----R.-------TSL---N-V..................... 393
F2.CM.02.02CM_0016BBY KN.II---E-I----------F-E--K-----------S-..RI---QV-....YAT-EI--DI.....-K---T-NGTL--A--NRV-AEVKNLTN.I..........-K-EP-----L-VT----S--.------DT-A---..................T-LL 393
F2.CM.95.MP257        -K.MI---E-IS--T------FKNP-K-----------RS-..HI---------....YAT-EI--DT.....-K------EKQ-YD--IK--TEFKD-YN.K..........-VG-QP-A---L--T------R.-------T-I---H-R..............VNDILS 399
F2.CM.97.CM53657      GD.I----E-IS---------F-R--A------T-I--RSM..RI----V-....YAT-TVL-DI.....-K-Y-T-NGTL--K--EGV-KEVQSHLN.K.........S-T-AP-----L-VT------R.-------TVA---A-N..................MT 395

G.BE.96.DRCBL         KD.II-S-E-IS----V---H--R--------------RSV..AI---Q--....Y-T-EV--DI.....-K----V-WT---E--RDVQA--Q-Y-I-K.........S-E-NS-----L--T-------.-------TSG---NSI...................LK 395
G.CM.01.01CM_4049HAN  -K.IR---E-----T-I------ETI--D-I--------S-..TF---Q--....YAT-DI--GI.....-------V-K---KTM--NVTTL-HKKYN.K.................S-A---L--T------R.-------TSE--...............NINISNS 396
G.CU.99.Cu74          K-.IM---E-L-N---V------ETI------------S-..RF---Q--....YAT-DI--DI.....-----EVNGT---E--QKVKE--K-SYNKKNITFG....N-T-E--T---L--TM---I-R.-------TSK----SX..................... 405
G.ES.99.X138          --.II---E-I---T-N-----KEAIN-T-I-------S-..SF---Q--....YAT-DI---I.....---Y---N-TE---MINNVSK--K-I-N-K.........S-T-NPP----L--T------R.-------TS-----SL..................-NS- 399
G.GH.03.03GH175G      --.II-Q-E-V---S-V----F-DT--------G----RS-..RF---Q--....YAT-DI---I.....-----IVNGE---QVIQKVKTH-E-IYN.K..........----SS-A---L--T------R.-------TSR---.......................N 401
G.KE.93.HH8793_12_1   G-.II-K-E-I---T-V------ET---T-V--------S-..HL---Q-L....YAT-DI---I.....-----DV-GRN-S-MIEKVKAQ--KI-N.K..........-T-DS-A---L--T------R.-------TSG---NET..................... 393
G.NG.01.01NGPL0669    -G.I---TE-TS--T-I------DTI------------E-..GI---QM-....YAT-EI---I.....-------V-GE--GDMSQKVKT--Q-V-N.K.........N-T-NS-----L--T--I---R.-------TSD---N-S..................... 402
G.PT.x.PT2695         G-.IM---E-I-N---N------ET-P-T-A--S----S-..RF---Q--....YATDAI--DI.....------------IR-E-MKQNVTA--EKI-.KK.........N-T-NPPA---L--T----I-R.-G-----TSA----SSLS..................... 423
G.SE.93.SE6165        GK.IKV--E-----T-V------KT-----------M---MGI---QT-....YAT-AI--DI.....-------VTKR--KEA-QNV-AE-GKI-NKSSE.......N-T-NS-A---L--T----I-R.-------TSG----SL...................L 403

H.BE.93.VI991         V-E-I---K-I---T-N------EP-Q-----TG----S-..RI---Q--....YAT-DI--DI.....-R-Y---GKQ--E--HKVIT--GSY-D-K..........----LQPPA---I--I-------.-------T-K------TN.........SSYTNDTYNS 412
H.BE.93.VI997         GQ.-I---K-IS--T-N-----DSPI--T---------G-..HF---Q--....YAT-DI---I.....-------V-EE---K--Q---TQ-SKY-V-R..........-L---PH----L-VT------R.-------TSG----S-..............TGDNI 403
H.CF.90.056           -Q.II--TK-IS--T-N-----K-P-N-T--------TS-..HL---------....YAT-DI--DI.....----------TD--K--H-VVTQ-GIHLN-R..........--S--PN----M-VR------R.-------TSG----S-.............EMHTNYT 396

J.CD.97.J_97DC_KTB147 K-.-I-K-K-IS-----------QT--------A-----G-..PI---QVL....YAT-AV---I.....---------GV------HKV-EE----EHFKNKT.......IVFAPAN----I--MM-T----.-------TSI----S-DEN.................... 402
J.SE.93.SE7887        GD.II---E-IS----N------KT---V-Y--------G-..HM---QVL....YAT-EI---I.....-ET-----ERD-S---RRV-T----H-N.K..........--N-TSP----I-----------.---L---TSK----S-..............DKNSIE 395

K.CD.97.EQTB11C       --.II---EDI-K-T-N------EA--------S-----S-..HI---------....YAT-DI--DI.....---------GGQ--K-VN-VKKE-GKH-N.K..........-----QP------QVTR-I---R.---S--DT-DTVDD-E..................... 400
K.CM.96.MP535         --.II---E-I---T-N------ET-Q-----------S-..HM---K--....Y-T-DI--DI.....---------GE---M--SRVKE--K-H-K-GT.........ITFKPPNP------L--M---A.-------T-K---E-G..................... 396
```

```
                                                   glycosylation NNT                                                                                          glycosylation NST
                                           glycosylation NCT     V3 tip                                                                     glycosylation NST
                       glycosylation NFT                                  V3                                                       CD4                    V4
                                     glycosylation NTS                                   glycosylation NNT     glycosylation NKT   CD4        glycosylation NST
                                                *                                        *                                             *     *
B.FR.83.HXB2                 EE.VVIRSVNFTDNAKTIIVQLNTSVEINCTRPNNNTRKRIRIQRGPGRAF....VTIG.KIGNM.....RQAHCNISRAKWNNTLKQIASKLREQFGNNK.........TIIFKQSSGGDPEIVTHSFNCG.GEFFYCNSTQLFNSTW...........FNSTWSTEG  404
01_AE.CF.90.90CF11697        -D.II---E-L-N-------H--K---------FK-M-TSA..RI---QV-....YKT-SIT-DI.....-K-Y-E-NGT---E----VTK----H-K-K.........----QP--------TM-H---R.-------T-K------...............T-NEIME  406
01_AE.CN.05.FJ051            --.II---E-L---------H--R--T------S----TS-..TI---QL-....YRT-EI--DI.....-K-Y-E-NKTT-HE----VVG----H-N.K.........----QPP----L--TM-H---R.-------T-N------..................NGTM  396
01_AE.CN.06.FJ054            G-.II---E-L-N-------H--K---------S----TS-..HI---QV-....YRT-EIT-DI.....-R-Y-K-NET---K--E-VRE----HIN.K.........--V-QPP----L-TTMLH---R.-------TSK---.......................  395
01_AE.HK.x.HK001             --.II---E---N-------H--E--Q-T----S----ESV..RI---QV-....YRT-EIT-DI.....-K-Y-Q-NAT--EKV---V-K------NKT.........N-R-QPH----L--TM-H---K.-------T----D-S-..................NTT  397
01_AE.TH.01.01TH_R2184       D-.II---E-L---T-----H--K---------S----TSM..TI---QV-....YRT-DIV--I.....-K-Y-E-NGTI--KV---V-D--K-H-N.K.........-TA-QPP----L--TM-H---R.-------T-K---N-...CINDTGNNETKGEC.....  409
01_AE.TH.02.OUR769I          --.II---E-I-N-------H--K---------S----TS-..HI---QV-....YRT-DIV-DI.......--Y-K-NGT---E----V-G----H-KGTE.........K---QPPL---L--TM-H-T-R.-------T-K------...............TNGTM  399
01_AE.TH.90.CM240            --.II---EDL-N-------H--K---------S----TS-..TI----V-....YRT-DI---I.....-K-Y-E-NGT---KV---VTE--K-H-N.K.........----QPP----L--TM-H---R.-------T-K---N-...C....LGNETMAGC.....  404
01_AE.US.00.00US_MSC1164     D-.II---E-L-N-------H--K---------S----TG-..TI---AV-....YRT-DIT-DI.....-K-Y-E-NET---E----V-R--K-H-N.K.........--N-QPP----L--TM-H---R.-------T-K-------..TGNESRDGCNGTTC.....  407
02_AG.CM.02.02CM_4082STN     --.IA---ESI-N---------VKP-K-D-I--------SM..RI---QT-....YAT-DI---I.........--Y--VNKTA-T---R-V-TQ--KY-K-ET.........--V-TN-----L--T-------.-------TSG----S-..............DSN---E  394
02_AG.EC.x.ECU41             -G.-----E-I-----------D-P-R-----------SV..RI---QT-....YAT-DI--DI.........------VN-TE--SS-QK-VT---KY-N-K.........----AN-----L-VT------.-------TSG------..............DVNNSTK  391
02_AG.GH.03.GHNJ196          R-.-----E--S----------AKP-Q----------TG-..HM-L--T-....YAT-DI--DI.........------V-AKA--D--Q-V-TQ-GKHY-G-T.........----TNH----V--M--T----.-------TSR----N-................KNGTA  405
02_AG.NG.01.PL0710           -N.-A---E-I-N----L----VNP-R----------I-GV..RI---QT-....YAT-KI---I.........--Y-TV-ETE-KEA-Q-V-TQ--KN-K-K.........----DKP----V--T-------.-------TSK-----G....................DT  390
02_AG.NG.x.IBNG              G-.-----E-I-N---------ANP-K-----------GV..HI---Q--....YAT-DI--DI.........------V-KTE--K--H-VVTQ-KTY-K-T.........----ANPL---V--T-------.-------TSK------................DN-NST  396
02_AG.SE.94.SE7812           --.I----E---N---I-----HE--K------G----SV..RI---QT-....YAT-DI--DI.........------V-WQQ--K--HDV-T----Y-N-T.........----DEP----L--T-------.-------TSN---R--NHNGTWNA...PGPFNDTEDK  406
02_AG.SN.98.MP1211           GD.I----E-ISN---------KP-W------S-----SV..RI---QT-....Y-T-AV--DI.........------V--E---S--QKVVT--GKH-NSS.........K---TN-----L--T-------.-------TAE------ASN....T...NGIWASNINA  401
03_AB.RU.97.KAL153_2         --.-----------T-------KEP-------------G-..HI--------....YAT-DIT-DI.........---------IT---------VI---K----K.........--V-N----------M------.-------T-K------.................NGTE  388
04_cpx.CY.94.CY032           --.-----K-I---T-N-----AKA-K------G-----SV..HI---LTW....YAT-EI--DI.........---------GND--D---V-SEE-KRL-P-K.........--K-APPV---L--T------K.-------T-P-----H..............MQNG-NI  405
05_DF.BE.x.VI1310            -G.II---Q-IS----N---H--E--H-----------S-..HL---Q--....YAT-DI--DI.....-K----V--EQ--K--I-V-KE-QSH-P-K.........--K-NS-----L--TM-----R.------DTSK---A-V.............--D-VFNAT  406
06_cpx.AU.96.BFP90           GN.IT-KTE-I---T-N------QP---R----G-----S-..SF---Q--....IAT-DI--DI.........------V---N-TDI-GEVKV--E-V-N-T.........H-T--S-A---L--T-------.-------TSN---TSD...............LFNT  408
08_BC.CN.97.97CNGX_6F        R-.II---E-L-N-V-----H--Q----V----------S-..RI---QT-....YAT-DI--DI.........---------KDA-YE--QRVSK--A-H-P-K.........--K-AS-----L--T------R.-------TSG---G-Y..............NGTYMLN  404
09_cpx.GH.96.96GH2911        -D.II---E-I-N-V-V-----ADP-T-D-V-TG----TSV..RI---QT-....YAT-DI--DI.....-KV--Q-NGT---EI-GKV-E--KGH-N-K.........----RP-----V--T-------.------DTSG---N--..................NGTW  392
10_CD.TZ.96.96TZ_BF061       --.II---E-L-----N------E--T---V-------RS-..RI---QT-....YAT.DI---I.........--YY---GTE--K--Q-V-K--GDL-NRP.........-----P--------T--T----.-------TSE------..................TSNN  387
11_cpx.GR.x.GR17             G-.-R---E-L-N----------ST-R-----------G-..HI---Q--....YAT-DI--DI.....K-----V---D-----Q-V-EQ-HNN-N.K.........--V-NEH----L-VT------.-------TSG------..............YANDN-ST  388
12_BF.AR.99.ARMA159          -D.II---Q-IS--T-----H--E--Q-----------S-..-L--------....YAT-DI--DI.....-K----V-G-Q-SK-VE-VKA---S--N.K.........--K-NS--------TM-----R.-------TSG---D-I....................  394
13_cpx.CM.96.1849            G-.IK--AK--S--VGN-----A-P-R-----------S-..RI---Q--....YAT-DI--DI.........---------Q------Q-V-AQ--KYLKD-T.........----ANP----L-VT-------.-------TSA------..............E-D-V--Q-  406
14_BG.DE.01.9196_01          K-.-I---E--------------K--V--XS-------S-..PI--------....YAT-QI--DI.........------L-K-Q-DK--R-V-GE--K--EED-.........--S-NK---------M-T----.-------T-------..................K-NST  398
15_01B.TH.99.99TH_MU2079     --.-I---S--N---I-----E-----------------S-..HI----SW....Y-T-EI--DI.........------L-ST--G---N--TE--------K.........--V-N-P-------IV------.-------TS------...............FQNGNST  391
16_A2D.KR.97.97KR004         GNKTI---A-I---T-N---FTKP-Q------DVGQ-RSV..RI------YTRQTYTRQAK-DI.........---Q----SRQ--D--QKV-EQ--KY-S-K.........----TN-----I--T---V---.------DTSG------.................PANAS  403
17_BF.PE.02.PE02_PCR0155     -D.II---Q-IF--T-----H--E--Q-T---------S-..RI-T-Q--....Y-T-XI--DI.....-K----V-G-Q--K--E-VKA--KSYWN.T.........--Q-DS-----L--T------R.-------TSG---D-G....................  399
18_cpx.CU.99.CU76            -K.-M---E-I---T-N-----T-P-N-T----S----RS-..HI---Q--....YAT-DI--DI.........---------EKD--D--HEV-TQ-GGHLNKS.........--Y--P-----I--T--T----.-------T-A------N.........KTXG--IKLD  405
19_cpx.CU.99.CU7             GK.-T---E-I----------FTXP-K-----------TSV..RI---QV-....YAT-..--DI.........------V-KKE-KEA-QEV-KQ-TVL-N.K.........--S-AN-----L--T-------.-------T-D----RR....................  363
20_BG.CU.99.Cu103            K-.IM---E-L-----V-----K-I-------------S-..GF---Q--....YAT-DIT-DI.........---F-VVNKT---EM-QKVKA--A-SYNK-IT.........N-K-SPPT---L--T------R.-------TSG----SN....................  393
21_A2D.KE.91.KNH1254         --.II--TE-ISR-E-S------E--T-----------RG-..HI---Q-I...YAT-KI---I.........---------E-D--K---GV-I--GDL-NTT.........K-N--P---------TK------.------DT-E------..............YRNG-LI  405
22_01A1.CM.01.01CM_0001BBY   K-.IM---E---N---I---XVKP-T---S--G----NS-..RI---QV-....YAR-DI--DI.....-K-Y-EVNKTE--E--HGVVKQ--KHWN.K.........----TN-----L--T------R.-------TSG------.................HDNT  389
23_BG.CU.03.CB118            K-.IM---E-L-N---V-----S-IT------------S-..TF---Q--....YAT-DI---I.....-E-F-VVN-TA--E-MQKVKA--A-SYNTSTE.........N-T-GPP----L--T----I-R.-------TSG----SN....................  401
24_BG.CU.03.CB378            KG.IM---E-L-N---V----F-K-ID------G-----S-..RF---Q--....YAT-AI--DI.........---Y-EVN-T--EEMIQ-VKAE-A-SYNKSIA.........N-T-SPPT---L--T------R.-------TSK---ESL...................LNS  404
24_BG.CU.03.CB471            K-.-M---E-L-N---V----F-K-IK-----------RSV..GI---QV-....YAT-DI--DI.........---Y-VVN-TE--DMIQKVKV--A-SYNKTTE.........N-T-STPT---L--T------R.-------TSG---SN....................  402
25_cpx.CM.01.101BA           --.I----E-ISN---S---HF-ET-T---I--------SV..RI---QT-....YAT-DI--DI.........------V--HQ--K--QEV-KQ-SKHLNIP.........E---ANH----V--T-------.-------TSA---GS-..............TNDTANN  415
27_cpx.CD.97.97CDKTB49       --.-IV-FE-L-N-------H--K--I------S----TSV..SI---QVL...FYRA-DI--DI.........---------NETQ-ED--YKV-AQ--KHLNKS.........SS---P-----I-VT--T---M.-------T-T------..................NDTT  398
28_BF.BR.99.BREPM12609       --.II---E-L-----I------ET-----------RN---..HL-R--IM....YVTERIQ-GI.........---------E--E---R-VVT--E---TKLGEQFKDFKNR--A-N-----------M-T----.-------T-L------.............KWNNTWNNT  415
29_BF.BR.02.BREPM119         --.-I---E-I---V-V------Q--------------S-..HM-W------....Y-T-EI-DDI.....-K------K---D-------N--K---N.K.........--V-N------L---M-R----.-------TSK------...............NATNME  395
31_BC.BR.02.110PA            D-.II---K-I---VQ------KEP-------------ES-..RI---QT-....YAT-DI--DI.........---Y------A-EK--LEVSK--S-H-P-K.........--N-TKH----L--T------K.-------TSR---RAYF......N...STXHXN--XN  416
32_06A1.EE.01.EE0369         -S.II---E-I---T--------Q--D-I----Y----RS-..HI---Q--....FAT-DI--DI.........---------G----EM-HKVIA--K-S-PGK.........N-T-NS-A---L--T-------.-------TSK----SS....................  402
33_01B.MY.05.05MYKL007_1     --.IM---E-L---T-N---H--K---------S-SI-TSM..TI---HV-....YRT-DIV-DI.........---Y-EVNGT---KV---VSE--K-H-K-K.........--V-RPPP---L--TR-H---R.-------TEK---N-...C....IGNETMGGC.....  406
34_01B.TH.99.OUR2478P        --.II---E-I-N-------H--K---------S----TSV..HI---Q--....YK--DI--DI.........---Y-E-NGT---E----V-G---KH-.-K.........--S--PP----L--TM-H---R.-------T-K------PEN..................  391
35_AD.AF.05.05AF095          G-.IIL--E-IS--------H--E--T------S-----S-..RI---QT-....YATD.I--DI.........------V---A--K--R-V-TQ--QHW.-K.........--V-SN-----V--T-------.------DTSG------NITDIG.............-P  391
36_cpx.CM.00.00CMNYU830      N-.TI---E-I-N------I--VKP-N-T-V--------GT..PI---Q--....YARNDV--DI.....-K---EVNATE--D--Q-VVQ---KFW.-T.........----TEP----L-VI-------.-------TSG------GYNNS..............NV  397
37_cpx.CM.00.00CMNYU926      GK.-M---E-I-N--------F-ET-------F------SV..RI---QT-....FAT-DI--DI.....-K-Q-AVNKTE--KA-QKVVAQ--NH---T.........--M-AN-----V--T------A.-------TSG------NTTTS..............TQ  397
38_BF1.UY.05.UY05_4752       G-.IM---Q-IS--------H--E--H-----------S-..HI--------....Y-T-DI--DI.....-K-----GEQ--K--ERVKE--KFY-P-K.........--Q-NS-----L-VT------R.-------TSK--.........................  397
39_BF.BR.03.03BRRJ103        -S.II---Q-IS--T-----H--EP-Q-------S----G-..HL--------....Y-T-EI--DI.....-K----V-G-Q--K-IE-VKR--KSH-P-A.........--K-NS-----L-VTM-----R.-------TSG---N-V....................  396
40_BF.BR.04.04BRRJ115        KD.I----Q-LS----I---H--E-IQ-----------S-..GI--------....YAT-DI--DI.....-K----V--TQ-RKK-R-VKGE-KS--NGA.........N-T-NS-----L--TM-----R.-------TSG---D-E................IEAN  403
42_BF.LU.03.luBF_05_03       -D.II---Q-XS--------H--E--Q-----------S-..HIA-------....YAT-DI--DI.........---Y--V--EQ--K--ERVRG---SL-.-K.........--K-XS-------VTM-N---R.-------T-E---D-E....................  393
43_02G.SA.03.J11223          --.I----K--S---QVL-----X-I----------X-S-..GF---QT-....YAT-DI--DI.....-E----V--TR--E---NVKAQ-HKI-N.K.........--L-NS-I---X--T-----XR.-------TSR--.......................-  387
U.CD.83.83CD003              --.-I---E-I-N---------ET-K------GSDKKI-QS-RI---KV-....YAK-GIT-..........------TDGE-R---Q-V-IA--R--N-K.........S---NS-----I--T--T----.-------TSE--TGI-NG.........TW.DKNC-ST  400
U.CD.90.90CD121E12           K-.-I---E-I-N--------F-ET-K---S-------QG-..HI----V-....YATSNI--DI.........------LTKWQ-KRA-Q-V-DA--RH-N.K.........--V-NS-----L--T-------.-------TSG------NG.........TWAENNC-NS  396
U.GR.99.GR303                G-.-I---E-------------E--S------S-----S-..HI---Q--....FATD.I--DI.........------V-G-T------RVVA--------K.........----NS-T---L--T---L-X-DR------TSG------................RPNDMQWS  405
U.NL.95.U_NL_95_H10986_D1    --.IM--TE-I---T-I------ETISL----------S-..PM--------....YAT-DI--DI.....-K-Y-EVN-T----M-RKV-V------NL-NK.........----NS-A---L--T------M.-------T-R---T-HL...........Y---GNDTQ  407
CPZ.CD.90.ANT                NTS--MNGRKNESVLVRFGKEFENL.T-T-I--G-R-VRNL..-I---MT-....YNVEIAT-DT.....-K-F-TVNKTL-EQARNKTEHV-A-HWKKVDNKTN.....AKTIWTFQD----VKV-W---Q.------DI-PW--A-Y.....................  398
CPZ.CM.01.SIVcpzCAM13        -N.ITV-VN-ASK-THDW----S-A-NLT-K-VG----GKV..-I---MT-....YNMDHIF-DT.....-K-F-ELNGTT--E--QKVRES-IKEIKA-ANGTY.....N-T-EP--------AN-M----.------DTRKM--.......................  405
CPZ.CM.05.SIVcpzEK505        SN.I---N..NSK..D-LL----E-IP------G-K--GQV..-I---MT-....YN-ENI--DT.........---Y-EVN-T.-EQIWNTTKQIII...N-R-..........N-T-IPNP---L-VTNLMI---.-------TS---TNQN....................  379
CPZ.CM.05.SIVcpzLB7          -K.IM--Y..KA...DKVL------IH-----VG-K-I-G-..PL---QL-....YGTETVV-DT.....-E---E-NQTH-YKI-N-VKRE-TTV-NESNK.........-VH-AN-------VANLH---.-------TSA---D-L...................LRNKT  389
CPZ.CM.05.SIVcpzMB66         KVT-LD-N..VSNDMD----K--ET-RL----TG---I-G-..PI--SQI-....YG-ETV--DT.........---F-QLNKTV-T--F-KVRHA-N-TYKGYLGNE.........--T-GP-T---L-VTNLHLI--.-------TSI---TSI....................  420
CPZ.CM.05.SIVcpzMT145        NQ.T-A-I..DPS.ENLA-I--KDP-K-T-R--G----GQ-..-I--AMT-....YN-ENVV-DT.....-K-Y-E-NGTQ-AKA-NETKEV--NILR.K.........N-S-MVP------VTN-H---.-------TSEII-I-K....................  387
CPZ.GA.88.GAB1               GN.ITV-VE-KSK-TDVW----VEA-SL--H--G----GEV..-I---MT-....YN-ENVV-DT.....-S-Y-K-NGTT--R-VEEVKKA-ATSSNRTAA.........N-TLNRA------VTH-M---.-------TS-I-TDNI....................  393
CPZ.TZ.00.TAN1               --.TKAYF.VN-SVNTPLL-KF-V-INLT-E-TG----GQV..-I---MT-....YN-ENVV-DT.....-K-Y-SVNATT-YRN-DWAMAAINTTMRARNETV.....QQT-QWQRD----VTSFW---Q.-------L-NWT-.................TWTA  396
CPZ.US.85.CPZUS              KN.-TV--K--A-I...-L--FSEG-NMT-I--G---VGNV..-L---MT-....YN-PKIV-DV.....-E------KLT-E....KQRKYTL-IIKKEANLT.....KVELIPNA-----V-NMML---.-------TIP---M-Y....................  375
CPZ.CM.x.SIVcpzDP943         NS.TXX-TEXMXXMDX-----FHKR-QMT-I--G--S-GQ-..-I--AMS-....YNLENI--DT.........---Y--X-A-E-EQRXXDTXRAIKSLKPGS.........N-T-TDQQ-------NMM----.-------T-P---NS-.................-NT  386
CPZ.CM.x.SIVcpzMB897         GN.KTVVR-RSKS-TE------A-AIY-----LG-K-IEG-..PI---QI-....YRTKTVV-DT.....-G-E-R-NGTA--E--R-VKEA-NNTYR-LNLSLT.....E-N-EGA----L-VT--Y----.-------TSN---.......................  388
N.CM.02.DJO0131              .D.I---NNSGG....NLL--W-EI-TM-----G---GGQV..-I--AMT-....YN-EKIV-DI.........------V.SNE-RSMWNKTKE-IKSLL---I.........-FKAQEKN-----VTHLM----.---I---TSR---ESM....................  384
N.CM.04.04CM_1015_04         DD.I-X-XXXD.....NML--W-XT-S------G--XGGQV..-I--AMT-....YN-EKI--DI.........------V.SEE-KSMWDRTKE-IKGLL---T.........-FKSRVNI-----VRHFM----.----L--TSR--......................  378
N.CM.04.04CM_1131_03         TD.I---N.DSGTILGN-L--W-KT-L------G---GGQV..-I--AMT-....YN-EKI--DI.........------V-EEE-KSMWNRTKE-IKDIL--KT.........-FEARGND-S---VTHLM----.----L--TSR---.....................E  397
N.CM.95.YBF30                DG.I---N.D...SHSNLL--W-ET-P------G---GGQV..-I--AMT-....YN-EKIV-DI.........---Y--V-KEL-EPMWNRTREEIKKIL-K-NI.........-FRARERNE--L-VTHLM---R.-------TSK---EEL....................  394
N.CM.97.YBF106               DD.I---H.....HGGNLL--W-ET-S------G---GGQV..-I--AMT-....YN-EKIV-DV.........---Y--V.SEE-GSMWNKTKK-IKRLL---T.........-FKAQDKN---L-VTHLM---X.-------TSR---ESE....................  383
O.BE.87.ANT70                GK.IRMMAKDILEGG-N---T--STLNMT-E--QID.IQEM..RI--.M-W....YSM-IGGTAGN..SS-A-Y-KYNATD-GKI---T-ERYL-LVN-TGSI.......NMT-NH-----L-VTHLH---H.-------TAKM--Y-F.SCNG......TTC-.....V  398
O.CM.91.MVP5180              -K.IR-MGK-I-ES--N---T---PINMT-I-EGIAEVQD-..YT--.MRW....RSMTL-RS-NTSPRS-V-Y-TYNKTV-E-A-Q-T-IRYLNLVNQTENV..............----SRT----A-VSHLH---H.-------TSGM--Y-FINCTK......SGCQ....-I  407
O.CM.96.96CMABB637           GK.IRMLGK-LSNT.DN-L-T--STLN-T-E--YEQLVQEM..KI--.M-W....YSMEVERNKTAA.TP-T-Y-TYNVSE-ESA-----ERYL-LINQTDNV.......AV--DR-T-----VSYLH---H.-------TSEM--Y-F.NCTP......TKCN.E-SYN  418
O.CM.98.98CMA104             GK.IR-MGKDI-AGSRN-L-T--STIN-T-E--F-L-VEQM..KI--.M-W....YS-DIGR--TS..DP-R-F-KFEET--KEA-X-T-ERYL-LVN-TIDNV.......SV--NR-ND-----AFLH---H.-------TSGM--Y-F.TCNK......TNCT.....N  405
O.CM.99.99CMU4122            GK.IR-MGK-I--SG-N---T-DSAIN-T-E--W-QSVQEL..KI--.MSW....YSM-ISSPLSN..AS-L-Y-RYNATE-KKA-----ERYL-LVN-TNNTXDNV..SM--SK-T---S--AHLH-S-H.-------T--M--Y-F.SCNG......TSCT..NNQS  405
O.FR.92.VAU                  GN.IT-MGKDIS-SGEN-LIT---DMT-A-E--GSQ-IRK-..MA--.M-W....YSMALSNTKGD...T-A-Y--Y-ATD--KA--N-TERYL-LVEY-QTDV.........--K-GNH----A-VTNFF---H.-------TN----H-F.SCKKNMTYNKINCT......  409
O.SN.99.SEMP1299             GN.IR-MGK-IS--M-N---T--STINMT-V-QG-QSVQE-..-I--.M-W....YSMSLAQEGKPN.NS-I-Y-KYNISD-EKA---T-ERYLDLRN-TNTV.......N-T-ER-I---S-VTHLH---H.-------TSKM--Y-F.SCIG......TNCT.....S  409
O.US.99.99USTWLA             GK.IR-MAK-ISS-SDN-L-T--STINMT-M--G--SVQEM..RI--.M-W....YSM-LGN-YTN..RS-I-F-TYNATE-KE--QG--ERYL-LVNYTEKI.......N-T--N-TD--I-VTHLH---H.-------TN-M--Y-F.ECNN......SNCR...IQN  407
```

```
                                V4                                CD4                                                
```

```
                                V4                                CD4                                                
```

```
                                                glycosylation NAS          glycosylation NHT
                             immunodominant region            glycosylation NKS      glycosylation NYT
B.FR.83.HXB2                SGIVQQQNNLLR.AIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWS..........NKSLEQIWN..HTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIKLFIMIVGGLVGLRIVFAVL 702
A1.GE.99.99GEMZ011          -------S----.------Q--K------------V--------------------------N----S---..........---YSE--D..NM---Q--K-VI---NI-YE--------------D--A-----G-----S-S-------I--------I--------- 698
A1.KE.00.KER2008            ------------.------Q--K------------V-------R------------------N----S---..........---QSE--D..NM--LQ---------DI-YR-L-D----------D--A--------T--DLS-------I--------I--------- 700
A1.KE.00.KNH1144            -------S----.-------M-K------------V-------R------------------N----S---..........---QDE---..NM--LQ--K--S---N--Y---------------D--A-G---N--T--D-S------RI--------I--------- 707
A1.KE.00.KSM4024            -------S----.---------K------------V-------R------------------N----S---..........----NE--D..NM--LQ--K--S---QI-YN---K-------------A-----G-----D--K------I--------I--------- 671
A1.KE.00.MSA4069            -------S----.------QI-R------------V----A--R------------------N----S---..........---YNE--D..NM--L--E---S---DV-YN-L-A----------D--A--------S--S---------I--------I--------- 698
A1.KE.00.NKU3005            -------S----.-------M-K------------V-------R------------------S----S---..........---YNE--D..NM--QQ-E---E---DY-YN--------------D--A--------S--S-SK------I--------I--------- 704
A1.RU.00.RU00051            -------S----.---------K------------V--------------------------N----S---..........---QSD--D..NL---Q-----H---DI-YD---K----------D--A-----------D-S----------------I--------- 692
A1.RW.93.93RW_024           -------S----.---------K------------V--------------------------N----T---.........K----SE--D..NM--LQ--K--S---N--YN--------------D--A-----N--T--D-S-------I--------I--------- 702
A1.SE.95.SE8891             -------S---M.---------K-----------QV-------R------------------T----S---..........DV-QSE--E..NM--LQ--K--S---QI-Y---------------D--A-----N-----D-SK------I--------I-----I--I 691
A1.SE.95.UGSE8131           ------------.-------M-R------------VM-------------------I---------ST--..........---YT---D..NM--LQ--K--S---DI-YQ------Y------K---------N-----D-S-------I--------I--------- 710
A1.TZ.01.A173               -------S----.---------K------------V--------------------------N----S---..........---HDD---..NM--LQ--K--S---Q--YN-------------------A-----------D-S------RI--------I----I---- 704
A1.UA.01.01UADN139          -------S----.---------R------------V--------------------------N----SN--..........S--HDE---..NM---Q--K-VI---NI-YG---K----------D--A--------S--D-S-------I--------I--------- 700
A1.UG.92.92UG037            -------S----.---------K------------V-------R----------------P-N----S---..........----DE--E..NM--LQ--K--S---IK-YE------I--R---KD--------------D-SK------I--------I--------- 696
A1.UG.99.99UGA07072         -------S----.---------K------------V--------------------------N----S---..........---QSE--D..NM--LQ--K--S---QI-YN------I--K----D--A-----N-----D-S-X----RI--------I------T-- 701
A2.CD.97.97CDKTB48          T------S---K.------QM-R------------V--L----Q----------------A-D-R--S---..........--TQ----K..NM--LQ--K--ST--DI-YM-L------------D--A-----N-----D--R------I--------I-----I-II 689
A2.CY.94.94CY017_41         -------S---Q.---------K------------V------------------------A-T----T---..........---QDE--D..NM--LQ--K--S---NI-YR-L------------D--A-----D--S----SH-----RI--------I-------II 704
A.SN.01.DDI579              -------S----.-------M-K------------V--I----------L-----------N----SN--..........R--Q-E--G..NM--LQ--K--S---DI-Y---------------D--A--------S--D-SK------I--------I--------- 700
A.SN.01.DDJ369              -------S----.------QM-K------------V--------------------------N----S---..........---YSE--D..NM--LQ--K--S---D--YN--------------D--A-------------S------RI--------I------T-- 705
A.SN.96.DDJ360              -------S----.---------K------------V-------R------L----------------S---..........---QNE--E..NM--LQ--K--S---DI-YN--------------D--A-----N-----D-SK------I--------I--------- 694
A.CD.02.02CD_KTB035         -------S---K.---------K---*-----------------------R-*----------N--*-Y-*-..........-R-Q-D-*R..NM-*-Q*-K-VSK--NT-YM------Q--R----D--A--E*---*-*-D-SK--*--RI--------I--------- 693
A.CD.97.97CD_KCC2           A------S---K.---------R------------V--------------------------T----S---..........---QDE---..NM---Q--K------EI-YG------K------F---------N-----E-S-------I--------I--------- 701
A.CD.97.97CD_KTB13          -------S---K.---------K------------V--------------------------N----S---..........---QGE--D..NM---Q-----S---DI-YK------E------LK--A--Q--------D-S------RI--------I--------- 723
B.AR.04.04AR151516          ------------.---X---M--------------V-------R------------------T-----------------------GD--E..NM---------S---D--YX-L------------D--A-----N-----D---------I--------I------T-- 698
B.AU.87.MBC925              ------------.----------------------L--------------------------T-----------------------DK---..NM--R--EK--D---D--YT---Q------------------------D---------I--------I------I-- 705
B.BO.99.BOL0122             ------------.-------M--------------L-------R------------------N----S---..........----SK--D..NM-----E---D---G--YT-------------------H---------D--K------I------------------ 696
B.BR.03.BREPM2012           ------------.-------M--------------V------F-R---------------------------..........----SE--D..NM-----EK--D---G--Y--L-D----------D--A-------------S-------I--------I------T-- 706
B.CA.97.CANB3FULL           ------------.----------------------L--------------------T------------T---..........----NE--D..NM---Q-E---DS--ND-YR---------------------------------S------R---------I--------- 696
B.CN.05.05CNHB_hp3          T------R----.----------------------V-----------------------------------..........---HDE---..NM-----E-Q-D---AE-YT---Q----------L----------------M-Q------I--------I----I---- 705
B.CO.01.PCM001              ------------.----------------------V--------------------------T-----------------------DM---..NM---D-EK--D---G--YT--------------K------------T--D--------R---I---S------I---- 696
B.GB.83.CAM1                ------------.----------------------V-----------------------------------..........----DK---..NM-----E---D---N--YT--------------KD-----T--------D---------I----I---I------TI- 702
B.GB.86.GB8                 ------------.---------R------------V--------------------------T----V---..........----DK---..NM-----E---D---NT-YT-------------------------N-----D---------I----I---I----I---I 695
B.GE.03.03GEMZ010           -------S----.----------------------V-------Q--R---L-----------T--------..........----NE--D..NM---Q-E---D---GI-YN---I--T------H---------------D-S-------I--------I------T-- 693
B.IT.05.SG1                 ------------.-------I---S----------V--I----R------------------S----T---..........--TQN---D..NM---Q-EK--D------YT-L----T---------*-*----NV--G-D--Q-----RI------C-I---V--V-P 660
B.JP.05.DR6538              -----------K.----------------------V--M----R-----------------------T---..........----GD--D..NM---Q-E---D---N--YT-L-----------------------------D--R------I--------I------VI- 710
B.KR.05.05CSR3              ------------.----------------------V-------R------------------T----S--N..........-RN-S---D..NM---Q-EK--D-V-V--YR---K-------------------N-----D---------I--------I------T-- 715
B.NL.00.671_00T36           ------------.----------------------V--I----Q------------------T--------..........----D---E..NM---Q-E---D------YT---D--K----------A--T--------S---------I---------------I-- 701
B.RU.04.04RU128005          -------X----.----------------------V-----------------------------------..........----NE--D..NM---Q-X---D---XT-YN-L-------------------------------D---------I--I--------------- 696
B.TH.00.00TH_C3198          -------R---S.------R---------------V-------------------------------N---.........N---Y-----..NM---Q-ET--D---KE-YT---Q-----------------N-----S--D--K-----RI--I--------------- 698
B.UA.01.01UAKV167           ------------.----------------------V-------G----------A--------T-----T--..........--T-K---D..NM-----E--------DI-YN---Q-------------------------S---------I--------I----I-T-F 711
B.US.04.ES10_53             ------------.----------------------V-------R------------------N----I---..........--N-S---D..NM-----E---E---QT-YN-------------K---------------D---------I--------I------T-- 692
B.US.99.PRB959_03           ------------.-------M--------------V-------R------------------S----S---..........-RN-TA--D..NM---Q-E---D---N--Y--L--------------------------S--S---------------I--------- 695
C.AR.01.ARG4006             -------S----.-------M------------T-V--L----R------------------N----S---..........----DN--D..NM---Q---------DT-YR-L--------Q--KD--A----QN-------S-------I----I---I----I---- 702
C.BR.04.04BR013             -------S----.-------M------------T-V--L----R-----------------------S---..........---K-A--D..NM---Q-----D---GT-YM-L------------D--A----QN-----D--S------I--------I----I---- 704
C.BW.00.00BW07621           ------------.-------M------------T-V--I--------------------------------..........---EGD---..NM---Q-----S---NT-YR-L-N--I---Q--KD--A--S-KN--S--D-S-------I--------I----I---- 689
C.CN.98.YNRL9840            -------S---M.--------------------T-V--I----------------------------S---..........D--QDD--G..NM---Q-----S---DT-YR-L--------E--KD--A--S-KN--S------------I--------I----I---- 707
C.ET.02.02ET_288            -------S----.-------M-----------RT-V--I----Q----------------P-T----S---..........---Q-E--E..NM---Q-----S----T-YR-L-V----------D--A----KD--S--S--Q------I--------I----I---- 696
C.GE.03.03GEMZ033           ------------.-------M--------------V--I----Q------L----------SF----ST--..........---QTD--E..NM---Q-----S---GT-YQ-L-DA-S------KD--A--S-N------S-S-------I--------I----I---- 704
C.IL.99.99ET7               -------S----.-------M------------T-V--I-----------------------N----S---..........---Q-D--S..NM---Q-----SK--NT-YG-L-V------E--K---A--R-KN-----D---------I--------I----I---- 692
C.IN.99.01IN565_10          -------S---K.--N----M------------T-V--I-----------------R-------*-S---..........----TD--D..NM---Q-----S---DI-YR-L-D---------KD-----S-KN-----S--H------I--------I----I---- 699
C.KE.00.KER2010             -------S----.-------M--------------V--I-----------------------N----S---..........---QTD---..NL---Q--K------DT-YR-L----S---R--KD--A--S-NN-----D--R---------I--VS-I----I-V-F 701
C.MM.99.mIDU101_3           -------S----.-------M--------------V--I----------------------------S---..........---Q-E---..NM------K------DT-YR-L-D------R--KD--A--S-KN---------------I--------I----I---- 704
C.MW.93.93MW_965            -------S----.-------M--------------V--I----R------M----------------S---..........--TYND--D..NM---Q-----S---DT-YR-L-D------R--KD--A--S-NN---------------I--------I----I---- 691
C.SN.90.90SE_364            -------S----.-------M------------T-V--I-----------------I---------T---..........-----D--D..NM---Q----------I-Y--L-------------KD--A--S-NN--------K------I--------I----I---- 687
C.SO.89.89SM_145            -------S----.-------M-----------R--V--I----E-----------------------S---..........-XTQ-E--E..NL---Q--K--S---DT-YR-L-V---------KD--A--S-G------D--K------I--I-----I----I-V-- 700
C.TZ.02.CO178               -------S----.----------------------V--I----------------------------S---..........---QTE--G..NM---Q-----S----A-YR-L-D--I---E--KD--A--S-NT-----D-SS-----RI--I-----I----I---- 708
C.UY.01.TRA3011             -------S---K.-------M------------T-V--I----------------------------S---..........---E-D---..NL---Q--K--S----T-YR-L-D---------KD--A----QN--T--D--K------I--------I----I---- 697
C.YE.02.02YE511             -------S----.--------------------T-V--I--------------------------------..........---QGD--D..NM---Q-----S---GI-YK-L--------E--KD--AS-S-NN---------------I--------I-V--I-T-- 701
C.ZA.04.04ZASK164B1         -------S----.--------------------T----------------L-----------S----T---..........---YTD--D..NM---Q-----DD--ET-YR-L--------R--KD--A--S-NN-------S-------I--------I----I---- 710
C.ZA.05.05ZAPSK240B1        -------S----.-------M--------------V--I----Q------M----------------S---..........--NQTE--D..NM---Q--K--S---YT-YR-L-D--I--------KD--A--S-NN-----S--K------I--------I----I---- 685
C.ZM.02.02ZM108             -------S----.-------M--------------V--I----------------------------S---..........-R-Q-E--D..NM---Q--K-VS--SDT-YK-L-A------Q--KD--A--S-----S------------I--------IS-------- 718
D.CD.83.ELI                 ------------.---------------------------------------------H----N----S---..........-R--NE--Q..NM-----E---D---G--Y-------T------K---------------S--Q------I----I---I--------- 699
D.CM.01.01CM_0009BBY        ------------.---------------------------------------------H----N----S---..........----DD--E..NM---Q-EK--E---GV-YN------I------K---------------S--------RI--------I--------- 711
D.KE.01.NKU3006             -----------K.-------M---------------------------------H----N--------..........----GY--E..NM-----E---D---G--YN-L----I----------Q-----------S--Q------I--------I--------F 715
D.KR.04.04KBH8              ------------.----------------------V--L-----------------------T----S---..........-R-VDT--D..NM---Q-E---D---GI-YD-L-----------KD--------------S--R------I--I-----I----I---- 688
D.TD.99.MN011               ------------.----------------------L-----------------------H----T----S---..........-R--ND--Q..NM---Q-EK--E---GV-YN------I------K-----------------SQ------I--------I--------- 691
D.TZ.01.A280                ------------.----------------------V----S-------------------H---------S---..........----DD---..NM-----EK--D---GV-Y-------V------K---------------S--K---------------I--K---T-F 698
D.UG.99.99UGK09259          -----------M.----------------------V--L-S-------------------H----N----S---..........--T-NK--D..NM-----EK--D---E--YN------T----------Q-----------S--K------I--------I--------- 701
D.YE.01.01YE386             ------------.---------------V------V----S-------------------H----N----S---..........--T-DE--G..KM--RQ-EK--D---GV-Y-------T----------Q-----------D-S------R---------I--------- 688
D.YE.02.02YE516             ------------.---------------------------Q-------------------H----S----S---..........-R--DA--Q..NM-----EK--D---G--YA-L----I--D---K---------------S-SQ------I-------SI--------- 692
D.ZA.90.R1                  ------------.----------------------V-------R-------------H----T---DPE--..........--TYDE--G..NM-----E---D---GY-YT------V------K-------------------------I--------I--------- 704
F1.AR.02.ARE933             ------------.-------M--------------L-------R------L-----------T----S---..........--TQKE--E..NM-----EK--S--SNE-YR---A-------------A-----------S------------I----FI--------- 691
F1.BE.93.VI850              ------------.----------------------V--------------------------N----S---..........---Q-E---..NM-----EK--S--SNI-YK------------------A-----------D-S-------I--------I--------- 677
F1.BR.01.01BR125            PYCS#RG-----fT---------------------L-----------------L-----------N----N---..........--NYSE--D..KM-----EK--S--SNE-YR------I------------A----T---S-----S-----RI--------I--------- 695
F1.ES.x.P1146               -----------K.-------M--------------V--L-----------------Y--------N----S---..........--TYD---D..NM-----EK--G--SNT-YR-L-Q-----I-------S--------S--D--------RI--------I--------F 707
F1.FI.93.FIN9363            -----------Q.-------M--------------V-----------------L-----------N----S---..........---QDE---..NM---Q-EK--S--SKT-YM---K--S---R------------D--S--D---------I--------I--------- 688
F2.CM.02.02CM_0016BBY       -----------K.--------------------------------------------------N----S---..........---Q-E--G..NM---Q-EK--D---DT-YR---A----------D--A---D--S--T----------RI---V----I--------- 691
F2.CM.95.MP257              -------S---K.-------------------------------------------------P-T----L---..........---QDE--G..NM-----EK--G---DT-YR--SA----------D--A---DN------S--R-----EI----I-S-I------T-- 696
F2.CM.97.CM53657            -----------K.--------------------------------------------------N----S---..........---QNE--E..NM---Q-EK--S---GT-YK--NA----------D--A---DN--S--T-------------------I--------- 695
G.BE.96.DRCBL               -------S----.-------------------R--V--L-----------------------N----T---..........---YNE--E..NM--I--E---D---YH-Y----Q--I-------D--A--Q-----S--S-S------RI-V------I--------- 693
G.CM.01.01CM_4049HAN        ------------.----------------------V------F-------------------N----T---..........---FDE---..NM--I--E-------QQ-YN---I--S----------A---D----S--D--K------I--------I--------- 695
G.CU.99.Cu74                -------S----.----------------------V--------------L-----------N--------..........---VDE--G..NM--I--E-------AQ-Y---------------D--A---------T-D-S-------I--------I------T-- 704
G.ES.99.X138                -------S----.----------------------V-------E------------R-----N--------..........---YNE--D..NL--VQ-E---S---QQ-YN-L-D----------D--A-----------D-SH------I--------I--------- 697
G.GH.03.03GH175G            -------S----.-------M--------------V-------R------------------N----T---..........--TYNE--D..NM--IQ-EK------KQ-Y------------------A--------S--D--K------I--------I------T-- 701
G.KE.93.HH8793_12_1         -------S----.----------------------V--L----R------------------N--------..........--TYND--D..NM--IQ-----S---QQ-Y---------------D--A--N-----T--D--K------I--------IS-K-I---- 689
G.NG.01.01NGPL0669          -------S----.----------------------V-------Q------------------T----T---..........---Y-D--G..NM--IQ-E---S---GQ-YY--Q--------------S--Q--N-----D-S-------I--------I--------- 695
G.PT.x.PT2695               -------S----.----------------------V--------------------R-----N--------..........---YN---D..NL--VQ-E---S---QQ-YT-L------------D--A----D------D-SR------I--------I--------- 722
G.SE.93.SE6165              -------G----.----------------------L--------------------------N----V---..........---YNE--D..NM--I--E-------YQ-Y--L------------D--A-Q---------G--R------I--------I--------- 705
H.BE.93.VI991               -------S----.--Q----M--------------V--------------------------N----S---..........----DE--D..NM------KQ-----DE-YR-L-V----------D--A-----N-----S--------RI-------II--------- 708
H.BE.93.VI997               -------S----.--Q----M-------V------V--------------------------N----ST--..........----AE--D..NM-------Q-D---EV-YR-L-L--T---Q---D--A---D-------S---------I--I---A-I----I---- 701
H.CF.90.056                 -------S----.--Q-R--M--------------V-------R------------------N----S---..........---QSE--D..NM------KQ-S---EE-YR-L-V--T-------D--A--------T--D-SH------I--------I----I---- 694
J.CD.97.J_97DC_KTB147       -------S---K.---------R---------------------------------------N----S---..........---HDE---..NM--V--E---D---RI-YN--V-----------D--A----T---S--K-S------RI--------I----I---- 702
J.SE.93.SE7887              -------S---K.---------K------------V--------------------------N--------..........---Y-D--E..NM--IQ-E-------GI-Y------A----N--KD--A----TN-------S-------I----I---I----I---- 697
K.CD.97.EQTB11C             ------------.------QM-----------R--V-------R------------------N----S---..........---QSE--E..NM---Q-EK--S-H--T-YR------I-------D--A-----------D-S-------I--------I------T-- 697
K.CM.96.MP535               -------S----.-------------------R-----------------------------N----S---..........---W-E---..NM-----EK--G--SDT-YK------T-------D--A-----------D--K------I----I---I----A---- 688
```

```
                                                                   glycosylation NAS          glycosylation NHT
                                                    immunodominant region            glycosylation NKS          glycosylation NYT
B.FR.83.HXB2               SGIVQQQNNLLR.AIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWS..........NKSLEQIWN..HTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIKLFIMIVGGLVGLRIVFAVL  702
01_AE.CF.90.90CF11697      -------S----.---------------------V----------KF--L------I----N----T--............---YAE--D..NM--I---K------NQ-YE-LTK----------D--A-----N-----S---------I---V----I---------  702
01_AE.CN.05.FJ051          -------S----.---------------------V----------KF--LR-----I---------RT--............---Y-E--D..NM--I--E---S----QVYDILT------DR--KD----NT--------D--K------I--I-----I----I----  694
01_AE.CN.06.FJ054          ----R--S---Q.---------------------V----------KF--L------I----N----ST--............--TYD---D..-M--T--E---S---D--YDIRT------DR--KD---------------D--R------I--------I----I----  695
01_AE.HK.x.HK001           -------S----.-------M-------------V----------RF--L------I---------NT--............---Y-E---..NM--TQ-E---S---DQ-YAILT------D---KD-----Q--------S--K------T--------I----I----  698
01_AE.TH.01.01TH_R2184     -------S----.---------------------V----------KF--L------I---------ST--............---Y-E---..NM--I--E---S---NQ-YEILTK--D--DR--KD---------------D---------I--------I----I----  702
01_AE.TH.02.OUR769I        -------S----.---------------------V----------KF--L------I---------YT--............---YDE---..NM--I--E---S----Q-YEILT------DR--KD-------T------D-SK------I--------I---------  696
01_AE.TH.90.CM240          -------S----.---------------------V----------KF--L------I---------ST--...........-R-F-E---..NM--I--E---S---NQ-YEILT------DR--KD---------------D---------I--------I----I----  696
01_AE.US.00.00US_MSC1164   -------S----.---------------------V----------KF--L------I----N----ST--............---Y-E--D..NM--I--E---S---NQ-YEILTK-----DR--KD-------------------------I--------I----I----  698
02_AG.CM.02.02CM_4082STN   -------S---K.---------K-----------V--L-G--Q-------------------------S---...........--NYTD--D..NM--LQ-EK--G---GI-YK-------------D--A-----N-----D--K------I--------I--------F  692
02_AG.EC.x.ECU41           -------S---K.-------M-R-----------V--L---------------------------N----------.........--NYTD--D..NM--LQ--K--S---EI-Y--------------D--A-----------D--K------I--------I--------F  688
02_AG.GH.03.GHNJ196        -------S----.--K---E--R-----------V--L-G--R------------R-----N----ST--............--TYND--G..NM--L------S---DI-YN---V----------D--A-------S--S---------I--------I---------  705
02_AG.NG.01.PL0710         -------S----.---------K-----------V--L----R----------------------T----S---...........--TY-D--E..NM--LQ--K--S---QI-YQ-------------KD--A-----------D---------I-------FI------T-F  687
02_AG.NG.x.IBNG            -------S---K.---------K-----------V--L----R----------------------T----S---...........--TFND--D..NM--IQ-EK--S---DI-YN-------R------D--A-----------D--------RI--------I---------  693
02_AG.SE.94.SE7812         -------S---K.---------R-----------V--L-A-------------------------T----S---...........--TYDH--G..NM--LQ--K--S---HI-YD-------------D--A---------D--S-SS-----RI--I-----I---------  705
02_AG.SN.98.MP1211         -----------K.---------K-----------V--L---------------------------T----S---.........N--TFKD--D..NM--LQ--K--S---EK-YT----A---------D--A-----------D--------LI-L-V----I---------  700
03_AB.RU.97.KAL153_2       ------------.---------------------V-------------------------------T---..............--P-DE-*-..NM-----E-------G--YN-------------I-A-----------D-SK------I-------------I----  685
04_cpx.CY.94.CY032         -------S----.---------R-----------V--L-S--------------------N----S---...........---YND--D..NM--LQ--K------QI-YG-L------------D--A-----------S--K------I--------I----I----  700
05_DF.BE.x.VI1310          -----------Q.---------------------V-------------L--S---H--S-S----S---...........---EGE--D..NM-----EK--S--S-T-YR---Q--I----------S--Q--------D--K------I--------I------T--  707
06_cpx.AU.96.BFP90         -------S----.---------------------------------------------P-N-L--T---...........--TYDE--G..NM--I----------QQ-Y----L--T-------D--A-----N--S--D---------I--------I--------I  705
08_BC.CN.97.97CNGX_6F      -------S----.-------M-----------T-V--I-----------------------------S---...........---QQE--D..NM---Q--K--S---NT-YR-L-D------R--KD--A--S-KN--S--D---------I--------I----I----  703
09_cpx.GH.96.96GH2911      -------S----.---------K-----------V------------------------N----S---...........---QHE--E..NM--L------S---Q--YR--------------D--A--S-----S--D--------R---------I----I----  688
10_CD.TZ.96.96TZ_BF061     ------------.---------------------V----S--------------RH---N----S---...........----GE--D..NM-----E---D---G--YN------I---------Q--------S--S---------I--------I-----C---  686
11_cpx.GR.x.GR17           -----------K.---V-----K-----------V------------------------N----V---...........--TYNE--ENENM--IK-E---D---QT-YT-------------D--A-----------D-S------RI--------I----I----  690
12_BF.AR.99.ARMA159        -----------Q.---------------------V-------------L---------A-N----S---...........---Q-E---..NM---Q-EK--S--SHA-YR---D--------------A-----------D-SK-----RI--------I---------  688
13_cpx.CM.96.1849          -----------M.---------K---------------I--------------------N----S---...........---QSE--D..NM--LQ--K-VS---EM-YR-L------------D--A--Q---------S-SK------I--------I---------  706
14_BG.DE.01.9196_01        ------------.-------M-------------V-----------------------T--------...........----DD---..NM---X-EK--D---G--YN---Q------R---------------S--S----X---RI----I--XI---------  695
15_01B.TH.99.99TH_MU2079   -------R----.---------------------V-----------------------------------...........----NE--D..NM-----E---S---KE-YT---Q--------------A-----------D---------I--------I---------  693
16_A2D.KR.97.97KR004       -------S---K.-------M-R-----------V-------Q-----------------F----S---..........-RTQDD---..NM--LQ-E---D---GT-YR-L------------KD--A-----------D-S-------I--------I-----M-II  701
17_BF.PE.02.PE02_PCR0155   -------S----.---------------------V-----------------------------------.........N--T----D..NM-----E---S--SRE-YR-----------------A-----------S-S-----RI--I-----I---------  696
18_cpx.CU.99.CU76          -------S----.-------M-K-----------V----------------------P-----S---...........--TYND---..NM--LQ--K--S---Q--YT-------------D--A---------T--D-X------RI--------I---------  703
19_cpx.CU.99.CU7           -------S---K.------Q--R-----------V--L-------------------------T----S---...........-T-YXD---..NM--L---K-----PQI-YN-----------------A----GK-----D---------I--------I---------  648
20_BG.CU.99.Cu103          -------S----.---------------------V-------------L-----------N----T---...........---YNE--D..NM--I--E-------QQ-YT-------------D--A-------------S---------------I------T--  688
21_A2D.KE.91.KNH1254       -------S----.---------------------V----------------------H--P-N----S---...........--T-NN--Q..NM---Q-EK--EK--GT-Y-------I---N--K-------------S--D--K-----RI---V----I---------  708
22_01A1.CM.01.01CM_0001BBY -------S----.-------M-K-----------V-------------------------T----S---...........---Q-E---..NM--LQ--K--S---DI-Y--------------D--A----SN--T--D-S-------I--------I---------  687
23_BG.CU.03.CB118          -------S----.---------------------V--L-----------L-----------N-------...........--T--E--D..NM--IQ-E-------HQ-YN------Q------RD--A-----------D-S-------I--I-----I-----M---  694
24_BG.CU.03.CB378          -------S----.---------------------V-----F--------L-----------N----S---...........---XNE---..NM--I--E---D---QQ-YT-L----T-------D--A---------T--D--K------I--------I------T--  703
24_BG.CU.03.CB471          -------G----.---------------------V-------------L-----------N----S---...........---FNE--D..NM--I--E-------TH-YT-------------D--A---------T--D-S-----RI--------I---------  701
25_cpx.CM.01.101BA         -------S----.---------------------V------------------------N----T---...........---FDD---..KM--IQ-E---D---LQ-YD---Q----------D--A-----------D-SK-----RI--------I-------L-  713
27_cpx.CD.97.97CDKTB49     -------S----.-------M-------------V-----------F--------I---------S---...........---Y-E---..NL--I---K--S----Q-YE-LVK--D------KD--A-----------D-S-------I--------I-----C--C  698
28_BF.BR.99.BREPM12609     ------------.---------------------------------------------N----S---.................D..NM---Q-EK--D---DM-YT---KA---------D--Q--T--------D---------I-------------C---  714
29_BF.BR.02.BREPM119       -------S---K.-------M-------------V--I----------------------T--------...........----D----..NM--R--E---D---G--YT-L-N-------------------N-----D--H------I--------I------T--  695
31_BC.BR.02.110PA          -------S----.-------M-----------T-V--I----------------I---------S---...........---Q-D--E..NM---Q---------DT-YR-L-V----------D--A--Q-QN-----G--K------I--------I----I----  714
32_06A1.EE.01.EE0369       -------S----.---------------------V--LX--V-----------T-----P-N----D---...........--TF-E---..NM--I--E--------H-YT----------------A---------S--D-S-------I--------I---------  696
33_01B.MY.05.05MYKL007_1   -------S----.---------------------V--L----Q--KF--L------T----T----ST--...........-R-Y-E---..NM--I--E--------Q-FD-LI-------R--KD-------T---S--S--------RI--I-----I----I----  700
34_01B.TH.99.OUR2478P      -------S----.---------------------------------F--L------I---------ST--...........-R-Y-E---..NM--IK-E---S-H-GQ-YQILT------D---RD-------T------D-S-------I--------I----I----  689
35_AD.AF.05.05AF095        -------S----.-------M-K-----------V------------------------T----T---...........--NQSE--D..KM--LQ--K--S---DT-YT-L--A---------D--A-----N-----D-SK-------------I---------  690
36_cpx.CM.00.00CMNYU830    -------S----.---------K-----------V--L----R-----------------N----S---...........--TYDS--E..NM--LQ-----S---QI-YE---------R---D--A--Q-------D--K------I--------I---------  695
37_cpx.CM.00.00CMNYU926    -------S----.---------K-----------V--L-------------------H-T----S---...........----ND--E..NM--LQ--K--E---QI-YN----------------A----TN--S--S-------RI--------I---------  695
38_BF1.UY.05.UY05_4752     -------S----.-------M-------------V-------------------------N-R--TT--...........--TY-E--G..NM-----E--V----NE-YR--G---I-------D--AF-------S--D-SR------I--I-----I----I----  697
39_BF.BR.03.03BRRJ103      ------------.---------------------V------------------------N----S---...........--QQE--E..NM-----E---S--SNE-YR---I--I----------AF-------S--D-SS-----R---------I------S--  696
40_BF.BR.04.04BRRJ115      -----------K.---------------------V-------Q------L-----------N----KN--......STWRT---DD--D..NM-----E---D---RT-YN-------------D--A-N--T---S--D--H------I-----A------------  708
42_BF.LU.03.luBF_05_03     -----------K.---------------------V-------Q------L-----------N----S---...........--TYQE--G..NM-----EK-----SQE-YR---D--------------A-----------D-S-------I---------------S--  677
43_02G.SA.03.J11223        -------S----.---------------------V------------------------N----T---...........---YDX--G..NM--IQ-E--VG---QH-YDX---X---------X--A-----X-----D-S-----RI--------I---------  685
U.CD.83.83CD003            ------------.---------------------V-------ES-----L-----------T----S---...........----DN--D..NL---------S---QV-YG-L-D--K----S-KD--------------D---------I--------I------T-F  697
U.CD.90.90CD121E12         ------------.---------------------V-------EG-----L-----------T----S---..........-R--ND--D..NM------K--S---QI-YG---D--K-------------------------D-S-------I--I-----I------T--  695
U.GR.99.GR303              ----R--S---Q.---------R-----------V-----------I-E--A---------T----S---...........---Q.ELTYGKNM--L---K--S---QQ-YE------I-------D--A--TX--------K-S-----RI--------I---------  704
U.NL.95.U_NL_95_H10986_D1  ------------.--------------------------M------------------------N----S---...........---Q-E--D..NM--I--E---SK--GT-YR-------------D--A--T-----S---S-------I--------I------T--  707
CPZ.CD.90.ANT              H------A---Q.---T-------S---V------M----K--R-----SL---AD-VT-H-T----N--V.NFTQTCAKNSSDIQC--E..NM--Q----LVQ-S-GQ-YNILQIAHE---R-KK--Y-----S------D--Q------I------AI-----LLVLV  704
CPZ.CM.01.SIVcpzCAM13      ------------.---------------V------L-------Q---I--L------S--Y-T---KT--..........G--MSD---..NL--QQ--KL-T---GT-FG-L--A-S------KD-----Q--------D---------I-L-A---II----IMS-V  703
CPZ.CM.05.SIVcpzEK505      ------------.-----------S----------V--I----R---I--L------SV-Y-N---TT--..........N-N-YDT--G..NM--QN--EQVR--SGV-FG-L-QA-E--SI--KS-----Q-S------D--K------I---V-A-I--I--ISIIM  674
CPZ.CM.05.SIVcpzLB7        T------S----.-----------S----------V--I----------------------S---RT--............--TYNE--D..NM--------VR---EI-YG---QA-D---N--KK-----H-T------D-SH------I----I---IVC--I----  684
CPZ.CM.05.SIVcpzMB66       -------S----.-----------S----------V---------------L-----------S---TT-T...........---YDD--Y..NM---Q--K-VS---DV-YN-L-KA-T---N--K----------------D--S------I--I-----I-------L-  719
CPZ.CM.05.SIVcpzMT145      T-----------.-----------S----------V-------R------L---T--T--P---R--KT-G.........NISDYQV----..NY--QQ----V----G--YT-L--ANT------K------S--N--S--D---------M-LIV---II----C--IG  694
CPZ.GA.88.GAB1             -----------K.-----------SI--V------L-------Q---I--L------AV-Y-T----N--P.........GSN-TDD--G..NL--QQ--KLVS---GK-FG-L--A-S------RD-----Q--------D--K------I-L-A---II----IMT-F  693
CPZ.TZ.00.TAN1             -------Q---Q.-----------S----------M----K-IR-----SL---AN--V-HSS----LT-AEDSTKCNHSDA-YYDC----..NL--Q----LVE-S-GT-Y--L-KA-T-----K----------S---D--D--Q------IA-I--A-------LMFIV  705
CPZ.US.85.CPZUS            ------------.-----------S----------V-----------I--L------T--Y-T---DT--..........N-L-YDA--G..NL--Q----KVR--SGT-F----QA-E--NT--KS-----Q-S------D---------I-LIV-AS---I---GVIF  675
CPZ.CM.x.SIVcpzDP943       ------------.-----------S----------V-----------I--L------S--Y-T---ST-G.........A-T-YDE---..NL--QD--KRVK--SGV-FD---QA-E-XNT--RS-----Q-S------D-S-------I-L-A-A--I-I---GVIM  691
CPZ.CM.x.SIVcpzMB897       -------S---K.-----------S----------V---------------L-----------N---NT--.........KKLGTNLSF-D..NM---Q-EK--D---ET-YE-LTR------V------A---S------D--Q---------V------I-I-----M-  683
N.CM.02.DJO0131            ----------V-.-----------SI------R-KV--I----R---I-SL------T--Y-T---ET--..........N-T-YDX--G..NL--QQ---KVR--SGV-FE--XKA-E--NT--KS-----Q---------S---------IA--V-A-II-I--ISVII  679
N.CM.04.04CM_1015_04       ------------.------Q----SI--------KV--I----R---I-SL------T--Y-T---ET--..........NHT-YDS--G..NL--QQ--EKVR--SGV-FD---QA-E--NT--KS-----Q-T------D-S-------IA--V-A-II-I--ISVII  672
N.CM.04.04CM_1131_03       ------------.-X-X--X----SI------X-KV--IX---R---I-XL------T--Y-T---ET--..........XXTXYXX--X..NL--XX--XKVR--SGX-FX-X-QA-E-XNT--KS-----Q--X-----D--S-------IA--V-A-II-I--ISVII  694
N.CM.95.YBF30              ---------I--.-----------SI--------KV--I----R---I-SL------T--Y-T---ET--..........N-T-YDT---..NL--QQ--EKVR--SGV-FG---QA-E--NT--KS-----Q-D---S--G--K------IA---A-I--I--ISI-I  691
N.CM.97.YBF106             ------------.-----------SI------R-KV--I----R---I-SL------T--Y-T---DX--..........S-T-YDT--X..NL--QQ---KVR--SGV-FD---QA-E--NT--KA-----Q---------D--K------IA--V-A-II-I--IS-II  677
O.BE.87.ANT70              K------D----.--Q---Q--R-SX---R--R--L--L-TL-QN---SL---K---V-Y-S-K--RT-I..........G.N-S--D..TL--Q----Q-S-IS-T-YEE-QKA-V---Q--KK-----E---I---LD--K------IA-I---A---V-VIMI--  698
O.CM.91.MVP5180            K------D----.--Q------R-S----R--R--LQ-L-TLIQN--R-NL---K-----Y-S-K--T-----------GRYNDDS--D..NL--QQ--QH---VS-I-YDE-QAA-D-----VKA-----E--------D--K------IA-I---A-I-I-VIMII-  711
O.CM.96.96CMABB637         K------D----.--Q---Q--R-S----R--R--L--L-TFIQN---NL---X-R---Y-S-K--ST-G..........GGN-S--D..SL--QQ--QQ-D-VS-I-YEE-QKA-E---Q--KR-----E-----S-LD--K------IA-I---A-I-V---MVI-  715
O.CM.98.98CMA104           K------D----.--Q--HQ--K-S----R--R--L--L-TFIQN---NL---K-N---Y-S-K--NT-G.........GDRD-GN--D..NM--Q---QQ-S-IS-T-YEE-QKA-V---H--KK----SE---I----D--K------IA-I---A---V---MI--  719
O.CM.99.99CMU4122          K------D----.--Q---Q--R-S----R--R--LQ-L-TLIQN---NL---K-R---Y-S-K--NT-A.........NGTN-NT--D..NL--QQ--XQ-D-ISDV-YGE-QKA-D---Q--KK----NE-----D-LD--K------IA-I---A-I-I---MI--  710
O.FR.92.VAU                K------D----.--Q------R-S----R--R--L--L-TFIQN--HNL---KNR---Y-S-K--KT-G..........GDN-S--D..EL--QQ--QQ---VS-F-YEK-Q-A-E------K------E---I---LD--K------IA-I---A-I-V-V-ML--  712
O.SN.99.SEMP1299           K------D----.--Q---Q--R-S----R--R--L--L-TLIQN---NL---K-----Y-S-K--TT-T......NCTNTNK-DD--D..KL--QQ--QQ-S-VS-I-YEE-RNA-V---Q--KK-----E---I---LD--K------IA-I---A-I-V---MI--  717
O.US.99.99USTWLA           R------D----.--Q------R-S----R--R--LQ-L-TLMQN----NL---K--S--Y-S-K--ET-G..........GNLS--D..SL--QQ--QQVS-VS-F-YDK-Q-A-E---E--RA-----E---I---LD--K------IA-I---A-I-V---MI--  710
```

glycosylation NGS — glycosylation NAT

B.FR.83.HXB2	SIVNRVRQGYSPLSFQTHLPT..PRG.PDRPEGIEEEGGERDRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVT.......RIVELLGR.......RGWEALKYWW.NLLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQGACRAIRHIPRRIRQGLERI	854
A1.GE.99.99GEMZ011	--I-----------L--LA-H..-E-.----GR-K-----QG--------S-F---A-----------------FIS-AA.......-T------SSLKGLRL---G---LG.---G--G----S--IN-IDTI-----GW------IGR-F-----N---------A-KA	857
A1.KE.00.KER2008	AVI--------------LT-N..--E.L---GR-------Q--G------S-F---A-------------H---FI--AA.........-G---..........------G---L-.---I--G----T------DT---V--GW----L-IG-RIG--FIN---------F--A	852
A1.KE.00.KNH1144	AVIK--------V---I-A-N..-G-.L---GR-------PG-G------S-F---A-----N------------FA---A.......-T-----HSSLKGLRL---G---L-.---V-------T--I--VDTI-----GW---A--IG--IG--FL------------A	866
A1.KE.00.KSM4024	A-I---------------S-S..---.L----R----D--QG-G--------F--IA-----------------FI---........-T------SSLKGLRL---G---L-.---L--GR---I--IN-V-TI--V--GW------IA--VG---L------------A	830
A1.KE.00.MSA4069	--IK--------------T-N..--E.L----R---G---QG----------F---A------------R----FIS---.......-T-----HSSLKGLRL---G---LG.---I--VR---I--IN-VDTI--V--GW------IG-RIG---L----------F--A	857
A1.KE.00.NKU3005	-VI---------------T-S..-G-.-----R-D----GQG--------S-F---A-----------------FI--AA.......-T-----HSSLKGLRL---G---LG.---V---R---I--IN--DTI-----GW------IG-RIG---LN----V----F--A	863
A1.RU.00.RU00051	-VI-----------X--LT-N..-EX.----GR--DG---XGK-------X-F---X----X-----L-------FI--A-.......-T-----HSSLKGLRL---G---L-.---V--XR---S--IN-VDTI-----GW------IG-RV----XX----X------F-KA	851
A1.RW.93.93RW_024	TVI-------------I-T-N..-GD.----GR-------Q--N------S-F---A-----------C------FT--AA.......-T-----HSSLKGLRL---G---LG.---L--G--------IN-FDTI-----GW-------G-RVG--LL-V--------F--A	861
A1.SE.95.SE8891	--I-------------I-T-S..-G-.L---GR-------Q--N--------F---A-----------C-R----FI---A.......-T-----HSSLRGLRL---G---LK.---S--GR---L--IN--DTI--VI-GW------IG---F----F#----SK---K-A	849
A1.SE.95.UGSE8131	TVIK------------I-T-S..--D.----GR-------QG--------S-F---A--------R---------FIS-A-..............-Q..........G---LG.---L--IR---I--I--FDTI-----GW------IG-RIG---L----------F--A	858
A1.TZ.01.A173	--I---------------T-N..--D.L---GR-------QG--------S-F---A-----------------FT---A.......-T-----HNSLKGLRL---G---LG.---V--G----S--IN--DTI-----GW------IG-RI----YN------------A	863
A1.UA.01.01UADN139	--I-----------L--LT-H..-E-.----GR-------QG---T----S-F---A-----------------FIS-AA.......-T------SSLKGLRL---S---LG.---G--G----G--IN--DTI-----GW-------G-RF----CN---------A-KA	859
A1.UG.92.92UG037	-VI---------------T-N..---.L---GR-------Q--G------S-F---A-----N-----------FI--AA.......-T-----HSSLKGLRL---G---LG.---L--GR---I--IN--DTI-----GW-----T--RLG---LN---------F--A	855
A1.UG.99.99UGA07072	--I---------------T-N..---.L---GR-------Q---------S-F---A-----------------FI---A.......-T-----HSSLKGLRL---G---LG.---I--G----IN-IN--DTI---T--W------IG-RL----L---------A---A	860
A2.CD.97.97CDKTB48	-V-K------------IPT-N..-E-.L---GR-------QG--------S-F---A-----------------CI--AA.......-----V-HSSLKGLRL---G--HL-.---V--G----T--IR--DTI-V-----W------IG-R------N------------A	848
A2.CY.94.94CY017_41	TV----------V---IPT-S..-E-.----R-T--G---QG----------FF--A-----------------CI--AA.......-T-----HCSLKGLRL---G--NL-.---L--GR------I--FDTI-V-----W------IG-R-F---LN------------A	863
A.SN.01.DDI579	--I---------------LT-N..---.-----R-------Q---------S-F---------------------FT--AA.......-T-----Q..........G---L-.-V-L--GR---S--I---------V---W------IG-R-G--LL-----V------A	848
A.SN.01.DDJ369	--I----K----------S-S..-G-.L---R--------Q---------S-F---A-N---------------FT--AA.......-T-----HSSLKGLRL---G-Q-LG.---L--GR------I---D-I--T--GW-------G-R-G--LL------------A	864
A.SN.96.DDJ360	--I---------------T-N..-G-.L---GR-------Q---------S-F---V-----------------FT--AA.......-T-----Q..........G---L-.---L--GR---S--I---------V---W-------G-R-G--LL------------A	842
A.CD.02.02CD_KTB035	-VI----K--------IPT-N..-V-.V---GR---GD--Q---------S-F---A-----------------FI--AA.......-T-----HSSLKGLRL---S---L-.---V--G-Q--S--I---DTL-----AW------IG-RIG--FLNV--------F--A	852
A.CD.97.97CD_KCC2	-----------------PI-S..-E-.----GR---G---QG-T------S-F-E-A----------V---S--FI--AA.......-T-----HSSLKGLRL---G---L-.---V-------T----AF--L--T---W---A---G-RIG--FLN---------F--A	860
A.CD.97.97CD_KTB13	-V----------------LT-S..-E-.L---GK---G---Q--T-----AS-F---A-E---------------FI--AA.......-T-----HSSLKGLRL---G---L-.---T--C----T--I--FDTL------W---I---GHRIG--FLN---------F--A	882
B.AR.04.04AR151516	--X-----------------A.----.H---GX-------------R---T-F-----------F----RH---------...............-H.......-X-----L-.-X-----------------------------L---L-SIG---L------------S	850
B.AU.87.MBC925	-V---------------RF-A.----.-G-------------------E---D-F---F-V----------------------.......-------.........-----------------.------------------V----------------L-RIY---I---T-----F--A	857
B.BO.99.BOL0122	-----------------R---.----.--------------------T-RP--H-L------------------------.......-------.........-----I-----.------G-----------------T--------I--I--R-Y---L---T--------A	848
B.BR.03.BREPM2012	-V---------------IR---.Q--.----G----------------E---T-F-----------F------------AA.......-------.........----------L-.-----I-------I----Y-------------I-----RTG---L------------A	858
B.CA.97.CANB3FULL	-V----------I----L--A.----.-----------------------G----F-----------F-------------.......-------.........-----------------.----------------------I--I-------------I-----R-F--VL-V----------L	848
B.CN.05.05CNHB_hp3	--------------L---F--.Q--.-G---------------G--E----F-T---E-----------------A.......-------.........-----------G.-----I-------I------------G-----I---A-R-W---LN--T--------A	857
B.CO.01.PCM001	--------------------A..A--.---------G-----------G---H-F-GV-------------L---SI---A.......-------.........-------R---.S------------------V-----------I--G--R------NV-----------A	848
B.GB.83.CAM1	-L---------------RF-V.----.-------------G----T-G---T-F-------------------------A.......-------.........------------------.----------R-------FDTI-----------------R-----L------------L	854
B.GB.86.GB8	-------R------L------.----.----------------Q----------F---F----------------------.......-G-----.........-----------------.-----I-------I----T------------------R-Y---L---T--------A	847
B.GE.03.03GEMZ010	--A---------------R--A.Q--.-------T------Q-----GP-AT-F-E---I------------------AA.......-------.........-----------------.-----G--------N----A-----------IV----RIG---I----------F--S	845
B.IT.05.SG1	-------------------FAA.----.-------------------N--G-S---D-F--I--V--W---R--------------LA.......-T-------.........-----L-----.-------------------TI--------------T-RR-----L---V--------A	812
B.JP.05.DR6538	--------------L--LI-V..S--P----D------------G----GV----FF----N-----------RH---F---AA.......-------.........--------.----------S--IN---T-------------------R-G---L------------	863
B.KR.05.05CSR3	--------------------A.----.-------------------------SQS-H-F-EI--V---------L-----------A.......-------.........-----I-----.S----------------FD---------------A--R-G--LL------------A	867
B.NL.00.671_00T36	-------K------------A.----.-------------------G-GP----F--I--V---------------------V.......-------.........--------------.------I---RG-----F--I-------------TI-R-F---L------------	853
B.RU.04.04RU128005	--------------------A..R--.------T----------K-G---D-F--I--V--X--------------------------.......-------.........-----I-----.-----------I------------------I--LL-R-G---I---T--------A	848
B.TH.00.00TH_C3198	AV------------L---F-A..Q--.-G-----------G---G--E-------T---E-------------S-----A-.......-------.........---------C-.------I-----------------V----R---I---GLR-Y---L---T--------L	850
B.UA.01.01UAKV167	-----------------R---..Q--.---------G-----Q-----GP--T-F--I--V-------------------A.......-------.........-----------------.------------S--I--F--I---------KI--IA-TIG---L------------F	863
B.US.04.ES10_53	-----------------IRP-A..R--.---------G-----N-----GP--D-F--I--V-------------------------.......-------.........-----------------.-----I-------I--F--I----------------I--RV-----------------S	844
B.US.99.PRB959_03	---K----------L--L--A..---.---------G-----Q---------D-L---F-----------------------A.......-G---.........-R-I---------------I-----I------------------------I---L-RIG---L---------F--A	847
C.AR.01.ARG4006	-----------------LI-N..----.-GGLG--------Q-K-------S-F-S-A-------------RQ----I--AA.......-AA-R---.......S---I---LG.S-V---GL---K--I---DTI----------I---I-RIWEI-CN---------F-AA	854
C.BR.04.04BR013	--------------L--LT--..----.---LG--------Q---------S-F-P-A-------------------I--AA.......-A------SSLRGLQ----I---LG.S-V---GL---K--I---DTI----------I--LI--IW---LS------H-F-AA	863
C.BW.00.00BW07621	-----------------LT-N..Q--.---LG--------Q---------S-F---F-----------------FI-VAA.......-A------SSLKGLQ----I---LG.I-V---GL---K--I--FDTI----------I--AI-RI----CN---------F-AA	848
C.CN.98.YNRL9840	-V---------------LT-N..-G-.---LGR-------Q-KNG-------F---A-----N--I--------FI-VTA.......-V------NSLRGLQK--Q----LG.S-V---GL---K--I---DTI--V-------I-KL-YSI---------------F-LA	866
C.ET.02.02ET_288	-L---------------LT-N..--E.---LR--------Q-----------F--IA---I---------------I--AA.......-A------SSLKGLQ----I---LG.SIV---GL---K--I----T---V-------I--L--RFW---CNA--------F-AA	855
C.GE.03.03GEMZ033	-----------------LT-N..--E.L--LR--------Q---------S-F---A-----------------FI--AA.......-AA-----SSLKGLQ--------LG.S-V---GL---K--IN--DTI----------I--LI-RI--G-YNV--------F-AA	863
C.IL.99.99ET7	--A--------------LI-N..Q--.L--LGR---G---Q--N--------F---A-------------H---FI---A.......-AA-----SSLRGIQ--------LG.S-V---GL---K--I---DTI--T-------I-----RGW--LCN---------F-AA	851
C.IN.99.01IN565_10	--I----------L---LT-N..---.---LR--------Q-K---V-----F---A-----------------FI-VTA.......-G------SLLRGLQ----I---LG.S-V---GL---K--I--FDTI----------IL-L--RI----YN---------F-AA	858
C.KE.00.KER2010	-----------------LT-N..---.---LGR-------Q-K-G-------F---A-----------------CI---A.......-V------SSLRGLQK-------LG.-CV---G----R--I---DTI----------I--LI-RI----LN--T--------AA	860
C.MM.99.mIDU101_3	-V---------------LT-N..-E-.---LRR-------Q-KS--------F---A-----N----C----T-FISVTA.......-V------SSLRGLQG-------LG.G-V---GL---K--I---DSI--T-------I-A--HRI-----N---------F-AA	863
C.MW.93.93MW_965	-----------------LT-N..----.L--LG--------Q---------S-F-P-A--------------Q---FI---A.......-A------SSLRGLQ--------LG.S-V---GL---R--I---DTI--T-------L-----RIR---CN---------F-TA	850
C.SN.90.90SE_364	-----------------LI-N..----.---LGR-------Q-----------F--IA-------------R----FI---A.......-A----IQ.......----T---LG.S-X---GL---K--I---DTI--T-------I--L--RI----SN---------F-AA	839
C.SO.89.89SM_145	-----------------LT-N..----.---LGR-------Q-K---------F---A-----------------FI-VAA.......-A------SSLRGLQ--------LG.--G---GL---K--I---DTI--T-------I--LA-RI--GV-N--T------F-AA	859
C.TZ.02.CO178	----------------ILT-N..----.---LG--------Q-----------F---A-N---------------FI--AA.......-A------SSLRGLQ----T---LG.GCG---GL---K--I--FDTI----------I--L--RIG--.......-----F-TA	860
C.UY.01.TRA3011	-----------------LT-S..----.---LGE-------Q---------S-F-P-A------------------I--AV.......-A------SSLRGLQ----I---LG.S-V---GL---K--I--FD------------I---L-RIW---CN---------F-AA	856
C.YE.02.02YE511	--I----------L---LT-N..S--.---LG--------Q--N---Q-A--F--VV-N---------------FT---A.......-A------SSLKGIQ--------LG.S-V---GL---K-----VDTI----------I--I--RS---LCN---------F-AA	860
C.ZA.04.04ZASK164B1	A-------------L-IPSQN..----.L--LGR---G---Q-KE-------S-F---A-------------Q---FI-VIA.......-TL------.........----T---LG.S-I---GL---R--I---DTI--TT------I--L-LRI-----N--A------F-AA	862
C.ZA.05.05ZAPSK240B1	-L------------L--LI--P.T--.---LG--------Q-K-------S-F-S-------N-----------FI--AV.......-A-----HSSLKGLQ----I---LG.S-V---GL---K---N--DT-----------L--LI-SF--G--N---------W-AA	845
C.ZM.02.02ZM108	-----------------LI-N..----.---L-R-------Q--G-------S-F---A-E---N----G------FI---.......-A-----HSSLKGLQ----I---LG.S-V---GL---K------DTI----------IL-LL-RI-----NT--------F-AA	877
D.CD.83.ELI	-L---------------L--A..----.------T-------G----V--L--FS----------------------I--AV...............-------.........----DI---L-.-----------R---S--FD-I--------------II-R----VLN-----------S	851
D.CM.01.01CM_0009BBY	-M---------------L--A..----.--------------Q--GT-------FS--------N-------------I---A.......-T------.........-------------L-.------G-------I--Y----V-----------LG-R-F--AL---T--------A	863
D.KE.01.NKU3006	-L---------------L--A..----.--------------QG-G--------LS--------N-------------I--A-.......-------.........----------I--L-.------I-------I---D----T--G----A--I--RVV---LN--T--------A	867
D.KR.04.04KBH8	-L---------------PF-A..----.-------T--G---Q--GK-----S-LS--------N-----------------A-.......-------.........----------V--LG.------I-------IR-FD-----------------RR-F--VL----------F--A	840
D.TD.99.MN011	-M------------L--L--A..----.--------------Q--G--------FS--------N------Q----I--AA.......-T-N---.........----------L-.--------R------I---D-----------I--I--R-A---LN------------A	843
D.TZ.01.A280	-L------------L--L--A..S--.--------------QG-G--------FS--------N------RH----I--A-.........---G---H.........-----I--L-.------I-------I----V-----------I--II-R-F--VL-----V-------A	850
D.UG.99.99UGK09259	-L---------------L--A..----.--------------PG-GN-------FS--------N---------------AA.......-------.........----------I--L-.------I-------I--FD---V---------I--I----F--.#NV-------SA--A	851
D.YE.01.01YE386	-L-KK------------L--A..----.----------------G--------FS--------N-------------I--A-.......-------.........----------I--L-.------I-------I--F-T---V-------I--IA-RVG---I-------------F	840
D.YE.02.02YE516	-M------------L--L--A..----.--------------PGKGT-------FS--F-E---N---Y---Q----I---A.......-T-D---.........-------------L-.--------R---I--I--F-------------I--I--R-G---LN------------A	844
D.ZA.90.R1	-V---------------L--A..Q--.---------------G--K-------FYS-------N------R---------AA.......-------.........-------------L-.--------------I---DTI---T-------------R----VLNV-V---------L	856
F1.AR.02.ARE933	-------K------L--LI-S..--E.---------G---QGK---V---T-F---A-----N------RH---FI--AA.......-T-D.......RGLR-------LLG.--TL------RI--I--F-T---V----------A--R-V--VCN-----------A	843
F1.BE.93.VI850	-------K------L--LI-S..---.---------G---QGK---V---T-F---A-----N------RH---FI--AA.........---D.......RGLR--------LG.--TR----------I--F-T---V-------I---L-R-G--VLN---------A--A	829
F1.BR.01.01BR125	-------K------L--RI-S..-EE.--------GG--TQGK-T-V-----F---V-----N------RH---FI--AA.........---N.......KGLR---I--LLG.--AL-----I----IN---T---V----------AL-R-G--VLN---------T--A	847
F1.ES.x.P1146	--L----K------L--LI-S..---.---------G---Q-K---------FF--V-E--WN---C-CRL---FI--AA.......-T-D.......RGLK--------L-.--T---GR------I--F-TI--V-----------L-RIG---INV--------F--S	859
F1.FI.93.FIN9363	-------K------L--LI-A..-TE.---------G---QGK---V-----F---V-----N------RH---FI--AA.........---D.......RGLR--------LG.-II------------I--F-T---V----------AL-R-V--VLN-------RV--A	840
F2.CM.02.02CM_0016BBY	--I-----------L--LI-S..----.----G--------Q-K---V---S-F---A-------------RH---FI--AA.......-T-D.......RGLKG---V---L-.--A---GR---I--I----T---V----------IL-R-G---L----------F--A	843
F2.CM.95.MP257	--I-----------L--LI-N..S--.-E--G--------Q-K-------S-F---A---F----V----C--NFI--AA.......-T-D.......KGLK----V---L-.--A---G-------I---DR-----------I--IL-R-G--VLN------------A	848
F2.CM.97.CM53657	AVI-----------L--LT-S..R-E.-E--G--------Q-KT--V---S-F---A-------------RH---FI--AA.......-T-N.......KGLI----I---LG.--A---GR-I----ID---T---V-------I---L-R-G---L---------A--A	847
G.BE.96.DRCBL	-----------------L-HH..Q-E.----A---G---Q---------S-F---A-----------------FI--AA.......-T------NSLKGLRL-------L-.---L--AR------IN--DTI-----NW-------A-R-G--VLN------------A	852
G.CM.01.01CM_4049HAN	-----I-----------LTHH..Q-E.---LGR-------Q-KN------S-F---A-----------------CV--AA.......-TA----HSSLKGLRL---G---CG.---LH-G-------IN--DTL-----NW-------A---G---L------------A	854
G.CU.99.Cu74	-LI----K---------PTHH..Q-E.-----R---G---Q--N------S-F---A---X--------------V--AA.......-T-----HSSLKGLRL---G---L-.---V--G-------IN-VDTV-V---NW------I--RT---FLN---------F--A	863
G.ES.99.X138	--I----K---------LTHH..Q-E.----GR----D--Q-K-K-----S-F---A-------------------A-.......-TX-X--HRXLKGLKL--DG---XG.-V-LN-G-------IN-FDXL-----XW---A--LG-RVGX--LN--V---X-----	856
G.GH.03.03GH175G	-----------------LTHQ..Q-E.---L-RT--G---Q-----V---S-F---A-----------------FA--AA.......-T------SSLKGLRL---G---L-.---L--V-------I---DTV-----NW-------I-R-G---LN--T------F--A	860
G.KE.93.HH8793_12_1	-------K---------LTHH..Q-E.-----R-------Q-K-------S-F---A-----------------FI--AA.......-T-----HNSLKGLRL---G---L-.---L--GR------IN--DTI-----NW------I--R-F--FLN--T---------A	848
G.NG.01.01NGPL0669	-----------------LTHH..Q-E.---------R---QG----------F--IA---V-------------FIW--A.......-T------SSLQGLRL---G--HL-.---L--G----I--I---DTV-----NW------LA-R-G---LN---------A--A	854
G.PT.x.PT2695	--I----K---------LTHH..Q-E.----GR---GD--Q---K-----S-F---A-------------H---FI--AA.......-T------SSLKGLSL---G---L-.---L--GR------IN-IDTV-----NW---A-----RVG---LN--V---------L	881
G.SE.93.SE6165	-----------------LTHH..Q-E.---------G---QG-G--V---S-F-P-------------------SI---A.......-T------SSLKGLRL---G---L-.---L--GR------I---DTV-----NW-------A-R-----LN--T---------A	864
H.BE.93.VI991	--------------L--LI-N..Q--.----RE-------Q-----------F-P-V-E---N------R------S--A.......-T---.......GR-------LLG.---L--G-------I----T-----------I--L--R-W---L----------F--A	860
H.BE.93.VI997	---G-------------LI-N..---.---------------Q--G--V-----F-PIV-------------RL---S----I.......-T---.......GR--R-----L-.-------G-------IN---T---V-------I--I--R-W--VL-------------	853
H.CF.90.056	-----------------LV-N..---.------T--G---Q-----V-----F-PVV-------S----RL---------V.......-T---.......GR--R-----L-.-------G-------ID---T----------GI-VI--R-W---L----------F--S	846
J.CD.97.J_97DC_KTB147	A-------------L--LI-N..-T-.V---GE---G---QG-T------S-F---A-----------------FV---A.......-T--T-.......GR----I---LG.--VC--G-------I-----------A----I--I---IF---L--------X....	850
J.SE.93.SE7887	A----------------LI-N..-TE.A---G----G---QG-T--------F---A-----N-----------FV--AA.......-T-GT-.......GL----I---LV.--VW--G-------I----T------------I--IA-R-F---L------------A	849
K.CD.97.EQTB11C	-V---------------LT-S..----.---------G---Q-K---V---S-F---A-----N------RH----V--A-.........-LD.......RGLKGS------L-.--IL--G--I----IN---T-----------I--I-YR-F--LL----------F--L	849
K.CM.96.MP535	-V---------------LI--..S--.A------------Q-KN--V---S-F---A-----N------RQ--N-I----..........-L-.......RGLRG-------L-.--V-----------I----T------G----I--IG-R-F--LL------------A	840

```
                              glycosylation NGS                                                                        glycosylation NAT
                                                            *                                                                                        *
B.FR.83.HXB2              SIVNRVRQGYSPLSFQTHLPT..PRG.PDRPEGIEEEGGERDRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVT.......RIVELLGR.......RGWEALKYWW.NLLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQGACRAIRHIPRRIRQGLERI 854
01_AE.CF.90.90CF11697     -----------------LTHH..Q-E.-----R---G---EGK-------S-F-S-A-----------------FI--AA.......-T-----HSSLKGLR----G---LG.---S--G----T--IT-FD-I-V---GW---------R-W--LI-----------A 861
01_AE.CN.05.FJ051         -----------------PIHH..Q-E.---------G---QG----V---S-F---A---V---------L---FIS-AA.......-T-----HSSLKGLR----G---LG.---L--G----I--I---DT------GW-------A-RIW--FL-----------A 853
01_AE.CN.06.FJ054         ----------------IPSHH..Q-E.-----K-------Q-----V---S-F---A------------R----FISS--.......-A------SSLKGLR----G---LG.---I-------I--I--FD-S---I-GW------IA---W---L-----------A 854
01_AE.HK.x.HK001          ----------------IP-HQ..Q-E.----GR------GQ-----V---S-F---L-----N-----------FI---........-T-----HSSLKGLRL---G---LG.---L--G----I--I----T-------W-------A-R-W---L-----------T 857
01_AE.TH.01.01TH_R2184    -----------------PTHH..Q-E.-----R---G---Q-----V---S-F---A-----------------FI--AA.......-T-----HSSLKGLR----G---LG.---L--GK---I--I---DT------GW------AA-R-W---L-----------A 861
01_AE.TH.02.OUR769I       ----------------IPSHH..Q-E.-----R---G---Q-----V---S-F---A-----------------FI--AA.......-T-----HSSLKGLR----S---LG.---V--G---RI--I--FD-I-----GW------AA-R-W---L-----------A 855
01_AE.TH.90.CM240         -----------------PSHH..QKE.---------G---QG----V---S-F---A------------------T--AA.......-T-----HSSLKGLR----G---LG.---L--G----I--I---D-----A-GW-------A---W---L-----------T 855
01_AE.US.00.00US_MSC1164  -----------------LSHH..Q-E.---------G---QG----V---S-F---A-----------------FIS-AA.......-T-----HSSLKGLR----G---LG.---L--G----I--I---D-I---I-GW-------A-R-W-------T--------A 857
02_AG.CM.02.02CM_4082STN  N-I--------------HHH..QGE.-G--GR-D-----Q-----V---S-F---A-----------------FV--AA.......-TA----HNSLKGLRL-------LG.---S--G-------IN-FDTI-----NW------IG-R-G----N--T--------A 851
02_AG.EC.x.ECU41          T-I-------------LTHH..Q-E.-----R---G---Q---K-V---S-F---A---------C--Q---FV---V.......-T-----HSSLKGLRL-------LG.-V-S--G-------IN---T------N-------L--RISG---N-----------A 847
02_AG.GH.03.GHNJ196       T-I-------------LTHH..Q-D.-G---R-------QA-A--V---S-F---A-----------------.................-HSSLKSLQL-------L-.---T--G---R---I---DTI-----NW------IG-RIA---CN-----------A 857
02_AG.NG.01.PL0710        N-I-------------LTHH..QGE.---R-R---G---Q-----V---S-F---A-----N---------K--V--AA.......-T-----SSLKGLRL-------L-.-I-S--G-------IN--DT------NW------IG-R-GG---N-----------T 846
02_AG.NG.x.IBNG           T-I-------------LTHH..Q-E.-----R---G---Q-K---V---S-F---A-----------------I--AA.......-T-----HNCLKGLRL--G----L-.--IS--V-------IN---TI--V--NW---A--IG-RVG----N---------F--A 852
02_AG.SE.94.SE7812        A-I-------------LTHH..Q-E.-G---R---G--GQ-----V---S-F---A-----------------FVS--A.......-T-----H.......--------L-.---S--G-------I---DTI--V--NW------L--R-G---LN---------F--A 857
02_AG.SN.98.MP1211        A-I-------------LTHH..Q--.A--------G---Q-----V---S-F---A---------L------FV--AA.......-T-----HSSLKGLRL------HL-.---S--G-------IN--DT--V---NW------I--RTG---CN-----------A 859
03_AB.RU.97.KAL153_2      -----G----------R--A..Q--.----------D------T------RF-----------F-I--H-------AA.......-----------.......---------.------I-----S--IN-IGTI-----GW------IG-RF---M-N---------A-KA 837
04_cpx.CY.94.CY032        --------------L--LI--T.Q--.L---G-T------Q--S--------F-P-------N------RH--N-----A.......-T----.......I--------L-.-F-L--G-------IN-F-T-----------I--A--R-----CN-----------A 853
05_DF.BE.x.VI1310         -------K---------PF-A..---.-----ET--G---Q--G--------FS--------N-------H----T---V.......---------.......---------L-.S-P----R-------I----T--VV----------AL-R-G---LN-----------A 859
06_cpx.AU.96.BFP90        --------------L--LI-N..-T-.A---GE---G---QG-T---------F---A-----------------FG--AA.......-T--I---.......----I---LG.--IC--G---Q---I--FD-A-----NW----A----RIF--FLNV--------F--A 857
08_BC.CN.97.97CNGX_6F     ----------------ILT-N..-G-.-G-LGR-------Q-KTS-------F---A-----N-----------FI-LTA.......-G------NSLRGLQ--------LG.S-V---GL---K-TI--VD-I-----------I-NR---I----HNV--------F-AA 862
09_cpx.GH.96.96GH2911     -V--------------LT-N..--E.---------G---Q-KV------S-FF--A-----N-----------FI--AA.......-T--I--H.......-V-QI---LG.--A---G-------I----T---V-----------L-R-G---L----------F--A 840
10_CD.TZ.96.96TZ_BF061    -L------------LH-L--A..---.--------------G-G------D-F-T-------N-------------I--A-.......-----------.......M----I--L-.-------I-------I---DTI--E--GW------I--R-V---LN--T--------A 838
11_cpx.GR.x.GR17          -----C----------ALT-S..QQE.----G-TK-G---QG-TS-----S-F---A-----N----L--Q---FI---A..........-T---.......-----S---LG.--A---G----S--I-----------------I---AHR-L---LN---------F--A 842
12_BF.AR.99.ARMA159       -------K------L---I-S..--E.---------G---QGKV--V-----L-T---------------------V--.......-----------.......----------.------R-------I----T---V--G----IL-AL-RIG--VLNA---V-------A 840
13_cpx.CM.96.1849         ---KS-----------LTHP..K-E.-----R---G---Q---------S-F---A-----------------FI--A-.......-T-----HSSLKGLRL---G---L-.---V--GR---I--IT--DS-------W---I--TG-RIG-G-L----------F--A 865
14_BG.DE.01.9196_01       X-I----K--------LTHH..Q-E.----GR-------Q-K-K-----S-F---A---X-------------FI--AA.......-T------GSLKGLRL---G---L-.---L--G-------IN-FDII-----NW--XA--I--RIG--FLN--T-------XX 854
15_01B.TH.99.99TH_MU2079  --------------L--PIHQ..Q-E.A----E---G---QG-NG-V---S-F---A---------------K--VS-A-.......-T------SSLKGLRL---G---LG.---L--G----I--I---DT------GW-------A-R-W---LR---------F--A 852
16_A2D.KR.97.97KR004      -V---------V---IPP--..-ED.---HGR--DG---QG----V---S-F---V-E----------------CIS-AA.......-T-----HSSLKGLRL---G---L-.---L--GR------I--F----------W------I--R-----IN-----------A 860
17_BF.PE.02.PE02_PCR0155  -------K----------P-S..--E.---------G---QG----V-----F---F---------C--Q---------.......------------.......-T-------.--------------I----T---V----------AL-RVG---FN------------ 848
18_cpx.CU.99.CU76         ---K------------LI-N..QL-.------T-G----QG-G--V---S-F-T-A-E-----------------W-AA.......-T------SSLKGLRL-------L-.---V--G--I-S--IN--DTV---T-NW------I--RTG---LN---------A--A 862
19_cpx.CU.99.CU7          -----------------LTRH..RG..---H-R---G---Q-K---------F---A-----------------FIS-AA.......-T-----H.......S---G---L-.---L--G----T--I----T---V--GW---A--II-R-W--VL-----------A 799
20_BG.CU.99.Cu103         --I--------------THH..Q-E.--------GG---Q-N---V---S-F-P-A---------------------AA.......-T------SSLKGLRL---G---L-.---G--G-------IN--DTV-----NW---I--I--RGF--FL----------F--A 847
21_A2D.KE.91.KNH1254      -L--------------L--A..---.----------------------FS--------N-------------I---V.......-----------.......-Q-G---L-.--------R-------I------------------L-IL-R-G--FLN---------F--A 860
22_01A1.CM.01.01CM_0001BBY --I-------------LT-N..QE-.L---GR---G---Q--NK-----S-F---A-----------------FI--AA.......-T-----HSSLKGLRL--G---L-.---I--G-------IT--D----T--GW------IT-R-W--FL---T--------A 846
23_BG.CU.03.CB118         --I-------------LTHH..Q-E.---------G---Q--G-----AS-F-PIV---------------------AA.......-TL-----SSLKGLRL---G---L-.---L--G-------IN--DTV-----N----I--I--RGF--FL----------F--A 853
24_BG.CU.03.CB378         --I-------------LTHR..Q-E.-----E---G---Q-KG------S-F-P-A---------C-----------AA.......-A-----HSSLKGLRL---G---L-.---L--GR------IN--DTV-----N----I--I--RVF--FL--------------L 862
24_BG.CU.03.CB471         --I--------------THH..Q-E.-------------Q--A--V---S-F-P-A----------------------A.......-A-----HSILKGLRL---S---L-.---L--GR------IN-FDTV-----N----I--I--RGF--FL----------F--A 860
25_cpx.CM.01.101BA        -----------------THH..Q-E.-------------Q-S---V---S-F-P-V-----N-----------FI--AA.......-T-----HSSLKGLRL---G---LG.---L--G-------IN-VDTI-----NW------GI-R-G---Y----------F--A 872
27_cpx.CD.97.97CDKTB49    --I---------V---I-T-N..-E-.LN--GR-G-----QG----T--A--F---V-----N-----------FV--...............-HSSLKGLKL---G--HL-.-I-L--GR------IN--DTI-----NW---------R-G---LN---------F-IA 850
28_BF.BR.99.BREPM12609    -----------------R--V..Q-E.----G---------N------AD-F--I--E------------------------.......-----------.......---------.------G-----------------------------IL-RI-----N--T--------A 866
29_BF.BR.02.BREPM119      --I----K--------L--A..A--.---------------KG------H-F---A--------I----L------T--.......-T-G------.......---------.--------------------TI-------------AG-RVG--VL---T--------A 847
31_BC.BR.02.110PA         -----------------LT-X..---.---LG--------E-K-------S-F---A-----------------I--AA.......-A------------.......-----I---LK.S-V---GL---K--I---DTI-----------I-----R#----YN----f-f-F-AA 863
32_06A1.EE.01.EE0369      ------------V-L--RI-N..-T-.A---GE---G---QGNTK----AT-F---V-----------------FG---A.......-T--T--H.......----I---LK.--VC--G-------IN-IDTI-----DW-------G-RVF--FLN-----------IA 848
33_01B.MY.05.05MYKL007_1  -----------------PIHQ..Q-E.---L-R-------QG----V---T-F---A---------I-------I--AA.......-T-----K....GLRQ---G---LG.---S--GL---I--I-----------GW-------A-R-W---L-----------S 855
34_01B.TH.99.OUR2478P     -----------------PTHH..Q-E.-----R---G---QG-N--V---T-F-S-A---------------------A.......---------H.......-----R---.------I------------------I-----R-Y---L-----------A 841
35_AD.AF.05.05AF095       A-I---------------T-S..-GD.L---GR-------QG-T------S-F---A-------------H---FI---A.......KT-----HNSLKGLRL---G---L-.---L--GR---T--T--VDTL-----GW--S-V-IG-RFG---LN-----------A 849
36_cpx.CM.00.00CMNYU830   NVI----------P---LTRH..Q-EE-----R-K-G---Q---K-V-----F---A-----------------FI--AA.......-T-----HNSLKGLRL---G---L-.---V--G----T--IT--D---VT--GW-------A-RTG---LNV-T--------A 855
37_cpx.CM.00.00CMNYU926   -----------------LTHH..Q-E.-----R---G---QG----V---S-F---A-----------------FI--AA.......-T--I--HSSLKGLRL---G---L-.---V--G----I--IN--DT---V--GW------IA-R-----L----------FGKG 854
38_BF1.UY.05.UY05_4752    --------------L-I-A-S..S-E.---------G---QGK---V-----F-T-V-----N----L--L---FI--AA..........---N.......RGLN----I----G.-------R--R---I----T---V----------G-L-RIG--VLN---------F--A 849
39_BF.BR.03.03BRRJ103     -------K------L--LI-S..--E.---------G---QGK---V-----FF--V-----N------RH---FI--AA..........---G.......RGLT-------LLG.-I-L--G-------I----TI--V-----------AL-RFG-G-L----------F--A 848
40_BF.BR.04.04BRRJ115     -----------------R--A..---.------T-G---------G---T-F-------------------------A.......--L----------.......----L---L-.S-----------------TI----------I--AI-RIG-G-L----------F--S 860
42_BF.LU.03.luBF_05_03    ---------X-------RH-A..Q--.------X-----------RX--D-F-T---V--------I----------AA.......-----------.......---------.------R-------I-------------------XL-RXG--XL---T--------A 829
43_02G.SA.03.J11223       ---X------------LTHH..Q-E.---------G---Q-----------F---A-E-V-N-----------FI--AA.......-T------SSLQGLRL---G---L-.---L---R------IN-VDTI-----GW-------A-R-----IN---------F--A 844
U.CD.83.83CD003           --I-------------L---.---.----GRT------E-NN--V-----F---A-E------I----------I---V.................KGLR--------LG.--VL--G-------I--------V-------I---G-RI----LN---------F--A 845
U.CD.90.90CD121E12        --I-------------LI--.---.----G-T------K-NT--V-----F---A-----------------FIS--A.................KGLR--------LG.--V-----------I--------V-------I--L--RVG---L----------F--A 843
U.GR.99.GR303             -V-K------------IPI-S..---.L-----------Q-K-------S--FS-A-----N-----------FI---A.......-T--I--T......#G--QX---LGfTF-L-Gf---QK--T---DI---V--GW------TG-RIG---L----------F--A 855
U.NL.95.U_NL_95_H10986_D1 --I-------------LF-S..S--.------T--R---Q-----------F---A-----N------RH--------V.........--L-.......RGLRGS------L-.S-V------------IN---T-----------I--IG-R-F--VL----------F--A 859
CPZ.CD.90.ANT             -CLRK-----H-----IPTQN..QQD.-EQ--E-R----RK--I-WRA-QH-FF--L-V--T-IIQWI-QIC-TC--NLW.......AVLQH-C-ITFRLCNHLENN-STL-.TII.......RTEIIKNIDRL--W-G-K--SILLAL-TIV-I--EV----------IA 856
CPZ.CM.01.SIVcpzCAM13     -VIR------------L--LI-A..Q-E.-GTLGETGG----PGSG--V---T-CWP-------N-VIW--RS-TS-AC--W.......-QL-K--QLLINILRLIQ-K-TLLR.GII---GR--RT--T---D------G-----I-RA--IVF-I-GN-----------T 862
CPZ.CM.05.SIVcpzEK505     -M-A-----------L--LI--..T--.------RT--DA--L-NG--V---S-F---A-E-F-N-L--L----T-C-S-LR.......-TL---RQNIHKGLQLLN-LRI-L-.GIIA--GR---I--IN--DT--V---------I--L--RIG-G-L-----------A 833
CPZ.CM.05.SIVcpzLB7       A---------------LS-I.----.LAQ-G-T------QGT---V--LD-F---A-N---D----L----T--V--AR.......-TL-IV-QYTLKGLRLV----L-LQ.GI----GK---T--T---DT-----------I--I--RIG-G-L----------F--A 843
CPZ.CM.05.SIVcpzMB66      ----------------LI-A..R-D.R----E---G---P-NV------S-F---A-N---D----L-----------L.......-TL--V-QTLLKGLR--R---IHLR.GI----G----T--I---DT-----------I--IA-RFG-G-LN-----------A 878
CPZ.CM.05.SIVcpzMT145     -LI----K------L--LI-A..NQ-.--GLGET-KG---NV-G------S-F-P-V-E---N-LS-L--Q---CASLIW.......-LL-I--QYSLRGVQQIGTL-HQLR.GT----TT-I----I---DT-----G-----IL-AITRLG-G-L-------------A 853
CPZ.GA.88.GAB1            -V-R-----------L--LI-V..Q-E.QG-LGE-D-G---Q--S--V---E-C-P-------N-GIW--QS-TS-ACN-W.......-QLKT--HLILHSLRLLR-R-CLLG.GII---GK---I--I---D------------I--AF-VTL-I--N-----------A 852
CPZ.TZ.00.TAN1            NV-KQ-----T--FS-IPTQA..EQD.-EQ-G--AGG--G--NI-WTPSPA-FFSIV-E---N-LIWI-QTFQNFIW-LWISLQALKQGIIS-AHSLVIVHRTIIVGVRQII.EWSSNTYAS-RVLLIQAIDRL-NFTGWW--LI--G-VYIA-G--N-----------LA 871
CPZ.US.85.CPZUS           -L-AK----------L--LF--..T-E.------T--GA-KT-NV--T---S-F---V-E---N-LI-L----E------LR.......-T-QI--QNINKGLQLLN-LRARC-.GVIA--AR---V--T---DT-----------I--LTRRLFLG-I-------------S 834
CPZ.CM.x.SIVcpzDP943      -LIAK----------L--LI--..T--.----G-T------PG-GT-T---S-F---A-----N-L--L----I---S-LR.......-TL-I--QNINKGLQWLRDISRHL-.GVIA--GR--QI--T---DT-----------I-----RIG-G-L-------------S 850
CPZ.CM.x.SIVcpzMB897      ----------------PF-V..R-D.---------G---N-NV--V---S-F---A-----N-------------W--R.......-TL-IV-QKFXAGLRLV--T-T-LK.SI----G----T--I---D------------X--IA-RVG-G-L-------------A 842
N.CM.02.DJO0131           T-IA----------L--LI--TTT--.-----ET--GV-GQG----V---S-F-T-V-E-F-N-LI-L----T-S---LQ.......-TL-----SLIRGLQLLN-LRTRL-.GIIA--GK---D--I----TI--V-------L--LA-RIG-G-L-------------A 840
N.CM.04.04CM_1015_04      T-IA----------L--LI--..T--.------T--GV--Q--G--V-----F---V-E-C-N-VTLL----------TQ.......-TL-I--QSLNRGLQLLN-LR-RL-.GI-A--GK---D--I----T------GX---L--LA-RIG-G-L-------------T 831
N.CM.04.04CM_1131_03      T-IA----------L--LI--..T--.-----XT--GV-XQ--G--V---X-F---V-E-X-N-VILLCX--X-----XQ.......-TL----QSLSRGLQLLN-LXXRX-.GX-A--GK---D--I---XTX-----------X--LA-RIG-G-L-V-----------T 853
N.CM.95.YBF30             T-IA----------L--LI--..A--.-----ET-GGV--Q--G--V---S-FS--V-E---N-LI-L----T-S---LR.......-TL----QSLSRGLQLLN-LRTHL-.GI-A--GK--RD--I----T---V-------I--LA-RIG-G-L-------------A 850
N.CM.97.YBF106            T-IA----------L--LI--A.A--.-----ET--GV-GQ--G--V---S-F-----E---N-LI-L----A-S---IR.......-TL-I--QSLSRGLQLLN-LRIRL-.GIIA--GK---D--I----T---V-------F--LA-RIG-G-L-------------A 837
O.BE.87.ANT70             N--KNI----Q---L-IPNHH..QEE.AGT-GRTGGG---EG-P-W-PSPQ-F-P-LYT---TII-WT--L-SN-ASGIQ.......KVISY-RLGLWILGQKIINVCRICA.AVT---L---Q---T---DTL-V---NW--GI-AGI-RIGTG--N-----------S 857
O.CM.91.MVP5180           NL-KNI----Q---L-IPV-H..RQE.AET-GRTG-----G--PKWTA-PP-F-QQLYT---TII-WT--L-SN-ISGIR.......-LIDY--LGLWILGQKTI--CRLCG.AVM---L-------TN--DTI-VS--NW--GI-LGL-RIGQGFL----------A--- 870
O.CM.96.96CMABB637        NL-KNI----Q---L--LNQH..QPE.AET-G-TGG---GE--H-WRP-QQ-F-Q-LYT---TII-W--RL-SS-ASGIR.......NVTSY--LGLEILGQ-TI--CRQTG.AVX---L---QK--T--VDTI----GNW--SI-LGI-RIG-G-LNX--------F--C 874
O.CM.98.98CMA104          NL-RNI----Q-----IPIHH..QSE.VGV-G-TG-G---G--P-L-P-QQ-F-Q-LYT---TII-WT--L-SN-ASGTQ.......-VISH--IGLWILGQKIIDGC-LCK.AVI---L---Q---T---DTI-V--GKW--DI-LGI-RIG-G-LN--------X---S 878
O.CM.99.99CMU4122         NL-RNI----Q-----IPNHH..QSE.AGA-G-TG-----G--P-L-P-PP-F-H-LYT---TII-W---L-SN-ASGIQ.......TVISH--LGLWILGQKIIN-CRICA.AVV---L---Q---TN-XDTL-V---NR--SIFLGI-RIG-G-LN-----------QA 869
O.FR.92.VAU               NL-KNI----Q---L-IPIQQ..QAE.VGT-G-TG-G--DE--R-WTP-PQ-F-H-LYT---TII-WI--L-SN-ASEIQ.......KLIRH--LGLWIIGQ-TI--CRLFK.AII---L---QT--TN--DTV-V---NW--SI-LGI-SIG-G-LN------------L 871
O.SN.99.SEMP1299          NL-KNI----Q----K--IHH..QPE.AEA-G-TG-G----GMPTL-PWPQ-F-P-LYT---TII-W---L-SN-ASGIQ.......TVISH--LGLWTLGQKIIS-CRLCI.AVI---L---Q---T---DTI-V---NW-VTI-LGI-RIG-G-LN------------S 876
O.US.99.99USTWLA          NL-KNI----Q---L-IPNHH..QQE.VGT--RTG-G---EG-R-W-ASPP-F-Q-LYT---TII-W--QI-SS-TSGIQ.......KLIS---LGLCTLWQ-IT--CRLCA.AVT---L----I--T---DTV-V--GSW--NI-SG--RIG-G-LN---------A--- 869
```

	gp41 end / Env end	
B.FR.83.HXB2	.LL*	856
A1.GE.99.99GEMZ011	.-Q-	860
A1.KE.00.KER2008	.---	855
A1.KE.00.KNH1144	.---	869
A1.KE.00.KSM4024	.---	833
A1.KE.00.MSA4069	.---	860
A1.KE.00.NKU3005	.---	866
A1.RU.00.RU00051	.-Q-	854
A1.RW.93.93RW_024	.---	864
A1.SE.95.SE889I	.-Q-	852
A1.SE.95.UGSE8131	.---	861
A1.TZ.01.A173	.---	866
A1.UA.01.01UADN139	.-Q-	862
A1.UG.92.92UG037	.---	858
A1.UG.99.99UGA07072	.---	863
A2.CD.97.97CDKTB48	.---	851
A2.CY.94.94CY017_41	.---	866
A.SN.01.DDI579	.---	851
A.SN.01.DDJ369	.---	867
A.SN.96.DDJ360	.---	845
A.CD.02.02CD_KTB035	.---	855
A.CD.97.97CD_KCC2	.---	863
A.CD.97.97CD_KTB13	.---	885
B.AR.04.04AR151516	.---	853
B.AU.87.MBC925	.---	860
B.BO.99.BOL0122	.---	851
B.BR.03.BREPM2012	.-Q-	861
B.CA.97.CAN3FULL	.---	851
B.CN.05.05CNHB_hp3	.---	860
B.CO.01.PCM001	.---	851
B.GB.83.CAM1	.---	857
B.GB.86.GB8	.-Q-	850
B.GE.03.03GEMZ010	.---	848
B.IT.05.SG1	.M--	815
B.JP.05.DR6538	.---	866
B.KR.05.05CSR3	.---	870
B.NL.00.671_00T36	.---	856
B.RU.04.04RU128005	.---	851
B.TH.00.00TH_C3198	.-V-	853
B.UA.01.01UAKV167	.---	866
B.US.04.ES10_53	.---	847
B.US.99.PRB959_03	.---	850
C.AR.01.ARG4006	.-Q-	857
C.BR.04.04BR013	.-Q-	866
C.BW.00.00BW07621	.-Q-	851
C.CN.98.YNRL9840	.-Q-	869
C.ET.02.02ET_288	.---	858
C.GE.03.03GEMZ033	.-Q-	866
C.IL.99.99ET7	.-Q-	854
C.IN.99.01IN565_10	.---	861
C.KE.00.KER20I0	.---	863
C.MM.99.mIDU101_3	.-Q-	866
C.MW.93.93MW_965	.---	853
C.SN.90.90SE_364	.-Q-	842
C.SO.89.89SM_145	.---	862
C.TZ.02.CO178	#-Q-	863
C.UY.01.TRA3011	.-Q-	859
C.YE.02.02YE511	.-Q-	863
C.ZA.04.04ZASK164B1	.---	865
C.ZA.05.05ZAPSK240B1	.---	848
C.ZM.02.02ZM108	.---	880
D.CD.83.ELI	.---	854
D.CM.01.01CM_0009BBY	.---	866
D.KE.01.NKU3006	.---	870
D.KR.04.04KBH8	.-I-	843
D.TD.99.MN011	.---	846
D.TZ.01.A280	.---	853
D.UG.99.99UGK09259	.---	854
D.YE.01.01YE386	.---	843
D.YE.02.02YE516	.---	847
D.ZA.90.R1	.---	859
F1.AR.02.ARE933	.F--	846
F1.BE.93.VI850	.---	832
F1.BR.01.01BR125	.---	850
F1.ES.x.P1146	.---	862
F1.FI.93.FIN9363	.-I-	843
F2.CM.02.02CM_0016BBY	.---	846
F2.CM.95.MP257	.---	851
F2.CM.97.CM53657	.---	850
G.BE.96.DRCBL	.---	855
G.CM.01.01CM_4049HAN	.---	857
G.CU.99.Cu74	.-Q-	866
G.ES.99.X138	.---	859
G.GH.03.03GH175G	.---	863
G.KE.93.HH8793_12_1	.---	851
G.NG.01.01NGPL0669	.---	857
G.PT.x.PT2695	.---	884
G.SE.93.SE6165	.---	867
H.BE.93.VI991	.---	863
H.BE.93.VI997	.---	856
H.CF.90.056	.---	849
J.CD.97.J_97DC_KTB147	....	850
J.SE.93.SE7887	.---	852
K.CD.97.EQTB11C	.---	852
K.CM.96.MP535	.---	843

	gp41 end Env end	
B.FR.83.HXB2	.LL*	856
01_AE.CF.90.90CF11697	.---	864
01_AE.CN.05.FJ051	.-Q-	856
01_AE.CN.06.FJ054	.---	857
01_AE.HK.x.HK001	.-V-	860
01_AE.TH.01.01TH_R2184	.---	864
01_AE.TH.02.OUR769I	.---	858
01_AE.TH.90.CM240	.---	858
01_AE.US.00.00US_MSC1164	.-V-	860
02_AG.CM.02.02CM_4082STN	.---	854
02_AG.EC.x.ECU41	.---	850
02_AG.GH.03.GHNJ196	.-I-	860
02_AG.NG.01.PL0710	.---	849
02_AG.NG.x.IBNG	.---	855
02_AG.SE.94.SE7812	.---	860
02_AG.SN.98.MP1211	.-Q-	862
03_AB.RU.97.KAL153_2	.-Q-	840
04_cpx.CY.94.CY032	.---	856
05_DF.BE.x.VI1310	.---	862
06_cpx.AU.96.BFP90	.---	860
08_BC.CN.97.97CNGX_6F	.-Q-	865
09_cpx.GH.96.96GH29I1	.---	843
10_CD.TZ.96.96TZ_BF061	.---	841
11_cpx.GR.x.GR17	.---	845
12_BF.AR.99.ARMA159	.---	843
13_cpx.CM.96.1849	.-Q-	868
14_BG.DE.01.9196_01	.X--	857
15_01B.TH.99.99TH_MU2079	.---	855
16_A2D.KR.97.97KR004	.---	863
17_BF.PE.02.PE02_PCR0155	.-Q-	851
18_cpx.CU.99.CU76	.-Q-	865
19_cpx.CU.99.CU7	.---	802
20_BG.CU.99.Cu103	.---	850
21_A2D.KE.91.KNH1254	.---	863
22_01A1.CM.01.01CM_0001BBY	.---	849
23_BG.CU.03.CB118	.---	856
24_BG.CU.03.CB378	.---	865
24_BG.CU.03.CB471	.---	863
25_cpx.CM.01.101BA	.---	875
27_cpx.CD.97.97CDKTB49	.---	853
28_BF.BR.99.BREPM12609	.---	869
29_BF.BR.02.BREPM119	.-Q-	850
31_BC.BR.02.110PA	.-Q-	866
32_06A1.EE.01.EE0369	.---	851
33_01B.MY.05.05MYKL007_1	.---	858
34_01B.TH.99.OUR2478P	.-I-	844
35_AD.AF.05.05AF095	.---	852
36_cpx.CM.00.00CMNYU830	.---	858
37_cpx.CM.00.00CMNYU926	#-Q-	857
38_BF1.UY.05.UY05_4752	.---	852
39_BF.BR.03.03BRRJ103	.---	851
40_BF.BR.04.04BRRJ115	.---	863
42_BF.LU.03.luBF_05_03	.---	832
43_02G.SA.03.J11223	.---	847
U.CD.83.83CD003	.---	848
U.CD.90.90CD121E12	.---	846
U.GR.99.GR303	.---	858
U.NL.95.U_NL_95_H10986_D1	.---	862
CPZ.CD.90.ANT	.-N-	859
CPZ.CM.01.SIVcpzCAM13	.---	865
CPZ.CM.05.SIVcpzEK505	.---	836
CPZ.CM.05.SIVcpzLB7	.---	846
CPZ.CM.05.SIVcpzMB66	.---	881
CPZ.CM.05.SIVcpzMT145	.---	856
CPZ.GA.88.GAB1	.---	855
CPZ.TZ.00.TAN1	.-N-	874
CPZ.US.85.CPZUS	.--X	837
CPZ.CM.x.SIVcpzDP943	.---	853
CPZ.CM.x.SIVcpzMB897	.---	845
N.CM.02.DJO0131	.---	843
N.CM.04.04CM_1015_04	.---	834
N.CM.04.04CM_1131_03	.-V-	856
N.CM.95.YBF30	.-I-	853
N.CM.97.YBF106	.---	840
O.BE.87.ANT70	.---	860
O.CM.91.MVP5180	.-V-	873
O.CM.96.96CMABB637	.---	877
O.CM.98.98CMA104	.---	881
O.CM.99.99CMU4122	.-V-	872
O.FR.92.VAU	.---	874
O.SN.99.SEMP1299	.---	879
O.US.99.99USTWLA	.-X-	872

HIV-1/SIVcpz Proteins

```
                      myristoylation
                       Nef start                                                                     acidic cluster   phos
                                                                                                                  poly-P helix  phos            HXB2 premature Nef end
B.FR.83.HXB2          MGGKWSKS.SVIGWPTVRERMRR...................AEPAADRVGAASRDLEKHGAITSSNTAATNAACAWLEAQEEEE..VGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLWIYHTQGYF.PD*QNYTPGP.GVRYPLTFGWCYKL  144
A1.GE.99.99GEMZ011    --------.-IV---Q----L--...................APA---PG--PV-Q--D-Y--V---------D-----------.-----K----------G-F-----------D---Y-KK--E-----V-------.--W-------.-I-F----------  146
A1.KE.00.KER2008      --------.-IV---QI---I--...................TP---PG---V-Q--D---------..VNHPS-T-V-------.-----R------------L-----------------Y-RK--E-----V-----F-.--WH------.-I-F-------F--  143
A1.KE.00.KNH1144      --------.-IV---EI------...................TPA-P-VG---V-Q--A---------..VNHPS------Y-D--.-----R----------G-I-----------D--VY-RK--E-----V-------.--W-------.-T--------F--  144
A1.KE.00.KSM4024      --S----R.-TME--E------C...................AS--PG---V-Q--D----V-----..VNHPS-V------K--.-----R------------F---F-------D---Y-KK--------V-N---F-.--W-------.-I--------F--  142
A1.KE.00.MSA4069      --------.-IV---E-------...................VP---SG---V-Q--D-Y--V-----..VNHPS-V----R---.-----R----------G-I------------------F-RK--E----V-N-----.--W-------.-----------FE-  143
A1.KE.00.NKU3005      --S-----.-IV---EI---L--...................TPA--NG---V-Q---R---------..TNHPS-V-----D--.-----R----------G-L-X---------D---Y-RK--E-----V-------.--W-------.-T-F-------F--  143
A1.RU.00.RU00051      --------.-IV---E----I--...................APA---RG--PV-Q--D-X-XV--N-------------------.-----K----------G-X-----------D---Y-RK-XE---X-VH----X-.--W-------.-IX----------  146
A1.RW.93.93RW_024     --------..KM---EI------...................AP-PS--TG---V-Q--AR---V-----..TNHPS-V---------E.-----R-----------------------------Y-R--K------V-----F-.--W-E------.-T--------F--  145
A1.SE.95.SE8891       --------.-IV---K-----A-...................TP---KG---V-Q--D----V-----..INHPS---V-------.-----R----------G-----------D---Y-R---E-----V-N-----.--W-------.-T-F-------F--  143
A1.SE.95.UGSE8131     --N----G.....--E----I-QAR............APAHTP-PT--TG---V-Q--------------..INHPS-T-----AQ-DEE-----R----------G-L-------------------Y-RK--E-----V-------.--WH------.-I--------F--  149
A1.TZ.01.A173         --------.-IV---EI---I--...................TP--VRG---V-Q--D------N--..INHPS-V------D-D..-----R-------------F---Y--N..........--R---------V-------.--W-C-----.-I--------F--  134
A1.UA.01.01UADN139    --------.-IV---QI---I--A..............PAPAAPA---TG--PV-Q--D----V-------N--D-----------.-----K----------G-F-----R-----D---Y-RK--E-----V-------.--W-------.-I------------  151
A1.UG.92.92UG037      --N-----.CIV---E----I-QTPT.........AARERTRQ-PT--KG---V-Q--D----V-----..VNHPS-V-----------.-----R-------------F--GF-------D---Y-KK--E-----V-------.--W-------.-I--------F--  154
A1.UG.99.99UGA07072   --------.-IV---EI---I--...................TPA--KG---V-QG-DQY--------..INHPS--------D--.-----R--I-------G-F-----------D---Y-RK--E-----V-----F-.--W-------.-T--------F--  143
A2.CD.97.97CDKTB48    -------R.TIV---EI------TPP..........AAEGVRPTP---EG---V-Q--AR---V-------N-PD-----------.-----R------A---G-----------D---Y-------------V-N-----.--W-------.-A-F-------F--  155
A2.CY.94.94CY017_41   -------R.-IP---AI------TPP..........TAQRTEAVS---PG---V-Q--AT---V---------PD---V------SE.-----R---------F-G-F---F-------D---Y--K------M-V-------.--W-------.-I--------F--  156
A.SN.01.DDI579        ---------.RQE--HK--D---Q..................TPPT---EG---V-Q--S----V--T-..QAS-S-T----------.-----R----------G-F-----------D---Y--Q-------V-------.--W-------.-T--------F--  145
A.SN.01.DDJ369        ---------..-.--AEI-----K..................TPPAPT---KG---V-Q-PD--R-V--E-..-NS-S-TR--E------.-----R----------G-F-----------D---------E-----V-N-----.--W-------.-I--------F--  145
A.SN.96.DDJ360        ---------.-LA---K--D---Q..................TPPT---KG---V-Q--D-Y--V--T-..QAS-SSI----H-D--.-----R---------F---F-----------D-------Q--------V-------.--W-------.-T-F-------F--  145
A.CD.02.02CD_KTB035   ---------.RLG---A-W---Q-...................TP---EG---V----DR------------SD--------G-..-----K--G----A---PF-----V-----DR--W--K--------V-------.--W-------.E----------F--  145
A.CD.97.97CD_KCC2     ---------.-LV---A--D----...................TPA--EG---V-Q--D-Y--V--N------PD-----------.-----R---T-------------------------D---W--K--------V-------.--W-------.-I-F-------F--  145
A.CD.97.97CD_KTB13    -------R.-LG--SAI------...................TP---EG---V-Q--D-Y--V--N------SD---V-------E.-----R----------G-F-----------D---W--K--------V-N-----.--W-------.-------------  146

B.AR.04.04AR151516    -------Y.IRG---A----IL-...................T----G--------R------------D--------G-E.-----R------------------------------VW--K--------VX---F-.--WH------.-T----C----F--  146
B.AU.87.MBC925        ---------.-IF----------...................------G----------E-------TN--D-------D--.-----K-------------------------------Y--K----------------.--W-------.-I---I-----F--  145
B.BO.99.BOL0122       ---------.K..--AA------K..................------G------------------E---DV-----------.-----R------------I----R------VY----V---N--V-----F-.--WH------.-T-W-------F--  143
B.BR.03.BREPM2012     --S----H.IGE---A--D----ARP..........AAERRRQ------G--------AER------------D------H-D--.-----R-------------F---Y-----------------Y-----------V-------.--W-------.-T-F-------F--  155
B.CA.97.CANB3FULL     ----L--R..MRE-A------L-...................------G---V-----R------------D--------D--.-----K------------------------------VY--K-K-------V-------.--W-------.-T-F-------F--  144
B.CN.05.05CNHB_hp3    -------G.CMA---S-----E-...................------G---V---GR---------..TA-D--V-----------.-----R----------G-L-----R----------Y--K--E-----V-----F-.--WH------.-I--------F--  143
B.CO.01.PCM001        -------R.KAAE-AD-----E-...................-----EG---------------------D--X---H-D--.-----R---------R---I-M-------------Y--K-K------V---E---.--W-C-----.------------F--  145
B.GB.83.CAM1          -------R.-LG--SA-----Q-...............AEPR-----EG---V-------------------D------------.-----R-------------L-I-----------I-Y--------------------.--W-------.-I--------F--  149
B.GB.86.GB8           ----L--R.-MF--SR--D--QQ...................-----EG---V-------------------D------------.-----R-----------------------Q---D---Y-PK--E-----V-------.--W-------.-T-F-------F--  145
B.GE.03.03GEMZ010     -------R.LTG---A-----EQ..................A---EG---------Y--L-----------S-----------E.-----R----------G-L-------------IVW--Q--------V-N---F-.--W----G--.-I--------F--  145
B.JP.05.DR6538        ---------.KLG---I------...................T----G---V-----Y----------N--D------------.-----R---------F-E-F-------------------Y--K--E-----V-----I-.--WH------.-I-F-------F--  145
B.KR.05.05CSR3        -------R.....--A-----E-...................------G--------RY--L-------N-PDL-----------.-----R----------------------------------K--------VH----FL.--W-G------.-----------F--  142
B.NL.00.671_00T36     -------H.KGG---A-----Q-...................------G--PV---------------..TN----V-----------.-----R----------G-L-------------VY--K--------V-------.--WD------.-I----I----F--  143
B.RU.04.04RU128005    -X-----R..XVE--A-----EK...................V-----G-------------------P----D---A-----E.-----R----------------------------------K--------V-N-----.--W-------.-T--------F--  145
B.UA.01.01UAKV167     -------R.-TG---A-----E-...................-----EG---V----R---------..VNS---T-----------.-----R-------------L-----------I-----------------V-N-----.--W-------.-I--------F--  143
B.US.04.ES10_53       -------...MSE-SI-------A.............EPRMRR----.AG-------G------------------D--V-----G-E.-----R----------G-----------------Y--Q--------V-N-----.--W-------.-T--------F--  150
B.US.99.PRB959_03     -------C.-MG--TA------K..................TEP-----AG---V-----------N-..VN---T--A--Q-D-E.-----R----------G-L--------------Y--Q--------V-N-----.--W-------.-I--------F--  147

C.AR.01.ARG4006       --N----...L.---A-------APP.........GVGDXRRQDA---EG--P--Q--DRY--L-----P-N--D-------Q---E.-----R---------F-G-F---F-----------------Y--Q-K------V-N-----.--W-------.-----------F--  154
C.BR.04.04BR013       --N----C.-PV--SAI---I--...................-----EG--P----DR---L-----P-N--D--------D-D..-----R---------F---F---F---------------Y-KK--------V-----F-.--W-------.-----------F--  145
C.BW.00.00BW07621     ---------.-IV---A----I--..................TD---EG----Q--DR----------PT-------------D-..-----R----------G-L--GF----------------Y-TK-KE-----V-------.--W-------.-T--S-----PF--  145
C.ET.02.02ET_288      --SA---C.-PR--S-I-----Q..................TP---EG---------Q---L-T---PEN-------Q------E.-----R-------------F--GF-------------------Y--K-KE-----V-N---F-.--W-------.-----------F--  146
C.GE.03.03GEMZ033     -------N.-IV---A----I--..................T----EG--------D---L-T----EN--D----K------E.-----R----V------G-F---F-------D---Y-KQ--------V-N---F-.--W-------.-------------  146
C.IL.99.99ET7         ---------.-LV---A--D-IS-..................TD---EG----Q--D---L-----V---KD--------Q--E.-----R----------G-----W------------Y-KK--E-----V-----F-.--W-------.-----------F--  146
C.IN.99.01IN565_10    -------R.-..---AI------..............AEPA-----E-----Q--D--...-----DTA------PRE-D---EG----R----------G-F---F-------D---Y-KK--E-----VH----FL.--W-------.-----------F--  146
C.KE.00.KER2010       ---------.-IV---A------...................-----EG----Q--D-Y--L-----GN--D---Q------E.-----R---------R-G-I---F-------D---Y-KK--------V-N---F-.--W-------.-----------F--  146
C.MM.99.mIDU101_3     -------R.-IV---AI------..................T----EG----Q--D-Y--L-----DTS--D---R------E.----IR---------F-G-F---F-----------------Y-KK--E-----V-------.--W-------.-----------F--  146
C.MW.93.93MW_965      ---------.-I----D----I-Q..................TS---EG----Q--D---L-T---VSN------------DGD.-----R-------------F---F-------D---Y-KK--------V-N---F-.--W-------.-------------  146
C.SN.90.90SE_364      -------C.-IV---E----I--..................TN---TG----Q--D-Y--L--------E---Q-----Q--E.-----R----------G-F-----------------Y-KK--E-----V-N-----.--W-C-----.-----------F--  146
C.SO.89.89SM_145      ---------.RIV--ANI---I--..................T----EG----K--D-Y--L-N---PTN--D---Q------E.-----K-------------F---F-----------------Y-KK--E-----V-------.--W-------.-----------F--  146
C.TZ.02.CO178         -------R.-IV---EI---I--T..............RPAAEGV---EG----Q--D----------G----DS--------S-E.-----R----V------S----F-----------------Y-KK--------V-N---F-.--W-C-----.-----------F--  152
C.UY.01.TRA3011       --N----C.-PV---DI---I--..................TD---EG--P--Q--D-Y--L-----P-N--D-------Q-A-E.-----R---------F-G-F---F-----------------Y--K--------V-N-----.--W-------.---F-------F--  146
C.YE.02.02YE511       ---------.-IV---V--D-I--TNPA........AAGVGAASQDS--AG-----Q--A----L-----PTN------------QD-E.-----R----------G-----W------------Y-KK--E-----M-N-----.--W-------.-T--------F--  158
C.ZA.04.04ZASK164B1   ---------.-IV---AI---I--T..............RPAAEGAG--EG--T--Q--D----L-T----TN-------------D-D.-----R----------G-F--GF-----------------Y-KK--E-----V-------.--W-------.----------PF--  152
C.ZA.05.05ZAPSK240B1  ---------.K..R--EI-----Q..................TRP-A--G-----Q--D----F-T----SN--DR---Q------D.-----R----V------G-F--GF-------D---W-KK--------V-N-----.--W-------.----------PF--  146
C.ZM.02.02ZM108       --N-------.....--A----I--T..............RTGAAT----EG-----Q--D-----------T---D-----------.-----R-------------F---F-----------------W-KKG-E-----V-----F-.--W----S--.-----------F--  147

D.CD.83.ELI           ---------.-IV---AI---I--..................TN----G---V---------------S---D---------SDE.-----R----------E-L-----------------W-KK--E-----V-N---I-.--W-------.-I---------E-  146
D.KE.01.NKU3006       ---------.-I----AI---I--..................ID---EG---V-----R---V------Q---D---------DGE.-----R----------G-L-----------------W-KK--E-----V-N-----.--W-------.-I--------FE-  146
D.KR.04.04KBH8        -------.....--AI------..................T-----G---------R------R-------D-----------E.-----R---T--------L--------D--------VW--K--------V-N---F-.--WH------.-I--------FE-  141
D.TD.99.MN011         ---------.-IV---AI---I--..................TD---EG---V-----R-----------T-----------G--..-----R---------F-----------------------VW--K--------V-------.--W-------.-T--------FE-  145
D.TZ.01.A280          ---------.-IV---AI---I-K..................-D----G---V-----R-----------H---D---------ADE.-----R--A------------L--------D--------VW------------V-----F-.--WH------.-T--------FE-  146
D.UG.99.99UGK09259    ---------.KMG--SA-------...................-----EG---V------Y-----T---Q--PD-----------E.-----R----------G------------------------VW--K--------V-------.--WD------.-----------F--  146
D.YE.01.01YE386       ---------.-IV---AI---I--..................T----EG--------------------Q--SD---------DNA.-----R----------G-L-----------------W-PK--E-------------.--W-------.-I--------FE-  146
D.YE.02.02YE516       ---------.-IV----I--KI-K..................TD----G---V---------------NNPA..--------D--..-----R------------------------------V------------V-------.--W-------.-I-H-------FE-  143
D.ZA.90.R1            ---------.-IV---AI---I--..................TD----G---V---------------S---D---------S-E.-----K---------F-E-----------------W-KK--E-----V-N---I-.--W-------.-I--------F--  146

F1.BE.93.VI850        ---------.-IV---A-G----Q..................TPT--EG---V----DRR--------RT--PDL-----------..-----R----V------------------------Y-KK-G-T----V-------.--W-------.-I-----L---F--  145
F1.BR.01.01BR125      --N----R.-IV---AI---I-Q..................TPT--EG---V-Q----R--------RDN-PDL-----------..-----R----------G-----------------Y-KK--E-----V-------.--W-------.-T-F-------F--  145
F1.ES.x.P1146         ------R-..LA---A----I--TPPT........AERVERRQNP---EG--------ASR--------G---PDV-------D-D..-----R----V----------------R------------K--------V-----F-.--W-------.-T-W-------F--  156
F1.FI.93.FIN9363      ---------.-IV---AI------..................PPPA--EG---V-Q---RR--------G---PDL-------D--..-----R----------G-F---Q-XX--------XY-KK--E-----L------#--W-------f-----------F--  146
F2.CM.95.MP257        ---------.-IV---AI---I-..................--EG---V-Q--D-R---N---G---DL-----------..-----R-------------L-----------------Y-RK--E-----V-------.--W-------.-P-F-------F--  140
F2.CM.97.CM53657      ---------.-KV----I---I-Q..................TPA--PG---V-Q--D--------Q-RD--SDL-------D--..-----R----------G-F---F-----------------Y-K---E-----V-------.--W-------.-T--------F--  145

G.BE.96.DRCBL         --N----R.K-A---E----L-Q..................HPA--EG---V-Q--DR------R---G--PD---------DSE.-----R-------------F---F-------D--VY-KK--E-----V-----F-.--W-------.-T-V-------F--  146
G.CU.99.Cu74          ---------.-IV---A----I-Q..................TP---AG---V-Q--AR-----------N-PD----Q----DSE.-----R-------------F---F-------D---Y-KK--------V-N---F-.--W-------.-T--------F--  146
G.ES.99.X138          -------C.-IV---XI---IK-..................XP---VGX--V-Q--AR-----------N-PD---------DSE.-----R---X------G-F---F-------D---Y-KQ--------V-N---F-.--W-S-----.-T-F-------F--  146
G.GH.03.03GH175G      I------R.-IKE--K-K-QI-Q..................-PA--EG-----Q--D----L-------N-PD---------DSE.-----K-------------F---F-------D---W-KK--E-----V-N---F-.--W-------.-T-L-------F--  146
G.KE.93.HH8793_12_1   -------R.......---A----I-Q..............TPIRQTP---EG---V-Q--AR----------TN-PD---------DSE.-----R----V--------F---F-------D---Y-KK--------V-N---F-.--W-------.-T-L-------L--  148
G.NG.01.01NGPL0669    --S-C--...L.---A----I-QT..............PPTEER-GA--EG-----Q--AR-------------PD--V--Q-DSE.-----R----------G-F---F-------D-----------------V-N-----.--W-------.-T-F-------F--  150
G.PT.x.PT2695         -------C.-IV---E----I--..................-PS--VG---V-Q--AR-----------N-PD-------N-DSE.-----R----------S-L---F-----------------Y-KQ--------V-N---F-.--W-------.-T-F-------F--  146
G.SE.93.SE6165        ---------.-IV---E----I-N..................TPT--EG---V-Q--DR-----------N-PD---------DSE.-----R---------F-G-F---F-------D---Y-KK--E-----V-N-----.--W-------.-T-F-------F--  146

H.BE.93.VI991         -------G.CIS---A----I-Q..................T----EG---V-Q--DRR--V-IN-I-SN--DS----------E.-----R----------G-F-----------------Y-KK--E-----V-N-----.--WH------.-E---------F--  146
H.BE.93.VI997         ---------.-IV---A----I--..................-Q----G---V----DRR--V-IN-----PDV--------A-E.-----R-------------L-----------D---Y-KK--E-----V-N-----.--W-------.-EG--------F--  146
H.CF.90.056           ---------.RMG--S-I-------..................---V-EG---V----DRR--V-IN---S--RDA-------DG-E.-----R----------G-F-----------D---Y-KQ--------V-N-----.--W-------.-E-F-------F--  146

J.CD.97.J_97DC_KTB147 ..........TA--SEI------....................ATPDA---EG-----Q--DR---------VE--S--------T---E.-----R----------G-F--GF-------D---Y-KK--E-----V-N-----.--W-------.-I------------  140
J.SE.93.SE7887        --N-------.....-----Q-------..................APA----G---V-Q--A-----------------D-------T---..-----K--I-------G-----F-------D---Y-KK--E-----VHN-----.--W-------.-I------------  142

K.CD.97.EQTB11C       ---------.-IV--S-------K..................TP----G---V-Q--D----V------FN-PD--------D-D..-----R---------F-G-F--GF-------D---Y-K---E-----V-----F-.--W-------.-I------------  145
K.CM.96.MP535         ---------.-IV---AI------AR...........PAADRVGTQ----G---V-Q--AR---V-----SHN-PD-----------..-----R-------------F--GF-------D---Y-KK--E-----V-----F-.--W-------.-I------------  154
```

```
                          myristoylation
                           Nef start                                                                                    acidic cluster  poly-P helix  phos (phos)                    HXB2 premature Nef end
B.FR.83.HXB2               MGGKWSKS.SVIGWPTVRERMRR...................AEPAADRVGAASRDLEKHGAITSSNTAATNAACAWLEAQEEEE..VGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLWIYHTQGYF.PD*QNYTPGP.GVRYPLTFGWCYKL 144
01_AE.CF.90.90CF11697      -------N.RIV---Q----I--..................TPA--EG-----Q--D---------..MN--D----------..-----R----------G-F---F--------D---Y-KK--E-----V-------.--W-------.-I----C------- 143
01_AE.CN.05.FJ051          ----C--G.-KV--SQI---I-Q.................TP--TEG---V-Q-PD----V-----..MN--DSV--R------EE-----R---------F-G-F---F--R-----D---Y-KK--E-----V-N---F-.--W-------.-I----C----F-- 145
01_AE.CN.06.FJ054          --N-----.....--R----I-Q.................TPT-TEG---V-Q--D----V-----..MS--D-V----------..-----R--------AF-G-F---FL-R-----D---Y-KK--E-----V-N---F-.--WH------.-I----C----F-- 139
01_AE.HK.x.HK001           ---------.-IV---Q----I-Q.................TP--TEG---V-Q--D----V--N-..MN-DDSV--R----D-EG-----R-----------D-F---F--------D---Y-KK--E-----V-N---F-.--W-------.-I----C----F-- 145
02_AG.GH.03.GHNJ196        --S-L---.RIV--AR----L--TPPTAERVRRPPPAAEGAGATSQ--VG-----Q--AR---------SS---D-----------EE-----R-------------------------------Y-KK--E-----V-----FL.--W-------.-I---------F-- 167
02_AG.NG.01.PL0710         ---------.-IV---KI-D-I-Q...............TPPEP----KG-----Q--DR---L-T---SQN-PD---------GGE.-----R----------G-L-----------------Y--K--E-----V-------.--W-------.-I-----L---F-- 150
02_AG.NG.x.IBNG            ---------.-IV---K-MK---Q.................TPT--TG-----Q--DR----------Q--PD--------D-N..-----R----------G-------------D---Y-KK--E-----V-N---F-.--W-------.-T-F-------F-- 145
02_AG.SE.94.SE7812         ---------.-IV---QI-D-I-Q.................TP---RG-----Q--DR---------------------------..-----R-------------------------------Y-KK--E-----V-----F-.--W-------.-T-F-------F-- 145
02_AG.SN.98.MP1211         ---------.-LV---Q----I--................TQPTPS--IG---V-Q--DR------R---H---D-----------..-----R---------------------Q-------Y-KK--E-----V-----F-.--W-------.-I------------ 148
03_AB.RU.97.KAL153_2       ---------.-IV---Q----I--.................APA---RG--PV-Q--D-Y--V-------N--D-----------..-----R----------G-F-----------D---Y-KK--E-----V-------.--W-...-...I-F---------- 141
04_cpx.CY.94.CY032         ---------.-IV---EI------ARAE........PERMRRAQ-----AG---V-Q--D------IN------PDKT---------E.-----R---------F-G-L-----------D---Y-KK--E-----V-----F-.--WD------.-E-F--C----F-- 158
05_DF.BE.x.VI1310          ---------.--V---AI--K---..................TP---EG---VFQ--DNR--V-IR--V---PDM--V-------E.-----K-------------------------------Y-KK--------V-N----S.--W-------.-I-----L---F-- 146
06_cpx.AU.96.BFP90         --N----.....--SQ-------..................TPPT-R--EG---V-Q--D-----------T------V---T---..-----R----V--------F---F--------D---Y-KK--E-----V-N---F-.--W-------.-T-F---------- 144
08_BC.CN.97.97CNGX_6F      ---------.-IV---AI---I--..................T-----G---V----------------D---D-----T-----..-----R---------F-G-F---F-------------Y-KK--E-----V-------.--WH------.---F-------F-- 145
09_cpx.GH.96.96GH2911      ---------.XIV---EI---I--................TSXS--.EG-----Q--GX---L------S--HD--------D-D..-----R----------G-----F--------D---Y-KK--E-----V-------.--W-------.---F-------F-- 146
10_CD.TZ.96.96TZ_BF061     ---------.-IV---Q-------..................TS---EG-----Q--A-F--L-----PGN--D--------A--E.-----R---------F-G-F--GF--------D---Y--K-K------V-N---F-.--W-------.-T---------F-- 146
11_cpx.GR.x.GR17           ---N-----.-IV---EI---L--.................TPPTA--EG-----K---R---------SQN-----------D-G..-----R----------G-F---F--------D---Y-KK--E-----VHN---F-.--W-------.-I----C------- 147
12_BF.AR.99.ARMA159        ---------.-IV---EI---L--..................-P---EG---V-Q---NR--------R-N-P-L-----------..-----R---------F-G-F-----------------Y-KK--E--------------.--W-------.---L-------F-- 145
13_cpx.CM.96.1849          -------R.-TV---A----I.....................--TG---V-Q--DR---V-----..LA-KDYV-------G-..-----K----------G-F--GF--------D-----KK--------V-------.--W-------.------C-E--F-- 138
14_BG.DE.01.9196_01        -------X.-LT---E----I--..................TP---VG---V-Q--A----------XXX-PX-------T-DSE.-----R----------X-F---F--------D---X--X--------V-N---F-.--W-------.-X-F-------F-- 146
15_01B.TH.99.99TH_MU2079   --S-----.....--Q---KIKQ..................TP---EG---V-Q--DN---V-----..MN--DSV--R-E-D--..-----R---------F-E-F---F--------D---Y-KK--E-----V-N---F-.--W-------.-I----C----F-- 139
16_A2D.KR.97.97KR004       -------R.-LP---AI------TP...........PAAERTPPAA--EG---V-Q--ATR--V-IN-------DS----------E.-----R----------G-F------R-----D------K--------V-----F-.--W-G------.-I---------F-- 155
17_BF.PE.02.PE02_PCR0155   --S-----.-IV---EI------AG............PGTRRAEPA--EG---V-Q--G-R--------S-N-PDL-------D--..-----R----------G-I-----------------Y-K---E-----V-N-------.--W-------.-I-----L---F-- 153
18_cpx.CU.99.CU76          ---------.-LV---EI---I--TQEL........PRTPPAESQ-ER...........---SG-N-S-----PTHGR--EE----..-----R----------G-F---F--------D---FF-K-PE-----VFN---FL.--W-------.---F-------F-- 145
19_cpx.CU.99.CU7           #--------.-I----A-----Q-..................------G---V---------------R----D---------ADE.-----X----------G-L--------------IVW--K---------V-------.--WH------.-I---------FE- 145
20_BG.CU.99.Cu103          --N-----.....--A----L--..................TP---AG---V-Q--A------------N-PD---------DSE.-----R---------F---F---F--------D---Y-KQ---------V-N-----.--W-------.-T-F-------F-- 142
21_A2D.KE.91.KNH1254       -------R.-IV--------I-K..................TD----G----------QR-----------N-PD---------N-E.-----R---------F-G-F---F--------D---Y-PK--E-----V-------.--W-------.-I------------ 146
23_BG.CU.03.CB118          --N-----.....--AA-------.............APPAXNP---AG---V-Q--AR---------------DR--V-----DSE.-----R---------------F---F--------D---W-KQ---------V-N---F-.--W-------.-I---------F-- 147
24_BG.CU.03.CB378          --S------..-P--SVI------TPPTRT...RGGASQNLASHEHS--VG---V-Q--AR---------V-N-PD---------DSE.-----R---------F---F--GF--------D---Y-KQ---------V-N---F-.--W-------.-T---------F-- 162
24_BG.CU.03.CB471          --I-C--P.....---EI------.............TPARQEP---VG---V-Q--AR-------------N-PD---------DSE.-----R-----------D-F---F-------------Y-K----------V-N---F-.--W-------.-I-F-------F-- 148
25_cpx.CM.01.101BA         ---------.-IV---EI------...................TPTA--EG---V-Q--AR------N---S--S----------S-E.-----R----------G-F--GF--------D-----K---------V-------.--W-------.-I------------ 147
27_cpx.CD.97.97CDKTB49     -------R.-IV---A------Q..................TPV---G---V-Q--AR---V--R---TN-PD-------Q--SE.-----R---------------------------D--TY-KQ---------V-N-----.--W-------.------------- 146
28_BF.BR.99.BREPM12609     -------R.GTG--EA-------..................TRPA------G---V-----R----------HN--D-----------..-----R----------G-L------R-----------PK--E-----V-------.--W-------.-T-F-------F-- 149
29_BF.BR.02.BREPM119       -------N.-I----A----I--..................TP---EG---V-----R----------SN--D-----------..-----R----------S-L-------------VW--K---------V-------.--W-------.-I---------F-- 145
31_BC.BR.02.110PA          --N----C#VQV---AI------..................-----EG--P--Q--DRY--L-----P-----------fNRRKM.RR-XXXXXXXXXXXXXXXf----F-----------fY-KK---------V-N-----.--W-------.-T---------F-- 143
32_06A1.EE.01.EE0369       ---------.-IV---EI---L--..................TP---EG---V-Q--D--------H-PT-----------D--..-----R----------G-F--GF--------D---Y-KK--E-----V-------.--W-------.-I------------ 145
33_01B.MY.05.05MYKL007_1   ---------.-IV---K----I-Q..................TP---GG---V-Q--D----V-----..MN--DSV--R-----NEE-----R---------F---F---F--------D---Y-KK--E-----V-N---F-.--W-------.-T-F--C----F-- 145
34_01B.TH.99.OUR2478P      ---------.-IV---Q----I--..................TP---EG---V-Q--D----V-----..MN--D-V--R----N-EE-----R---------F---F---F-------------Y-K---E-----V-N---F-.--W-------.-T-F--C----F-- 145
35_AD.AF.05.05AF095        ---------.-KV---E---K--Q..................-PA--QG---V-Q-------------..MNHPS-V----------..-----R---S-------G-F---F--------D---W-RK--E-----V-N-----.--W-------.---F---------- 143
36_cpx.CM.00.00CMNYU830    ---------.-I----QI-D---Q..................TP---IG---V----DR---------TS---DY--------G-..-----R----------G-F---F--------D------K---------V-----F-.--W-------............... 131
37_cpx.CM.00.00CMNYU926    ---------.-LV---Q----IKQ..................TP---EG---V-Q--A-------R--ST--PD-------G-DPD.-----R-------------F---F--------D------K-----------N-----.--W-------............... 132
38_BF1.UY.05.UY05_4752     --S----P..RV--SEI------.................AEPAN----EG---V-----RR--------N-N-PDV-----------..-----R----------G-L-----------------Y-KK---------V-------.--W-------.-I-F-------F-- 148
39_BF.BR.03.03BRRJ103      -------R.NKVE--A---KIEQ#...............PNRTNP......f-V-Q---RR--------R-N-PNL--------DDE.-----R----------G-L-----------Q--VW--K---------V-------.--W-------.-T-F-------F-- 143
40_BF.BR.04.04BRRJ115      -------G.R-P---EI------.................AEPA-----RG----------Y------------PD--------D--..-----K----------G-------------------Y-KK--E-----V-N---F-.--W-------.-----------F-- 149
42_BF.LU.03.luBF_05_03     ---X-----.-LV---A------K..................T----VG--P----------L-T----TK---V-----------..--X--R----------G-L---------------VY-------------------.--W-------.-I---------F-- 145
43_02G.SA.03.J11223        -------R.-MG--SEI---I--TP............PTAAEG-GT--EG-----Q--AR---------PQ---D--------D--..-----R-R-----------F--GF-----------VY-KK--E-----V-----F-.--WH------.-I---------F-- 153
U.CD.83.83CD003            --N----Q......--AI------.................ARPA------G----------Y---------E--E---Q---AQ-DGE.-----R-------------F--GF--------D---Y-KK--E-----V-N-----.--W-------.-I--------PF-- 146
U.CD.90.90CD121E12         --N----Q......--AI-----Q.................ARPA------G---V----DR-------------S---Q-------GD.-----K------------------F-------------Y-KK--E-----V-N-----.--W-------.-I-----L---F-- 146
U.GR.99.GR303              ---------.-IV---EI---I--T.............PPKDRQPL---EG---V-Q--DR---V-----R-N--NL----------..EA---R--L--------E-L------------W-----Q--------F-N-----.--W-------.-I------------ 152
U.NL.95.U_NL_95_H10986_D1  ---------.-IV---A----I--T.............PAAEGRRA----G---V----DR---------------D--------T---..-----R-------------F---F-------D------K---------V-------.--W-------.-I------------ 152
CPZ.CD.90.ANT              --SA---I.KWV-ARQAIRKI.H....................-TNP-DI-PCGNE-ASR--L---TIGTEKDVITYS-DHT--G..T----R----M----E-L-----X--------DX--XXLK-AA---M-MFN---I-.--W----EE-.------CR---F-- 143
CPZ.CM.01.SIVcpzCAM13      --AQ-----.-LV---EI-K-IKE...................-PI--EG--EV-K----Y-----RH-PS--QTL---DEMKHH-EE-----R----------G-F------RG-------VY--K--E-----V-----F-.--W-------.-E---PC----F-- 147
CPZ.CM.05.SIVcpzEK505      --KV-----.-LV---Q-ID----..................QQDP-EG---V-Q--ANR----IR--KEN-Q-L-----E-K--A..-----C----------E-F---F--R--------VW-TK--E-----V-----I-.--W-------.-----------F-- 145
CPZ.CM.05.SIVcpzLB7        --S------.-LV---E--R-I-Q..................ETV--EG--P--Q--AR---------PQ--ETI-----E-K----..IW---K---------F-E-F--GY-----------VY-KK--E----GV-----I-.--W-------.-I---I---G-F-- 145
CPZ.CM.05.SIVcpzMB66       ---------.-LV---E--D-L--..................TQT--EG--PV----AR-------R--SQ--ETL-----EVQN--..-----R---------F-G-F-----------------VY-KK--E-----V-----F-.--W-------.-I---------F-- 145
CPZ.CM.05.SIVcpzMT145      --A------.-LV---EI---IK-..................-PT---G--EV-Q-IASR--V-IR--PQ--ETL-----EMQD--..-----K---------F-G-F-------------------W-RK--E-----V-------.--W-------.-----------F-- 145
CPZ.GA.88.GAB1             --T------.-LV---E--R-I-E..................-PT--EG--EV-K---R-------R--PE--QTL-----EMDN--..-----R----T--------F-----------------VY-R---E-----V-----F-.--W----T--.-T-F--C----F-- 145
CPZ.TZ.00.TAN1             --NIF..G.RWP-ARKAI-DL.H...................NTSSEP--Q--Q--QNK-GL-TNTLGTSADVLEYSADHT---..-----R-A--M----E-L-I---W--------D--FF-PK-AA---T-M-N---V-.--W-------.-I----CR--LF-- 141
CPZ.US.85.CPZUS            --N------.-IV---E--N-L-Q................TQTTA--EG--PV-Q--AE-----TR--PQN-QTL---DEMTNH-SE-----R----------Q-F--GF-----------VY-R---E-----V-----I-.--W-------.--------Y---F-- 149
CPZ.CM.x.SIVcpzDP943       --N------.-LV---E--A-I-Q..................-Q---PG--SV-----A-----TR--PQN-QTL----EMQGH-EE-----R---------F-G-F--GY-------------Y-R---E-----V-----F-.--W-------.-----------F-- 147
CPZ.CM.x.SIVcpzMB897       --N------.-LV---A--D-L--..................TP---EG--PV-----R-------R--PQ--ETL-----E-RD--..-----R-------------F-----------------Y-----E-----V-----F-.--W-------.-----V-----F-- 145
N.CM.02.DJO0131            --KI-----.-IV---EI------Q.............RPHEPAV--.-VG-----Q--ANR--L-T---RTN-PTV--V------GE.-----R-------------------F--------------W-RK---V----V-----F-.--W-------.-T-F-------F-- 152
N.CM.04.04CM_1015_04       --KS-----.--V---EI------Q.............RQTQAA-A..-VG-----Q--ANR----T---RDN-ETV--V--------E.-----R---------F---F---F------------VW-RK--E---------N-----.--W-------.-----------F-- 151
N.CM.04.04CM_1131_03       --KIL--N.R-A--SEI------Q..............TQEPAQ----VG-----Q--ANR----T---RDN-ETVT----------E.-----R---I--------F---F------------VW-RK--E-----V-----F-.--W-------.-----------F-- 152
N.CM.95.YBF30              --KI-----.-LV---EI------Q..............TQEPAV--.-VGA----Q--ANR----IR--RDN-ESI-----------E.-----R-------I---Q-F---F---D-------VW-RK---------M-----IL.--WH------.-I---V-----F-- 151
N.CM.97.YBF106             --KI-----.-LV---EI---I--...................QTP--.-VG---V-Q--ANR----T---KDN-QTV--------X-..-----R-------------F---F------------VW-RK--E-----V-----F-.--W-------.-----XC----F-- 146
O.BE.87.ANT70              --NALR-G.KFE--AA-------T..............RTFPES--C-PG--QI--E-AAR-G-P--H-PQN---L-F--SHQ---..-----A----------G-F---F-----------------Y-HK-AE-----V-N---F-.--W-------.-T-F------LF-- 151
O.CM.91.MVP5180            --NA-----.KFA--SE--D----..............SSSDPQQ-C-PG---V--E-ATR-G-S--H-PQN---L-F-DSHKD-D..-----R---------F---F---F--------D---Y-HK-AE-----------F-.--W-C-----.-P-F------LF-- 150
O.CM.96.96CMABB637         --NA---N.KFA--SA--D----TR...........TSRTSPNS--C-PG--EI--Q-ADR-G-PI-Y-PQN---L-L--SHQ-S-E.-----R----------G-F--GF--------D---Y-PE-AE-----V-----F-.--W-------.-T--------PFR- 155
O.CM.98.98CMA104           --NVLG--.KFA--S--------..................APXPEP--C-PG--XL--E-AAR-G-P--Y-PQN---L-F--SHQD-DE.-----R---V---F-G-F---F--------D---Y-DA-AE-----VHN---FX.--W-------.-T-F------MF-- 151
O.CM.99.99CMU4122          --NALX--.RFP--SAI-X---K.................TSPEP--C-XG--PV--E-AAR-G-P--Y-PQN---L-F--SHSD--..-----R---V---F---F---V----N---------Y-HG-AE--------N---F-.--W-------.-T-F------LF-- 149
O.FR.92.VAU                --NVLG--.KWA--SE--R---QTS...........PDPQSSDTQP-PG-R-V--A-ADRRG-P--Y-PQN-Q-L-F-GSHQGKG..I----R----------G-F---F--------D---Y-HK-AE-----VH----F-.--W-------.-T-F------LF-- 153
O.SN.99.SEMP1299           --NVLG-D.IFK--SA------G.................TSPDP--C-PG--QI--E-AAR-G-P--Y-PQN---L-F--SHQD--..-----R----------G-F---F--------D---Y-HK-AE-----V-----F-.--W-G-----.-P-F------LF-- 149
O.US.99.99USTWLA           --NILT--.KFE---A--R---Q.................TSPDPQ-C-PG--EV--A-AGR-G-PN-Y-PQN---L-F---YQD-D..-----R-----Q---F---H--------N---D---Y-HK-AE-----V-----F-.--W-------.-P-------LF-- 149
```

```
                                                                  normal Nef end
B.FR.83.HXB2        VPVEPDKIEEA.NKGENTS.LLHPVSLHGMDDPEREVLEWRFDSRLAFHHVARELHPEYF.KNC*  205
A1.GE.99.99GEMZ011  ---D-SEV---.TE---N-.----ICQ-----E-----M-K------LT-R--------Y.-D--  208
A1.KE.00.KER2008    ---D--EV-R-.TE---N-.----ICQ-----E-----K-Q------T-Q---I----Y.-D--  205
A1.KE.00.KNH1144    ---D--EV-K-.TE---NC.----ICQ-----E-G---R-K---H--LE-I--KK---FY.-D--  206
A1.KE.00.KSM4024    ---D--EV-K-.TE---N-.----ICQ-----E-K---Q-K---H--RK-R-L-----FY.----  204
A1.KE.00.MSA4069    ---D--EV-K-.TE---N-.----IC----E-E-K---K-K------LR-R---M---Y.-D--  205
A1.KE.00.NKU3005    ---A-XEV-K-.-E---N-.----MCQ-----ED--T-M-K------LK-R-Q-----FY.-D--  205
A1.RU.00.RU00051    ---D-XEV---.TE---NC.----ICQ-----D-G---M-K------LK-R-Q--P-GFY.-D--  208
A1.RW.93.93RW_024   ---D--EV-KE.TE---N-.----ICQ-----E-K---K-V------LK-R-Q-----FY.--*-  206
A1.SE.95.SE889I     ---D--EV-K-.TE---N-.----ICQ-----E---T-M-K------LT-R-------FY.-D--  205
A1.SE.95.UGSE8131   ---D--EV-K-.TE---N-.----MCQ-----E---T-M-K--PH---K-R-F------Y.--*-  210
A1.TZ.01.A173       ---D-EEV-K-.TE----R.----I-Q-----D---T-R-T---K--LK-I------DFY.-D--  196
A1.UA.01.01UADN139  ---D-AEV---.TE---N-.----ICQ--V--E-K---M-K------LT-I---K---FY.-D--  213
A1.UG.92.92UG037    ---DE-EV---.TG---N-.----ICQ-----E-K-T-R-K---S--RV-K-------FY.-D--  216
A1.UG.99.99UGA07072 ---D--EV-K-.TE---N-.----ICQ-----E---T-M-K------LK-R-L-----FY.-D--  205
A2.CD.97.97CDKTB48  ---D-SEV---.TE---N-.----ICQ--AE-------K-K------LR-L---Q---FY.-D--  217
A2.CY.94.94CY017_41 -----SEV---.TQ---N-.----ICQ--V--------R-E--RS--RR-R--------Y.-D--  218
A.SN.01.DDI579      ---DQASV---.TE--.--.-----CQ-----K-K---M--------LK-T---M---FY.-D--  204
A.SN.01.DDI369      ---DQAEV--T.TG---N-.-----CQ--V--T-G---M-K------LK-I-A-KY-DWY.-DR-  207
A.SN.96.DDI360      ---DQAAV---.TG---N-.----ICQ-----E-K---M-------LK-I-Q-----FY.-D--  207
A.CD.02.02CD_KTB035 ---DS-AV---.TER--N-.----INQ-R----K---M-Q-----LK-L-Q-Q---FY.-D--  207
A.CD.97.97CD_KCC2   -----EAV---.TG---N-.----ICQ----------M-K------LK-R-K-----FY.-D--  207
A.CD.97.97CD_KTB13  -----EAV-R-.TE---NC.----ICQ--AE-------M-K------LK-L-Q-M---FY.--*-  207
B.AR.04.04AR151516  ---DKEQV-K-.-E---NC.----G-----E-S-K---I-K---------M-------Y.-D--  208
B.AU.87.MBC925      -----E-----.-E---N-.----M-----E----------------M-------Y.-D--  207
B.BO.99.BOL0122     --LD--QV-R-.-E---N-.----M-Q-----K---V-K--------M-------FY.-D*L  204
B.BR.03.BREPM2012   --L--EEV-K-.-E---N-.----I-----E-A-G---V-------LR-L---K---FY.-D*-  216
B.CA.97.CANB3FULL   -----E-----.-E---NV.----M-----E---G---V-K--------M---K----Y.-D--  206
B.CN.05.05CNHB_hp3  ---D-EQV----.--E--NC.----M-Q--T--------M-K--------M---K---FY.-D--  205
B.CO.01.PCM001      -----EEV-K-.-E---N-.----M---------K---Q-K----------------F-..-D--  207
B.GB.83.CAM1        -----EQV----.--R--A-.----M-Q-------K---M-K--------M---K---FY.-D--  211
B.GB.86.GB8         -----EEV-K-.-E---NC.----M-Q--IE---K---V-K-N-------M-------FY.-D--  207
B.GE.03.03GEMZ010   ---DQEQV-K-.-E---NC.----I-Q--V----K---M-K---K-------K------Y.--*-  206
B.JP.05.DR6538      --M--EQV----.-----NC.----M-----E-------V-K---S--L--R---Q---FY.-D--  207
B.KR.05.05CSR3      ----KE-V--V.TER----.----I-Q--IE-------M-K------------D-----Y.----  204
B.NL.00.671_00T36   ---D--QV-K-.-E---NC.----M-Q--I----K---V-K--------M---I----Y.----  205
B.RU.04.04RU128005  ---DQE-V----.-E---N-.----M--------K---M-K--------I---K------.--*-  206
B.UA.01.01UAKV167   ---D--QV-K-.-E---NC.----M-Q-------K---M-K--------M---I----Y.-D--  205
B.US.04.ES10_53     ----TE-V----.-E---N-.----I---------K---Q-K-------R---------Y.----  212
B.US.99.PRB959_03   -------V----.-E---NI.----I-----E---K---Q-K---H---R-M--------Y.-D--  209
C.AR.01.ARG4006     ---D-REX---.XT---N-.----M-----E-SDG---Q-K---L--RR-M--------Y.-D--  216
C.BR.04.04BR013     ---D-KEV----.-E-D-N-.----M-----E-EHG---Q-K---L--RR-M--------Y.-D--  207
C.BW.00.00BW07621   ---D-KEV----.-E---NC.----M-Q-RI--DH----M-K---S--RR---------Y.-D--  207
C.ET.02.02ET_288    ---D-SEV----.-E---NC.----I-------EDK---M-K--IS--HR-L--------Y.-D*-  207
C.GE.03.03GEMZ033   ---D-GEV--D.------NC.----M-Q--V---DK---M-K---S--RR-L-------FY.-D--  208
C.IL.99.99ET7       ---D-KEV----.-E---NC.----M----IE-ED----R-K---H--RR-M------DFY.-D--  208
C.IN.99.01IN565_10  ---D-REV----.-E---NR.----MCQ--IE-GD----M-K---H--RK-M-------FY.-D--  208
C.KE.00.KER2010     ---D-REV----.--E--NC.----M-Q--IE-E-----R-Q---L--HR-L---Q---Y.-D--  208
C.MM.99.mIDU101_3   ---D-RDV----.T--DNC.-----CQ---E-EHG---Q-K---Q--LR-R-------FY.-D--  208
C.MW.93.93MW_965    ---D-REV----.-E---NC.----M-Q--IE-ED--I-R-K---Q--RR-L---Q---FY.-D--  208
C.SN.90.90SE_364    ---D-REV----.-E---NC.----M-----E-EX----M-K--IQ--HR-M-------FY.-D--  208
C.SO.89.89SM_145    ---D-REV--D.------N-.----M-Q---E-ED----M-K---H--RR---------Y.----  208
C.TZ.02.CO178       ---D-SEV----.-E---NC.----M-----E-ED----K-K---L--RR-M--------Y.-D..  212
C.UY.01.TRA3011     ---D-REV----.-E---N-.----M-M---E--HK---Q-K---L--HR-M--------Y.-D--  208
C.YE.02.02YE511     ---D-REV----.-E---N-.----I-----E-GDG---M-K------HR---------FY.-D--  220
C.ZA.04.04ZASK164B1 ---D-REV----.SS--TNC.----INQ-----ED----V-K------HI-M-------WY.-D*.  212
C.ZA.05.05ZAPSK240B1 ---D-REV----.-E--DNF.----M-----E-E-G---K-K---S--RR-L---I----Y.-D--  208
C.ZM.02.02ZM108     ---D-GGV----.SE--DNC.----ANQ-----EH----K-K---L--RR-M--------Y.-D--  209
D.CD.83.ELI         ---D-QEV--D.TE--TN-.----ICQ---E----Q--K---N-----E-K---M--FY.--*.  206
D.KE.01.NKU3006     ----IKEV--N.TE---NC.----MCQ---E-------M---N-----E-K--TM----Y.-D--  208
D.KR.04.04KBH8      ----SEEV----.-E---NC.----M-Q--IE----Q--K---N-K---E-K--M----FY.-D--  203
D.TD.99.MN011       ---D-KVV--E.TE---NC.----MNQ--G----K---V---N-----E-K-K-K---F-..-D--  207
D.TZ.01.A280        ---D-KEV--T.TE---SC.----I-Q---E-T-K---V---N-----E-K--S-----Y.-D--  208
D.UG.99.99UGK09259  -----E-V----.TE---NC.----INQ---E---K---I---------E-K-QQ-----YN-DR-  209
D.YE.01.01YE386     ---D-KEV--D.TE---SC.----MCQ---E-Q-K---M---N-K---E-K-KM----FY.-D--  208
D.YE.02.02YE516     ---D-KMV--E.-E--DKC.----MHQ---G---K---V---N-K---E-K--KT---F-..-D--  205
D.ZA.90.R1          ---D-QEV-Q-.TE--TNC.----MNQ---E-------K-K-------E-K--------Y.-D--  208
F1.BE.93.VI850      ---D-EEV-K-.-E---NC.----M-Q---E-ED----R-K---S--LR-I---R---FY.QD*.  205
F1.BR.01.01BR125    ---D-EEV-K-.SE---NC.----INQ---E-ED----Q-K---------M--------Y.--*-  206
F1.ES.x.P1146       ---P--EV-K-.-E--TNC.----M-Q---E-ADG---M-K---M---R---K-T---FY.QD*.  216
F1.FI.93.FIN9363    -----EEV-K-.-E---NC.----M-Q---E-ED----K-K------LK-I---R---FY.RD*.  206
F2.CM.95.MP257      ---D-EEV-K-.-E---NC.----M-----E-DDK---K-Q------LR-I---R----Y.-D..  200
F2.CM.97.CM53657    -----EEV-K-.-E---NC.----M-Q---E-ED----K-K---S--RR-R---M----Y.-D*.  205
G.BE.96.DRCBL       --M--SEV----.-----N-.----ICQ---E-EDG---V-----S--RR-L--------Y.-D--  208
G.CU.99.Cu74        --LD-AE----.T----N-.----ICQ---E-DD----I-----S--RR-L------DFY.-D--  208
G.ES.99.X138        E-MD-AEV---.TE---N-.----ICQ---E-ADN---I-K---X--RR-M------NFY.-D--  208
G.GH.03.03GH175G    --ID-AEV----.----TN-.----ICQ--IE-EDK---V-K---S--RR-L--------Y.----  208
G.KE.93.HH8793_12_1 ---D-AVV----.TTE--N-.----IWQ---E-ED----I-----S--RR-L-------FY.----  210
G.NG.01.01NGPL0669  --MN-AE-----.----N-.----LCQ---E-AD----V-----S--RR---------Y.--*-  211
G.PT.x.PT2695       --MD-AEV----.-----N-.----ICQ---E-ADK---V-----S--RR-I------NFY.-D--  208
G.SE.93.SE6165      --MD-AEV----.-----N-.----ICQ---E-ED----V-----S--RR-I--------Y.-D--  208
H.BE.93.VI991       ---D-QDV-K-.-E---N-.----MCQ--IE-------M-K------LR-R-K-----FY.-D--  208
H.BE.93.VI997       I--D-QEV-R-.-E---NC.--Y-ICQ---E-E-G---M-K------T-T---K---FY.-D--  208
H.CF.90.056         ---N-QEV-Q-.-E---N-.----M-----E-DG----M-K------LT-L--VK---.Y.-D--  207
J.CD.97.J_97DC_KTB147 ---D-GEV-R-.-E--DNR.----ICQ---E-ED----M-K--TS--RR-M-------FY.----  202
J.SE.93.SE7887      ---D-SEV----.-E---NC.----ICQ--IE-E-----Q-K---S--RR-I--------FY.-D--  204
K.CD.97.EQTB11C     ---D-REV----.TE---NC.-----NQ---E-EH----K-K---S--RK-----M----Y.-D--  207
K.CM.96.MP535       ---D-AEV--T.TE--DNC.----INQ---E-EH--I-M-K---S--RR--------D-Y.-D--  216
```

```
                                                                normal Nef end
B.FR.83.HXB2          VPVEPDKIEEA.NKGENTS.LLHPVSLHGMDDPEREVLEWRFDSRLAFHHVARELHPEYF.KNC*   205
01_AE.CF.90.90CF11697 ---D-KEV--D.--D-SNC.----M-Q--V--E-----M-K---S--RR-I----R---Y.-D--   205
01_AE.CN.05.FJ051     I--D-KEA--D.T----SC.----M-Q---G-EHG---M-K---S--RK-I---KY-QFY.-D--   207
01_AE.CN.06.FJ054     ---D-KDV--N.T-----C.----I-Q--LE-E-G---M-K---A--RR-----M----Y.-D--   201
01_AE.HK.x.HK001      ---D-REV--D.-----NC.----A-Q--I--E-----M-K---A--RK-L-------FY.-D--   207
02_AG.GH.03.GHNJ196   --ID-AEV---.-E---NV.----ICQ-----ED----V-K-------T-T-------FY.-D--   229
02_AG.NG.01.PL0710    --MD-AEV---.TE---N-.----ICQ---E-DH-Q--V-K---Q---R---H------Y.-D--   212
02_AG.NG.x.IBNG       --MD-AEV---.-E---N-.----ICQ---E-DD----I---------R-T--------Y.-D--   207
02_AG.SE.94.SE7812    --MD-AD--KD.TE---N-.----ICQ---E-ED----V---------T-K---M---FY.-D--   207
02_AG.SN.98.MP1211    --MD-AEV---.-Q---N-.----ICQ---E-EDK---V---------R-T--------Y.-D--   210
03_AB.RU.97.KAL153_2  ---D-AEV---.TE---N-.----ICQ-----E-K---M-K------LT-R-------FY.-D--   203
04_cpx.CY.94.CY032    ---D-QEV---.TE----C.----I-Q---E-E-----K-K------YK---------FY.-D--   220
05_DF.BE.x.VI1310     ---N-EEV-K-.-E--DNC.----M-----E-DD----Q-K---S--LR-I---R---FY.QD*#   206
06_cpx.AU.96.BFP90    ---D-EEV--L.T----NC.----ICQ--AE-E-----K-K---S--RR-I---K---FY.-D--   206
08_BC.CN.97.97CNGX_6F ---D-REV---.-E--DNC.-----CQ---E-EH----K-K---Q--HR-R-------FY.-D--   207
09_cpx.GH.96.96GH29I1 ---D-KEV---.TG---N-.----MCQ--V-S.GG---M-----T--LKXI---K---FY.-D--   207
10_CD.TZ.96.96TZ_BF061 ---D-REV---.-----NC.----M-----E-KHG---V-K---T--HK-I--------Y.-D--   208
11_cpx.GR.x.GR17      ---D-REV---.-E---NC.----M-Q--IE-ED----R-K---S--RR-I------DFY.-D*.   207
12_BF.AR.99.ARMA159   ---D-VEV-K-.-E---NC.----M-Q---E-ED----I-K------LR-L---K---WY.-D*.   205
13_cpx.CM.96.1849     --ID-KEV---.-E---NC.----M-Q--IE-A-----M-K---S--RK-I-------FY.-D--   200
14_BG.DE.01.9196_01   ---D-AE----.T----N-.----ICQ---E-XDN---R-T---S--RR--------DFY.-D--   208
15_01B.TH.99.99TH_MU2079 ---D-SDV--N.-----NC.----M-Q---E-E-----K-Q---A--RK-I--------Y.-D--   201
16_A2D.KR.97.97KR004  ---S-AEV---.TE--TN-.----ICQ--IE------K-V---H--LV-K-------FY.-D--   217
17_BF.PE.02.PE02_PCR0155 ---N-EEV-K-.-E---NC.----M-Q---E-DD----M-K---X---R-L---K---FY.QD*.   213
18_cpx.CU.99.CU76     ---DSNEV---.-E---N-.----ICQ-----E-----K-K---A--RR-I---Q---FY.-D--   207
19_cpx.CU.99.CU7      ---D-KEA-KD.TE---NC.----ACQ-------K---V---N-----X-K--IK---FY.-D--   207
20_BG.CU.99.Cu103     --MD-AEV---.-N---N-.----ICQ---E-N-----I-----S--RR-L-------FY.-D--   204
21_A2D.KE.91.KNH1254  ---D-SEV--E.TQ---N-.----ICQ--AE------R-E---Q-GLR-----M---FY.-D--   208
23_BG.CU.03.CB118     ---D-ADV---.SN---N-.----ICQ---E-DDK---M-----S--RR-I---R--DFY.-D--   209
24_BG.CU.03.CB378     --MD-AEV---.-R---N-.----I-Q---E-D-G---V-----S--RR-I-------FY.-D--   224
24_BG.CU.03.CB471     --MD-AEV---.-----N-.----ICQ---E-A-G---V-----S--RR-I-------FY.-D--   210
25_cpx.CM.01.101BA    --MD-EE----.--ED-N-.----LCX-----AH----V-----S--RR-I---K---WY.-D--   209
27_cpx.CD.97.97CDKTB49 ---D-QEV---.-----S-.----ICQ---E-GDG---K-V------LK-L---T---FY.-D--   208
28_BF.BR.99.BREPM12609 ---D-EQV-A-.-E---N-.----M-Q-------K---M-K---------M--------Y.-D*.   209
29_BF.BR.02.BREPM119  ---D-E-V---.TE---NC.-----N----E---K---K-K---K-----L--------Y.----   207
31_BC.BR.02.110PA     ---D-REV---.-E---N-.----M-----E-EHG---Q-K---L--HK-I--------Y.--Y-   205
32_06A1.EE.01.EE0369  ---D-REV---.T---NC.----IC---AE-E-G---M-----S--RR-I---K---FY.-D--   207
33_01B.MY.05.05MYKL007_1 ---D-KEV---.--E--NC.----M-Q--I--ED----M-K---A--RK----------Y.-D--   207
34_01B.TH.99.OUR2478P ---D-GEVK-......--SC.----M-Q--IE-E-----M-K---A--RK-I-------FY.-D--   203
35_AD.AF.05.05AF095   ---D-VEV-K-.-E---N-.----MCQ-----E-G---M-K------LK-R-Q-----FY.-D--   205
36_cpx.CM.00.00CMNYU830 ..............................................................   131
37_cpx.CM.00.00CMNYU926 ..............................................................   132
38_BF1.UY.05.UY05_4752 ---D--EV-K-.-E---NC.----M-Q--I--AD----Q-K-------T---K-K---FY.Q-*.   208
39_BF.BR.03.03BRRJ103 ---D--EV-R-.-E---NC.----M-Q-----ADK---K-V---H---R-M-X............   194
40_BF.BR.04.04BRRJ115 ---N-EEV-K-.-E---NC.----M-Q---E-EDK---Q-K---K--C----LIM------.-DR-   211
42_BF.LU.03.luBF_05_03 ----X-QV-K-.TE-----.----M-----E-T-----V-K---K-------H-M----Y.-D--   207
43_02G.SA.03.J11223   --MD-ADV---.--E--N-.----ICQ---E-ED----V-----S--RR----------Y.-D--   215
U.CD.83.83CD003       ---D-KEV---.-E---NC.----ICQ-----E-K---M-K---S--RR-L-------FY.-D--   208
U.CD.90.90CD121E12    ---D-VEV---.-E---N-.----ICQ---E-E-K---M-K---H--RS-K----------.-D--   208
U.GR.99.GR303         ---D-REVRGG#-E---NCf----I-Q---E-T-----I-K------LR-R--GE----Y.-D--   214
U.NL.95.U_NL_95_H10986_D1 ---D-REV---.-E---NC.----M-Q-----AD----R-E---S--RV-Q--V-----Y.-D--   214
CPZ.CD.90.ANT         ----SPPDD......-RNI.----ACT--DG--HK-I-R-E--AS-MRR-I---R------.RD*.   198
CPZ.CM.01.SIVcpzCAM13 --LTTEQV-A-.-E-D-NC.----ICQ---E-ESK---I--------LR-I---K----Y.-D*.   207
CPZ.CM.05.SIVcpzEK505 --LS-EEV---.-Q-D-NV.----MCQ---E--DK---V-----S--RV-R-------FY.Q---   207
CPZ.CM.05.SIVcpzLB7   --LT-EEV---.-----NL.----ICQ---E-EDK---I-KY--Q--LR-I-----S--Y.----   207
CPZ.CM.05.SIVcpzMB66  --L--EEV-R-.-E-D-NI.----ICQ--QE-EA----V-T------LK-R--------Y.-D--   207
CPZ.CM.05.SIVcpzMT145 --LP-EEV-K-.-E---NC.----MCQ---E-E-----V-T---K--LR-L---K----Y.RD*.   205
CPZ.GA.88.GAB1        --LTEEQV-Q-.-E-D-NC.----ICQ---E-EDK---V--------LR-I---Q----Y.-D*.   205
CPZ.TZ.00.TAN1        ---D-PEDD-K......NI.----ACS--TT--DG-T-I-----S--RR-I---RY-----.*G.   196
CPZ.US.85.CPZUS       --LTEEEV-Q-.----TNI.----MCQ---E-EHG---I-Q--TE--RR-R-K--------.R-*.   209
CPZ.CM.x.SIVcpzDP943  --LT-EEV-K-.-E-DTNI.----ICQ-EIE-EHG---V--Y-----LR-I---Q----Y.QD*.   207
CPZ.CM.x.SIVcpzMB897  ---D-TDV---.-E-D-NV.----MCQ--QE-EH----M-Q---H--LT-R--------Y.-D--   207
N.CM.02.DJO0131       --LSAEAV---.-E-D-NA.----LCQ--V--GHK---Q-Q---S--RR-I------DFY.----   214
N.CM.04.04CM_1015_04  --LSAEAV---.-E-D-NA.----ICQ--V--DHK---V-----S--RR-I------DFY.----   213
N.CM.04.04CM_1131_03  --LSAEAV---.-E-D-NA.----ICQ--V--EHK---V-----S--RR-L------DFY.----   214
N.CM.95.YBF30         --LSAEEV---.-E-D-NA.----ICQ--A--DHK---V-----S--RR---------FY.----   213
N.CM.97.YBF106        --LSEEAV---.-E-D-NA.----ICQ--V--DHKQ--V-----S--RR---K----DFY.----   208
O.BE.87.ANT70         ---SEEEA-RLG-TC-RAN.----ACA--FE-THK-I-M-K--RS-GNT---MIT---L-.QKD-   214
O.CM.91.MVP5180       ---SAEEA-RLG-TN-DA-.----ACN--AE-AHG-I-K-Q--RS-GLT-I-LQK---L-.PK*.   211
O.CM.96.96CMABB637    ---SEAEA-RLG-TC-RAM.----ACN--S--AHG-IPK-Q--TS--RT---LIT---L-.NKD-   218
O.CM.98.98CMA104      ---SEAEA--LG--H-RAK.----ACA--FG-AHG-I-K-Q--RS-GLT---KIT---L-.PKD-   214
O.CM.99.99CMU4122     ---SEAEA--LG--CVTAK.----ACN--SE--HG-I-K-Q--RS-GNI---KVT---L-.LKD-   212
O.FR.92.VAU           ---SEAEA-ALG--C-RA-.----ACN--YE-QHK-I-K-Q--RS--NT---LIT---L-.SKD-   216
O.SN.99.SEMP1299      ---SEAEA--LG--C-RA-.----ACN--FE-NHGQI-K-Q--RS-GST---MVTN--L-.NKD-   212
O.US.99.99USTWLA      ---TEQEA-QLG--D-GAM.----ACN--SE-HHG-M-K-Q---S-GIT---MVT---L-.IKN-   212
```


HIV-1/SIVcpz Proteins

VI

HIV-2/SIV Proteins

Contents

VI-1	Introduction	**365**
VI-2	Annotated features	**366**
VI-3	Sequences	**367**
VI-4	Alignments	**369**
VI-4.1	Gag	369
VI-4.2	Pol	373
VI-4.3	Vif	380
VI-4.4	Vpx	382
VI-4.5	Vpr	383
VI-4.6	Tat	384
VI-4.7	Rev	385
VI-4.8	Env	386
VI-4.9	Nef	392

VI-1 Introduction

The number of HIV-2 and SIV sequences is growing much more slowly than for HIV-1, particularly for sequences of complete or near-complete genomes. In selecting the set of sequences to present here, we eliminated some sequences derived from the same sample, patient, or isolate. We further omitted a few that were very closely related. Within each of the protein alignments, we have included a slightly different set of sequences, drawing on both genome-length and subgenomic fragments in order to best represent viral diversity. Thus for HIV-2, we selected A and B genomes and subgenomic fragments that best represent the diversity of the virus in that region. We also included a single AB recombinant and both genomes available from other HIV-2 groups (G and U).

In these alignments, as in the database, we label sequences according to their native host. Thus, transfers of SIVsmm into humans have produced HIV-2 groups A through G. Unintentional transfers of SIVsmm into captive macaques are named according to the receptive host (Rhesus macaques = mac, Stump-tailed macaques = stm, and Pig-tailed macaques = mne). However, intentional, experimental cross-species transfers of virus remain labeled as being from the original host (e.g., HIV-1 injected into a chimpanzee is labeled HIV-1, not SIVcpz; and SIVsmm injected into a Rhesus macaque is labeled SIVsmm, not SIVmac).

VI-2 Annotated features

Features of HIV-2 annotated in the alignment that follows.

Feature	Protein	Location	Page
Gag p15 start	Gag	1	369
p15	Gag	135	369
p27	Gag	136	369
p27	Gag	364	371
p2	Gag	365	371
p2	Gag	381	371
p8	Gag	382	371
p8	Gag	433	371
p1	Gag	434	371
p1	Gag	447	371
p6	Gag	448	371
PTAP motif	Gag	458-461	371
PSAP in HIV-2 B, U	Gag	476-479	371
p6 end	Gag	510	372
Gag end	Gag	510	372
Pol p15 start (-1 from Gag)	Pol	1	373
p15	Pol	67	373
protease	Pol	68	373
protease	Pol	166	374
p51 RT start	Pol	167	374
D catalytic site	Pol	276	374
DD catalytic site	Pol	351	375
p51 RT end	Pol	605	376
p15 RNase H start	Pol	606	376
p15 RNase H end	Pol	725	377
p31 Integrase start	Pol	726	377
p31 Integrase end	Pol	1018	379
Pol end	Pol	1018	379
Vif start	Vif	1	380
Vif end	Vif	214	381
Vpx start	Vpx	1	382
Vpx end	Vpx	112	382
Vpr start	Vpr	1	383
Vpr end	Vpr	101	383
Tat start	Tat	1	384
exon 1 end	Tat	99	384
exon 2 start	Tat	100	384
Tat end	Tat	130	384
Rev start	Rev	1	385
exon 1 end	Rev	24	385
exon 2 start	Rev	25	385
Rev end	Rev	107	385
Env start	Env	1	386
signal peptide end	Env	22	386
gp120 start	Env	23	386
V3	Env	311-344	387
gp120 end	Env	525	389
gp41 start	Env	526	389

Feature	Protein	Location	Page
Env end	Env	879	391
Nef start	Nef	1	392
R17Y mutation	Nef	17	392
max HIV-1 similarity	Nef	153-182	392
premature stop in Mac239	Nef	93	392
normal Nef end	Nef	262	393

VI-3 Sequences

Sequences included in the HIV-2/SIV protein alignments.

Name	Accession	Proteins	Author	Reference
MAC.US.x.239	M33262	All	Kestler, H.	*Science* **248**(4959):1109-1112 (1990)
H2A.IN.NNVA.NEW	EU980602	All	Desai	Unpublished
H2A.GM.x.MCR35	AY509260	All	Schmitz, C.	*J Virol* **78**(4):2006-2016 (2004)
H2A.IN.95.CRIK_147	DQ307022	All	Santhosh, C.V.	*ARHR* **24**(10):1315-1317 (2008)
H2A.CI.88.UC2	U38293	All	Barnett, S.W.	*Virology* **222**(1); 257-61 (1996)
H2A.DE.x.BEN	M30502	All	Kirchhoff, F.	*Virology* **177**(1):305-311 (1990)
H2A.DE.x.PEI2	U22047	All	Kraus, G.	*ARHR* **14**(1); 65-77 (1998)
H2A.GH.x.GH1	M30895	All	Hasegawa, A.	*ARHR* **5**(6):593-604 (1989)
H2A.GM.87.D194	J04542	All	Kuehnel, H.	*Proc Natl Acad Sci USA* **86**(7):2383-2387 (1989)
H2A.GM.x.ISY	J04498	All	Franchini, G.	*Proc Natl Acad Sci USA* **86**(7):2433-2437 (1989)
H2A.GM.x.MCN13	AY509259	All	Schmitz, C.	*J Virol* **78**(4):2006-2016 (2004)
H2A.GW.86.FG	J03654	All	Zagury, J.F.	*Proc Natl Acad Sci USA* **85**(16):5941-5945 (1988)
H2A.GW.87.CAM2CG	D00835	All	Tristem, M.	*J Gen Virol* **72**(PT 3):721-724 (1991)
H2A.GW.x.ALI	AF082339	All	Azevedo-Pereira	Unpublished
H2A.GW.x.MDS	Z48731	All	Becker	Unpublished
H2A.SN.85.ROD	M15390	All	Clavel, F.	*Nature* **324**(6098):691-695 (1986)
H2B.CI.88.UC1	L07625	All	Barnett, S.W.	*J Virol* **67**(2):1006-1014 (1993)
H2B.CI.x.EHO	U27200	All	Rey-Cuille, M.A.	*Virology* **202**(1):471-476 (1994)
H2B.GH.86.D205	X61240	All	Dietrich, U.	*Nature* **342**(6252):948-950 (1989)
H2B.JP.01.KR020	AB100245	All	Kusagawa, S.	*ARHR* **19**(11):1045-1049 (2003)
H2G.CI.92.ABT96	AF208027	All	Brennan, C.A.	*ARHR* **13**(5):401-404 (1997)
H2U.FR.96.12034	AY530889	All	Damond, F.	*ARHR* **20**(6):666-672 (2004)
H2AB.CI.90.7312A	L36874	All	Gao	Unpublished
H2AB.CM.03CM_510_03.NEW	EU028345	All	Yamaguchi, J.	*ARHR* **24**(1):86-91 (2008)
MAC.US.x.MAC239_87082	AY600249	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.251_1A11	M76764	All	Marthas, M.L.	*J Med Primatol* **18**(3-4):311-9 (1989)
MAC.US.x.251_32H_PJ5	D01065	All	Rud, E.W.	*(in) Brown,F, Chanock,RM and Ginsberg,HS(Eds); VACCINES* **92**: MODERN APPROACHES TO NEW VACCINES INCLUDING PREVENTION OF AIDS: 229-235; Cold Spring Harbor Laboratory Press, New York (1992)

Name	Accession	Proteins	Author	Reference
MAC.US.x.251_BK28	M19499	All	Hirsch, V.	*Cell* **49**(3):307-319 (1987)
MAC.US.x.17EC1	AY033233	All	Anderson, M.G.	*Virology* **195**(2):616-626 (1993)
MAC.US.x.17EFR	AY033146	All	Flaherty, M.T.	*J Virol* **71**(8):5790-5798 (1997)
MAC.US.x.1937	AY611495	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.2065	AY611493	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.80035	AY611486	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.81035	AY599200	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.85013	AY611490	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.92050	AY603959	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.92077	AY599201	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.93057	AY611492	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.93062	AY607704	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.95058	AY611494	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.95086	AY607703	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.95112	AY588946	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96016	AY607701	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96020	AY611488	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96072	AY611491	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96081	AY597209	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96093	AY611489	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96114	AY588945	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96123	AY611487	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.96135	AY607702	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.97009	AY599199	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.97074	AY599198	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.MM142	Y00277	All	Chakrabarti, L.	*Nature* **328**(6130):543-547 (1987)
MAC.US.x.r80025	AY576480	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.r90131	AY576481	All	O"Connor, D.H.	*J Virol* **78**(24):14012-14022 (2004)
MAC.US.x.SMM142B	BD131285	All	Alizon, M.	Patent: JP 2002030099-A 2 29-JAN-2002; INSTITUT PASTEUR
MNE.US.82.MNE_8	M32741	All	Kimata, J.T.	*J Virol* **72**(1):245-256 (1998)
MNE.US.x.MNE027	U79412	All	Kimata, J.T.	*J Virol* **72**(1):245-256 (1998)
SMM.SL.92.SL92B	AF334679	All	Chen, Z.	*J Virol* **70**(6):3617-3627 (1996)
SMM.US.x.F236_H4	X14307	All	Hirsch, V.M.	*Nature* **339**(6223); 389-92 (1989)
SMM.US.x.H9	M80194	All	Courgnaud, V.	*J Virol* **66**(1):414-419 (1992)
SMM.US.x.PBJ14_15	L03295	All	Dewhurst, S.	*Nature* **345**(6276):636-640 (1990)
SMM.US.x.PBJA	M31325	All	Dewhurst, S.	*Nature* **345**(6276):636-640 (1990)
SMM.US.x.PBJ_143	M80193	All	Courgnaud, V.	*J Virol* **66**(1):414-419 (1992)
SMM.US.x.PBJ_6P6	L09212	All	Novembre, F.J.	*J Virol* **67**(5):2466-2474 (1993)
SMM.US.x.PGM53	AF077017	All	Novembre, F.J.	*J Virol* **72**(11):8841-8851 (1998)
SMM.US.x.SME543	U72748	All	Hirsch, V.	*J Virol* **71**(2):1608-1620 (1997)
STM.US.x.STM	M83293	All	Novembre, F.J.	*Virology* **186**(2):783-787 (1992)

```
                            Gag p15 start                                                                                                                      p15_p27
MAC.US.x.239                MGVRNSVLSGKKADELEKIRLRPNGKKKYMLKHVVWAANELDRFGLAESLLENKEGCQKILSVLAPLVPTGSENLKSLYNTVCVIWCIHAEEKVKHTEEAKQIVQRHLVVETGTTETMPKTSRPTAPSSGRGGNYPVQQ.IGGNYVHLPLSPRTLNAWVKLIEEKKFGAE 169
H2A.IN.NNVA.NEW             --AK----R--------R-----G-----RV---------------T-----S--------T--D-------------F----------------D-----KL------A----A-K--N-------P-EG---F---R.-----T-V--------V----V-------- 169
H2A.GM.x.MCR35.AY509260     --A-----K--------T-----G-----R---I------------------S-----R--T--D-------------F----------------D--G---------AA----A-K--N-------P--K---F----.V----T-V-------------V-------- 169
H2A.IN.95.CRIK              --AKS---R--------R-----G-----R------------K---T-----S--------T--D-------------F----------------D-----K-------A----A-K--N-------P--G---F---H.-----T-V-------------V-------- 169
H2A.CI.88.UC2.U38293        --A-----R--------------G-----R---I----------R---N---S--------T--D-------------F--------L-------D-----KL------A----A-K--N-------P--K---F----.A------V-------------V-------- 169
H2A.DE.x.BEN.M30502         --A-----R---------V----G-----R---I--------K---------S--------R--D-------------F--------L-------D-----KLA-----A----A-K--N-------P--KR-------.A------V-------------V-------- 169
H2A.DE.x.PEI2.U22047        --A-S---R---V----------G-----R---I-------GK---------S-------IT--D-------------F--------L-------D--G----------A----ADK--S----A--P-----------.-A---S-V-------------V-------- 169
H2A.GH.x.GH1.M30895         --A-----R--------------S-----R---I--------K---------S--------T--D-------------F--------L-------D-----KL-----GA----A-K--S-------P----R-F----TG----I-V-------------V-D------ 170
H2A.GM.87.D194.J04542       --A-----R---------V--------R-R----------------------S--------K--E-------------F--------L-------D-----KLA-----A----A-K--NI------P--K---F----.A----I-V-------------V-------- 169
H2A.GM.x.ISY.J04498         --AK----R--------------G-----R---I------------T-----S-------I---E----------------T-----L-------D-----R--G----A--E-A-K--NI------P--K---F----.---------------------V-------- 169
H2A.GM.x.MCN13.AY509259     --A-----K--------T-----G-----R---I------------------S-----R--T--G-------------F----------------D--G---------AA----A-K--N-------P--K---F----.V----T-V-------------V-------- 169
H2A.GW.86.FG.J03654         --A-----R--------------G-----K---I------------------S--------T--D-------------F----------------D--G----------A----A-K--N-------P--K..-F----.VA---T-I----G--------V-------- 167
H2A.GW.87.CAM2CG.D00835     --A-----R---------V----G-----K---I------------------S-----R--K--D-------------F----------------D-----R-AL---AA----A-K--D-------P--K-------S.-----T-V-------------V-------- 169
H2A.GW.x.ALI.AF082339       --A-----R-R------R-----G-----Q---I------------------S-----R--K--E-------------F-----V--V-------D--G----I----AA-IE-A-K--S-------P-EQ---F----.VA---T-V-------D-----V-------- 169
H2A.GW.x.MDS.Z48731         --A-S---R---------V----G-----R---I------------------S--------T--D-------------F------F---------D-----KFAR----A----A-K--N-------P----------N.-----T-V-------------V-------- 169
H2A.SN.85.ROD.M15390        --A-----R--------R-----G-----R---I-----K------------S--------T--D-M-----------F----------------D--G-----R----A----A-K--S---------EK-------H.V----T-I-------------V-------- 169
H2B.CI.88.UC1.L07625        --A-S-------T-----V----G---R-C---II--V--------------S----H---T----------------F------Y-L-------D-----K-A----AAD-...-K--A-------P-..--------.-A-----M-------------V-------- 164
H2B.CI.x.EHO.U27200         --A-G-------T-----V----G---R-----I---V---E-------R-GS----R--RK--G--------------------F-L-------D-----K-A----AAD-...-K--AM-K-SK-T-RL..A-----.-A---S---------------V-------- 164
H2B.GH.86.D205.X61240       --A-G-------T-----V----G-------------V--------------S--------K----------------F-I----F-L-------D-----K-A----AAD-...-K--A-NK----P-..--------.LA-------------------V-------- 164
H2B.JP.01.KR020.AB100245    --A-G-------T-----V----G-----C---I---V-----Y--------S----H---T--E-------------F------Y-L---V---D--D--K-A-S--AAN-...-KA-A--K----P---..------.VA---T---------------V-------- 164
H2G.CI.92.ABT96.AF208027    --A-S-------------V----G----------------------------S----------------------------------L---Q----------V-------------KV-A----I--P-----------.V--------------------V-------- 169
H2U.FR.96.12034.AY530889    --A--#------V-----V---------------------------------A----------------------------------L--------------------A---E-A-K--A-----X-P--GR-------.-----A---------------V-------- 168
H2AB.CI.90.7312A.L36874     --A-G-------T-----V----G----------I-------------N---S--------T-----M----------F--------L---K---D-----KLA-S----D-...-KT-TADK-A-T----..-S----.VA-----V-------------V----L--- 164
H2AB.CM.03CM_510_03.NEW     --A-S-------T-----V----G---R------I--V----------T---S--------K----------------F------F-L-------D-----K-A----A.DA...-K-TA-NK----P-..--------.-A-----V-------------V-------- 163
MAC.US.x.MAC239             --A--------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.251                --A---------------------------------------------------------------------------------------------------------------A-I----------------------.------------------------------ 169
MAC.US.x.251_32H            --A--------------------G------------------------------------------------------------------------------------------A------------------------.------------------------------ 169
MAC.US.x.251                --A--------------------G------------------------------------------------------------------------------------------A------------------------.------------------------------ 169
MAC.US.x.17EC1.AY033233     -------------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.17EFR.AY033146     -------------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.1937.AY611495      --A------------------------------------------X-----X-------V-------M-----------------------------------------------------------------------.---------------------V-------- 169
MAC.US.x.2065.AY611493      --A----------------------------------------------------------------X-------------X---------------------------------------------------------.------------------------------ 169
MAC.US.x.80035.AY611486     --A--------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.81035.AY599200     -------------------------------------------------------------------X-----------------------------------------------------------------------.------------------------------ 169
MAC.US.x.85013.AY611490     --A----------------------------------------------------------------X-----------------------------------------------------------------------.------------------------------ 169
MAC.US.x.92050.AY603959     --X------------------------R---------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.92077.AY599201     --A--------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.93057.AY611492     --A--------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.93062.AY607704     --A--------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.95058.AY611494     --A----------------------------------------------------------------X-----------------------------------------------------------------------.V----------------------------- 169
MAC.US.x.95086.AY607703     --A--------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.95112.AY588946     -------------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.96016.AY607701     --A--------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.96020.AY611488     -------------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.96072.AY611491     --A----------------------------------------------------------------L-----------------------------------------------------------------------.------------------------------ 169
MAC.US.x.96081.AY597209     -------------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.96093.AY611489     --A----------------------------------------------------------------M-----------------------------------------------------------------------.------------------------------ 169
MAC.US.x.96114.AY588945     -------------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.96123.AY611487     --A--------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.96135.AY607702     --X--------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.97009.AY599199     --A---------T------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.97074.AY599198     -------------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.MM142.Y00277       --A--------------------G-------------------------------------------------------------------------------------M----A------------F-----------.-----T------------------------ 169
MAC.US.x.r80025.AY576480    --A----------------------------------------------------------------X-----------------------------------------------------------------------.------------------------------ 169
MAC.US.x.r90131.AY576481    -------------------------------------------------------------------------------------------------------------------------------------------.------------------------------ 169
MAC.US.x.SMM142B.BD131285   --A--------------------G-------------------------------------------------------------------------------------M----A------------F-----------.-----T------------------------ 169
MNE.US.82.MNE               --A--------------------G------------------------------------------------------------------------------------------A------------------------.V----T------------------------ 169
MNE.US.x.MNE027.U79412      --A--------------------G-----------------------D--------------A---------------------------------------------------A---------------KR-------.-----T------------------------ 169
SMM.SL.92.SL92B.AF334679    --A-G-------------V----G-R-------II---R------S------S-----R--A-----M----------FS----V--L---M---D-----KT--S-----S--A-KL-AQ------P-..--------.V-N----T-------------V-------- 167
SMM.US.x.F236               --A--------E------V--------------------------------D--------------------------------------------------------------ADR--A-------P-----------.V--------------------V-------- 169
SMM.US.x.H9.M80194          -----------------------G---------I------------------------------------------------R-L-----------------------------ADK--A-------P-----------.V----T---------------V-------- 169
SMM.US.x.PBJ14              --A--------------------G---R-K---I--------------------------------------------------L-----------------------------ADK--A-------P--K--------.-----T------------------------ 169
SMM.US.x.PBJA.M31325        --A--------------------G---R-Q---I--------------------------------------------------L-----------------------------ADK--A-------P--K--------.-----T------------------------ 169
SMM.US.x.PBJ                -----------------------G---------I------------------------------------------------R-L-----------------------------ADK--A-------P-----------.V----T---------------V-------- 169
SMM.US.x.PBJ                --A--------------------G---R-K---I--------------------------------------------------L-----------------------------ADK--A-------P--K--------.-----T------------------------ 169
SMM.US.x.PGM53.AF077017     --A-------------------------------------------------T-------I--------A--------------------------------------------ADK--I-------P-----------.-----T---------------V-------- 169
SMM.US.x.SME543.U72748      --A------------------------------------------------D--------------------------------------------------------------ADK--A-------P-----------.V--------------------V-------- 169
STM.US.x.STM.M83293         --A-S-------------V----G----------------------------S-------IT--E-------------F-----------------------V-K---------ANK--A-------P-----------.V--------------------V-------- 169
```

```
MAC.US.x.239               VVPGFQALSEGCTPYDINQMLNCVGDHQAAMQIIRDIINEEAADWDLQHPQPAP...QQGQLREPSGSDIAGTTSSVDEQIQWMYRQQNPIPVGNIYRRWIQLGLQKCVRMYNPTNILDVKQGPKEPFQSYVDRFYKSLRAEQTDAAVKNWMTQTLLIQNANPDCKLVLK 336
H2A.IN.NNVA.NEW            -----------------------------------E----------V---I-G-LP..A------R---------T-E---E--F-A---V-----------I----------------I-------------------------P--------------------II-- 337
H2A.GM.x.MCR35.AY509260    -----------------------------------E----------V---I-G-LP..A------R---------T--------F-P---V-----------I----------------I-------------------------PT----------V------------ 337
H2A.IN.95.CRIK             ---R-------------------------------E----------A---I-G-LP..A------R---------T-E------F-A---V-----------I----------------I-------------------------P-----------V------------ 337
H2A.CI.88.UC2.U38293       --------------------------Q--------E----------A---I-G-LP..A----D-R---------T--------------V-----------I-------------V---------S----------------A-P-----------V------------ 337
H2A.DE.x.BEN.M30502        -----------------------------------E----------S---I-G-LP..A----D-R---------T----------P---V-----------I-------K--------I-------------------------P------------------------ 337
H2A.DE.x.PEI2.U22047       -----------------------------------E----------V---I-G-LP..A------R---------T-E------F-A---------------I------------------------------------------P-----------V------------ 337
H2A.GH.x.GH1.M30895        -----------------------------------E---D------A---I-G-LP..A----D-R---------T-E--------P---V-----------I------------------------------------------P------------------------ 338
H2A.GM.87.D194.J04542      -----------------------------------E----------A---I-G-LP..A----D-R---------T-----------P--V-----------I-----------------------S------------------P------------------------ 337
H2A.GM.x.ISY.J04498        -----------------------------------E----------V---I-G-LP..A----D-R---------T-E---E-----E--V-----------I----------------I------S------------------------------V-.S--------- 336
H2A.GM.x.MCN13.AY509259    -----------------------------------E-V--------V---I-G-LP..A------R---------T--------F-P---V-----------I----------------I-------------------------P-----------V------------ 337
H2A.GW.86.FG.J03654        -----------------------------------E----------VA--I-G-LP..A------R---------T-E------F-P---V-----------I----------------IN------------------------P------------------------ 335
H2A.GW.87.CAM2CG.D00835    -----------------------------------E----------AN--I-G-LP..A----D-R---------T-E------F-A---V-----------I----------------I------S------------------P-----------V------------ 337
H2A.GW.x.ALI.AF082339      -----------------------------------E----------VA--I-G-LP..A------R---------T-E------F-PR--V-----------I----------------I-------------------------P-----------V------------ 337
H2A.GW.x.MDS.Z48731        -----------------------------------E----------A---I-G-LP..A------R---------T--------F-A---V-----------I------------------------------------------------------------------- 337
H2A.SN.85.ROD.M15390       -----------------------------------E-------E--V---I-G-LP..A------R---------T-E------F-P---V-----------I----------------I-------------------------P-----------V------------ 337
H2B.CI.88.UC1.L07625       -----------------------------------E----------Q---I-G-LP..A----D-R---------T-E--------A---V-----------I----------------I-------------------------P------------------------ 332
H2B.CI.x.EHO.U27200        -------------------------E---------E----------Q---S-G-MP..A------R---------T-E--------P---V----------------------------I-------------------------P------------------------ 332
H2B.GH.86.D205.X61240      -------------------------E---------E----------Q---S-G-MP..A----D-R---------T-E--------A---V----------------------------I-------------------------P------------------------ 332
H2B.JP.01.KR020.AB100245   -------------------------E---------E---D------Q---S-G-MP..A----D-R---------T-E--------P--------------------------------I-------------------------P-----V------------------ 332
H2G.CI.92.ABT96.AF208027   ------------L------------E---------E-------X-------QG-PP..A--X---T---------TIE-----TH----------X-----X---------X--------------------------X------P------------------------ 337
H2U.FR.96.12034.AY530889   -------------------------E---------E---D------T---NQG-PP..A------R---------T-E--------P--------------------------------I-------------------------P------------------------ 336
H2AB.CI.90.7312A.L36874    -------------------------E------V--E----------Q---V-G-LP..A----D-R---------T-E--------G-S------------------------------------------------------A-T--RA---E---V------------ 332
H2AB.CM.03CM_510_03.NEW    -----------------------------------E---D------Q---S-G-LP..A----D-R---------T-E--------G---------------I----------------I-------------------------P---S-------V------------ 331
MAC.US.x.MAC239            ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.251               ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.251_32H           ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.251               ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.17EC1.AY033233    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.17EFR.AY033146    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.1937.AY611495     ------------C-----------------------V-----------------...------------------T----------------------X---------------------X------------------------P------------------------ 336
MAC.US.x.2065.AY611493     ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.80035.AY611486    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.81035.AY599200    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.85013.AY611490    ------------A-----------------------V-----------------...------------------XX--------------------------------------------------------------------------------------------- 336
MAC.US.x.92050.AY603959    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.92077.AY599201    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.93057.AY611492    ------------I-----------------------V-----------------...-------------------A--------------------------------------------------------------------------------------------- 336
MAC.US.x.93062.AY607704    -------------X----------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.95058.AY611494    ------------I-----------------------X-----------------...-------------------AE-----T----------------X-----------------------------------------------------------------X--- 336
MAC.US.x.95086.AY607703    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.95112.AY588946    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.96016.AY607701    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.96020.AY611488    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.96072.AY611491    ------------------------------------------------------...---------------------------------------------X------------------------------------------------------------------- 336
MAC.US.x.96081.AY597209    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.96093.AY611489    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.96114.AY588945    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.96123.AY611487    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.96135.AY607702    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.97009.AY599199    ------------------------------------------------------...---------------------------------------------------------------------------V------------------------------------- 336
MAC.US.x.97074.AY599198    ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.MM142.Y00277      --S---------L---------------------------------------Q--...------------------T-E-------------------------------------------------------------------P------------------------ 336
MAC.US.x.r80025.AY576480   ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.r90131.AY576481   ------------------------------------------------------...----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.SMM142B.BD131285  --S---------L---------------------------------------Q--...------------------T-E-------------------------------------------------------------------P------------------------ 336
MNE.US.82.MNE              -------------------------E---------E---------------Q--...------------------T---------------------------------------S-----------------------------P------------------------ 336
MNE.US.x.MNE027.U79412     ------------I------------E---------E---------------Q--...------------------T---------------------------------K---------IR------------------------P------------------------ 336
SMM.SL.92.SL92B.AF334679   -------------------------E---------E--------------RGQQPAQPA-G--------------TPS---E----A---V---D--------------------------------------------------P------------------------ 337
SMM.US.x.F236              -------------------------E---------E----------------G-LP..A------R---------T---------------------------------------------------------------------P------------------------ 337
SMM.US.x.H9.M80194         --------------X----------E---------E-------------X--G-IP..A------X---------TX-------X------------X----------------------------S------------------P-X---------------------- 337
SMM.US.x.PBJ14             -------------------------E---------E----------------G-IP..P------R---------T--------------------------------------------------------------S------P------------------------ 337
SMM.US.x.PBJA.M31325       -------------------------E---------E----------------G-IP..P------R---------T---------------------------------------------------------------------P------------------------ 337
SMM.US.x.PBJ               --------------X----------E---------E-------------X--G-IP..A------X---------TX-------X------------X-----------------------------------------------P-X---------------------- 337
SMM.US.x.PBJ               -------------------------E---------E----------------G-IP..P------R---------T--------------------------------------------------------------S------P---S-------------------- 337
SMM.US.x.PGM53.AF077017    -------------------------E---------E----------------G-IP..A----D-R---------T-E-------------------------------------------------------------------P------------------------ 337
SMM.US.x.SME543.U72748     -------------------------E---------E----------------G-LP..A------R---------T-E-------------------------------------------------------------------P------------------------ 337
STM.US.x.STM.M83293        -------------------L-----E--------KE----------M-----PGPLP.A-----------------PE------------------------------------V-----------T----------------A-PS------R---------------- 338
```

```
                                              p27_p2              p2_p8                                                                        p8_p1          p1_p6     PTAP motif          PSAP in HIV-2 B, U
MAC.US.x.239                 GLGVNPTLEEMLTACQGVGGPGQKARLMAEALKEALAPVPIPFAAAQQR..GPRK.PIKCWNCGKEGHSARQCRAPRRQGCWKCGKMDHVMAKCPDRQAGFLGLGPWGKKPRNFPMAQVHQGLMPTAPPEDPAVDLLKNYMQLGKQQREKQRE..............SRE  489
H2A.IN.NNVA.NEW              ---M-----G-----------------------MG-T----V-----.....-.T-R-----------------K--------PG-I--N----------M--R--------V--IP---T-----M-------EK---Q-RK---QRERPYKEVTE.DLLHLEQG-  500
H2A.GM.x.MCR35.AY509260      ---M------------------------------T-A---------.....KR.T---------------------------PG--I-N-----V----M--R--------V---P---T-----V-------EK---Q--R---QRERPYKEVTE.DLLHLEQG-  500
H2A.IN.95.CRIK               ---M-------------------------------P---------.....--.T-R--------------------------TG---N--E-------IS-R--------V---P---T-----M-------EK---Q-RK---QRERPYKEVTE.DLLHLEQGG  500
H2A.CI.88.UC2.U38293         ---M------------------------------T-P---------.....-R.T-R----------K--------------PG-I--N----------M-SR--------V--AP---I-----A---A---EK-L-Q-RK---QRERPYKEVTE.DLLHLEQG-  500
H2A.DE.x.BEN.M30502          ---M-----------------------------MG-S---------.....--.A-RY------------------------PG-I--N--E-------------R--------VT-AP---I-----A---AE--ER---Q-RK---QRERPYKEVTE.DLLHLEQ--  500
H2A.DE.x.PEI2.U22047         ---M------------I------------------A---------.....-R.T-------D--------------------SG---N--E-------I------------VTR-P---T-----A---A---EK-L-Q-RK-K-QKMRPYKEVTE.DLLHLEQG-  500
H2A.GH.x.GH1.M30895          ---M------------------------------T-P---------.....--.V-R--------------------------TG-------E-------M------------V--APP--I-----A-------ER---Q-RE---QRERPYKEVTE.DLLHLEQGK  501
H2A.GM.87.D194.J04542        ---M-------------S----------------T-A---------.....-R.A-R----------K--------------SG-I--N--E-------M--R--Q-----A--AP---I-----I-------EK---Q-RK---QRERPYKEVTE.DLLHLEQG-  500
H2A.GM.x.ISY.J04498          ---M------------I-----------------MR-A---------.....KR.A---------------------------SG-I--N-------------------------VVPSS---T-----M-------EK---Q-RK---QRQRPYKEVTE.DLLHLEQG-  499
H2A.GM.x.MCN13.AY509259      ---M------------------------------T-A---------.....KR.T---------------------------PG--I-N-----V----M--R--------V---P---T-----V-------E----Q--R---QRERPYKEVTE.DLLHLEQG-  500
H2A.GW.86.FG.J03654          ---M--------T--------------------VM--A---------.....--.TF-------------WS-----------SG---N----------I-H----A----A---P---T-----L-------EK---Q--R---QRERPYKEVTE.DLLRFEQA-  498
H2A.GW.87.CAM2CG.D00835      ---M-----------------------------MG-P---------.....-R.T---------------------------PG-I-TN----------I------------V---P---T-----L-------EK---Q--R---QRQRPYKEVTE.DLLHLEQG-  500
H2A.GW.x.ALI.AF082339        ---M-----------------------------MT-A---------.....-R.T------------------K--------PG-L--N--E-------------------VTR-P---T-----AE--A---EQ---Q-RK---QRERPYKEVTE.DLLHLEQG-  500
H2A.GW.x.MDS.Z48731          ---M-----------------------------MT-A---------.....-R.T---------------------------TG-I-------------I------------V---P---T-----A-------EK---Q-RK---QRERPYKEVTE.DLLHLEQGK  500
H2A.SN.85.ROD.M15390         ---M----------------------------VIG-A---------.....--.AF---------------------------PG-I-TN-----------------------V---P---T-----V-------EK---Q--R---QRERPYKEVTE.DLLHLEQG-  500
H2B.CI.88.UC1.L07625         ---M------------I-----------------T-A---------K..AGKRGTVT------V--T-K--------------QG-I-S---E---------------------T--P--VT-S---M---EGMTPRGATPSAPPADPAV-MLKSYMKMGRQQRE---  500
H2B.CI.x.EHO.U27200          ---M------------I-----------------T-STN------P-..AGKR.TVT------A--T----K-----------QQG-I-S---E-------F-----------V.-AP--IV-S---MN--FGMTPQGAIPSAPPADPAE-MLKNYMQLGKKQKEN--  498
H2B.GH.86.D205.X61240        ---M------------I-----------------T-A------V--K..AGKRGTVT------Q--T-----------------TG-I-S---E---------------------T--P--VT-S---MN--EGMTPRGATPSAPPADPAV-MLKSYMQMGRQQRE---  500
H2B.JP.01.KR020.AB100245     ---M------------I-----------------T-P-L----V---..AGKRGAVT------I--T----K-----------QG-I-S---E-------F-----------V--AP--IL-S---MN--ENMTPQGAMPSAPPADPAV-MLKDYMQLGRKQKGG--  500
H2G.CI.92.ABT96.AF208027     ---M------------X-----------------N-TAL-------K.T-GKRST----------TV----------------PG-I-----E-------F-------------T--P---T-S---M-----------X--RK-K-..................Q-N  488
H2U.FR.96.12034.AY530889     ---M------------I-----------------Q-M------V---.G---R.-VR---------T----K----------E-PG-N--S--ETKV--------E------------P---T-----AE------TPTAPPADPAVDLLKSYMQQGKK....QKEN--  500
H2AB.CI.90.7312A.L36874      ---PH-----------I-----------------T-A---------K..AG-RGTVT------G--T-K--------------PG-N--N--E---------------------------P-EIV-S---MNT-EGKTHQGAIPSAPPADPAV-MLKSYMQLGKQQREKQG  500
H2AB.CM.03CM_510_03.NEW      ---M------------I-------K---------T-A------F--K..TGKRGTVT------G--T----K-----------QG-I-S---E--V------------------T--P--VT-S---MN--EXMTPQGVTPSAPPADPAV-MLKSYMQMGKQQRE---  499
MAC.US.x.MAC239              ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.251                 ------------------------------------------------..-----.------------------------------------------------------------------T----------------------------------..............---  489
MAC.US.x.251_32H             ------------------------------------------------..-----.------------------------------------------------------------------T----------------------------------..............---  489
MAC.US.x.251                 ----------------------------------------------K-..-----.------------------------------------------------------------------T------------------------------....................---  485
MAC.US.x.17EC1.AY033233      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.17EFR.AY033146      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.1937.AY611495       -------------------------------X----------------..-----.--------------------------------------------XX--------------------------------------------------X-..............---  489
MAC.US.x.2065.AY611493       ------------------------------------------------..-----.-----------------------------------------------------------------X-----------------------------------..............---  489
MAC.US.x.80035.AY611486      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.81035.AY599200      ------------------------------------------------..-----.-----------------------------------------X---------------------------------------------------------..............-K-  489
MAC.US.x.85013.AY611490      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.92050.AY603959      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.92077.AY599201      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.93057.AY611492      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.93062.AY607704      ------------------------------------------------..-----.T--------------------------------------------------------------------------------------------------..............--X  489
MAC.US.x.95058.AY611494      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.95086.AY607703      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.95112.AY588946      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.96016.AY607701      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.96020.AY611488      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.96072.AY611491      ------------------------------------------------..-----.-----------------------------------------V---------------------------------------------------------..............---  489
MAC.US.x.96081.AY597209      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.96093.AY611489      ------------------------------------------------..-----.X-----------------------------------------------------------------------------------------------K-..............---  489
MAC.US.x.96114.AY588945      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.96123.AY611487      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.96135.AY607702      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.97009.AY599199      ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.97074.AY599198      ------------------------------------------------..-----.-------------------------------------------X-------------------------------------------------------..............---  489
MAC.US.x.MM142.Y00277        ---T-------------------------------A---------K..-----.--------------------------------------N--------------------------T------E---------H-------.................--G  485
MAC.US.x.r80025.AY576480     ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.r90131.AY576481     ------------------------------------------------..-----.---------------------------------------------------------------------------------------------------..............---  489
MAC.US.x.SMM142B.BD131285---T-------------------------------A---------K..-----.--------------------------------------N--------------------------T------E---------H-------.................--G  485
MNE.US.82.MNE                ---M-------------------------------G-L-------K..-----.-----------------T----------Q-G--------------F-------------M----T---------------------------.................-KR  485
MNE.US.x.MNE027.U79412       ---M-------------------K-----------G-L-------K..-----.----------------------------Q-G--------------F-------------M----T---------------------------.................-KR  485
SMM.SL.92.SL92B.AF334679     ---M-------------------------M-D--TGSLVAAQFRGAAKGQGN-PI-R-F----T-----------K------EEGRIQ-N--NQK-----------------.-T.TS-T-S---D..PAARIVKEYLEKA-REKT.................R-S  486
SMM.US.x.F236                ---M-------------------------------R-DQL----V--K..-Q--.T------------K---------------TG-------E----------------------MP---I----------------KM-RK---.................N--  486
SMM.US.x.H9.M80194           --XI-------------X-----X--------D--TQG-L--X-V--K..-QX-.I-X-X--X--------F-----X-------AG-------E----------------------MP---T---X------------KV-RR---.................N--  486
SMM.US.x.PBJ14               ---I----------------------------D--TQG-L----V--K..-Q--.I----------------------------AG-------E----------------------MP---T----------------KM-RR---.................N--  486
SMM.US.x.PBJA.M31325         ---I----------------------------D--TQG-L----V--K..-Q--.I----------------------------AG-------E----------------------MP---T----------------KM-RR---.................N--  486
SMM.US.x.PBJ                 --XI-------------X-----X--------D--TQG-L--X-V--K..-QX-.I-X-X--X-------------X-------AG-------E----------------------MP---T---X------------KV-RR---.................N--  486
SMM.US.x.PBJ                 ---I----------------------------D--TQG-L----V--K..-Q--.I----------------------------AG-------E----------------------MP---T----------------KM-RR---.................N--  486
SMM.US.x.PGM53.AF077017      ---M-------------------------------T-GQL----V--K..----.T------------K---------------TG-------E----------------------MP---T------------R---KM-RK---.................N--  486
SMM.US.x.SME543.U72748       ---M------------I------------------R-DQL----V--K..-Q-R.T-----------------------G---TG-------E-------F--------------MP---T----------------KM-RK---.................N--  486
STM.US.x.STM.M83293          ---M----------------------------VFQ-D-L--------...QG-R.TV-------A--T-K--KG---------QQG-Q-----E--V----F-------H------IP---T---------A---RS------K---.................--K  486
```

```
                                        p6 end
                                      Gag end
MAC.US.x.239                    KPYK.EVTEDLLHLNSLFGGDQ*   510
H2A.IN.NNVA.NEW                 A-CG.-T-----------K-L-    522
H2A.GM.x.MCR35.AY509260         A--R.-T----------*K---    521
H2A.IN.95.CRIK                  A-CG.-M-----------K---    522
H2A.CI.88.UC2.U38293            T-RR.-A-----------K---    522
H2A.DE.x.BEN.M30502             T-HR.-E-----------K---    522
H2A.DE.x.PEI2.U22047            T-H-.-A-----------K---    522
H2A.GH.x.GH1.M30895             A-HR.-A-----------K---    523
H2A.GM.87.D194.J04542           T-HR.GA-----------K---    522
H2A.GM.x.ISY.J04498             T-HR.-T-----------N---    521
H2A.GM.x.MCN13.AY509259         A-CR.-T----------*K---    521
H2A.GW.86.FG.J03654             T-CR.-T-----------K---    520
H2A.GW.87.CAM2CG.D00835         A-CR.-T-----------K---    522
H2A.GW.x.ALI.AF082339           T-H-.-------------K---    522
H2A.GW.x.MDS.Z48731             T-C-.-T-----------T---    522
H2A.SN.85.ROD.M15390            T--REPP-----------K---    523
H2B.CI.88.UC1.L07625            R---.-------------E---    522
H2B.CI.x.EHO.U27200             R---.-------------E---    520
H2B.GH.86.D205.X61240           R---.-------------E---    522
H2B.JP.01.KR020.AB100245        ----.-------------E---    522
H2G.CI.92.ABT96.AF208027        ----.---XX---S----D---    510
H2U.FR.96.12034.AY530889        R---.-------------N---    522
H2AB.CI.90.7312A.L36874         R---.-------------E---    522
H2AB.CM.03CM_510_03.NEW         R---.-------------E---    521
MAC.US.x.MAC239                 ----.------------------   511
MAC.US.x.251                    ----.-M----------------   511
MAC.US.x.251_32H                ----.------------------   511
MAC.US.x.251                    ----.------------------   507
MAC.US.x.17EC1.AY033233         ----.------------------   511
MAC.US.x.17EFR.AY033146         ----.------------------   511
MAC.US.x.1937.AY611495          ----.------------------   511
MAC.US.x.2065.AY611493          ----.------------------   511
MAC.US.x.80035.AY611486         ----.------------------   511
MAC.US.x.81035.AY599200         ----.------------------   511
MAC.US.x.85013.AY611490         ----.------------------   511
MAC.US.x.92050.AY603959         ----.------------------   511
MAC.US.x.92077.AY599201         ----.------------------   511
MAC.US.x.93057.AY611492         ----.------------------   511
MAC.US.x.93062.AY607704         ----.------------------   511
MAC.US.x.95058.AY611494         ----.------------------   511
MAC.US.x.95086.AY607703         ----.------------------   511
MAC.US.x.95112.AY588946         ----.------------------   511
MAC.US.x.96016.AY607701         ----.------------------   511
MAC.US.x.96020.AY611488         ----.------------------   511
MAC.US.x.96072.AY611491         ----.------------------   511
MAC.US.x.96081.AY597209         ----.------------------   511
MAC.US.x.96093.AY611489         ----.------------------   511
MAC.US.x.96114.AY588945         ----.------------------   511
MAC.US.x.96123.AY611487         ----.------------------   511
MAC.US.x.96135.AY607702         ----.------------------   511
MAC.US.x.97009.AY599199         ----.------------------   511
MAC.US.x.97074.AY599198         ----.------------------   511
MAC.US.x.MM142.Y00277           ----.------------------   507
MAC.US.x.r80025.AY576480        ----.------------------   511
MAC.US.x.r90131.AY576481        ----.------------------   511
MAC.US.x.SMM142B.BD131285       ----.------------------   507
MNE.US.82.MNE                   ----.-------------E---    507
MNE.US.x.MNE027.U79412          ----.--A----------E---    507
SMM.SL.92.SL92B.AF334679        R---.-------------E---    508
SMM.US.x.F236                   R---.-------------E---    508
SMM.US.x.H9.M80194              R---.-------------E---    508
SMM.US.x.PBJ14                  R---.-------------E---    508
SMM.US.x.PBJA.M31325            R---.-------------E---    508
SMM.US.x.PBJ                    R---.-------------E---    508
SMM.US.x.PBJ                    R---.-------------E---    508
SMM.US.x.PGM53.AF077017         R---.------------------   508
SMM.US.x.SME543.U72748          R---.-------------E---    508
STM.US.x.STM.M83293             T---.-------------E---    508
```

```
                              Pol p15 start (-1 from Gag)                                                           p15_protease
MAC.US.x.239                  FFRPWSMGKEAPQFPHGSSASGADANCSPRGPSCGSAKELHAVGQAAERKAER.............KQREALQ.GGDRGFAAPQFSLWRRPVVTAHIEGQPVEVLLDTGADDSIVTGIELGPHYTPKIVGGIGGFINTKEYKNVEIEVLGKRIKGTIMTGDTPINIFGR 155
H2A.IN.NNVA.NEW               ---DGPP------L-R-PNPA--NT-ST-N-S-S-PIG-IY-AREK--GAERETIQGGDR.GPTAPRAG-ST-R.-N--RL--------K------Y-------------------A-----SD-N--------------------------KVRA-------------- 168
H2A.GM.x.MCR35.AY509260       ---DGPP------L-R-P-PA--NT-ST-SRS-S-PTG-IY-ARKKTK-TERETIQRSDR.GLTAPRAG-STI-.RD---L-------LK------YV------------------A-----SN-S-------------------K------KVRA-------------- 168
H2A.IN.95.CRIK                ---DKPP----S-L-R-P-ST--NT-STSN-S-S-PTG-IY-AREK--GAERETIQGGDR.GPTAPRAG-GT-R.-N---L--------K------Y-------------------A-----SN-S----------------N---------KVRA-------------- 168
H2A.CI.88.UC2.U38293          ---DGLT------L-R-P-S----T-ST-SRSGS-PVR-IF-A-EK--GAEGETIQGGDG.RLTAPRAG-DTS-.R----L--------K------Y--D----------------A-----DN----V------------------K--N--VRA-------------- 168
H2A.DE.x.BEN.M30502           ---VGPT----S-L-RDP-P----T-STSGRS-S-TVG-IY-AREK--GAEGETIQRGDG.GLAAPRAE-DTS-.R----L--------K------Y--D----------------A-----DN-----------------------K--N--VRA-------------- 168
H2A.DE.x.PEI2.U22047          ---D-P-----S-L-RDP-PA---T-ST-SR--SRP-R-VL-AREE---AENETIQGGDR.GLTAPRTR-DTT-.R-------------K------YV------------------A-----SN-S---------------------K--N-KV-A-------------- 168
H2A.GH.x.GH1.M30895           ---DG--------L-R-P-S----T-ST-SRS-S--IGKIY-A-ER--GAEGETIQRGDG.RLTAPRAGKSTS-.R----L--------K------Y--V----------------A--Q--DN-V---------------I--I--K--N--VRA-------------- 168
H2A.GM.87.D194.J04542         ---DGPT--A---L-R-P-S----T-ST-NRS-S-PVG-IY-AREK---AEGETIQGGDG.GLTAPRAG-D-P-.R----L-T------K------F--D----------------A-----DN-----------------------K--N--VRA-------------- 168
H2A.GM.x.ISY.J04498           ---A-T-------L-R-PKFA--NT-ST-N-S-S-PTG-V--AREKT--AETKTIQRSDR.GLAASRAR-DTT-.RD---L--------K------Y--D----------------A-----SN-S-----------------D---R--N-KVRA-------------- 168
H2A.GM.x.MCN13.AY509259       ---DGPP------L-R-P-PA--NT-ST-SRS-S-PTG--Y-ARKKTK-TERETIQRSDR.GLTAPRAG-GTM-.RD---L-------LK------YV------------------A-----SN-S-------------------K------KVRA-------------- 168
H2A.GW.86.FG.J03654           ---D-PL---G--L-R-P-PA--NT-ST-I-S-S-PTG-IY-ARKK-KGAERETVQGSDR.GLTAFRAG-DTM-.-D---L--------K-------------------R-N----A-----SN-S---------------------------V-A-------------- 168
H2A.GW.87.CAM2CG.D00835       ---D-PL--------R-P-ST--NT-ST-I-S-S--TG-IY-AREK--GAETETIQRGDR.GLTAPRTR-GPM-.-DN--L--------K--------------------------A-----SN-S---------------------------VRA-------------- 168
H2A.GW.x.ALI.AF082339         ---A-P-----S-L-RNP-SA-INT-ST-SRA-S-P-GAVY-A-EK-K-AEREAIQRGDG.GLTAPRAG-DTT-.R----L--------K------Y-------------------A-----SN------------------ED---K--N--V-A-------------- 168
H2A.GW.x.MDS.Z48731           ---D-PL------L-R-PGSA--NT-ST-SRS-S-PTG-IY-AREK--GAERETIQRGDR.GLAAPRAGKDTM-.-DN-----------N--------------------------A-----SN-S---------------------------VRA-------------- 168
H2A.SN.85.ROD.M15390          ---TGPL------L-R-P-SA---T-ST-S-S-S--TG-IY-AREKT--AERETIQGSDR.GLTAPRAGGDTI-GATN--L--------K------Y-------------------A-----NN-S------------------------N-KVRA-------------- 169

H2B.CI.88.UC1.L07625          ---VRTL----S-L--DP----S-TI-T-D---R-HDTSGGDTIC-PC-SSSGDAEKLHEDGETAEREP--T--.-------------------R-C-------------------A-----SN-------------------D-----V---VRA-------------- 169
H2B.CI.x.EHO.U27200           ---VRPL----S---R.PGTP-DS-I-A-DE--IRHDTSGCDSICTPC-SSRGDAKELHATREEAEGE---T--.-------------------K-T----S--------------A-----SN---------------N---------V---VRA-V------------ 168
H2B.GH.86.D205.X61240         ---VRTL----S-L--DP----S-TI-T-DE--R-HDTSGGDTIC-PC-SSSGDAEKLHADGETTEREP--T--.-------------------K-C----S--------V-----A-----SN-------------------D-----V---VRA-------------- 169
H2B.JP.01.KR020.AB100245      ---I-PL----S-L-CSP-TP-DT-I-TTNE--REHDTSGCNAIC-PC-SSSGDVEGLHAAREKAERG-G-T--.-------------------K-Y----S--------------A-----SD-------------------------V--K-RA-------------- 169

H2AB.CI.90.7312A.L36874       ---A-TL----S-L---P----DS-I-A-DEH-R-QDTSGSDTIC-PC-SGSGDAEKLHATREAAEREAG----.-----L-------------Q-Y-------------------A-----SN--------------------I----V--KVRS-------------- 169
H2AB.CM.03CM_510_03.NEW       ---VRTL----S-L--DPG---S-TI-T-DE--RXHDTSGGDTIC-PC-SSSGDAEKLHADGETTEREP--T--.-----L-------------X-C------------X------A---X-SN-------X-----------D-----V--TVRA-------------- 169

H2G.CI.92.ABT96.AF208027      ---V-TL----S-L--DP-P--S-SISTTD---SRPTE----X-EE-KG.................AEKQT--.----XX----L----------Y--S----------------A-----SN-S---------------------K-V----QA-V------------ 151

H2U.FR.96.12034.AY530889      ---A-PL--K-S-L---P-PA-T-P-SPSSRT-S---DPDSPSSGPSSGS-KELHATGQE....TKGE---TI-.---G----------K--I---Y------K---N--------A--K-E-Y----T--E-E------KFH--N-K----T-R--------------- 165
MAC.US.x.MAC239               -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.251                  --------------------------------------------------T--.............-------.-D-----------------------------------------------------------------------------R--------------- 155
MAC.US.x.251_32H              -----------------------------------------A-----------.............-------.-----------------------------------------------L----------------------------------------------- 155
MAC.US.x.251                  -------------------------------------------------.................-------.----------------------------------------------------------------------K------------------------ 151
MAC.US.x.17EC1.AY033233       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.17EFR.AY033146       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.1937.AY611495        ---X-X----------------------------------------------X.............E------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.2065.AY611493        -----------------------X-----------------------------.............-------.-----------------------------------------------X----------------------------------------------- 155
MAC.US.x.80035.AY611486       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.81035.AY599200       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.85013.AY611490       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.92050.AY603959       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.92077.AY599201       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.93057.AY611492       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.93062.AY607704       -----------------------------------------------------.............-X-----.----------------------------------------------------------------------------------------------- 155
MAC.US.x.95058.AY611494       -----------------------------------------------------.............-------.------------------------------------------------X---------------------------------------------- 155
MAC.US.x.95086.AY607703       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.95112.AY588946       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.96016.AY607701       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.96020.AY611488       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.96072.AY611491       -----P-----------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.96081.AY597209       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.96093.AY611489       -----------------------------------------------------.............-------.------------------------------------------------X---------------------------------------------- 155
MAC.US.x.96114.AY588945       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.96123.AY611487       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.96135.AY607702       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.97009.AY599199       -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.97074.AY599198       -----X-----------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.MM142.Y00277         -----PL----------------------RT----------L-------.................-------.----------------------------------------------------------------------------------------------- 151
MAC.US.x.r80025.AY576480      -----------------------------------------------------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.r90131.AY576481      ----------------------------------------------K------.............-------.----------------------------------------------------------------------------------------------- 155
MAC.US.x.SMM142B.BD131285     -----PL----------------------RT----------L-------.................-------.----------------------------------------------------------------------------------------------- 151

MNE.US.82.MNE                 ---L-P-----------PN-------------------K------T---.................--K----.---------------------------A-------------A----------------------------K------------------------ 151
MNE.US.x.MNE027.U79412        ---L-P-----------PN-----T-------------K------T---.................--K----.--G------------------------A-------------A----------------------------------------------------- 151

SMM.SL.92.SL92B.AF334679      -----P-----S-L--.AD.NII-TVST--..-SSPDSQGVSRESTEGKD................-EEQ---.---G---------R-----R-L-----L-------------A------G---R--------------RT-K------V-RE-L------------ 148
SMM.US.x.F236                 ---A-P-----------PD-----T------S----TE---ED--K--G.................E---T--.---------------------Y--E----------------A------N-------------------D-K-K----V----------------- 151
SMM.US.x.H9.M80194            ---A-P-----------PD-----T---X--S----TE---EG--K--G.................E---T--.--NG-----------------Y--E--I-------------A------N-------------------D-K-K----V----------------- 151
SMM.US.x.PBJ14                ---A-P-----------PD-----T------S----TE---ED--K--G.................E---T--.--NG-------------I---Y--E----------------A------N-------------------D-K-K----V----------------- 151
SMM.US.x.PBJA.M31325          ---A-P-----------PD-----T------S----TE---ED--K--G.................E---T--.--NG-------------I---Y--E----------------A------N-------------------D-K-K----V----------------- 151
SMM.US.x.PBJ                  ---A-P-----------PD-----T---X--S----TE---EG--K--G.................E---T--.--NG-------------I---Y--E----------------A------N-------------------D-K-K----V----------------- 151
SMM.US.x.PBJ                  ---A-P-----------PD-----T------S----TE---ED--K--G.................E---T--.--NG-------------I---Y--E----------------A------N-------------------D-K-K----V----------------- 151
SMM.US.x.PGM53.AF077017       ---A-P-----------PD------------S----TE---ED--K--G.................E---T--.---G---------------A-Y--E----------------A------N-------------------D-K-K----V----------------- 151
SMM.US.x.SME543.U72748        ---V-P-----------PD-----T------S----TE---ED--K--G.................E---T--.---------------------Y--E-----------------------N-------------------D-K-K----V----------------- 151

STM.US.x.STM.M83293           ---L-P-------L---PNT-------PS-R-----EK---A--E----.................E-EDT--.-----------------------------------------A-----LQ----V------------F---N------K----------------- 151
```

```
                                 protease  p51 RT start                                                                                                      D catalytic site
MAC.US.x.239                NLLTALGMSLNFPIAKVEPVKVALKPGKDGPKLKQWPLSKEKIVALREICEK.MEKDGQLEEAPPTNPYNTPTFAIKKKDKNKWRMLIDFRELNRVTQDFTEVQLGIPHPAGLAKRKRITVLDIGDAYFSIPLDEEFRQYTAFTLPSVNNAEPGKRYIYKVLPQGWKGSP 324
H2A.IN.NNVA.NEW             -I---------L-V--I--I--K----------R----T----E--K-----.--RE-------------------R-----------------K---E---I------------KR------V---------H-D---------------------------------- 337
H2A.GM.x.MCR35.AY509260     -I---------L-V--I--I-IM----------R----T----E--K-----.--RE-------------------------------------K-------I------------K-----------------H-D-----------I---------------------- 337
H2A.IN.95.CRIK              -I--V------L-V--I--I-IM----------R----T----E--K-----.---E-------------------------------------K---E---I------------KR------V---------H-D---------------------------------- 337
H2A.CI.88.UC2.U38293        -I---------L-V-RI--I-IT--------R------T---VE--K-----.---E---------------------------------------------I------------K-------V---------Y-D--P-----------V------------------- 337
H2A.DE.x.BEN.M30502         -I---------L-V--I--I--T--------R------T----E--K-----.---E-------------------------------------K-------I------------K---SI--V---------H-D----------A---M------------------- 337
H2A.DE.x.PEI2.U22047        -I---------L-V---D-I--I---------VR----T----E--K-----.--RE-------------------------------------K---E---I------------KR----------------H-D----------T----------------------- 337
H2A.GH.x.GH1.M30895         -I---------L----I--I--T--------R-R----T----E--------.---E---------------------------------------------I------------K-------V---------H-D---------------------------------- 337
H2A.GM.87.D194.J04542       -I-AT------L-V--LD-I--T--------R------T----E--K-----.--RE---------------------------------------------I------------K-------V---------H-D-----------------E---V------------ 337
H2A.GM.x.ISY.J04498         -I---------L-V--I-----T---------QR----TR---E--------.--RE-------------------------------------K--------------------KR------V---------Y-D---------------------------------- 337
H2A.GM.x.MCN13.AY509259     -I---------L-V--I--I-IM----------R----T----E--K-----.--RE-------------------------------------K-------I------------K-----------------H-D-----------I---------------------- 337
H2A.GW.86.FG.J03654         -V---------L-V--I--I-IM--------R------T----E--K-----.---E-------------------------------------K-------I------------KR------V---------H-D---------------------------------- 337
H2A.GW.87.CAM2CG.D00835     -I---------L-V--I--I-IM--------R-R----T----E--K-----.---E-------------------R-----------------K-------I------------KR------V---------H-D---------------------------------- 337
H2A.GW.x.ALI.AF082339       -I---------L-V--I--IE-R----------R----T----E--K-----.T-RE-------------------------------------K-------I------------KR------V---------H-S---------------------------------- 337
H2A.GW.x.MDS.Z48731         -I---------L-V--I--I-IM----------R----T----E--K-----.--RE-------------------R-----------------K-------I------------KR------V---------H-D---------------------------------- 337
H2A.SN.85.ROD.M15390        -I---------L-V-----I-IM----------R----T----E--K-----.---E-------------------------------------K-------I------------KR------V---------H-D--P------------------------------- 338

H2B.CI.88.UC1.L07625        -I-NT---T----V--I-----K---------IR---------L--K-----.---E---------------------R---------------K-------------------EKR------V----------PN-----------I---------------------- 338
H2B.CI.x.EHO.U27200         -I-NS---T----V-RI-----Q---E-----IR---------L--K-----.---E-------------S-----------------------K---E---------------SK-------V------V---PD----------A----------L------------ 337
H2B.GH.86.D205.X61240       -I-NT---T----V--------E---------IR-----R---L--K-----.---E-------------------------------------K--------NWVF-.TRQV-EKR----I-V----------PN---------------------------------Q 337
H2B.JP.01.KR020.AB100245    -I-NT---T----V--I-----K---------IR---------L--K-----#---E---------------------R---------------K-------------------E-------------------PD---------------------------------- 338

H2AB.CI.90.7312A.L36874     -I-NT---T----V--I-----Q---------IR---------L--------.---E------S------------------------------K---E---------------EKR----I-V------V---PD-----------I---------------------- 338
H2AB.CM.03CM_510_03.NEW     -I-NTX--T----V-XX-----EX---X-X--IR-----RX--L-XK-----.-X-E---------------------------------X---K--------X----------E-R----I-V----------PX--------X--T-XT------------X---X-- 338

H2G.CI.92.ABT96.AF208027    -I-VK-----------------R----M----IR---------Q------D-.--QE-----------------T---------------------------------------EKR-----------------VD---------------------------------- 320

H2U.FR.96.12034.AY530889    ----E---------G-EAL-IGG----------R--------LE--K-----.---E-------------------------------------------S-------------EK-------V---------C-D-----------I---------------------- 334
MAC.US.x.MAC239             ---------------------X------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.251                -----------L----------------V-----------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.251_32H            -----------L----------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.251                -----------L----------T-----V-----------------------.--------------------------------------------------------------------------------------------------------------------- 320
MAC.US.x.17EC1.AY033233     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.17EFR.AY033146     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.1937.AY611495      ----------------------------N-----------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.2065.AY611493      ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.80035.AY611486     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.81035.AY599200     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.85013.AY611490     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.92050.AY603959     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.92077.AY599201     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.93057.AY611492     ----------------------------E---X-------------------.----------------------------------------------------------------------------------X---------------------------------- 324
MAC.US.x.93062.AY607704     -----------XX---------------X---X-------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.95058.AY611494     ----------------------------E-----------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.95086.AY607703     ----------------------------E-----------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.95112.AY588946     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.96016.AY607701     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.96020.AY611488     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.96072.AY611491     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.96081.AY597209     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.96093.AY611489     -----------L----------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.96114.AY588945     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.96123.AY611487     ----------------------------E-----------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.96135.AY607702     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.97009.AY599199     ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.97074.AY599198     -----------L----------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.MM142.Y00277       -----------L---------SP-----------------------------.--------------------------------------------------------------------------------------------------------------------- 320
MAC.US.x.r80025.AY576480    ----------------------------E-----------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.r90131.AY576481    ----------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 324
MAC.US.x.SMM142B.BD131285   -----------L---------SP-----------------------------.--------------------------------------------------------------------------------------------------------------------- 320

MNE.US.82.MNE               ------E---------------T----------R------------------.--------------------------------I--------K----------------------------V---------------------------------------------- 320
MNE.US.x.MNE027.U79412      ----------------------T----------R------------------.-----------------------------------------K----------------------------V---------------------------------------------- 320

SMM.SL.92.SL92B.AF334679    -I-A---V---Y-V----YT--K--E-M---R-----------Q--T-----.---E----R-----------------N-D------------KM--E--------------KEME-------------V---P---------I-----Q------------------- 317
SMM.US.x.F236               -----M-----L-------I--T-----E----R---------I--------.-----------------------------------------K---------------------R------V---------------------------------------------- 320
SMM.US.x.H9.M80194          -----M-----L-----X-I--T----------R---------I--------.--------X---------------XX-G-------------K---------------X-----R------V---------------------------------------------- 320
SMM.US.x.PBJ14              -----M-----L-------I--T----------R---------I--------.-----------------------------------------K---------------------R------V---------------------------------------------- 320
SMM.US.x.PBJA.M31325        -----M-----L-------I--T----------R---------I--------.-----------------------------------------K---------------------R------V---------------------------------------------- 320
SMM.US.x.PBJ                -----M-----L-----X-I--T----------R---------I--------.--------X---------L-----XX-R-------------K---------------X-----R------V---------------------------------------------- 320
SMM.US.x.PBJ                -----M-----L-------I--T----------R---------I--------.-----------------------------------------K---------------------R------V---------------------------------------------- 320
SMM.US.x.PGM53.AF077017     -----M-----L-V-----I--T----------R------------------.---------------------------------------------------------------R------V---------------------------------------------- 320
SMM.US.x.SME543.U72748      -----M-------------I--T-----E----R---------I--------.-----------------------------------------K---------------------R------V---------------------------------------------- 320

STM.US.x.STM.M83293         -------------V--------T---------I----------E--K-----.-----------------------------------------K-------I-------------R------V-----------G---------------------------------- 320
```

```
                                                  DD catalytic site
MAC.US.x.239              AIFQYTMRHVLEPFRKANPDVTLVQYMDDILIASDRTDLEHDRVVLQSKELLNSIGFSTPEEKFQKDPPFQWMGYELWPTKWKLQKIELPQRETWTVNDIQKLVGVLNWAAQIYPGIKTKHLCRLIRGKMTLTEEVQWTEMAEAEYEENKIILSQEQEGCYYQEGKPLEA 494
H2A.IN.NNVA.NEW           --------KI--------Q--III------------S-P---K----L-----GL-----D--------YY----------------Q---K-V-----------------------------------L----------L----L---R---------H--K-E-E--- 507
H2A.GM.x.MCR35.AY509260   --------Q---------L--III------------------K----L-----NL-----D--------YH----------------Q---KDV-------------------------R---K----------------L----L-------------H----E-G--- 507
H2A.IN.95.CRIK            --------QI--------Q--III-----------------------L-----GL-----D--------Y-----------------Q---K-V---------------------R-------------L----------L----L---R---------H----E-E--- 507
H2A.CI.88.UC2.U38293      --------QI-----------I-I-------------G----K----L-----GL--------------------------------Q---K-------------I----------------------------------L----L---R---D-----H----E-E--- 507
H2A.DE.x.BEN.M30502       --------Q------------I-I-------------G----K----L-----GL-----D-------------C-----------LQ---KDI--------------------S-------------------------L----L-------------Y----E-E--- 507
H2A.DE.x.PEI2.U22047      ----H---Q------------I---------------------T---L-----GL-----D--------YK----------------Q---K-V----------------------------------------------L----L------------------E-E--- 507
H2A.GH.x.GH1.M30895       ----H---Q------------I-I-------------G----K----L-----GL-----D--------L----------------LQ---K-I--------------------------------K-------------L----L-------------Y----E-E--- 507
H2A.GM.87.D194.J04542     ----FM--QI-----------I-I-------------G----K----L-----GL-----D--------------------------Q---K-I-----------------------------K----------------L----L-------------S----EEE--- 507
H2A.GM.x.ISY.J04498       --------Q------------II-------------------K----L-----GL-----D--------Y-----------------Q---K-V-----------------------------K-------P--------L----L-------------H----E-E--- 507
H2A.GM.x.MCN13.AY509259   --------Q---------L--III------------------K----L-----NL-----D--------YH----------------Q---KDV-------------------------R---K----------------L----L-------------H----E-E--- 507
H2A.GW.86.FG.J03654       --------QI--------E--III------------------K----L-----GL-----D--------YR----------------Q---K-V----------------------------------------------L----L---R-----K---H----E-K--- 507
H2A.GW.87.CAM2CG.D00835   --------Q---------S--III------------------K----L-----NL-----D--------YR----------------Q---K-V----------------------------------------------L----L---R---------H----E-E--- 507
H2A.GW.x.ALI.AF082339     ----H---QI--------Q--I-I------------------K----L-----GL-----D--------YK----G-----------Q---K-V--------H-------------------------------G-----L----L---R---------H----E-E--- 507
H2A.GW.x.MDS.Z48731       --------Q---------Q--III-----------------------L------L-----D--------YR----------------Q---K-V-----------------------------------------I----L----L---RV--------H----E-E--- 507
H2A.SN.85.ROD.M15390      ----H---Q---------K--III-----------------------L-----GL-----D--------YH----------------Q---K-I------------------L---------------------------L----L---R---------H----E-E--- 508
H2B.CI.88.UC1.L07625      -----S--K--D------S--III------------S--------S-L-----DM---------------K--------KR------Q--EK-V------------------LF-----R-I-K----------------L----LQ------E-----S--K--V---- 508
H2B.CI.x.EHO.U27200       -------AK--D------N---II-------V----S--------S-L-----NM---------------K--------K-------Q--EK-V------------------LF-----R-I-K----------------L----FQ------E-----S--K--V---- 507
H2B.GH.86.D205.X61240     S-C--S--K--D------S--III------------S--------S-L-----DM---------------K--------K-------Q--EK-V----A-------------LF-----R-I-K----------------L----LQ------E-----S--K-RV---- 507
H2B.JP.01.KR020.AB100245  -----SVGK--D------D---I--------V----S----------L-----D----------------K--------K-------Q--E--R------------------LF-----R-I------------------L----FQ------E-----S--K--I---- 508
H2AB.CI.90.7312A.L36874   -----S--K--D---R--N---II-------V----S-----K--S-L-----DM------------------------K-------QI-E--V------------------LF-----R-I----K-------------L----MQ------G-----S--K--V---- 508
H2AB.CM.03CM_510_03.NEW   -----S--K--D------S---II------------S--------S-L-----NM-------X-XX----R-----X--K-------Q--EK-V------------------LF---X-R-M-X-----------X----L----LQ-X----E-----S--K--V---- 508
H2G.CI.92.ABT96.AF208027  ----H---K--D------E-XIXI-------V----------Q----L-----KX-----D--------Y------X--K-------N-XEK---------------X----L-------N-XK----X-----------L----LQ------X-X---A--K-XE-X-- 490
H2U.FR.96.12034.AY530889  T-------NL--------E-AVII------------N-----QL-S-L-----KT-----D--------HK--------K----------E-DV------------------V----------K--K------------------FA------------A--R-EEN--- 504
MAC.US.x.MAC239           -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.251              -------------------------------------N---------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.251_32H          -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.251              -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 490
MAC.US.x.17EC1.AY033233   ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.17EFR.AY033146   ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.1937.AY611495    -----------------------------------------------L---------------------------------------------X---------------------------------------------------------------------------- 494
MAC.US.x.2065.AY611493    -----------------------------------------------L---------X---------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.80035.AY611486   -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.81035.AY599200   -----------------------------------------------L----------------------------------X----------X---------------------------------------------------------------------------- 494
MAC.US.x.85013.AY611490   -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.92050.AY603959   -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.92077.AY599201   -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.93057.AY611492   -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.93062.AY607704   -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.95058.AY611494   -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.95086.AY607703   -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.95112.AY588946   -----------------------------------------------X-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.96016.AY607701   -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.96020.AY611488   ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.96072.AY611491   -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.96081.AY597209   ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.96093.AY611489   -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.96114.AY588945   -----------------------------------------------X-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.96123.AY611487   -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.96135.AY607702   -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.97009.AY599199   -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.97074.AY599198   ----------------------------------------------XX-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.MM142.Y00277     -----------------------------------------------L----------S---------------------------------------------------------------------------------------------------------S----- 490
MAC.US.x.r80025.AY576480  -----------------------------------------------L-------------------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.r90131.AY576481  -----------------------------------------------X---------X---------------------------------------------------------------------------------------------------------------- 494
MAC.US.x.SMM142B.BD131285-----------------------------------------------L----------S---------------------------------------------------------------------------------------------------------S----- 490
MNE.US.82.MNE             ----H---N----------------------V---------------L-------------------------------------------K------------------------------------------------------------------------------ 490
MNE.US.x.MNE027.U79412    ----H---N----------------------V---------------L-------------------------------------------K------------------------------------------------------------------------------ 490
SMM.SL.92.SL92B.AF334679  ----A---Q------------L-------L--G-N-GLT---KM-T-LRDM--NL------D----N--L-----L-Y-K----------EK-R--------------------------N--KM---------------L----LA--R---N-----R--R-DE---- 487
SMM.US.x.F236             --------N--------------I-----------------------L-----G-----------------------------------------------------------------------------------------------------------------I-- 490
SMM.US.x.H9.M80194        ---XH-X-N--------------I-----------------------L--------------------------------------------------N-----------------------------------------------------P---P------------- 490
SMM.US.x.PBJ14            ----H---N--------------I-----------------------L----------S--------------------------------------------------------------------------------------------------------------- 490
SMM.US.x.PBJA.M31325      ----H---N--------------I-----------------------L-------------------------------------------------------------------------------------------------------------------------- 490
SMM.US.x.PBJ              ---XH-X-N--------------I-----------------------L---------------------------------------------------------------------------------------------------------P---P------------- 490
SMM.US.x.PBJ              ----H---N--------------I-----------------------L-------------------------------------------------------------------------------------------------------------------------- 490
SMM.US.x.PGM53.AF077017   ----H---N--G-----------I-----------------------L---------------------------------------------------------------------------K---------------------------------------------- 490
SMM.US.x.SME543.U72748    --------N--------------I-----------------------L-----G-----------------------------------------------------------------------------------------------------------------I-- 490
STM.US.x.STM.M83293       --------NI-----R-------I-----------------------L-----NL-------------------------------------DV-----------------------------K------A---------------A------------Q--R-D----- 490
```

p51 RT end p15 RNase H start

```
MAC.US.x.239              TVIKSQDNQWSYKIHQEDKILKVG.KFAKIKNTHTNGVRLLAHVIQKIGKEAIVIWGQVPKFHLPVEKDVWEQWWTDYWQVTWIPEWDFISTPPLVRLVFNLVKDPIEGEETYYTDGSCNKQSKEGKAGYITDRGKDKVKVLEQTTNQQAELEAFLMALTDSGPKANIIV  663
H2A.IN.NNVA.NEW           --Q-D-----T--V--GE------.-Y--V--A---------Q-V-------L----RI--------RET-----D---------D---V--------A--------P-A--I----------------V----R-R----------------A---------V----  676
H2A.GM.x.MCR35.AY509260   --Q-D-----T--V--GE------.-Y---------------Q-V-------L----RI--------RET-----D---------D---V--------A--------PRT--F-------R---------V----R---RM-------------A--------------  676
H2A.IN.95.CRIK            --Q-D-----T--V--GE------.-Y--V--A---------Q-V-------L----RI--------RET-----D---------D---V--------A--------P-T--F-------R---------V----R-------------S-----A--------------  676
H2A.CI.88.UC2.U38293      -IQ-------T------E------.-Y---------------Q-V-------L----RI--------RET-----DN-----------V--------T----G---P-T--F-------R---------V----R---R---------------A---A-----V----  676
H2A.DE.x.BEN.M30502       -IQ---GH--T------E------.-Y---------------Q-V-------L----RI--------RET-----DN-----------V--------T----G---P-A--F-------R---------V----------------------V-R---A-----V----  676
H2A.DE.x.PEI2.U22047      --Q-D-----T-----GE------.-Y-----------------V-------L----RI--------RET-----DN-------D---V--------A--------P----F-------R---------------R---RI-------------A--------------  676
H2A.GH.x.GH1.M30895       -IQ-N-----T------E------.-Y---------------Q-V-------L----RI--------RET-----DN-----------V--------T----G---P-A--F-------R-------R-V----R---R---R-----------A-T-------V----  676
H2A.GM.87.D194.J04542     ----------A-----GERV----.-Y---------------Q-V-------L----R---------R-T-----DN------V----V--------T----G---P-T--F-------R---------V----R-R-R----S---------A---A-----V----  676
H2A.GM.x.ISY.J04498       --Q-D-----T--V--GE------.-Y---------------Q-V-------L----RI--------RET-----DN-------D---V--------A--------P-A--F-------R------------------RI-------------A--V-------V--V-  676
H2A.GM.x.MCN13.AY509259   --Q-D-----T--V--GE------.-Y---------------Q-V-------LI---RI--------RET-----D--------D---V--------A--------PRT--F-------R---------V----R---RM-------------A--------------  676
H2A.GW.86.FG.J03654       --Q-D-----T--V--GE----G-#-Y----------f----Q-V-------L----RI--------R-T-----DN-------D---V--------A----GE-VP-A--F-------R---------------R-R---------------A--------------  675
H2A.GW.87.CAM2CG.D00835   --Q-D-----T------E------.-Y----H------K---Q-V-------L--*-RI--------RE------DN-------D---V--------A----G---P-T--F-------R---------V----R----I-------------A--------------  675
H2A.GW.x.ALI.AF082339     --Q-D-----T-----GE-----E.-Y--M------------Q-V-------L----RI-R------RET-----D---------D---V--------A--------L-A--F----P--R---------------R-----------------AL-V-----------  676
H2A.GW.x.MDS.Z48731       --Q-D-----T--V--GE-T----.-Y--V------------Q-V-------L----RI--------REI-----D---------D---V------------------P-A--F-------R---------V--K-R----------------A--------------  676
H2A.SN.85.ROD.M15390      --Q-D-E---T------E------.-Y--V-------I----Q-V-------L----RI--------REI-----DN--------D---V--------A----G---P-A--F-------R---------V---------K------------A---------V----  677
H2B.CI.88.UC1.L07625      --Q-NLA---T-----GNR-----.-Y--V--------------V-------L----EI-V------RET-D------------------V--------AY------L-KV----------RA--------V-----------------------AL--Q----QV----  677
H2B.CI.x.EHO.U27200       --Q-NLA---T-----G-------.-Y--V--------------V-------L----EI-M------RET-D------------------V-----I--AY------L--V-----------A--------V---------P-------------AL--Q----QV----  676
H2B.GH.86.D205.X61240     --Q-NLA---T-----GN-V----.-Y--V--------------V-------L----EI-V------RET-D------------------V-----I--AY------L--R----------RT--------V-----------------------AL-----E-QV----  676
H2B.JP.01.KR020.AB100245  --Q-NLA---T-----G-R--R--.-Y--V----------------------L-T--EI-M------RET-D------------------V-----I--AY------L--V----------RT--------V---------A-------------AL--Q----QV----  677
H2AB.CI.90.7312A.L36874   --Q-NLA---T-----GN-V----.-Y--V--------I-----V-------L-----I-M------RET-D------------------V--------AY------L-KV----------RT--------V----R----P-------------AL--Q----QV----  677
H2AB.CM.03CM_510_03.NEW   --Q-NLA---T-----GXRM----.-Y--V-------XX-----V-------L----EI-I-----XRET-D-----X------------V-----I--AY--------XK-----------A--------V-------X--------------AL-XQ-X--QV----  677
H2G.CI.92.ABT96.AF208027  --Q-NL----T--V--GNR-----.-Y-----------------V-X-----L----KL-F------R-T---------------D---V---H-----Y----E-L-QT-----------T--X-----V--------X---------X-----A---Q-----V---X  659
H2U.FR.96.12034.AY530889  --L-N-----T-----G-RT----.----V--------------V------------KI-M------RE-------------VLD---V--------AY----E-LL.-----------RA--Q-----V------R--P--------------A---E-----V--VT  672
MAC.US.x.MAC239           ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.251              ----N-------------------.--------------------------------------------------------N-------------------------------------------------------------------------------------  663
MAC.US.x.251_32H          ------------------------.-----------------------------------------------------------------------------K-------------------------------------------------------------------  663
MAC.US.x.251              ------------------------.------------------------------------------R-------------------------------------------------------------------------------------------T----  659
MAC.US.x.17EC1.AY033233   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.17EFR.AY033146   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.1937.AY611495    ------------------------.------------------------------------------R-----------------X-----------------X-------------------------------------------------------------------  663
MAC.US.x.2065.AY611493    ------------------------.------------------------------------------X-----------------------------------------------------------------------------------------------------  663
MAC.US.x.80035.AY611486   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.81035.AY599200   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.85013.AY611490   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.92050.AY603959   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.92077.AY599201   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.93057.AY611492   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.93062.AY607704   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.95058.AY611494   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.95086.AY607703   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.95112.AY588946   ------------------------.--------------------------X-------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.96016.AY607701   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.96020.AY611488   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.96072.AY611491   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.96081.AY597209   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.96093.AY611489   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.96114.AY588945   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.96123.AY611487   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.96135.AY607702   ------------------------.------------------------------------------X-----------------------------------------------------------------------------------------------------  663
MAC.US.x.97009.AY599199   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.97074.AY599198   ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.MM142.Y00277     ------------------------.---------------------------------------------------------------------------------------------V----S-------------------------------------------------  659
MAC.US.x.r80025.AY576480  ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.r90131.AY576481  ------------------------.----------------------------------------------------------------------------------------------------------------------------------------------  663
MAC.US.x.SMM142B.BD131285-------------------------.---------------------------------------------------------------------------------------------V----S-------------------------------------------------  659
MNE.US.82.MNE             ----N-------------------.------------------------------------------------------K---------------------K------V-------------------------------------------------A-----------  659
MNE.US.x.MNE027.U79412    ----N-------------------.---------------------------------------------------------------------------------V----------------------R-------------------------------------  659
SMM.SL.92.SL92B.AF334679  --L-N-----------G-R-----.------------I----N-V------SL----KT-F------RE--D--------A--------------I------------K--V--I-----RN--------V-----E--LP---A-------Q-L-L--K--PS-V-VVT  656
SMM.US.x.F236             --------------------V---.----V--------------V-------L----E---------REI---------------D---V--------------E--Q-A--F-V-----R--R------V----R--A-L--------------YL--A-----------  659
SMM.US.x.H9.M80194        P-----------------------.-------------S-----V---------I-------R------XEI-------------------V--------------E--Q-A--F-V-R---R--R------V----RG-T-L--------------YL--A-----------  659
SMM.US.x.PBJ14            ------------------------.-------------------V-------------R------REI-------------------V--------------E--Q-A--F-V-----R--R------V----R--A-L--------------SL--A-----------  659
SMM.US.x.PBJA.M31325      ------------------------.-------------------V-------------R------REI-------------------V--------------E--Q-A--F-V-----R--R------V----R--A-L--------------YL--A-----------  659
SMM.US.x.PBJ              P-----------------------.-------------------V-------------R------XEI-------------------V--------------E--Q-A--F-V-R---R--R------V----RG-T-L--------------YL--A-----------  659
SMM.US.x.PBJ              ------------------------.-------------------V-------------R------REI-------------------V--------------E--Q-A--F-V-----R--R------V----R--A-L--------------YL--A-----------  659
SMM.US.x.PGM53.AF077017   ------------------------.-------------------VE-------------------REI-------------------V--------------E--Q-A--F-V-----R--R------V----R--A-P--------------YL--A-----------  659
SMM.US.x.SME543.U72748    --------------------V---.----V--------------V-------L----E---------REI---------------D---V--------------E--Q-A--F-V-----R--R------V----R--A-L--------------YL--A-----------  659
STM.US.x.STM.M83293       --V-D-------------------.---------------------------I--------RE------A-------------V--------------E-L--T--F-V-----R---------------N---A--------------A---A--------V-  659
```

```
                                                     p15 RNase H end_p31 Integrase start
MAC.US.x.239 DSQYVMGIITGCPTESESRLVNQIIEEMIKKSEIYVAWVPAHKGIGGNQEIDHLVSQGIRQVLFL.EKIEPAQEEHDKYHSNVKELVFKFGLPRIVARQIVDTCDKCHQKGEAIHGQANSDLGTWQMDCTHLEGKIIIVAVHVASGFIEAEVIPQETGRQTALFLLKLAG 832
H2A.IN.NNVA.NEW --------VA-Q-----N-------------EA-----------------V--------------.----------E---N-----SH-----KL------N--AQ-Q---------V-AE--------------V--------------------S------------S 845
H2A.GM.x.MCR35.AY509260 --------VA-Q-------I-----------EA-----------------V-Q------------.----------E---------SH-----KL------N--AQ-Q---------VDAE-----------------------------------S------------S 845
H2A.IN.95.CRIK --------VA-Q-----NKI-----------EA-----------------V--------------.----------E---------SH--R--KL------N--AQ-Q---------V-AE--------------V---------------------------------S 845
H2A.CI.88.UC2.U38293 --------VASQ-------I------D----EAV----I-----------V--------------.----------E-----I---IH---I-QL------N--AQ-Q---------V-AEI-V--------------------------------S------------S 845
H2A.DE.x.BEN.M30502 --------VA-Q-----N-I-----------EAV----------------V--------------.----------E----II---TH---I-LL------NS-AQ-Q---------V-AEI-V----Y---------------------------S------------S 845
H2A.DE.x.PEI2.U22047 --------VA-Q------K------------ETL----------------V--------------.----------E---------SH-----KL------N--AQ-Q---------VDAE------------------------------------------------S 845
H2A.GH.x.GH1.M30895 --------VV-Q-------I------D----EAV----------------V--------------.-R--------E-----M---TH---I-QL------N--AQ-Q---------V-AEI-V--------------------------------S------------S 845
H2A.GM.87.D194.J04542 --------VA-Q-----N-I------D----EAV----------------V--------------.----------E-----I---TH---I-QL------N--AQ-Q---------V-AEI-V--------------------------------S------------S 845
H2A.GM.x.ISY.J04498 --------V--Q-A-----I--K--------EA--------------------------------.-R--------G---------AH-----NL------N--AQ-Q---------V-AE-----------------------------------S------------S 845
H2A.GM.x.MCN13.AY509259 --------VA-Q-------I-----------EA-----------------V-Q------------.----------E---------SH-----KL------N--AQ-Q---------VDAE-----------------------------------S------------S 845
H2A.GW.86.FG.J03654 --------VA-Q-----N-I-----------EA-----------------V--------------.----------E-----I---SH---I-KL------N--AHVQ---------V-AE--------------V--------------------S------------S 844
H2A.GW.87.CAM2CG.D00835 --------VA-Q-----N-I-----------EA-----------------V--------------.----------E---T-----CH--DI-QL------N--AQYQ---------V-AEV----------------------------------S------------S 844
H2A.GW.x.ALI.AF082339 --------VA-Q-----N-I-----------EA-----------------V--------------.----------E---------SH-----NL------N--AQ-Q---------V-AE--------------V--I-----------------S------------S 845
H2A.GW.x.MDS.Z48731 --------VV-Q-----N-I-----------EA-----------------V--------------.----------E---------SH-----QL------N--AQ-Q---------V-A---------------V---------------------------------S 845
H2A.SN.85.ROD.M15390 --------SASQ------KI-----------EA-----------------V--------------.----------E---------SH---I-NL------NS-AQ-Q---------V-AE-----------------------------------S------------S 846
H2B.CI.88.UC1.L07625 --------VA-Q---T--P------------EA---G-----R-L-----V--------------.----------E---G------H-----QL--K---NS----Q-----V---V-AE--------------V---------------------------------S 846
H2B.CI.x.EHO.U27200 --------VAAQ---T--PI-RE--------EK---G-------L-----V----------I---.----------E---N------H---I-QL------NS----Q---------V--E--------------V---------------------------------S 845
H2B.GH.86.D205.X61240 ---------AAQ---T--PI-AK--------EAV--G-------L-----V--------------.----------E---G------H---I-QL--K---NS----Q---------V-A-------------------------------------------------S 845
H2B.JP.01.KR020.AB100245 ---------AAQ---T--PI-KT-V------EA---G-------L-----V--------HI----.-N--------E---G------H---I-QL--K---NS----Q---------V-AE------------------------------------------------S 846
H2AB.CI.90.7312A.L36874 --------VA-Q---T--PI--K--------EA---G-------L-----V--------------.----------E---G------H---I-QL--K---NS----Q-----V---V-AE--------------V---------------------------------S 846
H2AB.CM.03CM_510_03.NEW ---------AAX---T--PIXTX-------XEA---G-------L-----V--------------.----------X---G---X--H-X-X-QLX-K---NX----QX-X------V-AE--------------------------X---X----S------------S 846
H2G.CI.92.ABT96.AF208027 ------X--AXQ---T--P---K----X---EXL--G-------------V-X------------.X---------EXF-------TH---I-QL--K---NSYHS-Q---------V-AE--------------V-----------V---------------------- 828
H2U.FR.96.12034.AY530889 ----I----A-Q---T--P--SK----L-R-EAV--G-------------V--------------.----------E---------QH---I-QL--K---N--HV-Q-----M---V-TEV-------------VV--------------------------------- 841
MAC.US.x.MAC239 -----------------------------------------------------------------.---------------------------------------------------X---------------------------------------------------- 832
MAC.US.x.251 -------------------------K------------------------V--------------.---------------------------------------------------V---------------------------------------------------- 832
MAC.US.x.251_32H -----------------------------------------------------------------.---------------------------------------------------V---------------------------------------------------- 832
MAC.US.x.251 -----------------------------------------------------------------.---------------------------------------------------V------------------V--------------------------------- 828
MAC.US.x.17EC1.AY033233 -----------------------------------------------------------------.----------------------------------------------------------------------------------------------------------------------------------------------------- 832
MAC.US.x.17EFR.AY033146 -----------------------------------------------------------------.----------------------------------------------------------------------------------------------------------------------------------------------------- 832
MAC.US.x.1937.AY611495 -----------------------------------------------------------------.--------------------------------X---------X--------V--X------------------------------------------------- 832
MAC.US.x.2065.AY611493 -----------------------------------------------------------------.---------------------------------------------------V---------------------------------------------------- 832
MAC.US.x.80035.AY611486 -----------------------------------------------------------------.---------------------------------------------------V---------------------------------------------------- 832
MAC.US.x.81035.AY599200 -----------------------------------------------------------------.---------------------------------------------------X---------------------------------------------------- 832
MAC.US.x.85013.AY611490 -----------------------------------------------------------------.---------------------------------------------------V---------------------------------------------------- 832
MAC.US.x.92050.AY603959 -----------------------------------------------------------------.---------------------------------------------------X---------------------------------------------------- 832
MAC.US.x.92077.AY599201 -----------------------------------------------------------------.---------------------------------------------------V---------------------------------------------------- 832
MAC.US.x.93057.AY611492 -----------------------------------------------------------------.---------------------------------------------------V---------------------------------------------------- 832
MAC.US.x.93062.AY607704 -----------------------------------------------------------------.---------------------------------------------------V-------------------------X-------------------------- 832
MAC.US.x.95058.AY611494 -----------------------------------------------------------------.----------X----------------------------------------V---------------------------------------------------- 832
MAC.US.x.95086.AY607703 -----------------------------------------------------------------.---------------------------------------------------V---------------------------------------------------- 832
MAC.US.x.95112.AY588946 -----------------------------------------------------------------.---------------------------------------------------X---------------------------------------------------- 832
MAC.US.x.96016.AY607701 -----------------------------------------------------------------.---------------------------------------------------V---------------------------------------------------- 832
MAC.US.x.96020.AY611488 -----------------------------------------------------------------.----------------------------------------------------------------------------------------------------------------------------------------------------- 832
MAC.US.x.96072.AY611491 -----------------------------------------------------------------.---------------------------------------------------V---------------------------------------------------- 832
MAC.US.x.96081.AY597209 -----------------------------------------------------------------.----------------------------------------------------------------------------------------------------------------------------------------------------- 832
MAC.US.x.96093.AY611489 -----------------------------------------------------------------.---------------------------------------------------V---------------------------------------------------- 832
MAC.US.x.96114.AY588945 -----------------------------------------------------------------.----------------------------------------------------------------------------------------------------------------------------------------------------- 832
MAC.US.x.96123.AY611487 -----------------------------------------------------------------.---------------------------------------------------V---------------------------------------------------- 832
MAC.US.x.96135.AY607702 -----------------------------------------------------------------.---------------------------------------------------X---------------------------------------------------- 832
MAC.US.x.97009.AY599199 -----------------------------------------------------------------.---------------------------------------------------V---------------------------------------------------- 832
MAC.US.x.97074.AY599198 -----------------------------------------------------------------.---------------------------------------------------X--------I------------------------------------------- 832
MAC.US.x.MM142.Y00277 -------------------------------T---------------------------------.----------S-----I-----------L--K-------------------V------------------V--------------------------------S 828
MAC.US.x.r80025.AY576480 -----------------------------------------------------------------.---------------------------------------------------V---------------------------------------------------- 832
MAC.US.x.r90131.AY576481 -----------------------------------------------------------------.---------------------------------------------------X---------------------------------------------------- 832
MAC.US.x.SMM142B.BD131285-------------------------------T---------------------------------.----------S-----I-----------L--K-------------------V------------------V--------------------------------S 828
MNE.US.82.MNE -------------------------------T---------------------------------.----------------------------L--K-------------------V---------------------------------------------------S 828
MNE.US.x.MNE027.U79412 -------------------------------T---------------------------------.----------------------------L--K-------------------V---------------------------------------------------S 828
SMM.SL.92.SL92B.AF334679 -----LN----Q-S--D-DI-A----QLVQ-EAV-IG-----------N-V-R------------.-S------D-----------AQ-YNI-QL--K---NA-N--Q---------T-AEV-------------V-------------------R-----------I-S 825
SMM.US.x.F236 ---------A-Q-------------------EA-----------------V--------------#K---------E-----------------L--K-------------------V-AE------------------------------------------------- 828
SMM.US.x.H9.M80194 --------VA-Q---------------K---EA----------X--E---V------E-------.----------E-------K-----X---L--K----------L---X----V-XX----------------------G-----------X-------------S 828
SMM.US.x.PBJ14 --------VA-Q-------------------EA-----------------V--------------.----------E-----------------L--K-------------------V-AE------------------------------------------------S 828
SMM.US.x.PBJA.M31325 --------VA-Q-------------------EA-----------------V--------------.----------E-----------------L--K-------------------V-AE------------------------------------------------S 828
SMM.US.x.PBJ --------VA-Q-------------------EA----------X------V--------------.----------E-------------X---L--K--------------X----V-XX----------------------G-----------X-------------S 828
SMM.US.x.PBJ --------VA-Q-------------------EA-----------------V--------------.----------E-----------------L--K-------------------V-AE------------------------------------------------S 828
SMM.US.x.PGM53.AF077017 ---------A-Q-------------------EA-----------------V--------------.----------E--------------I--L--K------H------------V-AE------------------------------------------------S 828
SMM.US.x.SME543.U72748 --------VA-Q-------------------EA-----------------V--------------.----------E--------------I--L--K-------R-----------V-AE------------------------------------------------- 828
STM.US.x.STM.M83293 -----------Q------K------------EA-----------------V--------------.----------E--------------I--L--K-------------------V-AE------------------------------------------------S 828
```

```
MAC.US.x.239                RWPITHLHTDNGANFASQEVKMVAWWAGIEHTFGVPYNPQSQGVVEAMNHHLKNQIDRIREQANSVETIVLMAVHCMNFKRRGGIGDMTPAERLINMITTEQEIQFQQSKNSKFKNFRVYYREGRDQLWKGPGELLWKGEGAVILKVGTDIKVVPRRKAKIIKDYGGGKE 1002
H2A.IN.NNVA.NEW             ---------------T----------V---QS--------------------------------T-------------------------S--------A------L-A----L------F------------------D----V-------II--------R----RQ- 1015
H2A.GM.x.MCR35.AY509260     ---------------T----------V---QS------------------------S-------T--------T--------------------------------LHA----L------F------------------D----V----------------------RQ- 1015
H2A.IN.95.CRIK              ---------------T----------I---QS------------------------S-------T-------------------------S---------------L-A----L------F------------------D----V-------II--------R----RQ- 1015
H2A.CI.88.UC2.U38293        ------------P--T----------V---QS------------------------S-------T-----------------------------------------L-R---N--K-Q--------------------------V---A-------------R----RQ- 1015
H2A.DE.x.BEN.M30502         ------------P--T----------V---QS------------------------S-------TI----------------------------------------L-R---N----Q--------------------------V-----------------R----RQ- 1015
H2A.DE.x.PEI2.U22047        ---------------T----------T---QS------------------------S-------TM----------------------------------------LHA----L------F------------------D----V-------I---------R----RR- 1015
H2A.GH.x.GH1.M30895         ------------S--T----------I---QS------------------------S-------TI----------------------------------------L-R---N----Q---------------------D----V---A----I--------R----RQ- 1015
H2A.GM.87.D194.J04542       ------------P--T----------I---QS------------------------S-------TI----------------------------------------L-R---N--K-Q---------------------D----V---A-------------R----RQ- 1015
H2A.GM.x.ISY.J04498         ---------------T----------V---QS------------------------E-------TM------------------------V---V-----------L-A----L------F----N---Q---------D----V--------I--------R---PRQ- 1015
H2A.GM.x.MCN13.AY509259     ---------------T----------V---QS------------------------S-------T--------T--------------------------------LHA----L------F------------------D----V----------------------RQ- 1015
H2A.GW.86.FG.J03654         ---------------T----------V---Q---------------------------------T----------------------------I------------L-A----L------F------------------D----V----E-----------------RQ- 1014
H2A.GW.87.CAM2CG.D00835     ---------------T----------V---Q-------------------------S-------T-------------------------S---------------L-A----L------F------------------D----V-------II--------R----RQ- 1014
H2A.GW.x.ALI.AF082339       ----------S-V--T----------V---QS------------------------S-------T---------------------------------S-------L-T--L-----P---------------------D----V-----------------R----RQ- 1015
H2A.GW.x.MDS.Z48731         ---------------T----------I---QS------------------------S---D---T-------------------------S---------------L-A----L------F------------------D----V-------I---------R----RQD 1015
H2A.SN.85.ROD.M15390        ---------------T----------I---QS------------------------S-------TI-------I----------------S---------------L-A----L-D----F----------------------LV-------II--------R----RQ- 1016
H2B.CI.88.UC1.L07625        ---------------T--D---A---I---Q-----------------------------D--V-I--V----A--------------------V-----------F-A--L--Q--Q-------------------------LI----E---I--------RH------ 1016
H2B.CI.x.EHO.U27200         ---------------T--D---A---I---Q--------E--------------------D--V-I--V----T-------------------IV-----------L-T--L--Q------------------D----------I----E---I--------RN------ 1015
H2B.GH.86.D205.X61240       ---------------T-PS-------V---Q---------------------------L-D--V-I--V----T--------------------V-----------F-A--L--Q--Q--------------------------I----E------------RH-----G 1015
H2B.JP.01.KR020.AB100245    ---------------T--D---A---I-V-Q-----------------------------D--V-I--V----A-------------------IV----A------L-T-KL--Q-----F-----------------------I----E---I---------N------ 1016
H2AB.CI.90.7312A.L36874     ---------------T--D-------V---Q-----------------------------D--V-I--V----A--------------------V-----------F-A--L--Q--Q--------------------------I----E---I--------RH------ 1016
H2AB.CM.03CM_510_03.NEW     X--------------T--D-------V---Q-------------------------S-L-X--X-I--V-X--T--------------------V-----------F-A--L--Q--QX-------------------------X----E------------RH-----G 1016
H2G.CI.92.ABT96.AF208027    ---------------T----------V---QA--------------S------K---K--D----I---------------------------I--------------T-----------------------------------I----E--------------------  998
H2U.FR.96.12034.AY530889    ---------------T----------V-----------------------------E-------TM-------A-------------------I----S-------L-------------------------AD---------LI--------I--------R------- 1011
MAC.US.x.MAC239             -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.251                ---------------------------------------------------------------------------S--------T------------------------------------------------------------------------------------- 1002
MAC.US.x.251_32H            ---V---------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.251                --------------------------------------------------------------------------------------------------------------------------------------------------------------------------  998
MAC.US.x.17EC1.AY033233     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.17EFR.AY033146     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.1937.AY611495      ------------------------------------------------------------------------------------------X------X------------X--------------------------------V-------------------------- 1002
MAC.US.x.2065.AY611493      -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.80035.AY611486     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.81035.AY599200     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.85013.AY611490     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.92050.AY603959     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.92077.AY599201     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.93057.AY611492     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.93062.AY607704     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.95058.AY611494     -------X-------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.95086.AY607703     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.95112.AY588946     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.96016.AY607701     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.96020.AY611488     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.96072.AY611491     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.96081.AY597209     -----------------------------------------------------------------------------------------------------------------------------------------------------------------X-------- 1002
MAC.US.x.96093.AY611489     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.96114.AY588945     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.96123.AY611487     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.96135.AY607702     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.97009.AY599199     -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.97074.AY599198     -------------------------------------------------------------------------------------------------------------------X------------------------------------------------------ 1002
MAC.US.x.MM142.Y00277       --------------------------------------------------------------------------------------------------------------------------------------------------------------------------  998
MAC.US.x.r80025.AY576480    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.r90131.AY576481    -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1002
MAC.US.x.SMM142B.BD131285-------------------------------------------------------------------------------------------------------------------------------------------------------------------------  998
MNE.US.82.MNE               -----------------------------------------------------------------M----------------------------L------------------------------------------------V--------------------------  998
MNE.US.x.MNE027.U79412      ---------------T-------------------------------------------------M----------------------------L---------------------------------------------------------------------------  998
SMM.SL.92.SL92B.AF334679    ----K----------T----------L-V-QS----------------DL---KN--K-----E----L----A-------------------IV------L-T-YLN-Q----Q-----------------A----------VI----E-----------------RQ-  995
SMM.US.x.F236               ---------------T--------------Q----------------------T-----------I----------------------------V------------------------------------------------------E--------------------  998
SMM.US.x.H9.M80194          ---------------T-----------X--Q--X--------X----------T-----------I-------I--I-----------------V------------------------------------------------------E--------------------  998
SMM.US.x.PBJ14              ---------------T--------------Q----------------------T-----------IVS--------------------------V------------------------------------------------------E--------------------  998
SMM.US.x.PBJA.M31325        ---------------T--------------Q----------------------T-----------I----------------------------V------------------------------------------------------E--------------------  998
SMM.US.x.PBJ                ---------------T-----------X--Q--X--------X----------T-----------I-------I--------------------V------------------------------------------------------E--------------------  998
SMM.US.x.PBJ                ---------------T--------------Q----------------------T-----------IVS--------------------------V------------------------------------------------------E--------------------  998
SMM.US.x.PGM53.AF077017     ---------------T--------------Q----------------------T-----------I----------------------------V------------------------------------------------------E--------------------  998
SMM.US.x.SME543.U72748      ---------------T--------------Q----------------------T----------------------------------------V----------------------------------R-------------------E--------------------  998
STM.US.x.STM.M83293         ---V-----------T--------------Q----------------------T------D---T---V-----------K---L---------V-------------------------------------------------V-------------------------  998
```

```
                                  p31 Integrase end
                                      Pol end
MAC.US.x.239                      VDSSSHMEDTGEAR.EVA*                1019
H2A.IN.NNVA.NEW                   ---GP-L-G-R-DG.----                1033
H2A.GM.x.MCR35.AY509260           L--G--L-GAR-DG.-M--                1033
H2A.IN.95.CRIK                    M--GP-L-G-R-DG.----                1033
H2A.CI.88.UC2.U38293              L---P-L-GAR-DG.----                1033
H2A.DE.x.BEN.M30502               L---P-L-GAR-DG.-M-C                1033
H2A.DE.x.PEI2.U22047              -----L-G-R-DG.----                 1033
H2A.GH.x.GH1.M30895               L---HLEGARE-DG.----                1033
H2A.GM.87.D194.J04542             L-----L-GAR-DG.----                1033
H2A.GM.x.ISY.J04498               M--G--L-GAR-DG.-M--                1033
H2A.GM.x.MCN13.AY509259           L--G--L-GAR-DG.-M--                1033
H2A.GW.86.FG.J03654               M--G--L-GAR-DG.-M--                1032
H2A.GW.87.CAM2CG.D00835           L-----L-GAR-NG.----                1032
H2A.GW.x.ALI.AF082339             L--GP-L-GAR-DG.----                1033
H2A.GW.x.MDS.Z48731               M--GP-L-G-R-DG.----                1033
H2A.SN.85.ROD.M15390              M--G--L-GAR-DG.-M--                1034
H2B.CI.88.UC1.L07625              L-CGTD----RQ--.-M-Q               1034
H2B.CI.x.EHO.U27200               L-C-ADV---MQ--.---Q               1033
H2B.GH.86.D205.X61240             L-C-AD----RQ--.-M-Q               1033
H2B.JP.01.KR020.AB100245          L-C-AN-----Q--.-M-Q               1034
H2AB.CI.90.7312A.L36874           L-C-TDV---RQ--.-M-Q               1034
H2AB.CM.03CM_510_03.NEW           L-C-AD----RQ-S.-M-Q               1034
H2G.CI.92.ABT96.AF208027          ----TN----RQTG.----               1016
H2U.FR.96.12034.AY530889          L----NL-GAEKV-.-M-L               1029
MAC.US.x.MAC239                   --------------.----               1020
MAC.US.x.251                      M-------------.----               1020
MAC.US.x.251_32H                  --------------.----               1020
MAC.US.x.251                      --------------.----               1016
MAC.US.x.17EC1.AY033233           --------------.----               1020
MAC.US.x.17EFR.AY033146           --------------.----               1020
MAC.US.x.1937.AY611495            --------------.----               1020
MAC.US.x.2065.AY611493            --------------.----               1020
MAC.US.x.80035.AY611486           --------------.----               1020
MAC.US.x.81035.AY599200           --------------.----               1020
MAC.US.x.85013.AY611490           --------------.----               1020
MAC.US.x.92050.AY603959           --------------.----               1020
MAC.US.x.92077.AY599201           --------------.----               1020
MAC.US.x.93057.AY611492           --------------.----               1020
MAC.US.x.93062.AY607704           --------------.----               1020
MAC.US.x.95058.AY611494           --------------.----               1020
MAC.US.x.95086.AY607703           --------------.----               1020
MAC.US.x.95112.AY588946           --------------.----               1020
MAC.US.x.96016.AY607701           --------------.----               1020
MAC.US.x.96020.AY611488           --------------.----               1020
MAC.US.x.96072.AY611491           --------------.----               1020
MAC.US.x.96081.AY597209           --------------.----               1020
MAC.US.x.96093.AY611489           --------------.----               1020
MAC.US.x.96114.AY588945           --------------.----               1020
MAC.US.x.96123.AY611487           --------------.----               1020
MAC.US.x.96135.AY607702           --------------.----               1020
MAC.US.x.97009.AY599199           --------------.----               1020
MAC.US.x.97074.AY599198           --------------.----               1020
MAC.US.x.MM142.Y00277             M-------------.----               1016
MAC.US.x.r80025.AY576480          --------------.----               1020
MAC.US.x.r90131.AY576481          --------------.----               1020
MAC.US.x.SMM142B.BD131285         M-------------.----               1016
MNE.US.82.MNE                     --------------.----               1016
MNE.US.x.MNE027.U79412            --------------.----               1016
SMM.SL.92.SL92B.AF334679          MG--AS---Q*KNSGAL--               1013
SMM.US.x.F236                     L--G--L-------.----               1016
SMM.US.x.H9.M80194                L--G--L-------.----               1016
SMM.US.x.PBJ14                    L--G--L-------.----               1016
SMM.US.x.PBJA.M31325              L--G--L-------.----               1016
SMM.US.x.PBJ                      L--G--L-------.----               1016
SMM.US.x.PBJ                      L--G--L-------.----               1016
SMM.US.x.PGM53.AF077017           L--GP-L-------.----               1016
SMM.US.x.SME543.U72748            L--G--L-------.----               1016
STM.US.x.STM.M83293               ---G--L---R--G.----               1016
```


HIV-2/SIV Proteins

```
                              Vif start
MAC.US.x.239                  MEEEKRWIAVPTWRIPERLE.RWHSLIKYLKYKTKDLQKVCYVPHFKVGWAWWTCSRVIFPLQEGSHLEVQGYWHLTPEKGWLSTYAVRITWYSKNFWTDVTPNYADILLHSTYFPCFTAGEVRRAIRGEQLLSCCRFPRAHKYQVPSLQYLALKVVS..DVRSQGENPT  167
H2A.IN.NNVA.NEW               ---G---V-----V-G-M-.K----V---------ER------H---------------NNN-N--I-A--N---------S--------TEK-------D---S-I-G---A-------------KV----NY-Q---T---T--F---V--QQ.ND---RDST-  168
H2A.GM.x.MCR35.AY509260       ---G-N--V-----V-G-M-.K----V-----R----ED------H---------------KGN----I-A--N---------S-S--M---TER-------DC--S-I-----S-------------K-----NY-Q--RA------F---V--QQ.ND-P-RNGTP  168
H2A.IN.95.CRIK                ---D---V----KV-G-M-.K----V---------E-------H---------------KDN--I-I-A--N----E---S--------TE--------DCT-S-I-----A-------------KI----NY-Q--RT---T--F---V--QQ.ND---RDGT-  168
H2A.CI.88.UC2.U38293          ---G-S--V-----V-G-M-.-----V-----R----EG-R----H---------------KGE----I-A--N---------SHS--L---TEK-------DC----I-----C--A-----------K-----NY-Q---A----------V--QQ.ND-P-RKG-A  168
H2A.DE.x.BEN.M30502           ---DRN--V-----V-G-M-.K--A-V-----R----EE-R----H----------------GK----I-A--N---------SH---L---TEK-------DC----I-----S-------------K-----NY-Q---A----------V--QQ.ND-P-RKGTA  168
H2A.DE.x.PEI2.U22047          ---GE---V-----V-G-M-.K----V----HR----EG------H----------------GN----I-A--N---------S--------TER-------DC--S-I-----S-------------K-----NY-Q--RSK--L--F---V--QQ.NG-P-KNST-  168
H2A.GH.x.GH1.M30895           ---G-N--VPTWRVPGR-M-.-----V-----R-R--EE-R----H---------------KGE----I-A--N---------SHS------TER-------D-----I-----S-------------K-----NY-Q---V----------V--QQ.ND-P-RKGTA  168
H2A.GM.87.D194.J04542         ---G-N--V-----V-G-M-.-----V-H---R----EE-R----H---------------EGE----I-A--N---------SHS--L---TEK-------DC--S-I-----S-------------K-----NY-Q---A----------V--QQ.NG-P-RKGAA  168
H2A.GM.x.ISY.J04498           -DQG---------V-G-M-.K---------R----EQ-R----H---------------KGN----I-A--N---------S-S--M----EG-------DC--T-I-----S-------------KS----NY-Q---SK-----F---V--QQ.NDKP-RD-T-  168
H2A.GM.x.MCN13.AY509259       ---G-N--V-----V-G-M-.K----V-----R----ED------H---------------KGN----I-A--N---------S-S--M---TER-------DC--S-I-----S-------------K-----NY-Q--RA------F---V--QQ.ND-P-RNGTP  168
H2A.GW.86.FG.J03654           ---G---V--I--V-G-M-.-----V-----R----E-------H---------------K-N----I-A--N---------SHS------TEK-------DC--T-I-----S-------------K-----KY----RS------F---V--QQ.ND----NSA-  168
H2A.GW.87.CAM2CG.D00835       ---G-S--------V-G-M-.K----V-------G--EQ------H---------------RGD-R--I-A--N---------S----M---TEK-------DC--T-I-----S-------------K-----KY----RS------F---V--QQ.ND-P-RDRT-  168
H2A.GW.x.ALI.AF082339         ---G-S--V-----V-G-M-.K----V-----R----E-------H----------------GR----I-A--N---------S--------TEK-------DC--S-I-G---S-------------K-----NY-Q---S------F---V--QQ.NGKP-RNST-  168
H2A.GW.x.MDS.Z48731           ---G-T--V-----V-G-M-.K----V----------EE-R----H---------------KGN----I-A--N---------S----L---TEK-------DC--S-I-----T------A-------K-----NY-Q---S---T--F---V--QQ.NG-P-RDST-  168
H2A.SN.85.ROD.M15390          ---D---V-----V-G-M-.K----V----------E-------H---------------KGN----I-A--N---------S-S------TEK-------DC--V-I-----------------K-----NY----RA------F---V--QQ.ND-P-RDST-  168
H2B.CI.88.UC1.L07625          --G--N--V-------G---.K----V----HR--E--Q-S----H---------------K-EAY-------N----R-F--S----L---KRS-Y-----DV--Q---GS--S----N---------KI--Y-NY-S--EG------F---R-IQEGKDG----SA-  169
H2B.CI.x.EHO.U27200           -----N----------C---.-----------R-----Q-S----H---------------K--A--------N----R-F--S----L---ERS-Y-----DV--R---GS--SS---N---------KI--H-NY-S--TG------F---R--QEGKDG----ST-  169
H2B.GH.86.D205.X61240         -----D--V-------G---.-----------R-GE--Q-S----H-----------I---NK-AW-------N----R-F--S----L---ER--Y-----DV--Q---GS--S--S-N---------KI--Y-NY-S--EG------F---R--QEGKNG----SA-  169
H2B.JP.01.KR020.AB100245      -----N----------G---.K----V----HR--E-EQ-T----H---------------K--AY--I----N----R-F--A----L---K-S-Y-----DV--H---GS--F----N---------KI--Y-NY-S--RG------F---R--QEGGNGP---SAA  169
H2G.CI.92.ABT96.AF208027      ----------------G---.K------F--F-------AV----H---------------TKEA---I----N---------Q----L---TRK-Y-----ET--Q---GS--D---------------I----NY-T---R------F---Q--QK.GHG-K--SX-  168
H2AB.CI.90.7312A.L36874       --G--N--V----K--G---.K----V----HR-GE--Q-S----H---------------KG-AC-------N----R-F--S----L---EES-Y-----DV--Q---G---S----N---------KI--Y-NY-S--EG------F---R--QEGKNG----SA-  169
H2AB.CM.03CM_510_03.NEW       -----D--V----K--G--A.-------------GE--Q-S----H-X-------------DKXAX-X-----N----R-L-------L--XER--Y-----DV--Q---GS--S----N---------KI--Y-NY-S--EG------F---R-IQEGKNG---K-T-  169
H2U.FR.96.12034.AY530889      -----N--V-----V-R---.-------H-#-N--E---------H---------------KK-------A--N---------S----L---TRG-------DC--Q---GS--S-------------------N--S--RQ------F---RALQ..-G--KR----  166
MAC.US.x.MAC239               -------------------.------------------------------------------------------------------------------------------------------------------------------..----------  167
MAC.US.x.251                  -------------------.--------------------------------------------------------------V---------------------------------------------------------------..----------  167
MAC.US.x.251_32H              ---G---------------.--------------------Y-------------------------------R-------------R--------D-------------------V----------K-----R-----------R---..Y---------  167
MAC.US.x.251                  -------------------.----------------------------------------------------R-------------R--------D------------------------------K-----R---------------..----------  167
MAC.US.x.17EC1.AY033233       -------------------.------------------------------------------------------------------------------------------------------------------------------..----------  167
MAC.US.x.17EFR.AY033146       -------------------.------------------------------------------------------------------------------------------------------------------------------..----------  167
MAC.US.x.1937.AY611495        -------------------.--------------------------------------X-----------E----A------------------------S-------------------------------------------..----------  167
MAC.US.x.2065.AY611493        -------------------.------------------------------------------------N----E----X-----------------X--X----------------------------------------------..----------  167
MAC.US.x.80035.AY611486       -------------------.------------------------------------------------------------------------------X-----------------------------------------------..----------  167
MAC.US.x.81035.AY599200       -------------------.------------------------------------------------------------------------------------------------------------------------------..----------  167
MAC.US.x.85013.AY611490       -------------------.------------------------------------------------N--------------------------X-------------------------------------------------..----------  167
MAC.US.x.92050.AY603959       -------------------.------------------------------------------------------------------------------------------------------------------------------..----------  167
MAC.US.x.92077.AY599201       -------------------.------------------------------------------------------------------------------------------------------------------------------..----------  167
MAC.US.x.93057.AY611492       -------------------.--------------------------------------------------------------------------D---XI----------------------------------------------..----------  167
MAC.US.x.93062.AY607704       -------------------.------------------------------------------------N-------------------------------------------------------------------------------..----------  167
MAC.US.x.95058.AY611494       -------------------.------------------------------------------------X----------------X-------DX--X----------------------------------------------..----------  167
MAC.US.x.95086.AY607703       -------------------.------------------------------------------------N--------------------------X---X----------------------------------------------..----------  167
MAC.US.x.95112.AY588946       -------------------.------------------------------------------------------------------------------------------------------------------------------..----------  167
MAC.US.x.96016.AY607701       -------------------.-------------------------------------------------------------------------X--------------------------------------------------..----------  167
MAC.US.x.96020.AY611488       -------------------.------------------------------------------------------------------------------------------------------------------------------..----------  167
MAC.US.x.96072.AY611491       -------------------.------------------------------------------------N----E--------------------D-------------------------------------------------..----------  167
MAC.US.x.96081.AY597209       -------------------.------------------------------------------------------------------------------------------------------------------------------..----------  167
MAC.US.x.96093.AY611489       -------------------.------------------------------------------------N-------------------------------------------------------------------------------..----------  167
MAC.US.x.96114.AY588945       -------------------.------------------------------------------------------------------------------------------------------------------------------..----------  167
MAC.US.x.96123.AY611487       -------------------.------------------------------------------------------------------------------X-----------------------------------------------..----------  167
MAC.US.x.96135.AY607702       -------------------.------------------------------------------------------------------------------S-----------------------------------------------..-X--------  167
MAC.US.x.97009.AY599199       -------------------.------------------------------------------------------------------------------------------------------------------------------..----------  167
MAC.US.x.97074.AY599198       -------------------.-------------------------------------------------------------A----------------------------------------------------------------..----------  167
MAC.US.x.MM142.Y00277         --------V----------.-----------------A-----H------------------------N----R--------------D------E-------------------------R-----------H---------R---..H---------  167
MAC.US.x.r80025.AY576480      -------------------.------------------------------------------------N--------------------------X---X----------------------------------------------..----------  167
MAC.US.x.r90131.AY576481      -------------------.------------------------------------------------------------------------------------------------------------------------------..----------  167
MAC.US.x.SMM142B.BD131285     --------V----------.-----------------A-----H------------------------N----R--------------D------E-------------------------R-----------H---------R---..H---------  167
MNE.US.82.MNE                 -------------------.------------------------H-----------------K-Q-------N----R-----H-------R-------DC---------------------------------T---------R---..Y----R----  167
MNE.US.x.MNE027.U79412        -------------------.------------------------H-----------------E-Q-------N----R--------------R-------D----------------------------------N---------R---..Y----R----  167
SMM.SL.92.SL92B.AF334679      ---G---V---V---SR-IVE----C--FH----RE-E-A-----H------Y-A-------E--------V--N---------S---------EK-------DV--Q-A--------A-HA--Q------V--Y-GYAV--HSS-Q---L------LQ.ND-PK-K---  169
SMM.US.x.F236                 -----N--V-----------.-------H---N-----MA-----H---------------RDET--------N-A-----------------R--------D---T----------SE----K-----K-----K--K---N----------T---..H------D--  167
SMM.US.x.H9.M80194            -----N--V-------G---.K------H---N------A-----H---------------RDEA--------N----X-------------R-----A--D---T---G------SE---------K-----K--K---N----------T---..H-X----X--  167
SMM.US.x.PBJ14                -----N--V-------G---.K------H---N------A-----H---------------KDEA--------N-------------------R--------D---T---G------SE---------K-----K--K---N----------T---..H---------  167
SMM.US.x.PBJA.M31325          -----N--V-------G---.K------H---N------A-----H---------------KDEA--------N-------------------R--------D---T---G------SE---------K-----K--K---N----------T---..H---------  167
SMM.US.x.PBJ                  -----N--V-------G---.K------H---N------A-----H---------------RDEA--------N----X-------------R-----A--D---T---G------SE---------K-----K--K---N----------T---..H-X----X--  167
SMM.US.x.PBJ                  -----N--V-------G---.K------H---N------A-----H---------------KDEA--------N-------------------R--------D---T---G------SE---------K-----K--K---N----------T---..H---------  167
SMM.US.x.PGM53.AF077017       -----N--V----K--G---.K---------FR------A-----H---------------RDE---------N-A-----------------RK-------D---T---G------SE---------K-----K--K---N----------T---..H----R-D--  167
SMM.US.x.SME543.U72748        -----N--V-----------.-------H---N-----MA-----H---------------RDKT--------N----------H--------R--------DC--T-----------SE---Q------K-----K--K---N----------T---..H----R-D--  167
STM.US.x.STM.M83293           --------V-------G---.-------H---N--E-S-A-----H----------------GEA--------N---------E--------TR---S----DC--Q---G-----------------K--------TK---N-------------E..H----R--TA  167
```

```
                                                Vif end
MAC.US.x.239                WKQWRRDNRRGLRMAKQNSRGDKQRGGKPPTKGANFPGLAKVLGILA*   214
H2A.IN.NNVA.NEW             R--R---Y-----L-REDG-SH----SES-AQ--H---V---------   216
H2A.GM.x.MCR35.AY509260     R------Y----QL-R-DD-SH----SES-APR-Y---V-E--E----   216
H2A.IN.95.CRIK              R--R---Y-----L-R-DG-SH-----ES-AQ-TH-----E--------   216
H2A.CI.88.UC2.U38293        R------HW----V-R-DY-SLE-------APR-H---V---E----   216
H2A.DE.x.BEN.M30502         R------HW----V-REDH-SL--G-SE-SAPR-H---V----E----   216
H2A.DE.x.PEI2.U22047        R-R--SNYW--F-L-RKDG--H----SE--AS--Y---V----E----   216
H2A.GH.x.GH1.M30895         R------HW----V-R-DY-SL----SE-SAPR-H---V----E----   216
H2A.GM.87.D194.J04542       R------HW----V-R-DY-SL--G-SE-SAPR-H---V---------   216
H2A.GM.x.ISY.J04498         R-----NY-----L-R-DG-SH----SE--AQ--Y---V----E----   216
H2A.GM.x.MCN13.AY509259     R------Y----QL-R-DG-SH----SES-APR-Y---V-E--E----   216
H2A.GW.86.FG.J03654         R--R-G-Y-------R-D---Y----SES-PTR-H-----E--E----   216
H2A.GW.87.CAM2CG.D00835     R------Y-----L-R-D--SY----SES-AP--Y---V----E----   216
H2A.GW.x.ALI.AF082339       R------Y-----V-R-D---L-----ES-AP--H---V----E----   216
H2A.GW.x.MDS.Z48731         R--R---Y-A---L-R-D--SH-----ES-APR-Y---V----E----   216
H2A.SN.85.ROD.M15390        R--R---Y-----L---D--SH---SSES--PRTY---V-E--E----   216
H2B.CI.88.UC1.L07625        R--R--N---SI-L-RK-NNRAQ-GSSQ-LAPRTH-----E-------   217
H2B.CI.x.EHO.U27200         R--R--NS---I---RD-I-TSQ-SSSQSLAQ-TY-----E-------   217
H2B.GH.86.D205.X61240       R--R--NS--SI-L-RK-NNRAQ-GS-Q-FAPRTY-----E-------   217
H2B.JP.01.KR020.AB100245    R--R--N---SI---RK-NNRAQ-SSCQSSAQ-TY-----E------R   217
H2G.CI.92.ABT96.AF208027    R--R-G-----I---RK--XR-Q-DSSQSF-Q--Y-------------   216
H2AB.CI.90.7312A.L36874     R--R---S---I---RK-NSRTQ-GSSQ-FAPRTY-------------   217
H2AB.CM.03CM_510_03.NEW     RE-R--N---SI-X-RK-NNRAQ-GSSQ-LAPRTH-------------   217
H2U.FR.96.12034.AY530889    R-FR--NC---FQL-RK-CERYQ-GSSATSSQ-TY--------A----   214
MAC.US.x.MAC239             ----------------------------------------------   215
MAC.US.x.251                ----------------------------------------------   215
MAC.US.x.251_32H            -------S--S-----------------------------------   215
MAC.US.x.251                -------------------------S-------D-------------   215
MAC.US.x.17EC1.AY033233     ----------------------------------------------   215
MAC.US.x.17EFR.AY033146     ----------------------------------------------   215
MAC.US.x.1937.AY611495      ----------------------------X-R---------------   215
MAC.US.x.2065.AY611493      -------------------------------V-------------   215
MAC.US.x.80035.AY611486     ---------------------X------------------------   215
MAC.US.x.81035.AY599200     ---------------------X------------------------   215
MAC.US.x.85013.AY611490     -------------------------------D-------------   215
MAC.US.x.92050.AY603959     ----------------------------------------------   215
MAC.US.x.92077.AY599201     ----------------------------------------------   215
MAC.US.x.93057.AY611492     ----------------------------------------------   215
MAC.US.x.93062.AY607704     ----------------------------------------------   215
MAC.US.x.95058.AY611494     -------------------------------D-------------   215
MAC.US.x.95086.AY607703     ----------------------------------------------   215
MAC.US.x.95112.AY588946     ----------------------------------------------   215
MAC.US.x.96016.AY607701     ----------------------------------------------   215
MAC.US.x.96020.AY611488     ----------------------------------------------   215
MAC.US.x.96072.AY611491     ------------------------------R---------------   215
MAC.US.x.96081.AY597209     ----------------------------------------------   215
MAC.US.x.96093.AY611489     ----------------------------------------------   215
MAC.US.x.96114.AY588945     ----------------------------------------------   215
MAC.US.x.96123.AY611487     ----------------------------------------------   215
MAC.US.x.96135.AY607702     ----------------------------------------------   215
MAC.US.x.97009.AY599199     ----------------------------------------------   215
MAC.US.x.97074.AY599198     ------------------------------R-D-------------   215
MAC.US.x.MM142.Y00277       ----------S--V-----------------E---------------   215
MAC.US.x.r80025.AY576480    ----------------------------------------------   215
MAC.US.x.r90131.AY576481    ----------------------------------------------   215
MAC.US.x.SMM142B.BD131285---------S--V-----------------E---------------   215
MNE.US.82.MNE               ----------S--------------S------VD-------------   215
MNE.US.x.MNE027.U79412      ----------S--------------S-------D-------------   215
SMM.SL.92.SL92B.AF334679    R-----N-------VT-Y-G-PESNSSTT---RV-----E------GR   217
SMM.US.x.F236               ------N--K---------RN--GSS-S-AE----------------   215
SMM.US.x.H9.M80194          ------NX---X-L-T--X-RN--GSSESFAX---------X------   215
SMM.US.x.PBJ14              ------N------L-R----RN--GSSESFAE-T------------V-   215
SMM.US.x.PBJA.M31325        ------N------L-R----RN--GSSESFAE-T--------------   215
SMM.US.x.PBJ                ------NX---X-L-T--X-RN--GSSESFAX---------X------   215
SMM.US.x.PBJ                ------N------L-R----RN--GSSESFAE-T--------------   215
SMM.US.x.PGM53.AF077017     ------N----I--------RN--GSSESLAE----------------   215
SMM.US.x.SME543.U72748      -----GN-------------RN--GSS-S-AE----------------   215
STM.US.x.STM.M83293         R-----G--GSI-V-T--G--H-P--S--S-E-TD------------   215
```

```
                          Vpx start                                                                                        Vpx end
MAC.US.x.239              MSDPRERIPPGNSGEETIGEAFEWLNRTVEEINREAVNHLPRELIFQVWQRSWEYWHDEQGMSPSYVKYRYLCLIQKALFMHCKKGCRCLGEGHGAGGWR.PGPPPPPPPGLA*   112
H2A.IN.NNVA.NEW           -T----TV--------------A--E----A---------------------G---------A--T-------M---M-I-ARR--T---G---G----.S--L------WV-   113
H2A.GM.x.MCR35.AY509260   -T----TV----------E---N--D--I-A---------------------R---------Q--T-----------M-T---R--T---G---P----.-----------V-   113
H2A.IN.95.CRIK            -T----TV--------------A--D----A---------------------G---------A--T-------M---I-I-VRR--T---G---P----.S----------V-   113
H2A.CI.88.UC2.U38293      -T-----V-----------------D--I-AL--------------------R-------L-S--T-------M---M-I-L----T---G---P----.-----------V-   113
H2A.DE.x.BEN.M30502       -T-----V-----------------E--I-AL--------------------R---------A--T-------M---I-T-F-R--T-W--DM-RE-LEDQ----------V-   114
H2A.DE.x.PEI2.U22047      -A---KTV----G--V------A--E-M--A---------------------R----DL---Q--T-----R-M-Y-M-I-V----T---G---P----.--------.--V-   112
H2A.GH.x.GH1.M30895       -T-----V-----------------D--I-AL--------------------R----D-------T-------M---V-I-F-R--T---G---P----.S----------V-   113
H2A.GM.87.D194.J04542     -A-----V-----------------D--I-AL--------------------A---------T--T------IM---VYI-F----T---R---P----.-----------V-   113
H2A.GM.x.ISY.J04498       -TN---T-----------E---D--D----A---------------------R---------R--T-------M---V---F----T-R-----P----.S----------V-   113
H2A.GM.x.MCN13.AY509259   -T----TV----------E---N--D----A---------------------R---------Q--T-----------M-T---R--T---G---P----.-----------V-   113
H2A.GW.86.FG.J03654       -T----TV----------E---A--D----A--------H------------R---------T--T---------M-MY--A-RDGT---G-M-QK-GD.Q----------V-   113
H2A.GW.87.CAM2CG.D00835   -T----TV----------E---A--D----A---------------------R---------Q--T-------M---V-I-F-R--T---G---P----.-----------V-   113
H2A.GW.x.ALI.AF082339     -AN---TV-----------------D----AL--------------------R---------Q--T-------M---MYT-FM---T---G---P----.S----------V-   113
H2A.GW.x.MDS.Z48731       -A----TV---------EQ--A--D----A---------------------R---------T--T-------M---MYI-GR---T-R-----P----.-----------V-   113
H2A.SN.85.ROD.M15390      -T----TV--------------A-------A---------------------R---------E--T------I----VY--VR---T---R---P----.-----------V-   113
H2B.CI.88.UC1.L07625      KM-----V-----D--------D--E--IT-L--V----------------C-A--RE-----S--T-----L-M---M-V-YT------Q----P----.S------------   113
H2B.CI.x.EHO.U27200       KM-----V---------V-------ET-L-HL--V---------------K--A--RE-----I--T-------M---M-I-FA---G--R----P----.S------------   113
H2B.GH.86.D205.X61240     KM-----V-----D---V----A--E--IT-L--V------------------A--RE-----I--T-----L-M---M-V-YT------Q----P----.S------------   113
H2B.JP.01.KR020.AB100245  EM-----V-----------------E--IT-L--V------------------A---E----TI--T-------M----Y--LAR--A--R----P----.-------------   113
H2G.CI.92.ABT96.AF208027  DM-----X--E--D-----------E--XT----I--------------R-------------A--T-----L-M---M-V------T--QK---P----.Q------------   113
H2U.FR.96.12034.AY530889  -G-----------E---V----S--E---RD---A--Q---------------Q--R------E--T-------------------Q-------P----.S----------I-   113
H2AB.CM.03CM_510_03.NEW   KM-----------D--------X--E--IT-L--V----------------C-A--R------S--T-----L-M---M-V-YT---T-RQ----X----.S------------   113
H2AB.CI.90.7312A.L36874   KM-----V-----D---V----A--E--IV-L---------------------A--R-D----I--T-----L-----M-V-FA---T--RG---P----.R------------   113
MAC.US.x.MAC239           ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.251              ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.251_32H          -----------------V-----------------------------------------------Q----------------------------------.-------------   113
MAC.US.x.251              -----------------------------------------------------------------Q---------M------------------------.-------------   113
MAC.US.x.17EC1.AY033233   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.17EFR.AY033146   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.1937.AY611495    ----------------------------------------X-----------------------------------------------------------.-------------   113
MAC.US.x.2065.AY611493    ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.80035.AY611486   --------------------------------X-------------------------------------------------------------------.-------------   113
MAC.US.x.81035.AY599200   ---------------------------------X------------------------------X-----------------------------------.-------------   113
MAC.US.x.85013.AY611490   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.92050.AY603959   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.92077.AY599201   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.93057.AY611492   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.93062.AY607704   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.95058.AY611494   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.95086.AY607703   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.95112.AY588946   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.96016.AY607701   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.96020.AY611488   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.96072.AY611491   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.96081.AY597209   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.96093.AY611489   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.96114.AY588945   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.96123.AY611487   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.96135.AY607702   -X--------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.97009.AY599199   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.97074.AY599198   ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.MM142.Y00277     -----------------------------------------------------------------Q--T-------------------------------.-------------   113
MAC.US.x.r80025.AY576480  ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.r90131.AY576481  ----------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.SMM142B.BD131285-----------------------------------------------------------------Q--T-------------------------------.-------------   113
MNE.US.82.MNE             -----------------------------------------------------------------Q----------------------------------.-------------   113
MNE.US.x.MNE027.U79412    ----------------------------------------------------------------------------------------------------.-------------   113
SMM.SL.92.SL92B.AF334679  -T----------------------HN---AL-QT--Q-----------R-C----V----Y----A----VQ-M---M-Q-FR---T-R----SQ----.T------------   113
SMM.US.x.F236             -------------------K--------------A---------------R---------M---E--T-------------V----------E-------.T------------   113
SMM.US.x.H9.M80194        -XX---X---------XX--X-D--H--X-----A-----X---------R-X-------M---V--T-X----------------------GE-------.-------------   113
SMM.US.x.PBJ14            ----------------------D--D--------A---------------R------Y--M---V--T-----------M-----------GE-------.-------------   113
SMM.US.x.PBJA.M31325      ----------------------D--D--------A---------------R---------M---V--T-----------M-----------GE-------.-------------   113
SMM.US.x.PBJ              -XX---X---------XX--X-D--H--X-----A-----X---------R-X-------M---V--T-X----------------------GE-------.-------------   113
SMM.US.x.PBJ              ----------------------D--D--------A---------------R---------M---V--T-----------M-----------GE-------.-------------   113
SMM.US.x.PGM53.AF077017   ----------------------------------A---------------R------R--M---E--T-------------V---R-------E-------.S------------   113
SMM.US.x.SME543.U72748    ------------------E----------G---A-------------------------M---E--T-----------------------------------.T------------   113
STM.US.x.STM.M83293       ----------------A-E------H----D------------------R------------PG-----------------S-R----------------.-------------   113
```

```
                          Vpr start                                                                                    Vpr end
MAC.US.x.239              MEERPP.ENEGPQREPWDEWVVEVLEELKEEALKHFDPRLLTALGNHIYNRHGDTLEGAGELIRILQRALFMHFRGGCIHSRIGQPGGGNPLSAIPPSRSML*   101
H2A.IN.NNVA.NEW           PT-L--.-DRT-P---G----I-I-R-I-----R-----------RY--T-Y---P---R----V------T---A--G-------R---------TP---R-   102
H2A.GM.x.MCR35.AY509260   PT-F--.-DGT-P---G---II-I-RKI-K---------------Y-HT---------R---NV---------A--RL-----T--RT-FP-TSTP-T-Q-   102
H2A.IN.95.CRIK            PT-L--.-DGT-P---G----I-I-R-I-----R-----------IY--D-Y------R----V------T---A--G-------R------T--TP---R-   102
H2A.CI.88.UC2.U38293      ................#-GRS*-MIV*I----*R-------I--D-Y------N-----R---K------S--L-V-GSR----*TRRRT-CP-T-TP-G-H-    82
H2A.DE.x.BEN.M30502       PT-F--.-DGT-R-DLGSD--I-T-R-I-----R-------I---YY-H---------R---KT------V---A--NR----*TRRRT-CP-A-TP-A-H-   101
H2A.DE.x.PEI2.U22047      PA-F--.-D-T-P-G-G----IGI-R--R------------T---Y-CA-------S-R---NV------V---A--KI-----TR-ET-F----TP-G-Q-   102
H2A.GH.x.GH1.M30895       PT-F--.-DGT-R--LGGD--IRI-G-I------------I----Y-HS-----P---R-----------V-L-A--NR---S-TRRRT-FP-A-TP-G-Y-   102
H2A.GM.87.D194.J04542     PT-F--.-DGT-R--LGST--I-T-K-I---------C--I----Y------------R----V------V-I-A--DR--K--TRRRA-CP-A-TP-G-H-   102
H2A.GM.x.ISY.J04498       PA-F--.-DGT-P---G----I-I-R-I----------------YY--T---------R----V------T---A--G-------R-R-------TP-N-Q-   102
H2A.GM.x.MCN13.AY509259   PT-F--.-DGT-P---G---II-I-RKI-K---------------Y-HT---------R---NV---------A--RL-----T--RT-FP-TSTP-T-Q-   102
H2A.GW.86.FG.J03654       PT-L--.-DRT-P---G-A--I-I-R-IE----R-----------RY--T---------R-----------A---A--G------TR---------TP-G-H-   102
H2A.GW.87.CAM2CG.D00835   PT-L--.-DGT-P---G----I-I-RDI----------------G---A-------R-R----V------T---A--N------TR---------TP-R-*-   101
H2A.GW.x.ALI.AF082339     PT-F--#R-GT-f---Gf---I---R-I-----R-----M-I---GY--T-------R-R---NA----------A--GR--V--TR-R-------TP-N-Q-   100
H2A.GW.x.MDS.Z48731       PT-L--.-DGT-P---G----I-I-R-IT--------RS------R---T---------R------------T---A--G------TR---------TP---Q-   102
H2A.SN.85.ROD.M15390      PT-L--.VDGT-L---G---II-I-R-I------------I---KY--T---------R---KV------T---A--G------TR---------TP-N-Q-   102
H2B.CI.88.UC1.L07625      AP-T--.---S------E---ED-M--I-Q---R-----------F--S------A------K-------L---A--Q------------------*--R-   101
H2B.CI.x.EHO.U27200       VP-I--.-DKN------EQ---D----I-Q---------------F------N---------KL------L------Q------------------*G-Q-   101
H2B.GH.86.D205.X61240     AP-I--.---N------E--IG-I---I-Q---------------F--S-----A------K-------L---A--Q------S-------T---P*G-R-   101
H2B.JP.01.KR020.AB100245  AP-T--.---D------E---ED------Q---R-----------F--G-----IA--------------L------Q---------------V----*--QK   101
H2G.CI.92.ABT96.AF208027  I...--.-D-A---------------I-----------------Y--D-------------K-------L----D-RC----GNA------T-----GVF-    99
H2U.FR.96.12034.AY530889  #A-I--.-DGA-----G----R-I--------VR-----------Y---------------K-----I-L------R-----SSR-----TTV---*GV--   100
H2AB.CI.90.7312A.L36874   AS-T--.---D------E---ED----I-Q---R-----------F--S------A------K-------L------Q---------------T---EG-Q-   102
H2AB.CM.03CM_510_03.NEW   XXXI--.---S----Q-E-RXG-I---I-Q--X--------X----F--S-Y----A-----VK-------L---A--R--------------X---P*G-Q-   101
MAC.US.x.MAC239           ------.------------------------------------------------------------------------------------------------   102
MAC.US.x.251              ------.-----*----G------------G---------------------------------------------------------R--------------   101
MAC.US.x.251_32H          ------.-----------------------------------------------------------------------N--------------T-----GV--   102
MAC.US.x.251              ------.-----------------------------------------------------------------------N--------------T----*GV--   101
MAC.US.x.17EC1.AY033233   ------.------------------------------------------------------------------------------------------------   102
MAC.US.x.17EFR.AY033146   ------.------------------------------------------------------------------------------------------------   102
MAC.US.x.1937.AY611495    ------.-----------------X--X--------------------------------------------------T----------------------X-   102
MAC.US.x.2065.AY611493    ------.-----------------------------------------------------------------------T------------------------   102
MAC.US.x.80035.AY611486   ------.-----------------------------------------------------------------------T------------------------   102
MAC.US.x.81035.AY599200   ------.-------------------------------------X----------------------------------------------------------   102
MAC.US.x.85013.AY611490   ------.-----------------------------------------------------------------------T------------------------   102
MAC.US.x.92050.AY603959   ------.------------------------------------------------------------------------------------------------   102
MAC.US.x.92077.AY599201   ------.------------------------------------------------------------------------X-----------------------   102
MAC.US.x.93057.AY611492   ------.-----------------------------------------------------------------------T---X--------------------   102
MAC.US.x.93062.AY607704   ------.----X------------------------------------------------------------------M----------------X-------   102
MAC.US.x.95058.AY611494   ------.----------------------------------------------------------------------XT--X---------------------   102
MAC.US.x.95086.AY607703   ------.-----------------------------------------------------------------------T------------------------   102
MAC.US.x.95112.AY588946   ------.------------------------------------------------------------------------------------------------   102
MAC.US.x.96016.AY607701   ------.-----------------------------------------------------------------------X---X--------------------   102
MAC.US.x.96020.AY611488   ------.-----------------------------------------------------------------------T------------------------   102
MAC.US.x.96072.AY611491   ------.-----------------------------------------------------------X-----------F------------------------   102
MAC.US.x.96081.AY597209   ------.------------------------------------------------------------------------------------------------   102
MAC.US.x.96093.AY611489   ------.------------------------------------------------------------------------------S------------X----   102
MAC.US.x.96114.AY588945   ------.------------------------------------------------------------------------------------------------   102
MAC.US.x.96123.AY611487   ------.-----------------------------------------------------------------------T-----XX-----------------   102
MAC.US.x.96135.AY607702   ------.-----------------------------------------------------------------------T------------------------   102
MAC.US.x.97009.AY599199   ------.------------------------------------------------------------------------------------------------   102
MAC.US.x.97074.AY599198   ------.-----------------------------------------------------------------------T------------------------   102
MAC.US.x.MM142.Y00277     ------.------------------K---------------------------------------------I---S--S--------------T---------   102
MAC.US.x.r80025.AY576480  ------.-----------------------------------------------------------------------T------------------------   102
MAC.US.x.r90131.AY576481  ------.------------------------------------------------------------------------------------------------   102
MAC.US.x.SMM142B.BD131285 ------.------------------K---------------------------------------------I---S--S--------------T---------   102
MNE.US.82.MNE             ------.-D------------------------------------------------------K--------------T------S-------T-----R---   102
MNE.US.x.MNE027.U79412    ------.-D-------------------------------------Y--D-------------K--------------N------S-------T---------   102
SMM.SL.92.SL92B.AF334679  #-H.A-.-D-TNP-------IRD-----------------------YV-DTY---I-----I-K---K---L---H--T-------R-P---GS--SA-DV--   100
SMM.US.x.F236             -A----.-D-A-----------------------------------Y--D---------------------I---S--A------S............-GV--    90
SMM.US.x.H9.M80194        -T----.-D-A---------X------X-----NX------X----Y--D-----------------G---I------R------S--X----T---X-GV--   102
SMM.US.x.PBJ14            -T----.-D-A----------------I-----N------------Y--D---------------------I------R------S-------T-----GV--   102
SMM.US.x.PBJA.M31325      -T----.-D-A----------------I-----N------------Y--D---------------------I------R------S-------T-----GV--   102
SMM.US.x.PBJ              -T----.-D-A---------X------R-----NX------X----Y--D---------------------I------R------S--X----T---X-GV--   102
SMM.US.x.PBJ              -T----.-D-A----------------I-----N------------Y--D---------------------I------R------S-------T-----GV--   102
SMM.US.x.PGM53.AF077017   -T----.-D-A-----------K----V------------------Y--D---------------------I------A------SR------T-----GV--   102
SMM.US.x.SME543.U72748    -A----.-D-A----------------I------------------Y--D---------------------I---S--A------SR------T-----A---   102
STM.US.x.STM.M83293       -TH---.-D------------------I-Q---R-------S----Y--D-------------K--------------R-------------AT---T-GV--   102
```

```
                                Tat start                                                
```

```
                          Rev start          exon 1 end exon 2 start                                                        Rev end
MAC.US.x.239              MSNHEREEELRKRLRLIHLLHQTNPYPTGPGTANQRRQRKRRWRRRWQQLLALADRIYSFPDPPTDTPLDLAIQQLQNLAIESIPDPPTNTPEALCDPTEDSRSPQD*  107
H2A.IN.NNVA.NEW           -TERAG--GPERK----R---------Q-----R---N-R----Q--I-I-----S--T-----A-S-V-QT-----G-T-QEL-----HP--S.......Q-LAEA-  101
H2A.GM.x.MCR35.AY509260   -TGRAD-----RK-------------Q-L---S---N-R---KQ--R--V----K--T-----A-S---RT--H--E-T-QDL-----SF--SSEGTDSNQ-LAKN-  108
H2A.IN.95.CRIK            -NERAD--G-QRK----R---------Q-----S---N-R----Q--L-I-T---S--T-----ANS-V-QT-----G-T-QEL-----HP--S.......Q-LAET-  101
H2A.CI.88.UC2.U38293      -CEKAD----QRE----R---------YR----G---N-R--R--Q-LR------KLHTA----A-SS--W---H--G-T-REL-----DL--S....DSNQGLAET-  104
H2A.DE.x.BEN.M30502       --ERAD--G-QGK---LR---------Q-----S---N-R--R--Q-LR-V---NKLCAV------S---R---H--R-T-QEL-----DL--S....NSNQGLAET-  104
H2A.DE.x.PEI2.U22047      -NGRAD--G-QRKQ---R---------Q-L---R---N-R--RKQH-R--V---NS--T-----A-S---R---R--G-T-QEL------L--SSESTNNNQGLAETY  108
H2A.GH.x.GH1.M30895       -HEKADG---QE-----R---------H-----S---N-R--Q----LR-V----KL-T-------S---R---D--R-T-HEL-----DL--S....NSNQGLAET-  104
H2A.GM.87.D194.J04542     -RDRAD--G-QEK----R----I----H-Q---S---N-R--R--Q-FR-V---TKLHTI------S---R------G-T-QEL-----DL--S....NSNQGLAET-  104
H2A.GM.x.ISY.J04498       -TERAD--GV-RK----R---------Q-L---R---N-R---E---K-I--------T-----A-P---QT-----G-T-QTL-----...........TQ-LAETQ  97
H2A.GM.x.MCN13.AY509259   -TGRAD-----RK-------------Q-L---S---N-R---KQ--R--V----K--T-----A-S--GRT--H--E-T-QDL-----SF--SSEGTDSNQ-LAKN-  108
H2A.GW.86.FG.J03654       -TERAD--G-QRK----R---------Q-----S---N-R---KQ--R-I-----S--T-----A-S---R---H--G-T-QDL------L--SPESTNSNQ-LAEA-  108
H2A.GW.87.CAM2CG.D00835   -TERAD--G-QRK----R---------Q-----S---N-R----Q--R-I--------T-----A-P---RT--H--E-T-QDL-----HP--S.......Q-LAEA-  101
H2A.GW.x.ALI.AF082339     -TERAG--D-QRK----R---------Q---------N-R----Q--G-IV------FT-----ASS---R-V-H--G-T-QDL-----DL--SSESADNNQGLAET-  108
H2A.GW.x.MDS.Z48731       -NERAD--G-QRK----R---------Q-----R---N-R----Q--R-I-----S--T-----A-S---QTV----G-T-QDL-----HL--S.......Q-PAAT-  101
H2A.SN.85.ROD.M15390      -NERAD--G-QRK----R---------Q-----S---N-R---KQ--R-I-----S--T-----A-S---QT--H--G-T-QEL-----HL--S.......Q-LAET-  101

H2B.CI.88.UC1.L07625      -TTR-KD..-Q-G---L----------QT----S---N-R---K--GL-I-------R-LS-S--EE-----V-R--E-TV-DL-N---S-...........PTA-AF  95
H2B.CI.x.EHO.U27200       -NAR--D..-Q-G---L----------Q-----S---N-R---KQ-GL-I-------HPL--S--EG-------R----I-KDL-N---S-...........PTA-AS  95
H2B.GH.86.D205.X61240     -TAR-GD..-Q-E---L----------Q-----S---N-R---K--GL-I-------H--STA-AEE--N----R----TV-DL-N--P-LNQS........PTT-AP  98
H2B.JP.01.KR020.AB100245  -TTR-KD..-Q-G---LR---------Q-----S---N-R------GL-I-----Q-HPL--S--EE--N-------K-TV-DL-T---SI...........PTA-TS  95

G.CI.92.ABT96.AF208027    ---L-E--.---------F--------Q---X-----N-R-K-KQ--L-I-----------X--ARE----------S-T-QDX--S--AV-KN-Q--PSN*-TSPAR  106

H2U.FR.96.12034.AY530889  ---P-EG..VQR------------------------N-R----Q--L-I-----K--T-------S-V--------G---QDL-E--AVV--H.......Q-TSAP-  99

H2AB.CI.90.7312A.L36874   -TAR-KD..-Q-GI--L----------Q-----S---N-R---K--GL-I-------HPL-NS-AEE------RR----TV-DL-N---SS...........PTT-AP  95
H2AB.CM.03CM_510_03.NEW   -TTR---....-E---L----------H---------N-R---K--GL-I-------R--SAA-AEE--N----R----TV-DL-N------...........TT-AP  93

MAC.US.x.MAC239           ---------------------------------------R---------------------------------------X--------------------------  108
MAC.US.x.251              --S------------------------------------R---------------------------------------------------------K--------  108
MAC.US.x.251_32H          --S------------------------------------R------------------------------------------------------G---N-------  108
MAC.US.x.251              --S---------------------S--------------R---------------------------------------------------------KG-------  108
MAC.US.x.17EC1.AY033233   -N--------------------------------------------------------------------------------------------------------  108
MAC.US.x.17EFR.AY033146   -N-------------------------------------R------------T-----------------------------------------------------  108
MAC.US.x.1937.AY611495    ----XX-----R---------------------------R------------------------------------X------------I---X------------  108
MAC.US.x.2065.AY611493    ---------------------------------------R---------X---------------------------------------I---------X------  108
MAC.US.x.80035.AY611486   ---------------------------------------R------------------------------------------------------------------  108
MAC.US.x.81035.AY599200   ---------------------------------------R-----------------------------------------------------------X------  108
MAC.US.x.85013.AY611490   --X------------------------------------R------------------------------------------------------------------  108
MAC.US.x.92050.AY603959   ---------------------------------------R-------------------------------------------------I----------------  108
MAC.US.x.92077.AY599201   ---------------------------------------R------------------------------------------------------------------  108
MAC.US.x.93057.AY611492   --------------------------X------------R------------------------------------------------------------------  108
MAC.US.x.93062.AY607704   ---------------------------------------R-------------------------------V----------------------------------  108
MAC.US.x.95058.AY611494   ---------------------------S-----------R------------------------------------------------------------------  108
MAC.US.x.95086.AY607703   ---------------------------------------R------------------------------------------------------------------  108
MAC.US.x.95112.AY588946   ---------------------------------------X------------------------------------------------------------------  108
MAC.US.x.96016.AY607701   ---------------------------------------R------------------------------------------------------------------  108
MAC.US.x.96020.AY611488   ---------------------------------------X------------------------------------------------------------------  108
MAC.US.x.96072.AY611491   ---X-----------------------------------R-----------------------------------------------------X---X--------  108
MAC.US.x.96081.AY597209   ---------------------------------------R------------------------------------------------------------------  108
MAC.US.x.96093.AY611489   ---------------------------------------R------------------------------------------------------------------  108
MAC.US.x.96114.AY588945   ---------------------------------------R-------------------------------X----------------------------------  108
MAC.US.x.96123.AY611487   ---------------------------------------R-------------------------------V----------------------------------  108
MAC.US.x.96135.AY607702   ---------------------------------------R------------------------------------------------------------------  108
MAC.US.x.97009.AY599199   ---------------------------------------R--------------------X---------------------------------------------  108
MAC.US.x.97074.AY599198   ------------------------------X--------R------------------------------------------------------------------  108
MAC.US.x.MM142.Y00277     -RS-TG-----R-----------------S-S------KR----Q--------------------------------------------I--------N-----A-  108
MAC.US.x.r80025.AY576480  ---------------------------------------R------------------------------------------------------------------  108
MAC.US.x.r90131.AY576481  ---------------------------------------R------------------------------------------------------------------  108
MAC.US.x.SMM142B.BD131285-RS-TG-----R-----------------S-S------KR----Q--------------------------------------------I------LRRIR---A-  108

MNE.US.82.MNE             --S-AE-----R-------------------------N-R----Q-------------------N-----V------------------I--I-H---NP------  108
MNE.US.x.MNE027.U79412    --S-AE-----R------------------S------N-R----Q----F------V--------------------------------I--I-H---SP------  108

SMM.SL.92.SL92B.AF334679  -Q-P-E-..--R-------I-------A-----S---N-R---K---I-IV------------A--DI------FDS-S-Q-L-E---TV--T-R-QSA-*-AC-KR  105
SMM.US.x.F236             --ST-E-..--R------F--------Q---------R-R----Q----I-------------V------------G----EL-----SA--P-N-VAKSP*-H-N-  105
SMM.US.x.H9.M80194        --SN-E-..--R------F-------XX---------R-R---XQ----I--X-----------A------------G----XL-N--ASA--P-K-AA-SP*-H-S-  105
SMM.US.x.PBJ14            T-SNGE-..--R-----------S---D---------R-R----Q----I-------------V------------R----EL-N--ASA--P-K-AA-SP*-H-SS  105
SMM.US.x.PBJA.M31325      --SN-E-..--R---------------D---------R-R----Q----I-------------V------------R----EL-N--ASA--P-K-AA-SP*-H-SS  105
SMM.US.x.PBJ              --SN-E-..--R-------------XX---------R-R---XQ----I--X-----------A------------G----XL-N--ASA--P-K-AA-SP*-H-S-  105
SMM.US.x.PBJ              --SN-E-..--R---------------D---------R-R----Q----I-------------V------------R----EL-N--ASA--P-K-AA-SP*-H-SS  105
SMM.US.x.PGM53.AF077017   --SSAG-----R------F--------Q---------R-R----Q----I-------------A-S-----V----D-SL-N------SV---...L--NP--R---  105
SMM.US.x.SME543.U72748    --S.TE--.---------F--------Q---------R-R----Q----I-----------N--------------G----DL-----SA--T-K-AAKS-*-H-N-  105

STM.US.x.STM.M83293       --DQ-E-..--------QF--------Q---------N-R----Q--N------N---------AS------V----G-S-QDL----P-L-KD-Q-TA-N*--R-TR  105
```

```
                                 signal peptide end gp120 start
                             Env start
MAC.US.x.239                 MGCLGNQLLIAILLLSVYGIYCT..LYVTVFYGVPAWRNATIPLFCATKNRDTWGTTQCLPDNGDYSEVALNVTESFDAWNNTVTEQAIEDVWQLFETSIKPCVKLSPLCITMRCNKSETDRWGLTKSITTTASTTSTTASA.......KVDMVNETSSCIAQDNC.TGL  160
H2A.IN.NNVA.NEW              -TRERA---VV---T-TCL---KQ.Q-------I---K--S---------------------D--Q-I------A----D-------V----N-----T------T---VA-E-KNR--GNATT-APT--PEINET..............TEIS-NT--VRG---.S--  154
H2A.GM.x.MCR35.AY509260      -MGGR----V----A-TCL-----..N-------I------S---------------I------DE-Q-IT-----A----D------------N------------T---VA----NTDARNTTTPTTASPRT.............IKP-TEIS-N-T--RAN--.S--  153
H2A.IN.95.CRIK               -AHE-T-----F--T-ACL---K..Q-------I---K--S-------R-------I------D--Q-IV---A-A----D------------N-----T------T---V-----T-T-TTTTAPT-TSAGSTA-P............-PI----NT--MHAN--.S--  155
H2A.CI.88.UC2.U38293         -EPGR----AV---T-ACL---K..Q---------V----S---------------I------D--Q-IP-----A----D------------R------------T---VA-N--PVTGNNTNA-AKP-AARP--NP..........SYLTII--S-T-VGA---.---  157
H2A.DE.x.BEN.M30502          -EPGR---FVV---T-ACLV--S..Q-------I---K--S---------------I------D--Q-II-----A------------V----H------------T---VA-N-SRVQGNTTTPNPRTSSSTTSRPP-.........SAASII----N--ENNT-.A--  158
H2A.DE.x.PEI2.U22047         IMDSR---IV----T-ACL---A..Q-------I---K--S-------R-------I------D--Q-IP-----A------------V----N-----V------T---VQ-E--STS-ESSNS-SEGS-VP.................EIL---T---TNNS-.SD-  149
H2A.GH.x.GH1.M30895          KM-GKSL-CV-S--A-A-LV---..Q---------V----S---------------I--K---D--Q-IT-----A----D-------V----S------------T---VA-S--STTNNTTTTGSTTGM...................SEI--.T-PSYS---.---  146
H2A.GM.87.D194.J04542        -EPGR----V----T-ACL---K..Q-------I------S---------------I------D--Q-IT-----A----D------------R------------T---VA-N--ITSGTTATPSPP....................NITIID-N-T--GDN--.---  146
H2A.GM.x.ISY.J04498          TMSGKI---V-F--T-ACL-----..K---------V-K--S---------------I------D--Q-IP-----A----D-I-----V----N------------T---V--N--A-TESAVAT-SPSGP......................--I-D-DP--QLN--.S--  147
H2A.GM.x.MCN13.AY509259      -MGGR----V----T-TCL-----..N-------I------S---------------I------D--Q-IT-----A----D------------N------------T---VA----NTDARNTTTPTTASPRT.............IKP-TEIS-N----RAN--.S--  153
H2A.GW.86.FG.J03654          -KGSK-------V-A-A-L-H-K..QF------I------S---------------I------D--Q-IT-----A------------V----N------------T---VA-N-TRNM-TWT-R-DTQ....................NITIINDT-HARA---.---  146
H2A.GW.87.CAM2CG.D00835      -ERGR---------A-ACL---R.QQ-----------K--S---------------I------D--Q-IP-----A----D--I---------N------------T---VA-K--I-TS-TTMIRTTTPS--.................-EAPISDN-P--RTN--.S--  152
H2A.GW.x.ALI.AF082339        -MSSR----VT---A-ACLV--K..Q-----------K--S---------------I------D--Q-I------A----D-------V----R------------T----A-K-SNIS-ESTTTSP-PGS.................TLKPLI--SDP--KA---PR--  151
H2A.GW.x.MDS.Z48731          -TRKMH--------T-ACL---KPQQ-------I------S---------------I------D--E-I------A----D-------V----S------------T---VA-N-SRLSDSASTRNTTTNA--A--TIR-TTIAP....RNTTIS-N-P--RA---.S--  165
H2A.SN.85.ROD.M15390         -...M---------A-ACLV----..Q---------T-K---------R-------I------D--Q-IT-----A-----------------H------------T---VA-K-SST-SSTGNN-T-KS-STT--TP-D.........QEQEIS-DTP-ARA---.S--  155
H2B.CI.88.UC1.L07625         -AHTS-H-F-LL--I----FLGHKKN-------I-------V------T-------V---------T-ISV-I--A-----------VD---S------------T---VA----NTG-NTTTKPITTPI-TTKP.............SENLL-D--P--KN-T-.P-I  156
H2B.CI.x.EHO.U27200          -AHVN-Y--VTL--I-I--YMGK..NF------I---K--S-------R-------V------D--T-IQ--I--A----D----D--TK---S------------T---V--K---.TWSSASKETTTSSASLRS--.............QTLL--D-K--QN-S-.A-I  154
H2B.GH.86.D205.X61240        -AYFSSR-P--L--IGIS-FV-K..Q-------I-------V--I---T-------V---------T-IR--I--A----D----Q--VD---R------------T---VA-N-S-T--NPGNASSTT--KPT---R...........GLKTI---DP--KN-S-.---  156
H2B.JP.01.KR020.AB100245     -AHINKY-FA-L--I-I--FMGE..DF------I---K---V----------I---V------D--D-I---I--A----D-------VA---N------------T---V--N-SNLN-TANTT-ANNI-A...............SKNIS-L-GND---RD---.--I  152
H2G.CI.92.ABT96.AF208027     -AY--------TX---T--F---..K-------I------SV---------------------D----L------A----D-------V----N-----T------T----A-----T--NK----GTT-V-PA-VP--KM.......VTAEL--S--Q-LMY---.--I  160
H2U.FR.96.12034.AY530889     -A---------L-C-NACV----..Q-------I-------------N------------------L------A------------V----S------------T----A---S-T--N-----GTAA--T-PAMQES-K..........VI---EP--RNNS-.S--  157
H2AB.CI.90.7312A.L36874      KM-GK-L-FV-S--A-A-L----..K---------V----S---------------I------D--Q-I------A------------V----S------------T---VA-S--STTATTTPPSTTNN--TTEPT-............GGPEI---FP-MRT---.---  156
H2AB.CM.03CM_510_03.NEW      -AYSSSC-P-TL----IC-FTXK..Q-X-----I---X--XX--X-----------I------D--Q-IT-----A----X------------R------------T-XXVA-N-TTT-ATPQNT-TRN.....................TTIIEDNDP--SRN--.---  146
MAC.US.x.MAC239              -----------------------..-----------------------------------------M----------------------------R-------------------------------------------------......SA-----------------.---  160
MAC.US.x.251                 -----------------------..-----------------------------------------L----------E-------------------------T-------------------S---TAAPKAMSE.........-IN-I-------VH---.---  159
MAC.US.x.251_32H             -----------------------..Q----------------E-----------------------L----------E-----------------------------------K-------L----P-AP-A--.........-I----------TH---.---  159
MAC.US.x.251                 -----------------------..Q----------------------------------------L----------E-----------------------------------------------S--ITTAAP-S-PVSE.....-I------------N--.---  162
MAC.US.x.17EC1.AY033233      -----------------------..----------V------------------------------M----------------------------------------------------------------------------RV......-----------------.A--  160
MAC.US.x.17EFR.AY033146      -----------------------..----------V------------------------------M----------------------------------------------------------------------------RV......-----------------.A--  160
MAC.US.x.1937.AY611495       ------------X----X---X-..-----------------------------------------M-----------------------------------------------------------X-X----.......SAX--------X--------.---  160
MAC.US.x.2065.AY611493       -----------------------..-------------------------------X---M----------X-------X------------------------------------------X-------.......SA-----------------.---  160
MAC.US.x.80035.AY611486      ------X----------------..-----------------------------------------M-------------------------------------X----------------------------------.......SA-----------------.---  160
MAC.US.x.81035.AY599200      -----------------------..Q----------------------------------------M-----------------------------------------------------------------------T.......SAX----------------.---  160
MAC.US.x.85013.AY611490      ---------------------X..X----------------------------------------X------------------------------------------------------------------------.......SAE----------------.---  160
MAC.US.x.92050.AY603959      -----------------------..-----------------------------------------M-----------------------------------------------------------------------.......SA-----------------.---  160
MAC.US.x.92077.AY599201      -----------------------..-----------------------------------------M-----------------------------------------------------------------------.......SA-----------------.---  160
MAC.US.x.93057.AY611492      -----------------------..Q----------------------------------------M-----------------------------------------X-----------------PX--.......SAX---------X-------.---  160
MAC.US.x.93062.AY607704      -----------------------..-----------------------------------------M-----------------------------------------------------------------------.......SA-----------------.---  160
MAC.US.x.95058.AY611494      -------------------X---..-----------------------------------------M-----------------------------------------------------------------------.......SA-------X---------.---  160
MAC.US.x.95086.AY607703      -----------------------..Q----------------------------------------L-----------------------------------------------------------------------.......SAE----------------.---  160
MAC.US.x.95112.AY588946      -----------------------..-----------------------------------------M-----------------------------------------------------------------------.......SA-----------------.---  160
MAC.US.x.96016.AY607701      -----------------------..-----------------------------------------M--------------------------------------------------------XA-----------.......SAX----------------.---  160
MAC.US.x.96020.AY611488      -----------------------..-------------------------------------------------------------------------------------------------------------------.......SA-----------------.---  160
MAC.US.x.96072.AY611491      ---P---------X---------..Q----------------------------------------L-----------------------------R---------------------------X-----T.......SA------------X---.---  160
MAC.US.x.96081.AY597209      -----------------------..-------------------------------------------------------------------------------------------------------------------.......SA-----------------.---  160
MAC.US.x.96093.AY611489      -----------------------..------X----------------------------------L-----------------------------------------X-----------------X--.......SAE----------------.---  160
MAC.US.x.96114.AY588945      -----------------------..-----------------------------------------M-----------------------------------------------------------------------.......SA-----------------.---  160
MAC.US.x.96123.AY611487      -----------------------..X---X------------------------------------L-----------------------------------------X-----------------X-X.......SA-----------------.---  160
MAC.US.x.96135.AY607702      -----------------------..-----------------------------------------M-----------------------------------------X-----------------X--.......SA-----------------.---  160
MAC.US.x.97009.AY599199      -----------------------..-----------------------------------------M-----------------------------------------------------------------------.......SA-----------------.A--  160
MAC.US.x.97074.AY599198      -----------------------..-----------------------------------------M-----------------------------------------------------X-------------.......SA-----------------.---  160
MAC.US.x.MM142.Y00277        ---------------------I..Q--------------------------------D----L----------E-----------------------------------K------S-------T--TAK....SVETR-I-----P-VVH---.---  163
MAC.US.x.r80025.AY576480     -----------------------..Q----------------------------------------L-----------------------------------------------------------------------.......SAE----------------.---  160
MAC.US.x.r90131.AY576481     -----------------------..-----------------N-----------------------M-----------------------------------------------------------------------.......SA-----------------.---  160
MAC.US.x.SMM142B.BD131285-------------CSKCLW.-I-I..Q--------------------------------D----L----------E-----------------------------------K------S-------T--TAK....SVETR-I-----P-VVH---.---  162
MNE.US.82.MNE                -------------F--A-----I..Q-----------------V-R-------------D----L---I-------E------------H------------T------K-------K------S----P-AIP-.......KAEAIKV---N-P--NH---.---  160
MNE.US.x.MNE027.U79412       -------------F--A-----I..Q-----------------V-R-------------D----L---I-------E------------H------------T------K-------K------S----AP-TK-TTT......KEIEV---N-T-VNR---.---  161
SMM.SL.92.SL92B.AF334679     -A-P-LH---D--F---L-TW-A..Q---I---I--------------Q-------V-----------L------A----D------------N------------T----A-K---N--------RAA---S-P-T-SPLTA..ASPSGEEI--D-M--TKNN--.S-I  165
SMM.US.x.F236                -----------L--V--LE-C-V..Q----------K-----------------------D----L-I----A----D------------N------------T----A-----T---------GNAG--TTAIT---.....TPSVAENVI--SNP--KNNS-.A--  162
SMM.US.x.H9.M80194           -----------LS-X-AS----V..Q-----------------V------R-----------------L-I----A----D------------N------------T------------XX-----GXPAP-TTQ---.....PPSPIIAKV--DSDP--RSN--.---  162
SMM.US.x.PBJ14               -----------L----AS----V..Q-------I-------V-------------------------L-I----A----D------------N------------T-------------------GTPAP-TTQ-T--QASTTPTSPITAKV--DSDP--KIN--.---  167
SMM.US.x.PBJA.M31325         -----------L----AS----V..Q-------I-------V-------------------------L-I----A----D------------N------------T-------------------GTPAP-TTQ-T--QASTTPTSPITAKV--DSDP--KIN--.---  167
SMM.US.x.PBJ                 -----------L--X-AS----V..Q-------I-------V------R------------------L-I----A----D------------N------------T------------XX-----GXPAP-TTQ-T--QASTTPTSPITAKV--DSDP--RSN--.---  167
SMM.US.x.PBJ                 -----------L----AS----V..Q-------I-------V-------------------------L-I----A----D------------N------------T-------------------GTPAP-TTQ-T--QASTTPTSPITAKV--DSDP--KIN--.---  167
SMM.US.x.PGM53.AF077017      ----------------A-------..Q----------K---------------------D----L-I----R-----------------N------------T---------T---------NEIP-TT--T--K..SPKAEAITAKVI--SDP--SNN--.---  165
SMM.US.x.SME543.U72748       -----------L--V--LE-C-V..Q----------K----------R-----------D----L-V-I--A-----------------N------------T----A-----T---------GRAE--TTAK---S...TTTTTVTPKVI--GD---KNNS-.A--  164
STM.US.x.STM.M83293          -A-P------------ACLT----..Q----------------------------------------L-I----A----D------------N------------T----A-----N---K----GKTV--VTP-AAA-A.......T-PEL-------VSNN--.---  160
```

```
                                                                                                                                                     V3
                                                                                                                                              |-------------------|
                              *                      *                *                 *                 *                 *                   *                 *                                                      *
MAC.US.x.239                EQEQMISCKFNMTGLKRDKKKEYNETWYSADLVCEQGNN.TGNESRCYMNHCNTSVIQESCDKHYWDAIRFRYCAPPGYALLRCNDTNYSGFMPKCSKVVVSSCTRMMETQTSTWFGFNGTRAENRTYIYWHGRDNRTIISLNKYYNLTMKCRRPGNKTVLPVTIMSGLV  329
H2A.IN.NNVA.NEW             RE-EIVI-Q------I---P-R-------S-V---PDSTTSTRNNSSSGQ-MR-WYMRD-------DM---------------------E-N-T---AA-----------------------------S-----------H---SIY-K-------V-I-H-P---  324
H2A.GM.x.MCR35.AY509260     GK-EVVN-Q------E-----Q---A---K-V---GNGT....T-T-----------T---------M---------F-------------A-N-----AAT-----------------------F---------------------IH-K-------V-I-L----R  319
H2A.IN.95.CRIK              GE-EVVN-E------V---P-T-------R-VD--PDST..T-SRK-----------T---------D-----------------------A-N-----A-T-----------------------L---SKSD------------SIH-K-------T-I-L---YK  323
H2A.CI.88.UC2.U38293        GD-G-VN--------EQ--I-G-TD----D-V--DST-K.--TNTT---R-------K---------SMK----T----------------A-N-P---AA------------------------------------H-----H-K-------V-I-L---HR  326
H2A.DE.x.BEN.M30502         GY-E-MQ-E---K--EQ---RR-KD---LE-V--DNTT.....AGT---R-----I-K---------M---------F-------------E---T---AA-------------------------------------------R-K---------I-L-----  323
H2A.DE.x.PEI2.U22047        GS-EVVD-R------QL--PQQ-S-----K-V--DTT-G...TSRK-----------T---------M---------LC------------E---P---AAT-----------------------------------TH-----H-K-----S---I-LR--R-  316
H2A.GH.x.GH1.M30895         GK-EIVN-Q-Y----E-----Q-------K-V---SN-T.KDGKN------------T----------K----------------------E-------A-T--------------------------------------------SIH-K-------V-I-L-----  315
H2A.GM.87.D194.J04542       GK-EVVE-E------EQ---RK--DA---R-V--DKT-G....TGT---R-------K---------MK--------F-------------E-------AA--------------------------K-----------------H-K-------V-I-L---RR  312
H2A.GM.x.ISY.J04498         RE-D-VE-Q------EL----Q-S-----K-V---SD-S..TDRK------------T---------M---------FV------------E-N-----A-T--------P---L-------------------------------IL----E----V-I-L---RR  315
H2A.GM.x.MCN13.AY509259     GE-EVVN-Q------E-----Q-S-----K-V---GNGT....TDT-----------T---------M---------F-------------A-N-----AAT-----------------------F---------------------IH-K-------V-I-L----R  319
H2A.GW.86.FG.J03654         KE-E--D-Q-S----E---R-Q-T-A---K-V--DNN...-SSQ-K-----------T---------M---------F-------------A-N-----AAT--------------------------K---------NF-----H-K----------I-F---FK  313
H2A.GW.87.CAM2CG.D00835     -E-KIVK-H------E-----Q-------S-V--DNSTDQ-T--TT-----------T---------M---------F-I-----K-----A-N-----A-T--------------------------K----------H---S-Y---------V-I-L---QR  322
H2A.GW.x.ALI.AF082339       GD-E-VN-R------Q---P-Q-------K-V---PF-T.-T-QT------------T--------------------------D-I----A-N-----AAT-------------------------------------H-----H-K-------V-I-L----I  320
H2A.GW.x.MDS.Z48731         -E-E-VK-Q------QI----Q-------S-V--GEM--.-T-Q-----S-------K---------TM--------F-------------Q-N-----A-T--------------------------K---------AS---SLH-K-------V-I-L-----  334
H2A.SN.85.ROD.M15390        GE-ET-N-Q------E-----Q-------K-V---T.---.ST-QTQ----------T-------------------------------A-N-----A-T-------------------------------------------SLH-K-----I-KQIML---H-  323
H2B.CI.88.UC1.L07625        GL-NTVD-Y------R--E--Q-KD---EK--E-NGNS....TSTI---RT---------------SL---------------------------------------------------T--------T-----M---SK--------------H----------I-I------N  322
H2B.CI.x.EHO.U27200         GL-E--D-Q-K-------ES-Q-KD---KQ-----K-T..RS---K--IKT----I-----------SL--------F-------K------N-------.LY-------------------------------K---------S------H-K-----M-V-IRTV--IL  321
H2B.GH.86.D205.X61240       GE-EIMQ-N-S----R--EL-Q-KD----E--E-NNTR...KYT----IRT---TI-----------SL--------FF-------------N-----A----------S------------------EK---------T----SIH-K-------V-IRTV---L  323
H2B.JP.01.KR020.AB100245    -F-N-VT-R-------V-EP-T--D---AE-VQ-........-G---HIKT----I-----------QL--------F-FM----S------N-------------------------------------KS---V-----K----IH---------V--------I  314
H2G.CI.92.ABT96.AF208027    QS-S-VG-----------QR---------Q--QX--S-K.SE-------R-------------------------F-------------A-N-------T--------------------------------------------IX-----------I--------  329
H2U.FR.96.12034.AY530889    ---PLV--R-----------Q--K-----H-I----NT..SES--K----------------------------------------------N-------T---RG------------S------------------------Y------------------  325
H2AB.CI.90.7312A.L36874     GE-E-VD-Q------E---T-Q-S-----K-V---SN-A.SDGRD------------T--------------------F----------------N---------------------------------M---SK---------------IH-K-------V-I-L-----  325
H2AB.CM.03CM_510_03.NEW     -N-E-VT-N-S----EQ----R-------Q-V--DKN-SSETLNKT---R-------K---------TM-----T---------------Q-N-----AA----I--------------------------------------------H-K-------V-I-L-P---  316
MAC.US.x.MAC239             ---------------X-----------------------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.251                ------G-----------------------------S.--------------------------------------------------------------------------------------------------------I-----------------Q--  328
MAC.US.x.251_32H            ------G------------T---------T-------S.-D-------------I---------T-------------------------------------------------------------------------------------------------  328
MAC.US.x.251                ----------T--------T---------T-------S.-D-----------------------T-------------------------------------------------------------------------------------------------  331
MAC.US.x.17EC1.AY033233     ---------------N-----------------------.-------------------K------------------------------------------------------------------------------------I---------------------  329
MAC.US.x.17EFR.AY033146     ---------------N-----------------------.-------------------K------------------------------------------------------------------------------------I---------------------  329
MAC.US.x.1937.AY611495      ---------------------------------------.------------------------------------X-----------------------------------------------------------------------------------------X-  329
MAC.US.x.2065.AY611493      -----X---------------------------------.-------------------------------------------------------------------------------------------------------------X----------------  329
MAC.US.x.80035.AY611486     ---------------------------------------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.81035.AY599200     ---------------------------------------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.85013.AY611490     -----------------------------X---------.-D-----------------------X---------------------------------------------------------------------------------------------------  329
MAC.US.x.92050.AY603959     ---------------------------------------.-------------------------------------------------------------------------------------------------------X----X----------------  329
MAC.US.x.92077.AY599201     ---------------------------------------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.93057.AY611492     -----------------X-X-------------------.--X-----------------------------------------X----------------------------------------------------------------------------------  329
MAC.US.x.93062.AY607704     ---------------------------------------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.95058.AY611494     -----------------------------X---------.-X------------------------------------------NX---------------------------------------------------------------------------------  329
MAC.US.x.95086.AY607703     -----------------------------X---------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.95112.AY588946     ---------------------------------------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.96016.AY607701     --------------------X------------------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.96020.AY611488     ---------------------------------------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.96072.AY611491     --------------------X------------------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.96081.AY597209     ---------------------------------------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.96093.AY611489     ---------------------------------------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.96114.AY588945     ---------------------------------------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.96123.AY611487     ---------------------------------------.----------------------------------------------------------------------------------------------------H------------------------  329
MAC.US.x.96135.AY607702     ---------------------------------------.----------------------------------------------------------------------------------------------------H------------------------  329
MAC.US.x.97009.AY599199     -----------------X-------------X-------.--------X--------------------------------------------------------------X-------------------------------------------------------  329
MAC.US.x.97074.AY599198     -------------------------X-------------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.MM142.Y00277       ---P--------------------------------S.--------------------C---D-----C---------------------N--------------------R---------------------------H---------------------A--  332
MAC.US.x.r80025.AY576480    -----------------------------X---------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.r90131.AY576481    -----------------X---------------------.-----------------------------------------------------------------------------------------------------------------------------  329
MAC.US.x.SMM142B.BD131285---P--------------------------------S.--------------------C---D-----C---------------------N--------------------R---------------------------H---------------------A--  331
MNE.US.82.MNE               ---P---------------RR---------------S.-E----------------------------------------------N-------------------------------------SK---------------------------------------  329
MNE.US.x.MNE027.U79412      ---P----------------R---------------S.-ED----------------------------------------K------N-------------------------------------SK-----------N---------------------------  330
SMM.SL.92.SL92B.AF334679    ---P--G-Q---------Q-RQ-------R-------G-...ES-----------------------------Y---------------A-N-------------------------------------GS-------------I-------------------------  332
SMM.US.x.F236               ---P--G--------N-------------R--I---SA-..ES--K---H-----------------------------------S--L--A-N-------------------------------------KS---------------R----E--------------  330
SMM.US.x.H9.M80194          ---P-V--------------R--------R------NS-GNETD-K-----------------------X-------------------A-N-T-----------------------------------X-S-----------X----R--------------------  332
SMM.US.x.PBJ14              ---P-V--------------R--------R------NS-..ETD-K---------------------------------------S-----A-N-T-------------------------------------S----------------R--------------------  335
SMM.US.x.PBJA.M31325        ---P-V--------------R--------R------NS-..ETD-K---------------------------------------S-----A-N-T-------------------------------------S----------------R--------------------  335
SMM.US.x.PBJ                ---P-V--------------R--------R------NS-ENETD-K-----------------------X-------------------A-N-T-----------------------------------X-S-----------X----R--------------------  337
SMM.US.x.PBJ                ---P-V--------------R--------R------NS-..ETD-K---------------------------------------S-----A-N-T-------------------------------------S----------------R--------------------  335
SMM.US.x.PGM53.AF077017     ---P-V--R-------K------------R-I----STT.SR---K-------------------------------------------A-N-------------------------------------S-S---------------IR------------------I  334
SMM.US.x.SME543.U72748      ---P--G------------I---------R--I---PA-..-S--K---Q-----------------------------------S-----A--------------------------------------NS---------------------------------------  332
STM.US.x.STM.M83293         -E-SLVG-------------R--------S--I---NVT..-E------R-----------------L-------------------T--A-N---------------------------------------------------------S--------------------  328
```

```
                              V3
                                   *                                                      *    *                                        *                            *
MAC.US.x.239                FHS..QP.INDRPKQAWCWFGGKWKDAIKEVKQTIVKHPRYTG.TNNTDKINLT.APG.GGDPEVTFMWTNCRGEFLYCKMNWFLNWVEDRNTAN.....QKPKEQHKRNYVPCHIRQIINTWHKVGKNVYLPPREGDLTCNSTVTSLIANIDW...IDGNQTNITMSAE 485
H2A.IN.NNVA.NEW             ---..--.--T--R------K-R--E-MQ------AE----R-.--I-E--KF-.---K-S----AY-----------N-T-----I-N-TNE...........T-H-------K--------A-RH--------E---D-----I----EE.........-V-F--G 471
H2A.GM.x.MCR35.AY509260     ---..--V--K--R------K-E--G-MQ---E-LA-----K-.--E-KN--F-.---K-S----AY-----------N-T-----I-NKTHR-...............------K-------------------E---------I-----A....NE-N-D--F--- 466
H2A.IN.95.CRIK              ---..R-V--T---------K-R----MQ---K-VAE---AV..-KDPKN-TFA.---K-S----EY-----------D-T---D-I-N-P-RK..........AW------------------H--------E---------I-----V..NEIDKR----F--- 475
H2A.CI.88.UC2.U38293        ---..-AV--KK-R------K-N--G-MQ-----LAG----K-.--D-S---FV.K--V-S-----Y---------F--N-T-------N-TSQK...........Q---A--------------QY--------E---------I-----....T--------F--- 476
H2A.DE.x.BEN.M30502         ---..--.--T--R---R---R-RE-MQ-----L-Q----K-.I-D-G---F-.K--A-S----A------------N-T--------K-QTR...................--K-------------------E-A-E-----I-----I..DKNRTH----F--- 471
H2A.DE.x.PEI2.U22047        ---..R-I--E----------D--K-MQ-----L------R-.--D-Q---F-.Q--K-S-A--VY-----------N-TR----I-N-AHPQ...............---A-----------R--Q-I-------E-V-------I-----M....FD---S--F--- 465
H2A.GH.x.GH1.M30895         ---..--.--T--R------K---RE-MQ-----LI-----K-.--D-KN--F-.K--R-S----AY-----------N-T-------N-PNQT...........QH--A--------------------------Q---------I-----V....NS------F--- 464
H2A.GM.87.D194.J04542       ---..R-VY-KK-G------Q-N-IE-MR-----LA-----G-.--D-G---F-.K--I-S-----Y-----------N-T-------NKTNQT...........HG--A--------------T---------E---------I-----S....----------F--- 462
H2A.GM.x.ISY.J04498         ---..-KI--KK-R----R-K-E-RE-MQ-----L------K-.--D-N---F-.--EKDS----AY-----------N-T-------NKTGQQ...........H-------E-------------------E-S-E-----I-----V...DGD-R----F--- 465
H2A.GM.x.MCN13.AY509259     ---..--V--K--R------K-E--G-MQ---E-LA-----K-.--E-KN--F-.---K-S----AY-----------N-T-----I-NKTHR-.............-------------------------E---------I-----A....N--N----F--- 466
H2A.GW.86.FG.J03654         ---..--V--KK-R------E-Q--E-MQ---E-LA-----K-NRSR-EN-KFK.---R-S-----Y--------S---N-T-------N-TGQK...........Q---A--R---------R----L-------E---------I-----A......-D-----F--- 463
H2A.GW.87.CAM2CG.D00835     ---..R-I--K--R------K-N-TE-MQ-----LAE----K-.-K-ITD-TFK.--ER-S-----Y--S------F--N-T-------NKPNTT...............---A---------------------------E---------I-----E...R-NQT----F--D 473
H2A.GW.x.ALI.AF082339       ---..--.--K--R------K-E-RK-MQ---E-L------K-.--D-NQ--F-.K--R-S-A--VY----------H-N-T-------NKTGQE...........QH--A---K---I--A-------------E---------------T....----------F--- 469
H2A.GW.x.MDS.Z48731         ---..--.--R--R------K-N-RR-MQ---E--I-----K-.-S-ITN-TF-.---K-S----AY-----------N-T-------N-TDKP...........WH------K----------R---------E-V-------I-----D....KD------F--D 483
H2A.SN.85.ROD.M15390        ---HY--.--K--R------K------MQ---E-LA-----R-.--D-RN-SFA.---K-S----AY-----------N-T-----I-NKTHR-...............---A---K----------R---------E-S-------I-------...QNN------F--- 472
H2B.CI.88.UC1.L07625        ---..--.L-T--R------K-N-IE--R---E--I-----K-.----ER-R-V.G-SA-S----RH---------F--N-T-------N-TGTT...........QK---T---K--V--------Y--------T-S---S---------VYYDGNDTK-------- 475
H2B.CI.x.EHO.U27200         ---..--.--K---------K-N-TE--Q---E--KN----S-.-T-ISQ-R-A.EHARSS----RY-----------N-TF------N-TGLK..............---AS------V-----I-R----------E-S---------------.....-K-L----V--- 470
H2B.GH.86.D205.X61240       ---..--.--K--R------K-N-TE------R--I-----K-GAK-ITSVK-VSEH-K-S---T-Y-----------N-T-------NKTNTT...........R---A------------------I-------E-S------------NS.....-NST---SV--- 474
H2B.JP.01.KR020.AB100245    ---..--.--Q--R--------N-TQ------E--A--------.--D-A--K-A.EHSQ-S----RY-----------N-T-------N-TGIK............K--------V---------------------H---------I-------...TNN-R-------- 463
H2G.CI.92.ABT96.AF208027    ---..--.--E--R-----------E-MQ---E-V----S-K-.--D-KX-TF-.T--E-S----K--------------T-----I----MTL.....L--Q-RQR---------------------------------------------------EE..NRSN-H---IF--- 487
H2U.FR.96.12034.AY530889    ---..-A.--NN---------N--G--Q---E--AN----H-.-K-ISQ---A.E-A.---S--K--------------T-------N-SMEG.....-TSR-R------------V--------R-----------------M-----R....-NK-----S---- 481
H2AB.CI.90.7312A.L36874     ---..--.--K--R------K-E-RE-MQ-----LI-----K-.--D-RN-TF-.K--T-S----AY-----------N-T-------N-TGQT...........QH--A---K-------------------Q---------------V..DVGN-R----F--- 476
H2AB.CM.03CM_510_03.NEW     ---..--.--R--R------R-N-TE-M------L------K-.--D-T---F-.K--E-S----AY-----------N-T-------N-TGT............QH--A---K-------------------E---------I-----T....K-------F--- 464
MAC.US.x.MAC239             ---..--.------------------------------------.-----------.--R.--------------------------------------.....----------------------------------------------------------...------------- 485
MAC.US.x.251                ---..--.------------R-----------------------.-----------.---.--------------------------------------.....--------------------------------------I---------------...T------------ 484
MAC.US.x.251_32H            ---..--.--------------N-------M-------------.-----------.--R.--------------------------------DVT-.....-R---R-R---------------------------------------------------...T------------ 484
MAC.US.x.251                ---..--.------------------------------------.-----------.---.--------------------------------DVTT.....-R---R-R---------------------------------------------------...T-----S------ 487
MAC.US.x.17EC1.AY033233     ---..-R.-----R------------------------------.-----------.--R.--------------------------------------.....----------------------------------------------------------...------------- 485
MAC.US.x.17EFR.AY033146     ---..-R.-----R------------------------------.-----------.--R.--------------------------------------.....----------------------------------------------------------...------------- 485
MAC.US.x.1937.AY611495      ---..--.----------R-X-E---------------------.-----------.---.-----------------------------------T-....QKPK#-XXXXX---------------------------------------------------...------------- 485
MAC.US.x.2065.AY611493      ---..--.------------------------------------.--X--------.--X.-----------------------------------T-.....------X-----------------------------------------------------...-HE---------- 485
MAC.US.x.80035.AY611486     ---..--.------------------------------------.-----------.---.--------------------------------------.....--X-------------------------------------------------------...------------- 485
MAC.US.x.81035.AY599200     ---..--.------------------------------------.-----------.---.--------------------------------------.....-XXXX-X---------------------------------------------------...------------- 485
MAC.US.x.85013.AY611490     ---..--.------------------------------------.-----------.---.-----------------------------------T-.....---T-------------------------------------------------------...------------- 485
MAC.US.x.92050.AY603959     ---..--.-X-X-X-----------------------X------.-----------.--X.--------------------------------------.....------X----------------------------------------------------...------------- 485
MAC.US.x.92077.AY599201     ---..--.------------------------------------.-----------.---.--------------------------------------.....----------------------------------------------------------...------------- 485
MAC.US.x.93057.AY611492     ---..--.--------------X---------------------.--D--------.---.-----------------------------------T-.....--XX-------------------------------------------------------...------------- 485
MAC.US.x.93062.AY607704     ---..--.------------------------------------.-----------.---.-------------------------------------------..----------------------------------------------------------...------------- 479
MAC.US.x.95058.AY611494     ---..--.----------R-E-X---------------------.-----------.---.--------------------------------------.....---TXR-----------------------------------------------------...------------- 485
MAC.US.x.95086.AY607703     ---..--.------------------------------------.-----------.---.-----------------------------------T-.....--X--------------------------------------------------------...------------- 485
MAC.US.x.95112.AY588946     ---..--.------------------------------------.-----------.---.--------------------------------------.....----------------------------------------------------------...------------- 485
MAC.US.x.96016.AY607701     ---..--.------------------------------------.-----------.---.-----------------------------------T-.....HXT--------------------------------------------------------...------------- 485
MAC.US.x.96020.AY611488     ---..--.------------------------------------.-----------.---.--------------------------------------.....----------------------------------------------------------...------------- 485
MAC.US.x.96072.AY611491     ---..--.--------------E---------------------.--D--------.---.-----------------------------------T-.....--X----------------------R--------------------------------...------------- 485
MAC.US.x.96081.AY597209     ---..--.------------------------------------.-----------.---.--------------------------------------.....----------------------------------------------------------...------------- 485
MAC.US.x.96093.AY611489     ---..--.--E-------R---E--E------------------.--D--------.---.----------------------------------XT-.....--S--------------------------------------------------------...------------- 485
MAC.US.x.96114.AY588945     ---..--.------------------------------------.-----------.---.----------------------------------XH.....H---X-------------------------------------------------------...------------- 485
MAC.US.x.96123.AY611487     ---..--.--------------X--X------------------.-----------.---.-----------------------------------T-.....--S------X-------------------------------------------------...------------- 485
MAC.US.x.96135.AY607702     ---..--.--------------X--X------------------.-----------.---.-----------------------------------T-.....--S--------------------------------------------------------...------------- 485
MAC.US.x.97009.AY599199     ---..--.------------------------------------.-----------.---.X-N---------------------------------.....-------------------------------------X--------------------...------------- 485
MAC.US.x.97074.AY599198     ---..--.------------------------------------.-----------.---.--------------------------------------.....--R..KN----------------------------------------------------...------------- 483
MAC.US.x.MM142.Y00277       ---..--.V-E-------R---N--E------------------.-----------.--R.--------------------------------SLTT.....-----R-------------------------------------------------N-...T-----S------ 488
MAC.US.x.r80025.AY576480    ---..--.------------------------------------.-----------.---.-----------------------------------T-.....--X--------------------------------------------------------...------------- 485
MAC.US.x.r90131.AY576481    ---..--.------------------------------------.-----------.---.X-N---------------------------------.....----------------------------------------------------------...------------- 485
MAC.US.x.SMM142B.BD131285---..--.V-E-------R---N--E------------------.-----------.--R.--------------------------------SLTT.....-----R-------------------------------------------------N-...T-----S------ 487
MNE.US.82.MNE               ---..--.----------R-E-N--E------------------.-----------.---.-------------------------------K-LTG...TT---Q-R-------------------R----------------------------------...------------- 487
MNE.US.x.MNE027.U79412      ---..--.--E-------R-E-N--E------------------.-----------.--R.-------------------------------K-LTG...TT---Q------------------------------------------------------...------------- 488
SMM.SL.92.SL92B.AF334679    ---..--.--E--R-----------RE-MQ---K----------.--D-R------.---.----------------------------------SSPRWTT-TK------------------R-----------------------------------...--N-E-------- 493
SMM.US.x.F236               ---..--.--E----------E-S--K--Q---E-L--------.--D-R------.--A.--------------------------------DQKGGRWKQ-NR---Q-K-------------------------------------E----...-NS-E-------- 491
SMM.US.x.H9.M80194          ---..--.--E------------E--E--R---XXL------K-.--XXXQ-K---.---.------------------------------XIKNGS.RWTS-NQ--RYQK-------X------------X------------------E----...-N--K-------- 492
SMM.US.x.PBJ14              ---..--.--E---------------K--Q---E-L--------.--K-EQ-K---.---.-------------------------------IQNGS.RWTS-NQ--RQR------------------------------------------E----...-N--E-------- 495
SMM.US.x.PBJA.M31325        ---..--.--E------------E--K--Q---E-L--------.--K-EQ-K---.---.------------------------------NIQNGS.RWTS-NQ--RQR------------------------------------------E----...-N--E-------- 495
SMM.US.x.PBJ                ---..--.--E------------E--E--R---XXL------K-.--XXXQ-K---.---.------------------------------XIKNGS.RWTS-NQ--RYQK-------X------------X------------------E----...-N--K-------- 497
SMM.US.x.PBJ                ---..--.--E------------E--K--Q---E-L--------.--K-EQ-K---.---.-------------------------------IQNGS.RWTS-NQ--RQR------------------------------------------E----...-N--E-------- 495
SMM.US.x.PGM53.AF077017     ---..--.--E------------S-RE--Q---E----------.--E-K------.---.---------------------------------SGSLWTQ-TKADEK----------------R-------------------------D----...------------- 495
SMM.US.x.SME543.U72748      ---..--.--E-------R---N-SE--Q---E-L---------.--D-R------.--A.-------------------------------DQNSNRWKQ--KP--Q-------------------------------------------E----...-NN-E-------- 493
STM.US.x.STM.M83293         ---..--.--E------------E-RG------E-L--------.--D-A--RIV.---.----------------------------I-N-S-SEMRDWNKNK---Q-----------V---------------Q----------I-------...TNN-E----A--- 489
```

```
                                                  gp120 end  gp41 start                                                                                            *
MAC.US.x.239               VAELYRLELGDYKLVEITPIGLAPTDVKRYTTGGTSRNKRGVFVLGFLGFLATAGSAMGAASLTLTAQSRTLLAGIVQQQQQLLDVVKRQQELLRLTVWGTKNLQTRVTAIEKYLKDQAQLNAWGCAFRQVCHTTVPW....PNASLTPKWNNETWQEWERKVDFLEENI  651
H2A.IN.NNVA.NEW            --D------------------F---PE---SFA.SR-QT---------------D----------S--P-----------------------M------------A----------A----DS---------------....E-DT-V-N-K-M-CK---Q--RH--A--  636
H2A.GM.x.MCR35.AY509260    ---------------------F---AE---SST.PM-----------------------------S--------------------------M------------A----------------S---------------....V-DT---E---M------G-IRD--A--  631
H2A.IN.95.CRIK             ------V---------V----F---SE---SS-#-G-HT---L--E-------------------S--P-----------------------M------------A----------------S---------------....V-D----D-K-M--*---Q--RY--A--  639
H2A.CI.88.UC2.U38293       --------------I------F---SE---SSA.PA----------L------------------S-------------------I-------------------A----------------S---T-----------....V-D----R---M------KQ-RY--A--  641
H2A.DE.x.BEN.M30502        --------------I------F----QR--SST.PV------------------------R----S--------------------------M------------A----------H-----S---------------....V-D--S-D-K-M------KQ-RY--A--  636
H2A.DE.x.PEI2.U22047       ---------------------F---SE---SSA.PQ----------V------------------S-HPGLYW-------------------M----------------------R---R--S---------Y---L-....E-N-IV-D---M------QQTRD--A--  630
H2A.GH.x.GH1.M30895        --------------I-V----F---RE---SSA.PV-----------------------------S--------------------------M------------A----------------S---------------....V-D--S-D---M------KQ-RY--A--  629
H2A.GM.87.D194.J04542      --------------I-V---PF---KE---SSA.PV-----------------------G-----S--------------------------M------------A----------------S---------------....V-D----D---M------KR-HY--A--  627
H2A.GM.x.ISY.J04498        ----------------V----F---AE---SSA.PG-H----L--------T---A---------S------FR------------------M------------A---------A---R--S---------------....V-DT---E---M------H-IR---A--  630
H2A.GM.x.MCN13.AY509259    ---------------------F---AE---SST.PM-----------------------------S--------------------------M------------A----------------S---------------....V-DT---E---M------G-IRD--A--  631
H2A.GW.86.FG.J03654        A--------------------F---S----SSA.HQ-HT--------------------------S--------------------------M------------A----------------S-----------S---....V-DT---D---M------Q--RY--A--  628
H2A.GW.87.CAM2CG.D00835    ---------------------F---SQ---SPA.HG-P---A---------T---V---T-----S------------------------------------I--A----------------S---------------....A-E----D---M------Q--RY--A--  638
H2A.GW.x.ALI.AF082339      ---------------------F---SER--SST.PR-----------------------T-A---S--------------------------M------------A-------------R--S---------------....V-N--K-D-D-M------QQ-RY--A--  634
H2A.GW.x.MDS.Z48731        ---------------------F---PE---SSA.HG-H-------------T---A---------S--------------------------M------------A----------------S---------------....V-S--E-D-E-M------Q--RY--A--  648
H2A.SN.85.ROD.M15390       ---------------------F---KE---SSA.HG-HT-------------------------VS---------------------------------------A---------Q---R--S---------------....V-D--A-D-D-M------KQ-RY--A--  637
H2B.CI.88.UC1.L07625       -G-------------------F---EI---SST.-P------M------L--M-------T----S-----------------------------------------------------L--S---------------....--ET---D-E-M---Q--KR-N--DA--  640
H2B.CI.x.EHO.U27200        -S---K---------------F---SI---SSV.-P------L----------------------S-----------------V---------------------A----------------S---------------....V-E--K-D---M---Q---Q-R--DA--  635
H2B.GH.86.D205.X61240      -S-------------------F-----R--SSV.KP------M---------M-------T----S----------------PV---------------------A----------------S---------------....--ET---N---M---Q--KQ-H--DA--  639
H2B.JP.01.KR020.AB100245   -S-------------------F---S----SSA.-P------P----------------------S---------------------------------------A-------------L--S---------------....--KNF--N-D-M---Q--NQ-R--D---  628
H2G.CI.92.ABT96.AF208027   ---------------------F---N----SSV.-PK---------------M-------T----S--------------------------M--------------------------R--S---------------DALGA-KT-E-Q---M------KQIN---D--  656
H2U.FR.96.12034.AY530889   -----------------R---F---P----SSV.-P-----------------------------S--------------------------M--------------------------S------------------....I-DT---N-D-M------E--NY-----  646
H2AB.CI.90.7312A.L36874    --------------I-V----F---SE---SST.PG-H-------------T---A---------S--------------------------M------------A----------------S---------------....V-D----D-D-M---Q--KQIRD--A--  641
H2AB.CM.03CM_510_03.NEW    --------------I------F---TEG--SST.PG-----------------------------S---------------------------------------A-------------X--S------------X--....V-E----N-D-M---Q--E--RD--A--  629
MAC.US.x.MAC239            -------------------------X----------------------------------------------------------------------------------------------------------------....------------------------X---  651
MAC.US.x.251               -------------------------N-R-----------------------------------------------------L--------------------------------------------------------....----------------------------  650
MAC.US.x.251_32H           ------------------------------------------------------------------------------------------------------------------------------------------....-------D---D----------------  650
MAC.US.x.251               ------------------------------------------------------------------------------------------------------------------------------------------....-------D---D----------------  653
MAC.US.x.17EC1.AY033233    --------------------------------------------------------------------------------------------------------------------------T---------------....-----I----------------------  651
MAC.US.x.17EFR.AY033146    --------------------------------------------------------------------------------------------------------------------------T---------------....-----I----------------------  651
MAC.US.x.1937.AY611495     -------------------------X---------------------X----------------------------------------------------------------------------------------X-....--XNX--R-D------------------  651
MAC.US.x.2065.AY611493     -------------------------G-------------------------------------------X--------------------------------------------------------------------....---------D------------------  651
MAC.US.x.80035.AY611486    -------------------------X----------------------------------------------------------------------------------------------------------------....------------------------X---  651
MAC.US.x.81035.AY599200    -------------------------N----------------------------------------------------------------------------------------------------------------....----------------------------  651
MAC.US.x.85013.AY611490    -------------------------X----X---------------------------------------------------------------------------X-------------------------------....---X---X----------------X---  651
MAC.US.x.92050.AY603959    ------------------------------------------------------------------------------------------------------------------------------------------....----------------------------  651
MAC.US.x.92077.AY599201    -------------------------X----------------------------------------------------------------------------------------------X-----------------....-----I----------------------  651
MAC.US.x.93057.AY611492    -------------------------H----------------------------------------------------------------------------------------------------------------....---------X--------------X---  651
MAC.US.x.93062.AY607704    -------------------------G----------------------------------------------------------------------------------------------------------------....----------------------------  645
MAC.US.x.95058.AY611494    -------------------------X----------------------------------------------------------------------------------------------------------------....-----X---D------------------  651
MAC.US.x.95086.AY607703    -------------------------N----------------------------------------------------------------------------------------------------------------....-----X----------------------  651
MAC.US.x.95112.AY588946    ------------------------------------------------------------------------------------------------------------------------------------------....----------------------------  651
MAC.US.x.96016.AY607701    -------------------------X----X-----------------------------------------------------------------------------------------------------------....-------X-X------------------  651
MAC.US.x.96020.AY611488    ------------------------------------------------------------------------------------------------------------------------------------------....----------------------------  651
MAC.US.x.96072.AY611491    -------------------------H----A-----------------------------------------------------------------------------------------------------------....---------D------------------  651
MAC.US.x.96081.AY597209    ------------------------------------------------------------------------------------------------------------------------------------------....----------------------------  651
MAC.US.x.96093.AY611489    -------------------------X----------------------------------------------------------------------------------------------------------------....-------E--------------------  651
MAC.US.x.96114.AY588945    ------------------------------------------------------------------------------------------------------------------------------------------....---X------------------------  651
MAC.US.x.96123.AY611487    -------------------------G----------------------------------------------------------------------------------------------X-----------------....-----X------------------X---  651
MAC.US.x.96135.AY607702    -------------------------G----------------------------------------------------------------------------------------------------------------....---N------------------------  651
MAC.US.x.97009.AY599199    -------------------------X----------------------------------------------------------------------------------------------------------------....---------S------------------  651
MAC.US.x.97074.AY599198    -------------------------G----------------------------------------------------------------------------------------------------------------....----------------------------  649
MAC.US.x.MM142.Y00277      -------------------------N--------------------------------------V-------------------------------------------S-----------------------------....-------D----------------A--  654
MAC.US.x.r80025.AY576480   -------------------------N----------------------------------------------------------------------------------------------------------------....---X-X----------------------  651
MAC.US.x.r90131.AY576481   ------------------------------------------------------------------------------------------------------------------------------------------....----------------------------  651
MAC.US.x.SMM142B.BD131285- ------------------------N--------------------------------------V-------------------------------------------S----------H------------------....-------D----------------A--  653
MNE.US.82.MNE              -------------------------N----------------------------------------------------------------------------------------------------------------....---N---N--------------------  653
MNE.US.x.MNE027.U79412     -------------------------N---------P------------------------------------------------------------------------------------------------------....-------N--------------------  654
SMM.SL.92.SL92B.AF334679   ---------------------M---H------S.--K-------------------------------------------------------------------------------------S-------------L-....--D--V-D---M------K--E---A--  658
SMM.US.x.F236              --------------I----------S-R----T-A---------------------------V--S--------------------------------------------------------S---------------....--ET-V-N---M-------Q-----A--  657
SMM.US.x.H9.M80194         ---------------------F---S------T-A------------------------------S-----------------------------------------------------XX-S-----------X---....--DT---X---MX-----KQ-N---A--  658
SMM.US.x.PBJ14             --------------------AF---S------T-A---------------------------V--S----------------------------------A---------------------S---------------....--DT---N---M------KQ-N---A--  661
SMM.US.x.PBJA.M31325       --------------------AF---S------T-A-----------------------S---V--S----------------------------------A---------------------S--------------R....--DT---N---M------KQ-N---A--  661
SMM.US.x.PBJ               ---------------------F---S------T-A------------------------------S-----------------------------------------------------XX-S-----------X---....--DT---X---MX-----KQ-N---A--  663
SMM.US.x.PBJ               ---------------------F---S------T-A---------------------------V--S----------------------------------A---------------------S---------------....--DT---N---M------KQ-N---A--  661
SMM.US.x.PGM53.AF077017    -------------------------S-R----.-A---------------------------V-RS------------------------------------------------R-------S---------------....-----V-N---M-------Q--D-A--  660
SMM.US.x.SME543.U72748     ---------------------------R----T-A------------------------------S------------------------H-------------------------------S---------------....--D--V-N-D-M------G------A--  659
STM.US.x.STM.M83293        -------------------------N------S.---T-----------------------------------T------------------------------------------------S---------------....--D--V-D---M-------------A--  654
```

```
MAC.US.x.239                 TALLEEAQIQQEKNMYELQKLNSWDVFGNWFDLASWIKYIQYGVYIVVGVILLRIVIYIVQMLAKLRQGYRPVFSSPPSYFQQTHIQQDPALPTREGKERDGGEGGGNSSWPWQIEYIHFLIRQLIRLLTWLFSNCRTLLSRVYQILQPILQRLSATLQRIREVLRTELT 821
H2A.IN.NNVA.NEW              SQS--Q---------------I-------T------------------I-A--V---V--L-CSF-K----------G-I--I--HK-QEQ-D---T-E-V-DN--YR----P-A------LL------R-YNS-KD----SFLT------N-.......-DW--LRTA 799
H2A.GM.x.MCR35.AY509260      SQQ--Q-----------------S-----T------------II-IVV-----------SR--K----------G-L--I--HK-WEQ-A--ET-E-V-NNV-D-L*--P-R------H-------G-YNI--N----ISLT-R-AF-S-QRA-TA--DW---DAA 800
H2A.IN.95.CRIK               SQS--Q-------S-------------------T----------C----I-V------V----SR--K----------G-I--I--HK-LEQ-A---I-E-V-DN--YR-*--P-A------HL------R-YNS-KD----SFLT------N-.......-DW--LKTA 801
H2A.CI.88.UC2.U38293         SQS-------------------------------T---------------I-A---A--V--L-SRF-K----------G-L--I--HT-RGQ-AN-ET-G-A-DAS-YDF---P-N--QL--HL-T----G-Y-I--D---ANSPTRRL-S-N-....TA--DW--LKAA 807
H2A.DE.x.BEN.M30502          SQS-----------------------IL------T--V-------H----I-A---A--V--L-SRF-K----------G-L--I--HK-RGQ-AN--T-E-V-GDS-YDL---P-N-VQ---HL-T---IG-YNI--D---KNSPTRRL-S-S-....TA--DW--LKAA 802
H2A.DE.x.PEI2.U22047         SRS--Q----------------------------T-----------VII-I-A------V--L-SR--K----------G-I--I--HK-WEQ-D--ETDE-A-NSI-D-----P-A--------------G-Y-V-KD----SFPT--L-F-S-QRA-TT--DW--LTIA 800
H2A.GH.x.GH1.M30895          SQS--Q----------------------------T----------------V---A-----L-SR--K----------G-L--I--HT-RGQ-AN--T-E-DRDDD-YDL*--P-N------HL-T----G-YKI--D---TNSPTHRL-S-N-....TA--DW--LKAA 794
H2A.GM.87.D194.J04542        SQS--Q----------------------------T---------------I-G---A-----L-SR--K----------G-L--I--HT-RGQ-AN-ET-E-A-DDS-FGL---PLN--Q---HL-T----G-YNS--G---KNSPTRRL-S-S-....TA--DW--LKAA 793
H2A.GM.x.ISY.J04498          SES--Q----------------------------T----------M----IVA------V----SR--K----------G-I--I--HK-WEQ-D--ET-E-V-NDV-SR----P---------L------R-YNS--D----L-L---............L-DW--LKAA 789
H2A.GM.x.MCN13.AY509259      SQQ--Q----------------------------T-----------II-IVV-----------SR--K----------G-L--I--HK-WEQ-A--ET-E-V-NNV-D-----P-R------H------AG-YNI--N----ISLT-R-VF-S-QRA-TA--DW---DAA 801
H2A.GW.86.FG.J03654          SQS--Q-----------------T--L-FT--VR-------V---IVA-----------SR--K----------G-I--I--HK-QEQ-A--ET-E-V-SN--DR----P-A-------L------G-YNI--D----ISP-----F-S-QRA-TA--DW--LKAA 798
H2A.GW.87.CAM2CG.D00835      SQS-----L-------------N----T-----T---S-----------I-V------V----SR--K--------S-G-I--I--HK--EQ-AS-ET-E-V-GN--DR-----------------------G-YNI--N----IF-T------N-.......-DW--PKVA 801
H2A.GW.x.ALI.AF082339        SEQ--R--------T------------T--L--TA-V---------I--IVA------V----SR--K----------G-I--I--HK-QEQ---GET-E-V-DNV-DRL---P-A-L----HL-A---IG-Y-I--D----ISP-----FRS-QRA-TT--DW--LKAA 804
H2A.GW.x.MDS.Z48731          SQK---------Q------------I-------T---------------I-V------V----SR--K----------G-I--I--HK-REQ-D--ET-E-V-GN--DR----P-A-------L------R-YNS--D----IFLT----F-N-.......-DQ--LRIA 811
H2A.SN.85.ROD.M15390         SKS--Q-------------------I-------T--V-------L-I-A--A------V----SR--K----------G-I--I--HK-RGQ-AN-ET-E---SN--DRY---P-A--------------R-Y-I--D----SFLT--L-Y-N-.......-DW--LRTA 800
H2B.CI.88.UC1.L07625         ------------R-------------------FT--MA--RL-L-V-A-L-V-------M----R--K------------T--IP-RKHRGQ-AN-ET-DE--NE-AYR--------A-------RN--I--YNG--N--LKTS-----A-...........QP--LS-A 799
H2B.CI.x.EHO.U27200          -K--------------------Q--I-S----FT--MA--RL-L---I-IVV---A---I----R--K------------T--IP-RK-RGQ-AN-ET-EG--NNE-YR-----------P----RD--I--Y-G-----KTF-T----V-...........QP--LPPA 794
H2B.GH.86.D205.X61240        --------------------I-------------T------HL-L---A-LVV---V--------R--K------------T--IP-RK-RGQ-AN-ET-EG--ND-DYR------------L---RN--I--YNG-----LKTF---HQ-STN-........QP--LPVA 801
H2B.JP.01.KR020.AB100245     -K---V------E---K-----Q----S----FT---A---I-L-VI--LVV-------L----R--K------------V--IP-HKGQGR-AS-ET-EG--DDEDYR------R---------RN--I--YNS--S--LKIF-P--LVF...........QQ--LP-A 787
H2G.CI.92.ABT96.AF208027     -R--------------------------------T--V--VYL-L-V-A---V------V--I-GR--------X-------V--I--RK-QEX--K-EI-GES-NK--YR------------X---GN-----Y----A----IX-T---LF--I-R---A---H--L-AA 826
H2U.FR.96.12034.AY530889     -Q---A----------------N--I--------T--V--VYL-L-V-A-I-I------V--L-GS--K--------H---VE-IP--R-QEQ--KGEI-EGA-DS--YR------R------H---------Y-S--D----IC-S---LF-SI........-R-HL-IA 809
H2AB.CI.90.7312A.L36874      SES--Q-----------------------------------V------------IVA--VI--V---IGR--R----------G----IR-HK-QEQ-AN-ET-EG--ND--YR----------------RN--I--YDG-----LKTF-T---A-...........QP--LLFA 800
H2AB.CM.03CM_510_03.NEW      SQI--Q-----------------S-----T---R---------A--VV-X-A------XSR--K---------SG-V--I--HT-QGQ--S-ET-EG--ND-DYR--------V---L---RN--I--YNG-----LKTS--------...........P--LPFA 788
MAC.US.x.MAC239              -------------------------------------------------------------------------------------------------G----------------------------------------X---------------------------A 821
MAC.US.x.251                 ----------------------------------------------------------------------------------*------Q---------G----------------------------------------------------------R----------- 819
MAC.US.x.251_32H             ---------------------------------------------------------------------------------------------------G------------------------------------------A-------------A------------- 820
MAC.US.x.251                 -------------------------------------------I-V-----------------------------------*--T-------------G------------------------------------------A---------------R-V--------- 822
MAC.US.x.17EC1.AY033233      ------------------------N------------------------------------V-------------------------------------------------------------------------------------------------------------- 821
MAC.US.x.17EFR.AY033146      ------------------------N------------------------------------V--------------------------------------G-----------*--------------------------------------------------------- 820
MAC.US.x.1937.AY611495       -----------------------------------------------------------X---------------------------------------G------------------------------------X-------------F---X--------------- 821
MAC.US.x.2065.AY611493       ---------------------------------X------------------------------------------------------------------G--------X-----------------------------------------F----------X-------X 821
MAC.US.x.80035.AY611486      ---------------------------------------------------------------------------------------------------G---------------------------------------------------------------------- 821
MAC.US.x.81035.AY599200      ---------------------------------------------------------------------------------------------------G-------------------------------------------------------------X-------- 821
MAC.US.x.85013.AY611490      ---------------------------------X-----------------------------------------------------------------G---------------------------------------------------------------------- 821
MAC.US.x.92050.AY603959      ---------------------------------------------------------------------------------------------------G----------------------------------------L---------F------------------A 821
MAC.US.x.92077.AY599201      ---------------------------------------------------------------------------------------------------G---------------------------------------------------------------------- 821
MAC.US.x.93057.AY611492      -------------------------------------------------------------------------------------X-------------G---------------------------------------------------------------------- 821
MAC.US.x.93062.AY607704      ---------------------------------------------------------------------------------------------------G---------------------------------------------------------------------- 815
MAC.US.x.95058.AY611494      --------------------------------------------------V-----------------------------------H------------G---------------------------------------------------------------------- 821
MAC.US.x.95086.AY607703      -------------------------------------------X-------------------------------------------------------G---------------------------------------------------------------------- 821
MAC.US.x.95112.AY588946      ---------------------------------------------------------------------------------------------------X---------------------------------------------------------------------- 821
MAC.US.x.96016.AY607701      ---------------------------------------------------------------------------------------------------G---------------------------------------------------------------------- 821
MAC.US.x.96020.AY611488      ---------------------------------------------------------------------------------------------------X---------------------------------------------------------------------- 821
MAC.US.x.96072.AY611491      ---------------------------------T------------------------------X----------------------------------G----------------------------------------L-------------X---X-------I--- 821
MAC.US.x.96081.AY597209      ---------------------------------------------------------------------------------------------------G---------------------------------------------------------------------- 821
MAC.US.x.96093.AY611489      ---------------------------------T-----------------------------------------------------------------G---------------------------------------------------------------------- 821
MAC.US.x.96114.AY588945      ---------------------------------------------------------------------------------------------------G-------------------------------X-------------------------------------- 821
MAC.US.x.96123.AY611487      ---------------------------------------------------------------------------------------------------G---------------------------------------------------------------------- 821
MAC.US.x.96135.AY607702      ---------------------------------------------------------------------------------------------------G---------------------------------------------------------------------- 821
MAC.US.x.97009.AY599199      --------------------------------------------------------------------------X------------------------G---------------------X------------------------------------------------ 821
MAC.US.x.97074.AY599198      -----------------------------------------------------------------------------------------X---------G---------------------------------------------------------------------- 819
MAC.US.x.MM142.Y00277        ---------------------------------T---------I--I-----------------R-----------------*--T--------K---KG---GS-------------------------------------A-------F-------R-------L--- 823
MAC.US.x.r80025.AY576480     -------------------------------------------X-------------------------------------------------------G---------------------------------------------------------------------- 821
MAC.US.x.r90131.AY576481     ---------------------------------------------------------------------------------------------------G---------------------------------------------------------------------- 821
MAC.US.x.SMM142B.BD131285    ---------------------------------T---------I--I-----------------R-----------------*--T--------K---KG---GS-------------------------------------A-------F------YGEFG----L--- 822
MNE.US.82.MNE                -------------------------------------R--------------------------------------------*---R--Q----K--T-G---GS-------------------------------------A-------F--F-T-------------- 822
MNE.US.x.MNE027.U79412       -------------------------------------R-------------------------------------------------R--Q----K--T-G---DS-------------------------------------A-------F--F-T----V--------- 824
SMM.SL.92.SL92B.AF334679     -QM----RL------------------------T--VR------FL-I-IV-------V----SR---------------H--I---R-QE--AK--T-GE--N---YRL-----------------I-------YN-LTR-A--A--N--QLC----EIS-P---LV-R-AG 828
SMM.US.x.F236                -Q-----------------------I-------T---R------L--L---G------V-----R-------------A-V--IP-HKGQEP--K--E-G---DR--SR--------------------------S--DW-L-S------V--S--T----V---I-I-IA 827
SMM.US.x.H9.M80194           -XX----------X---------XX-XX---X--TX---------L--L---G------V-----R-------------AXVX-IP-XTGQE---K--E-G---XR---R---X----------------------S--DW-L-XC-----V--S--R----A---I-V-I- 828
SMM.US.x.PBJ14               -QS-----------T----------I-------T-----------L--L---G------V-----R-------------A-V--VP--TGQE---K--E-G---GR---R--------------------------S--DW-L-NC-T---V--S--R----A---I-VQIA 831
SMM.US.x.PBJA.M31325         -QS-----------T----------I-------T-----------L--L---G------V-----R-------------A-V--IP--TGQE---K--E-G---GR---R--------------------------S--DW-L-NC-T---V--S--R----A---I-VQIA 831
SMM.US.x.PBJ                 -XX----------X---------XX-XX---X--TX---------L--L---G------V-----R-------------AXVX-IP-XTGQE---K--E-G---XR---R---X----------------------S--DW-L-XC-----V--S--R----A---I-V-IA 833
SMM.US.x.PBJ                 -QS-----------T----------I-------T-----------L--L---G------V-----R-------------A-V--IP--TGQE---K--E-G---GR---R--------------------------S--DW-L-NC-T---V--S--R----A---I-VQIA 831
SMM.US.x.PGM53.AF077017      -QA----------------------I-------T-----------L--L--VG------V---I-R-------------V-V--IP-HKGQEQ--K--E-G---DR---R------------------------N----Y-L-I------VF---...-------V---IS 827
SMM.US.x.SME543.U72748       -Q-----------------------I-------T---R-------L--L--VG------V-----R----------P--A-V--IP-HK-QEP--K--E-GE--DR--SR--------------------------S--DW-L-I------V-----R----V---I-I-I- 829
STM.US.x.STM.M83293          -Q------V------------------------T--VR-------L-I-LVM--VA---M-L--R--K-----------CR--IP-HKGQEQ--K--T-EG--DR--IN--------T-------V------YN-F-AC---I--T-H-TF--I-RI-------V-LGAA 824
```

```
                                                                Env end
MAC.US.x.239                YLQYGWSYFHEAVQAVWRSATETLAGAWGD.LWETLRRGGRWILAIPRRIRQGLELTLL*   879
H2A.IN.NNVA.NEW             F----CEWIQ--F--AA-TTG----S-CRS.--G----I--G---V-------A--A---   858
H2A.GM.x.MCR35.AY509260     -----CEWIQG-F--FA-ATR-----T-R-.--GA-Q-I--G---V-------A-IA---   859
H2A.IN.95.CRIK              F----CEWIQ--F--FA-T-GK---S-GRN.-CG-VE-I--G----------A-VA---   860
H2A.CI.88.UC2.U38293        -----CEWIQ--F--IA-T-R-------RG.-CKAVQ-I--G---V-------A-IA---   866
H2A.DE.x.BEN.M30502         Q----CEWIQ--F--FA-TTR--------W.---AA--I--G---V-------A--A---   861
H2A.DE.x.PEI2.U22047        -----CEWIQ-VL-VLA-TTR----S--R-.--GAMG-I--G---V-------A--A---   859
H2A.GH.x.GH1.M30895         -----GEWIQ--F--FAKTTR----S---G.-CAAVQ-V--G---V-------A-IA---   853
H2A.GM.87.D194.J04542       -----CEWIQ--FR-FA-T-R--I----RG.-C-AAQ-I--G---V-------A-IA---   852
H2A.GM.x.ISY.J04498         -----CEWIQ--F--LA-VTR---TS-GRS.--GA-G-I--G---V-------A-IA---   848
H2A.GM.x.MCN13.AY509259     -----CEWIQG-F--FA-ATR-----T-R-.--GA-Q-I--G---V-------A-IA---   860
H2A.GW.86.FG.J03654         -----CEWIQ--F--LA-TTR------GR-.--RA-Q-I--G---V-------A--A---   857
H2A.GW.87.CAM2CG.D00835     F----CEWIQ--F--AA-A-R------CR-.V-GM-Q-I--G---V-------A--A---   860
H2A.GW.x.ALI.AF082339       -----CEWIQ--FR-FA-I-R---TNT-R-.--GAVQWV--R---V-------A-IA---   863
H2A.GW.x.MDS.Z48731         F----CEWIQ--F--FA-VTR---TS-CRG.--G--G-I--G---V-------A-IA---   870
H2A.SN.85.ROD.M15390        F----CEWIQ--F--AA-ATR------CRG.--RV-E-I--G---V-------A-IA---   859
H2B.CI.88.UC1.L07625        -----I-W-Q--I--AT-A-R----NTGRA.--KA---TAEA-I------------A---   858
H2B.CI.x.EHO.U27200         --R--I-W-Q--I--AA-A-G----S-ART.S-GV---AAGE-I---------A--A---   853
H2B.GH.86.D205.X61240       -----I-W-Q--LR-AA-ATG----S-GET.---A---AA-A-I----------------   860
H2B.JP.01.KR020.AB100245    -----IGW-Q--I--AT-A-G---TNTGRA.---A---TA-A-V------------A---   846
H2G.CI.92.ABT96.AF208027    -FS--FRWLQ--CT-AT-A-Q---TST-RA.--K--G-V--G------------------   885
H2U.FR.96.12034.AY530889    ------Q--K--F--FGKA-R---SRTGRE.-----G-V---LR---------F--A---   868
H2AB.CI.90.7312A.L36874     -----IGW-Q-----AAGATG----STGRT.---A---TA-G-I-V----------A---   859
H2AB.CM.03CM_510_03.NEW     -----IGW-Q--LR-AA-ATR----S-GET.---A--GAA-A--------------A---   847
MAC.US.x.MAC239             ------------------------------.----------------------------   880
MAC.US.x.251                ------N-------VD--------------.-*-A------------------------   877
MAC.US.x.251_32H            ---------Q----VG--------------.----------------------------   879
MAC.US.x.251                ---------------G------------R-.----------------------------   881
MAC.US.x.17EC1.AY033233     ------------------------------.----------------------------   880
MAC.US.x.17EFR.AY033146     -------------V----------------.----------------------------   879
MAC.US.x.1937.AY611495      ---------Y----------------------.----------------------------   880
MAC.US.x.2065.AY611493      ---------X-----------------X-.X--------X------------------   880
MAC.US.x.80035.AY611486     ------------------------------.----------------------------   880
MAC.US.x.81035.AY599200     ------------------------------.----------------------------   880
MAC.US.x.85013.AY611490     ---------------X--------X-----.----------------------------   880
MAC.US.x.92050.AY603959     ------------------------------.----------------------------   880
MAC.US.x.92077.AY599201     ------------------------------.----------------------------   880
MAC.US.x.93057.AY611492     ------------------------------.---X--------------X--X------   880
MAC.US.x.93062.AY607704     -----C---Q--------------------.----------------------------   874
MAC.US.x.95058.AY611494     ------------------------------.----------------------------   880
MAC.US.x.95086.AY607703     ------------------------------.----------------------------   880
MAC.US.x.95112.AY588946     ------------------------------.----------------------------   880
MAC.US.x.96016.AY607701     ------------------------------.----------------------------   880
MAC.US.x.96020.AY611488     ------------------------------.----------------------------   880
MAC.US.x.96072.AY611491     ---------Y-----F--------------.----------------------------   880
MAC.US.x.96081.AY597209     ------------------------------.-----------------X----------   880
MAC.US.x.96093.AY611489     --------------X---------------.-----------------X-----X---   880
MAC.US.x.96114.AY588945     ---------X--------------------.X---------------------------   880
MAC.US.x.96123.AY611487     ------------------------------.----------------------------   880
MAC.US.x.96135.AY607702     ------------------------------.----------------------------   880
MAC.US.x.97009.AY599199     ---------------------------X--.----------------------------   880
MAC.US.x.97074.AY599198     ------X--X--------------------.----------------------------   878
MAC.US.x.MM142.Y00277       ---------Q-----AQ-----------E.---A-Q----------------------   882
MAC.US.x.r80025.AY576480    ------------------------------.----------------------------   880
MAC.US.x.r90131.AY576481    ------------------------------.----------------------------   880
MAC.US.x.SMM142B.BD131285---------Q-----A#------------Ef---A-Q----------------------   880
MNE.US.82.MNE               ---------Q----VA--------------.-----G-V--------------------   881
MNE.US.x.MNE027.U79412      ---------Q----VA--------------.-----G-V------------E-------   883
SMM.SL.92.SL92B.AF334679    -IR---N--I--C-EA----Q-AIV----L.I----G-V--G-A------------M-N-   887
SMM.US.x.F236               ------R--Q-----W-KL-R----S----.I----G-V--G------------------   886
SMM.US.x.H9.M80194          ------R-LQ--AXXW-KFXR----S--R-.-----G-V--G----XX------------   887
SMM.US.x.PBJ14              ------R-LQ--A--W-KLVR----S--RH.-----G-V--G------------------   890
SMM.US.x.PBJA.M31325        ------R-LQ--A--W-KFVR----S--R-.-----G-V--G------------------   890
SMM.US.x.PBJ                ------R-LQ--AXXW-KFXR----S--R-.-----G-V--G----XX------------   892
SMM.US.x.PBJ                ------R-LQ--A--W-KFVR----S--R-.-----G-V--G------------------   890
SMM.US.x.PGM53.AF077017     -----CRW-Q-----A--ARG----S--R-.S----G-V---V------------A---   886
SMM.US.x.SME543.U72748      ---------Q--A--W-KF-R----S--R-.I----G-V--G-------V------A---   888
STM.US.x.STM.M83293         -----CIWIQ--A--A--A-G----S-GR-.-----G-V--R-G----------------   883
```

```
                           Nef start    R17Y mutation                                        premature stop in Mac239                                     max HIV-1 similarity
MAC.US.x.239               MGGAISMRRSRPSGDLRQRLLRARGET.YGRLLGEVEDGYSQSPGGLDKGLSSLSCEGQK..YNQGQYMNTPWRN.PAEEREKLAYRKQNMDDIDE*DDD.LVGVSVR.PKVPLRTMSYKLAIDMSHFIKEKGGLEGIYYSARRHRILDIYLEKEEGIIPDWQDYTSGPG  163
H2A.IN.NNVA.NEW            --ASG-KKH---PRG--E------AGA.C-GRCDALGGE--RFQEESGR-Q--P----RQNHQGEYADRPWRN#PSTERPNDWYGEQNKfDV-S-DD-...----P-T.SRTQ--E-T--F------L--D------M---E--RE-I-L------------N------  163
H2A.GM.x.MCR35.AY509260    --ASG-KE-----QG--E---------.C-GHCSGSGGE-L--QEESGREQR-P-Y---R..-Q--DF----*-T.--K-G--EL-KQ--R--V-SD---.--K--T.-R----E-THR-----------------M---E----------K----------N--H---  163
H2A.IN.95.CRIK             --ASG-KKH-K--RG--E-----QA--.CVGQWK-LGGE----QE-S---QR-P-----Q..-Q--DF------T.--T-KG-DL--Q-----V-SD---.-----T.-R------T-RM------L---------MF--E-L-#........................  139
H2A.CI.88.UC2.U38293       --ASG-KKL-KQ-QG--E-------GG.-V-QCSASGGES---Q--SGREQK-P-----Q..-R--D-------T.--M-GQ-ES-KQ-----V-SD-N-.Q-----T.SR----P-T----V--------------MF--RD------L-------V-----N--H---  164
H2A.DE.x.BEN.M30502        --ASG-KKL-KH-RG--E-------DG.--KQRDASGGE---FQEESGREQN-P-----Q..-Q--E---S----.--T--Q-DL--Q-----V-SD---.-I--P-T.-R--R-E-T-----------------Q-MF--R----------------------N--H---  164
H2A.DE.x.PEI2.U22047       --ASG-KKC--SLQG--E---------.C-GQWDGSAGE-L-FQE-SGR-QNLP-----R..-Q--DF------T.--AG--GTL-KQ-----V-ADN-N.-I--P-T.-R----A-T----V-I---LN-----D-M---E---------M-----------N--H---  164
H2A.GH.x.GH1.M30895        --ASG-KKH-KH-QR--E-----H-GG.-VQQCNASGGE----QE-SG--QK-P-----Q..-R--DF------T.--I-GQ-KL-KQ-------SS---.----P-T.-R----A-T----V-------KR.--D-MF--RD------L-------------N--H---  163
H2A.GM.87.D194.J04542      --ASG-KK--EH-QG--E-------GG.-VKQRNASGGES---QE-SGREQK-P-----Q..-Q--EF------T.--AIGQ-NS-KQ-----V-SD---.----P-M.-R----E-T----------------------F--RE------LF------------N--H---  164
H2A.GM.x.ISY.J04498        --ASG-KK-----RG-QE-------GA.C-G-WD-S-G----FHE-SGREQKLP-----R..-Q--DF------T.--T-K--ES--Q-----V-SD---.-----DT.SR----A-T-RM-V---DL--D-------M---E-------------------------N--H-L-  164
H2A.GM.x.MCN13.AY509259    --ASG-KE-----QG--E---------.C-GHCSGSGGE-L--QEESGREQR-P-Y---R..-Q--DF------T.--K-G--EL-KQ--R--V-SD---.------T.-R----ELTHR----V-------------M---E-----------------A---N--H---  164
H2A.GW.86.FG.J03654        --ASG-KK--K-LQG-QE---Q-----.C-GRCN-SGG--L--HE-SGREQN-P-----R..-Q--DFV-----T.--A----EL-KQ-----V-LD---.Q--FP-T.-R----P-TF---V---------------LF--Q-------L--D-................  148
H2A.GW.87.CAM2CG.D00835    --ASG-KK----LQG-QE------AG-.C-ECYNAL-GESLR-QE-S-REQN--------R..-Q--DF------A.--A-GK-N---Q-------SD---.----PAT.-R-------T----V---------------LF--E-----------------A---N------  164
H2A.GW.x.ALI.AF082339      --ASG-KK--G-LQG--E---QTP---.C-GQCSGSGG-----Q--SGR-QKLP-----R..-Q--DF------T.--T----EL-KQ-----V-LD---S------T.-R-Q----T----V----L---R-----MF--E-----------------------N--H---  165
H2A.GW.x.MDS.Z48731        --ASG-KK-----RG-QE------AGA.C-GHWD-LGGE--R-QE-S---QR-P-----Q..-Q--DF------T.--T-G--DS--Q-----V-SD---.---Y--T.-RT---A-T----V---------------MF--E-----------------------N--H---  164
H2A.SN.85.ROD.M15390       --ASG-KKH---PRG-QE------AGA.C-GYWN-SGGE--RFQE-S-REQK-P----RQ..-Q--DF-----KD.--A----NL--Q-----V-SD---.Q-R---T.------P-THR-------L--TR-----MF--E---K--N----------A---N--H---  164
H2B.CI.88.UC1.L07625       --S-G-KKQ-KQQRG--E----TQE-P.--K-SEGQRKQS-R----S--D-N-P----R..............NA.-RA-GGG....Q-DT--..SDE-N.E---Y--.-NR---S-T--M-------------------E-------T---N----VSG--N--Y---  146
H2B.CI.x.EHO.U27200        --S-G-KKQ-KQQPG--E------RGP.R-ESS--RQERSL-Y---S----N-P--D--..............KT.LGA-GGG......--DS-E..DDE-N.E---R--.-G----P-TF---V-----L----E----F--E---K---T---N----VSG--N--H---  146
H2B.GH.86.D205.X61240      --S-G-KK--ERQQG--EK---VPERP.----S--RREQS-R---ES--D-N-P-----..............NA.RGA-GGG....Q-DA-E..SDE--.E--AICKT-I----P-T----V--------Q-----M---E-------T-F-N----VSG--N--H---  147
H2B.JP.01.KR020.AB100245   --S-G-KKQ-KQRRG--E---QTQEGP.----SE-QRGQS-R----S--DSN-P----R..............SA.-GAQGGG....Q-GIE-..SDE-N.E---Y--.-NR---P-T---------W-------------E----L--T---N----VSG--N--H---  146
H2G.CI.92.ABT96.AF208027   --S-G-KK-APQQLG--K---Q----P.--K-WEGL-E----YREESG-----------Q..-T---F--------.--T--A-----X--N--V-SD-N-.----P-Y.------V-------------------------E---------M---Q-------N--A---  164
H2U.FR.96.12034.AY530889   ---S--KKH-KR--K--E---A----N.----W-GL---FELYQ--SG-D---RF--A-GQG-SE-EF------T.--TGKD--Q-KQ-----V--E-N-.----A--.-R-A----T----V------------D--F--Q-----------N---------N------  166
H2AB.CI.90.7312A.L36874    --S-G-KKQ-KQQRG--E-----Q-GP.----S-GRRGES----EES-R--N-P-----..............TA.-GA-GGG....Q-DIEE..DDE-N.E---Y--.-Q----P-T---------------------F--E-------T-F-N-R--VGG--N--Y---  146
H2AB.CM.03CM_510_03.NEW    --S-G-KK--EQQRG--E----V-ERP.--K-S-GRRERS-R--EES--D-N-P----..............NA.LGA-GGG....Q--A-ES--D-E..--AY--.-SR---P-T--T-------L------D-----E-------T-F-N----VSG--N--H---  145
MAC.US.x.MAC239            ---------------------------.--------------------------------..-------------.--------------------E---.-------.--G------------------------------------------------------L---  164
MAC.US.x.251               ---T-------ST--------------.-E---------S---L-E--------------..---E---------.--------------------E---.-------.-------------------------------------------------------------  164
MAC.US.x.251_32H           ------R---KSA--------------.---------------L----------------..-------------.-----------------V--E---.----P-M.-R-----------V--------------------------M--------------------  164
MAC.US.x.251               ----------K-A-----K--------.-----------S---L---G-----R------..-------------.----K---------------E---.-------.------A-T-------------------------------M--------------------  164
MAC.US.x.17EC1.AY033233    ---------------------------.--------------------------------..-------------.--------------------E---.-------.-------------------------------------------------------------  164
MAC.US.x.17EFR.AY033146    -----------S---------------.--------------------------------..-------------.--K-----------------Q---.-------.---------------I------------------------------K--------------  164
MAC.US.x.1937.AY611495     ---------------------------.--------G-----------------------..-------------.--------------------EX--.-------.T-----A------------------------------------------------------  164
MAC.US.x.2065.AY611493     -----XX-----------------X-X.--------X-----------------------..-------------.--------------------E---.-------.-------------------------------------------------------------  164
MAC.US.x.80035.AY611486    ---------------------------.--------------------------------..---X---------.--------------------E---.-------.-------------------------------------------------------------  164
MAC.US.x.81035.AY599200    ---------------------------.--------------------------------..-------------.--------------------E---.-------.-----------M-------------------------------------------------  164
MAC.US.x.85013.AY611490    -----------X---------X-----.--------------------------------..-------------.--V-----------------E---.-------.-------------------------------------------------------------  164
MAC.US.x.92050.AY603959    ---------------------------.-----------------E--------------..-------------.--------------------E---.-------.-------------------------------------------------------------  164
MAC.US.x.92077.AY599201    ---------------------------.--------------------------------..-------------.--------------------E---.-------.-------------------------------------------------------------  164
MAC.US.x.93057.AY611492    ---------------------------.--X----------------X--X---------..---E---------.--V-----------------E---.-------.-------------------------------------------------X-----------  164
MAC.US.x.93062.AY607704    --A---R--------------------.C-------------------------------..-------------.--V-----------------E---.-------.-------------------------------------------------V-----------  164
MAC.US.x.95058.AY611494    ---------------------------.--------------------------------..------------X.--------------------E---.-------.------A------------------------------------------X-----------  164
MAC.US.x.95086.AY607703    ---------------------------.--------------------------------..-------------.--------------------E---.-------.-------------X-----------------------------------------------  164
MAC.US.x.95112.AY588946    ---------------------------.--------------------------------..-------------.--------------------X---.-------.-------------------------------------------------------------  164
MAC.US.x.96016.AY607701    ---------------------------.--------------------------------..-------------.--------------------E---.-------.-------------------------------------------------------------  164
MAC.US.x.96020.AY611488    ---------------------------.--------------------------------..-------------.--------------------E---.-------.-------------------------------------------------------------  164
MAC.US.x.96072.AY611491    ---------------------------.--------------------------------..----E--------.---X----------------E---.-------.------------------------------------------------------E------  164
MAC.US.x.96081.AY597209    ---------------------------.-----------------X--------------..-------------.--------------------E---.-------.-------------------------------------------------------------  164
MAC.US.x.96093.AY611489    ------T-----X--------------.C------------------X-----X------..-------------.--------------------E---.-------.-------------------------------------------------V----E------  164
MAC.US.x.96114.AY588945    ------X--------------------.X-------------------------------..-------------.--------------------E---.-------.-------------------------------------------------------------  164
MAC.US.x.96123.AY611487    ---------------------------.--------------------------------..-------------.--------------------E---.------K.-------------------------------------------------------------  164
MAC.US.x.96135.AY607702    ---------------------------.--------------------------------..-------------.-----X--------------E---.------X.-X-----------------------------------------------------------  164
MAC.US.x.97009.AY599199    ------------------------X--.--------------------------------..-------------.--------------------E---.-------.-------------------------------------------------------------  164
MAC.US.x.97074.AY599198    ---X--X--------------------.--------------------------------..-------------.--------------------E---.-------.-------------------------------------------------------------  164
MAC.US.x.MM142.Y00277      ------KK--K-PR------------N.----FKG----S---L----------------..----E--------.-----K--P-----I-----E---.---IP-E.AR----------------------------------------------------I.H----  163
MAC.US.x.r80025.AY576480   ---------------------------.--------------------------------..-------------.--------------------E---.-------.-------------X-----------------------------------------------  164
MAC.US.x.r90131.AY576481   ---------------------------.--------------------------------..-------------.--------------------E---.-------.-------------------------------------------------------------  164
MAC.US.x.SMM142B.BD131285------KK--K-P#------------Kf----FKG----S---L----------------..----E--------.-----K--P-----I-----E---.---IP-E.AR----------------------------------------------------I.H----  162
MNE.US.82.MNE              ----T-K---KSP--------------.----WE-L-------L--S---S-L-------..-----F-----K-.--G-----------I-----E-N-.----P--.-R----II-----V------------------E---K---M--------M----N------  164
MNE.US.x.MNE027.U79412     ----T-KK--KSP--------------.----WEGL-------L--S--S--L-------..-----F-----K-.--G-----------------E--G.----P--.-R----A-T----V------------------E---K------------V-----------  164
SMM.SL.92.SL92B.AF334679   ---IT-SKPA-KH----KK--S---DS.----WDGL-E-SQ--Q--S-RA-N-CLTRDGYK.--V-EF-------.--T--Q-QQ--Q-C------D--E.-A--V-W.------A-------------------D-----E---------------------N------  165
SMM.US.x.F236              ------KKQYKRG-N--E---Q-----.----WEGL-E-----L-ASG---------P--..-SE----------.--T--A--G--Q-----V-DE---.-I----H.-R----A-T----------------------NE-------M-------------N------  164
SMM.US.x.H9.M80194         ---VT-KKQRXAG-N-XE---Q-----.----WEGL-GE----XDASG---------P--..-CE--F-------.--T-GA--G--Q-----V-NE..-.---CP-S.-R--V-I-T-------------------V---I---K---X---X---------N--A---  162
SMM.US.x.PBJ14             ---VT-KKQR-RG-N-YE---Q---D-.----WEGL-GE----QDASG---------P--..-CE--F-------.--T--A--D--Q-----V-SA..-.---CP-S.-R--V-I-T-----------------------D---K---L-------------N--A---  162
SMM.US.x.PBJA.M31325       ---VT-KKQR-RG-N-YE---Q-----.----WEGL-GE----QDASG---------P--..-CE--F-------.--T--A--D--Q-----V-SA..-.---CP-S.-R--V-I-T-----------------------D---K---L-------------N--A---  162
SMM.US.x.PBJ               ---VT-KKQRXAG-N-XE---Q-----.----WEGL-GE----XDASG---------P--..-CE--F-------.--T--A--D--Q-----V-SA..-.---CP-S.-R--V-I-T-------------------V---I---K---X---X---------N--A---  162
SMM.US.x.PBJ               ---VT-KKQR-RG-N-YE---Q-----.----WEGL-GE----QDASG---------P--..-CE--F-------.--T--A--D--Q-----V-SA..-.---CP-S.-R--V-I-T-----------------------D---K---L-------------N--A---  162
SMM.US.x.PGM53.AF077017    --A-G-KKQ--QR-G-GEK--Q-----.H-K-WEGL-------R-E-GRDWNLH-F---G..-SE--F-------.--R-----K--Q-----V-DD--E.-I----H.------A-------------------------E----------------T----N------  164
SMM.US.x.SME543.U72748     ------KKQH-RG-N--E---------.----WEGL-E-----L-ASG---------P--..-SE----------.-TA-KA--G-KQ-----V-DE---.------H.------A-T----------------------NEK------M-M-----------N--L---  164
STM.US.x.STM.M83293        --ASG-KKQRKQH-E--E---------.--K--EGLGE-SGP-Q-AS----N-H---P-R..--E--F-----K-.--A-SA--E--Q-----V--E--N.----A-H.-R----E-T------L-----S----------E-------M--------V----N--A---  164
```

```
                          max HIV-1 similarity                                                                       normal Nef end
MAC.US.x.239              IRYPKTFGWLWKLVPVNVSDEAQ....EDEEHYLMHPAQTSQWDDPWGEVLAWKFDPTLAYTYEAYVRYPEEFGSKSGLSEEEVRRRL..TARGLLNMADKKETR*  262
H2A.IN.NNVA.NEW           ----MFH-----------PQ-GE....DT-T-C--------KF--RH--T-V-R-------D-V-FRLH-----H----P---WKA--..K----PFN*RRQ-QLY  262
H2A.GM.x.MCR35.AY509260   ----MF----..........................................#----HE-K-FIL------H----P--DWKAK-..K---IPFS*fTGTAIF  219
H2A.IN.95.CRIK            .......................................................-FRL-L----H----P---WKA--..K---IPFS*ETGT-IL  178
H2A.CI.88.UC2.U38293      V---MC-----------M-Q--E....D--TNC--------RY--IH--T-V-R-NSM---E-K-FTL------H----P-K-WKAK-..K---IPYSE#TGT-IL  263
H2A.DE.x.BEN.M30502       V---MY--------S-EL-Q--E....---ANC-V------RH--EH--T-V-Q--SM---N-K-FTL------H----P-K-WKAK-..K---IPYSE#TGT-IL  263
H2A.DE.x.PEI2.U22047      V----F----------D-PQG.......E-D-C-L------GS---H--T-M-R---R---E-T-FN-------Y----P---WKAK-..K---IPFS*#TGTAIF  259
H2A.GH.x.GH1.M30895       V---MC----------D--Q--E....D--TN--T------RH--EH--T-L-R-------D-K-FILH-----H----P-K-WKAK-..K---IPYS*#TGTAIL  261
H2A.GM.87.D194.J04542     T---MY----------DI-Q--E....-V-TNC-V------RY--EH--T-V-R---M---S-K-FILH-----H----P-K-WKAK-..K---IPYSE#TGT-IL  263
H2A.GM.x.ISY.J04498       V---MF----------T-PQ-GE....DT-TLC---S--V-RF---H--T-V-----M--HE-TTFIL------H--ME-DDWKAK-..K---IPFS*#TGT-IL  262
H2A.GM.x.MCN13.AY509259   ----MF----------D-TRQEE....D-GT-C-L------RF---H--T-I--------HD-K-FILH-----H----P--DWKA--..K---IPFS*#TGTAIF  262
H2A.GW.86.FG.J03654       .......................................................-FTL------HN---P-K-WKA--..K---IPFS*#TGTAIL  186
H2A.GW.87.CAM2CG.D00835   V---MF----------DT-Q-GEDTETDT-T-C-L------RH--MH--T-V----SM--LK---FT-------H----P-D-WKAK-..K---IPFS*#AGT-IH  266
H2A.GW.x.ALI.AF082339     ----MF----------D-PQ-GE....DT-T-C-L--V---RH--TH--T-V-R---K--HD-K-FILH-----Y----P-D-WKA--..K---IPFSKNRNS*.F  263
H2A.GW.x.MDS.Z48731       T----C----------DLPQDEN.....T-T-C-V----I-KF--QH--T-V-R--SL---E-T-FKL------HQ---P---WKA--..K---IPFN*ETGTAIL  262
H2A.SN.85.ROD.M15390      V---MF----------D-PQ-GE....DT-T-C-V------KF---H--T-V-E---L---S---FI-------H----P---WKA--..K---IPFS.......-  257
H2B.CI.88.UC1.L07625      ----R-----------DIPE-ER....GA-TSC-V----I-S---IH--T---R---L--HD-V-FN-------YQ---P-K-*KA--..K---IPTD.......-  238
H2B.CI.x.EHO.U27200       V----F---------I-MIA-PE....DE-T-C-V------A----HE-T-V-Q--SL---D-V-FS-F-----YQ--MP-K-WKAK-..R---IPTE.......-  239
H2B.GH.86.D205.X61240     -----Y----------E-PAATR...E-E-T-C------I-S---IH--T-I-Q--SL---D-V-FN-F-----YQ---P---WKA--..K---IPTD.......-  241
H2B.JP.01.KR020.AB100245  V---------------IPE-DK....-A-TSC------A-P---VH--T-I-Q--SL---D-I-FN-------YQ---P-K-WKV--..K---IPTD.......-  239
H2G.CI.92.ABT96.AF208027  ----X-----------T------.......---T-C-V------P----ET-------------D-R-FILH----RW----P-AVWKEK-..KQ---PIE.......-  257
H2U.FR.96.12034.AY530889  T-W-MC-------E--DAN-K--.......---RL--VGS----CEE-H---A-V----SS---S-Q-FIKC--------------K---.-----PVKNC.....-  261
H2AB.CI.90.7312A.L36874   -----H----------E-AAVTR...E-E-T-C-V-----AA----H--T-V-Q--SL---SD--FN-F-----YQ---P-K-WKA--..K---IPTE.......-  240
H2AB.CM.03CM_510_03.NEW   -----Y----------ETPT--R...EGE-T-C--------S---YH--T-V-----L--HD-V-FN-F-----YQ---P-K-WKAK-..K---IPTD.......-  239
MAC.US.x.MAC239           --X--------------------....---X-X-------------------------------------------------------..----------------  264
MAC.US.x.251              -----------------------....-----------------------P-------------------------------------..----------------  264
MAC.US.x.251_32H          -----------------------....---R----Q-----K-------------------------------------P--------..----------------  264
MAC.US.x.251              -----------------------....---R----Q-----K------------------------A-----#---------------..-----------R----  263
MAC.US.x.17EC1.AY033233   -----------------------....-------------------------------------------------------------..----------------  264
MAC.US.x.17EFR.AY033146   --F--------------------....-------------------------------------------------------------..----------------  264
MAC.US.x.1937.AY611495    T----H----------D------....-----C-I--------------I-----------------------------P--------..A---------------  264
MAC.US.x.2065.AY611493    T----------------------....-----C-X-----------------------------------------------------..----------------  264
MAC.US.x.80035.AY611486   -----------------------....X--X---------------------------------------------------------..----------------  264
MAC.US.x.81035.AY599200   -----------------------....K------------------------------------------------------------..----------------  264
MAC.US.x.85013.AY611490   -----------------------....-------------------------------------------------------------..----------------  264
MAC.US.x.92050.AY603959   --F--------------------....-----C--------------------------X---------------------------..----------------  264
MAC.US.x.92077.AY599201   -----------------------....-------------------------------------------------------------..----------------  264
MAC.US.x.93057.AY611492   -----------------------....X--XX--------------------------------------------------------..----------------  264
MAC.US.x.93062.AY607704   -----------------------....---RX------------------------S-------------------------------..----------------  264
MAC.US.x.95058.AY611494   -----------------------....-------------------------------X-----------------------------..----------------  264
MAC.US.x.95086.AY607703   -----------------------....X------------------------------------------------------------..----------------  264
MAC.US.x.95112.AY588946   -----------------------....-------------------------------------------------------------..----------------  264
MAC.US.x.96016.AY607701   -----------------------....X--X---------------------------------------------------------..----------------  264
MAC.US.x.96020.AY611488   -----------------------....-------------------------------------------------------------..----------------  264
MAC.US.x.96072.AY611491   P----------------------....---R---T-------------------------------I---------------------..----------------  264
MAC.US.x.96081.AY597209   -----------------------....-------------------------------------------------------------..----------------  264
MAC.US.x.96093.AY611489   -----------------------....---RXX----X-------------------------X------------------------..----------------  264
MAC.US.x.96114.AY588945   -----------------------....X------------------------------------------------------------..----------------  264
MAC.US.x.96123.AY611487   -----------------------....-------------------------------------------------------------..----------------  264
MAC.US.x.96135.AY607702   -----------------------....------X------------------------------------------------------..----------------  264
MAC.US.x.97009.AY599199   -----------------------....-------------------------------------------------------------..----------------  264
MAC.US.x.97074.AY599198   T----------------------....-----X-I-----------------------------------------------------..----------------  264
MAC.US.x.MM142.Y00277     ---L-M-------I---------....-------V-------------------------------I--------------K--K---..A-----E---R---S-  263
MAC.US.x.r80025.AY576480  -----------------------....X------------------------------------------------------------..----------------  264
MAC.US.x.r90131.AY576481  -----------------------....-------------------------------------------------------------..----------------  264
MAC.US.x.SMM142B.BD131285 ---L-M-------I---------....-------V-------------------------------I--------------K--K---..A-----E---R---S-  262
MNE.US.82.MNE             P----------------------....-G--N--L----------------V--------------I---------------------..------K----R--S-  264
MNE.US.x.MNE027.U79412    P----------------------....-G--N--L----------------V--------------I---------------------..------K-------S-  264
SMM.SL.92.SL92B.AF334679  -----F----Q---D-----N....N--T-C-V-----Y-HS----K------N-H-------F--H-----W----T----E---ANKPKPQKK------X..  265
SMM.US.x.F236             ----MHY---------D------....---T-C-V-----Y----------------E---S-K-FIK----------------K---..------K#------S-  263
SMM.US.x.H9.M80194        ----MF-------X---------....---T-C------------------------X---X-K-F-EH------Q----K---Q---..------K-----K-S-  262
SMM.US.x.PBJ14            ----MF-----------------....---T--------------------------K---N-K-F-EH------Q----K---Q---..------K-----K-S-  262
SMM.US.x.PBJA.M31325      ----MF-----------------....---T--------------------------K---N-K-F-EH------Q----K---Q---..------K-----K-S-  262
SMM.US.x.PBJ              ----MF-------X---------....---T-C------------------------X---X-K-F-EH------Q----K---Q---..------K-----K-S-  262
SMM.US.x.PBJ              ----MF-----------------....---T--------------------------K---N-K-F-EH------Q----K---Q---..-----FK-----K-S-  262
SMM.US.x.PGM53.AF077017   ----MF----------SA-----....---T---V----I----------------SQ---R---FI-H---------------K---..------K-------S-  264
SMM.US.x.SME543.U72748    T---MY----------D------....---T-C-V-----H----------------E---S-K-FIK----------------K---..----IYK----R--S-  264
STM.US.x.STM.M83293       -----Q----------DM-N---...ED-GT---V-----H----------V-----L--H----F--H----------PK---E---..------K-------S-  265
```

VII

PLV Proteins

Contents

VII-1 Introduction **395**
VII-2 Sequences **396**
VII-3 Alignments **403**
VII-3.1 Gag 403
VII-3.2 Pol 407
VII-3.3 Vif 414
VII-3.4 Vpx 417
VII-3.5 Vpr 418
VII-3.6 Tat 419
VII-3.7 Rev 421
VII-3.8 Vpu 423
VII-3.9 Env 424
VII-3.10 Nef 431

VII-1 Introduction

The selection of Primate Lentivirus Protein sequences for the following alignments was based on the sequences in the complete genome alignment as a starting alignment, and complete or nearly complete genes from other isolates were added if they increased the diversity of samples represented. For example, several diverse African green monkey virus isolates have been sequenced only in a region of the *env* gene, and recent Gorilla sequences are only available in the pol and env genes. When necessary, some of the more common sequences (such as HIV-1 M group) were removed to make room on the Compendium pages for these diverse virus sequences. More complete alignments are available from our web site `http://www.hiv.lanl.gov/content/sequence/NEWALIGN/align.html` where space limitations are not an issue.

The annotation is mainly based on knowledge from HIV-1, and should therefore be taken with a grain of salt for HIV-2 and SIV sequences.

VII-2 Sequences

Sequences included in the PLV protein alignments.

Name	Accession	Proteins	Author	Reference
H1B.FR.83.HXB2	K03455	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	van Beveren, C.P.	*(in) Weiss,RL, Teich,N, Varmus,H and Coffin,J(Eds); RNA TUMOR VIRUSES, SECOND EDITION,* **2**, Vol 2: 1102-1123; Cold Spring Harbor Laboratory, Cold Spring Harbor (1985)
H1A1.UG.85.U455	M62320	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Oram, J.D.	*ARHR* **6**(9):1073-1078 (1990)
H1B.US.90.WEAU160	U21135	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Wei, X.	*Nature* **422**(6929):307-312 (2003)
H1C.ET.86.ETH2220	U46016	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Salminen, M.O.	*ARHR* **12**(14):1329-1339 (1996)
H1D.CD.84.84ZR085	U88822	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Gao, F.	*J Virol* **72**(7):5680-5698 (1998)
H1F1.BE.93.VI850	AF077336	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Laukkanen, T.	*Virology* **269**(1):95-104 (2000)
H1G.SE.93.SE6165	AF061642	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Carr, J.K.	*Virology* **247**(1):22-31 (1998)
H1H.CF.90.056	AF005496	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Gao, F.	*J Virol* **72**(7):5680-5698 (1998)
H1J.SE.93.SE7887	AF082394	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Laukkanen, T.	*ARHR* **15**(3):293-297 (1999)
H1K.CM.96.MP535	AJ249239	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Triques, K.	*ARHR* **16**(2):139-151 (2000)
H101_AE.TH.90.CM240	U54771	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Carr, J.K.	*J Virol* **70**(9):5935-5943 (1996)

Name	Accession	Proteins	Author	Reference
H102_AG.NG.x.IBNG	L39106	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Howard, T.M.	*ARHR* **10**(12):1755-1757 (1994)
H103_AB.RU.97.KAL153_2	AF193276	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Liitsola, K.	*ARHR* **16**(11):1047-1053 (2000)
H104_cpx.CY.94.CY032	AF049337	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Gao, F.	*J Virol* **72**(12):10234-10241 (1998)
H1N.CM.95.YBF30	AJ006022	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Simon, F.	*Nat Med* **4**(9):1032-1037 (1998)
H1O.BE.87.ANT70	L20587	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Vanden Haesevelde, M.	*J Virol* **68**(3):1586-1596 (1994)
H1O.CM.91.MVP5180	L20571	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Gurtler, L.G.	*J Virol* **68**(3):1581-1585 (1994)
CPZ.CD.90.ANT	U42720	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Vanden Haesevelde, M.M.	*Virology* **221**(2):346-350 (1996)
CPZ.CM.01.SIVcpzCAM13	AY169968	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Nerrienet, E.	*J Virol* **79**(2):1312-9 (2005)
CPZ.CM.05.SIVcpzEK505	DQ373065	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.CM.05.SIVcpzMB66	DQ373063	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.CM.98.CAM3	AF115393	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Corbet, S.	*J Virol* **74**(1):529-534 (2000)
CPZ.GA.88.GAB2	AF382828	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Bibollet-Ruche, F.	*ARHR* **20**(12):1377-1381 (2004)
CPZ.CM.98.CAM5	AJ271369	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Muller-Trutwin, M.C.	*J Med Primatol* **29**(3-4); 166-72 (2000)

Name	Accession	Proteins	Author	Reference
CPZ.CM.05.SIVcpzMT145	DQ373066	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.TZ.00.TAN1	AF447763	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Santiago, M.L.	*J Virol* **77**(3):2233-2242 (2003)
CPZ.CM.05.SIVcpzLB7	DQ373064	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.US.85.CPZUS	AF103818	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Gao, F.	*Nature* **397**(6718):436-441 (1999)
CPZ.GA.88.GAB1	X52154	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Huet, T.	*Nature* **345**(6273):356-359 (1990)
GSN.CM.99.CN166	AF468659	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Courgnaud, V.	*J Virol* **76**(16):8298-8309 (2002)
GSN.CM.99.CN71	AF468658	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Courgnaud, V.	*J Virol* **76**(16):8298-8309 (2002)
MUS_1.CM.01.1085	AY340700	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Courgnaud, V.	*J Virol* **77**(23):12523-12534 (2003)
MUS_1.CM.01.CM1239	EF070330	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
MUS_2.CM.01.CM1246	EF070329	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
MUS_2.CM.01.CM2500	EF070331	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
MAC.US.x.239	M33262	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Kestler, H.	*Science* **248**(4959):1109-1112 (1990)
H2A.GW.x.ALI	AF082339	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Azevedo-Pereira	Unpublished

Name	Accession	Proteins	Author	Reference
H2A.DE.x.BEN	M30502	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Kirchhoff, F.	*Virology* **177**(1):305-311 (1990)
H2A.SN.x.ST	M31113	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Kumar, P.	*J Virol* **64**(2):890-901 (1990)
H2B.GH.86.D205	X61240	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Dietrich, U.	*Nature* **342**(6252):948-950 (1989)
H2B.CI.x.EHO	U27200	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Rey-Cuille, M.A.	*Virology* **202**(1):471-476 (1994)
H2G.CI.92.ABT96	AF208027	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Brennan, C.A.	*ARHR* **13**(5):401-404 (1997)
H2U.FR.96.12034	AY530889	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Damond, F.	*ARHR* **20**(6):666-672 (2004)
MAC.US.x.251_BK28	M19499	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Hirsch, V.	*Cell* **49**(3):307-319 (1987)
MAC.US.x.251_1A11	M76764	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Marthas, M.L.	*J Med Primatol* **18**(3-4):311-9 (1989)
MAC.US.x.EMBL_3	Y00295	Gag, Pol, Vif, Tat, Rev, Env, Nef	Franchini, G.	*Nature* **328**(6130):539-543 (1987)
SMM.US.x.H9	M80194	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Courgnaud, V.	*J Virol* **66**(1):414-419 (1992)
SMM.US.x.PGM53	AF077017	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Novembre, F.J.	*J Virol* **72**(11):8841-8851 (1998)
SMM.US.x.SIVsmH635F_L3	DQ201172	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Kuwata, T.	*J Virol* **80**(3):1463-75 (2006)
SMM.SL.92.SL92B	AF334679	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Chen, Z.	*J Virol* **70**(6):3617-3627 (1996)

Name	Accession	Proteins	Author	Reference
STM.US.x.STM	M83293	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Novembre, F.J.	*Virology* **186**(2):783-787 (1992)
MNE.US.x.MNE027	U79412	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Kimata, J.T.	*J Virol* **72**(1):245-256 (1998)
GRV.ET.x.GRI_677	M66437	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Fomsgaard, A.	*Virology* **182**(1):397-402 (1991)
VER.KE.x.TYO1_patent	DJ048201	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Omori, T.	Patent: WO 2007049749-A 13 03-MAY-2007; DNAVEC CORPORATION
VER.KE.x.9063	L40990	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Hirsch, V.M.	*J Virol* **69**(2):955-967 (1995)
VER.KE.x.AGM155	M29975	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Johnson, P.R.	*J Virol* **64**(3):1086-1092 (1990)
VER.DE.x.AGM3	M30931	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Baier, M.	*Virology* **176**(1):216-221 (1990)
TAN.UG.x.TAN1	U58991	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Soares, M.A.	*Virology* **228**(2):394-399 (1997)
SAB.SN.x.SAB1C	U04005	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Jin, M.J.	*EMBO J* **13**(12):2935-2947 (1994)
OLC.CI.97.97CI12	FM165200	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
DRL.x.x.FAO	AY159321	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Hu, J.	*J Virol* **77**(8):4867-4880 (2003)
RCM.NG.x.NG411	AF349680	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	*J Virol* **75**(24):12014-12027 (2001)
RCM.GA.x.GAB1	AF382829	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
MND_1.GA.x.MNDGB1	M27470	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Tsujimoto, H.	*Nature* **341**(6242):539-541 (1989)
MND_2.x.x.5440	AY159322	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Hu, J.	*J Virol* **77**(8):4867-4880 (2003)

Name	Accession	Proteins	Author	Reference
MND_2.CM.98.CM16	AF367411	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Takehisa, J.	*ARHR* **17**(12):1143-1154 (2001)
MND_2.GA.x.M14	AF328295	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Souquiere, S.	*J Virol* **75**(15):7086-7096 (2001)
DEB.CM.99.CM40	AY523865	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
DEB.CM.99.CM5	AY523866	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
DEN.CD.x.CD1	AJ580407	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Dazza, M.C.	*J Virol* **79**(13):8560-8571 (2005)
TAL.CM.01.8023	AM182197	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	*Virology* **349**(1):55-65 (2006)
TAL.CM.00.266	AY655744	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	*Virology* **349**(1):55-65 (2006)
MON.CM.99.L1	AY340701	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Courgnaud, V.	*J Virol* **77**(23):12523-12534 (2003)
SUN.GA.98.L14	AF131870	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	*J Virol* **73**(9):7734-7744 (1999)
WRC.CI.97.97CI14	AM745105	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
WRC.CI.98.98CI04	AM713177	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
WRC.GM.05.Pbt_05GM_X02	AM937062	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Locatelli, S.	*Virology* **376**(1):90-100 (2008)
LST.CD.88.524	AF188116	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	*J Virol* **74**(8):3892-3898 (2000)
LST.CD.88.485	AF188115	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	*J Virol* **74**(8):3892-3898 (2000)
LST.CD.88.447	AF188114	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	*J Virol* **74**(8):3892-3898 (2000)
LST.KE.x.lho7	AF075269	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Hirsch, V.M.	*J Virol* **73**(2):1036-1045 (1999)

Name	Accession	Proteins	Author	Reference
SYK.KE.x.SYK173	L06042	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Hirsch, V.M.	*J Virol* **67**(3):1517-1528 (1993)
SYK.KE.x.KE51	AY523867	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
COL.CM.x.CGU1	AF301156	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Courgnaud, V.	*J Virol* **75**(2):857-866 (2001)
MON.NG.x.NG1	AJ549283	Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Barlow, K.L.	*J Virol* **77**(12):6879-6888 (2003)

```
                           membrane binding                                                                                                                                        phosphorylation site
                           Gag start, p17 start  nuclear localization                                                                                            nuclear localization                 p17 end  p24 start
H1B.FR.83.HXB2            MGARASV.LSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEEEQ.......................NKSKKKAQQAAADTG.HSNQVSQNYPIVQNIQGQMVH  144
H1A1.UG.85.U455           -------.---KK--S----------N---R---L--------K-TL-------A---Q--------A----T-----------V--------DV-------N----M-.......................--N-QRT-----N--.....S---------A---P--  140
H1B.US.90.WEAU160         -------.-------------------N-----------------------------------------------------V------K--V--------------.......................--------------E.NR-----------L------  144
H1C.ET.86.ETH2220         ------I.-R-EK--A----K-------H-M---L---N----K--L--D--D--A--K--IK----A----T---K--F----------K-------------.......................-E-QQ-T---G-ADR...GKD---------M------  142
H1D.CD.84.84ZR085         -------.----K--A----------------------------L----------K--ME----A-K--------I----------K-------------.......................-N---R-----G--.N-S-----------L------  144
H1F1.BE.93.VI850          ------I.----K--E----Q-------R--M--LI----------LD---------QK-IR-----------K--F---V--Y---AGVT--------L----.......................---QQ-T------K-....------------L------  140
H1G.SE.93.SE6165          -------.-T--K--A----------R-S--I--L-----------L--D----A---Q--MR---------T--IK----A-----------V-------EEV-KI-.......................K--QE-I----M-K-.N------------A------  144
H1H.CF.90.056             -------.----K--A--------------R---L-----------L-------P---L--IE-I--AIK--T---K--F-L--V-----RK-DV------------I-.......................---QQ-T------KE.KD-K----------A------  144
H1J.SE.93.SE7887          ------I.----K--D-------------RI--L-------D---L------SAK--Q---V----A----TQ-IK---------------------M---E----I-.......................--N-QQ--K-ET-KK.DNS----------L---P--  144
H1K.CM.96.MP535           -------.----K--A------------------L-----------L-------T------IT-I---I------IK----I-V--F---K--V---------L----.......................---QR-T--.E-A...DKG---------L------  140
H101_AE.TH.90.CM240       -------.----K--A----------R---R---L-----------L--SF---A---Q--IE---ST-K--L---K--F------W-------V-----------V-.......................---QR-T-----G--.S-SK---------A----A-  144
H102_AG.NG.x.IBNG         -------.----K--A--------------R---L-----------L-------A---Q-LME---SA-R------K--F--I---W------D---------M--I-.......................----Q-V--T--A--....SS---------AK---T-  141
H103_AB.RU.97.KAL153_2    -------.----K--A-----------E--RI--L-----------L--S--------Q---E----T-K------K-----------------------------I-.......................----Q-T----TG--.S-SK----------A----T-  144
H104_cpx.CY.94.CY032      -------.----K--A--R-----------R---L-----------L-------A---Q-LME---ST-K-----------IT--W------DVQ-----------I-.......................S---Q-T-----AA-.G-SN----------A------  144
H1N.CM.95.YBF30           -------.-T--K--Q--S-Y---------RM--L-----------C----MD-AD--AKL-N--E-A-K-----------AL-V-----S--Q-HN-Q------K-K-E.......................QHKPEPKNPE-GAAAATDSNI-R---L--TA------  146
H1O.BE.87.ANT70           ---S---.-T-SK--A--Q---K--S----R---L-----------C--E----A--NEKL-Q--E-A-K---DS-Q--W-AIVV-W---N-YK-G--QQ-IQ-LK-VMG.......................SRKSAD--KEDTSAR-AG------S-A------  142
H1O.CM.91.MVP5180         -------.-T-SK--A--R------S--A-R---L---------Y-C-------A--TE-L-Q--E-A-K----D-K--W-AI-V-W---N-FD-R--QQ-IQ-LK-VMA.......................SRKSAE--KEETSPR-T-------T-A------  142
CPZ.CD.90.ANT             ---G---.-R-EK--T--S-----------MI--L---RS--Q---LSSS--------EKAIH--S--IEIR-P-II--F--ICV-W---KGEK----EQ-VKTVKMKV........MQTQAETGSSQTASRGMLLRLLLLNKQWCQR-LSGEGR----IVDAG-IAR-  160
CPZ.CM.01.SIVcpzCAM13     -------.-T--K-----V------R---MM--L-----------C----M-SA---LKL-Q--E-A-K---G----F--L-V-W---SN--VE--QK--EQLKKHH.......................GEQVR-NENTENNESRQGASA-SGT-G-------A------  149
CPZ.CM.05.SIVcpzEK505     -------.-T--K--Q----Y---------MM--L-----------C----MD-A---A-L-R--E-A-K---G----F--L-V-----NN-KVQN-Q---E-LR-KM.......................KAEQKEPEPE---GAAAAPESSI-R---L---A------  147
CPZ.CM.05.SIVcpzMB66      -------.-T--K-------Y---------MM--L-----------C--S-M--T---K-L-Q--E-A-K---G----F--IVV-W----G-PV--------QLQ..........EAQQKGKQEVAAATADG.................TST--R-F---A-A------  142
CPZ.CM.98.CAM3            -------.-T--R--A--------------MM--L-------D---C----M--A---E-L-R--E-A-K---G----F--L-V-W---K--SVE--QQ--T-LK-AV.......................TASREQEV-QPQQQQQDSA--R---V---A---L--  144
CPZ.GA.88.GAB2            ------I.-T--R--A----------R-R-MM--------------CD---M--K---Y-L-K--E-A-K---G----F--L-V-W---A---VE--QQ--Q-LK-FVQK.......................ETATAQEA--KQSESNTTEI-R---V---A----I-  146
CPZ.CM.98.CAM5            -------.-T--R--A------X-------MM--L-----------C----M--A---ENL-K--E-A-R---G----F--L-V-W---K--QVE--QQ--S-LK-TV.......................AAQEQEV-QPQQQQQESSI-R-------A----I-  143
CPZ.CM.05.SIVcpzMT145     ------I.-T--R--A--------------MM--L-----------C----M-NA---EKL-K--E-A-K---G----F--L-V-W---K---VE--QQ--Q-LKAAVNP.......................EQGETPSA-A-F-V-R-A---LT-  134
CPZ.TZ.00.TAN1            -------.-R-DK--T--S---KSR-R---LI--L---GS--Q---M----M-NV---WK-IL-----VDI--P-II--F--ICV-----AGERVQ--E--VKIVKMKL..TVQKNNSTATSSGQRQNAGEKEETVPPSGNTGNTGRATETPSG-RL--VITDA--VAR-  167
CPZ.CM.05.SIVcpzLB7       ------I.-T--K--D---V------R-R-MM--L-----------C----M--T---E-LIT--E-A-K---N-K--F--IVV-W-------A---------TLTAVR.......................QKQQ-EKE---AAARDNDSG-R-----T-A---P--  144
CPZ.US.85.CPZUS          -------.-T--R--A--------------MM--L-----------C----M--AD--L-L-K--E-A-K---G----F--L-V-W---S-VTVE--QQ--V-LK-V.......................VQLQ-QQEEKE-QQQEASG.--IG-S---VI--A------  146
CPZ.GA.88.GAB1            -------.-T--K-----V------R-R-MM--L-----------CD---M-SK---TKL-Q--E-A-K---G----F--L-V-W-I-SD-TVE--QK--EQLKRHHGE.......................QQSKTESNSGSREGG-SQGASA-AGI-G---L---A------  152
GSN.CM.99.CN166           ----H-AM-T-TK---Y--V----K-----MI--L----K----Y-LSDA---NQ----R--EIVF-LEP----A-K--FGI-SV---I-AD-QVE--EQ-KQQVRIRCHLEE.......................T-G-E-QNKDPPGAAGGQAV------VIR-A---YQ-  150
GSN.CM.99.CN71            ----H-AM-T-TK---Y--VS---R-----AI--L----K----Y-LSGS---SK---Q-V-EVTS-LEPN---S-K---GI-SV---I-AD-WVE--EQ-KKQVQIRCHLVG.......................T-E-ETQNKDPPGAAGGQAV--------R-A---FQ-  150
MUS_1.CM.01.1085          ----H-AM-T-NK---Y--V----K-----LI---I------D--GLSDS---SRD--QK-IEVIL-LEPA---S-K--FGI-SV-W-I-AEVHVE--E--KKQ-RARCHLEE.......................KKGEKESQAQAA.-A---VIR-A---FQ-  141
MUS_1.CM.01.CM1239        ----H-AM-T-TK---Y--V----K---R-LI-------K-----GLSDS----K---QK--EV-I-LEP----S-K--FGITSV-W-I-AEV-VE--EQ-RKQ-RIRCHLAD.......................KQGEEKTAQ--T-A----IR-P---FQ-  141
MUS_2.CM.01.CM1246        ----H-AM-K-TK-NKY--V----K-----LI-------------GLSDT---SKD--QR--EVIC-LEE----A-K--FGI-SV-W-I-ADVTV---E--KKQ-RIRCHLAE.......................-KEDSKG-S-AKATTSG----VIRTA---YQ-  145
MUS_2.CM.01.CM2500        ----H-AM-T-TK--KY--V----K---R-LI-------K--D--GLSDS----K---QR--EVIL-LEP----S-K--FGI-SV---I-AEVSVE--E--KKQ-RMRCHLGE.......................--E--EQR-QPSGAGAS-----VIRTA--TYQ-  146
                          Gag p15 start                                                                                                                                                                  p15  p27
MAC.US.x.239              --V-N--.---KKA-EL-------N-----M---V---AN--D--GLAES---NK---QK--SV-A-LVP----N-K------CVIW-I-AEEKV-H-E--K.Q-VQRHLVVE.......................TGTTETMPKTSRPTAP-SGRGG---VQ-IG.-NY--  146
H2A.GW.x.ALI              ----N--.-R-RKA-EL-R-----------Q-------AN--D--GLAES---SK---QR--KV-E-LVP----N-K--F---CVVW---AEEKV---EG-K.Q-IQRHLAAE.......................IETAEKMPSTSRPTAPPSEQGG-F-VQ-VA.-NYT-  146
H2A.DE.x.BEN              ----N--.-R-KKA-EL--V----------R-------AN--DK-GLAES---SK---QK--RV-D-LVP----N-K--F---CVIW-L-AEEKV---E--K.-LAQRHLVAE.......................TGTAEKMPNTSRPTAPPSGKRG---VQ-AG.-NY--  146
H2A.SN.x.ST               ----N--.-R-KKA-EL-------------R-------AN--D--GLAES---SK---QK--TV-D-LVP----N-K--F---CVIW-I-AEEKA---E--K.QKVQRHLVAE.......................TKTTEKMPSTSRPTAPPSGNGG-F-VQ-VA.-NYT-  146
H2B.GH.86.D205            ----G--.---KKT-EL--V----------M---V---VN--D--GLAES---SK---QK--KV-A-LVP----N-K--F-I-CVIF-L-AEEKV---E--K.--AQRHLAAD..........................TEKMP-TNKPTAPPSGG.----VQ-LA.-NY--  141
H2B.CI.x.EHO              ----G--.---KKT-EL--V--------R-M-------VN-----GLAESR-GSK----K-RKV-G-LVP----N-K------CVIF-L-AEEKV---E--K.--AQRHLAAD..........................TEKMP-MSKPSKPTSRL..A--VQ-IA.-NYS-  141
H2G.CI.92.ABT96           ----S--.---KKA-EL--V----------M---V---AN--D--GLAES---SK---QK--SV-A-LVP----N-K------CVIW-L-AEQKV-H-E--K.QVVQRHLVVE.......................TGTTEKVP-TSRPIAPPSGRGG---VQ-VG.-NY--  146
H2U.FR.96.12034           ----N#-.---KKV-EL--V----N-----M---V---AN--D--GLAES---AK---QK--SV-A-LVP----N-K------CVIW-L-AEEKV-H-E--K.Q-VQRHLAVE.......................TETAEKMP-TSRPTXPPSGGRG---VQ-IG.-NYA-  145
MAC.US.x.251_BK28         ----N--.---KKA-EL-------------M---V---AN--D--GLAES---NK---QK--SV-A-LVP----N-K------CVIW-I-AEEKV-H-E--K.Q-VQRHLVVE.......................TGTAETMPKTSRPTAP-SGRGG---VQ-IG.-NY--  146
MAC.US.x.251_1A11         ----N--.---KKA-EL-------N-----M---V---AN--D--GLAES---NK---QK--SV-A-LVP----N-K------CVIW-I-AEEKV-H-E--K.Q-VQRHLVVE.......................TGTAEIMPKTSRPTAP-SGRGG---VQ-IG.-NY--  146
MAC.US.x.EMBL_3           ----NA-.---KKA-EL-------------M---V---AN--D--GLAES---NK---QK--SV-A-LVP----N-K------CVIW-I-AEEKV-H-E--K.Q-VQRHLVVE.......................TGTAETMPKTSRPTAP-SGRGG---VQ-IG.-NY--  146
SMM.US.x.H9               --V-N--.---KKA-EL-------------M-------AN--D--GLAES---NK---QK--SV-A-LVP----N-K------RV-W-I-AEEKV-H-E--K.Q-VQRHLVVE.......................TGTADKMP-TSRPTAPPSGRGG---VQ-VG.-NYT-  146
SMM.US.x.PGM53            ----N--.---KKA-EL------N-----M---V---AN--D--GLAES----K---QK-ISV-A-LVPA---N-K------CVIW-I-AEEKV-H-E--K.Q-VQRHLVVE.......................TGTADKMPITSRPTAPPSGRGG---VQ-IG.-NYT-  146
SMM.US.x.SIVsmH635F_L3    ----N--.---KKT-EL------N-----M---V---AN--D--GLAES--DNK---QK--SV-A-LVP----N-K-----ICVIW-I-AEEKV-H-E--K.Q-VQRHLVVE.......................TGTADKMP-TSRPTAPPSGRGG---VQ-VG.-NY--  146
SMM.SL.92.SL92B           ----G--.---KKA-EL--V------R---M----I--A---D--GSAES---SK---QR--AV-A-LMP----N-K--FS--CVVW-L-AEMKV---E--K.-TVQSHLVVE.......................SGTAEKLP-QSRPTAPPSGG..---VQ-VG.NNY--  144
STM.US.x.STM              ----S--.---KKA-EL--V----------M---V---AN--D--GLAES---SK---QK-ITV-E-LVP----N-K--F---CVIW-I-AEEKV-H-E--K.QVVKRHLVVE.......................TGTANKMP-TSRPTAPPSGRGG---VQ-VG.-NY--  146
MNE.US.x.MNE027           ----N--.---KKA-EL-------------M---V---AN--D--GLADS---NK---QK--SA-A-LVP----N-K------CVIW-I-AEEKV-H-E--K.Q-VQRHLVVE.......................TGTAETMPKTSRPTAP-SGKRG---VQ-IG.-NYT-  146
GRV.ET.x.GRI_677          --GGH-A.---RS--TF-------N-----QI--LI--GK-M---GLHEK----K---QK-IEV-T-LEP----G-KA-F-LCCVIW-I-AEQKV---E--VVTVKQHYHLVD...............KNEKAA.......KKKNE-TAPPGGE-R---V-NQN.NAW--  146
VER.KE.x.TYO1_patent      ---AT-A.-NRRQ--QF-------N-----QI--LI--GK-M---GLHER----E---KR-IEV-Y-LEP----G-K--F-L-CV---L-KEQKV---E--VATVRQHCHLVE...............KEKSATETSSGQKKNDKGIAAPPGG.---F-AQ-QG.NAW--  152
VER.KE.x.9063             ---AT-A.-K-RQ--QF-------N-----QI--LI--GK-M---GLHDR----ED--KK-IEV-F-LEP----G-K--F-L-CV---I-REQKV---E--VATVRQCCHLVE...............KEKRATEPSSGEKKSNRE-TAPPGG.---F-AQ-QG.NAW--  152
VER.KE.x.AGM155           ---AT-A.-NRRQ--EF-H-----N-----QI--LI--GKKMD--GLHEK----E---KK-IEV-S-LEP----GMK----L-CV-L----EKKV---E---AIVRQCCHLVD...............KEKTAVTPPGGQ-KNNTGGTATPGG.---F-AQ-QG.NAW--  152
VER.DE.x.AGM3             ---AT-A.-NRRQ--KF-H-----T-----QI--LI--GK-M---GLHER---SE---KK-IEV-Y-LEP----G-K--F-L-CV-F---KDK-V---E--VAIVRQCCHLVEKERNAERNT............TETSSG-KKNDKGVTVPPGG---F-AQ-QG.NAWI-  156
TAN.UG.x.TAN1             ---GH-A.---RN--TF-------N-----Q---LI--GKKM---GLHEK----K---QK-IEV-S-LEP----G-K--F-LCCVIW-I-AEQKV---E--VVLVKQRCHLVE...............KEKTAAAPSGGQ--NYNTAATP-GR.HG---V--QN.N-W--  152
SAB.SN.x.SAB1C            ---SN--.---RK--AF-SV----N-------R-L----K--D--SLSAN----K--VVK--SV-L-LVP----N-IA-F-LCCV-A-I-AE-KV---E--KA-VK--VPAEMTESATATSSGQTK...........ELQAKKKNEPTVTPSGG-R-----SVN.N-W--  157
OLC.CI.97.97CI12          --GTE-K.IDPEL-KKL-RV---ET---C-RRA-VK--VQ--KHLCID-DD-G-K---KR-VEK-L-LEEG-TQK-K--L-ICIVIL-LQEGLPVRN-I--KKRVKVKE.......................................KKEEQLAM--IR-A--IFE-  128
DRL.x.x.FAO               ---S--G.-R-EK--AF--V----S---------L--VAK--D--GLHER---SQ---QK---V-F-LVP----N-I--F--CCVVW---SKVKVT--E--VQ-VKQHCHLVD.......................-NENA-SKNENGETAT-SGR-R----QVVN.Q-A--  147
RCM.NG.x.NG411            ------I.---KK--A--QV---------M---L---C------GLSDT---SA--IQK-I-VIL-LVP----G-K--F-LCCV-W---KEVKV---E--VAHVKKQCHLVDKDENAGEQEKGAT......................VT-SGQRG-----TIN.Q-PE-  146
RCM.GA.x.GAB1             ------L.---KK--A--SV---------M---L---CKK-NK-GLSDH----AT--EK---V-L-LVP----G-K--F-LCCV-W---KEVKV---E--VA-VK-CCHLVEKAENTTEKEKGAT......................APPSGQRG----ITIN.Q-PE-  146
MND_1.GA.x.MNDGB1         --NGN-A.-L-TD--KF-----KR----C-R---LC-CKG--D--GLSDK----QQ--EK--SVCW-LYDQ--DN-KA-VG--CVVA-I-AG----S-QD--K-LKVITRKEEKQEDE..............................-K-F-VQRDAA--YQY  133
MND_2.x.x.5440            ---S--G.-R-EK--EL-------S-----Q---VI-V-K--D--GLHEK---SQ---EK--SV-F-LVP----N-I-----CCCIW---AKVKVT--E--KE-VKQRYHLVVERENAASEEEKGA.....................TATPAVR-K----QVIN.QTP--  147
MND_2.CM.98.CM16          ---S--G.-R-EK--EL-------S-----Q----I-V-K--D--GLHEK---SK---EK--SV-F-LVP----N-I-----CCCVW---AKEKVT--E--KE-VKQKLHLVAEKENAASEKEQR.....................AIVTPSGR-K----QIIN.QTP--  147
MND_2.GA.x.M14            ---S--G.-R-EK--EL-T-----S-----Q----I-V-K--D--GLHEK----Q---EK--SV-F-LVP----N-LA----CCCIW---AKVKVT--E--KK-V-QCFHLAAKGESAASEKEEK.....................ATATPSGR-K----QIIN.QTP--  147
DEB.CM.99.CM40            ---GR-AP-T-AA---F--V----K-----MI--L---GK--D--GLAAA---NK---QE--TR---LEA----N-K---GI-CVIWAC-ANLSVD--E--KKAVYRRIKQVEQEEELEMALIQRQKEKEVEQKQQ-QQE-RQQEPQAAPQTTP-V----LRQG.Q-F--  169
DEB.CM.99.CM5             ---GR-AP-T-AA---F-------K-----LI--L-------D--GLAGA----K---QE--TR---LEAS---G-K---GI-CVVWAC-ANLSVE--E--KKAVARRIKQAE....QEEELEMALIQKKAEKELE--QKT----ADTQQKTNQPA-F--LRQG.Q-Y--  165
DEN.CD.x.CD1              --SGS--PMT-AA---F--V-----S----MI--LI-SKK-M---GLGEQ----Y---QT-IET-A-LEPQ---S-K--FR-CCVVW-C-RGL-VA--E---REVRKRIGAKKKNIVIGNADQGDDQPPQGAAGGGSSAE-G-S-GGSSFRDPN-------RTA---YQ-  170
TAL.CM.01.8023            ----H-AM-T-RA--KY--V----K-----EVR-V----K-M---GLSDA----K---VK-IEAIS-LEPN---GIK--FGI-CV---I-AG---E--EQ-KAEVKKRCHPV.................EKKEESKGTTE-G-VRSAIVT-SGH-G----IR-A-Q-YQ-  153
TAL.CM.00.266             ----H-AM-T-RA--KY--V----K-----MV-------K-M---GLSDA----K---VK-IEAIYMLEPN---GIK--FGI-CV---I-AG-D-E--EQ-KAEVRKRCHLGEKKETTEEEK....................VRKATVT-SGQTG----IR-A-Q-YQ-  149
MON.CM.99.L1              ----H-AM---TK--KY--V----R-----LI------AK--D--GLSDS----QD--KK--EVIL-LQP----SIK--FGIASV---I-AG---E--EQ-KQQVKIRCHLA.....................GEQGEQKA---AAPPTGGVP-G---V-RTQG-GFQ-  148
SUN.GA.98.L14             ---GN-T.VDREVVRSF-RVA-K-----T-QAR-V---GK--D--GL-KE--R-V---QK--SVCW-LYAS--KN-KA-VG--CV-A-C--G-PV---Q---K-VRL-PAKNEEAQA...............................---F-VQREG.QNYI-  131
WRC.CI.97.97CI14          --SGN-IS-APHICEGLQRVH--KDS----QK-----LVK---KLCLDSA-IS-A--VKKLMSH-A-L-PDA-K-VK-IWGMLCV-A-L-E--PCRN-Q--KE-VNV--KKP...................................---M-VLRTA.QGPQW  127
WRC.CI.98.98CI04          --SGN--S-APHICEGLQ-VH--K-S----QK--VI-LVK---KLCLDST-IA-ET-TKKLMEH-A-LVPD--K-VK--WGMMCV-A-L-A--PCTN-Q--KE-V-VKDTKQKEKP.............................---M-VLRTA.QGPQW  131
WRC.GM.05.Pbt_05GM_X02    --SGN-IS-APEICESLQRV---KDS--T--R--L--LVK---KLCLDGK-ISSQ--VKKLCEHIA-LVPE--KTI--IWGMISV-A-LYNN-QVS--QA-IK---VRPREEKA...............................-V-M-VLRGP.QGPQW  129
LST.CD.88.524             --SGN--.--RQIEKDFCNV-----S--T-QKR-VD--TK--D--GLGAN----A---KK---VRW-LYP---KN-KA-VG--CVIC-C-LG-RVN--Q--IN-VKI--NK.................................PVPE-K-F-LIREN.QRW--  129
LST.CD.88.485             --SGN--.--RQIEGEFCS------S--T-QKR-----TK--D--GLGAH----AD--KR---VCW-LYS---KN-KA-VG--CV-C-C-LG-K-AN-A--MQ--KI-PTQE.................................DRQ-K-G---LIREN.QRW--  131
LST.CD.88.447             --SGN--.--RQIEGEFCS-----DS--T-QKR-L---T---D--GLGAH----AD--KK---VCW-LYS---KN-KA-VG--CV-C-C-LG-KVA--Q--MQ--KI-PTQE.................................ERQPK-G---L-REN.QRW--  131
LST.KE.x.lho7             --SGN--.--RQIEKDFCSV-----S--T-QKR-VE--TK--D--GLGSQ----A---KK--SVCW-LYA---KN-KA-VG--CVIC-C-LGVR-S--Q--IS-VKI-PAPEAAGKK.............................Q-TGG---LIREN.QRW--  134
SYK.KE.x.SYK173           ---AG-AI-T-R----Y-------K---R-LVR-L---KK--D--GLSDQ-M-SK---EK--TV-L-LEAN---N-K--FGIISVVWA--AKK-VE--EQ-KQ-VK-ACNWKDDPPATSGGQSENSSQNM.ASET-SGQKVVQQEKQKAATPPPRG---LLR-P-N-WI-  169
SYK.KE.x.KE51             ---GG-AI-T-RK--KY-------K---R-LVR-L---GK--D--GLSDQ---SK---E--IKTIL-LEKH---N-K--FGIT-VVWA--AEKDVE--EQ-KQ-VK-ACNWEDEETVTSSGQKENSNDTA....T-SGREGKMQLPAAMP-SGG-G---LIR-P-N-WI-  166
COL.CM.x.CGU1             --NEQGL.-GKKT-EDLQ-V--KK-K-GC--I--VR-MCT-VS-CVLIFE--KSAT-VA---EKVT-LVD----V-----G.CSVC--L-RKWN-E--Q--EK-V--AYKKQAMIEMASKEEEKAK.......................KEAEKLDM---TGP.QGP--  143
```

```
                                                                                          CyPA binding                                                             major homology region
H1B.FR.83.HXB2               QAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVH.AGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTN.....NPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQ 308
H1A1.UG.85.U455              --L-------------D-----------------------------M---V------------D-----------L----.----P------------------V--------G.....--------D--R----------------V--------------------------F------- 304
H1B.US.90.WEAU160            --L-----------I------------------------------------------------------------L----.-----------------------------A----......------------------------------V---------------------------------- 308
H1C.ET.86.ETH2220            -P--A-------------------------T---------------------------------D-----------L----.---V------D--------------------A---G.....---V---D------------------------V-----K------------------F------- 306
H1D.CD.84.84ZR085            --L-----------I-----N------------------------------------------------------L---Q.---V----------------------------S.....--------------------------------V---------------------------------- 308
H1F1.BE.93.VI850             -SL-----------I-----------------------T---------------------D-----------L----.---AP---------------------Q---G.....---V---D------------------------V-----K------------------F-V----- 304
H1G.SE.93.SE6165             ---T-------------------------------------L----------------D-----------M--QQ.---FP---I---------S-----T---G.....----------------------------------V----------------------------F-C----- 308
H1H.CF.90.056                -----------------------------------------A----------------D----------------.----P----------------------A---G.....--A----D------------------------V-----K------------------F------- 308
H1J.SE.93.SE7887             --L-----------I-------------------------------I-----------D----------------.--------V------------------------G.....------------------------------V-----------------------------F-A----- 308
H1K.CM.96.MP535              --L-----------I---------------T-----------------------------D-----D-------L----.----P----------------------A---S.....---V------------------------------V------------------------------F------- 304
H101_AE.TH.90.CM240          -PL----------------G-N-------------------M---I--------------------P----------.----P----------------------------.....-------D------------------------V---------------------------------- 308
H102_AG.NG.x.IBNG            -SM-----------I---G------------------------M---I----------D----------------.----P----------------------------S.....----------------V-------------V------------------------------F------- 305
H103_AB.RU.97.KAL153_2       -SM-----------I--------------------------M---I----------D-----------L--AQ.---FP--------------SS-----------S.....-------D------------------------V------------------------------F------- 308
H104_cpx.CY.94.CY032         -S------------I--------------------------M---I----------D-------D---T----.----P----------------------------S.....---V-------------------------T---I------------------------------F-C----- 308
H1N.CM.95.YBF30              -PLT----------I-----------------M--------S-----------------------V------D---T---P.V--LP---L-D-------------A--VA---A.....---V---D--R---V----R-------V---E-K---------------------------- 310
H1O.BE.87.ANT70              -------------A------N--I----M------ISY-I-----AI----G-L-V---V-----V----T--PP.V--LP---I---T----------Q----H-T-RP....-Q-----D--RK--V-----M-K----V-----K---------------------------- 307
H1O.CM.91.MVP5180            -------------A------N--I----M------V-Y-I-----AI----G-L-V---V----------T--PA.M--LP---I---T----------Q----I-T-RG....ANS----D--RK--V-----M-K----V---------------------------------- 307
CPZ.CD.90.ANT                -PLT---------C----N-N----------------H-------A--D--G---V---V----------L--T-.---VQA--L---T-----------V---MQ--STP..QQ-GGV---D-------M----V--X---V---E-K------------------I---- 327
CPZ.CM.01.SIVcpzCAM13        TP------------------N--------------L-L-V-----A-----G---V-R-V----------M----.--------L---------------------------A-A.....-------D--R--V-----RV-K--C-V-----K-------K--------------M---- 313
CPZ.CM.05.SIVcpzEK505        -PL-----------------N-------M--------------------------------V----------G----.M---P---V-------------------A--VA---A.....---V---D--R---V--------------A-----K-----T-------------------- 311
CPZ.CM.05.SIVcpzMB66         -P--------------------------M--------------------------------V----------L----.---VP------------------I---V----S.....-------D-----------------K--C-V-----K-----S-------------------- 306
CPZ.CM.98.CAM3               -PM-----------I---N-N-------M---------V-----AI----G---V---V----------T----.---V----L--A------------A---A---A.....---V-------R--VV-----V----C-V-----K---------------------------- 308
CPZ.GA.88.GAB2               --M-----S----A--------------M---------V-----AI----G---V---V----------L--L-.---V-------------------------V----S.....---V-------R--VV-----V----C-V-----K-------------------V----- 310
CPZ.CM.98.CAM5               --M-----------I---N-N-------M---------V-----A-----G---V---V----------L----.--------L--A---------------VA---A.....-----------R--VV-----V-K--C-V-----K---------------------------- 307
CPZ.CM.05.SIVcpzMT145        NP-----------A------N-------M----------------------I-----I---V--D-------T--T-.-------------------------S-E--V----A.....-------D--R---V-------K----V-----K---------------------------- 298
CPZ.TZ.00.TAN1               -P----------R-I---G-N----------------Y---S---A--E-----------V----------A--A-.---QQA--ML--T-A------------VL---TPQA..QGGV---D------------L------V-----K------------------I---- 334
CPZ.CM.05.SIVcpzLB7          --L-------------------------A--A-----------------------I---V-----Q----L--I-.M-------------------V--------S.....-------------------M----V------V-----K---------------------------- 308
CPZ.US.85.CPZUS              --M----------A------N-------M---------V-----AI----G---V---A------D*--T--L-.-----------------------------V----A.....-------D--R--VV-----V-K--C-VG----K---------------------------- 309
CPZ.GA.88.GAB1               -----------------------------------L---V-----A-----G---V---V----------L--T-.--------L----------------------T-A.....-------DV-R--V------V----C-V---------------------------------- 316
GSN.CM.99.CN166              --LNS-I-KT--SI----K-A--TVAL-Q-----CI-Y-M-Q---AI-DY-G-V-II-DV---Q-----LL---P.-A-QPVAGL-D-S-A----V---PN---E-I-RQ....-Q-VN-SD--RK------QRC-K--N-VN----K-------KE-----F-C----- 315
GSN.CM.99.CN71               --LNS-V-KT--SI----K-A--TVAL-Q--T--CI-Y-M-Q---A--DY-G-V-II-DV---Q-----LL--QP.-A-QPVAGL-D-S-A----V---PN---E-I-RQ....-N-VN-AD--RK------QRC-K--N-VN--G-K-------KE-----F-C----- 315
MUS_1.CM.01.1085             --LNA-I-KT--SI----K-AA--VA--Q--A---I-Y-I-QL--AI-E--G-I-II-DV---Q-----LLN-QP.QQ-QPNAGL-N-T------VS--P----E-I-RA....-N--N-A----K-V-M--QRC-K--N-VN----K-------K-----F-C----- 306
MUS_1.CM.01.CM1239           -SLNA-I-KT--SI----K-AA--VA--Q--A--TI-Y-M-QL--AI-D--G-I-II-DV---Q-----LL--QP.QQ-QPNAGL-S-S------V---PN---E-I-RQ....-D--N-AD---K---M--QRC-K--N-V-----K-------K--------C----- 306
MUS_2.CM.01.CM1246           -SL---L-KT--ST----K-A---VAL-Q-----CI-Y-V-Q---AI-D--G-I-LI-DVV--Q--DR-VL--QP.TQ-QPNAGL-Y-S-A-V--VS--PA---E---RQ...QD-VN--N--RK------QRC-K--N-VN----K-------K-----F-C----- 310
MUS_2.CM.01.CM2500           -SL---I-KT--STI---K-A---VAL-Q-----CI-Y-V-Q---A--D--G-V-LI-DVV--Q-----LL--QP.-Q-QPNAGL-Y-S-A----VS--PA---E---RQ...QD-VN--D--RK------QRC-K--N-VN----K---R---K-----F-C----- 311

MAC.US.x.239                 LPL----------LI---K-GA--V-G-Q-----C--Y-I-Q---C--D------IIRDI------D--LQ--QP..A-.QQ--L---S---------SVD---Q--YRQ....QN-----N--R---Q---Q-C----N--N---VK-------QS-------S----- 309
H2A.GW.x.ALI                 VPL-----D----L----K-GA--V-G-Q-----C--Y-I-Q---C--D------IIR-I------D--VA--IP..--LPA--L-------------VE---Q--FRP....RN-V---N--R---QI--Q-C----N--N----K-------QS-------S----- 310
H2A.DE.x.BEN                 VPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--D------IIR-I------D--SQ--IP..--LPA--L-D-----------VD---Q--YRP....QN-V---N--R---QI--Q-C--K-N--N----K-------QS-------S----- 310
H2A.SN.x.ST                  VPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--D------IIR-I------D--AQ--IP..--LPA--L-------------VE---Q--FRP....QN-V---S--R---QI--Q-C----N--N----K-------QS-------S----- 310
H2B.GH.86.D205               LPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D--QQ--SP..--MPA--L-D-----------VE---Q--YRA....QN-V---N--R---Q---Q-C----N--N----K-------QS-------S----- 305
H2B.CI.x.EHO                 LPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D--QQ--SP..--MPA--L-------------VE---Q--YRP....QN-V---N--R---Q---Q-C----N--N----K-------QS-------S----- 305
H2G.CI.92.ABT96              LPL----------L----K-GA--V-G-Q-----CL-Y-I-Q---C--E------IIR-I------X--LQ--QQ..--PPA--X---T---------IE---Q-THRQ....QN-----NX-R---X---Q-C---XN--N---VK-------QS-------SX---- 310
H2U.FR.96.12034              LPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I--D---D--TQ--NQ..--PPA--L-------------VE---Q--YRP....QN-----N--R---Q---Q-C----N--N----K-------QS-------S----- 309
MAC.US.x.251_BK28            LPL----------LI---K-GA--V-G-Q-----C--Y-I-Q---C--D------IIRDI------D--LQ--QP..A-.QQ--L---S---------SVD---Q--YRQ....QN-----N--R---Q---Q-C----N--N---VK-------QS-------S----- 309
MAC.US.x.251_1A11            LPL----------LI---K-GA--V-G-Q-----C--Y-I-Q---C--D------IIRDI------D--LQ--QP..A-.QQ--L---S---------SVD---Q--YRQ....QN-----N--R---Q---Q-C----N--N---VK-------QS-------S----- 309
MAC.US.x.EMBL_3              LPL----------LI---K-GA--V-G-Q-----C--Y-I-Q---C--D------IIRDI------D--LQ--QP..A-.QQ--L---S---------SVD---Q--YRQ....QN-----N--R---Q-R-Q-C----N-IN---VK-R-----QS-------S----- 309
SMM.US.x.H9                  LPL----------L----K-GA--V-G-Q-----C--X-I-Q---C--E------IIR-I------D--LQ-XQP..---PA--L---X----------XD---Q--XRQ....QN-----N--X---Q---Q-C----N--N---VK-----S-QS-------S----- 310
SMM.US.x.PGM53               LPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D--LQ--QP..---PA--L-D-----------VE---Q--YRQ....QN-----N--R---Q---Q-C----N--N---VK-------QS-------S----- 310
SMM.US.x.SIVsmH635F_L3       LPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D--LQ--QP..--LPA--L-------------VE---Q--YRQ....QN-----N--R---Q---Q-C----N--N---VK-------QS-------S----- 310
SMM.SL.92.SL92B              TPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D--LQ--RGQQPAQPA-GL---S----------PS---E--YRA....QN-V---D--R---Q---Q-C----N--N---VK-------QS-------S----- 310
STM.US.x.STM                 LPL----------L----K-GA--V-G-Q-----C--Y-I-QL--C--E------II--I------D--MQ--QP.P--LPA--L---S---------SPE---Q--YRQ....QN-----N--R---Q---Q-C----N-VN---VK-----T-QS-------S----- 311
MNE.US.x.MNE027              LPL----------LI---K-GA--V-G-Q-----CI-Y-I-Q---C--E------IIR-I------D--LQ--QQ..A-.QQ--L---S----------VD---Q--YRQ....QN-----N--R---Q---Q-C-K--N--N------------QS-------S----- 309
GRV.ET.x.GRI_677             -PL----------C----RWGA--V---Q-----CLSY-V-Q---VI-D--G-L-I---V----------T-RPP.---LPA--L-D-T---------SI----E-TF-A....--R-D--AQ-RK-V----Q-V-Q--N-QKV-----------Q--------A----- 311
VER.KE.x.TYO1_patent         VPL----------A----K-GA-IV---Q-----C--Y-I-Q---VL-D--G-L-IV--I------Q--VT--LP.---LPA--L-D-----------SV---LE-IYTA....--RVD--A--R-------Q-C-K--N-V-V-----------K--------AI---- 317
VER.KE.x.9063                VPL----------A----K-GA-IV---Q-----C--Y-I-Q---VL-D--G-L-IV--I------Q--VT--PP.---LP---L-D------------V---LE-IYTA....--RVD--A--R-------Q-C-K--N-V-V-----------K--------AI---- 317
VER.KE.x.AGM155              VPL----------A----K-GA-IV---Q-----C--Y-I-Q---VL-D--G-L-IV--I------Q--VT--PP.---LPA--L-D-G----------V---LE-IYTA....--RVD--A--R-------Q-C-K--N-V-V-----------K--------AI---- 317
VER.DE.x.AGM3                VPL----------A----K-GA-IV---Q-----C--Y-I-Q---VL-D--G-L-IV--I------Q--IA--PP.---LPA--L-D------------V---LE-IYTA....--RVD--A--R-------Q-C-K--N-V-V---------A-K--------AI---- 321
TAN.UG.x.TAN1                TPL----------T----R-GA-IV---Q-----CLSY-I-Q---VI-D--G---II--V--D---Q--IT--PP.---LPA--L-D-----------SVA---E-TF-A....--RVD--R--RG-V----Q-C-K--N-I-V-----A-----K-------QA----- 317
SAB.SN.x.SAB1C               -PL-----------I---K--A--V-------A---I-Y-I-Q---A--E--G-L-IV-DV------D--LR--PP.QQ-P-Q-VL-D-Q----------I----E-T-RA....QNAVN--N---G------Q-C-K--N-VN----K------K--------A----- 322
OLC.CI.97.97CI12             VPLAT---QS---AL-DRG-KS-IV-L---FA--LL-EEI--L--V--DY-G---V---A-----E-Y--Q--QP.I--LPV--IPP--A---V----N--Q-LA-VNAG.......-D--AL----V-I--Q-L-K----VTV------A----S------F-A----- 290
DRL.x.x.FAO                  --------------I---K--A--V---L-----CISY-I-Q---AI-D--G-L-IV-DI------D---R--QV.---LPQ-VL-N-S---------SIE---E-T-RA....QDS-N--A--RQ-VV---QRC-T--N-VN---VK-------K--------A----H 311
RCM.NG.x.NG411               -P----------------K-GA--V---------CI-Y-V-Q---AI-E--G-L-IV--V--D---D--LR--NP.---LPA--L---T---------NIA---A-T-RA....-N--A--N--RN--V---Q-C-K--N-VN----K-------K--------A----- 310
RCM.GA.x.GAB1                NP----------------K--A--A---------CI-Y-I-Q---AI-E--G-L-IV--V------D--AR---P.----PA--L---T---------SIA---A-T-RA....-N--N--NL-RN---V--Q-W-K--N-VN----K-----S-K--------A----- 310
MND_1.GA.x.MNDGB1            TP----IIQT---T----KWK-----L----T---ISH---I---A--D--G---V--DV---Q-----LT--QQ.QPAQPG-GL-T-S----------VE--LA--NMQ....QNA-N--T---S-----M-RL-KSHC-I--T-V------A-K-------NVM---- 298
MND_2.x.x.5440               -G----------CI---K----IV---I-----CI-Y---G---AI-D--G-L-IV-DV------D--LR--PV.---MPQ-VL-N-T-N-------SIE---E-T-RQ....QDQVN--G---Q--V---Q-C-S--N-VN----K-------K--------A----R 311
MND_2.CM.98.CM16             -G----------CI---K----IV---I-----CL-Y---G---AI-D--G-L-IV-DV------D-NLR--QV.---MPQ-VL-N-T---------SIE---E-T-RE....QDAVN--G---Q--V---Q-C-S--N-VN----K-------K--------A----R 311
MND_2.GA.x.M14               -G----------CI---K----IV---I-----CL-Y---G---AI-E--G-L-IV-DV------D--LR--QV.---LPQ-VL-N-T---------SIE---E-T-RQ....QEQVN--A---Q--V---Q-C-S--N-VN----K-------K--------A----R 311
DEB.CM.99.CM40               MPL---IVKT-INA----K----IV-L-QV-A--C--Y-I-GL--AI-DL-G---II-DV---------LQ--QQ.QP..PQ--L---S-A-----N--VE---A---RPAGQGQG--D--Q--R--V----QRC-K--N--N---VK-------K-------------- 336
DEB.CM.99.CM5                MPL---IVKT-INA----K----IV-L-QV-A--C--Y-I-GL--AI-DL-G---II-DV---------LQ--QQ.QP..PQ--L---S-A-----N--VE---A---RPAGQGAG--D--Q--R--V----QRC-K--N--N---VK-------K-------------- 332
DEN.CD.x.CD1                 --L-T-V-KT-I-E--DRK----IV-L--V-A--LI-Y-I-QL--AI-E--G---V--DV-----HD--LA--ANLQQ-LP-AAQ-D-----------SIED--T---RQ....QG--N--D-----V-M--Q-V-K--Q-V-----K----S-K--------A----- 336
TAL.CM.01.8023               --L---L-KT-IATI---K-A--TVAL----A--CI-Y-I-Q---A--E--G---II-DI---Q-----QM--Q..I--LPA--L---A----------PA--VE-T-RQ....QN-VD--S-----V-I--QRC-K--N-VN---VK---R---KE------RC----- 317
TAL.CM.00.266                --L---L-KT-IATI---K-A--TVAL----A--CI-Y-I-Q---A--E--G---II-DI---Q-S---QM--Q..I--LPA--L---A----------PA--VE-T-RA....QN-VD--A-----V-M--QRC-K--N-VN---VK---R---K-------RC----- 313
MON.CM.99.L1                 --VE--L-KT--Q-I---K-A---VAL-Q-----MI-Y-I-QL--AI-D--G-L-II-DV------N--L---QP.PQ-QPNAGLGD-T-A----VS--P-Q--E-I-RA....-N-VQ-SD--RK-V----QRC-K--N-VN----K-------K--------C----- 313
SUN.GA.98.L14                -PL----VQT---I----GWK--TVA--A--MT--I-E-I-V---AI-E--G----I-DH-V--S-----M--QQ.QPAQPG-GL-T-N-T----V---VE---Q-IYGL..NGATRVA-QD---G-V-E-MERV--LHQT--VVE-----------T---F-A----S 298
WRC.CI.97.97CI14             DPL----IQT---HI-DRG-GA-SV-F-A--TDR-LSH-I--L--NI-D------II-DIV---IQ----T--QA.----------Q-T------V-ASVEDELA-LAQN.....--VD--S-----VLR--Q-V-QI-V-V-----K-----D-KA-------V----- 291
WRC.CI.98.98CI04             DPL----IQT---HI-DRG-GA-CV-F----TDR-LAH-I--L--NI-D------II--IV---IQD---T--QP.----------Q-T------V-ASVEDELA-LAQN.....---D--AV----VLR--Q-V-QI-V-V-----K-----D-KT-------V----- 295
WRC.GM.05.Pbt_05GM_X02       EPL----IQT---HI-DRG-GA-CV-F----TDR-LTH-I--L--NI-D------IV--IV---IT----T--PP.V--LP--A--N-T------V---ME-ELA-LAQQ.....---DI-A-------R--Q-V-QI-V-------K-----D-KS-----F-V----- 293
LST.CD.88.524                TPL----IQT---I--DRGWK--TVA-----T-K-L-E---V---AI-D--G---II-DH-A--G-----Q--QQ.QPAQPG-GL-T-NAT-V--I---VE--QA-T-AD.....T--D--K---E-V-QAME-V--IHQ-V-VM--K-------K--A---F-A----G 293
LST.CD.88.485                TPL----IQT---I--DRGWK--TVA-----T-K-L-D---V---AI-D--V---II-DH-V--G-----T--QQ.QPAQPG-GL-T-NAT-V--V---VE--LA-T-AD.....T-VD--K---E-V-QAME-V--IHQ-V-VM--K-------KE-A---F-A----G 295
LST.CD.88.447                TPL----IQT---I--DRGWK--TVA-----T-K-L-D---V---AI-D--G---II-DH-V--G-----K--QQ.QPAQPG-GL-T-NAT-V--V---VE--LA-T-AD.....T-VD--K---E-V-QAME-V--IHQ-V-VI--K-------KE-A---F-A----G 295
LST.KE.x.lho7                TPL----IQT---I--DRGWK--TVA-----T-K-L-D---V---AI-D--G---II-DH-V--G-----Q--QQ.QPAQPG-GL-T-NAT-V--V---VE--LA-T-AD.....T-VD--K---E-V-QAME-V--IHQ-V-VM--K-------KE-A---F-A----G 298
SYK.KE.x.SYK173              TGVPV---KT--EA-NS-K-DASIV-L-Q--T--FI-Y---G---A--D--G---VI-DI----G-A--LE--QP.QQ-APQAGL-D-SA---------SIA---E-I-RQ....-N-VQ-----R-------Q-C-QV-N-V-----K-----I-K-------HC----- 334
SYK.KE.x.KE51                VGVNT---KT--EA-NS-K-DASLV-L-QI-T--FI-Y---D---AI-D--G---VI-DV----GS---LQ--QP.QQQQPVAGL-D-SA---------I----E-I-RQ....-N--Q--Q--RQ------Q-C-QV-N-VN----------T-K-------HC----- 331
COL.CM.x.CGU1                -PL-----G----C--.GGIA-SLA---L-Y-T--IAY-M-L---ILDT--GFP-V--DE--KK-E-Y-LL---..QQ-QQQ-AL-Q-TA---T-N--SVA--VA-GE..........--..AN--RG--VQS-E-VIQIAR-S-V------S--D-KS------SA----P 300
```

```
                             p24 end p2 start              p2 end p7 start                                   Zn motif                           Zn motif
H1B.FR.83.HXB2               ASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNSAT.....IMMQRGNFRNQR..................................KIVKCFNCGKEGHTARNCRAPRKK.........GCWKCGKEGHQMKDCT..  427
H1A1.UG.85.U455              -T-D-------------------S--R----G----------------------------QQTS..........-------GP..................................RRI----------L-K--------.........----------------..  421
H1B.US.90.WEAU160            -T-----------------------------G----D------------------------T-.....M--------SP-..................................-TI----------I----------.........------Q---------..  427
H1C.ET.86.ETH2220            -T-D------D--------------R----G-S--------------A------------N-TT.....-----KS--KGPK..................................RAI----------L----------.........----------------..  424
H1D.CD.84.84ZR085            -------------------------------G---------------S-----------A-SASAA.....-----KS--KGT-..................................-------------I----------.........----------------..  428
H1F1.BE.93.VI850             ---D--G---D------------------TG---------------S------------ANSA.....-----KS--KG--..................................RV-----------I----------.........------R---------..  421
H1G.SE.93.SE6165             ---D--G---D--------------R---QG-S--------------S------------ASGA-AA.....------S--KGP-..................................RTI----------L----------.........----------------..  428
H1H.CF.90.056                -T-D---------------------R---QG-SI-------------S---------------TN-A.....----K---KG--..................................-F-----------I----------.........------R---------..  428
H1J.SE.93.SE7887             -T-D------D------------------SG------------------------------TN.....------------DHK..................................R---------Q--I-K--------.........----------------..  426
H1K.CM.96.MP535              -T--------D-------------------G-S--------------S----I---------PV.....V---K---KGH-..................................-------------I----------.........----------------..  422
H101_AE.TH.90.CM240          -T---------------------S-----TG----------------S-----------HAQHAT.....-----------KG-..................................-RI------R---L---------Q.........----------------..  425
H102_AG.NG.x.IBNG            -T---------------------S--R---TG---------------------------QQAN.....V--------G-..................................RTI----------L----K----R.........----------------..  422
H103_AB.RU.97.KAL153_2       -T-D---------------------R---SG--------V--------------------Q-AN.....-----KS---GP..................................-RI----------L----------.........----------------..  425
H104_cpx.CY.94.CY032         -T---------------------S-----TG----------------S------------AS-A-AA.....-----KSK-KG--..................................RTI----------L----------.........----------------..  428
H1N.CM.95.YBF30              -T---------------------QL------G----------------A-------------QQPT-S....VFA-----KGI-..................................-PI----------L----K---RG.........------Q--------KNE  432
H1O.BE.87.ANT70              -T---------------------Q---S---G-------V--------T---------ATAQQDLKGGYTAVF----QNPIRK..................................GTI----------I----------.........------Q--------RN.  432
H1O.CM.91.MVP5180            -T---------------S-----Q-------E-------V--------T---KI-----ASAQQDLKGGYTAVF----QNP-RK..................................GPI----------I-K-------R.........------Q--------KN.  432
CPZ.CD.90.ANT                ---P--A-------I--------H------TG-S----L---------A----------ASAN-AQGT...AVFL----GNRGGK..................................RPL---------------K---R-.........---R--Q----L-N-PAT  451
CPZ.CM.01.SIVcpzCAM13        ------T---D------------Q--------G-S--------------A----------TM-QSQ-RT...D-FF-K-PGATP-..................................RKI--Y-------L----K---R-.........------Q--------S..  434
CPZ.CM.05.SIVcpzEK505        -T---------------------N--R----G-S---------------A----------T-AQTATS.....VF------KGI-..................................-TI----------L----K-----.........------Q--------RS.  431
CPZ.CM.05.SIVcpzMB66         -T--------D------------S--R----G----GDEP-F------S------------AQH-NDA.....KR-FKGPK..................................R-------------I----K---R-.........------Q-----RN--N.  423
CPZ.CM.98.CAM3               -T--------D------------N--R----G-----------------A------------LQ-PT.....GVFL----NGKPT..................................RKI----------L--------R-.........------Q------N-PA.  428
CPZ.GA.88.GAB2               ---D-------------------Q--R----G-----------------A------------IQQTS.....S-F---AGG-TPP..................................RR-----------I----K---R-.........---------------QN.  430
CPZ.CM.98.CAM5               -T--------D------------S--R----G-----------------A-------------Q--S.....NVF-----NGRVN..................................-KI--Y-------V--------R-.........------Q---N--E--TG  428
CPZ.CM.05.SIVcpzMT145        ---D------D---I--------N--R----G-----------------S-------------NKT......E-F----GQNRPPN..................................-KI----------L--------R-.........------Q--------NQN  419
CPZ.TZ.00.TAN1               ---P--T-------------------H------QG------L---------S---KI-----ASA-AGG.....VN-L-G-KRPPLK..................................KGQLQ------V-------------.........---R--Q---------TR  456
CPZ.CM.05.SIVcpzLB7          -T--------D------------S--R----G-S---------------A------------ANAA.....TV--------KGPK..................................RVI----------I----K---R-.........------Q---R-Q--NM.  428
CPZ.US.85.CPZUS              --P---T------------------R----G------V--------A----------C-MK-PS.....SVFL-K--AGKPG..................................RKI----------L--------R-.........------Q---------AG  430
CPZ.GA.88.GAB1               ----------D------------Q--------G-----------------S------------M-Q-QGRA...DVFF-K-QGAGPK..................................RKI----------L----K---R-.........---R--Q----------..  437
GSN.CM.99.CN166              -D-A------QS--------E--L----M.-G-N----L---------T--S-L-----ATALKGTSS...YN-V-K-PPGR-..................................G-TP--Y---QF--L--D-PK-KER.........K-F---RA--FS-P-R..  435
GSN.CM.99.CN71               -D-A------QL--------E--L----M.-G-S----L---------T--S-L-----ATALKG.SS...YN-V---PPGR-..................................G-TP--Y----F--I--D-PK-KER.........K-F----A--LARQ-KT.  435
MUS_1.CM.01.1085             SD-A------S---------E--L---SM.-G------LQ---------T--S-LM---LASAF.KQAG.TL.N-V-GAKG.........................ARGRGNGSQGNRGNPR-Y---QY--V--D-PK--N-.........K-F---R---LARQ-R..  434
MUS_1.CM.01.CM1239           SD-A------S---------E--L---SM.-G------LQ---------A--S-LM---ASAL.RQNS.QL.N-V-G...........................ARG.KGSQGGPRGNPR-Y---QF--M--S-PK-KTR.........K-F---R---LA-Q-R..  430
MUS_2.CM.01.CM1246           -D-A------QS--------E--LV---M.-G---K--LQ----IE--M--S-LM----TNAL.RQNTINTIN-V---.QSPR.............................GVMGKKRENSTR-Y---QF--L--D-PK-KST.........R-F-------LARQ-R..  437
MUS_2.CM.01.CM2500           -D-A------QS--------E--LV---M.-G------LQ---------I--SKLM---ASAL.RQNTIETIN-V-G..QGPR.............................GGAGGGPRRTPR-----QF--I-KD-PK--VR.........K-F---R---LA-Q-R..  438
                                                                        p27 p2                 p2 p8
MAC.US.x.239                 TDAA------Q---I--------LV--G--VNP-----L----------Q---LM---LKEALAPVPIP..FAAA-QRGP..................................R-PI--W-------S--Q-----RQ.........-------MD-V-AK-P..  428
H2A.GW.x.ALI                 TDPA------Q------------LV--G--MNP-----L----------Q---LM---LKEAMTP-PIP..FAAA-Q......................................RRTI--W-------S--Q------Q.........-------P--L-AN-P..  426
H2A.DE.x.BEN                 TDPA------Q---I--------LV--G--MNP-----L----------Q---LM---LKEAMGPSPIP..FAAA-Q......................................R-AIRYW-------S--Q-----RQ.........-------P--I-AN-P..  426
H2A.SN.x.ST                  TDPA------Q------------LV--G--INP-----L----------Q---LM---LKEAMAP-PIP..FAAA-Q......................................RRTI--W-------S--Q-----RQ.........-------A--I-AK-P..  426
H2B.GH.86.D205               TDPA------Q---I--------LV--G--MNP-----L-----I----Q---LM---LKEALTP-PIP..FAAV-QKAGK..................................RGT-T-W----Q-----Q-----RQ.........-------T--I-SK-P..  425
H2B.CI.x.EHO                 TDPA------Q---I--------LV--G--MNP-----L-----I----Q---LM---LKEALTPS-NP..FAAA-PRAG...................................KRT-T-W----A-----Q-K---RQ.........------QQ--I-SK-P..  424
H2G.CI.92.ABT96              TDPA------Q---I--------LV--G--MNP-----L-----X----Q---LM---LKEAL-PTALP..FAAA-QKTGG.................................KRSTI--W--------V-Q-----RQ.........-------P--I-AK-P..  431
H2U.FR.96.12034              TDPA------Q---I--------LV--G--MNP-----L-----I----Q---LM---LKEALQPMPIP..FAAV-QRGGP..................................RRP-R-W----------Q-K---RQ.........-----E-P--N-AS-P..  429
MAC.US.x.251_BK28            TDAA------Q---I--------LV--G--VNP-----L----------Q---LM---LKEALAPVPIP..FAAA-KRGP..................................R-PI--W-------S--Q-----RQ.........-------MD-V-AK-P..  428
MAC.US.x.251_1A11            TDAA------Q---I--------LV--G--VNP-----L----------Q---LM---LKEALAPVPIP..FAAA-QRGP..................................R-PI--W-------S--Q-----RQ.........-------MD-V-AK-P..  428
MAC.US.x.EMBL_3              TDAA------Q---I--------LV--G--VNP-----L----------Q---LM---LKEALRPVPTP..FAAA-QRGP..................................R-PI--W-------S--Q-----RQ.........R------MD-V-AK-P..  428
SMM.US.x.H9                  TDPAX-----Q---I--------LV--G-XINP-----L------X---Q-X-LM---LKDALTQGPLP..FXAV-QKGQ..................................X--IX-X--X----S--QF----RX.........-------A--V-AK-P..  429
SMM.US.x.PGM53               TDPA------Q---I--------LV--G--MNP-----L----------Q---LM---LKEALTPGQLP..FAAV-QKGP..................................R-TI--W-------S-KQ-----RQ.........-------T--V-AK-P..  429
SMM.US.x.SIVsmH635F_L3       TDPA------Q---I--------LV--G--MNP-----L-----I----Q---LM---LKEALRPDQLP..FAAV-QKGQ..................................RRTI--W-------S--Q-----RQ.........---G---T--V-AK-P..  429
SMM.SL.92.SL92B              TDPA------Q---I--------LV--G--MNP-----L----------Q---LM----KDALTGSLV....AA-FRGAAKGQ.................................GNKP-IR------T--S--Q-----R-.........------E--RIQAN-P..  432
STM.US.x.STM                 -DPS------R---I--------LV--G--MNP-----L----------Q---LM---LKE-FQPDPLP..FAAA-QQG...................................RRT---W----A----KQ-KG--RQ.........------QQ----AK-P..  429
MNE.US.x.MNE027              TDPA------Q---I--------LV--G--MNP-----L----------Q--KLM---LKEALAPGPLP..FAAA-QKGP..................................R-PI--W-------S--Q-----RQ.........------QM--V-AK-P..  428
GRV.ET.x.GRI_677             -P-D------Q---I--------L---G--MNP-----LI--------Q---KLMV-M--NGQ.........N-V-V-PQKKGP..................................RGPL-------F--MQ-E-K---QI.........K-F----I--MA---K..  427
VER.KE.x.TYO1_patent         --G---Q----S--I--------V---G--MHP-----L---------SY--K-M--M-QTMQ-Q.......N-V-Q-GPKR-..................................RPPLR-Y----F--MQ-Q-PE---T.........K-L----L--LA---R..  434
VER.KE.x.9063                --G---Q----S--I--------V---G--MHP-----L---------SY---M--M-QNMQTQ.......N-V-Q-AAGGV-................................RQRPPL--Y----F--MQ-Q-TE---I.........K-L----L--LA---R..  437
VER.KE.x.AGM155              --G---Q----S--I--------V---G--MHP-----L---------SY--K-M--M-QNLQSQ.......N-V-Q-GG-GRP..................................RPPP--Y----F--MQ-Q-PE---I.........K-L----P--LA---R..  435
VER.DE.x.AGM3                --G---Q----S--I--------V---G--MHP-----L---------SY--K-M--M-QNMQSQ.......N---Q-GQ-GRP..................................RPP---Y----F--MQ-Q-PE---M.........R-L----P--LA---R..  439
TAN.UG.x.TAN1                TP-D----------I--------LV--G--IHP-----L-------------KLMV---Q-MQG.......VN-V-GAPRGGRG...............................RGRGPPR--K--QI--IQKD-PRAGPN.........K-L----P--LA---R..  437
SAB.SN.x.SAB1C               TDPA------QS--I----------V--G--MNP-----L-----I--AQ----LM----TAAFQQQ-V..GN-FV-Q-ARPRGP..........................LGGRGRPLNPNI--Y----P--L--F-K---RQ.........------SPD------Q..  453
OLC.CI.97.97CI12             G-----T----K--I--------GV----.VKPS----L----------Y--------LRP-QMVQQRPAP..............................................GGGPQQGRLR-Y---RF---KN--SQKGP.........T-Y----P--IA-N-R..  405
DRL.x.x.FAO                  TDAA------Q---I--------V---G--MNPS----LL--------------M----REQQAA...........----QNPP-GPP...........................RGPRGPPP-NPR-P---QF---L-Q-TK---.........----FR--ALD-MLRN-...  435
RCM.NG.x.NG411               -DPA------QS--I----------V--G--MNP-----L-----A---Q----------QM-Q.......SN--A-Q-GP-KGP...............................PKLGGGPRFL--Y----T-----Y-KTS---.........---R--E---L----P..  434
RCM.GA.x.GAB1                -DPA------QS--I--------MV--G--MNPS----L---------Q---------QMMQSN.........--A-QSAN.............................RGPPRRSGGNPNLR-Y----P--IS-Y-K---R-.........------SPD-LL-----..  433
MND_1.GA.x.MNDGB1            --G---M--QQH--IE----E--Q--RS--KG------LE---------Q----LM--M-RT-VGQ..S....QNFVQQRGPQR.............................GPVRQPTGRKPI----N----V--FFK---R-.........----N--AMD--KAQ-P..  423
MND_2.x.x.5440               TDPQ--T---Q---I--------AT--G--MNP-----LL---------KY--QMM----Q-AQAA.......V----NSGGPPRG..............................PPRQPPRNPR-P----F--VL-D------R.........----F---DP--L-RN-...  433
MND_2.CM.98.CM16             TDPQ--T---Q---I--------S---G--MNPS----PL---------KY--QMM----KEAQS.......AV----NSGGPPRG..............................PPRQPPRNIR-P----F--GL-D-IS----.........----F---DL--I-RN-P..  434
MND_2.GA.x.M14               TDPQ--T---Q---I--------S---G--MNPS----LL---------KY--QMM----QKMQS.......EV----NSGGPPRG..............................PPRQPPRNPR-P---RY--VL-D--L----.........----F---DT--M-RN-P..  434
DEB.CM.99.CM40               -D-A------T--MI-------RI---G--QNP-----LH---------QQ---LM----ASALKE-GS...LG-V-QRRGPPGS................................RRQIR-----QI--LQKD-KR---T.........K-F---Q---IA-N-G..  458
DEB.CM.99.CM5                -D-A------T--MI-------RI---G--QNP-----LH---------QQ---LM----ASALKE-GS...LG-V-QRGGQGGG...............................PRRQLR-----QI--VQ-D-KK---V.........K-F-------IA-N-G..  455
DEN.CD.x.CD1                 -D-S--L---TS-----------Q--E----NP-----LQ---------Q---PT---LASAFQKQGLCHGSRKGAQPPVER.....................................RGLR-----QI--L-KD--K-KRVT.....QPGA-F----L---A-N-R..  462
TAL.CM.01.8023               TD-A------Q---------E--A----M--Q------LQ---------Q--S-LM----AEALK.QNTQQVIA-V-.-QGGP.........................KGRGGPRRTPPGQIR-Y----F--I-K--P--Q-PRG....PPGS-F---QM--RAAQ-R..  453
TAL.CM.00.266                TD-A------Q---------E--A----M--G------LQ---------Q--S-LM----AEALK.QNTQQVIA-V-Q-QGGP.........................RGRGGPRRNPPGQIR-Y----F--I-K--P--PRQKV....PPGT-Y----P--IA-Q-S..  450
MON.CM.99.L1                 TD-A------Q------------L----M.-G-S--D-LQ---------M--S-I-----AGAIAN...MPMN-V-................................ARGPPQRKGQP-------F--M-K--K--QRR.........K-YN--QP--LA---PQP  434
SUN.GA.98.L14                G-E---E--K-KM----------LVI----ETPS-----R---------S--GKI-----ASAFRQQVG..RQA-V-QNLPPRNSQ...........................GRFVRIGGGGPR-PMT------P--L-NQ--E-K-G......PPGS------M--KQAQ-P..  436
WRC.CI.97.97CI14             GAP---Q---DK--I-------------G-.QQP-----L-----------Y-SKIM-Q-LQEAL.H.SQ....A-V-QAGAKP..................................KGPIR-----QI--M-KE-PK-K-PVPMRRGQGPT--G--EV--VQRN-P..  416
WRC.CI.98.98CI04             GAP---Q---DK--I--------A---G-.QQP-----L-----------Y-SKIM-Q-LQEAL.-.SQ....A-V-QQPVKP..................................KGPIR-----QI--M-KD-PK-K-PVPLRRGQGPT--G--EI--VQRN-P..  420
WRC.GM.05.Pbt_05GM_X02       GVP---Q---DK--I--------Q---G-.QSP-----L-----------Y-SKIM-Q-LQEAL.-QSQ....A-V-QQAVK...................................RGPLR-----QI--M-KD-KK-KRSPQQR..GPPT--G---T--IQ-N-P..  416
LST.CD.88.524                G-H---E--K-KM----------MVI----EG-------K---------A--RKI-----ATAMQGQMK...HL-V-QTPP--AQ............................GRFVRTGGGGPRRPLT------P--S-KM--Q--QE.........----N--AKD-R-AQ-PKP  428
LST.CD.88.485                G-H---E--K-KM---------RMVI----EG-S--D--K---------A--GKI-----ATAMQ-QMR...QN-V-VTPP--EQ............................GRFVRTGGGGPR-PLT------P-----M--Q--QE.........----N--SKE-RFAQ-PKP  430
LST.CD.88.447                G-H---E--K-KM---------RMVI----EG-------R---------A--G-I-----ATAMQ-QMR...QN-V-VTPP--AQ............................GRFVRTGGGGPR-PLT------P-----M--Q--QE.........----N--SKE-RFAQ-PKP  430
LST.KE.x.lho7                G-H---E--K-KM---------RLVI----EG-S-----K---------A--GKI-----ASAIQQQVR...QN-V-VTPL--AQ............................GRFVRTGGGGPR-PLT------P-----M--Q--QE.........----N--SKE-RFAQ-P..  431
SYK.KE.x.SYK173              -DPS--G-L-Q---I-----E-RQ----MVK.P-----LQ---------L---KLM----VMAQQ-V.......N-V-GPSKG..................................RSMI--Y---QI--MQKD-KK-L-A.........K-FN---T--LARA-RQP  451
SYK.KE.x.KE51                -DPA----L-QS--I-----E-RS----MVK.P-----LQ---------QY--KL-----VMTQH-L.......G-I-GPRQGSNP................................RRGPTR-----QL--LQKD-PR-K-L.........K-FN--GT--IARQ-RQP  452
COL.CM.x.CGU1                -AG-I-A--ANN--I-H------R---G-QK.PS--D-LA--------D---K------Q-FQQER.....TN-IEVKTAKCFNCQGIGHLARMCPKRPIGGAGRGRGRGRGGFRGAPRRP-R--T-NQ---MQ-D-P..............................  432
```

```
                          p7 end_p1 start      p1 end_p6 start      Vpr binding                                                     Vpr binding
                                                                                                                                    p6 end, Gag end
H1B.FR.83.HXB2           .......ERQANFLGKIWPS....YKGRP...GNFLQSRP..........EPTAPPEESFRSGVETT....................................TPPQK.QEPIDKELY..PLTSLRSLFGNDPSSQ*     500
H1A1.UG.85.U455          .......--------------....N----...---P----..........------A-I-GM-EKM....................................-S-A-.--LK-R-QT..--V--K-------L---     494
H1B.US.90.WEAU160        .......-----------S-....Q----...---P---L..........-----------FRE---....................................--S--------------.------K----------     501
H1C.ET.86.ETH2220        .......--------RL---....N----...--------..........-------SLRP.......EPTAPPPE..........................SFRFEEA--SP-.--LK-R-....A----K------HLL--     505
H1D.CD.84.84ZR085        .......--------------....H----...--------..........------A-ASGL-RR*P...................................#SQK-E.-KDK-----.--A--K-------L---     501
H1F1.BE.93.VI850         .......--------------....N----...--------..........------A---GFRE-I-...................................-SP-.--QK-G---.P--A--K-------*---     494
H1G.SE.93.SE6165         .......--------------....N----...-----N-T..........------A--LGF-E-IA...................................-SP-.--MKE-------...P--K----S--*---     498
H1H.CF.90.056            .......--------------....S----...--------..........------A---GF-E-M-...................................-SP-QEQLK---PP...-A-------S--LL--     501
H1J.SE.93.SE7887         .......--------------....S----...--------..........------A--LGL-E-I....................................-SP-.---K-----.-----K----S--L---     498
H1K.CM.96.MP535          .......--------------....H----...--------..........------A---GF-E-I-...................................-SPR.--TK---QS.P-----K-------L---     496
H101_AE.TH.90.CM240      .......--------------....N----...---P----..........------A-NWGM-E-I-.............................GEEI-SLP-.--QK---HP.P--V--K-------L---     504
H102_AG.NG.x.IBNG        .......--------------....S----...---P----..........------A---GM-E-I....................................P-SPQ.---R--G--.P-----K-------*L--     495
H103_AB.RU.97.KAL153_2   .......--------R-----....S----...---P----..........--S---A-N-GM-E-I....................................--SL-.--QK-R-QH.P-SI--K----D--L---     499
H104_cpx.CY.94.CY032     .......--------RM----....S----...-----N--..........------A-CLERKE---...................................SSL-.---R-----.-----K----S--L---     501
H1N.CM.95.YBF30          .......G--------S-SP....F----...---P-TTTRK..........------L--YGFQE-..........................KSTQGKEM-EN--RTENS--.P---------------     513
H1O.BE.87.ANT70          .......GK-------Y--P....GGT--...--YV-RPA..........H-S---..................................MEEEVKGQENQE-KGGPN----..-FA--K----T-Q*L-Q     501
H1O.CM.91.MVP5180        .......G--------Y--P....GGT--...--YV-KQV..........S-S---..................................MEEAVKEQENQS-KGDQE----..-FA--K----T-Q*--Q     501
CPZ.CD.90.ANT            ......NTGKV-----PT-T...WWGC--...---V-KEEVV..........------I-IYQEEH...............................KR-QKGLKG-EELPPS-......--K----K-Q*WKF     525
CPZ.CM.01.SIVcpzCAM13    .......G--V-----G----....RS----...---V-N-T..........------A--YGYQEK..............................EKNQEEQEEGNR--.P-----K----S------     507
CPZ.CM.05.SIVcpzEK505    ......G---------V--L....S-----...---P-TTTRK..........------I--YGYQ...............................EEKTTQGTEREEK-KTESS--.P-----K----S---L--     512
CPZ.CM.05.SIVcpzMB66     .......---------L----....S-----...-----K--..........------I---GFRE-............................EPAQKEEV-SE-TE.NK--.PQ----K-------L---     500
CPZ.CM.98.CAM3           ......G---------G-SP.WSGGSK--...-----G-K..........------L--YGYQEA-G............................KRQEEEREK-KE..P--.P-----K----S----E-     508
CPZ.GA.88.GAB2           ......EG--V----GG---..R..S-----...---P-RKV..........------P-T-GYAEQ-Q...........................TGKSQKGEESTT-KE.SP--....P--K----S------     508
CPZ.CM.98.CAM5           .......G--------D-SP.WSGRSK--...---P-GGK..........------M--YGFQET-M..............................EK-KE-KE..P--.PQ----K----S---LE-     503
CPZ.CM.05.SIVcpzMT145    N......G-.I-----N--P....WG-----...--YV-G-K..........------M---GFQ...............................EEKKV-EEVREESKPPS.-----K----------     492
CPZ.TZ.00.TAN1           N.....NSTGV-----RT-L....WGC---...---V-NT-...EKGKAQEQE--QTP..........VVPTAPPLE......................................MTMKGGF--K-I--S-Q**--     526
CPZ.CM.05.SIVcpzLB7      .......GK.V-----L-----....SQ---...---P-----..........------I---GFQPK.........................TTQKQEDREEDKTENK--.P-----K-------L---     504
CPZ.US.85.CPZUS          .......N--------H-SPSWSGGSK--...-----EN-K..........------I-D-GYQE--V..............................-QEKQGK-KE...PF..Q----K----S----E-     506
CPZ.GA.88.GAB1           .......G--V-----G----..R..S-----...---V-N-T..........------I--YGYQE-E.............................KSQEKKEGESS---..-P---K----S------     509
GSN.CM.99.CN166          ...TGTPIKELY--DLGRGR.....RRA-...R--PVTFL..........L----A-.............................EEEPQQLNLSTPVGEPQSKGQK-ESRNT--.....P--S----E-Q...-     513
GSN.CM.99.CN71           ....GTPAKEL---DLGRGR......RA-...R--PVTSL..........L-S--A-.............................EVEPQQLNLSPPVGEPQNKEQKAEKRNT--.....P--S----E-Q...-     511
MUS_1.CM.01.1085         ....SDNTKAV----DPFGL...Q-R--...R--P-TSL..........L------QEGELVNLSPAV..................................GPGEGKQDR-S--.....P--S---QD-Q-D--     510
MUS_1.CM.01.CM1239       ....SEGAKS-----DPLGL...QRR--...R--P-TSL..........L------QEGDLVNLSPPV..................................GP-DGASDR-S--.....P--S---QD--LEK-     506
MUS_2.CM.01.CM1246       ...TDTGKSAV-----DLGF...RSRK-...R--PVTSL..........M-----DTAEPVNFSPLVGPKE...............................EPPKDRKDR-T--.....P--S----D-Q-K--     517
MUS_2.CM.01.CM2500       ...TNSDKGAV-----GFGF.....HRK-...R--PVTSL..........M-S---DTAEIVNFSPEVGPRKEN.............................QTPQAP-DR-N--.....P--S----D---E--     519

                                                                                                                                    p6 end
                                  p8_p1              p1_p6             PTAP motif                          PSAP in HIV-2 B, U         Gag end
MAC.US.x.239             .......D---G---LGPWG......KK-...R--PMAQ......VHQGLM-----..............................EDPAVDLLKNYMQLGKQQREKQRES..REKPYKEVTED-LH-N----G-Q...-     511
H2A.GW.x.ALI             .......----G---LGPWG......KK-...R--PVT-......VPQGLT-----.................AEPAADLLEQYMQQGRKQREQRERPYKEVTEDLLHLE....Q..GETPHKEVTED-LH-N----K-Q...-     522
H2A.DE.x.BEN             .......----G---LGPRG......KK-...R--PVTQ......APQGLI-----.................ADPAAELLERYMQQGRKQREQRERPYKEVTEDLLHLE....Q..RETPHREETED-LH-N----K-Q...-     522
H2A.SN.x.ST              .......----G---LGPWG......KK-...R--PVAQ......IPQGLT-----.................IDPVEDLLEKYMQQGKRQREQRERPYKEVTEDFLQLE....-..QETPCRETTED-LH-N----K-Q...-     522
H2B.GH.86.D205           .......----G---LGPWG......KK-...R--PMTQ......VPQGVT-S---MNPAEGMTPRGATPSAPPAD...........PAVEMLKSYMQMGRQQRE...S..RERPYKEVTED-LH-N----E-Q...-     522
H2B.CI.x.EHO             .......----G---FGPWG......KK-...R--PVQA.......PQGIV-S---MNPAFGMTPQGAIPSAPPAD...........PAEEMLKNYMQLGKKQKE...N..RERPYKEVTED-LH-N----E-Q...-     520
H2G.CI.92.ABT96          .......----G---FGPWG......KK-...R--PMTQ......VPQGLT-S---...................MD...........PAVDLLKNYMXLGRKQKE...Q..RNKPYKEVTXX-LH-S----D-Q...-     510
H2U.FR.96.12034          .......-TKVG---LGPWE......KK-...R--PMAQ......VPQGLT-----....AEPAVDLLTPTAPPAD...........PAVDLLKSYMQQGKKQKE...N..RERPYKEVTED-LH-N------Q...-     522
MAC.US.x.251_BK28        .......D---G---LGPWG......KK-...R--PMAQ......VHQGLT-----..............................EDPAVDLLKNYMQLGKQQRE...S..REKPYKEVTED-LH-N----G-Q...-     507
MAC.US.x.251_1A11        .......D---G---LGPWG......KK-...R--PMAQ......VHQGLT-----..............................EDPAVDLLKNYMQLGKQQREKQRES..REKPYKEMTED-LH-N----G-Q...-     511
MAC.US.x.EMBL_3          .......D---G---LGPWG......KK-...R--PMAQ......VHQGLT-----..............................EDPAVDLLKNYMQLGKQQRE...S..REKPYKEVTED-LH-N----G-Q...-     507
SMM.US.x.H9              .......----G---LGPWG......KK-...R--PMAQ......MPQGLT---X-..............................EDPAVDLLKNYMKVGRRQRE...N..RERPYKEVTED-LH-N----E-Q...-     508
SMM.US.x.PGM53           .......----G---LGPWG......KK-...R--PMAQ......MPQGLT-----..............................EDPAVDLLRNYMKMGRKQRE...N..RERPYKEVTED-LH-N----G-Q...-     508
SMM.US.x.SIVsmH635F_L3   .......----G---FGPWG......KK-...R--PMAQ......MPQGLT-----..............................EDPAVDLLKNYMKMGRKQRE...N..RERPYKEVTED-LH-N----E-Q...-     508
SMM.SL.92.SL92B          .......NQK-G---LGPWG......KK-...R--PMQT.......TSLT-S---..............................DPAARIVKEYLEKAQREKT...R..RSRPYKEVTED-LH-N----E-Q...-     508
STM.US.x.STM             .......---VG---FGPWG......KK-...H--PMAQ......IPQGLT-----..............................EDPAADLLRSYMQLGKKQRE...S..RKTPYKEVTED-LH-N----E-Q...-     508
MNE.US.x.MNE027          .......D---G---FGPWG......KK-...R--PMAQ......MHQGLT-----..............................EDPAVDLLKNYMQLGKQQRE...S..KRKPYKEVAED-LH-N----E-Q...-     507
GRV.ET.x.GRI_677         .......NG------YGHWG.....GAK-...R--V-Y-......GDTVGL------.........................METAYDPAKKLLQQYAEKGQRLREER...EQTRKQKEKEVEDV--S----G-Q...-     514
VER.KE.x.TYO1_patent     ........G-V----YGRWM.....GAK-...R--PAATL........GA--S---PP-GTTP........................YDPAKKLLQQYAEKGKQLREQKR...NP-AMNPDWTEGYSLNS.---E-Q...-     520
VER.KE.x.9063            ........G-V----YGRWT.....GTK-...R---AATH........GV--S---PPTP...........................YDPAKKLLQQYAEKGKQLREQGK...RT-PTNPDWAEGYSLNS.---E-Q...-     520
VER.KE.x.AGM155          ........G-V----YGRWM.....GTK-...R--PAATL........GA--S---PPNN.......................STPYDPAKKLLQQYAEKGKQMRNQNR...NP-ANNPDWNEGYSLNS.---E-Q...-     521
VER.DE.x.AGM3            ........G-V----YGRWM.....GAK-...R--PAATL........GV------PP-P...........................YDPAKKLLQQYADKGKQLREQRK...KP-AVNPDWTEGYSLNS.---E-Q...-     522
TAN.UG.x.TAN1            .......SG------RMPTW..G...TK-...R---EQGGA.........V-----MPAHGFP.........................TGSPAAGAYDPARKLLEQYAKKG..DQLRKQKEKELEDY--S----G-Q...-     524
SAB.SN.x.SAB1C           ........K-V----FGPWG.....R-K-..RNFP-T-...........IR-----M-RDY-R..PEENWYADRPPTRGPGPDDPATALLKQYAVQGKRQKQQWQ......NHSPQQSP-EEAYS-------E-Q...-     555
OLC.CI.97.97CI12         .VGGQQQQSKV-----WRGG.....-AE-PLNMQQRTEMD.........IF-S--TL-..................................QILEKGTR......EKREM--EKKPLYP--A----E-Q...-     481
DRL.x.x.FAO              .......PK-V----NTPWG.....S-K-...R--P-AS..........LT-S---LPG.....................YLQEDPAERMLQKYMEQGAQQKRQQR...-QQQKRGP-EEAYN--S----S-QLQ.-     522
RCM.NG.x.NG411           ........KEV----NAPWK.....Q-K-...R-LPLTS..........LR----GR-EG.................WNPQYDPAEEMLRKYLALGRQHKQEQRKE....NKEKVGRA-EDA-S--N----S-L...-     522
RCM.GA.x.GAB1            ........K-I----RLPWG.....Q-K-...R--PLTS..........LT-S--GM--NYDP......................AEEMLKNYLRRAGEQKRQQRQE...ESKKREGA-QEA----N----S-QLQ.-     519
MND_1.GA.x.MNDGB1        ...KPAQQQRV----YGPWG.....PSK-....--YPAQE..........VT-----L.........................EEKPLQKTLSTYQKLGRGL......RQKMKE-KREEDFH--ST--QE-Q...-     503
MND_2.x.x.5440           .......PKMV----NTPWG.....S-K-...R--PAMP..........LT-S---MPGL...................EDPAEKMLLDYMKKGQQQRAAAGA....-KEEK-GP-EAAYN--S----T-QLQ.-     519
MND_2.CM.98.CM16         ........KMV----NTPWG.....S-K-...R--PAMP..........LT-----MPGM...................EDPAERMLLDYMKKGQQQKAESK......--KKERGP-EAAYN--S----T-QLQ.-     518
MND_2.GA.x.M14           ........KMV----NAPWG.....S-K-...R--PAMP..........LT-S---MPGM...................EDPAEKMLLDYMRKGQQQRAAKET....KQEK--GP-EAAYN--S----T-QLQ.-     519
DEB.CM.99.CM40           ......QTPR-----NTPWA.....SR--...R---EGIG......KYIP-S--DQVY-T.............................TRVPVQGQENNPQRE-TTRRGK-.PD-MG-K----E-Q...-     536
DEB.CM.99.CM5            ......QVPR-----NAPWA.....SR--...R---EGIG......KYIP-S--DQVY-T.............................TRVPVQGERNAPQEDTTTRRGK-.PD-MG-K----E-Q...-     533
DEN.CD.x.CD1             ..SNTNDQG-V----NGLGR.....-NP-...R--V-YY-Q......MNQDLPWS-VRV...................QQQPQYPQNFQEGGNRQSCQEDKRQNKQS...HQEDEETTRESLYP--K----D-Q...-     559
TAL.CM.01.8023           .......QP-------AFLG.....-IP-...VKGPRNF-.........VVQ-----I.........................EPLTASWQEGTTTWTPPTKD....AAAPKQQTRESLYP--A----D-Q...-     532
TAL.CM.00.266            .......SA-------SFLG.....-IP-...VRGPRNF-.........VVQ-S---I.........................EPLTASWQEGTTTWTPPAKK.....PQETTTTRESLYP--A----D-Q...-     528
MON.CM.99.L1             P....KQNKGV----NPFGP.....-KG-...R--PLT..........SVQ-S--TLPA.............................GTAEPVNLNQEIGS-...KTGTP--TRRDLYP--A----E-Q...-     511
SUN.GA.98.L14            ......QGK-V----YGPWN.....R-P-....---P............VM-S---.....................LEDLTLGNRMTPPQSKAERALETYRLLGQGLR....AQQKR--RGECQEPC-NM--PE--...-     522
WRC.CI.97.97CI14         .....KIR--V----NGRGQ.....MPSNNYPVQKLRSLNNFGP..TTTM-S---MREDPNI.....................QELEKYLLKGAQLREQQEHKK......T-GQKR-VGVPEWP--K---EP-Q...-     512
WRC.CI.98.98CI04         .....KSR--I----NGRGQ.....MPSNNYPVQKVRSLNNFGP...ATI-S---MREDPNI.....................QELEKYLLKGAQLKEQAQAK.......K-EKKR-TGLPEWP--K---ET-Q...-     514
WRC.GM.05.Pbt_05GM_X02   .....ETR--V----NGRGQ.....RPSNNYPVQKMNAGVNTFG...PVV----V.........................DPTMQELEQYLLRGAQMKRE......K-KEMQKEGKVEYP--K---EP-Q...-     503
LST.CD.88.524            .......KGKV----YGPWR.....NG-S..---PIMGGAAG......VI-S---.......................MEKVPPTRAERAIETYRNLGQQLRK.....QQQVPQKCVDEPCLNFFF-PD-Q...-     512
LST.CD.88.485            .......KGKV----YGPWR.....NG-P..---PLMGGNAG......VV-S---........................MERNSTKAERALETYRNLGQQLRK.....QQQVPQKCVDEPCLSFFF-PD-Q...-     513
LST.CD.88.447            .......KGKV----YGPWR.....SG-P..---PLMGGNAG......VV-S---........................MERSPTQAERALETYRNLGQQLRR.....QQQIPQKCVDEPCLSFFF-PD-Q...-     513
LST.KE.x.lho7            .....KPKGKV----YGPW-.....SK-P..--YPLLGGAAG......RI-S---........................MESAPTKAERALETYRTLGQQLKR....-QQQVPQKCVDEPCLNSLF.PD-Q...-     516
SYK.KE.x.SYK173          KRNQGPPVA-------G-GV.....SR--P..A--PV..........RS--S---L-..........DIEDGPWLTWSAQMSQQAQAKAQNSPSKKPPTNREVLS-...K-SSG--ETKSLYP--S----E-Q...-     555
SYK.KE.x.KE51            RKGQGNPPP-------G-GN......RK-P..A--PV.........MRET-----A-D.....WPWQQQTWGNYASPQQKHIAMATPLQPQSSPKTVKDLLVPGK-S...EKKEEMKEKGPLYP--Q----D-Q...-     562
COL.CM.x.CGU1            .......NK------ATRGV..E.......................LQTAIF--..........................................KMSKD...LP-RREKGES.LYP--K----D-Q...-     483
```

```
                          Gag-Pol TF start                                                                                   Gag-Pol TF end protease start
H1B.FR.83.HXB2            FFREDLAF..LQGKA...REFSSEQT..RANSPTRRELQVWGRDN..............................NSPSEAGADRQGTV.....SFNFPQVTLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMSL....PGRWKPKMIGGIGGFIKVRQYDQILIEIC  123
H1A1.UG.85.U455           ----N---..Q--E-...--------..-------S-N-WDG-K-..............................DL-C-T--E----D.......--S---I---------V------I---------------DIN-.....--K----I--------------------  122
H1B.US.90.WEAU160        ------V-..PK---...--------..-T---------Q---................................--L-----N---A-.........--------I------------E-----------------D-N-.....--K-------------------VP----  123
H1C.ET.86.ETH2220         ----T---..Q------...----P---..--------ESQTRANSPTTR..........................ELQVRGSNTFSEA-AERQG..-L----I-------------------------------IN-.....--K--------------------I----  130
H1D.CD.84.84ZR085         ----N---..P-----...G-L-----..-------S-SFGFGEEIT............................#LPETGTEGQRQGTV..-LS---I------V---------------------------DIN-.....--K-------------------H------  123
H1F1.BE.93.VI850          ----N---..Q--E-...-K-P-----..-------S---R-QRG--............................PLSEAGAERRGTV.P-LS---I----------------I------------------DIN-.....--K---------------K---N------  123
H1G.SE.93.SE6165          ----N---..Q--E-...--------D..-T----C-KPR-RRG-S.............................PLPEA-DEGKG..AISL--I---------V------I---------------------IN-.....--------------------------VP---S  120
H1H.CF.90.056             ----N---..Q-RE-...-K--P--A..-T----S---R-RRG-D..............................PLSE-GAAEGQGT...-LS---I---------V--E---R------------------IN-.....--K-------------------E-VA----  122
H1J.SE.93.SE7887          --------..Q-RE-...--L-P-----..-------S--PRARRG-............................PLPETGAEGQGTV..-S----I----------R-----R-----------------DID-.....-RK-------------------NEVP---E  121
H1K.CM.96.MP535           ----N---..P--E-...--------..-------S---R-R-G--.............................PLSEAGDQRQGTE.P------I------I----V----R------------------IN-.....--K------------------V-----  123
H101_AE.TH.90.CM240       ----N---..Q------...G-------..-------S-K-GDG---.........................GGRDNLLTEAGAERQGTS.S--S---I---------V---------------------DIN-.....--K---------------K----------  128
H102_AG.NG.x.IBNG         ----N---..Q--E-...-K-------.GT--S-S---WDG---...............................TSLSTAGTEGQGAI.S------I---------VR-E---I---------------DIN-.....--K--------------------------  123
H103_AB.RU.97.KAL153_2    ----N---..Q-RE-...-K--------..--I---S-K-WDG---.............................NPLPETGTEGQGTA.S------I---R-----VR-----------------------DIN-.....--K--------------------------  123
H104_cpx.CY.94.CY032      ----NV--..Q-RE-...-K-----A..-----A-GM-REERG--..............................LLSEAGTEGQGTI..------I-----------L---IR-----------------IN-.....--K--------------------P----  122
H1N.CM.95.YBF30           ----E-VS..--RET...-KLPPDNNKE--H--AT---W-S-G.............................EEHTGEGD--EPGEDRELSV.PT-----I------VI-V---KEVR-----------I--LQ-....E-K-------------------N-TVD-Q  130
H1O.BE.87.ANT70           ---QI--S..GGHE-...-QLCA-TS..TPI---..............................DGGGSEGTGESGTER-PERAL..-VCL--IP--D--I--ARV--H-C-V-----------NNIQ-....E-K---------------KE--NVTV--E  119
H1O.CM.91.MVP5180         ----V--S..GGHE-...-QLCA-TS..VPI---..............................NGGGSEGTRESESEG-SGRAV..PICL--IP--D--I--A-V--H-C-------------NNIQ-....E---T-------------KE-NNVTV-VQ  119
CPZ.CD.90.ANT             ----TDPH....VVGVQT--LCA-GGSSG---S-H-D-..S-GA..............................Q-DSEGS--GGGTT..-LV--EIP-----MMEVL-Q--KCQ----------V-GIH-....Q-N----T---------S-Q--NKVP-Q-G  123
CPZ.CM.01.SIVcpzCAM13     ----G---..P-RE-...-QLCT--D..------S---W-PRE...............................G-EPRGTGRGEQTI.P-S----I-------IPVR-E---C-------------IQDIQ-....Q-I---------------K---NVS---E  122
CPZ.CM.05.SIVcpzEK505     ----G--S..--RE-...-K-PPDNNKE-----SN---W-S...........................GGEDHTGDRE-RKGEDRELSV.PTL----I------IL-V----EI-------------I--IQ-....E-K---------------K---NVI---Q  129
CPZ.CM.05.SIVcpzMB66      ----T---..Q-------...----P--E-..-T--S-N---R-Q-G............................GTC-EGGSEE-GDREQAV.S-A----IS-----V--VR-E--------------------IE-....G-------------------------NVT-D--  126
CPZ.CM.98.CAM3            ----R-VS..VERGIKET--LPAG-E..GTH-SAP---W-S-SNR.............................K-T-GGEGEGKGTPVS-V-L--I-------LGV--E---I-------------IDNVQ-....T-K-R---------------K---N-A---E  128
CPZ.GA.88.GAB2            ---GG---..PERE-...--L-PKEG..GT--S-S-D-R-CRA-T.........................DRKESE-GGEYNRKGEP..-LS---I-------LPVRVA---C-------------DNIQ-....E-K-R--L-------------K---NVE---E  126
CPZ.CM.98.CAM5            ----G-VS..VEREIKKT--L-AGRE..GTH-S-HG--W-S-NN-.............................GEKE-GEGTPVS-I-L--I-------LGV-VE---I-------------IDNVQ-....T-K-R-R---------------K---N-V---E  124
CPZ.CM.05.SIVcpzMT145     ---KK--S..-G-E-...--LCAR-E..GT--S-YGI-RIS.............................GGE-SQGGGE-GEQAAV.------I-------IPV--E---I-------------I-DLN-.....--K-S---------------K-FERVN---E  122
CPZ.TZ.00.TAN1            ----THPL.....VGVQT--LCA-HP...RERE......................GSGAGDSTDTSGANCP-TGDDDERRVL---N-----MM-V-VQ--VCQ-----------S-FCNIK-....K-Q-T--T---------VP-SE-YN-PVQ-G  119
CPZ.CM.05.SIVcpzLB7       ----T---..-P-----...-KL-------..--H--AN---W-SAK.............................NHPETGGQGRG-DREQAV.P-L----IS-----I----V--Y--G------------------DLE-....G---------------------K---N-P---G  126
CPZ.US.85.CPZUS           ----T-VPI.VERGIKET--LPGK-E..G-H-S-N-GFR-SR--S..............................DTGETGKGKGAL..-A----I-------LR--VA--IV-------------DNIQI....E-T-R---M---------K---HVN---E  125
CPZ.GA.88.GAB1            ----R---P..-RE-...-QLCA--N..-T-G--D---W-P-GR..............................-EPGEER-REQSI..-T-L--I-------IPV-VE---C-------------I-RIQ-....Q-L-----------------K-F-NVH---E  121
GSN.CM.99.CN166           ---PWPGA....PPSP..Q-L--NFPAANSPGG...............................GGRTPAAQSLDT-RGTPEQGTEGGEQEH-LSL-VIS--G--MKEVS----VVSM---------IVQDSAI..ELDTP-T---V-----L-Q--EHKHVEVIFN  128
GSN.CM.99.CN71            ---PWPGA.....TSP..Q-L--NFPAAISPGG...............................GGGTPAAQPV-AGGGAPEQGAEGRKEEH-LSL-VIS--G--MREVSM---VVSM---------IVQ-SAI..ELDTP-T--TV-----L-Q-KEHRHVEVIFN  127
MUS_1.CM.01.1085          ---GS-RP.....PEKEAE-LPPDLSLANCSAR-GGGIGQSLSSS.......................GSR-GETGQEE-LS--VIP-SG--VRPVE-E--VVSM---------IIQDAHI..HLE-P-N--SV-----N-S-KE-RGVQVCFN  124
MUS_1.CM.01.CM1239        ---GSSWP.....PEKEAK-LPPDLAFTNGPARAGGRPGQSLPAS.......................GS--WSERQEE-LS--VIS-SG--IREVM-E--VVHM---------IVQDSNI..QIDKP-N--LV-----N-S--E-RGVQVSFN  124
MUS_2.CM.01.CM1246        ----GSRL.....PVSEAE-LP-YIAHAYGTPRYS-AGELLP.TSRAQG....................RATQGQERQKD-LS--VIS--G--IK-VD-N--PVSM---------IIQDSDI..KIS-ELT--VV--L--L-N-K--KGVIVKFN  127
MUS_2.CM.01.CM2500        ----RFR......LPQEAE-LPCYIAHAISSPRYS-DS-LLP-GGSTEGE...................PNPPSPTGQEE-LS--VIP--G--IR-VMVH--PVSM---------IIQ-SEI..EIDKE-T--VV--L--L-N-K--KAVKV-FN  128
                          Pol p15 start (-1 from Gag)                                                                    p15 protease
MAC.US.x.239              ---PWSMG....KE-...PQ-PHGS......-ASGADANCSP.................RGPSCGSAKELHAVGQAAERK-ERKQREALQGGDRG-AA--FS--R--V--AH-E--PV-V--------SIVTGIE-....GPHYT--IV-------NTKE-KNVE--VL  135
H2A.GW.x.ALI              ---AWPMG....KE-...SQLPRNP......-SAGINTNSTP.SRASSGPAGAVYAAGEKAKRAEREAIQRGDGGLTA-R....AGRD-TQRGDRGLAA--FS--K--V--AY-E--PV-V--------SIVAGIE-....GSNYT--IV-------NTKE-EDVE-KVL  148
H2A.DE.x.BEN              ---VGPTG....KE-...SQLPRDP.......--SGADTNSTS.GRSSSGTVGEIYAAREKAEGAEGETIQRGDGGLAA-R....AERD-SQRGDRGLAA--FS--K--V--AY-ED-PV-V--------SIVAGIE-....GDNYT--IV-------NTKE-KNVE-KVL  148
H2A.SN.x.ST               ---VGPMG....KE-...PQ-PCGP.....N-AGADTNSTP.-RPSRGPTREVHAAREKAERAEREAIQRSDRGLPAAR....ETRD-MQRDDRGLAA--FS--K--V--AHVE--PV-V--------SIVAGVE-....GSNYS--IV-------NTKE-KNVE-RVL  148
H2B.GH.86.D205            ---VRTLG....KE-...SQLPHDP......-ASGSDTICTPDEPSRGHDTSGGDTICAPCRSSSGDAEKLHADGETTER....EPRE-LQGGDRG-AA--FS--R--V-KAC-E--SV-V-----V--SIVAGIE-....GSNYT--IV-------NTKE-KDVE--VV  149
H2B.CI.x.EHO              ---VRPLG....KE-...SQ-PRPG.....T-GDSAICAP.DEPSIRHDTSGCDSICTPCRSSRGDAKELHATREEAEG....EQRE-LQGGDRG-AA--FS--R--V-KAT-E--SV-V--------SIVAGIE-....GSNYT--IV-------NTNE-KNVE--VV  148
H2G.CI.92.ABT96           ---VWTLG....KE-...SQLPHDP.......--SGSDSISTT.................DGPSSRPTEELHAXGEEAKG....AEKQ-LQGGDRXXAA--LS--R--V--AY-ES-PV-V--------SIVAGIE-....GSNYS--IV-------NTKE-KNVE-KVV  131
H2U.FR.96.12034           ---AWPLG....K--...SQLPHGP.......--AGTDPNSP....SSRTSSGSADPDSPSSGPSSGSAKELHATGQETKG....EQRE-IQGGDGG-AA--FS--K--I--AY-E--PVKV--N-----SIVAGIK-....EPYYT--TV-E-E---NTKKFHNVN-KVL  145
MAC.US.x.251_1A11         ---PWSMG....KE-...PQ-PHGS......-ASGADANCSP.................RGPSCGSAKELHAVGQAAERKTERKQREALQGDDRG-AA--FS--R--V--AH-E--PV-V--------SIVTGIE-....GPHYT--IV-------NTKE-KNVE--VL  135
MAC.US.x.251_BK28         ---PWSMG....KE-...PQ-PHGS......-ASGADANCSP.................RGPSCGSAKELHAVGQAAER....KQREALQGGDRG-AA--FS--R--V--AH-E--PV-V--------SIVTGIE-....GPHYT--IV-------NTKE-KNVK--VL  131
MAC.US.x.EMBL_3           ---PWSMG....KE-...PQ-PHGS......-ASGADANCSP.................RGPSCGSAKELHAVGQAAER....KQREALQGGDRG-AA--FS--R--V--AH-E--PV-V--------SIVTGIE-....GPHYT--IV-------NTKE-KNVE--VL  131
SMM.US.x.H9               ---AWPMG....KE-...PQ-PHGP.....DASGADTNCSX.................RGSSCGSTEELHEGGQKAEG....EQRE-LQGGNGG-AA--FS--R--V--AY-EE-PI-V--------SIVAGIE-....GPNYT--IV-------NTKE-KDVK-KVL  131
SMM.US.x.PGM53            ---AWPMG....KE-...PQ-PHGP.....DASGADANCSP.................RGSSCGSTEELHEDGQKAEG....EQRE-LQGGDGG-AA--FS--R--V-AAY-EE-PV-V--------SIVAGIE-....GPNYT--IV-------NTKE-KDVK-KVL  131
SMM.US.x.SIVsmH635F_L3    ---VWPMG....KE-...PQ-PHGP.....DASGADTNCSP.................RGSSCGSTEELHEDGQKAEG....EQRE-LQGGDRG-AA--FS--R--V--AY-EE-PV-V--------SIVTGIE-....GPNYT--IV-------NTKE-KDVK-KVL  131
SMM.SL.92.SL92B           ---PWPMG....KE-...SQLPHAD........NIIDTVSTP.................RSSSPDSQGVSRESTEGKD....KEEQALQGGDGG-AA--FS-RR--V-RAL-E--PL-V--------SIVAGIE-....GPGYT-RIV-------NTKE-RTVK--VL  128
STM.US.x.STM              ---LWPMG....KE-...PQLPHGP.....NTSGADANCPS.................RRPSCGSAEKLHAAGQEAER....EQED-LQGGDRG-AA--FS--R--V--AH-E--PV-V--------SIVAGIE-....GLQYT--VV-------NTKEFKNVN--VL  131
MNE.US.x.MNE027           ---LWPMG....KE-...PQ-PHGP.....NASGADTNCSP.................RGPSCGSAKKLHAVGQTAER....KQKEALQGGGRG-AA--FS--R--V--AH-E--PA-V--------SIVAGIE-....GPHYT--IV-------NTKE-KNVE--VL  131
GRV.ET.x.GRI_677          ---VWPLG....RSE.TKKFCAIQRRHSWSGTNSPP.................NGNSLRSSKEAPPAVCREGTAPERGE....RTDKETEGERSGGC-LELP--R--MKRVI-E-TPVQ----------IIQ-KD-HFPPHKP-RS-VV-----G-H-KE-QGVQVQLE  143
VER.KE.x.TYO1_patent      ---VWTVD....-GK.TEKFSRRYS.......WSG..TECASSTERHHP.........IRPSKEAPAAICRERETTEGAK-...EST-NESGLDRGIF-EL.P--R--IK-VY-E-VPIK----------IIK-ND-..QLS-P-R--I------GLN-KE-NDREVK-E  141
VER.KE.x.9063             ---VWTVD....-DK.TEKLSRCHS.......WSG..TERAPSTDT............IRPSKEAPAAVCREGKTVEGTR-...ENTTNKSRLGGRIF-EL.P--R--IK-VI-E-VPVKV---------IIK-TDI..QLT-P-R--V------GLN-KE-SNVEVK-E  138
VER.KE.x.AGM155           ---VWPVD....-DK.TKKFSR-HS.......WGG..TKCAPSTEQ.........LYTLRPSKEAPAAVCRERETNEKSEQ...KPPSEQSRLERGIF-EL.P--R--IR-CI---TAVK----------IIKDTD-..QLR-S-R--IV-----GLN-KE--NVEVQLE  141
VER.DE.x.AGM3             ---VWPVD....-SE.TQKFSRRYS.......WGG..ANCAPSTES............IRPCKEAPAAICRQGEAVEGTK-...KTTSSESRLDRGIF-EL.P--R--IK-VY-E-VPIR----------IIK-AD-..QLS-T----I------GLN-KE-SDREVRLE  138
TAN.UG.x.TAN1             ----NANL....-NK.AAQF-RARRSR.....SDGSPDACARIP..............HRFASCGSIRPSQETAGAVCQKG..RSIEKAEGEGTGGL-LEFP--R--MSK-Y-Q--PVQV---------IIK-EDI..HLE---R--VV-----G-S-K---N--VN-D  141
SAB.SN.x.SAB1C            ---VWPLG....QRE.TQEFP-DLHQTNSSPNG-GLQQAGGKLVCRQTSDQRTRARRSSNSPVKAVCCSGETAETAVA......KPLA-TEPLRGGLQL---S--R--MK-VY-E--KVT---------S-IQGIE-....GDN---RI------C-N-KA-HNQEVK-E  155
OLC.CI.97.97CI12          ----MEGR...ESR-SPQHATEDRDGHLSLCTYPG..................................TDFGEG-K......GKERDEGGEET-LSL-SIS--R--MIELY-E--QVT-----------IKGKDI..YME-NGI-RI------A-R-M--RCRIRIF.  123
DRL.x.x.FAO               ----YSLG......QRETQ----GL..................ADPISSSIARIPPGGSGREDAPKVHGAGSSAEETT-...TTTAEEGPLRGGLQL--FS-RL--TTIVEVE--KV------------IQGLE-....T-K---QI------T-N----FNCE-K-A  137
RCM.NG.x.NG411            ----CSLE.....AGETEEPSLDFLETY.....CS-...........ERGRMESPVRSSRRDAQEIPSIGETAQTGAEER....EQGESGESLRGCVVL--FS--K--VTWVE-E--KV---------G--IQNLD-....GEN----I------Y-N-K-FNHCK-T-A  141
RCM.GA.x.GAB1             ----TPLG.....SGEAEELSFDFLDSLCSRDGEQLRP..................CRRDAKELSEEGRGTKETTEAG....RE-EERGSISG-L-L--FA--K--TTIAQ-E--KV-V----------I-GIE-....GND-T--I------Y-N-K--KNCE---A  137
MND_1.GA.x.MNDGB1         ---VWPLG......SLQTG-L-.........................GTRGDSNSSTIRGETSAENSEHLSEIRERA.....QAEDEGGEERGG-S--EYS-SR--IEEVSVD-VTIR----------IFN-RNI..KLK-N-Q--I------NLR-K---NVYV--R  130
MND_2.x.x.5440            ----YPLG......QWQAQ-LPCHA..................IDPISTPDARVGRPGRKDAVRLYEERATTEGS-R-....PER-KEGSLRG-LQL--FS--N--TTVVE-E--KV------------IKDLE-....T-N--RI------A-R----FNCKV--A  136
MND_2.CM.98.CM16          ----YPLG......QWQTQ-LP...GDAIDPNGP................SNARDGRPSRENAVRLHEERATAESGKQ.....TGKEREGSIRG-LQL--FS--N--TTVVE-E--KV------------IKDLD-....K-N---QI------S-N-K-FFNCKVT-A  135
MND_2.GA.x.M14            ----CSLG......QWQAQ-L-....CHATDPISTP...............DARDGGSSREDVIRLYEKGATTEGSKGD....QAGEGQGSLRG-LQL--FS--N--TTVVE-E--KV------------IKDVD-....T-N---QI------A-R-K--FNCK-TVA  136
DEB.CM.99.CM40            ----YSLG.....LQET.Q-L-RRNWEVHPGEHS-SGILD.............................YQGTSTGAGEQPPEGGNNEEREV.PG--GAEIS--R--MT-VH-E-RKVQV-I-------IIS-KDI..DLEAP-T--TV--L----N-KC-PG-E-AMA  131
DEB.CM.99.CM5             ----CSLG.....LPET.Q--FGRDWEIHPREHS-SG-FN.............................NQGTCARGKECSPGGHNDQEREV.PR--GAEIS--R--MTAVYVE-KRVQV-I-------IIS-KD-..DLDAP-T--TV--L----N-KC-PG-E-KMA  131
DEN.CD.x.CD1              ----WSWE....EESSQELCTVLPSDESGFAMEPSEG..............PAATTVPTKFSRGGKQTILSGRQEAEQAV...PSG-RGDNEGE-LS-VE-S--R--IGQ-WVQ--RV-V-I-------IIA-QDI.NLGTEG-I--TV-----Y-N-QA-PG-E-KLM  147
TAL.CM.01.8023            ----GFFR......EDPPGQGTE-LPSGATHRTAD.......................RASHSIMAGGDYHMDPSDQG....C-STQATDEGE-LSLVSQS--R--I-KVL-E-REV-V---------IIQ-GDI..ELE-VPHQ-TV--LA-Y-Q-QS-TGVE--WE  133
TAL.CM.00.266             ----VFFR......ENPTGQRTE-FSSGATIGTAN.......................RASHSIMAGGDYHVDPT-QE.....APRNNHNKGE-LSLFGQS--R--I-KVL-E--EV-V---------IIQVGEI..ELE-VPHQ-TV--LA-Y-Q-QS-SA-E-QWE  132
MON.CM.99.L1              ----SFRS......QEGAK-LP-HLCAAQRTDSAS.............................RDCGASQSQPRDWFP...EDGDPKGDEAG-LSLLSIS--G--IRSVL----KVSM---------IIQDQHV..NLGEN-T--VV-----M-E-K--KA-QVIFE  128
MON.NG.x.NG1              ----SFW.......PRQEGAKELSSNLLA----SGTN.....................PXGXQGELLTSSGTTRDP.......SADSPEGDEEG-LSLLSIP--G--VR-VE-A--SVSM----X----IVQXQDI..NLGEN-T---V-----L-Q-K-FKGVTVSFQ  132
SUN.GA.98.L14             ---VWPLE......QGAA--L--HAFCSA.................SRGSDIGEQNDTTPVKGRESSGDLQTVRTG-Q....STAEEEGKGGM-GAMLEYA-SR---EKVF-N--PVT----------IVS-D-V..SIE-H---RVV-----Q-R-KE-ADVFV--R  139
WRC.CI.97.97CI14          ----WQRA....DALEQLSSSEA-ELEQFWANNNNAFCSSNE-GSKH.........SGAGKIFIEGSTIEGTTGTQE......--GTEEGSGGAGMAL--IS-RT--MIQCI-E-HSVS-----------IDVNL.....NKN-Q-RI------TMN-KLHR-VK-R-L  145
WRC.CI.98.98CI04          ----WERA....DALKQLPSSKG-ELEQFWAG.NNTVCSSNEGGSKH.........TRIREIPPERSSIEGAGTGK.......KGREEERNGAARVA---IS-RN--MIQCI-E-NSVS-----------IDIHLN....NKN-Q-RI------TMN-KL-RK-Q-R-L  144
WRC.GM.05.Pbt_05GM_X02    ----WPRA....EALQQLSS.............SKN-CRGEYLWASG......ANSTSGSYHARTRAISSERSSDEEG......E-ERNAEGGESRISL--IS-RT--MIQCI--RESVS-----------VDVQL.....E-N-Q-RI------TMQ-KL-KDVR-K-L  136
LST.CD.88.524             ---VWSLE......EWTF-QL-NHGRGC-SD-FSAT...................HGESASNKSREGNRDIQEPRTTTEE.....ATTSSTEVCGRALPEFLL-SR--IEKVLVN--EVT----------IFS-N-V..RIE-PYT-RVV-----Q---KEFRDVF-Q-A  138
LST.CD.88.485             ---LWTLE......EWSS-KL-FDGGEC-SSTISSP...................NGKELNKSRESLRNIQEPRTTTEE.....ATAGSTEMCGRAMPEFLL-SR--IEN-LVNR-PVS-----R----IFS-N-V..KIE-PYT-RVV-----Q-R-KE-KDVF---A  137
LST.CD.88.447             ---LWTLE......EWSS-KL-FDGGEC-SSTISSP...................NGEEPNTSRESLRNIQEPRTTAEE.....ATADPTEMCGRAMPEFLL-SR--IEN-LVN--PVS----------IFS-N-V..RIE-PYT-RVV-----Q-R-KE-KDVF---A  137
LST.KE.x.lho7             ---IWSLE.....QQTSRQLS-TGRS....CRE.NSISSP................NGECSNKSRESTGDLQDFRTTAEK...AAAASPTEMCGRALPEFTL.SR--IES-LVN--PVS----------IFS-N-V..RIE-PYT-RVV-----Q-R-KE-RDVF---A  137
SYK.KE.x.SYK173           ----RMGG.....LQEAPSQLP--ERAICSP-G.........GHRRWAMAHMVSTDEPTSTSEGTELSLEETPHQQG-ALA...EGEQWEGGDQE-LSLS-LS--R--MIEVDVE-D-VQM-V-------IIR-EDI..QLHQP-S--IV--L--N-T----RN-RFTVV  151
SYK.KE.x.KE51             ----RVGE......QETPSK--CDEGDTNSPTSGG..-AMAAT-MGKLCFTPTEAHSNGHPATATVLSKDSQGPLGPWETF...REERRDEGEGT-LSL-SIS--R--MME-NVE-TKVTM---------IIR-QDI..QLHQP-T--IV--L--N-R----N--RFQ-E  157
COL.CM.x.CGU1             ---GNQGGG.....-TDSHLPP..............................................EDEQR...PATQEGEGGE.-LSLLE-S-RR--I-IVEVE--KCQ-----------FK-EDV..QLT-A-TAVQ-Q----A-R-K--RDKILKMG  105
```

```
                                                               protease end p66, p51 RT start                   M41L                    K70R
                                                                                                                              D67N                                              D110 catalytic site
H1B.FR.83.HXB2          GHKAI.....GTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETV...PVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKS......VTVLDVGDAYFSVPLDEDF  279
H1A1.UG.85.U455         -K-T-.....---------------M----------------...------E--------------------T---N-------------------------------------------------------------T--------......-----------------S-  278
H1B.US.90.WEAU160       -----.....---------------------------------...----------------------------------------------------------------------------------------------S-------......------------------  279
H1C.ET.86.ETH2220       -K---.....---------------M---L-R----------...--------------------------TA--E---Q----R-------------------------------------------------------------......----------------G-  286
H1D.CD.84.84ZR085       -----.....---------------------------------...--------------------------T----D-------R---------I-------------------------------------------------......I----------I--C---  279
H1F1.BE.93.VI850        -----.....---------------M----------V------...--------------------------T---L------------------------------S-------K-----------------------------......---------------K--  279
H1G.SE.93.SE6165        -K---.....--I-----I------M-----------------...----------R---------------T---K---E-----------------I-----------------------------------------------......------------------  276
H1H.CF.90.056           -K---.....---------------I-----------------...--------------------------T-------------R-------S--I-----------------------------------------------......-S-------------KE-  278
H1J.SE.93.SE7887        -K---.....----I----------M---L-------------...-----------I--------------TQ--A---E----RV--------------------------------------------------------......---------------Y---  277
H1K.CM.96.MP535         -Q---.....---------------------------------...--------------------------T-------------------------------------------------------------------------......---------------K--  279
H101_AE.TH.90.CM240     -K---.....---------------M---------------D--...--T----------------------T---K---E------------------------------------------------------------------......-----------------S-  284
H102_AG.NG.x.IBNG       -K---.....---------------M-----------------...--------------------------TD------------------------------------------------------------------------......---------------K--  279
H103_AB.RU.97.KAL153_2  -K---.....---------------M---L-------------...--T-----------------------TD--K--------------------------G------------------------------------------......---------------Q--  279
H104_cpx.CY.94.CY032    -K---.....---------------M---L-------------...--------------------------T----D----------------I--------------------------------------------------......---------------PE-  278
H1N.CM.95.YBF30         -R--V.....---------------------------------...-----------------T---E--R-------------R---------I-----------------------------------------Q---......------------C---K--  286
H1O.BE.87.ANT70         -REVQ.....---------------I--GL---------AP-...----------------SK---E--TA--Q---Q----R---------I------G-----------E-----------G---Q-Q-......------------C---P--  275
H1O.CM.91.MVP5180       -KEVQ.....-------------L---I--GL---------AP-...----------------SR---E--TA--Q---Q----R---------I-------------------------G---QRQ-......------------C---P--  275
CPZ.CD.90.ANT           DRTVL.....A---L--N--------V-CLL---------KV---...-----E----R------SK---E--K---DKL-A-N---R---D-----I------TS---------------I----------Q---......------------I---Q--  279
CPZ.CM.01.SIVcpzCAM13   -K-VQ.....---------------I--------V----S-----...-------------RI-----SA------T---Q-----------------I-------S-------------------------E---......------------C---K--  278
CPZ.CM.05.SIVcpzEK505   -K--V.....---------------F---------------I...---------------R---------T---------------R---------I----------------------------------------......------------C----N-  285
CPZ.CM.05.SIVcpzMB66    -KR-V.....---------------I-------------------...--S-------R---------R--T-------------RV--------I-------------------------------Q---......------------C----N-  282
CPZ.CM.98.CAM3          -R-TT.....---------------I-------------------...------------R------A------T---Q-----------R---------I----------------------------------------......------------C----N-  284
CPZ.GA.88.GAB2          -KRVE.....S--------------V---L-----------TP-...-------T---R------SK---E--T---R------------R---------I---R------------------------------------......---P--------C----N-  282
CPZ.CM.98.CAM5          -R-TT.....---------------I------------V---...------------R------A------T---Q------------------------I------------E--------------------------......------------C----N-  280
CPZ.CM.05.SIVcpzMT145   -K-IT.....S--------------V--KL-----------D--...K-F----------------SK---E--T---K-------------------------I------------E--------------------------......------------C---P--  278
CPZ.TZ.00.TAN1          NKEVR.....A-----E--I------I-K-L-----------V-...K-Q--E-------------SK---E--T---KTL--------AV---------I-------TS-----------------L-----------R-RNM......------------I---P--  275
CPZ.CM.05.SIVcpzLB7     -K--L.....---------------I-------------------...--Q---------------------T--------Q----R---------------------------------------------------R-......------------C----N-  282
CPZ.US.85.CPZUS         -R--Q.....-S-------------I-------------S-K--...--R--------R------A------T---Q--------TRV---------I-----------------------------------R-......------------C---KE-  281
CPZ.GA.88.GAB1          -R-VV.....---------------I---L----V----S-----...-------------------SA------T---Q----------------------I----------------------------------------......------------C---K--  277
GSN.CM.99.CN166         EKRIK.....A----------L---C-SKL-V---MVQQK--PI...E-H--A-KE---I-----SR---E--TQ-TQ--LRL-QLE---------S---------KSQ--M-I---K--EA---LA-----------IPHPAGLEQKEH--II-MK---Y-I--Y---  290
GSN.CM.99.CN71          EKRIK.....A----------L---C-SKL-V---MVQQKL-P-...E-H--S-KE---I-----SK---E--TQ-TQ--L-L-QLE---------S---------KSQ--M-I---K--EA---LA------------EQ-EH........--II-MK---Y----Y-E-  283
MUS_1.CM.01.1085        -KTIR.....A--------I--M---C-SKF-I---MIQGKL-PI...K-T--E-AK--M-------Q---R--ED-TQQ-L-V-Q-EE--------S-C---R---KS---M-I---Q--EA--ELA------------IEHEH........--IV-IK--FY-----P--  280
MUS_1.CM.01.CM1239      EKTIK.....A--------I--M---C-SKF-I---MIQEK--PI...K-A--E-AK--M-------Q------EG-VQQ-L-LDQ-EE---D----S-C---R---KS---M-I---Q--EA--E-T-------------AEHEH........--IV-IK--FY-----PA-  280
MUS_2.CM.01.CM1246      EKRIV.....A----S---I--L---C-SKLEI---MVMAGQQLK.PTK-S--ERKK--M------SK---E--K--TQK-I-L-QLEEA--N----S-----R---K----M-I---K--EV--E-A---T-----S--AQ-AH........--IV-MQ--FY-----KE-  285
MUS_2.CM.01.CM2500      DKRI-.....A--------I--L---C-SKL-I---MVGGQL-PT...K-G--A-KE--M------SK---E--E--TQR-L-L-QLEET--Q----S-A---R---KS---M-I---K--EA----S---T-----S--S--AH........--IV-MQ--FY-I--Y-E-  284
                                                                protease p51 RT start                                                                                                D catalytic site

MAC.US.x.239            -KRIK.....--IMT-D--I--F------AL-MS-----AKV-P-...K-A----K----L-----SK---V--R---EK---D-QLEEAP-T------T-------KN---M-I------RV----T------------A-R-R......I----I------I----E-  291
H2A.GW.x.ALI            NKRVK.....A-IMT-D--I--F---I--AL-MS--L-VAK--PI...E-R----K----LR-----K---E--K---EKT-R--QLEEAP-T------T-------KN---M-I-------V----T-I-----------A--RR......I------------I--H-S-  304
H2A.DE.x.BEN            NKRVR.....A-IMT-D--I--F---I--AL-MS--L-VAK--PI...K-T----K---RL------K---E--K---EK-----QLEEAP-T------T-------KN---M-I-------V----T-I-----------A---R......ISI---------I--H---  304
H2A.SN.x.ST             NKRVR.....A-IMT-D--I--F---I--AL-MS--L-VAK--PI...KIM----K----LR-----K---E--K---EK--R--QLEEAP-T------T-------KN---M-I-------V----T-I-----------A---R......I-----------I--H---  304
H2B.GH.86.D205          -KRVR.....A-IMT-D--I--F---I-NTL-M-----VAKV-P-...K-E----K----IR----SR---L--K---EK-----QLEEAP-T------T-------KN---M-I-------V----T--N#-------fAE-RR......I--I---------I---PN-  303
H2B.CI.x.EHO            -KRVR.....A--MT-D--I--F---I-NSL-M-----VAR--P-...K-Q---EK----IR----SK---L--K---EK-----QLEEAP-T----S-T-------KN---M-I-------V--E-T-------------AS--R......I---------------P--  304
H2G.CI.92.ABT96         -KRIQ.....A--MT-D--I--F---I-VKL-MS-----AKV-P-...K-R----------IR----SK---Q--R---DK--Q--QLEEAP-T------T-T-----KN---M-I------RV----T------------AE-RR......I----I------I---V--  287
H2U.FR.96.12034         -KTIR.....--IMT-D--I--F------EL-MS-----GKEAL-...IGG----K----LR----SK--LE--K---EK-----QLEEAP-T------T-------KN---M-I------RV----S------------AE--R......I-----------I--C---  301
MAC.US.x.251_1A11       -KRIR.....--IMT-D--I--F------AL-MS--L--AKV-P-...K-A----KV---L-----SK---V--R---EK---D-QLEEAP-T------T-------KN---M-I------RV----T------------A-R-R......I----I------I----E-  291
MAC.US.x.251_BK28       -KRIK.....--IMT-D--I--F------AL-MS--L--AKV-P-...K-T----KV---L-----SK---V--R---EK---D-QLEEAP-T------T-------KN---M-I------RV----T------------A-R-R......I----I------I----E-  287
MAC.US.x.EMBL_3         -KRIK.....R-IMT-D--I--F------AL-MS--L--AKV-P-...K-A----KV---L-----SK---V--R---EK---D-QLEEAP-T------T-------KN---M-IH-----RV--ELYRSPIR-------A-R-R......I----I------I----E-  287
SMM.US.x.H9             -KVIK.....--IMT-D--I--F------AM-MS--L--AKVXPI...K-T----K----LR----SK---I--R---EK---D-QLEXAP-T------T----XX-GN---M-I-------V----T---------X--A-RRR......I-----------I----E-  287
SMM.US.x.PGM53          -KVIK.....--IMT-D--I--F------AM-MS--L-VAKV-PI...K-T----K----LR----SK---V--R---EK---D-QLEEAP-T------T-------KN---M-I------RV----T------------A-RRR......I-----------I----E-  287
SMM.US.x.SIVsmH635F_L3  -KVIK.....--IMT-D--I--F------AM-MS-----AKV-PI...K-T----KE----R----SK---I--R---EK---D-QLEEAP-T------T-------KN---M-I-------V----T------------A-RRR......I-----------I----E-  287
SMM.SL.92.SL92B         -KVIR.....E-LMT-D--I--F---I-AAL-VS--Y-VAKV-YT...K----E-----RL-----SK---Q--T---EK-----QLERAP-T------T------NKD---M-I-------M--E-T-------------EMER......I----I----------PE-  284
STM.US.x.STM            -K-IK.....--IMT-D--I--F------AL-MS----VAKV-P-...K-T----K----I-----SK---E--K---EK---D-QLEEAP-T------T-------KN---M-I-------V----T-I----------A-RRR......I-----------I----G-  287
MNE.US.x.MNE027         -KRIK.....--IMT-D--I--F------AL-MS-----AKV-P-...K-T----K----LR----SK---V--R---EK---D-QLEEAP-T------T-------KN---M-I-------V----T------------A-R-R......I-----------I----E-  287
GRV.ET.x.GRI_677        DKIIT.....-SI-I-S--I------I-A-A-MK-VMGVLSSQIE.ETK-Q--E-K----L-----SR---E--T---KQ--E---L-R--G---------------K-Q--M---------A----F-----------Q---Q......I--I-I----Y-I--CKE-  301
VER.KE.x.TYO1_patent    DKILR.....--I-L-A--I--------APA-AR-VMGQLSEKIP.VT-----E-AR--C-R----SK---E--Q---SQL-Q-----RV-G--A----I-C-----KSQ--M---------A----F-----------R-MRQ......I---------Y-I---PN-  299
VER.KE.x.9063           DKILR.....----L-A--I--------APA-AR-VMGQLSEQIP.IT-----D-AR--HI-----SK------T---EQL----------G--A----I-C-R---K-Q--M---------V----F-----------Q-M-Q......I----I----Y-I---PE-  296
VER.KE.x.AGM155         DKILR.....----I-A--I------F-A-A-AK-VMGQLSQTIP.IT--R--E-AR--RL-----SK---I--Q---KTL-E---L-RV-GD-A------C-R---KSQ--M---------A----F-------------M-Q......I-II------Y-I---PE-  299
VER.DE.x.AGM3           DKILR.....--I-I-S--I------I-APA-AK-VMGQLSEQIP.IT-----E-AR--FL-----SK------Q---DQL----------G--A------C-----KSQ--M---------A----F---------S-FE-MTE......I----I----Y-I---PE-  296
TAN.UG.x.TAN1           -KETR.....-N--I----I--V---I-APL-AR-IMGQLNEKIP.ITK-Q-----S---L-----S-------T--TQAL---------V-G---------C-R---K-Q--M-I---A---A----F-------------R--Q......I----I----Y-I---PE-  299
SAB.SN.x.SAB1C          DKTCK.....A-I---E---------V-A-L-V---LTQRE--PI...K-H----Q---RIR----SK---E--KA--EDL--Q-HLER-------------R---K-Q--I-M---Q---S----Q-----------QQREQ......I----I------C---P--  311
OLC.CI.97.97CI12        EKEIQ.....EA----D----VV---V-AAV-IR-VCAQLSDKIP.ITK----E-HKP-QI-----SK---E--EK--DRL-E----Q-AEAS-QW---I-C-----KNEY---I---A--AA------------------EREQ......I----I------I---PEY  281
DRL.x.x.FAO             -KITH.....AS--------------V-CKL-A-----V-QV-P-...K-T----------I-----SK---E--K---EQ-----Q--R-----------C-----G----------Q---A----F---------G---ACEQ......I-----------C---P--  293
RCM.NG.x.NG411          -KTTH.....AH--------------V-SKL-AS------KA--I...K-E----Q---R------SK---E--T---NA--------R----------I-C-------------------------F---------G---QCRQ......I----I------C--Y-P-  297
RCM.GA.x.GAB1           -KRTH.....AH--------------V-KKL-A-------Q---I...K-E--S-Q---R------SK---E--T---NA-------------------I-C-------------------------F---------G---QCER......I----I------CL-Y-P-  293
MND_1.GA.x.MNDGB1       -KGTF.....----I----ID-----IMEKL-GK-ILAQLSDKIP.ITK------V---RI-----SK---VG-QK--DRL-E-----RVD-G-N----I-------KNE----I-------L-------H-L---------I--C-R......I----I------I---P-Y  288
MND_2.x.x.5440          -KITH.....AA--------------V-KKL---------K----...K-T----T---RI-----SK---L--Q---NQ---------RV-----------C-----G-S-------Q---V----F------T--G---QCEQ......I----I------C---V--  292
MND_2.CM.98.CM16        -KTTH.....AS----------V---V-KKL-------V-KV---...K-T----T----I-----SK---L--Q---SQ-----Q---------------C-----G-S-------Q---V----F---------G---QCEQ......I----I------C------  291
MND_2.GA.x.M14          -KSTH.....AS--------------V-KKL---------K----...K-T---------RI-----SK---L--QT--D---------R------------C-----G-S-------Q--RV----F---------G---QCEQ......I-I--I------C----S-  292
DEB.CM.99.CM40          NKV-V.....SDL---D--I--L---Y-AKM-VSI---VLDYTP-......---E-A----I-----SR---E--KD-LKP-IEA-Q-VPAA-T----S---V-----KS---M-I-L-A---A--EV----T-----SA-PQMEQ......I----LA---Y-IA--PT-  285
DEB.CM.99.CM5           DKV-V.....SD----D--I--L---Y-AKM-VSI---VLDFTP-......---E-A---RI-----SR---E--K--IRP-VEA-Q-VPTE-T---SS---V-----KN---M-I-L-A---A---V----T-----S--PQMEQ......I----LA---Y-I---PQ-  285
DEN.CD.x.CD1            DKV-V.....D--------I--L---Y-SKL-V---MAVQQMRPQ.VT-------F---RI------T--L---QQ-V-Q-----RLE-AP-T--------V-P---K-QY-M-I---K--EA-----------------QQ-EH......I----LK---Y-----P--  305
TAL.CM.01.8023          HKR-R.....D---I----I--L---F-AKF-V---MVSTE-PVI...K----E-KEP--I------K---EG-QA-IDD-V-A-QLEE--------S-----R---K-Q--M-I---A--QA----A---Q-----S--PQ--Q......I----MK---Y-I---P--  289
TAL.CM.00.266           NKR-R.....D---I----I--L---F-AKFSV---MVS-D-PIT...K----E-KEP--I------R---EG-QI-IDD-V-A-QLEEV-------S-----R---K-Q--M-I---A--QA----A---Q-----S--PQM-Q......I----MK---Y-I---P--  288
MON.CM.99.L1            DREVW.....A--------I--L---V-AKM-V---MAAGGDLQPPSLE-T--A-KE---L-----SR---E--TQ-TQ--L-L-QLEPTE-N----S-C-------KS---M-I-------A----F-----------E-MDH........--IV-MK--FY-I--W-P-  287
MON.NG.x.NG1            -KTVR.....X---L-X--I--L---V-SRM-I---MXQGIAQSP.SLT-T----K----L-----SK---E--TT-XR--IEL-QLEEAE-T----S-C---R---K----M-I-------A----F-----------QNMAN......I-IV-MK--FY-I----N-  290
SUN.GA.98.L14           DKT-T.....-----------D-----I-SAL-GR-VLATLSEKIP.ITK----E-AA---------SR---EG-QK--DRL-A-----RAD-G------I-------KNE----I---V--GM----H-L------------C-Q......I--V---------I---P-Y  297
WRC.CI.97.97CI14        -KTIE.....ADIMT-K--I------V-NKM-MH-A-AQLSDKIP.ITK-S----CS--R------SKK--EG-QK--DRL-A-----RAD-G------I-------KNE----I---K--EV----F---------G--R-N-N......IS-V-I------I---PEY  303
WRC.CI.98.98CI04        -KTVE.....ADIMT-K--I------V-EKM-MH-A-AQLSDKIP.ITK-S----CT--R------SK---EG-QK--DRL-A------AS-G------I-------KNE----I---K--EV----F---------G--R-N-N......IS-V-I------I---P-Y  302
WRC.GM.05.Pbt_05GM_X02  -KEVR.....ANLMT-K--I------V-EKL-MH-A-AQLSDKIP.ITK-S----CQ--RI-----SK---EG-QK--DRL-A-----RAEAN------I---R---KNE----I---G--AV----F---------G--H--RN......I----I-----FI---PEY  294
LST.CD.88.524           -KTTV.....----L-----D-V---I-NPL-AR-VLAQLSEKIP.ITK-H----C---R------SK---EG-QA--DRL-------PVD-G------I-------KNE----I---K--EL----H-L---------I--C-Q......I--V-I------I---P-Y  296
LST.CD.88.485           -KTTV.....----L-----D-V---I--SV-AK-VLAQLSDKIP.ITK-R----C---R------SK---EG-QA--DRL--------ADAG------I-------KNE----I---Q--AL---LH-----------I--C-Q......I--V-I------I---P-Y  295
LST.CD.88.447           -KTTV.....----L-----D-V---I--NV-AR-VLSQLSDKIP.ITK------C---R------SK---EG-QA--DRL--------ADAG------I-------KNE----I---Q--AL----H-----------I--C-Q......---V-I------I---P-Y  295
LST.KE.x.lho7           -KTTV.....----L-----D-V---I--AM-AK-ILAQLSDKIP.ITK-S----C---R------SK---EG-QA--DRL-------PVD-G------I-------KNE----I---K--EL----H-L---------I--C-Q......I--V-I------I---PNY  295
SYK.KE.x.SYK173         KPSGKRKQVE--L---------L---I--KL-VK-VMVQTAL-P-...K-S---DKEL-RL-----SV--LE--KA-VED-L-A-QLE-AS-T---------V-R---KK---M-I---K--EA----F---------G---RQ-.......L-II-LK---Y-----KE-  311
SYK.KE.x.KE51           RPDGRIKEVE-SL---------L---I-SKL-AK-VMIAQKL-PM...K-Q----KEL--L------R--LE--RA-VSD-LEK-QLE-AS-T---------V-Q---KL---MPM---K--EA---------------G--ST-R.......M----LK---YT-----E-  317
COL.CM.x.CGU1           -KEYR.....-DI-I-H--I--V---M-KEGKLVVAQLSDR-PVT...K-T--E-----------SK---EG-QK--ERL-A----E-AELG------I-C-R---KNE----I------------M---------G--IE-TH......I----I------I---P-Y  261
```

```
                                                                                                   K219Q
                                                    polymerase motif                        T215Y
H1B.FR.83.HXB2               RKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDS..WTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEV  447
H1A1.UG.85.U455              -----------------V--------------S--------------S-H--------------------------A------A--S--FI------------------------------Q------.--------------------A----K--------A----DI  446
H1B.US.90.WEAU160            -----------------------------------------------------------------------------------------F---------D--------------------K----E-..--------------------A----K--------------I  447
H1C.ET.86.ETH2220            ----------T-------------------P------PQ------AP--E----------------------AP-----E---K--F------------------------------Q------.----------------------------------A----DI  454
H1D.CD.84.84ZR085            ------------------------------------I----------EV----------------------A---K--E-----------------------------------S-T----E-..-------------------------K--------------  447
H1F1.BE.93.VI850             K---------V-------------------C----------MK-------------------------------E------F----------------H------------Q--N---.--------------------------P-------A----DI  447
H1G.SE.93.SE6165             -----------------V------------------R------AN--EM----------------------A------E---K--F------------------------------Q--D-E-..-------------------------TH-------A----DI  444
H1H.CF.90.056                ----------------------------------------A---E---EM----------------------A------A---K--F----------------------------TVK------.---------------------N---K--------A----DI  446
H1J.SE.93.SE7887             ------------------------------C------K---ER--E------------------E---R--K---E---K--F------------------------------Q----ED..-------------------------K------K-A----DI  445
H1K.CM.96.MP535              -----------------V------------H-----------IK--EM--------------------P--------E---K--F------------------------------Q--D---.-------------------------K--------V----DI  447
H101_AE.TH.90.CM240          -----------------------------------------IK--EM-----K-----------------------A---S--F---------------------R------E-------.--------------------A----K--------A----DI  452
H102_AG.NG.x.IBNG            ----------V-------------------A----------TK--E----------------------A------G---K--F-----------------------------VE------.--------------------A---IK---R----A----DI  447
H103_AB.RU.97.KAL153_2       ----------T-------------------------------------E---------------------E-----E------F-----------------------------------.--------------------A--------------A------  447
H104_cpx.CY.94.CY032         ----------T-------------------C----------FK--E----------------------A------E------F------------------------------QPA----.-------------------------K--------A----DI  446
H1N.CM.95.YBF30              -------------------------------T---------EKH-E-I--------------LA---EAV-D--D---K--F------------------------------K-----V..--------------------R-K-----I---------  454
H1O.BE.87.ANT70              ----------V-------------------------D---RD--ELE-C------------PLTE--KR--L--E--YQ--F----------------------------S-Q--N--V..----------I----------Q--R--E----I----S----  443
H1O.CM.91.MVP5180            ----------V-----V-------------------D----S--EVE---I-----------PLAE--KRV-L--E--YQ--F------------------------------Q--D-EV..---------------------Q--R-KE---I----S----  443
CPZ.CD.90.ANT                ----------V--------C----------A------A---DKY-AVE-------------M--TA--EM--K-----QV---E------------Q---------------K-K---P-D..-------------------------TK---R-I--V-S--DR  447
CPZ.CM.01.SIVcpzCAM13        ----------V-----------------------------AK----TV-------------D-T---Q-V-------H--F------------------------------K-----EV..------L---I---------------IK-----I----R--D-  446
CPZ.CM.05.SIVcpzEK505        ----------V-------------------T----------N--EL-------------------T---EAV-R--S---T--F----------------------------T-Q-----T..--------Q------------------K-----I--A------  453
CPZ.CM.05.SIVcpzMB66         ----------V-------------------------------------E-I-------------K-EL--E-V----A---K--F------------------------------Q----E-..-----------I------C------R-K-----I---------  450
CPZ.CM.98.CAM3               ----------V-----------------------I---T---Q-H--LI--------------LRE--E-V-L------T--F------------------------------Q--Q-ET..-------------------------K-----IK------DI  452
CPZ.GA.88.GAB2               ----------V-----------------------------V-D---QK--EL---------------SLEE--KRV-Q--E------F------------------------------Q--Q-EN..------------------------IK-----I--A-----I  450
CPZ.CM.98.CAM5               ----------V------V-------X---------------D-----H--LI------------------K--E-V-L------T--F------------------------------Q--Q-EI..--------------------S---K-----IK---------  448
CPZ.CM.05.SIVcpzMT145        --F-----------------------------S--------------TKH-E-I--------------P-EE--KRV----A---K--F----------------------------Q-K---QEV..------------------------IK-----I--A-G----  446
CPZ.TZ.00.TAN1               ----------L--N---K-F------------------D----EH--VD--------I----NEEE--KL-KK------T---E-----Y-EK---M--------N-----N-T---PEQ..----H--------------H---TKE----I--V-G---P  443
CPZ.CM.05.SIVcpzLB7          -----------------------------------------R---------------------K-E---N-V----A---K--F------------------------------Q----E-..--------------------------IK-I-R----V-NS---  450
CPZ.US.85.CPZUS              ----------V---------------------------D-----H--VI--------------NLEK--E-V-L---Y--T--F--------E-------------------Q--Q-EI..-------------------------IK-----IK-A------  449
CPZ.GA.88.GAB1               -----------------V--------------S------------EK----T------------------D---K-V--------K--F------------------------------Q----EV..----------I---------------IK-----I----K--D-  445
GSN.CM.99.CN166              ------SV--V--QA-AE---FK-------A--T---NIVASL-RQI-D-E--V-LI------LI---RKL-E--QVVHKI-NL-TSYNIQ--EA---PDY-VK-L------KG-RL--VE--DQ-T..---------------V---H---RTK----CI--A-G--DQ  458
GSN.CM.99.CN71               ------SV--V--Q--AE---FK-------A--T---ATVASL-YQI-D-E--V--I------LI---RKLAE--QVVHKI-NL-TSYNIQ--EA---QDY-VR-L------KG-RL--VE--DQ-I..---------------V---H---RTK---RCI--A-G---Q  451
MUS_1.CM.01.1085             ------SL-QV--LG-AR---FK----------T---FTAA-L--EIK-A---VTLI------LI--NR-LSS--KVVAQI-NM--KL-IQ--QE-F-EDY-VQ-L------RG-KI--VA--QQEV..------V-----------A-A-----TK----CI--V-S---E  448
MUS_1.CM.01.CM1239           ------SL-QV--QG-AR---FK----------T---YTAA-L-QEI-EA----TLI------LI--NR--SG--RVVAQI-NM--KL-IQ--ED-Y--DY-VQ-L------RG-KI--VT--QQEN..-----------------A-A---V-TK----CI--V-S---E  448
MUS_2.CM.01.CM1246           -P----SV-AV--MG-AK-F-FK----------T---ATTA-L--QV--D-Q-AL-V------LI---R---E--RLVKKI-DL-SSK-IQ--ED---PDY-VE-L------KG-RIK-VE--DQ-T..----E-----------A-V-S---TKH--RCI--I-G---L  453
MUS_2.CM.01.CM2500           -P----SV--V--LG-AK---FK----------T---ATTA-L--QI--E-Q-AL-V------LI---R---E--RLVKRV-DL-TSK-IQ--EE---QDY-VE-L------GG-RIR-VE--DQET..----E-----------A---S---TKH---CI--V-----T  452
                                                    DD catalytic site
MAC.US.x.239                 -Q-----L--V--AE--K--I-K--------------YT-RHV------A---VTLV-----ILIA--RTDLE-DRVVLQSKEL-NSI-FS--EE-F--D---Q------W-T--KL-K-E--QRET..-----------V----A-------TKH--R-I--KMT---E  459
H2A.GW.x.ALI                 -Q-----L--V--AE--K--I-K--------------HT-RQ-------A-Q-VILI-----ILIA--RTDLE-DKVVLQ-KEL-NGL-FS---E-F--D--YK----G-W-T--KL-K-Q--Q-EV..---------H--V----A-------TKH--R-I--KMT---G  472
H2A.DE.x.BEN                 -Q-----L-AV--ME--K--I-K--------------YT-RQV------A---VILI-----ILIA--RTGLE-DKVVLQ-KEL-NGL-FS---E-F--D---Q---C--W-T--KL-KLQ--Q--I..-----------V----A---S---TKH--R-I--KMT---E  472
H2A.SN.x.ST                  -Q-----L-----AE--K--I-K-S------------YT-RQV------A----ILI-----ILIA--RTDLE-DRVVLQ-KEL-NGL-FS---E-F--D--YQ------W-T--KL-R-Q--Q-EV..-----------V----A-------T-N--R-I--KMT---E  472
H2B.GH.86.D205               -Q-----L--V--AE--K--I-K---------QS-C-Y--R-V-D----A-S-VI-I-----ILIA--RSDLE-DRVVSQ-KEL-NDM-FS--EE-F--D---K------W-K--KL-K-Q----EV..----A------V----A-LF----T-HI---I--KMT---E  471
H2B.CI.x.EHO                 -Q-----L-AV--AE--K--L-K--------------YT-A-V-D----A-N-VT-I-----IL-A--RSDLE-DRVVSQ-KEL-NNM-FS--EE-F--D---K------W-K--KL-K-Q----EV..-----------V----A-LF----T-HI---I--KMT---E  472
H2G.CI.92.ABT96              -Q-----L--V--AE--K--I-K--------------HT-R-V-D----A-E-XIXI-----IL-A--RTDLE-DQVVLQ-KEL-NKX-FS---E-F--D--YQ-----XW-K--KL-K-N-X--ET..-----------V-X--A-L-----TKN-X--I--XMT---E  455
H2U.FR.96.12034              -Q-----L-----AE--K--I-K----------T---YT-RNL------A-E-A--I-----ILIA--RNDLE-DQLVSQ-KEL-NKT-FS---E-F--D--HK------W-K--KL-K-E---R-V..-----------V----A-V-----TKH----IK-KMT---E  469
MAC.US.x.251_1A11            -Q-----L--V--AE--K--I-K--------------YT-RHV------A---VTLV-----ILIA--RTNLE-DRVVLQ-KEL-NSI-FS--EE-F--D---Q------W-T--KL-K-E--QRET..-----------V----A-------TKH--R-I--KMT---E  459
MAC.US.x.251_BK28            -Q-----L--V--AE--K--I-K--------------YT-RHV------A---VTLV-----ILIA--RTDLE-DRVVLQ-KEL-NSI-FS--EE-F--D---Q------W-T--KL-K-E--QRET..-----------V----A-------TKH--R-I--KMT---E  455
MAC.US.x.EMBL_3              -Q-----L--V--AE--K--I-K--------------YT-RHV------A---VTLV-----ILIA--RTDLE-DRVVLQ-KEL-NSI-FS--EE-F--D---Q------W-T--KL-K-E--QRET..-----------V----A-------TKH--R-I--KMT---A  455
SMM.US.x.H9                  -Q-----L--V--AE--K--I-K-------------XHTXRNV------A---VTLI-----ILIA--RTDLE-DRVVLQ-KEL-NSI-FS--EE-F--D---Q------W-T--KL-K-E--QRET..----N------V----A-------TKH--R-I--KMT---E  455
SMM.US.x.PGM53               -Q-----L--V--AE--K--I-K--------------HT-RNV-G----A---VTLI-----ILIA--RTDLE-DRVVLQ-KEL-NSI-FS--EE-F--D---Q------W-T--KL-K-E--QRET..-----------V----A-------TKH---I--KMT---E  455
SMM.US.x.SIVsmH635F_L3       -Q-----L--V--AE--K--I-K--------------YT-RNV------A---VTLI-----ILIA--RTDLE-DRVVLQ-KEL-NGI-FS--EE-F--D---Q------W-T--KL-K-E--QRET..-----------V----A-------TKH--R-I--KMT---E  455
SMM.SL.92.SL92B              -Q---------V--QE--K--I-K--------------AT-RQV------A---VLLV------LI--NRGLTE-DKMVTQ--DM-NNL-FS--ED-F--N--LQ----L-Y-K--KL-K-E----ER..-----------V----A-------TKN---MI--KMT---E  452
STM.US.x.STM                 -Q-----L--V--AE--K--I-K--------------YT-RN------RA---VTLI-----ILIA--RTDLE-DRVVLQ-KEL-NNL-FS--EE-F--D---Q------W-T--KL-K-E--QR-V..-----------V----A-------TKH----I--KM----E  455
MNE.US.x.MNE027              -Q-----L--V--AE--K--I-K--------------HT-RNV------A---VTLV-----IL-A--RTDLE-DRVVLQ-KEL-NSI-FS--EE-F--D---Q------W-T--KL-K-E--Q-ET..-----------V----A-------TKH--R-I--KMT---E  455
GRV.ET.x.GRI_677             ----------V--TG------F-C---------T---NTAAN---EIKRHT-GLE-V------WLA--HDETR-NQQVDIV-KM--EK--E-----V-R---WE----K---N---INK-E--PLEG.E----K---V--V-------------TKHT-AM---K-N-L-E  470
VER.KE.x.TYO1_patent         ---------TV--QG------F-C---------T---NTAAS---EIKRNL-ALT-V------W---QENEHT-DKLV-Q--TK-QA---E--E--V-----YE----K-W-H--ELSR-Q-E---E..----------------A-L---LRTKNI---I--K-N-L-L  467
VER.KE.x.9063                ----------V--QG--V---F-C---------T---NTAA----KIKEELKPLT-V------W---QEDEYT-DKLV-Q--VK-KT---E-----V--K--YE----K-W-Q--QLNS-N-E--EQ..----------------A-L---LRTK-I-R-I--K-N-L-E  464
VER.KE.x.AGM155              ---------TV---G------F-C---------T---NTAS----EIK-ELKQLT-V------W---QE-GPK-DQLVQT--NR-QE---E--E--V-R----E----K-W-H--KL-S-E-EK-EQ..------L----------A-L---LRTKNI------K-N-LD-  467
VER.DE.x.AGM3                ----------V--QG--T---F-C---------T---NTAAS---EIK-ELKPLT-V------W---QEDEYT-DRLV-Q--MK-SA---E-----V--K--YE----K-W-H--QISS-E-ED-EE..--------R--------A-L---LRTKN----I--K-N-L-T  464
TAN.UG.x.TAN1                -P-------QV--VG------F-C---------T---HTAQRV-KEIG-KVQ-LE-I-----MWI----PEKE-DE-V--V-KL--LR-FE-----V--R--YD----K-Y-SH-SLNK-EI--QEH..-----------V------L-----TKA-------AQD--A-  467
SAB.SN.x.SAB1C               Q---------V--RE--------K---------T---TTAN---QE--QK---VD-------MLIA--RPKAE-LVMVQQ--DY-ET--FK--E--F--D--Y-------Y-K--QL-E-T---REE..--------------------T---TKH--R-I--ARP---I  479
OLC.CI.97.97CI12             -E--------V--DR--R--I------------C----TVAHLM--W--E--E-DL--------------PRAE-LE-VKQ--EL-RSYN-E--P--Y-ED--YE----T-Y-KR-KIAR-K----EE..-----------Q--------S---TKA----IK-P-G-LDK  449
DRL.x.x.FAO                  ----------V--QG----------------------ATAD---Q---EK---VI--------F----RPKHL-EQM-K---T--KF---E-----F-DN--YE----Q---K-----E-R--D-EI..-----------V--------S---TKE----I--A-P-D-E  461
RCM.NG.x.NG411               --------LA---QG----------------------TTAN--------E---V---------F----RTAHE-SQM-KR--E---T--FE-----F-DK---E----V---------E-K-L--EE..--------------------S---TKE----I--V---DDK  465
RCM.GA.x.GAB1                ---------AV--QG--V-------------------A-AN---Q---EE---VI--------F----RTKLE-DKM-KQ--D---F--FE-----F-DK--Y---------KS----E-K----EE..--------------------S-LRTKE----I--A---D-K  461
MND_1.GA.x.MNDGB1            -P-----V--V--QA--K--M------------C---GTVASL--V---NH-TVQL-------F----YTAEE-EKA-V---AL-MT-N-E--E--Y------H----------KIEKVQ---LAEQ.P---E-----------A-L-----TK-----I--GLNI--K  457
MND_2.x.x.5440               ----------V--QG----------------------ATAD---KA-KEKH-EVL--------F-----SASE-SKM-V---E--QF---E-----F------E----V---K-----K-Q----E-..--------------------S---TKE----I--ARP-DKI  460
MND_2.CM.98.CM16             ----------V--QG----------------------ATAD---Q---ERH--V---------F----RVAPE-SRM-Q---D---F---E-----F------E----I---K-----KVQ----EE..--------------------S---TKE--R-I--A-P-D-K  459
MND_2.GA.x.M14               ----------V--QG----------------------ATAD---KT-KEEY-EVL--------F-----SASE-NKMVNR--E--RF---E-----F------E----V---Q-----K-Q----ET..--------------------S---TKE----I--A-P-D--  460
DEB.CM.99.CM40               A------V--V--IQ--E--EFR------SA------A-VGRQ-QI--EK--ELI-V-------L-----RKD--LK-VAQ---F--ER--R--PE-Y-ED---H--------K--RLM-VT--DEEE..---HK------Q--------S---TKH--RAI--VPG--DP  453
DEB.CM.99.CM5                AR--------V--TQ--E--EFR------SA------ATVGRQ-KL-KEKY--LI-V-------L---N-DKVS-LEQVKQ--KF--EK--R--PE-Y-ED--YH--------K--HLL--E---EEE..---HK------Q--------S---TKH---AI--VPG--DK  453
DEN.CD.x.CD1                 AQ-----V-------G--Q--IFK------V---T---ATAD-M-SQ---AH--VILI------L-A--RSLE---KMVK---DF-AV--FE---A-F-ET--IQ--------KH-QI-K-D--LPEDGQ----A------V------L-S---TKE---SI--V-D-G-T  475
TAL.CM.01.8023               AR--------L--DR--R---FK----------T---YTSDQL-RK--EKY-QVALF------LI---ESLE---RIVG---AL-NKV-IQ--EA-F-DK--IK-L----Y-E--K--N-E----QT..------L--------------S---TKA---C---V-G-L-E  457
TAL.CM.00.266                SQ--------L--DR--R---FK----------T---HTSDQL-RK--DKY-QVALF------L---NESLEE--RIVG---AL-NKV-IQ--EA-Y-DK--IK-L----Y----K----E--D-LT..-----EL---I------C---S---TKA--RY---V-G-L-E  456
MON.CM.99.L1                 ------S---L--AE-AK---FK----------TV--HTAAELPQEI--KY-EVTLI------LI--NHSLPE--KIVGVI-AT---K-IQ--PE-F-DDY-VQ-L------SH-KIAK-E--DQEQ..-----E----I------A---S---TKN---SI--ARS--DT  455
MON.NG.x.NG1                 ------S---V--AE-AV---FR----X-----T---Y-ASELXKXI-RKY-EVTLI------LI---HSLPE--KIVTAI-TX-SEK-IQ--IE-Y-DKY-LQ-L------TH----K-E---Q-Q..-----G-----------A-------TKH--XCIX-X----DI  458
SUN.GA.98.L14                ----------V--QA--K--V------------C---GTVASL--K--R-Y-EVQL-------LI---Y-KKK-EEIVKQ---L-ME-N-E--E--Y-G---YK----I----R-EIEK-K--PLEEE.P-------I--VI-----L-E-LRTKE----I--N-P-A-K  466
WRC.CI.97.97CI14             -Q-----V--V---S--K--V------------C----TVASL--I--GK--N-QVL-------I----DKTE-LN-V----KL--T-N-E--E--L-L---CN----I------QI-E-K---EVK...----V----------A-L-E--QTK---R-I--N-P---E  470
WRC.CI.98.98CI04             -Q-----V--V---S--K--V------------C---ATVASL-DV-K----TMQ-L-------I---F-KPR-LE-V-----L--T-N-E--E--L-L---CK----I----R-QI-E-K---EVR...----V----------A---E--QT------I--N-P---E  469
WRC.GM.05.Pbt_05GM_X02       -Q-----V------K--T--V------------C---ATVASL--I---KH-KVQ-V-----------Y-KKE-LQM-Q---DL--T-N-E--E--L-L---CK----I-----EI-E-K---EIR...----V---I------A---E---TKE----I--N-P--AE  461
LST.CD.88.524                -Q--------L--QE--K---------------C---ETVSGL-AE---L----LL-------FI---R-KEG-SQAVK---EL-IT-N-E--E--F-DK--YH----V----R-EIERVK---I-TRKT---E------V----A-L-D--RTKE----I--V-P-EQL  466
LST.CD.88.485                -Q-----V-----QE--R---------------C---ATVAGL-SE---L----IL-------LI---RDKKG-DQAVK---EL-MT-N-E--E--F-A---YH----V----R-EIEKVK--DI-LQKT---N------V----A-L-E--RTKE----I--V---E--  465
LST.CD.88.447                -Q-----V-----QE--R---------------C---ATVAGL-SE---L----IL-------FI---RDKKG-DQAVK---DL-VT-N-K--E--F-A---YH----V----R-EIEKTK--DI-LQKT---N------V----A-L-E--RTKE----I--V---D--  465
LST.KE.x.lho7                ----------L--QE--K---------------C---GTVAGL-SE---L---MI--------FI---R-RKG-DQAVK---EL-MT-N-E--E--F-A---YH----V----R-EIEK-K---M-LTKT---Q------V----A-L-D--RTKE----I--V-P-E-I  465
SYK.KE.x.SYK173              -P-----V-----AS--E---FT----------T----TINQ--Q----KYS-LTLI------LI-T-RSEKA-QEIVQQIVTA--KV-FKV-KE-W-DQY-MQ-L--T------QL-K-E--NI-DE.I---QL---I-V---------S---TKE---CI----Q----  480
SYK.KE.x.KE51                -Q-----V-----SS--E---FK----------T---ATVNQ--Q-I-EK-S--I-V-------L---NRSEKE-GQIVS-IVKS--AV-FSI-PE-W-DKF-LQ-L--T------SL-KVQ---ITES.P---EL--II-V-------------CKA-S-CI--A----DI  486
COL.CM.x.CGU1                Q---------V--LA--T--V-K----------T---GTVAR-------RGQL.Q-A-------L----PLEE--RAVQ---DA--YY--E--E--Y-AD--YK--------KQ-KL-KVKI--Q-K..-----------A---L--L----RTKEIS-CTK-K-H-L-E  428
```

```
                                                                                                                                                                                  p51 RT end p15 RNase H start
H1B.FR.83.HXB2              IPLTEEAELELAENREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEP.FKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTP.KFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTN  615
H1A1.UG.85.U455             VT-----------------D-------------V--------D----------.-----------K-S-----------V---VS---------I-.--R---------A--M---------------------------------D--A-------------------------D  614
H1B.US.90.WEAU160          --I-----------------------------------L-----------------.-----------V-----------------------------------.---------------------------------------------------------------------K------------  615
H1C.ET.86.ETH2220          VT---------------------F--------------ND---F-F----.---------F-KRGT----------AV----AL-----------.--R---------A---D-------------------------------A-V--------------I-------D  622
H1D.CD.84.84ZR085          --------------------M-------------L-----------------.-----------------------------AI-----------.--R------------ID-------------------------------I-----------K-----------D  615
H1F1.BE.93.VI850           V---A-------K-----R---------------------D---------N-.-----------KV-S----------------AL--------RSP------L----D----D-------------------------T---A--D-------S-----K--------D  616
H1G.SE.93.SE6165           VS--A---M---------R------------E----V----LD----------.Y---------KRGS-----------V----A-----------.------R-----I---D--------------------R--T---P----Y---------------------D  612
H1H.CF.90.056              ----K-------------R--I---------------R-------------------.----------K--T-----I---------S--------I-.--R------------------------H---------T---A----Y-I---------------------D  614
H1J.SE.93.SE7887           V---R---------K-------------SA-E----V----LD----------.-----------KR-S---------A-V----AL-A--------.--R---R--------D--------------------------M----------S-----T-------D  613
H1K.CM.96.MP535            V---A-------------------------------------ND----------.H-----------S------------------A--G-------.--R--------------------------------------T----------------H----K-R-----D  615
H101_AE.TH.90.CM240        V------------------T-------------V--V-----D----------.-----------RGS------R-----V---VA-----------.--R---R--------M-------------------------D--------------S-------------D  620
H102_AG.NG.x.IBNG          VA---------------------------T---V--L-----D----------.-----------KK-S-----------V---VAM----------.--R---R--------M-------------------------D-----------------------I-------D  615
H103_AB.RU.97.KAL153_2     ----A-----------------------------V---------------------.-----------L------------------A-----------.-------------------------------------------------------------------S-------D  615
H104_cpx.CY.94.CY032       V---T--------------------A---------------------------.H----------T-S------R----------AM-C---------.--R---------D---------------------------TD--A-----------------Q-------D  614
H1N.CM.95.YBF30            VNF-----------------L--------G-E-V-------------------L.H---------K--S-----I---V-V-R-VA-----------.--R--V---V--A---DH-------------------------T---S-----------------------F--D  622
H1O.BE.87.ANT70            V--SR------E----R--Q-------Q-D---WVN-----GE---------E.H--------T-QKAS----IR--A-VI--VSQ---I----L-.-----VTR-------AD-------------S----I----R--S---M----Y------------------E  611
H1O.CM.91.MVP5180          V--SK------E----K----------Q-D---WVS---H-E------V--DE.H-----------QKAS----IR--A-V---VSQ-A------L-.--R--VTR-------A--------------S----I-------T------------------N-----------E  611
CPZ.CD.90.ANT              VQM-R------E--KQ--QQKIE-Y--Q-GLP-K-T-----S---------NE.G-L--A-----PT-T---E-R--AGV----GL--------EV-.--Q---TR---DA--SD------------------IR---N-LAD--PE------------NSQ---------D  615
CPZ.CM.01.SIVcpzCAM13      V---P------------VST--------N-D-E----V----E--------F--Q.HR----------Q-ST----IR--A-VI---AV-----------.R-R--V---S-----A----------------I-----------S--TD--PT-D-Y-------------K-----I-D  614
CPZ.CM.05.SIVcpzEK505      VT--R-------------------A--N-D-E--------------------DL.H---------K--ST----IR----V---VAL-----------.--R--V---V--------------------D-----------------T---S----Y------------------F--D  621
CPZ.CM.05.SIVcpzMB66       VTF-T-------------------A------E------------------F--Q.Y-------------S-----------V---VAL---------V-.R-R-----------A---D-----------Y-------------QD--P---------------------------D  618
CPZ.CM.98.CAM3             VTM-Q---M--E--X---XD---------E-E----X----NS-------F--Q.H----------Q-S-----IR--A-V----A-------------.--R--V---V-----S--------D-X------------------T---P-------------------K-------D  620
CPZ.GA.88.GAB2             VEM-R------E------------T----E-E----V---NN--------F--T.GR----------Q-A-P---IR--A-V-P-VSP---IL--TV-.-----V-R-V--A---D--------------------R-----------------Q----Y-----------RR------I-  618
CPZ.CM.98.CAM5             VAM-Q---M--E-------D---------E-E----V----NS-------F--Q.H----------Q-S-----IR--A-V----A-------------.--R--V---V-----S--------D------------------P---P-------------------K-----I-D  616
CPZ.CM.05.SIVcpzMT145      VA--P-----------Q---------A----E-E------------------F---.H---------GKQ-A-----IR--A-T----A-------------.--R--V---M--A--AD----------------I-------------T---E-----Y-------------R--------  614
CPZ.TZ.00.TAN1             VEM-R------E--KQ----K-Q-A----KLP-Q-A-------------------E.G----------KSP-T---EIR--AGLI---GN---I---IV-.--L--VS----SQ---D---V--V-----I-----IR---N-LSD--PE-------------DS-K-R-------  611
CPZ.CM.05.SIVcpzLB7        VTF-R--------------N---------A------V-----S-----------Q.------------QGS----------A-VM---AI---------I-.--R-------A---A----D------------Y-----SH--F-----Q---P---------------------------D  618
CPZ.US.85.CPZUS            VNF-H---M--E------------------E-E-V--V----RS-------F--R.H----------Q-S-----IR--V-V----A----------V-.-----LV----V-----S--------D------------------K--T-A-E-----------Q-----K-------D  617
CPZ.GA.88.GAB1             V---P------------VST----------D-E----------NC-------F----.H----------Q-S-----IR--A-------A-------------.--R--V---S--A--A-------------I-----------S--T---PTTD-Y-------------T-------D  613
GSN.CM.99.CN166            VE-----QI----------QSSE-G---AE-P-VV--TSL-EQ--G-MFS--G..-M--SA-F-KT-N--I-SYQ-FAD-LT-VGR--Q-T---V-D--RI-VV--Q-DA--MN--------TI-AIH--H-LRQ--T-VS--LED-V-Y---------TS----------  626
GSN.CM.99.CN71             VE-----QI----------QSSE-S---AE-P-VV--TSL-EQ--G-MFS-DS..-M-RS--F-KT-T--I-SYQ--AD-MT-VGR--Q-T---V-D--RI-VV--Q-D---MN--------TI-A-H--H-LRQ--T-VS--LEN-V-Y---------TS----------  619
MUS_1.CM.01.1085           VQ-----QA-----Q---QQT-Q-A----EQP-V---VSL-DS--G-NFT-SK..GV-R---F-KV-SV---SYQ--AD-MA--GR--L-----P-D-VRI-VI--Q-DQ--IAH--VS---TI-A-H-TH-L-Q-FT-VP--LKE-P-Y---------VS----------  616
MUS_1.CM.01.CM1239         VQ-----Q------Q----QA-Q-A-F--EQP-V---VSL-DS--G-NFL-NK..GI--S--F-KV-SV---SYQ--SD-IAR-GR--L-----P-E-VRI-VI--Q-DQ----H--VS---DI-A-H-TH-LRQ-FT-VP--LSE-P-Y-------HKVS----------  616
MUS_2.CM.01.CM1246         VE-S-Q-Q-----------QEAG-A----E-P-VL--VSL-EQ--G-TFT-DR..NM-R---F-KI-T--S-PYQ--A--LSRASK-AL-C-----D-CRI-VV--Q-DN--ADS--T----DI-A-H--Y-LRQ-FT-VP---ET-P-Y---------NS---R-----D  621
MUS_2.CM.01.CM2500         VE---Q-Q----------M-QE-K-A----E-P--L--VSL-EQ--G-TFT--K..NM-----F-KV-T--I-TYQ--A--LA-AAR--L-C-----D-CRI-VV--Q-DH--A-S--T--V--I-A-H--H-LRQ-FT-VP--LEA-P-Y---------NS---R-----D  620
                                                                                                                                                                                p51 RT end p15 RNase H start
MAC.US.x.239               VQW--M--A-YE--KI--SQEQE-C--QEG-P-E-TVI-SQDN--S-K-H--D..-I--V--F-KIKNT---G-RL-AHVI---GK-A-----QV-.--H--VE-DV--Q---D---V------D-IS-----R-VFN-V-D--E-E--Y-T--SC-KQS-E-----I-D  626
H2A.GW.x.ALI               VQW--L--A--E---I--SQEQE-H--QEE-E-E-TV--DQDN----K-H-GE..-I--VE---K-KNT---G-RL-AQV----GK-AL----RI-.R-H--VER----Q--DD---V----D-D--S-----R-AFN-V-D--L------T--PC--QS-E-----I-D  639
H2A.DE.x.BEN               VQW--L--A--E--KI--SQEQE-Y--QEE-E-E-T---SQGH----K-H--E..-I--V----KIKNT---G-RL-AQV----GK-AL----RI-.--H--VER----Q--DN---V------D--S-----R-TFN-VGD--P------T--SC--QS-E-------D  639
H2A.SN.x.ST                VQW--L--A--E--KI--SQEQE-C--QEE-E-E-TV--DQDN----K-H-GG..-I--V----KVKNT---G-RL-AQV----GK-AL----RI-.--H--VERD---Q--DN---V----D-D-IS-----R-VFN-V-D--L------T--SC-KQSRE-----I-D  639
H2B.GH.86.D205             VQW--L--A--Q--KI--EQEQE-S--KERVP-E-TV--NLAN----K-H-GN..-V--V----KVKNT---G-RL-AHV----GK-AL----EI-.V-H--VER---DQ---D---V------D--S----IR-A-N-V-D-LE-R--Y-T--SC--TS-E-------D  638
H2B.CI.x.EHO               VQW--L--A-FQ--KI--EQEQE-S--KEGVP-E-TV--NLAN----K-H-GD..-I--V----KVKNT---G-RL-AHV----GK-AL----EI-.M-H--VER---DQ---D---V------D--S----IR-A-N-V-D-LE-V--Y-T--SC-KAS-E-------D  639
H2G.CI.92.ABT96            VQW--L--A--Q--KI--XQXQE-A--KEXEPXE-TV--NLDN----KVH-GN..RI--V----KIKNT---G-RL-AHV--X-GK-AL-----L-.F-H--VERD---Q---D---V----D-D--S--H--R-V-N-V---LEQT--Y-T--SC-KTS-X-------D  622
H2U.FR.96.12034            VQW--M--A-F---KI--SQEQE-A--REEEN-E-TVL-NQDN----K-H-GD..RT--V--F-KVKNT---G-RL-AHV----GK-A------I-.M-H--VER-V--Q---D---V--VLD-D--S-----R-A-N-V---LL.E--Y-T--SC--AS-Q-------D  635
MAC.US.x.251_1A11          VQW--M--A-YE--KI--SQEQE-C--QEG-P-E-TVI-NQDN--S-K-H--D..-I--V--F-KIKNT---G-RL-AHVI---GK-A-----QV-.--H--VE-DV--Q---D---V------N-IS-----R-VFN-V-D--E-E--Y-T--SC-KQS-E-----I-D  626
MAC.US.x.251_BK28          VQW--M--A-YE--KI--SQEQE-C--QEG-P-E-TVI-SQDN--S-K-H--D..-I--V--F-KIKNT---G-RL-AHVI---GK-A-----QV-.--H--VERDV--Q---D---V------D-IS-----R-VFN-V-D--E-E--Y-T--SC-KQS-E-----I-D  622
MAC.US.x.EMBL_3            VQW--M--A-YE--NI--SQEQE-C--QEG-P-E-TVI-SQDN----K-H--D..-I--VR-F-KIKNT---G-RL-AHVI---GK-A--V-QV-.--H--VERDV--Q---D---V------D-IS-----R-VFN-V-D--EVE--Y-T--SC-KQS-E-----I-D  622
SMM.US.x.H9                VQW--M--A-YE--KI-PSQEPE-C--QEG-P-E-PVI-SQDN--S-K-H--D..-I--V--F-KIKNT---G-SL-AHV----GK-A-I---QV-.R-H--VEX-I--Q---D---V------D--S-----R-VFN-V----Q--------RSC--QSRE-------D  622
SMM.US.x.PGM53             VQW--M--A-YE--KI--SQEQE-C--QEG-P-E-TVI-SQDN--S-K-H--D..-I--V--F-KIKNT---G-RL-AHV-E--GK-A-----QV-.--H--VER-I--Q---D---V------D--S-----R-VFN-V----Q---------SC--QSRE-------D  622
SMM.US.x.SIVsmH635F_L3     VQW--M--A-YE--KI--SQEQE-C--QEG-PIE-TVI-SQDN--S-K-H--D..-V--V--F-KVKNT---G-RL-AHV----GK-AL----EV-.--H--VER-I--Q---D---V----D-D--S-----R-VFN-V----Q---------SC--QSRE-------D  622
SMM.SL.92.SL92B            VQW--L--A------I--NQEQE-R--REDEP-E-TVL-NQDN--S-K-H-GD..RI--V--F-KIKNT---GIRL-ANV----GK--L--------.F-H--VER-V-DQ---D-----------D-IS----IR-VFN-V-D--EKE-VY-I--SC--NS-E-------D  619
STM.US.x.STM               VQW--M--A-Y---KI--SQEQE-Q--RED-P-E-TVV-DQDN--S-K-H--D..-I--V--F-KIKNT---G-RL-AHVI---GK-A-----QI-.--H--VER-V--Q--AD---V------D--S-----R-VFN-V---LE-T-------SC--QS-E-----I-D  622
MNE.US.x.MNE027            VQW--M--A-YE--KI--SQEQE-C--QEG-P-E-TVI-NQDN--S-K-H--D..-I--V--F-KIKNT---G-RL-AHVI---GK-A-----QV-.--H--VE-DV--Q---D---V------D-IS-----R-VFN-V-D--E-E--Y-----SC-KQS-E-----I-D  622
GRV.ET.x.GRI_677           -VW-----A-YKN-QG-VQ-TQE-T----L-E---TV----E------FT--G..AV--V-R--KQ-ET----LRT-AHL----CK-ALT---RL-.RVQ--VD-K--DM--QD---VS--------S--L------S-V----K-EDVY------SKV--------LSE  637
VER.KE.x.TYO1_patent       VTW-P---A-Y---A----TEQE-T--K-GIPIR-AV--LEG---S--FK--G..QV--V---TKQKNT---ELRT-AGL----CK-AL----IL-.VLE---ER-V--Q--AD---VS-----D--S----L-----T-T----PKEDVY-----C--NS-E-----ISQ  634
VER.KE.x.9063              VQW-----A-Y---K----TEQE-T--A-G-P-R-AV--LED---S--FK--G..-I--V----KQKNT---ELRT-AGL----CK-AL----EL-.I-E--VER-V--Q--AD---VS--------SV--------T-T----P-EDVY-----C--NS-E-----I-Q  631
VER.KE.x.AGM155            VEW-P---A-YE--K----TEQE-T--A-E-P-R-AV--L-D---S--FK--G..-I--V--F-KQKAT---ELRV-AGV----GK-AL----QL-.T-E--VERD---Q--AD---VS-----D--SV----T---T-T----P-EDVY-----C--QS-E-----I-Q  634
VER.DE.x.AGM3              VTW-----A-Y---K----TEQE-T--K-GRPIR-AV--LEG---S--FK--G..QV--V---TKQKNT---EFRV-AGL---LCK--L----EL-.VLE---ER-V--Q--AD---VS---D----S---------T-T----PKEDVY-----C--NSRE-----I-Q  631
TAN.UG.x.TAN1              VQW--Q--V-FRQ-Q-----EQQ-A-------S-R-T-T-LDD---G--FK--G..-V------TKSKNT---EFRV-AGM--RVCK-A-TF---L-.IMEV-VER-V--Q--SD---V--L--I--IS----I----N-V-D--PK-KVY-C---C--NS------FL--  634
SAB.SN.x.SAB1C             VQW--------E---Q--RQKQQ-Q----ALP-R-KVL-L-D---G------PE.N-I--V----KIKT----ELRM-AGL----GK-------QI-.IME--VER-L--Q--SD---V-------M-S--Q-IR---K-V-D--P-EAVY---------NS-E-----L-D  647
OLC.CI.97.97CI12           VEW----KE--E--------EQE--M-KEEEP-QCTLCYTKGTVAYRIWQKHG...T----RFQ-KMT----KFR-M-G-IE--GK--------L-.I-QV-TTRKD--A-AFDD--VH---DI---SNSENQRTFWE-VPA-M.EV--------Y--KS-E------SS  614
DRL.x.x.FAO                VEW-R-----YE--KL-I--QM-----Q-E-P-K-K---L-N---S---E-DD.N-P--------KVKNS---MRM-AGL----AK-AL----RL-.I-Y--VER-V--Q--Q----V----D----IS--H-I-----N-LSD--P-E-VY---------VS-E-------S  629
RCM.NG.x.NG411             VEF-R-----YE--KAV---KI-----N-E-P---KV--LEG---S---E-GD.S-P--------KQKT----EIRM-AGL----CK-A-----L-.--E---ER-V--Q--SD---VS---------S-----IR---N-V-D--P-E-VY---------NS---------D  633
RCM.GA.x.GAB1              VEM-K---I-YE--KM----KL-------EK-P-V-N---LEG---S---E--S.G-P-------KQKT----EIRM-AGL----AK-A-----RL-.T-R---ER-V-D*-RSQ---V----D-----S-----IR-G-N-V-D--P-E-VY---------NS-I---------  628
MND_1.GA.x.MNDGB1          VTM----R--YEQ-K---A-EQE-S----N-E-YVRF--TTG-DISF-WK-G..N-V-RA---GKQKT--S--LMK-AG-T--VGR--------FV-.-MQI-TTR-I--D--H----C-----V--IS--M-ERE--S-SP--LE-V--Y---------DS-M-----I-D  624
MND_2.x.x.5440             VEW-R-----YE--KL-VQ-E------Q-E-P-M-KV--LT----S---E--D.N-P--V-----TKN----ELRV-AGL---VAK-CL-----L-.--Y--LER-V-DQ--HD---V------D-IS----IR---N-L-D--P-EDVY---------NS-E-----Y-S  628
MND_2.CM.98.CM16           VEW-R-----YE--KL-VQ-E------Q-E-P-M-KV--LT----S---E--D.N-P--V-----TKN----ELRV-AGL----AK-AL-----L-.--Y---ER-V-DQ--P---------------S--H-IG---N-LR--VP-EDVY---------NS-E--------A  627
MND_2.GA.x.M14             VEW-R-----YE--KL-VQ-E------Q-G-P-M-KV--LT----S---E--E.N-P--V-----TKH----ELRV-AGL----AK-C-----EL-.--Y--LER-V-DQ--HD-------------IS----IR---N-L----P-EDVY---------NS-E-----Y-A  628
DEB.CM.99.CM40             VT-S---QA--R--K----QE-S----KEQEP----LS-L-K---G-V-R-PK..GI-----FS-DK---Y--FH--AK-MY--G-----F--RI-.Q-R--VV--E-DN--HNH---A---D--AIH--H------E-VS---PD-D-Y----------S---------E  620
DEB.CM.99.CM5              VS-S---QA--R--K-TP-QE-S----KEDEP--S--S-I-R---G-V-R-TK..GL-----FSKDK-T-Y--FH--AK-MY--G-----Y---V-.--R--VV--E-DN--HNH-------D--A-H--Q--R---E-VS---PD-D-Y----------S-M-----ISE  620
DEN.CD.x.CD1               -S---G--A--K-AQ-----SIS-S--R-E-P-----S-L-EN--G-T-K--K..HL------HKDK-S-Y-PYQ--AKVMA--GR-AL----RL-.T-R--VS-VD-DA--S----VN---D--M-S----LRNF-N-VS---DK-P-Y-T------Q--Q-R----AT  642
TAL.CM.01.8023             VT------A------A--A-ET--S--KEGEP-E--LT-LAE---G-MVK-GK..RL-----F-KQ-T--S-PYQ--VG-M---GK-A--F--QV-.I-RV-VV--E--Q--S----CS-L-TIVPIH-----R---N-VS---E----Y------H--S-E-------T  624
TAL.CM.00.266              VQ-S----A---------A-ET--T--QEGEP-E--LT-LAE---G-M-K-GK..RL-----F-KQ-T--S-PYQ--IG-M---GK----F--QV-.V-RI-VV--E--Q----H--CS---TIVPIH-----R---N-VQ---E--D-Y------H--S-E-------A  623
MON.CM.99.L1               VV-SDL-QA-----------EST-G----NEP-KV-LTSLAEN--G-RFF--K..SV-----F-KI-ST-S-TYQ--AD-LA-LGK-A--T--RL-.I-R--VV--Q-DA--ADN--IN-V-DI-A-YS-H-LRQ--T-VQD--E--P-Y---------TS-E---------  622
MON.NG.x.NG1               VE-SDL-QA---------SQQST-G----QL--KV--TSLXE---G-XYF--..XRV-----F-KIKN--S-TYQ--AD--T-LGK-AL-----L-.I-R--VN--Q-D---ADN--V-----I-P-HS-H-LRQ--T-VQX--EH-P-Y---------AS-E-------S  625
SUN.GA.98.L14              VKM----RE-YQS-Q-V-Q-S-S-S--E-D-E--CRV--VK--IL-F-WL-G..KQV-RV-R-QKRGA--E-PCQ--AA-L---GR-------FV-.-IQV---R-I-SQ--AD---C-----L---S--K-EQE--T-AT--VP-.D-Y------EKLE-R-----I-Q  632
WRC.CI.97.97CI14           VEW----LE--EI-KQ----EQS-A--KGKDLIAEL-KTGPGQWAYTIREIGE.K-P--V---G-Q-HT-S-ELRSMAH-M---G--------RV-.--RI--E-DI-DS--SD---VS---DT-H-S--F-ARE--K-VL--LPEV--Y-I--S---DS-E-----.WT  637
WRC.CI.98.98CI04           VTW----LE--EI-KS----EQM-A--TGKELVAEL-KTGSGQWAYSIREAGE.K-P--V---G-Q-HT---ELRS-AH-M---G---------V-.---I--E-DI-DS--SDH--VS---DT-H-S--Y-ARE--KIVL----EV--Y-I--S---DS-E----F.WT  636
WRC.GM.05.Pbt_05GM_X02     VEW-M---A--EI--A----SQQ-S--GGGDLIAEV-KTGTGQWAYTIKEAGE.T-P--V--FT-I-HT---ELRT-AHCI---G--A------L-.--RI-VE-DI-DM--S----------V-H-S--F-KRE--T-VS---E-V-EM-I--S--K-S-E-----.WT  628
LST.CD.88.524              VTW----LE-YET-KT----K-Q-A------E-VVRV-QNKK-VI-F-WR-G..NAI-RA-R-Q-QKS--M-PLQ--A------GK-------FV-.-IQV-VTR----N--INF--V----DL--IS--Q-EQE--S--A---E-TD-Y------E-VG-T-----I-Q  633
LST.CD.88.485              VTW-DA-LE-YEQ-QQ----KLQ-A----T-E-VVRV-QNKK-VI-F-WR-G..NSI-RA-R-Q-QKA----PLQK-A------GK-------HV-.-IQV-VTR-V-DQ--SDH--V-----L--IS--Q-EQE--S--A---E-VD-Y------EKVG-T-------Q  632
LST.CD.88.447              VTW-DA-LE-YEQ-QQ----KLQ-A----T-E-VVRV-QSKK-II-F-WR-G..NSI-RA-R-Q-QKS----PLQK-A------GK-------HV-.-IQV-VTR-V-DQ--SDH--V-----L--IS--Q-EQE--S--A---E-VD-Y------EKVG-T-------Q  632
LST.KE.x.lho7              -NW----LE-YGQ-K-V---KMQ-A----E-E--VRV-QNKK-II-F-WR-G..NNI-RA-R-Q-QKA----PLQK-V--I---GK-------FV-.-IQV-VTR-V--H--SDH--V----DL--IS--Q-EQE--IW-A---I-VD-Y------EKVG-T-----I-Q  632
SYK.KE.x.SYK173            LT---A--A--E---Q----EQA-S----K-P-E-H-T-L-SQ--G-M-K--QKGPP-I---T-KTFA--S--YQS-AQLLN--GIQ-LWY---V-.T-H--VKR-E--K---D------V--VK-IS-----RWY-N-VP---PE-V----------DS-T-N----AS  649
SYK.KE.x.KE51              VE------A-M--------VEQT-S--Q-E-P-E-H-S-L--Q--G-I-K-G-QE-P-I---TGKAYAT-H--YQA-AQLMN--GIQALW----I-.E-H--VKR-E--K---DH-----L--VKC-H--I--RWY-N-VS--V.E------------NS-E-Q---I-D  654
COL.CM.x.CGU1              VQ-SA---E--QD-KQ-I--E-Q----I-HE-IWVD-FRL-K---G-AVC-K..HGT-RR--HNTGKNQ-F-SMQE-ASVI---GR-A--T---V-.-M-V-AKR-D--Q--SD---SA----I---SSSYV---VWN-VL--LKE-P-YWT--GCS-KLGA----WINS  595
```

```
                                                                                                                     p66 RT, p15 Rnase H end p31 Integrase start
H1B.FR.83.HXB2              RGRQKVVTLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAG.IR..KVLFLDGID.KAQDEHEK.YHSNWRAMASDFNLPPVVAKEIVASCDKCQLK.GEAMHGQVDCSPGIWQLD  779
H1A1.UG.85.U455             -------S--E--------H--H-------S-------------------R----I------K--E------S------------------S-.--..---------.---ED---.--C---------------------N-----.-----------------  778
H1B.US.90.WEAU160           ----R--S-------------H-----------------------------------S------------------------------S-.--..---------.---E----.--------------------------------.-----------------  779
H1C.ET.86.ETH2220           -----I-S--E-----------Q-------S---------------L----K----I--------S--R---S------------------S-.--..---------.---E----.---------NE--I----P-----C-------.---I----N---------  786
H1D.CD.84.84ZR085           -------PF-------------N-------------------------K------S---------------------------------S-.--..---------.---E----.--------------------------------.-----------------  779
H1F1.BE.93.VI850            K-K----S--E-----A-------------S-------------------K----I---------Q--R---S--------------------.V-..-I-------.---E----.--N-------------I----------------.-----------------  780
H1G.SE.93.SE6165            K-K--II---E-----A-----Q------RS-------------------R--A-------------------S------------------S-.--..---------.---E---R.--N-------------I----------------.-----------------  776
H1H.CF.90.056               --K----S--E-----------------------------------------K-----------E--------S------------------S-.V-..---------.---E---R.--N----V--------I----------------.-----------------  778
H1J.SE.93.SE7887            K-----------------H------R--------------------------K-----------E--------S------------------S-.--..---------.---ED---.--------------------------------.-----------------  777
H1K.CM.96.MP535             -------SI-E-----A-----C-------S-------------------K---D--------------RI--S--------------------.--..---------.---E----.--N-------------I----------------.---I-------------  779
H101_AE.TH.90.CM240         -------S--E--------H--H-------S-------------------R----V------E--------S------------------S-.--..---------.---E---R.------T---------I------TN-------.-----------------  784
H102_AG.NG.x.IBNG           -------S--E--------H--H-------S-------------------R-----------K----D---S------------------N-.--..---------.---E---R.-----K----------I------------M-.----------G-------  779
H103_AB.RU.97.KAL153_2      -------S--------------H---------------------------K------S-------------------------------------..E---------.---E----.--G-----------------------------.-----------------  779
H104_cpx.CY.94.CY032        -------S-SE-------------------S-----------I-------R---D-----------R-D---S------------------N-.--..---------.---E----.--N------------S---------N-----.-----------------  778
H1N.CM.95.YBF30             -------SIA------A-----LM---E--RD----------M---HS---K------S----E-----R---S------------------S-.--..-I-----E.---ED-DR.-----K----------I----------------.---------N----V----  786
H1O.BE.87.ANT70             Q-K--IIK-DE-----A--M--L------KET-----------V-SS--T----PI-Q----E-T---Q---T-----------KI-----KD.--..R----E---.Q--ED---.-----K-L--E-G---------I---P--HI-.---I--------EV--I-  775
H1O.CM.91.MVP5180           Q-K-NIIK-EE-----A--M-VLI-----KEQ---------V----SS--T--D-PI-Q----E-T---R---T-----------KI-----KD.--..R----E---.Q--ED---.-------L----G---I-----I---P--HI-.---T-----Y--E---M-  775
CPZ.CD.90.ANT               ---SR-KH-QK----QA-----LM--E--TGP-----------VL-GT------P--EE--QK---R-QI--S------------------Q-.--..Q----E---.---ED-D-.------SL-DEY----I-----I-Q----HV-.---R-----------V-  779
CPZ.CM.01.SIVcpzCAM13       K-K--IIS-EN----QA--G-L-------EQQ---------V-----S---H-----------E-----D---S------------------S-.--..---------.---E---R.-----K-L-GE-----I--------N-----V-.---L---------T--M-  778
CPZ.CM.05.SIVcpzEK505       ------TSISE----QA----VLM----A-Q----------V----HS---K-----------E-----RI--S-------------I-----T-.--..---------.---E---R.-----K-----------I----------------.---I---IN----V----  785
CPZ.CM.05.SIVcpzMB66        K----IIS--E-----A-----Q------EV-----------------G---T----I------E---------S-------------I-----S-.--..---------.---E----.--N-------------I-------N---------.---I-------------  782
CPZ.CM.98.CAM3              -----I-N-EN-----A--T-V------AEHI---------V----HS----------L---E------S--S-------V-----------S-#-..----------f-HKE---Ef--N----------I-HI----L--R-----------.-----------------  784
CPZ.GA.88.GAB2              N-K---IS-EE-----A--E-LRI------Q---V------V----NS----------I-------E----------T-----------------S-.--..---------.---ED--R.---------DE-Q-----------T-------.---I-------------  782
CPZ.CM.98.CAM5              -----IIN-EN-----A--T-V------AEHI---------V----HS----------GS---E------S--S------------------S-.--..---------.---E----.--N--K---------------------------.-----------------  780
CPZ.CM.05.SIVcpzMT145       -----IHS-E------A----VLM------Q---V------V----L-SH--H---DI------E------I--S-----R-----------S-.--..--------E.---E---R.-----K---Q------I-------Q-------.-----------------I-  778
CPZ.TZ.00.TAN1              ---YRSKD-EN----QA--W-VD---K---AQ---------VM-VL-GL----D-PI-E---QK-TQ-TAI---------------E------KN.--..-I-----N.E--ED-D-.-----K-L-DEY-----------I-Q-P--HI-.---I-----Y--E---I-  775
CPZ.CM.05.SIVcpzLB7         --K--IIP--E-----A-----Q------P---------------L-------D---------N---------T------------------S-.--..---------.R--E----.--N-------------I----------------.-----------------  782
CPZ.US.85.CPZUS             -----II--EN-----A--T-V----K--ENT--V--G---V----HS---------------E------S-IS-------------I-----S-.--..---------.---ED-D-.-----T-----------------------P-.---I-------------  781
CPZ.GA.88.GAB1              K-K--IIS-EN----QA--K-LL------DQQ---------V-----S---H-----------E------I--S-----------------------.--..---------.R--E---R.-----K----------I-------H-----V-.-----------------V-  777
GSN.CM.99.CN166             T-KHRAIE-EE----QA--H-VL---KEGPPKM-L------VM--L-S--EV-T-A--E---QE-L---AI--S-----------TE------K-.--..Q---EN-E.P-VED---.-----KYLRDQYKI-ALL-----NK-S---VH.--PK----NAEL-M--M-  790
GSN.CM.99.CN71              T-KYRAIE-EG----QA----VL---KEGPPRM-L------VM--L-S--EV-T-P--E---QE-LG--A---S-----------TE------R-.--..Q--MEN-E.P-VED---.-----KYLRDQYKI-ALL-----NK-S---IH.--PK----NAEL----M-  783
MUS_1.CM.01.1085            T-KE--IS-ES-----A--E-VL---KEGPPRM-L------V---VAS--QE-T-P--E---QE-LS--AI--S------------E-----NS-.--..Q--MEQ-E.P-KED---.-----KYLRDKYQI-ALL-----NH-P---TH.--PKT---NAEI-V--M-  780
MUS_1.CM.01.CM1239          T-KE---S-EN-----A--E-VL---KEGPPSM--------V---VAS--QE-T-P--EE--Q--LT--A---S------------D------H-.--..Q--MEQ-E.P-KED---.-----KYLRDKY-I-ALL-----NL-P---TH.--PKT---NAEL-V--M-  780
MUS_2.CM.01.CM1246          --QERAIN-EN----QA----L---M-GPP-M-L-----------T-A-EV---P--E---Q-MLS-NAIFISR-----------E--H---R-.--..Q--MEN--.P-VED---.-----KYLRDKY-I-TIL-----NK-SA--TH.--PK---NADL-V--M-  785
MUS_2.CM.01.CM2500          --QE-AID-EN-----A-----LM--T-GPPKM------------T-A-ET---P--EK--Q--LT--AI--S-----------QE--H---K-.--..Q--MEN-E.P-IED-D-.-----KYLRDKY-I-TLL-----NR-SA--TH.--PK----NADI-V--M-  784
                                                                                                                         p15 RNase H end p31 Integrase start
MAC.US.x.239                --KD--KV-EQ----QA--E-FLM--T---PKA--IV----VM---TGC-TE---R-------EM---SEI-V-----------QEI-H---Q-.--..Q----EK-E.P--E--D-.----VKELVFK-G--RI--RQ--DT----HQ-.---I---ANSDL-T--M-  790
H2A.GW.x.ALI                ---D--KV-EQ----QA--E-FA--VT---PKA--IV----VM--VAG--TE--NRI------EM----AI-V-----------QE--H---Q-.--..Q----EK-E.P--E----.----VKELSHK-G--NL--RQ--NT-AQ--Q-.---I----NAEL-T--M-  803
H2A.DE.x.BEN                --KD--KV-EQ----QA--EVFRM--A---PK---IV----VM--VAG--TE--NRI------EM----A--V-----------QE--H---Q-.--..Q----EK-E.P--E----.---IIKELTHK-GI-LL--RQ--N--AQ--Q-.---I----NAEI-V--M-  803
H2A.SN.x.ST                 ---D--RL-EQ----QA--E-FAM-VT---PKA--IV----VM--VAG--TE---KI------EM----AI-V-----------QE--H---Q-.--..Q----EK-E.P--E----.----VKELSHK-G--KL--RQ--NT-TQ--Q-.---I----NAEL-T--M-  803
H2B.GH.86.D205              --KD--KV-EQ----QA--E-FA---T--EPQ---IV----VM---A---TET--PI-AK---EM----A--VG-------L---QE--H---Q-.--..Q----EK-E.P--E----.--G-VKELVHK-GI-QL---Q--N------Q-.---I----NADL-T--M-  802
H2B.CI.x.EHO                --KD--KP-EQ----QA--E-FA-------PQ---IV----VM--VA---TET--PI-RE---EM-----I-VG-------L---QE--H---Q-.--..QI---EK-E.P--E----.--N-VKELVHK-GI-QL--RQ--N------Q-.---I----NSEL-T--M-  803
H2G.CI.92.ABT96             --KD--XV-EQ----QX--E-FAM------PK---IX----VMX--AX--TET--P---K---EX----XL-VG----------QE--X---Q-.--..Q---XK-E.P--E---X.F---VKELTHK-GI-QL---Q--N-YHS--Q-.---I----NAEL-T--M-  786
H2U.FR.96.12034             --KDR-KP-EQ----QA--E-FAM--E---PK---------IM---AG--TET--P--SK---E--R--A--VG----------QE--H---Q-.--..Q----EK-E.P--E----.----VKELQHK-GI-QL---Q--NT-HV--Q-.---------NTEV-T--M-  799
MAC.US.x.251_1A11           --KD--KV-EQ----QA--E-FLM--T---PKA--IV----VM---TGC-TE---R------KEM---SEI-V-----------QE--H---Q-.--..Q----EK-E.P--E--D-.----VKELVFK-G--RI--RQ--DT----HQ-.---I----NSDL-T--M-  790
MAC.US.x.251_BK28           --KD--KV-EQ----QA--E-FLM--T---PKT--IV----VM---TGC-TE---R-------EM---SEI-V-----------QEI-H---Q-.--..Q----EK-E.P--E--D-.----VKELVFK-G--RI--RQ--DT----HQ-.---I----NSDL-T--M-  786
MAC.US.x.EMBL_3             --KDI-KV-.T----QA--E---HGIE---PKR--IVEL-VCY-NNNRF-TE---R-------EM--.VR--V-----LE-----QEIGP---Q-.F-..Q----EK-E.P--E--D-.----VKELVFK-G--RI--RQ--DT----HQ-.---I----NSDL-T--M-  784
SMM.US.x.H9                 ---G-TKL-EQ----QA--E-F----A---PKA--IV----VM--VAG--TE---R-------KM----AI-V-------X--E-QE--H---QE.--..Q----EK-E.P--E----.----VKKLVFK-X--RL---Q--DT----H--.--XI----NXXL-T--M-  786
SMM.US.x.PGM53              ---D-AKP-EQ----QA--E-F----A---PKA--IV----VM---AG--TE---R-------EM----AI-V-----------QE--H---Q-.--..Q----EK-E.P--E----.----VKELVFK-GI-RL---Q--DT-H--HQ-.---I----NAEL-T--M-  786
SMM.US.x.SIVsmH635F_L3      ---D-AKL-EQ----QA--E-F----A---PKA--IV----VM--VAG--TE---R-------EM----AI-V-----------QE--H---Q-.--..Q----EK-E.P--E----.----VKELVFK-GI-RL---Q--DT--R-HQ-.---I----NAEL-T--M-  786
SMM.SL.92.SL92B             --KE--LP-EQA---QA----LL---K--PSK--V------V-N--TG--SE-D-DI-A------VQ--A--IG----------NE--R---Q-.--..Q----ES-E.P--ED-D-.----VKEL-QKY-I-QL---Q--NA-N---Q-.---I---TNAEV-T--M-  783
STM.US.x.STM                --KN--KA-EQ----QA--E-FAM--A---PKA---V----VM---TG--TE---K-------EM----AI-V-----------QE--H---Q-.--..Q----EK-E.P--E----.----VKELVFK-GI-RL---Q--DT----HQ-.---I----NAEL-T--M-  786
MNE.US.x.MNE027             --KD--RV-EQ----QA--E-FLM--T---PKA--IV----VM---TGC-TE---R-------EM---TEI-V-----------QEI-H---Q-.--..Q----EK-E.P--E--D-.----VKELVFK-G--RL---Q--DT----HQ-.---I----NSDL-T--M-  786
GRV.ET.x.GRI_677            --KSRIRE-EN----QA--T-VKM--E---EN---------VMN-LT-C-QE-N-P--E---QA-M--RQ---Q-----------TEI-----K-.--..QI---R-E.E---D-A-.--N---S-VQE-G--NI-------A-P---IR.--PK-----A-IET--M-  801
VER.KE.x.TYO1_patent        Y-K-R-E--EN----QA--T--KM--E---PN----------M--LT---T--D-P--E---ALM-Q-QQI--Q------------EI-----K-.--..R----EK-E.E--E---R.--N--KNL-DTYG--QI-------M-P---I-.--PV-----A---T--M-  798
VER.KE.x.9063               Y-K-R-EK-EQ----QA--V--KM--E---EK----------M--LT---T--D-P--E---ALMV--RAI--Q------------EI-----Q-.--..R--IGR-E.E--E--DR.------KNL-DT-G--QI-------M-PNAKI-.--PI-----A---V--M-  795
VER.KE.x.AGM155             Q-K-R-QQ-EN----QA--T--KM--E---PK----------M--LT---T--D-P--E---A-MVQ--AI--Q------------EI-----K-.V-..RI--IGR-E.E--E--DR.-------NL-DT-G--QI-------M-P---V-.--PI-----A---V--M-  798
VER.DE.x.AGM3               Y-K-R-EK-EN----QA--M--KM--E---PN----------M--LT---T--D-P-IE---ALMVQ-HQI--Q----D-------EI-----Q-.M-..-I---EK-E.E--E---R.--N---NL-DTYG--QI-------M-P---I-.--PV-----A---V--M-  795
TAN.UG.x.TAN1               T-ESQ-IE-E-----QA----VL---KH-KSK---------VM--LT---T--D-V--------EM-N-DAI--T-----------QEI-----Q-.--..Q-----RME.E--ES-D-.--T--QFIRDA-GI-AL-------A-P---IR.--PI-----A-V-V--M-  798
SAB.SN.x.SAB1C              --D----A-EN-----A--E--L---R---SK---I------M---AGE-TE-DNNI-Q-----E-----A--I--------V----EI-----Q-.--..Q-----R-E.E--E--D-.--A---S-QQE-G--AI-------A-P---I-.--SV-----A---V--M-  811
OLC.CI.97.97CI12            G-KH-I-HIQESS---A-IR--IE--R--KS--------L--M--VLEK-EV--EP-IQEL-KE-E---RI-IG-------L---QE--Q---K-.--TKQ-M--AD-E.P--ES-N-.--Q-AQSLQQE-QI-LI---R--DE-EE-KG-.-VPR--I-AVGVYR--I-  780
DRL.x.x.FAO                 --KE--IAIEE-----A-----L---K---PK----------M---SSS-EI-DNPI-------M-S--A---N------------EE-----R-.--..Q-----NME.---E--DL.--N---SL-QE-G--GI-------Q-P---IH.--PI-----A---T--M-  793
RCM.NG.x.NG411              --KE--KE-EN-----A--G-VL---K---KR-------RHVF--LAGS--T---P--Q-----M-G--E---S------------------Q-.--..Q------G.---E----.----------EE-QI-QI-------Q-P---V-.----V-----A---T--M-  797
RCM.GA.x.GAB1               --KE--KE-EE-----A--E-VL---K---PK---------VY--LE----T-D-G--TE--N-M-G--A---S------------E------K-.--..Q------E.---E----.--N----L-E--QI-QI-------Q-P---V-.---I-----A---T--M-  792
MND_1.GA.x.MNDGB1           --F-R-EEYLN----Q---H-VK---E---SY---------VV--LASR-TETDHPI-KE---LMKG---I--S-L------------I-----S-.--..-----QN-E.P--E----.----EAQLREK-H--AL---Q--Q--S--CHH.--PIK--T-A-L-V--I-  788
MND_2.x.x.5440              --KS---A-EE-----A--T--KM------PR-------------LS-A----DNPI-RE---LM-G--EI--S-------------I-----K-.--..Q----E---.---E--D-.--N----LSQ------I-------Q-P---T-.--PV----NAD--T--M-  792
MND_2.CM.98.CM16            -NKSR-IA-EN-----A--E--KM------PK----------M--LS-A----DNPI-RE---LM-H--A-------------------------R-.V-..Q----E---.---E--D-.--N----L-Q--CI-NI-------Q-P---T-.--PI-----T---T--M-  791
MND_2.GA.x.M14              --KS--IA-EN-----A--E--L---K---PRA------------LS-T--R-DNPI-RE--N-M-A--A------------------I-----K-.--..Q----E---.---E--D-.--N---SLSQE-SI--I-------Q-P---V-.--PI-----AA--T--M-  792
DEB.CM.99.CM40              W-K-S-KC-EN-----A--E--L---EEGPSKM------------LEH-SET-HKI-EKV-QA-QG--QI-------------D-------K-.--..RI---ERVE.E--ED---.--G--KQLRDEYQ--TLI--Q-ITQ-P---TR.--PT-----A-M----M-  784
DEB.CM.99.CM5               W-K-G-KS-EN-----A--E--L---EEGPSKM-----------LLEH-TET-H-I-EK--QA-QE--Q---N-----------D-------R-.--..RI---ERVE.E--ED---.--G--KQLRDE-Q--TLI--Q-I-Q-P---I-.--P------A-M-V--I-  784
DEN.CD.x.CD1                N-K------EE-----A--IGVLM-----PNK-------------AGT-TT--HQI-E---QAMQG--EI-IT-----------TEI-----K-.--..QI---EQ-P.Q--E--D-.------KEIRDKYG--TM-----INV-P---TH.--PIR----A-L-T--M-  806
TAL.CM.01.8023              S-KEH--G-SN-----A--E-VL---KH--PK---------VY--LVGN-TETDNKI-ED--Q--LG--A---T-L------------------Q-.--..RI--VEQ-P.E--E---R.--N--KDLKAR-K--TI---A-IEA-P---VQ.--PKT--TNAAV-T--M-  788
TAL.CM.00.266               T-KEH-IK-EK-----A--E-VL---K---PK---------VY--LAGS-TETDNKIIED--QL-LS--A-----L------------------Q-.--..-I--VEQ-P.E--E---R.--N---DLKAR-K--TI---A-IEA-P---VQ.--PKT--NNAAV-T--M-  787
MON.CM.99.L1                F-AR-A---EQ-----A--E-VL---R-GPPQM--I-----V---LASC-EI---PI-EA--QE-L--DQ-F-S-----------E--R--GQ-.--..Q--MENLE.P-RED---.-----KYLRDTYHI-TLL-----NH-H---NH.--PKS---NAEV-V--M-  786
MON.NG.x.NG1                T-QR--IP-EQ-----A--E-VL----NGPASM--------VM--LASK-EV---PI-EK--Q--LG-KQ---S------------EX----XQ-.--..Q--MEN-E.X-RED---.-----KYLRDTYGI-TLL-----NH-P--HTQ.--PK----XSDI----M-  789
SUN.GA.98.L14               G--SR-KK-EN----QA--E--KM--E--RSS-----------RLLSKR-TETD----KE-V-LIRQ-DQ---G----------QEI-Q---Q-.--KRQ-M-IEK-E.P-VE--G-.F-N-AASLQEM-DI-L----Q--NE-AQ--Q-.---IT----A-V----I-  798
WRC.CI.97.97CI14            KSKE---SVQN----Q---I-LK---E-G-QKM--I-----VM--VKDH-TM---P--QE---L-KG--EI-ID-----------TEI-R------.--S.-IM--EN-T.P-IES-T-.--QG-K-LQEE-H--QLI--R-IEE-ST--G-G-AP---M--YEV-T--M-  803
WRC.CI.98.98CI04            NSKE--IN-YN----QA--A-LK---E-G-SKI--I-----VM--VKDH-TI---A--QE--DL-KQ--EI-IE-----------TEI--------.--S.--M--EN-T.P-IES-T-.--QG-K-LQEE-HI-QLI--R--EE-ST--G-G-APL--M--YEV-T--M-  802
WRC.GM.05.Pbt_05GM_X02      KT-N-IQKI-N----Q---I-LRM--E--SDK---I-----V----KDH-T----P-MQE--DL-KG--EI-IT-----------TEI-A------.--S.R-M--EN-T.P-IES-T-.--QG-G-LEEE-H--TIL--R--EE-AT--G-G-HPI--M--YEI-T--M-  794
LST.CD.88.524               S-KE-IKE-Q-----QA--E-VL------KST---------VMKVLSQR-TET-HPI-KN--KE-Q--DA---G----------QEI-Q--IK-.--TKQ-M--EK-E.P-VE--G-.F-N-AKDLEEKYGI--M---Q--NE-A---Q-.---IT----V-V------  799
LST.CD.88.485               S-KE--KE-------QA--E-VL---K--KSR---------VMKVLSQR-TET-HPV-K----ECK--DQ---G----------QE--H---K-.--QIQ----EK-E.P--E--S-.F-N-AKDLEEK-H---M---Q--N--N---T-.---IT----T------I-  798
LST.CD.88.447               S-KE--KE-------QA--E-VL---R--KSR---------VIKVLSQR-TET-HPV-K----ECK--DQ---G----------QE--H---K-.--QKQ----GK-E.P--E-YS-.F-N-AKDLEEK-HI--M---Q--N--N---T-.---IT----T------I-  798
LST.KE.x.lho7               S-KE--KE-------QA--E-VLM-----NSK---------VMK-LSQR-TET-HPI-KD----CKQ-DQ---G----------QE--H---K-.--QKQ-M--EK-E.P-VE--S-.F-N-AKDLEEK-----M---Q--ND-AN--K-.---IT----V-V------  798
SYK.KE.x.SYK173             D-T-R-QY-EQ----QA--EGLLM-----KDK----V----SY--LMTC-TNT-HPI-E---QEA----AI-VT-----------A------K-.--..-I---ER-P.Q--ED--R.----MEYLRQE-H--RQ---A-IQQ-P---NR.--PK-----VDIYN--M-  813
SYK.KE.x.KE51               --K-TLKR-EN-----A--E-VLM------P--------------LMNC-TS--HPV-E--MQKAME---I-IT----------QE--Q---K-.--..-I----N-P.---E----.--T-VEYIRQE-H--RQ---A--EM-P---IR.--PK-----TDL-T--M-  818
COL.CM.x.CGU1               K-EEE--N-HEGS--QA--TGVL---KHGPKRM-L----L------TG--YDQLPSSRGE--QAGMA--AIHV--C----------LI-QK-..-.V-..Q-MWI-K-E.A-EED-Q-.F---VQYLKEQ-G--T------WER-SE--N-.-Q-V---L-Y-Y-L----  757
```


PLV Proteins

```
H1B.FR.83.HXB2              CTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTI.HTDNGSNFTGATVRAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGER.IVDIIATDIQTKELQKQI.TKIQNFRVYYRDSR 946
H1A1.UG.85.U455             ------------------------------------I----------V-.---------S-A-K-V----N-Q------------------------------E--------------------------------.-I----------------.S------------- 945
H1B.US.90.WEAU160           -------I----------------------------I------------.---------ST--K-----------------------I------------------------------------------------.-I---------QQ-----.-------------- 946
H1C.ET.86.ETH2220           -------I--------------------------------------RV-.---------SNA-K--------Q------------------------------E--------------------R-----------.-I----S--------N--.L------------- 953
H1D.CD.84.84ZR085           --------------------------------A--------------VV.---------S---K------------------------------------------------------------------------.-I----S----R------.-------------- 946
H1F1.BE.93.VI850            -----------------E------------------I----------I-.---------S-A-K-S------Q-----------------I---------------------------------------------.-I---S-----R------.-------------- 947
H1G.SE.93.SE6165            -------I-I--------------------------I---------TV-.---------S-A-K------N-T---------------------------------------------------------------.-I----S-----------.-------------- 943
H1H.CF.90.056               ------Q-----------------------K----------S-----V-.---------S-A-K------D-Q---------------------------------------------------------------.-I----------------.SN--K--------- 945
H1J.SE.93.SE7887            --------------------------------A-F-I----------V-.---------SGA-K------D--------------------------------E--------------------------------.-I---------R------.-------------- 944
H1K.CM.96.MP535             -------I----------------------------I----------V-.-----T---STV-K-------V----------------------------------------------------------------.------------------.LN--K--------- 946
H101_AE.TH.90.CM240         -----------------------------------------------V-.---------S-A-K------NVQ------------------------------E--------------------------------.-I----------------.-------------- 951
H102_AG.NG.x.IBNG           -------I----------------------------I----------V-.---------S-A-K------NVT---------------------------------------------------------------.-I----S-----------.I------------- 946
H103_AB.RU.97.KAL153_2      -------I----------------------------V----------I-.---------STA-K-----------------------------Q--QT--------------------------------------.-I----------------.I------------- 946
H104_cpx.CY.94.CY032        ---------M--------------------------I----------M-.-A---P---S-A-K------D-N---------------------------------------------------------------.-I----S-----------.-------------- 945
H1N.CM.95.YBF30             -------I-----------L----------------I----------V-.---------S---K------N---------------A--------------I-----------------------------T----.-I---------TN--T--.L-V----------- 953
H1O.BE.87.ANT70             ---M---I-I--------F----------------------A-----V-.-----P---ST-MK------N-QH---------------A------S-Q----------R--------V------------T----.-I--L-SQ---T------.L-XHK--------- 942
H1O.CM.91.MVP5180           -------I-I-------DF----------------------A-----V-.-----P---S-AMK-----T--QH---------------A------S-Q-------------------V------------T----.LI--L-SQ---T------.L--N---------- 942
CPZ.CD.90.ANT               ---------I-----S--F-----MAD---KS---------S-------.-----A---S-A-K------N-Q--------------------Q--Q----I-----Q-----V---H-------------TP-Q-.-L--L------TQ--N--.L---Q---H----- 946
CPZ.CM.01.SIVcpzCAM13       -------I-I----I---------------------I----S-----VL.-----T---SNA-K------N--K-------------------------E------------------S------------T----.-I----------------.L--K---------- 945
CPZ.CM.05.SIVcpzEK505       -------I-----------L----------------I----------V-.---------S---K------N-Q-------------A--------------I-----------------------------T----.-I---------TK--T--.L-V----------- 952
CPZ.CM.05.SIVcpzMB66        -------I-----------M----------------I----------V-.---------SS--K--------Q----------------------------I----------------------------------.-I--L------TK--Q--.SN--K--------- 949
CPZ.CM.98.CAM3              -------------------L----------------I----------V-.-----P---SN--K--------Q----------------------------I-E-----R---------------------T----.-I---------TN-----.L-V----------- 951
CPZ.GA.88.GAB2              -------I-----------L----------------I----------K-.-----T---S-A-K-------------------------------------------------------------------T----.---ML---L--Q------.L---K--------- 949
CPZ.CM.98.CAM5              -------I-----------L----------------I----------V-.-----P---SN--K--------Q----------------------------I-----Q-----------------------T----.-I---------TN-----.L-V----------- 947
CPZ.CM.05.SIVcpzMT145       -------I-I--------------------------V------------.-----P---SNA-K--------R--------------------------E------------------------------------.-I----------------.L---Q--------K 945
CPZ.TZ.00.TAN1              ---------I--------F-------E---R-----I----------K-.-----P---STA-K------Q-QH-------------------Q--Q--E-------Q-R---I---Y-----------E-T----.LL--LT-N---Q------.L-V---------A- 942
CPZ.CM.05.SIVcpzLB7         -------------------M----------------I----S-----I-.---------STA-K------Q-Q---------------------------------------------------------------.-I---------T------.S-V-K-------R- 949
CPZ.US.85.CPZUS             -------I-I---------L----------------I----------V-.---------SS-----------Q----------------------------I-----------------------------T----.-I----SEL--DL-----.L-V----------- 948
CPZ.GA.88.GAB1              ---------I---------------------------------------.-----P---S-A-K------D-------------------L----------------------------------------T----.-I---------S------.L-V-K--------- 944
GSN.CM.99.CN166             -------------------VW-RI--Q---RQ--LK--E--AT---THL.-----P--VSKELE------D-QHTT-V-----------N---Q--ET-TKI-EEVTY-E---AQ-CY-----------DMCPT--.--NM-H-ELE-QH-NT-N.S-F-K------QGA 957
GSN.CM.99.CN71              -------------------VW-RI--Q---RQ--LK--E--AT---THL.-----P--ISKELE------N-QHST-V-----------N---Q--ET-QKI--EVTY-E---AQ-CY-----------DMCPT--.L-NM-H-ELE-QH-NT-S.S-F-K------QGA 950
MUS_1.CM.01.1085            -------I-----------MW-KIL-R---RQ--LG--E--SL--ATQ-.-----A--IS-EFE------S-EHTT-V-----------NK--Q--EA-HKI-EEVTY-E---AQ----L---K---V-DL-NA--.-INMLH-ELELQH--Q-K.Y-F--------TGS 947
MUS_1.CM.01.CM1239          -------I-----------TW-KIL-R---RQ--LG--E--AL---TQ-.-----A--ISEEFG------S-DHTT-V-----------NK--Q--ET-QKI-EEVTY-E---AQ----M---K-----DLA-A--.-INMLH-ELELQH--S-K.S-F--------TGS 947
MUS_2.CM.01.CM1246          -----------------SF-W-RI-SQ---RL--LE-SN--AT--ISQ-.-P---T---SKEFQ-VA---N-QHTT-V-----------NA--Q--ET-HK--EEVTY-E---AQ----L-N-------DMTPT--.L-NMLY-ELEIQQ--NHN.--FSK------RGA 952
MUS_2.CM.01.CM2500          -------I----------F-W-RI-SQ-N-RL--LE--N--AT--ITQ-.-----T---SKEFQ-VA-----EHST-V-----------NA--Q--ET-NK--EEVQY-E---AQ----L-N-------DMTPT--.L-NMLY-E-E-QQI-NTK...FSK------RGA 949

MAC.US.x.239                -------I-I--------F-------Q---RQ--L----------ITHL.-----A--ASQE-KMVA-----EHT--V-----------A--HH--NQ-DRI-E--NSVE-I-L---HCM----R----DMTPA--.LINM-T-EQEIQFQ-SKN.S-FK-------EG- 957
H2A.GW.x.ALI                ---------II-------F-------Q-S-RQ--L------S---ITHL.---S-V---SQE-KMVA--V--E-S--V-----------A--HH--NQ-SRI-E--NTVE-I-L---HCM----R----DMTPA--.LINM-S-EQEIQF--TKN.L-FK--P----EG- 970
H2A.DE.x.BEN                Y------I-I--------F-------Q-S-RQ--L------S---ITHL.-----P---SQE-KMVA--V--E-S--V-----------A--HH--NQ-SRI-E--NTIE-I-L---HCM----R----DMTPA--.LINM-T-EQEIQF--RKN.SNFK--Q----EG- 970
H2A.SN.x.ST                 -------I-I--------F-------Q-S-RQ--L------S---ITHL.-----A---SQE-KMVA--I--E-S--V-----------A--HH--NQ-SRI-E--NTVE-I-L---HCM----R----DMTPA--.LINMVTAEQEIQF--AKN.S-L------F-EG- 970
H2B.GH.86.D205              -------I-I--------F-------Q---RQ--L------S---ITHL.-----A---SPS-KMVA--V--E-T--V-----------A--HH--NQ-DRL----VSIE-V-L--THCM----R----DMTPA--.L-NM-T-EQEIQFF-AKN.L-F---Q----EG- 969
H2B.CI.x.EHO                ---------I--------F-------Q---RQ--L------S---ITHL.-----A---SQD-KM-A--I--E-T--V----E------A--HH--NQ-DRI----VSIE-V-L--THCM----R----DMTPA--.--NM-T-EQEIQF--TKN.L-F--------EG- 970
H2G.CI.92.ABT96             ---------I--------FV------Q---RQ--L----------ITHL.-----A---SQE-KMVA--V--E-A--V--------------HH---Q-DKI----NSIE-I-L---HCM----R----DMTPA--.-INM-T-EQEIQFQ-TKN.S-FK-------EG- 953
H2U.FR.96.12034             --------VI--------F-------Q---RQ--L----------ITHL.-----A---SQE-KMVA--V--EHT--V-----------A--HH--NQ-ERI-E--NTME-I-L--AHCM----R----DMTPA--.-INM-S-EQEIQF--SKN.S-FK-------EG- 966
MAC.US.x.251_1A11           -------I-I--------F-------Q---RQ--L----------ITHL.-----A--ASQE-KMVA-----EHT--V-----------A--HH--NQ-DRI-E--NSVE-I-L---HSM----R--T-DMTPA--.LINM-T-EQEIQFQ-SKN.S-FK-------EG- 957
MAC.US.x.251_BK28           -------IVI--------F-------Q---RQ--L----------ITHL.-----A--ASQE-KMVA-----EHT--V-----------A--HH--NQ-DRI-E--NSVE-I-L---HCM----R----DMTPA--.LINM-T-EQEIQFQ-SKN.S-FK-------EG- 953
MAC.US.x.EMBL_3             -------IVI--------F-------Q---RQHY.----------YLH-YTHS--A--ASQE-KMVT-----EAHLWV-----------A--HH--NQ-DRI-E--NSVE-I-L---HCM----R----DMTPA--.LINM-T-EQEIQFQ-SKN.S-FK-------EG- 951
SMM.US.x.H9                 -------I-I-----G--F-------X---RQ--L------S---ITHL.-----A---SQE-KMVA---X-E-T-XV-------X---A--HH--TQ-DRI-E--NSIE-I-L--IHCI----R----DMTPA--.L-NM-T-EQEIQFQ-SKN.S-FK-------EG- 953
SMM.US.x.PGM53              -------I-I--------F-------Q---RQ--L------S---ITHL.-----A---SQE-KMVA-----E-T--V-----------A--HH--TQ-DRI-E--NSIE-I-L---HCM----R----DMTPA--.L-NM-T-EQEIQFQ-SKN.S-FK-------EG- 953
SMM.US.x.SIVsmH635F_L3      -------I-I--------F-------Q---RQ--L----------ITHL.-----A---SQE-KMVA-----E-T--V-----------A--HH--TQ-DRI-E--NSVE-I-L---HCM----R----DMTPA--.L-NM-T-EQEIQFQ-SKN.S-FK-------EG- 953
SMM.SL.92.SL92B             ---------I--------F-------R---RQ--L----I-S---I-HL.-----A---SQE-KMVA--L-VE-S--V-----------A-DLH---N-DKI-E---SVE-L-L--AHCM----R----DMTPA--.--NM-T-ELE-QY-NS-N.S-F--------EG- 950
STM.US.x.STM                -------I-I--------F-------Q---RQ--L------S----THL.-----A---SQE-KMVA-----E-T--V-----------A--HH--TQ-DRI----NTVE-V-L---HCM---KR--L-DMTPA--.L-NM-T-EQEIQFQ-SKN.S-FK-------EG- 953
MNE.US.x.MNE027             -------I-I--------F-------Q---RQ--L------S---ITHL.-----A---SQE-KMVA-----EHT--V-----------A--HH--NQ-DRI-E--NSME-I-L---HCM----R----DMTPA--.LLNM-T-EQEIQFQ-SKN.S-FK-------EG- 953
GRV.ET.x.GRI_677            ---------I--------F-------R---K---H-----LA-----HL.-----P---SQN-A-V---GN-EHTT----------S-----RQ--E--S-I--DC-R-E------TH------------I-SA--.L-NMLT-QLELNT--N--.Q--L--K----EG- 968
VER.KE.x.TYO1_patent        --------VI--------F-------R---K---K----ILS---ITQL.-----P---SQE-A-I---GK-EHTT----------SI-----Q--E---KI--DCQYTE---L--CH------------QTSA--.LIN--T-QLEIQH--TK-.Q--L-------EG- 965
VER.KE.x.9063               ---V---IVI--------F-------R---R---K----ILS---IVQL.-----P---SQE-A-I---GK-EHTT-V--------S------Q--E---KI--DCQYTE---L--CH------------LTPA--.LINM-T-QLELQT--TK-.Q--L-------EG- 962
VER.KE.x.AGM155             ---I---IVI--------F-------R---K---K----II----ITHL.-----P---SQE-A-M---GKVEHTT-V--------SI-----Q--E---KI--DCQYTE---L--CH------------LT-A--.LINM-T-QLEINT--TK-.Q--L-------EG- 965
VER.DE.x.AGM3               ---------I--------F-------R---K---K----ILS---ITQL.-----P---SQE-A-M---GK-EHTT-V--------SI-----Q--E---KI--DCQYTE---L--CH------------LTPA--.LINM-T-QLELQH--TK-.Q--L-------EG- 962
TAN.UG.x.TAN1               -------I-I--------FL-----AR---K---K----IIS---ITKL.-----P--VSQE-QTI---GQVEHTT----------S-----RQ--ET-EKI-EDCAF-E---L--CH------------MTPA--.LINM-T-QLEIQHI-T-Q.Q--S--K----EG- 965
SAB.SN.x.SAB1C              -------I-I--------F-------Q---KA--H------S---ITQL.-----T---SQQ-A-I---GK-EHT--V---------------Q--E----I--D--R-E---I---H---------------A--.LIN--H-ELE--T--QK-.S----------EG- 978
OLC.CI.97.97CI12            ---------CL-A-TG--FLYG-ILR--D-EN-VK-V-R--SQ----I-.-----T--VNQKMQ-V---L--EH-T---W------II-AK-RV--E-L-RI-E--T--E--LA--I------N---Q--GTVA-E.L-NQ-Y--T.QI-YTQNK.F-KFSGYRV-YKNK 946
DRL.x.x.FAO                 -------I-I--------H-----------EK--------------SHL.-----P---SEK-ATV----K-EHTT-V--------------NQ-------I-----K-E-------L-----------E---A--.------S-LL-SK--QN-.L----------EG- 960
RCM.NG.x.NG411              -------I-I--------------------K---H------A-----HL.-----T---S-A-Q-D----Q-EHT--V--------------HQ--I----I-----K-E-------L------------------.-I----S-L--TK--N--.S--------F-EG- 964
RCM.GA.x.GAB1               -------I-I--------------------K---H------A----RKL.-----A---S-A-Q-V----Q-EHA--V---------------Q--I--E---E---K-E-------LV-----------------.-I------LA-NK--N--.S----------EG- 959
MND_1.GA.x.MNDGB1           -----NQI-I--------FMK----T----KK--E------AQ--ISKL.-----P---SQE-ETM---L--EHT--------------NK--Y--EL-EKI-EDCKE-----A--T------QR--L--MT----.--NM-N-ELEYQYQ-N--.S-NL--K--F-EG- 955
MND_2.x.x.5440              -------I-I---------L----------K---H------------HL.-----P--VSEK-ATV----Q-EHTT-V-----------AK-HH-----E-------R-E-------L-----------E--P---.-----T---L-TK-RQN-.S-----Q----EG- 959
MND_2.CM.98.CM16            ---M-----IA------R-L------T---K---H------------HL.-----P---SEK-ATV----Q-EHTT-----------I-AK-HH--Q----------K-E-------L-----------E--P---IVDI-ATDLLT--LQHN.-.Q----------EG- 958
MND_2.GA.x.M14              ---------I---------L----------K---H------------HL.-----P--VSEK-ATV----Q-EHTT-------------AK-HH--V--E-------R-E-------L-----------E--P---.-----T---L-TK--QN-.S----------EG- 959
DEB.CM.99.CM40              -----N-I-I-----------TKIL-----K---L---Q-GA----TQ-.-----P--ISQP-K-----L--EHTT-V------------K-RI--DT-K-I--D--R-E---A--TY-I----R----DM-PI--.L-NM-T-E-E--T--QK-.S-LLG-K----EGA 951
DEB.CM.99.CM5               -----N---I--------FM-TK-LS----K---L---Q-GA---I-Q-.-----P--ISQP-K-----L--DHTT-V------------K-RV--ET-K-I-ED--R-E---A--T--I----R----DMA-I--.L-NM-T--LE-QT--QK-.S-LLGYK----EGA 951
DEN.CD.x.CD1                -----N---I---NA----L-TKIL-----K---L---QI-A---ITKL.-----P--ISQAMQ------N--HTT-V--------I--NK--Q--EV-AL--ED--R-E--LA--TL-L----R--L-DMT-A--Y-NMLNTHLEIQH..T--..P-FS--K----EGA 971
TAL.CM.01.8023              ------Q--C-----------TKIL-R---R---L---QV-S---IEHL.-----P--VS-EMQ-TA--LK-EHTT-V--------S--NK--Q---T-Q-I--EVQY-S---AQ-T--L----R--L-DMCPA-A.LINM-Y-EL--TT--N--.HNFSD-K----KGA 955
TAL.CM.00.266               ------QI-C-----------TKIL-R---R---L---QV-S---ISHL.-----P--VS-EMQ-MV--LK-EHSA-V--------A--NK--Q---T-T-I--EVQY-S---AQ-T--L-Y--R--L-DMCPA-A.-INM-Y-EL--TQ--N--.QNFSD-K----KGA 954
MON.CM.99.L1                ------SI-----------VW-KIL-R---KC-GIA--E--AM---TQ-.-----P---SQEFE--A---N-HHTT-V-----------N--RQ--ET-KKI--EV-Y-P--LAQ-L--L---------DMAPVD-.FINM-H-ELELQTSNN--.--FSK----F-TGA 953
MON.NG.x.NG1                ---Y--SI-----------VW-KIL-R---K--GIA--E--AT--IHHV.-----P---SQEFAX-A---N-EHST-V-----------N---QI-TTL-KI--EVQF-S--LAQTLY-L----X----DI-PI--.L-NM-H--LELQHSNN--.--FSK----F-TGA 956
SUN.GA.98.L14               ---M-E---I---------MV---L-N-Q-KT--TW----CAM----Q-.-----P--ISKD-E-V---L--QHTT-------------AK--V--Q--SRI-ED-QE-----L--LH-----QR--L--MT-A--.FINM-NA-LE-QY---IN.S--LK-K----EG- 965
WRC.CI.97.97CI14            -------I--N--------L--KI-S--DAYQ--L-T-Q--S---I-K-.-----T--KSQKMQ-I---L--E--------------I--K-RR--D--HSI-ED--K-E--LA--L------E---L--.-PA--.L-NM-TS-LE-QQT-Q-K.L-FK--Q----TGA 969
WRC.CI.98.98CI04            -------I--N--------L--KI----DAYQ--L---QI-----M-K-.-----T--KSQKMQTI---L--E--------------I--K-RR--E--HRI-ED--K-E--LA--L------E---L--.-PA--.L-NM-TAELE-NQT-Q-K.L-FK--Q----TGA 968
WRC.GM.05.Pbt_05GM_X02      -------I--N--------L--KI--T-DAYQ--L-TIQ--------K-.-----T--KSQKME-I---L--E--------------I--K-RV--Q--QRI-ED--K-E--LA--IY-----E---L--.-PA--.--NM-QS--E-NQT-Q-K.F-FK--Q----TGA 960
LST.CD.88.524               ------QI-MN---M---FMV-----D---KT--N-----CS-----Q-.-----P--VSKD-Q-VT--LS-EHTT-------------AK--V--Q--ER--ED-QQ-----L---H-----QR--L--L-PA--.YINM-NA-LE-QYI--LN.S--LK-K----QG- 966
LST.CD.88.485               ---M--Q--IN-I---C-FMV-----D---KT--N-----CS----QQV.-----P--VSKE-Q-VT--L--EHST-------------AK--V--Q--ER--ED-QQ-----L---H-----QR--L--L-PA--.FINM-NA-LE-QY---LN.S--LK-K----QG- 965
LST.CD.88.447               ---M--Q--IN-------FMV-----D---KT--H-----CS----QQV.-----P--VSKE-Q-VI--L--EHTT-------------AK--A--Q--ER--ED-QQ-----L--TH-----QR--L--L-PA--.FINM-NA-LE-QY---LN.S--LK-K----QG- 965
LST.KE.x.lho7               ------Q--IN-------FMV-----D---KT-SH-----CS-----Q-.-----P--VSKE-Q-VT--I--EHTT-------------AK--V--S--ER--ED-QQ-----L---H-----QR--L--LTPA--.FINM-NAELE-QY---LN.S--LK-K----QG- 965
SYK.KE.x.SYK173             ---E-----C---NT------TKILKR---D---L--MQI-S---I-Q-.-----P--VSDKFK-----C--EHTT----------I---K-RY--EA-S-I--DVT--Q---A--T--L----------I-P---.YINMLY-EL-LQQNTTS..P-FS-------QGK 979
SYK.KE.x.KE51               ---------C---NT----T-TKILKR------GL---QI-A---I-H-.-----P--ISDAFA-------VEHTT----------I--NK-RQ--ET-N-I-EEV-R-E---A--T--L----------TTPAD-.YINMLY-EL-LQNTHT...Q-FS--K----HGT 983
COL.CM.x.CGU1               ---E-----------CTLFCW-TILKR---E--GRA-I---SQ-E-RQV.-----P--VSQHFK--V--L--AHTT-H-----------QR--DV-RK-KKMK----T-ESK-A---YAL-------L--K-PW--.Q-ERAIIELD-QN-T-LQNQ-FK--KA-WKEH. 924
```

```
                                                                p31 Integrase end
                                                                      Pol end
                                                                                 ↘
H1B.FR.83.HXB2          NPLWKGPAKLLWKGEGAVVIQDNS.DIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED*  1003
H1A1.UG.85.U455         D-I----------------------.------------------------M-G------  1003
H1B.US.90.WEAU160       D-----------------------.-------------------------------  1004
H1C.ET.86.ETH2220       D-I----------------------.----------------------A----G------  1011
H1D.CD.84.84ZR085       D-I----------------------.-------------------------------  1004
H1F1.BE.93.VI850        --V----------------------.E--I----------------V-----G------  1005
H1G.SE.93.SE6165        D-V--------------------N.E-------------------------G------  1001
H1H.CF.90.056           D-I----------------------.E-------E------------------------  1003
H1J.SE.93.SE7887        D-I-------P-------------.E-------------------------G------  1002
H1K.CM.96.MP535         E-I----------------------.E-------------------------G------  1004
H101_AE.TH.90.CM240     D-I----------------------.-------------------------G------  1009
H102_AG.NG.x.IBNG       D-I----------------------.-------------------------G------  1004
H103_AB.RU.97.KAL153_2  D-I--------------------N.--------------------------------  1004
H104_cpx.CY.94.CY032    E-I----------------------.----------------------N---G------  1003
H1N.CM.95.YBF30         D-I-------------------G.----------------------G----G--NQ  1011
H1O.BE.87.ANT70         D-I-----Q-------------KG.--------------E-------T-SM--G-T-SE  1000
H1O.CM.91.MVP5180       D-I-----Q-------------KG.----------------------T-SM-N--T-SE  1000
CPZ.CD.90.ANT           D-V-----Q-----------K-QE.E------------KE-R-KIEDR-DL-G--N---  1004
CPZ.CM.01.SIVcpzCAM13   D-I-----T-------------QG.E-------------------------------  1003
CPZ.CM.05.SIVcpzEK505   E-T-----------------G.----------------------G----G---SQ  1010
CPZ.CM.05.SIVcpzMB66    D-I---------------L--QE.E-----------------------------NQ  1007
CPZ.CM.98.CAM3          D-I-----R----------KERE.EV--I-------------------SM-GG---SQ  1009
CPZ.GA.88.GAB2          D-I-----S--------------G.-------------KH----V-----L-GG---NQ  1007
CPZ.CM.98.CAM5          D-I-----R----------KE-E.EV--------------------GSM-G--N--Q  1005
CPZ.CM.05.SIVcpzMT145   D-I-------------------QD.E--------------------A-S-------S-  1003
CPZ.TZ.00.TAN1          D-I-----R----------KEGE.------------KE-------AGGMDD--N-T-  1000
CPZ.CM.05.SIVcpzLB7     D-I---------------L--QE.E------------------E-AN-L-D----N-  1007
CPZ.US.85.CPZUS         D-I----------------KE-E.EV----------K--------A-SM-G----S-  1006
CPZ.GA.88.GAB1          D-I-----T-------------QG.EL----------------------------N---  1002
GSN.CM.99.CN166         ----Q---V----------V-TQAGE-IT---------KP--EAKTE-VGG-THTNN--  1016
GSN.CM.99.CN71          --H-Q---V----------V-TQAGE-IT---------KP--EAK-E-VGS-AHTSN-R  1009
MUS_1.CM.01.1085        D-S-----V---R-------KTD-GEVIT---------KP--AKEDVGSKSDTGDLRKE  1006
MUS_1.CM.01.CM1239      D-S-----A----------VKTEQGQVIT---------KP--AKENVGSKSNTGDHRKE  1006
MUS_2.CM.01.CM1246      --Y-LR--R--------L--KTKEGE-VT---------K---TRQNV-SEPD-VHVRKE  1011
MUS_2.CM.01.CM2500      --S-Q---Q--------I--KTDQGE-IT---------K---TKENV-SDPNPVY-RKE  1008

MAC.US.x.239            DQ-----GE---------ILKVGT.-------------K---GGKEV-SSSHMEDTGEA  1015
H2A.GW.x.ALI            DQ-----GE-----D---IVKVGT.-----------------GRQEL-SGPHLEGARED  1028
H2A.DE.x.BEN            DQ-----GE---------IVKVGT.-----------------GRQEL-SSPHLEGARED  1028
H2A.SN.x.ST             DQ-----GE-----D---IVKVGA.---II--------K---GRQEM-SGSNLEGARED  1028
H2B.GH.86.D205          DQ-----GE---------I-KVGT.E------------H--GGKGL-CSADMEDTRQA  1027
H2B.CI.x.EHO            DQ-----GD---------I-KVGT.E---I---------N--GGKEL-CSADVEDTMQA  1028
H2G.CI.92.ABT96         DQ-----GE---------I-KVGT.E------------K---GGKEV-SSTNMEDTRQT  1011
H2U.FR.96.12034         DQ-----D----------L-KVGT.----I------------GGKEL-SSSNLEGA-KV  1024
MAC.US.x.251_1A11       DQ-----GE---------ILKVGT.-------------K---GGKEM-SSSHMEDTGEA  1015
MAC.US.x.251_BK28       DQ-----GE---------ILKVGT.-------------K---GGKEV-SSSHMEDTGEA  1011
MAC.US.x.EMBL_3         DQ-----GE---------ILKVGT.-------------K---GGKEV-SSSHMEDTGEA  1009
SMM.US.x.H9             DQ-----GE---------ILKVGT.E------------K---GGKEL-SGSHLEDTGEA  1011
SMM.US.x.PGM53          DQ-----GE---------ILKVGT.E------------K---GGKEL-SGPHLEDTGEA  1011
SMM.US.x.SIVsmH635F_L3  DQ--R--GE---------ILKVGT.E------------K---GGKEL-SGSHLEDTGEA  1011
SMM.SL.92.SL92B         DQ------E----------KVGT.E------------K---GRQEMGSSASMEDQ*KN  1007
STM.US.x.STM            DQ-----GE---------IVKVGT.-------------K---GGKEV-SGSHLEDTREA  1011
MNE.US.x.MNE027         DQ-----GE---------ILKVGT.-------------K---GGKEV-SSSHMEDTGEA  1011
GRV.ET.x.GRI_677        D-V-----R-I--------KEGE.-------------K---ERKTM-SEGSMEGVREA  1026
VER.KE.x.TYO1_patent    D-V-----Q-I--------LK.DGS-L-----------K--EPKQRVGNEGDVEGTRGS  1023
VER.KE.x.9063           D-V-----Q-I--------K.EGEEL-I---------K--EPRKRMG-EGNMEGLRGP  1020
VER.KE.x.AGM155         D-V-----R-I--------LK.EGEEL-----------K--EPRKTLG-ETHLEGAGGS  1023
VER.DE.x.AGM3           D-V----GQ-I--------K.GGVEL-EY--------K--EPRKRMG-ESNLEGAGGA  1020
TAN.UG.x.TAN1           E-V-----T----------KEGE.EL-----------KE--ERKTVG-KTNMEG*...  1019
SAB.SN.x.SAB1C          D-V-------I----------EQG.EL-TI--------K------L-SQAPLEGNGRT  1034
OLC.CI.97.97CI12        EGQ-C---E---------LLKEGD.KYFS--L-----VKP--VDSSANV.....*....  994
DRL.x.x.FAO             DQ------E-V--------KEGT.-L--I--------K----NVDSHTNME-*.....  1012
RCM.NG.x.NG411          DQQ------I-E--------QE.-L-------C---K---RKEVDRETNMEGRQ-ES  1022
RCM.GA.x.GAB1           DQ--R-----I----------EETG-L-----------KE--RKDVDSEANLAGRQ-EN  1018
MND_1.GA.x.MNDGB1       DQ-----GI---------LKYQE.E--I---C---K---E.SGKNSQ-NLESV*..  1009
MND_2.x.x.5440          DQQ-----E-I--------KEGV.-L----------K----TVDSNPHMGR*.....  1011
MND_2.CM.98.CM16        DQQ-----E-I--------KEGT.-L----------------TVDS-PN-EA**....  1010
MND_2.GA.x.M14          EQ------E-I--------KEGT.-L----------------TVDSNPHMED--ETA-  1017
DEB.CM.99.CM40          D-T-----T----------CKTEVG----------V-KE---D-DSKVNTQNSNE*..  1007
DEB.CM.99.CM5           D-T-----T----------CKTEVG----------V-K----GVDSKTDTQNPSE*..  1007
DEN.CD.x.CD1            --Q-Q---E----------KTQAG--FP---------KP-NAE.SRKGEQNKGM-SET  1029
TAL.CM.01.8023          ----Q---H-V--------LRTDEGEVIT---------KP--QA-GNKTDLEGSKEQ-A  1014
TAL.CM.00.266           ----Q---H-I--------LRTDEGEVIT---------KP--QALGNKIDLE-SKEQ-A  1013
MON.CM.99.L1            D-Q-----R----------KTEEG--LT---------KP--TTENVGGDTNQYNLRKQ  1012
MON.NG.x.NG1            D-S-----H--------L--KTDQGEVIT----------P--SKEDVGSKPSAH-IREV  1015
SUN.GA.98.L14           D-Q----------------KQ.GEN-L--------LVK---GESSSVEM-G*......  1016
WRC.CI.97.97CI14        -QQ-Q--G--P------L--ETPEG.-IT-------L-KVWNGEGMDRSSSDKN*....  1022
WRC.CI.98.98CI04        -QQ-Q--G---------L-VETPEG.-IT--K------KVWNGEGMDRSSSNKD*....  1021
WRC.GM.05.Pbt_05GM_X02  SQQ-Q--G--V------L-V-TPEG-LLT---------KTWDGKGMDSSTTN..*....  1012
LST.CD.88.524           D-Q-R---Q----------VKEGE.N-FSI------LVK---EGPKDSQSSMDN-*...  1020
LST.CD.88.485           D-Q-----QI---------VKEGD.T-FS-------LVK---EGPKDSEGSLDNN*...  1019
LST.CD.88.447           D-Q-----Q----------VKEGD.T-FS-------LVK---EGPKDSEGGLHNN*...  1019
LST.KE.x.lho7           D-Q-----Q----------VKEGE.N-FS-------LVK---E.GPK-SESSLDNN*..  1019
SYK.KE.x.SYK173         .NE-----R----------V-TEEG--FA--------T-H-ER-DSGSH-ENDPKT--  1037
SYK.KE.x.KE51           .SD-Q---S------------TPDQQVIA---------SSD-ERVDSGTHLEITSKSN-  1041
COL.CM.x.CGU1           TGE-Q--GE-V---------RNSQGTLF-K----V--T-LQYGEDV-SENLL-NGQKEA  983
```

```
                          Vif start
H1B.FR.83.HXB2            MEN................RWQ..VMIVWQVDRM..RIRTWKSLVK...HHMYVSGKARGWFYRHHYESPHPRIS.SEVHIPL.GDAR.........LVITTYWGLHTGE.....RDWHLGQGVSIEWRKK.............................RYSTQV   98
H1A1.UG.85.U455           ---................---..----------..K----N----..------K--Q---------R-S-V-.-------.-E--.........--VR-----------.....K-----H-------L-.............................------   98
H1B.US.90.WEAU160         ---................---..----------..----------..----I-K--K--S------T-----.-------.-EGK.........-------------.....----------------Q.............................------   98
H1C.ET.86.ETH2220         ---................---..-L--------..K----N----..---HI-RR-N--V-----D-R--KV-.-------.-E--.........-I-K-----Q---.....------H-------LR.............................S-N---   98
H1D.CD.84.84ZR085         ---................---..----------..--N-------..Y--HI-K--K-------D----K--.-------.----.........--V-----------.....-E-----------R.............................------   98
H1F1.BE.93.VI850          ---................---.L----------..--N-------..Y-----K--K--S----FQ-R---V-.-------.EEVK.........----------P--.....-E----------QG.............................K-R--I   98
H1G.SE.93.SE6165          ---................---..----------..-----H----..------K--------P--A-R---V-.-------.---T.........--V-----------.....K--Q--H-------QR.............................--R---   98
H1H.CF.90.056             ---................---..----------..--N-------..Y--HI-R---------F--T------.-------.-E--.........----------N--.....-E----------L-.............................------   98
H1J.SE.93.SE7887          ---................---..----------..--N-------..Y--N--K---K-L-----D-N--K--.-------.-E-I.........----------Q--.....---------------QR.............................--R---   98
H1K.CM.96.MP535           ---................---..----------..--N-------..Y--HI-K--NR-Y-------R--K--.-------.---E.........--V-----L--.....---------------L-.............................--R---   98
H101_AE.TH.90.CM240       ---................---..----------..-----N----..----I-K--KK---------Q--KV-.-------.-E--.........---R-----Q--.....K-----H-------QR.............................K----I   98
H102_AG.NG.x.IBNG         ---................---..----------..-----N----..------K--K-------F--R--KVC.-------.----.........--VR-----------.....------H------KQ-.............................-----I   98
H103_AB.RU.97.KAL153_2    ---................---..----------..-----N----..--I-I-K-----V-K-----RN---.-------.---K.........--K------------.....--------A-----E.............................------   98
H104_cpx.CY.94.CY032      -A-................---..----------..K----N----..------K--K-*--------R--KV-.-------.-E--.........--VR-----QP--.....Q-----H-------LR.............................------   97
H1N.CM.95.YBF30           ---................---..--V-------..K--K-N----..------K--K--Y------TH--K--.------V.-Q--.........--TV-----T--.....QS----H-------LR.............................K-K---   98
H1O.BE.87.ANT70           ---................---..-L-------Q..KVKA-N-----..Y-K-R-R-TEN-W-------RN--V-.-S-Y--V.-V-H.........V-V------MP--.....--E---H-------Y-.............................K-K--I   98
H1O.CM.91.MVP5180         ---................---..-L----I--Q..KVKA-N-----..Y-K-M-K--AN-R-------RN-KV-.-A-Y--V.AE-D.........I-V------MP--.....-EE---H------QY-.............................E-K--I   98
CPZ.CD.90.ANT             I-K...............-*K..IEMI------..--NI--------..--IWETKVLKP-K------ND--KKG.E------PTLDKK.........--V-VF---QC--.....-P----H-------CG.............................K-I---   99
CPZ.CM.01.SIVcpzCAM13     ---...............K--..----------..--KA-N-----..Y---KTK--KD---------H-N--V-.------FKDNSK.........-I-----A-T--.....-A----H----Q--MG.............................S-V---   99
CPZ.CM.05.SIVcpzEK505     ---................---..--V-------..-V---N-----..------K-----------TDN-K--.--I-----.-E-K.........---V-----MP--.....-P----H-------QG.............................I-R--I   98
CPZ.CM.05.SIVcpzMB66      ---................---..----------..--K--N-----..------R---N---------NN-K--.-------.-E-K.........----------T--.....-A----H-------QR.............................S-R--I   98
CPZ.CM.98.CAM3            ---...............--L..----------..-VKA-N-----..C---KTK--KD---------S--KV-.--I---I.EK-K.........--V------MP-Q.....HA----H-------CG.............................S-----   98
CPZ.GA.88.GAB2            --S................---..-----L----..-----N-----..F---K-K-----------D--N-KV-.------I.R--K.........I-VC---N-MP--.....QP----H-------QR.............................T-R--I   98
CPZ.CM.98.CAM5            ---................---..--V-------..--K--N-----..T---RTK--KK---------DN-KV-.-------.DT-K.........--V------T--Q.....HA----H-------CG.............................S-K---   98
CPZ.CM.05.SIVcpzMT145     ---................---..-Q--------..-VK--N-----..Y---R-K--KN-------DGTN--MA.-QIY--I.NQ-K.........-IV------MP--.....------H--A----ER.............................G-----   98
CPZ.TZ.00.TAN1            ---................---..-QV--MI---..-L---T-----..--IFTTKCCKD-K------TDT-KRA.G-I-----.TERSK.........--VLH---AC--.....-P----H-IGL---QG.............................K----I   99
CPZ.CM.05.SIVcpzLB7       ---...............--K..-L----I---..--KA-N-----..Y-I---R-------------EN-KV-.--I-----.-E-K.........---V-----TP--.....------H------KQ-.............................K----I   98
CPZ.US.85.CPZUS           ---................---..-L--------..-VKA-N-----..Y---R-K--K-----------TN--VN.------I.DK-K.........I--I-----K---.....QA----H-------QG.............................T-K--I   98
CPZ.GA.88.GAB1            ---................---..----------..--K--N-----..Y-I-R-K-----------DH-N-KVA.--I---FRDYSK.........-IV----A-SP--.....-A----H----Q--LG.............................S-V---   99
MUS_1.CM.01.1085          -APR..............KM-V.VSP-QETSE-...K-DWIIRAT-Y..YILS.GR..AAFT-V---QLY-Q-F-Q-KIR---TQNT-ESGAVESTYIEL-VL-DVTNVGPASCSKSTWWK-SFIL--VYLRR...................HPQDREQDYIW---YL  126
MUS_1.CM.01.CM1239        -APR..............KM-V.VSP-QETTE-...KMDWIIRAT-H..-IFS.GK..TPFV-V---QVQ-Q-F-QNKIKLA-TQNTLESGETEVTFIE--VL-DVTNVGPASLSKSTYWK-SYIL--VYMRK...................GPRDRERDYVW-H-YL  126
MUS_2.CM.01.CM1246        -APG..............KM-I.VSP-QYMSE-...KMDWLIRCT-H..YIFS.GK..APFT-V---QLQ-Q-FTQNKIK---SINRIEGGPTETTYIE--ILLDVTNVGPASLS-STYWQ-SYILK--YIRA...................APQNREIDMVH-E-FL  126
MUS_2.CM.01.CM2500        -AQR..............KM-I.VTP-QFMTE-...-MDWLIRCT-H..-ILS.GK..APFS-I---QVQ-Q-FTQNQIK---TIKK-EEDKVESTYIQ--IL-DVTNVGPASLSKSTYWQ-SYIL--VYVRA...................APQDREVDMVW---YL  126
                          Vif start
MAC.US.x.239              --E...............EK--I.A-PTWRIPE-...LE-WHSLIKY...LKYK.TKDLQKVC-VP-FKVGWAWWTC-R-IF--QE.........GSH-EVQG--H-TPE......KG-LSTYA-R-T-YS-............................NFW-D-  100
H2A.GW.x.ALI              --E...............GKS-I.V-PTWRVPG-...MEKWHSLVKY...LKYR.TKDLEKVC-VP-HKVGWAWWTC-R-IF--Q-.........RSH-E-QA--N-TPE......KG-LSSYA-R-T-YTE............................KFW-D-  100
H2A.DE.x.BEN              --E...............DRN-I.V-PTWRVPG-...MEKWHALVKY...LKYR.TKDLEEVR-VP-HKVGWAWWTC-R-IF--Q-.........KSH-E-QA--N-TPE......KG-LSSHA-RLT-YTE............................KFW-D-  100
H2A.SN.x.ST               --E...............GK--I.A-PTWRVPG-...ME-WHSLIKY...LKYR.T-DLEKVC-VP-HKVGWAWWTC-R-IF--K-.........ESH-E-QA--N-TPE......KG-LSSYS-RLT-YTE............................KFW-D-  100
H2B.GH.86.D205            --E...............EKD-I.V-PTWRIPG-...LE-WHSLIKY...LKYR.T-ELQQVS-VP-HKVGWAWWTC-RIIF--NK.........GAW-EVQG--N-TPE......-GFLSSYA-RLT-YER............................NFY-D-  100
H2B.CI.x.EHO              --E...............EKN-I.A-PTWRIPC-...LE-WHSLIKY...LKYR.TKDLQQVS-VP-HKVGWAWWTC-R-IF--KE.........GAH-EVQG--N-TPE......-GFLSSYA-RLT-YER............................SFY-D-  100
H2G.CI.92.ABT96           --E...............EK--I.A-PTWRIPG-...LEKWHSLIKF...LKFK.TKDLQKAV-VP-HKVGWAWWTC-R-IF--TK.........EAH-E-QG--N-TPE......KG-LSQYA-RLT-YTR............................KFY-D-  100
H2U.FR.96.12034           --E...............EKN-I.V-PTWRVPR-...LE-WHSLIKH...L#YN.TKELQKVC-VP-HKVGWAWWTC-R-IF--KK.........GSH-EVQA--N-TPE......KG-LSSYA-RLT-YTR............................GFW-D-   99
MAC.US.x.251_1A11         --E...............EK--I.A-PTWRIPE-...LE-WHSLIKY...LKYK.TKDLQKVC-VP-FKVGWAWWTC-R-IF--QE.........GSH-EVQG--H-TPE......KG-LSTYA-RVT-YS-............................NFW-D-  100
MAC.US.x.251_BK28         --E...............EK--I.A-PTWRIPE-...LE-WHSLIKY...LKYK.TKDLQKVC-VP-FKVGWAWWTC-R-IF--QE.........GSH-EVQG--H-TPE......-G-LSTYA-R-T-YSR............................NFW-D-  100
MAC.US.x.EMBL_3           --E...............EK--I.A-PTWRIPE-...LE-WHSLIKY...LKYK.TKDLQKVC-VP-FKVGWAWWTC-R-IF--QE.........GSH-EVQG--H-TPE......-G-PSTYA-R-T-YSR............................DLLDRC  100
SMM.US.x.H9               --E...............EKN-I.V-PTWRIPG-...LEKWHSLIKH...LKYN.TKDLQKAC-VP-HKVGWAWWTC-R-IF--RD.........EAH-EVQG--N-TPE......XG-LSTYA-R-T-YSR............................NFW-DA  100
SMM.US.x.PGM53            --E...............EKN-I.V-PTWKIPG-...LEKWHSLIKY...LKFR.TKDLQKAC-VP-HKVGWAWWTC-R-IF--RD.........ESH-EVQG--N-APE......KG-LSTYA-R-T-YSR............................KFW-D-  100
SMM.US.x.SIVsmH635F_L3    --E...............EKN-I.V-PTWRIPE-...LE-WHSLIKH...LKYN.TKDLQMAC-VP-HKVGWAWWTC-R-IF--RD.........KTH-EVQG--N-TPE......KG-LSTHA-R-T-YSR............................NFW-D-  100
SMM.SL.92.SL92B           --E...............GK--V.A-PVWRISR-I..VE-WHSCIKF...-KYK.TRELEKAC-VP-HKVGWAWYTA-R-IF--EE.........GSH-EVQV--N-TPE......KG-LSSYA-R-T-YSE............................KFW-D-  101
STM.US.x.STM              --E...............EK--I.V-PTWRIPG-...LE-WHSLIKH...LKYN.TKELSKAC-VP-HKVGWAWWTC-R-IF--Q-.........EAH-EVQG--N-TPE......KG-LSEYA-R-T-YTR............................NFWSD-  100
MNE.US.x.MNE027           --E...............EK--I.A-PTWRIPE-...LE-WHSLIKY...LKYK.TKDLQKVC-VP-HKVGWAWWTC-R-IF--QE.........ESQ-EVQG--N-TPE......-G-LSTYA-R-T-YSR............................NFW-D-  100
GRV.ET.x.GRI_677          --R...............EKQ-I.VRVVWRVSE-..QIS-WRGIVTY...KIR..NKQLP.-E----WQVQWQFWTY-QFI---SK.........DDYIEVNI-HN-TPE......-G-LSSH--GLSYYHQKG...........................-K-E-  100
VER.KE.x.TYO1_patent      -NP...............NKE-V.MRVTWKVPGD..LITKWQGIVRY...WMR..QRNLK.-N-YM--QITWAWYTM-RYV--IGK.........HGEICVDL--H-TPE......QG-LSTYA-G-QYVSNLES...........................K-R-EL  102
VER.KE.x.9063             -NQ...............EKE-V.MRVTWKVSGD..LITKWQGIVKY...WMT..KRNLK.-Q-MM--QINWAWYTMCRYI--IGL.........EGEIQ-DC--H-TPE......-G-LSTYAIG-QYLGAVG............................NFR-EL  101
VER.KE.x.AGM155           -SQ...............EKH-V.MRLTWKVQEE..VITKWQGIVRY...WMN..KRNLK.-E-KM--QITWAWYTM-RYV---P-.........SGEIHVDI--H-APK......QG-LSTYA-G-QYVSLVND...........................K-R-EL  102
VER.DE.x.AGM3             -NQ...............EKE-V.MRVTWKVPEE..LITKWQGIVRY...WMR..TR-LD.-K--M--QITWAWYTM-RYE---GQ.........HGSIHVDL--H-TPE......KG-LSTYAEG-QYLSNRDP...........................W-R-EL  102
TAN.UG.x.TAN1             --R...............EKL-V.TRLTWRVSGE..HIDKWKGIVKY...-MR..NRLQD.-T-LM--QCGWAWYTC-RFL---G-.........EGKI-VDC--H-TPE......QG-LSTYA-A-SFENWQN............................T-K-E-  101
SAB.SN.x.SAB1C            --................KH-I.VRPLWKVTGG..QQE-WTSLVKY...-MHVSKQCVH.-R-TP-TKIRWNWY-YQ-WV---KD.........GALIKV-N--H-TPE......KG-LETYATG-GYS...KG...........................EWF-EL   99
OLC.CI.97.97CI12          --..................-I.A-PMYRLNNG..QVEHVKFMAK......LAYKFWGYH-IWQNC-KY-REWWTHGKIY--VP-GE..........IE--L--HIRH-......KGLQNSG--NFKYIVG.............................ISWMYL   94
DRL.x.x.FAO               --................KM-I.VIPTWRVERY..MVE-WNSLCKY...-KHKGEKYLER-H-AP-FQCS.GWWTH-QWT--FKD.........KSKI-V-VL-N-TPD......KG-LSKYAIT--HIQE............................KFC-LI   99
RCM.NG.x.NG411            -A................EK--I.ERP-WKVDR-..KVEQWHSLVKY...-QYKGKKE-KD-E-VP-FKV-WGWW-R-------GN.........NTKIKV----S-TIE......KG-LGTY--G-AYIDNRCDP..........................C-F-DI  105
RCM.GA.x.GAB1             -A................EKM-I.VRP-WRVDR-..KIEQWHSLVKY...-QYKGKKA-KE-E-VP-FKV-WGWW-H-------EE.........GSK-K-----N-TVE......KG-LGTY--G-L-I-G............................D-V-D-  101
MND_1.GA.x.MNDGB1         --................RVERI.VRLTWKVSSQ..RIEKWHWLVRRQMAWATANNEEGCW-L-P-FMAYNEWYTC.-K-V-IINR-I-..........-IVRS--H-QIE......VGCLSTYA----AVVRPPPFEK.......................EWC-EI  109
MND_2.x.x.5440            --................KQ-I.AIPTWE.DRQDVKLRDWNSIVKY...-KYKGEKHLEN-QLYP-FQCS.GWWTH-QKK--FKD.........GSKII--VL-N-TPE......KG-LSQYAFTV-YQ-E............................N-F-YI  100
MND_2.CM.98.CM16          --................KQ-I.VIPTWKLDRH..AVE-WNSLVKY...-KYKGEKHLDK-DLKA-FQCS.GWWTH-QKY---EE.........DEQII--IL-N-TPE......KG-LSTYAMT--YKFT............................N-Y-HI   99
MND_2.GA.x.M14            --................KQ-I.VIPTWRIDR-..QLEKWNSLVKY...-KYKEEKHLDK-ELF--FQCS.GWWTH-QKI--FKD.........GSKII--AL-N-TPE......KG-LSQYAITA-YN-G............................D-Y-HI   99
GSN.CM.99.CN166           --RQR.............QKM-V.VQP-QIMTE-..KVDWLLRATKH...-IWS.GK..TPFV-V---QLQ-Q-FTQNKIRLAMDTRKVG.EEVEATYIE--IL-DTTSHGPASLS-STYYQ-A-I---VYNRS....................KLGAREGDIIW---NL  127
DEB.CM.99.CM40            --................KT-I.VRLTHKIPMN..KLD-WCSAIKR...EKYK.TKELDTVQWI--F-LRNSYYTQTKIVF-ITK.........EGGV-VELL-C-TPE.....KGWLPTMA.-A-A-ET............................KTWK-EL   99
DEB.CM.99.CM5             --................KE-I.VRLTHKIPVN..KLDKWCSAIKR...GKYK.TKELQAIKWI----LRRAFYTQTKIVF-VTE.........EGKV-VEIF-S-APE.....KGWMPSMA.-G-S-EQ............................TGWQ-EL   99
GSN.CM.99.CN71            --RQR.............QKM-V.VQP-QVMTE-..KVEWLIRATKH...-IWS.GK..TPFV-I---QLQ-Q-FTQNKIRLA-DTRQVG.EEVEATYIE--IL-DTTSHGPASLS-STYYQ-A-T---VYNRS....................KLGARETDITW---NL  127
DEN.CD.x.CD1              -Q-LGRENR.........TKE-I.VRLTYKIGET..KLEHWMST-KS...IIYK.TNELK-TVIIP-FRITWEWYTMTKI----R..........AAYMEVKL--H-TPE.....KGYLSQVA.I--AYVHL...........................STGWA-E-  108
TAL.CM.01.8023            -D................KL-V.TRLTWKVPRN..RMQ-WVGTIRY...-MYK.TKYLQKLIWV--FQLRMASFTQAKWIL--EEKENR.....VEYIT--Q--C-APD.....KGYTSDWDTI----SVSDV.......................TLDAKV---EL  111
TAL.CM.00.266             -D................KL-V.TRLTWKVPRN..RMQ-WVGTIRF...-IHK.TKYLQKLIWV--FQLR-ASYTQAKWIL--EEKDTR.....VEY-T--Q--C-APD.....HGYTSDWDTI----SVSDI.......................VRDEKV-A-E-  111
MON.CM.99.L1              -APR..............KM-V.VTPTNIISES..KMDWLIRCTKW...-ILT.GS..APFV-V---QLHNQ-F-QNKIKL--DLGITQEGESWATY-E--I--D-TNVGPSALSPTVYNK-AYT---VYWMQKVERVP.....LDPVWNPNPR..DREV.RTW-A-HL  138
MON.NG.x.NG1              -APR..............KM-V.VSPVHIRSE-..WMEWLIRATKY...NIYT.GK..VPFV-V---QLT-Q-F-QCKIKL--EMGITQXXEQWATYXE--IL-DXTNVGPASLSPSVKWK-AYI---VYWRQ..PRVPQQEGVVDRVWFPSPRRPDREIDRVW-A-YL  144
WRC.CI.97.97CI14          --................KE-I.G-PLIRISKG..KFR-WTYYIA......NLIKAGQANFLWGSWHAAAECFTR.HI-K--IQ-HWE.........MQ-RC--N-SPE......KG-MDTM-FN-RMVYEN............................-WE-FI   97
WRC.CI.98.98CI04          --................KE-I.G-PLIKISKG..RFK-WAYWIA......HLMQN-KASFLWGTWHTQAECFTR.YLLR---Q-HWE..........-Q-RC--N-APE......KG-LDTM-FN-RMVYEN............................-WE-FI   97
WRC.GM.05.Pbt_05GM_X02    --................KE-I.A-PLIKISKG..RFR-WSYWIA......ELIQE-QARFLFGSWHVHAECFTR.HI-R---Q-HWE.........MQ-RC--N-APE......IG-IDTL-FN-RMVWEN............................KWE-FI   97
SUN.GA.98.L14             --ENPPQWRWWDEREWEDRRQYKIVRIVWLIDRI..AVEKFLDLRR...M-RET...-DDWVSMYGTGTGWEWYTY.NKII--VTYGT.........V-VRI-GH-TPA......KGLINQW-C-M--IYN............................S-Q-EI  113
LST.CD.88.524             --................RDQRI.VRVAWITNRS..LVEKFLDIKR...N-SES...-NDWVSLHGTGVGWEFYTY.NKLL---DTGT...........--VRI-AH-AA-.......-G-ISQWA-----IYG............................N-Q-E-   96
LST.CD.88.485             --................RDQRI.VRVAWITTRN..LIEKFLDIRR...A-KET...QVDWVSLHGTGVGWEFYTY.NKIV---EIGT...........--VRC-AH-AA-.......-G-ISQWA-----IYG............................N-Q-EI   96
LST.CD.88.447             --................RDQRI.VRVAYITTRS..LVEKFLDIRR...A-KET...QNDWVSLHGTGVGWEFYTY.NNIV---ETGT...........--VRC-AH-AA-.......-G-ISQWA-----IYG............................N-Q-EI   96
LST.KE.x.lho7             --................RDQRI.VRVAWITTRS..LIEKFLDIVR...LNTEA...-ENWVGLHGTGVGWEFYTY.HKLV---ESGT...........--VRI-AH-AA-.......-G-ISQWA-----VYR............................D-Q-EI   96
SYK.KE.x.SYK173           --................KE-I.V-PTWRMTP-..QID-LQHIIKT...-KYK.-KELEKAT-K---QIEWQWYTYCQWT--VG..........DGTIW--F-HN-APE.....RGWL-MQG.IR-QYQWN............................QWN-DL   98
SYK.KE.x.KE51             --................KE-I.V-LTWKLHPN..QIDKMQHI-KV...-KYK.-KQLEKAT-K---QITWEWWTRAQWE--VG..........QGV-I-KF-HN-TPE.....KGLLRTEG.-GLI-AHPS............................GW--E-   99
COL.CM.x.CGU1             --K.................M-A.VRICYPMGK-..KPKPLREWIK-.........QFPECQY-YLGKTESPEWWTEGKIEFY-GSAIK.........-S-IYLGT-TPDR.......Y-R-PP--Y-LQIG............................KWE-DL   93
```

```
                                                                                                                                                                    Vif end
H1B.FR.83.HXB2           DPELADQLIHLYYFDCF.SDSAIRKALLGH..IVSPRCEYQ..AGHN...KVG...SLQYLALAALI.............................TPKKIKPPLPSVTKLTEDRWNKPQKTKGHRGSHTMNGH*.......................... 192
H1A1.UG.85.U455          --D---H----H-----.-E----R-I--Q..--R------.----..---...-------K--V.............................--TRA-------K-------------R-----R-L-R--.......................... 193
H1B.US.90.WEAU160        --D---------H----.-E----N-I---..L-I------.----..---...-------T---.............................--------------K--------------------------.......................... 193
H1C.ET.86.ETH2220        --G---H---MH-----.AE------I--Y..R-----D--.----..---...-------T---.............................K---A-------S--V--K-------R-R--N-------.......................... 193
H1D.CD.84.84ZR085        --G-------M------.A-------I---..---------.----..---...-------T---.............................A---R-------K----------RQ---R----------.......................... 193
H1F1.BE.93.VI850         --G-------I------.-E------I---..RI----N--.----..---...-------T---.............................A-E-T-------Q--V---------E-R------------.......................... 193
H1G.SE.93.SE6165         E-D---H----H-----.---------I--Q..---------.----..Q--...-------KV-V.............................-S-RSR-------E-A----------R---ENP------.......................... 193
H1H.CF.90.056            E-G-------MH-----.-E------I--R..V-R---N-P..---K...Q--...T------T--V.............................A-----------R--V----------R------------.......................... 193
H1J.SE.93.SE7887         --G-------MC-----.---------I--Q..------D--.----..---...-------T---.............................R--RR-------Q--V----------T---E---------.......................... 193
H1K.CM.96.MP535          E-D-------I------.-E--V---I---..R------C-.----..---...-------T--V.............................A-RRP---V---K--V----------R-----Q------.......................... 193
H101_AE.TH.90.CM240      --D---K----Q--G--.---------I--Q..V-RR----P..S---..---...-------K--T.............................---R-R-------.EI-----------.R---ENP------.......................... 191
H102_AG.NG.x.IBNG        --D-----------N--.---------I--E..--R------.----..---...-------N--V.............................A-T-T--------R--A----KE----R-----RP-----.......................... 193
H103_AB.RU.97.KAL153_2   --N--------------.-E----N-I---..R---S---R..----..---...-----------R.............................---------------------R--D-------S---.......................... 193
H104_cpx.CY.94.CY032     --D-------MH-----.-E------I---..R---------.----..---...-----------..............................S---T-------K--V----------R-R-ENQI-----.......................... 192
H1N.CM.95.YBF30          ---M--K----H-----.TA----Q-V--R..P-L-----P..---K...Q--...T------T-WV.............................GA--R---------------EH--MQ----NPI-----.......................... 193
H1O.BE.87.ANT70          ---T--RM---H--T--.TA--V----I--Q..R-LTK---P..T--S...Q--...T--L---R-VV.............................KARSR-------Q---------HLRIRDQLK-PS-----.......................... 193
H1O.CM.91.MVP5180        ---T--RM---H--T--.TE------I--Q..R-LTK---L..---S...Q--...T--F---K-VV.............................KV-RN-------QR---------W-IRDQL---S-----.......................... 193
CPZ.CD.90.ANT            ---T---M--Q---P--.--Q-V-Q-I--E..RILTY-H-K..K--S...Q--...T-----FCKILEFRG.........................Y--GPRRQF--LSI---------RRMR---ENQ*.............................. 192
CPZ.CM.01.SIVcpzCAM13    --YV--K---SQ-----.A-----R-I---..--E---A-R..E--R...Q--...T--F---K--V.............................KE-RLR------A--------T--R----PEN---S---.......................... 194
CPZ.CM.05.SIVcpzEK505    ------K----------.TA----Q-V--R..P-I-K---P..---K...Q--...-------I-WV.............................GVQ-RR--------------RH--I-D------T--L-.......................... 193
CPZ.CM.05.SIVcpzMB66     --D--------------.AE------I--Q..--------P..T---..---...-------ST--.............................KS-PRR------R--A-----NS---RD-K-N-I-----.......................... 193
CPZ.CM.98.CAM3           ---M--K----Q-----.AA------V--K..Q-Y-K---P..---Q...Q--...-------R-WV.............................RVG-K---------------RH--IRV-QE--------.......................... 193
CPZ.GA.88.GAB2           --QT--R---C------.AA--V-R-I--E..--TKV--FP..R--P...Q-Q...------IVQVV.............................KT--K--------I-A----------RDQQENL--S---.......................... 193
CPZ.CM.98.CAM5           ---T--R----H-----.AA------V--K..Q-Y-K---P..T--Q...QX-...-------R-WV.............................GVQ-KT-----IX--A------Y---R--QEN----X--.......................... 193
CPZ.CM.05.SIVcpzMT145    ---T--R----H-----.TAT---N-I-EK..A-L-K--FP..E--S...---...------W--KQ-V............................KT--NR------A--A---------R-RD-QE-PMSS-C-.......................... 193
CPZ.TZ.00.TAN1           ---T------TR--T--.AAG-V-Q-I--E..RILTF-HF-..S--R...Q--...T--F--FRKVVESQDK..........................Q--GPRR--------------HRT-T-R-EN--LS-C-.......................... 199
CPZ.CM.05.SIVcpzLB7      ------HM------N--.AE------I--R..V------FP..---K...---...T-----IS--L.............................QRQPTR------S--A------S-----------I-S---.......................... 193
CPZ.US.85.CPZUS          ---M--K----------.TA------V--R..P-I-K---P..T---..---...----F---K--V.............................GQS-RR--------------C-----QPEN---S---.......................... 193
CPZ.GA.88.GAB1           --FT--R---SQ-----.AET---R-I--Q..L-A-----K..E--R...Q--...----F---K---.............................SERRHR------A---------H-R--V-QENL-R----.......................... 194
MUS_1.CM.01.1085         S-T--MRI--TQ--S--.YSQD-PRVIR-QH..TLGD--H....PDAHT.---PPP---V---L--KK.....................H..GKLAP.....QP-SALALQSGPNHHAPGT-SHVGS-RR-RKTLFQRRAPWNLE*.......................... 235
MUS_1.CM.01.CM1239       T--I-MKI--THH-S--.YSQD-PRVIR-QP..PLGD--H....PGAHT.---PPPT--V---LT-KK.....................H..GKSTT.....QPQTSMALQSGPNYHAAGPTGHVGP-RR-RKTLFQRRAPWNLE*.......................... 235
MUS_2.CM.01.CM1246       ---VTM-I--TH--S--.QQRD-QR-IR-EQ..LLRK--H....IKTHYP---TPL--EK--FF-Y-K.....................HTNGASAS.....KS-SLMAKSSNANNYGSGASRPVGT-RR-RKTLLQRQAPWNLG*.......................... 238
MUS_2.CM.01.CM2500       -----M-I--T-H-S--.QEQD-IR--R-ER..ILGI-QH....QTAHLP---TPP--EK--FF-Y-K.....................QGDGATPS.....KS-SPMAK-PGTNNHGAGTKRHVGS-RR-RKALLQGRAPWDLGPSHGAPSRP*............. 247
                                                                                                                                                                                   Vif end
MAC.US.x.239             T-NY--I-L-ST--P--.TAGEV-R-IR-EQ..LLSC-RFP..RA-KYQ.....VP-------KVVS.......................DVRSQGE.....NPTW-QWRRDNRRGLRMAKQNSRGD-QR-GKPPTKGANFPGLAKVLGILA*............... 214
H2A.GW.x.ALI             T-DC--S---GT--S--.TAGEV-R-IR-EK..LLSC-N-P..QA-KSQ.....VP---F---VVVQQ.....................NGKPQRN.....S-TR-QWRRDYRRGLRVARQDSRGL-QR-GESPAPGAHFPGVAKVLEILA*............... 215
H2A.DE.x.BEN             T-DC--I---ST--S--.TAGEV-R-IR-EK..LLSC-N-P..QA-KAQ.....VP-------VVVQQ.....................NDRPQRK.....G-AR-QWRRDHWRGLRVAREDHRSL-QG-SEP-APRAHFPGVAKVLEILA*............... 215
H2A.SN.x.ST              T-DC--S---ST--S--.TAGEV-R-IR-EK..LLSC-N-P..QA-KYQ.....VP---F---VVVQQ.....................NGRPQRD.....N-TR-QWRRNYRRGLRVARQDGRSH-QR-SEPPAPRAYFPGVAKVLEILA*............... 215
H2B.GH.86.D205           T-DV----L-GS--S--.-ANEV-R-IR-EK..ILSY-N-P..SA-EGQ.....VP---F---RVVQE.....................GKNGSQGE....SATR-QRRRNSRRSIRLARKN-NRAQQGSGQPFAPRTYFPGLAEVLGILA*............... 216
H2B.CI.x.EHO             T-DV--R-L-GS--SS-.TANEV-R-IR-EK..ILSH-N-P..SA-TGQ.....VP---F---RVVQE.....................GKDGSQGE....S-TR-QRRRNSRRGIRMARDNIRTSQQSSSQSLAQGTYFPGLAEVLGILA*............... 216
H2G.CI.92.ABT96          T--T----L-GS-----.TAGEV-R-IR-EQ..ILSC-N-P..TA-KRQ.....VP---F---QVVQK.....................G.HGSKGE....SXTR-QRRGDNRRGIRMARKNSXRDQQDSSQSFTQGAYFPGLAKVLGILA*............... 215
H2U.FR.96.12034          T-DC----L-GS--S--.TAGEV-R-IR-EQ..LLSC-NFP..SA-RQQ.....VP---F---R--QD.....................G..RSKRE....NPTR-FRRRNCRRGFQLARKNCERYQQGSSAT-SQGTYFPGLAKVLAILA*............... 213
MAC.US.x.251_1A11        T-NY--I-L-ST--P--.TAGEV-R-IR-EQ..LLSC-RFP..RA-KYQ.....VP-------KVVS.......................DVRSQGE.....NPTW-QWRRDNRRGLRMAKQNSRGD-QR-GKPPTKGANFPGLAKVLGILA*............... 214
MAC.US.x.251_BK28        T-DY--I-L-ST--P--.TAGEV-R-IR-EQ..LLSC-KFP..RA-RYQ.....VP-------KVVS.......................DVRSQGE.....NPTW-QWRRDNRRGLRMAKQNSRGD-QR-SKPPTKGADFPGLAKVLGILA*............... 214
MAC.US.x.EMBL_3          NTR-CRHFSCIALISL-.TAGEV-R-IR-EQ..LLSC-KFP..RA-RYQ.....VP-------KVVS.......................DVRSQGE.....NPTW-QWRRDNRRGLRMAKQNSRGD-QR-GKPPTKGADFPGLAKVLGILA*............... 214
SMM.US.x.H9              T-DY--T-L-GT--P--.-EGEV-R-IR-EK..LLSC-KFP..KA-KNQ.....VP-------TVVS.......................HVXSQGE.....XPTW-QWRRNXRRGXRLATQNXRRN-QGSSESFAXGANFPGLAKXLGILA*............... 214
SMM.US.x.PGM53           T-DY--T-L-GT--P--.-EGEV-R-IR-EK..LLSC-KFP..KA-KNQ.....VP-------TVVS.......................HVRSQRE.....DPTW-QWRRNNRRGIRMAKQNSRRN-QGSSESLAEGANFPGLAKVLGILA*............... 214
SMM.US.x.SIVsmH635F_L3   T-DC--T-L-ST--P--.-EGEVQR-IR-EK..LLSC-KFP..KA-KNQ.....VP-------TVVS.......................HVRSQRE.....DPTW-QWRGNNRRGFRMAKQNSRRN-QGSSKSPAEGANFPGLAKVLGILA*............... 214
SMM.SL.92.SL92B          T-DV----A-ST--P--.AAH-V-Q-IR-EQ..-LSY-G-A..VA-HSS.....VQ---L---KVVLQ.....................NDRPKGK.....NPTR-QWRRNNRRGLRMVTQYSGGPESNSSTTPTKRVNFPGLEKVLGILGR*............... 217
STM.US.x.STM             T-DC----L-GT--P--.TAGEV-R-IR-EK..LLSC-RFT..KA-KNQ.....VP-------KVVE.......................HVRSQRE.....N-AR-QWRRGNRGSIRVATQNGRGH-PR-SKP-TEGTDFPGLAKVLGILA*............... 214
MNE.US.x.MNE027          T-DY--I-L-ST--P--.TAGEV-R-IR-EQ..LLSC-RFP..RA-KNQ.....VP-------RVVS.......................YVRSQRE.....NPTW-QWRRDNRRSLRMAKQNSRGD-QR-SKPPTKGADFPGLAKVLGILA*............... 214
GRV.ET.x.GRI_677         --GT--RM------N--.T-R--QQ-IR-E...KYTW-TFK..E--KGQ.....VQ---L---V-Y.....TNGIR...............KRSKRT....FTRMAGNLGSRQGAMGRMAT-HAQGS-RRSQKALWNEHANPSMELLCRGGKET*.............. 219
VER.KE.x.TYO1_patent     --AT--SI--GH--N--.KER--QQ--R--...RFVF-QFP..E--KST...-QVP-------L-H.....QNGLR...............ERSKRG....KTRRSRNLGSKQGAVGRMAK-YVTRSQPG-EAAFWERTPVPSMELLSGGRRKTWYSHDGKGLQIL*. 235
VER.KE.x.9063            --AT--GM---H--S--.TER--Q--IR-E...RFVF-QFP..E--KTT...-KVP-------L-H.....QNGIR...............ERSQRK....QARRSRDLGSKQRTMGRMAKGHVRRSQSGSQITFWQRAPLPSLELLSGGRRETGSPNDGEGL*.... 231
VER.KE.x.AGM155          --NT--SM--CH--T--.T-R--QQ--R-N...RFIF-QFP..G--KLT...-QVP-------L-H.....QNGLR...............KRSQRG....ETRRTRNLGSQQGAVGRMAQ-YGRRNQQRSQTAFWPRTPIPSMELLSGGRGETGKTHSGKGI*.... 232
VER.DE.x.AGM3            --AT--S---TH--T--.TER--------Q...RFTF-QFP..E--KKT...-QVP-------L-H.....QNGLR...............QRSQRS....KTGGTRNMGFEQGAVGRMAK-HARRYQSGSQDAFWARAPVPSMELLSGGGRKESHSHARKGL*.... 232
TAN.UG.x.TAN1            T-DV--HM--CH--P--.T-R--QQ-IR-E...SFLW-T-K..E--VAENHW-QVR---F---TVYTDF..LRNGRR..............KRFQGK....KTRMVRNLGSQQGAVGRMIK-HGSRTQSGSTTPFWERTPLPSMELLSGRRGKEWGTNDRKGL*.... 238
SAB.SN.x.SAB1C           --WT--HI--WS--P--.T-R-VQQ-IR-E...KYLW-KH-..V--QPT...-QVP-------RVY......TNGLRRVAPTSRRGSSQGSPQESQRR....DTRMARNMGFAQRAVRRMAP-HVTGPQFR-PVPLPKESPFPSLVEYCGRTSH*............. 234
OLC.CI.97.97CI12         T--YV-FMV--V--A-W.TQ------I--MP..DHRY-QLET..--LG......VL---L---R-VRNG...............................RKSSRGSGATQGA-GGLGE-H-GRNYCR-KKALSQGVIGPSFSILCSKTYQ*............. 203
DRL.x.x.FAO              --TT--RI---E--P--.T-Q---GV---Y...RIQA-YFP..R--KGQ.....VG-------Q-H.....IKYLKH.................GRKTVSGESSSGAN..GSRRGRVRRMATEESLKNQPG-SITLPPRVPLPSLEHLHGALP*............. 217
RCM.NG.x.NG411           --TV--K-V-NQ--P--.T-K---Q-I--E...QVLL-GFE..K--RDQ.....VG------IK-W.....ARQQLKD..............GRGQRK....SARGPH..WGWRCRVCPLASQNANRSQLRSQVALSSRVHFPSVAYLCGTLA*............... 222
RCM.GA.x.GAB1            F-WT--S---KI--P--.T-R---QS---E...KVLV-AF-..G--RDQ.....VGT--F--IQ-W.....AKSQLD...............RYGRK....SPRGPH..WGWR-RVPALATGHARKGQLGSQVTLSSRVHFPSVAHLCGTLA*............... 216
MND_1.GA.x.MNDGB1        T--V--H----HFY---.M---VM--IR-E..E-LKV-RFP..---KAQ....GVL---F-C-RVIYGP.................................EER*ASIRRSRT.....SQRAIQSVASRYY-GNKGRSKKALPSHYPKCSIRILCAKHRE*..... 222
MND_2.x.x.5440           --VT--RI--GE--P--.T-Q------F-E...RLVA-YFP..W--RGQ.....VGT--F---Q-Y.....LRGRKH...................GRE....GTRGAR..RSRRGR-GAMAGNVTGENQPG-PVTLPPRVPFPSLEHMCRTLA*............... 213
MND_2.CM.98.CM16         --QT--RM--WK-LP--.TEQ---Q----K...RLTV-YFH..W--KGK.....VG-------LSY.....TAYCNN...................GRR....GPRDPR..RSRRGGS-YLVGKISGENQSG-KITLPPRVPFPSLERMHRTLA*............... 212
MND_2.GA.x.M14           --VT--RM--WE--P--.TAA-V--V-Y-E...RIVA-YSP..W--KGQ.....VGT--L---R-Y.....IKFCRN...................GRK....STRGTT..GGRR-G-RTVAGKIIGTSEQR-SITLPPRVPFPSLEHLCRTLA*............... 212
GSN.CM.99.CN166          T-GV-M-I--GK--P--.QNED--R-IR-EQ..ILGN--H....PEAHL.Q--TPQT-EQ--FF-YVK.....................H.YGKNPT.....QSSYALAQ-KGAHNHGYRSKRSVGN-RG-RETLLKGRAPWNLA*....................... 237
DEB.CM.99.CM40           T-D---H----R--P--.-QQ-V-Q-IR-EI..LLQQ-Q....QY-.RGPEGKGPP----I--RVAVK.....................HGKVPTLT........-TTFHLKNNTHD-ARVAAGHARS-RGSPKTF-QRRTIWSLESLCRLTRPTRLDGRTSMGCIGH* 230
DEB.CM.99.CM5            T-D---H----R--P--.-QQ-VTQ--K-EK..LLQQ-Q....QY-.RGPEGKGPP------RTAVK.....................YGKVSTLT........-TTFHIKNSTND-ASTTTSHAGS-RGSFKTL-KRRVMGSLESLCRPSCPTRLDRRTSMGGVSN* 230
GSN.CM.99.CN71           T-DV-M-I--TR--P--.QNEDV-R-IR-EQ..LLGY-DH....PEAHP.Q--TPHT-ER--FF-YVR.....................Y.YGKNPT.....KS-STLDQ-KGAHNHGSCSKRHVGN-RR-RKTLLKGRAPWNLA*....................... 237
DEN.CD.x.CD1             --VT--EI--KDH-S--.TRQ-V-Q--R-EV..-LST-ARKKAHY-PRTGPGREPP----IV-Q-A-R.....................YGKVTTIT........-TS-SFKNGRNHSKAATGSSMGYHRG-KKTLQQRRD-WNMGSLLVITSAPTLDRRTSHGSRRHR 245
TAL.CM.01.8023           T--V--H---TK--P--.TA--V-Q-IR-EV..LLAV--.......VRHKGRDSPP---F---KVAVR.....................HGKVATQS.......SSANNIKAHTN--SSIRKTHERS-GRDAHTL-KRR--WSLESLCRVTSRA*............. 227
TAL.CM.00.266            T-NV--H---TK--P--.TA--V-Q-IR-EV..LLAV-G.......VRHKGRDTPP---F---T-AVR.....................YGKVATQS.......SSTADLQTH-N--ASARKAHERSSRRNAYTL-KRR--WSLESLCRIASRP*............. 227
MON.CM.99.L1             T-D--N-IV-TH--A--.QQLDV-R-IR-EK..LLGQ-QH....LATHYP---TPLT-EK--FL--VR.......................YGATPT.....KSSAPLDL-DG-HR-TGPTGRHLGA-RRSPKTLLKGRAPWHLGTGNGAAC*............... 255
MON.NG.x.NG1             T-DV---I--TR--S--.QA-DV-R-IR-EA..XLGN-HH....IQTHYN---TPPT-EK--FL-CFK.......................YGKCTT.....EPSSTLDL-RG-HRARRSAGRHMGP-RG-PXALLKGRAPWHLGSGH*................... 257
WRC.CI.97.97CI14         M-SE-NKI--SQH-E-W.TQNVV-NTIRTGKIN-IDA-Q-H..K--AG.....VLT--T-C-QV-QHE.....................................QLQR.AGAFRRT........STNSRGPSRYARALC--KTAFPKGSPCKSISLLLSTERV*..... 207
WRC.CI.98.98CI04         M-SE--VM--SQHV--W.TQNVV-DTIRSGRLDIRNA-Q-H..K--AG.....VLT--T-C-QV-QHE.....................................QLQG.TSTYRRT........STNSRGPSRYARTLC--KTAFPKGSPCKSISLLLATEKL*..... 207
WRC.GM.05.Pbt_05GM_X02   TVK---V---SQH-E-W.TQ--VKNTIRQGYLHPIDS-K-H..Q--LG.....VLT--A-C-Q--QHE.....................................QLR-R-A-NRRT........PGNTRGSSRYARAVCH-KTAFPPGSPCKSISLLLATEKL*..... 208
SUN.GA.98.L14            --LV---M--FK----W.TSRC--R-M--E..KILHE-RN-..VA-KG....LVL---F-C-RV-HGQ.....................................QERAS...RTATRRRGATERAICTMAGRYH-RNQGRSGKAFPSSNSCPSLAILCGLYRVRKGRVY 234
LST.CD.88.524            --I----ML-VK----W.VHRD-VR-IR-E..KILHS-KMP..T--KG....CVL-----C-RQ-QHV.....................................QT-TAAQ-RKAPRG-WSSKRAMGSMASR-N-GNQGR-KTPFPRSNFGPSLGILCQQYRVRRGGLH 220
LST.CD.88.485            --I----M--CK--N--.-TRD-----W-E..KIMAI-AMP..T--KG....CVL---F-C-RQ-QHV.....................................QS-AA-ESRKTSRG-WTTKRAMGSMVSR-H-GNKGG-QTPFPRSHFGPSLGILCQQYRVRGGSLH 220
LST.CD.88.447            --I--N-M--CK--N--.-TRD-----W-E..KIMAF-AMP..T--KG....CVL---F-C-RQ-QHV.....................................QN-AA-ESRKTSRG-WATKGAMGSMASR-H-GNKRG-QTPFPRNNFGPSLRILCQQYRVRGESLH 220
LST.KE.x.lho7            --I----M--YK--N--.IQRD--R----E..RIT.M-CQP..L--KG....CVL---FIC-RQ-QHV.....................................QAEAA-KSRETTRG-WTTKGAMGSVVSR-H-GNKRG-KTPFPRSNVGPSLAILCRQYRVRRGGLH 219
SYK.KE.x.SYK173          T-AV--R---NF--P--.TAR-VNQ-VR-EL..LTSH-WT....P-TD.....QVP-------QVYL.....................KDGGGFLQSLPACARNTMVLHSKKCR-DPKR.-QCHCKGR-GSD-SIQAFYSSRNIWSLESILK.....RR....GRD*. 221
SYK.KE.x.KE51            T-QT--A-V-SQ--P--.--R-VQQ--R-EK..LTSH-WN....F-KE.....QVL-------QKYLS....................KDGTGFLQSLPAAARGTTILHSKKCQLGRGRSNKCHCQSR-GSP-SMQAFYSGRNIWSLESMLRGGSRSRRNTHHGLD*. 233
COL.CM.x.CGU1            SWQQ-VGYM-IIRGK-VRIEEEA---IR-LP...WNP-DF-..V--LG.....VLK--E-C-L--R.........................................................EKERTH--MRVLK-TSHGCRTR*......................... 171
```

H1B.FR.83.HXB2	.....	192
H1A1.UG.85.U455	.....	193
H1B.US.90.WEAU160	.....	193
H1C.ET.86.ETH2220	.....	193
H1D.CD.84.84ZR085	.....	193
H1F1.BE.93.VI850	.....	193
H1G.SE.93.SE6165	.....	193
H1H.CF.90.056	.....	193
H1J.SE.93.SE7887	.....	193
H1K.CM.96.MP535	.....	193
H101_AE.TH.90.CM240	.....	191
H102_AG.NG.x.IBNG	.....	193
H103_AB.RU.97.KAL153_2	.....	193
H104_cpx.CY.94.CY032	.....	192
H1N.CM.95.YBF30	.....	193
H1O.BE.87.ANT70	.....	193
H1O.CM.91.MVP5180	.....	193
CPZ.CD.90.ANT	.....	192
CPZ.CM.01.SIVcpzCAM13	.....	194
CPZ.CM.05.SIVcpzEK505	.....	193
CPZ.CM.05.SIVcpzMB66	.....	193
CPZ.CM.98.CAM3	.....	193
CPZ.GA.88.GAB2	.....	193
CPZ.CM.98.CAM5	.....	193
CPZ.CM.05.SIVcpzMT145	.....	193
CPZ.TZ.00.TAN1	.....	199
CPZ.CM.05.SIVcpzLB7	.....	193
CPZ.US.85.CPZUS	.....	193
CPZ.GA.88.GAB1	.....	194
MUS_1.CM.01.1085	.....	235
MUS_1.CM.01.CM1239	.....	235
MUS_2.CM.01.CM1246	.....	238
MUS_2.CM.01.CM2500	.....	247
MAC.US.x.239	.....	214
H2A.GW.x.ALI	.....	215
H2A.DE.x.BEN	.....	215
H2A.SN.x.ST	.....	215
H2B.GH.86.D205	.....	216
H2B.CI.x.EHO	.....	216
H2G.CI.92.ABT96	.....	215
H2U.FR.96.12034	.....	213
MAC.US.x.251_1A11	.....	214
MAC.US.x.251_BK28	.....	214
MAC.US.x.EMBL_3	.....	214
SMM.US.x.H9	.....	214
SMM.US.x.PGM53	.....	214
SMM.US.x.SIVsmH635F_L3	.....	214
SMM.SL.92.SL92B	.....	217
STM.US.x.STM	.....	214
MNE.US.x.MNE027	.....	214
GRV.ET.x.GRI_677	.....	219
VER.KE.x.TYO1_patent	.....	235
VER.KE.x.9063	.....	231
VER.KE.x.AGM155	.....	232
VER.DE.x.AGM3	.....	232
TAN.UG.x.TAN1	.....	238
SAB.SN.x.SAB1C	.....	234
OLC.CI.97.97CI12	.....	203
DRL.x.x.FAO	.....	217
RCM.NG.x.NG411	.....	222
RCM.GA.x.GAB1	.....	216
MND_1.GA.x.MNDGB1	.....	222
MND_2.x.x.5440	.....	213
MND_2.CM.98.CM16	.....	212
MND_2.GA.x.M14	.....	212
GSN.CM.99.CN166	.....	237
DEB.CM.99.CM40	.....	230
DEB.CM.99.CM5	.....	230
GSN.CM.99.CN71	.....	237
DEN.CD.x.CD1	FHKNHAERNMEAL*	258
TAL.CM.01.8023	.....	227
TAL.CM.00.266	.....	227
MON.CM.99.L1	.....	255
MON.NG.x.NG1	.....	257
WRC.CI.97.97CI14	.....	207
WRC.CI.98.98CI04	.....	207
WRC.GM.05.Pbt_05GM_X02	.....	208
SUN.GA.98.L14	ESNHFAQ*.....	241
LST.CD.88.524	EIHHFS*.....	226
LST.CD.88.485	EIHYHSK*.....	227
LST.CD.88.447	EIHYHSK*.....	227
LST.KE.x.lho7	EIHHPSQ*.....	226
SYK.KE.x.SYK173	.....	221
SYK.KE.x.KE51	.....	233
COL.CM.x.CGU1	.....	171

```
                       Vpx start                                                                                        Vpx end
MAC.US.x.239           MSD.PRERIPPGNSGEETIGEAFEWLNRTVEEINREAVNHLPRELIFQVWQRSWEYWHDEQGMSPSYVKYRYLCLIQKALFMHCKKGCRCLGEGHGAGGW.RPGPPPPPPPGLA*   112
H2A.GW.x.ALI           -AN.---TV-----------------D----AL--------------------R---------Q--T-------M---MYT-FM---T---G---P---.-S----------V-   113
H2A.DE.x.BEN           -T-.----V-----------------E--I-AL--------------------R---------A--T-------M---I-T-F-R--T-W--DM-RE-LEDQ----------V-   114
H2A.SN.x.ST            -AG.---T-----------------D----A---------------------R---------I--T-------M---M-I-S-R--T---G---P---.-S----------V-   113
H2B.GH.86.D205         -.-.----V-----D---V----A--E--IT-L--V-----------------A--RE-----I--T-----L-M---M-V-YT------Q----P---.-S------------   112
H2B.CI.x.EHO           -.-.----V---------V-------ET-L-HL--V--------------K--A--RE-----I--T-------M---M-I-FA---G--R----P---.-S------------   112
H2G.CI.92.ABT96        -.-.----X--E--D-----------E--XT----I--------------R-------------A--T-----L-M---M-V------T--QK---P---.-Q------------   112
H2U.FR.96.12034        -G-.----------E---V----S--E---RD---A--Q---------------Q--R------E--T-------------------Q-------P---.-S----------I-   113
MAC.US.x.251_1A11      ---.------------------------------------------------------------------------------------------------.-------------   113
MAC.US.x.251_BK28      ---.-----------------------------------------------------Q---------M----------------------------.-------------   113
SMM.US.x.H9            -XX.---X---------XX--X-D--H--X-----A-----X--------R-X-------M---V--T-X---------------------GE------.-------------   113
SMM.US.x.PGM53         ---.-------------------------------A--------------R------R--M---E--T------------V---R-------E------.-S------------   113
SMM.US.x.SIVsmH635F_L3 ---.---------------E----------G---A-----------------------M---E--T------------------------------------.-T------------   113
SMM.SL.92.SL92B        -T-.---------------------HN---AL-QT--Q-----------R-C----V----Y----A----VQ-M---M-Q-FR---T-R----SQ---.-T------------   113
STM.US.x.STM           ---.-------------A-E------H----D-----------------R-----------PG-----------------S-R--------------.-------------   113
MNE.US.x.MNE027        ---.------------------------------------------------------------------------------------------------.-------------   113
DRL.x.x.FAO            -AERQSVERA-AEPMGAGEV-LE---Q-SLLR--Q--RL-FHP-FL-RL-NTCM-HY--ALQL-FT-S-----L-L---M---FQQ--S--QGR-PPPLRPAGDRL-----PP.-   114
RCM.NG.x.NG411         -AE.G---V-EAPT-AGDVEF-PWLHRMLT-VNLEA.RL-FHP-F--RL-RTCV-H---RL-R-LE-AG----L-M-----I--QS--SQR...--Q-QAREA-ERIQIL--M.-   109
RCM.GA.x.GAB1          -AE....-A-EVPT-AGEAEFQPWLRDMLEKVNLEA.RL-FHP-F--RL-RTCV-H---VHQR-LE-AA----L-M-----I--QT--SQR...--PNPR.AV-ERITIL--M.-   105
MND_2.x.x.5440         -AE....-A-EAPE-AGE.VGLEQ--ETSL-R-----RL-FHP-FL-RL-NTCV-H---RHQR-LD-A-----L-MH--MYT-MQQ--P-RNG.RPR......-...-----M--   100
MND_2.CM.98.CM16       -AE....GA-EIPE-AGE.VDLNT--E-SL-K--Q--RL-FHP-FL-RL-NACI-H---RHQR-L--A-----L-MN--M-T-MQQE-P-RSG.-P.....-.-...-----MV-   100
MND_2.GA.x.M14         -AE....-A-EAPQ-AGE.VGLEQ--E-SL-QV----QL-FHP-FL-RL-NTCV-H---RLRRTLN-A-----L-M---M-V-MQQ--P-RSG.-S.....-.-...--------   100
```

```
                                                                                                                               Vpr end in HXB2
                                                                  amphipathic α-helix                                       frameshift in HXB2
                               Vpr start                oligomerization                                                     H(S/N)RIG motifs                                          Vpr end
H1B.FR.83.HXB2                 MEQA.........................PEDQGPQREPHNEWTLELLEELKNEAVRHFPRIWLHGLGQHIYETYGD........TWAGVEAIIRILQQLLFIH#FRIGCRHSRIGVTR.......................................QRRARNGASRS*   96
H1A1.UG.85.U455                ----.........................----------YA-*A--------H-D------Q----------N----........--E----------------.-----Q-----IIP.......................................G--G--------   96
H1B.US.90.WEAU160              ----.........................----------Y------------S--------V---S-----------........--T----------------.-----------I-P.......................................------------   97
H1C.ET.86.ETH2220              ----.........................----------Y---A--------Q--------P---N---Y-------........--S----L--T----M---.-----Q-----IL-.......................................------------   97
H1K.CM.96.MP535                ----.........................----------N------I-----R--------P---N------T----........--E-L--------------.-----H-----IIP.......................................---G---S----   97
H1D.CD.84.84ZR085              ----.........................-A--------Y------------S------------S---Y-------........-------------------.-----Q-----I-P.......................................-------S----   97
H1F1.BE.93.VI850               ----.........................-G--------Y---A--I--------------P----------N----........--E----------------.-----------IVP.......................................---V--------   97
H1G.SE.93.SE6165               ----.........................----------Y---A-----------------L----------N----........--E----------------.-----Q-----I-P.......................................R--V-D-PG---   97
H1H.CF.90.056                  ----.........................---------------------I----------V---Q------N----........--V----L--T--------.-----Q-----I--.......................................---V---P----   97
H1J.SE.93.SE7887               ----.........................---H------YH--------------------P---S---Y--S----........--E----------------.-----H-----IIP.......................................---G--------   97
H101_AE.TH.90.CM240            ----.........................--#f------Y---------------------P--Q----Y--NN---........--E-------M------V-.-----Q-----IMP.......................................G--G---TG---   95
H102_AG.NG.x.IBNG              --R-.........................----------S------------H--------P---------------........--E--K-----------V-.-----Q-----IIQ.......................................G--G----G---   97
H103_AB.RU.97.KAL153_2         ----.........................----------Y------------S--------V---S---Y-------........-------------------.-----Q-----IIQ.......................................R-----------   97
H104_cpx.CY.94.CY032           ----.........................----------N---------------------P----------N----........--E----------------.-----Q-----I-P.....................................#KGEE-GD------   98
H1N.CM.95.YBF30                --R-.........................---A------Y---A----------------------------N----........--E----------------.Y----Q-----I-P.......................................-.-R---T----   96
H1O.CM.91.MVP5180              ----.........................L-N---A---F------------E--------P--QAC--Y-------........--E--M-----------T-.Y----Q-----ILP...................................SNTRG-GR---S----  101
H1O.BE.87.ANT70                ----.........................--N---AK--F---A--------A--------P---A---Y-------........--V--M-----------T-.Y----Q-----INP......................................RG-GR---S----   98
CPZ.CM.98.CAM3                 ----.........................------P---Y---A-----------------A-------F--D----........-------------------.--L--Q-----IIP.......................................G.-R--------   96
CPZ.CM.05.SIVcpzLB7            ---V.........................----------Y---A--V-----Q--------P--Q----Y--D----........--E--Q-------------.Y----Q-----I-P.......................................---R-D--N---   97
CPZ.CM.05.SIVcpzMB66           ---L.........................----------Y------I-----R--------D--Q-------T----........-------------------.Y----H-----INI.....................................NQ---R----N---   99
CPZ.CM.05.SIVcpzMT145          ----.........................------S---YEQ-L-DT---I-Q--------P---A--E--------........S----Q-----------V-.-----H-----INP.......................................---R-D-TN-A-   97
CPZ.CM.05.SIVcpzEK505          ----.........................----------Y---A-----D-----L-----P-------YF-N----........--E-------T--------.Y----Q-----I-P.......................................-.-R--------   96
CPZ.GA.88.GAB1                 ----.........................------P---YQ--A--T--------------P---Q---F--D----........--V----------H-----.--L--Q-----ILP.......................................---RS--SN---   97
CPZ.GA.88.GAB2                 ----.........................------AG--Y---A--I--D-----------P---Q---L--NN---........--E----------H---V-.-----Q-----IE-......................................QR--R-D-SY---   98
CPZ.CM.01.SIVcpzCAM13          --H-.........................------P---Y---A--I--D------K-Y--S---Q------T----........--E--------------V-.--V--H-----IIP.......................................---RA---N-P-   97
CPZ.US.85.CPZUS                ---V.........................------A---Y---A-DV-----E--L-----P--Q----Y--S----........--E----------------.Y----N-----ISL..................................ARRTP-.GR---S----  101
CPZ.CM.98.CAM5                 ---I.........................------P---Y---X--I-----T--L-----P---R---Y-------........--I----------H---V-.--L--Q-----ILX.......................................E.-R---T--X-   96
CPZ.CD.90.ANT                  ----.........................-Q-E------M---L--T---I---------QPT-QH--NWV-AN---........S-R--QVL-IL--KA--T-.Y-H--A-----HGP...............................................R-R-   89
CPZ.TZ.00.TAN1                 ----.........................-N-N------YT--L-DI---I-Q---K----PI-Q-V-NWVFTI---........S-E--QEL-K---RA--T-.Y-H--I-----S....................................................-   84
RCM.NG.x.NG411                 -.........................EML---EE-----YD--LMDT-I-IQE--KK--TPEL-TQE-NY---QH--........SLE--K-M-TL-NKA--L-.--H--EG----GA-.................................GG.NPLR.SILHSRNIL-  101
                               Vpr start                                                                                                                                                      Vpr end

MAC.US.x.239                   -........................EERP--NE------WD--VV-V-----E--LK--DPRL-TA--N---NRH--........-LE-AGEL-----RA--M-.--G--I-----QPG.................................GG..NPLSAIPPSR-ML-  102
H2A.GW.x.ALI                   -....................TEAPTEFP-AGM--HQGARD--VI-V-R-I-E--L---DPRM-IA--GY--TRH--........-LERAREL-NA--RA--M-.--A--GR--V-Q--.................................GR..NPLSAIPTPRNMQ-  106
H2A.DE.x.BEN                   -....................TEAPTEFP---GT-R-DLGSD-VI-T-R-I-E--L---DPRL-IA--YY-HNRH--........-LE-AREL-KT--RA--V-.--A--NR----*--.................................RR..TPCPA-PTPRAMH-  105
H2A.SN.x.ST                    -....................TEAPTESP---RT-P---GD--VI-T-R-I-*--LK--DPRL-IT--NY--ARH--........-LE-ARGL-----RA-LL-.--A--GR----QP-.................................GR..NPLSAIPTPRGMR-  105
H2B.CI.x.EHO                   -....................AEAVPEIP--KN------WEQ-VVDV---I-Q--LK--DPRL-TA--NF--NRH-N........-LE-AGEL-KL--RA--L-.--G--Q-----QPG.................................GG..NPLSAIPPS*GMQ-  105
H2B.GH.86.D205                 -....................AEAAPEIP-NEN------WE--IG-I---I-Q--LK--DPRL-TA--NF--SRH--.........L--AGEL-K---RA--L-.--A--Q-----QSG.................................GG..NPLSTIPPP*GMR-  105
H2G.CI.92.ABT96                -.......................AEEIP---EA-----WD--VV-V---I-E--LK--DPRL-TA--NY--DRH--........-LE-AGEL-K---RA--L-.--GD--C----GNA.................................GG..NPLSTIPPSRGVF-  103
H2U.FR.96.12034                -........................AEIP---GA-----GD--VR-I-----E------DPRL-TA--NY--NRH--........-LE-AGEL-K---RAI-L-.--G--------SS-.................................GG..NPLTTVPPS*GVL-  101
MAC.US.x.251_BK28              -........................EERP--NE------WD--VV-V-----E--LK--DPRL-TA--N---NRH--........-LE-AGEL-----RA--M-.--G--N-----QPG.................................GG..NPLSTIPPS*GVL-  101
MAC.US.x.251_1A11              -........................EERP--NE--*---WG--VV-V-----EG-LK--DPRL-TA--N---NRH--........-LE-AGEL-----RA--M-.--G--I-----QPG.................................GR..NPLSAIPPSR-ML-  101
SMM.US.x.H9                    -........................TERP---EA-----WD--XV-V---X-E--LNX-DPRL-XA--NY--DRH--........-LE-AGEL-----GA----.--G--------QSG.................................GX..NPLSTIPPXRGVL-  102
SMM.US.x.PGM53                 -........................TERP---EA-----WD--VVKV---V-E--LK--DPRL-TA--NY--DRH--........-LE-AGEL-----RA----.--G--A-----QS-.................................GG..NPLSTIPPSRGVL-  102
SMM.US.x.SIVsmH635F_L3         -........................AERP---EA-----WD--VV-V---I-E--LK--DPRL-TA--NY--DRH--........-LE-AGEL-----RA----.--S--A-----QS-.................................GG..NPLSTIPPSRAML-  102
MNE.US.x.MNE027                -........................EERP---E------WD--VV-V-----E--LK--DPRL-TA--NY--DRH--........-LE-AGEL-K---RA--M-.--G--N-----QSG.................................GG..NPLSTIPPSR-ML-  102
STM.US.x.STM                   -........................THRP---E------WD--VV-V---I-Q--L---DPRL-SA--NY--DRH--........-LE-AGEL-K---RA--M-.--G--------QPG.................................GG..NPLATIPPTRGVL-  102
MND_1.GA.x.MNDGB1              -....................GQKRDEQVS-----P---Y-Q-LADTM--I-E--RK---L-I-NAVSEYCVQNT-SEE........EAC-KF-TLMNRAIWV-.LAQ--DGTF.RER..................................RPQLPPS.GF-PRGD-L-  105
MND_2.CM.98.CM16               -.........................EQP---EA-----Y---LIDT-A-VQE--LK--D-RL--AV-SW--K----........-LE--RQL-S---KA-SM-.Y-Y---E----QEG................................GGC.YPLR.SFPRSDNPL-  102
MND_2.GA.x.M14                 -.........................EQP---EA-----Y---LIDT-A-IQE--LK--D-RL--AV-SW-------........-LE--QKL-TL--RA--L-.--H---E----QEG.................................GKYNPLR.SFPRPNNPL-  102
GRV.ET.x.GRI_677               -ASGRDPRE...............PLPGWL-IWDLD---WD--LQDM-RD-NE--R---GMNM-IRVWNYCV-EGRR....HNTPWNEIGYKYY--V-KSM-V-.--C---RRGPFSPY..........................EERRNGQGGGA.....PPPPPGLA-  119
VER.KE.x.TYO1_patent           -ASGR...............DPREARPGEV-IWDLS---WD--LRDM-QD-NQ--RL--G-EL-FQVWNYCQ-EGER....HGTPMMERAYKYY-LV-KA--V-.--C---RRQPFEPY..........................EERRDGQGGGR....AN-VPPGLE-  120
VER.KE.x.AGM155                -ASGR...............DPREERPGGL-IWDLS---WD--LRDMV--IN---KL--G-EL-YQVWNYCQ-EGER....QGRPIAERAYKYY-LV-KA--V-.--C---RRQPFEPY..........................EERRNGQGGGR....PG-VPPGLD-  120
VER.KE.x.9063                  -ASGR...............GPRENRPGEV-IWDLN---WD--LRDM--D-NQ--KL--G-EL-FQVWNYCQ-EGER....RGAPMTERAYRYY-LV-KA--V-.--C---RRQPFEPY..........................EERRDGQGGGR....AN-APPGLD-  120
SAB.SN.x.SAB1C                 -ASGGWLPPVGGDPPKDPPKNPREEIPGWL-TWDLP---FD--LRDM-QD-NS--QC----NL-FR-WWN-V-EPAI...DHGQTRLE-WYKYC----KA--V-.MKGR-CKPKTHPAY......................GPGAGGPPPGLGGASGG..A-SAAPGL.-  141
TAN.UG.x.TAN1                  -AEGR...............DSRERRPGWL-IWDLS---WD-*LRDMVA--NQ--Q---G-EL-FQVWNFCQ-EGER....NGAPMIERAYRYY-LV-KA--V-.--C---RRTPFEPY..........................EERRNGVGGGR...DGREPPPGLA-  120
VER.DE.x.AGM3                  -ASGR...............DPREARPGEL-IWDLS---WD--LRDM--DINQ--KM--G-EL-FQVWNYCQ-EGER....NRTPMLERAYKYYKLV-KA--V-.--C---RRQPFEPY..........................EERRDGQGGGR....AG-VPPGLD-  120
SYK.KE.x.KE51                  -A--F..........FRVYQRQPGEPLFYIPRNVNWD-AGVISATAKA-QVAQ--CK--TAEEIF-VWNQCL-VEAGPEETPTMAWIRCMLDMH-A-NFM--E-.-AA--PQRTKYARH..................................RGYPHPS.........-  116
LST.KE.x.lho7                  -SRQR................QQRSPERP---F--P---WGQ-LADTM--I-E--R----EVM-AQVS-YCVDST-SEE........EACMKF-TLVNRA--L-.LP.T-PSYA.AAR............................AGTAHPTEAAPRPNEGDIQRGSD-  115
LST.CD.88.524                  -SRQR................QQRSPERP---F--P---WAQ-LADTM--I-E--R----EAI-AQVSEYCVSST-SEE........EACMRF-TLVNRA-YL-..LPQ-PSYAAAR.............................AGTAHPTEAAPRPTGRDIQRGHY-  115
WRC.CI.97.97CI14               -.....................SNS.REQAPSG--PPI-E...D-LGMQ--FAM-KQ----EV-VRVFLYCSQQREYEE.......VPTMHYFH-V--RG--M-.LTQ--PKQQEKREQ.............Q.......FQRAEQLYRNLERTAGDQGEVEQRDLISL-  117
WRC.GM.05.Pbt_05GM_X02         -.....................SNSGRER-PTG---AI-E...D-LGMQ--FAM-K-----EV-VRVFLYCSQQRNYEE.......Q-TMHYFH-V--RG-YL-.LIQ--PKEKENRE-.............D.......FQRAEQFYRNLERASGDQGGVEQRDLISL-  118
WRC.CI.98.98CI04               -.....................SNS.REQAPTG--PPT-E...D-LGMQ--FAM-KL----EV-VRVFLYCSQQRNYEE.......GPTMHYFH-V--RG--M-.LTQ--PKQQEKREQ.............Q.......FQRAEQFFRNLERAAGDQGEVEQRDLISL-  117
OLC.CI.97.97CI12               -........................AERV------P---LED-VRDTM--IIV--K-----EL-VRVFLFCAQRHINDE.......MSTSLLF--L--RA--Q-.Y-N--PGVELRRNP.........................................HR-SEGMQE-   97
COL.CM.x.CGU1                  -.................................RVLKGTS#-LQDQV-NITR--R----LTHISVIMWGCTRRHWGR.......P-TAGL-YV--MECA-YL-VHAGQ-...........................................RFPST-NP........-   78
DEN.CD.x.CD1                   -A...................................-TTQRQLQ-AVWDITE--RK--SKEEIT-IWD-CWSLPALPHWTEGQVMA-A-IDF---M-KEIWK-.Y-V--F-REAERV-...........HYPNIRPLRP.................RR...EEP....-   99
DEB.CM.99.CM5                  --RYPPSHP....................-HFTSRTVPMTRLALQQAMQD-NE--LK--T-EE-W-VWN-CVDLPAQPDWTGEQAWA-S-IDY-K-V-RM-WL-.L-EA-F-REREA--...........RYPNIRPLTG.................RN...-EVRDGE-  118
DEB.CM.99.CM40                 --RYPPSHP....................-HFTSRTIPMTRLGLQQAMQD-NE--LK--T-EE-F-VWN-CVDLPAQPDWTGEQAWA-S-IDY-K-V-KM-WL-.L-EA-F-REREA--...........GYPNIRPLTG.................RN...-EVRDGE-  118
TAL.CM.00.266                  --RL....................PPSHPLPLTSRLIPTPRPQLER-MR-VRD-MLI--T-DEVI-VWN-CV-LPA-PDWTGEQAWA-S-IDFA-KGN-M-LL-.-QQ--Y-EQQRR-P...........HYPNIRPLSG.................RG...ER.TMQ.-  116
GSN.CM.99.CN166                --RI....................PPSHPMPWLSRRVPTTMAIAQNA-W-INE--EK--S-DE-R-IWHDVT-LPA-PDWTVDQAAI-CAIDY--.V-T---R-.-KD--F-RVNLS.G...........RYPAIRPSRGTAPPDSNSVSHADPEQPRR..PSRYRMDE-  134
GSN.CM.99.CN71                 --RT....................PPSHPLPWISRRVPTTMAVAQSAMW-INE--EK--S-EE-R-IWHDVT-LPA-PDWTVDQAAI-CAIDYV-RV-T---R-.--D--F-RYNRIV-...........RYPVIRPLRGTAPPDSNSVPHADPEQPRR..PSRYRMDE-  136
MUS_1.CM.01.CM1239             --RV....................PPSHR-PWHSRVVPTTMQQAQQAMWD-NE--EK--S-EE-R-IWNDVT-LPA-PNWTVDQAAI-CAIDY--RT-T---R-.Y-E--Y-RYSNTI-...........RYPNIRPLRGTQAPPSNSMPNADPTPPLR..PSRYRMDE-  136
MUS_1.CM.01.1085               --SL....................PPSHPAPWHSRVVPTTMRQAQ-AMW-VNE--EK--S-EE-R-IWNEVT-LPA-PDWTVDQAAI-CAIDYV-RV-T-V-R-.--D--Y-RYNR-M-...........RYPTIRPLRGTQPPPSNSTPNADPVPPLS..PSRYRMDE-  136
MUS_2.CM.01.CM2500             ---H....................PPSHPLPWLSR-VPTTMAQAQNAMWD-NE--EK--S-DE-R-IWDRVT-LPA-PNWTVDQAAI-CAIDY--WV-G---R-.--E--Y-RYSQ-I-...........RYPQVRPLRGTQPPDSNSTPNVDPFPPRR..PSRYRMDE-  136
MUS_2.CM.01.CM1246             ----....................PPSHPLSWLSR-TPTTMVQAQAA-W--NE--EK--S-DK-R-IWDRVT-LPA-PQWSVDQAAITCAIDY-KSV-T---R-.Y-D--Y--YAQTI-...........RYPPLRPLRGTQPPDSNSMPNADPTPSLR..PSRYRMDE-  136
MON.CM.99.L1                   ---PPQSHP...............LHWTSRMV.....PIERQALQAAIW--NE--LK--S-EE-R-IWEQVT-LPA-PAWNADQAWA-CAIDYT-WV-TI-YR-.Y-E--Y-RYAEQI-...........RYPVLRPMRGTAPGPTSSVPQADPDNPRR..PSRYRMDE-  136
MON.NG.x.NG1                   --N-PPSHP...............PHWTSRVV.....PIERAALQAAIWD-NE--LX--SKEE-R-IWDQVT-LPA-PGWNADQAWA-XAIDYT-RV-TXI-R-.M-D--Y-RHAR-X-...........RYPXLRPLRGTAPGPTSSVPHADPERPLH..PSRYHHDE-  136
TAL.CM.01.8023                 --RL....................PPSHPLPITSRLIPTPRAQLER-MR-VRD-MLI--T-EEVI-VWN-CV-LPAEPDWTGEQAWA-S-IDFA-KGNEM--L-.-QQ--F-EQQRR-A...........HYPNIRPLSG.......................RRNTMQ-  115
RCM.GA.x.GAB1                  -M.................#PLFTEMELP---E------YD--LMDT-I--QE--KK--TYAL-TQI-DYV--QH--........SIE--Q-M--L--RA--L-.--N--AG----TS-................................GS..NPLR.SIPQTRNIM-  108
SMM.SL.92.SL92B                -.........................EHA---ETNP---WD--IRDV-----E--LK--DPRL-TA--NYV-D----........-IE-AGE--K---KA--L-.--H--T-----QP-................................GP..NPL.GSIPSARDVL-  101
MND_2.x.x.5440                 -.........................EQP---EA---G-Y-K-LVGT-A-IQE--LK--D-RL--AV-SWV--QQ--........-LE--QQL-S---RA--L-.--Y---G----QEG................................GRY.NPL.-SFPSPNNPL-  102
DRL.x.x.FAO                    -.........................ERP---E--P---YD--LVDT-V-IQE--LK--D-RL--NV-SWV--R---........SLE--HQL-IL--RA--L-.--Y--QE--L-.PA................................GRW.NPFR.TFPRANQQL-  101
SUN.GA.98.L14                  -AS..................RREPVEQP---E--P---FAQ-LADTM--IRD--E----AAIVAQVL-YCADST-SE........KE-CMRA-TL-NRA----.LP.M-PSYIRIRSG.........................SGNVNPRPPTRPRPGQGDIRRGLQD-  117
LST.CD.88.447                  -SRTR................QQRSPEKP---F--P---WGQ-LADTM--I-E--R----ETI-AQVSEYCVSST-SEE........KACMKF-T-VNRA----.LPT-PSYAAAR............................AGTA.HPTEAAPRPTK-DIQRGRY.-  115
LST.CD.88.485                  -SR-R................QQRSPERP---F--P---WGQ-LADTM--I-E--R----EAI-AQVSEYCVSST-SEE........EACMRF-T-VNRA----.LPT-PSYAATR..............................AGTAHPTEAAPKPTERDIQRG-Y-  115
SYK.KE.x.SYK173                -AE-F...........FNPSQHVQGTPWFFIPRNVELT-NVINVTVKA-LVVT--SK--TPQEIY-VWNQSLNEEAGTD.SPTMAWERTMLDMV-A-NLM--E-.-AA--PQRTRYARH...........................................-GYPHP--  114
```

```
                          Tat start                                                                disulfide bonding              NLS          exon 1 end exon 2 start
H1B.FR.83.HXB2            M.............E...PVDPRLEPWKHPGSQPK..........................TAC.TNCYCKKCCFHCQVCFITKALGIS.YGRKKRRQRR....RA.HQNSQTHQASLSKQPTS.QPRGDPTGPKE................................  86
H1A1.UG.85.U455           -.............-.......----N-----------T..........................---.S-----V--W---L--LK-G---.--K---KP--....GP.P-G-KD--TLIP---LPQSQ-.VSA-QE-SKKKVESKAKTDRFA..................  101
H1B.US.90.WEAU160         -.............-.....---H-------------..........................---.------R---------M--G---.--------------..-S.P-------D--------.-----------SKKKVERETETDPED..................  101
H1C.ET.86.ETH2220         -.............-.......----N----N-------..........................---.NQ------SY--L---L#-G---.----------------..--.P-S-KD--NLI----L-.HT------SE-SKKKVESKAETDPYA..................  100
H1D.CD.84.84ZR085         -.............D.......----NID--N------R..........................---.N--------Y--------G---.--------------...-P.PHS-----DPIP---S-.--------Q--..................................  86
H1F1.BE.93.VI850          -.............-.......----S-D--N------T..........................-P-.-K----R------W--T--G---.----------H....-T.P-S--V--N--P---L-.-A---------SKKEVESKAKTDPCA..................  101
H1G.SE.93.SE6165          -.............D.......----N----N-------..........................-P-.NK-F--V--W------LN-G---.------KH--...GT.P-S-KG--DPVP---LP.TT--N------SKKEVASKAEADQCD..................  101
H1H.CF.90.056             -.............D.......----K----N------Q..........................---.N--------Y--M--LK-G---.------S--H...-T.PASL-D--N-I----L-.RTH--------QKKEVASKTETDP....................  99
H1J.SE.93.SE7887          -.............-.......----NR---N-------..........................---.---------Y-----LQ-G---.--------------...S-.PPG-K---DLIP---L-.-TQRK----E-SKKEVESKAEPDRFD..................  101
H1K.CM.96.MP535           -.............D.......----NI---NQ------..........................---.NQ---R--Y---I--LK-G---.N------P----...TT.PY--EN--DP-R---L-.----EQ-D---SKKKVESKTKTDQFD..................  101
H101_AE.TH.90.CM240       -.............-...L---N----N------T..........................---.SK-------W---L--LK-G---.------KH--...GT.P-S-KD--NPIP---LP.II-RN--D---SKKEVASKAETDQCD..................  101
H102_AG.NG.x.IBNG         -.............-...L---S----N------T..........................---.SK----M--#---L--LN-G---.-------R--...GT.P-SR-D--NPVP---LP.TT--N--D---SKKEVESKTKTDPCD..................  100
H103_AB.RU.97.KAL153_2    -.............-.......-----------------..........................---.---------L------MK-G---.----------------..---.P-DN--D-V--P---A-.----------..................................  86
H104_cpx.CY.94.CY032      -.............-.......----D----N------T..........................-D-.NK-F-----W------LK-G---.------KH--...GS.L-G-KG--NLIP---L-Q--N--S---E-QKKKVASKTEADPFA..................  102
H1N.CM.95.YBF30           -.............-.......---------N-------..........................---.N----R--Y--LY--TK-G---.------S---...-T.P-S-KS--DLIPE--L-.-QQ--Q--Q-KQKEALESKTEADPCD..................  101
H1O.BE.87.ANT70           -.............D.......----EVP--H------Q..........................IP-.N----R--Y--Y---VR-G---.------GRP....A-..ASHPD-KDPVP--SPT.ITKRKQERQE-QEEEVEKKAGPGGYPRRKGSCHCCTRTSEQ...  114
H1O.CM.91.MVP5180         -.............D.......----EMP--H----K-Q..........................-P-.N----R--Y--Y---TK-G----.H------RPA....A-..ASYPDNKDPVPE-SL-.HTGRKQKRQE-QEKKVEKETGPSGQPCHQDSCNSCTRISGQ...  115
CPZ.CD.90.ANT             -.............D.......---AETP--L--PAT-A..........................-P-.N----C--Y--PL--TK-G---.----R-AR-...NR-TTAES-ENN-DPV--..........................................  73
CPZ.CM.01.SIVcpzCAM13     -.............-....-I--E--------------..........................---.N----A-----MY--TK-G---.-------.A--...-TAFAS-KNN-DPVR---L-.----NQERSEKQAEKVESQAAADP..................  99
CPZ.CM.05.SIVcpzEK505     -.............-.......----E----N-------..........................---.N--H--V--Y--VY--TK-G---.------S---...-T.P-SNKS--DP-P---L-.-RL--Q--Q--QKKTLES..................  93
CPZ.CM.05.SIVcpzMB66      -.............-....-I--N----N-------..........................---.N----Q--Y---L--TK-G---.---R--K----...-T.SES--N--DPVP---L-.--G-IE--Q-KSKKEVESQTTSDQFA..................  101
CPZ.CM.98.CAM3            -.............-.......----D----L------R..........................---.N----M-----I---EK-G---.--------G--...-LSAAS-THN-G--QQ--L-.RQQ-NQ--QE-QKKTLESQAKADPFA..................  102
CPZ.GA.88.GAB2            -.............D...-I--S----N-------..........................---.NS---R--Y--ML--TS-G---.-------.G--...-TALKG-KSN-DP-RQ--L-.-QQ--QKCQ--SEEEVESKTEADL..................  99
CPZ.CM.98.CAM5            -.............-......---XN----L-------..........................---.N----C----M---EK-G---.--------G--...HSLASS-THN-D-IRQ--L-.RQQ-NQ--EE-QKKTVESQAEADPRDWSAGS..................  107
CPZ.CM.05.SIVcpzMT145     -.............-....-IE-D----N----K-R..........................-E-.NK-F-----Y--IL--TK-G---.------GR-K...P-..SSN-DNKDP-R---L-.--IR-Q--SEKQKKEVERTSEADQSDW..................  101
CPZ.TZ.00.TAN1            -.............N...-I--QVA--E---AA-E..........................-P-.------------P---TK--------.---R-GRKS...AV..HST.NN-DPVRQ-SLP.KRSRIQNSQE-SQEEVEAETTSGGRPRQQDSSVSSGRTSGTSSS  117
CPZ.CM.05.SIVcpzLB7       -.............-....-I--S-D--N-------..........................-P-.N--F-----Y--PF--TK-G---.-----W-S--...-T.SKS-PHY-DPI----FP.PS---SLS-E-QKKKVESKAQTDQ..................  99
CPZ.US.85.CPZUS           -.............D.......----NI----Q-----..........................---.N-----------V---TK-G---.------GR--...-TTAKS-ADN-DP-R---F-.-QL-NQ--Q--QKKALEKQTTSDQRD..................  102
CPZ.GA.88.GAB1            -.............D...-I--D----------R..........................-V-.N----A--Y--IY--TK-G---.------TT--...-TAPAG-KNN-D-IP---L-.-S--NKE-SEKSTKEVASKTEADQ..................  100
MUS_1.CM.01.1085          -.............D...-SVEE-PKEQR--AA-A..........................-P-.-S-F---A---VL--QK------.---RF-KRS-.....ADSSLPS-SDPVQIKSL-.DR-R-QKRQ--QKETVEAEAETDP..................  98
MUS_1.CM.01.CM1239        -.............D...-SVEN-P-EQR--AA-T..........................-P-.SS----R-A---ML--QK------.---RF-KRP-.....SD--V-N-PNPVPIKSIP.NLQRQQSRQ--QKEALAAAAGPDPCDI..................  101
MUS_2.CM.01.CM1246        -.............D...-SVEG-P-DQR--AA-P..........................-P-.------C-A---IL--QK------.---RI-KRP-.....TDSLLPSD-NPVRF-SIP.DSQQNKSNK-KQKKALEEAAGTDPCRCFQKTSTT..................  108
MUS_2.CM.01.CM2500        -.............D...-EVEN-P--QR-S-A-T..........................-P-.-P----A-AY--MQ--L-------.---NRK-H--.....TPEGLPNN-DPIQLKSIP.E--R-QER--KQKKEVAPEAVSDPCAI..................  101
                          Tat start                                                                                                   exon 1 end exon 2 start        Tat end
MAC.US.x.239              -.............-TPLREQENSLESSNER-SCISEADASTPESANLGEEILSQLYRP..LE--YNT-------Y---F--LK-G---C.-EQSRK-R-.......TPKKAKANTS-A-NK-I-.NRTRHCQPE-AKKETVEKAVATAPGLGR................  130
H2A.GW.x.ALI              -.............-TPLKEPGSSLMPYNEP-SCTSEQDVAVQELAKQGEEILSQLYRP..LET-NNT----E--Y---L--LN-G---W.-D--G--R-.......SPKKIKA-SS-A-DKSI-.TRTRNSQPEEKQKKTLETTLGTDCGPGRSHIYIS..........  136
H2A.DE.x.BEN              -.............-TPLKAPESSLKPYNEP-SCTSERDVTAQELAKQGEELLAQLHRP..LEP-TNK---R-S----L--SK-G---.-E--G--R-.......TPRKTK-PSP-APDKSI-.TRT--SQPT--QKKTSEATVVTTCGLGQ................  130
H2A.SN.x.ST               -.............-TPLKAPEGSLGSYNEP-SCTSEQDAAAQGLVSPGDEILYQLYQP..LE--DNK-------Y--M--LN-G---W.-E--G--R-.......TPKKTKA-SS-A-DKSI-.TRT-NSQPE-KQKKTLETALETIGGPGR................  130
H2B.GH.86.D205            -.............-IPLQEQESSLKSSNEP-SSTSEPVVNTQGLDNQGEEILSQLYRP..LK--DNT-------Y---L--LK-G---C.-D-SR--.........SAKRAK-TAP-APDKSL-.ART--SQPT-KQKKEVETTGTTDPGPGR................  128
H2B.CI.x.EHO              -.............-IPLKEQESSLNSSSGH-SSTSEGVANTQGLDNRGEEILSQLYRP..LK--SNT------SY---L--LK-G---C.-E-SRK-.........SSKRAK-TTS-APNESL-.ART--SQPT-KQKKEVETTRATDLGPGRSNTSTSRFAN......  138
H2G.CI.92.ABT96           -.............-TPSREQENSLKSCREL-SCTFEETVDAPGLEGTQAEILYQLYRP..LE--SNK-------Y---L--LN-G--VC.-EQPR--R.......TPKKTKANSF-A-NESI-.ARTRNXQPK-K..................................  112
H2U.FR.96.12034           -.............-IPLKEQESSLKSCKEQ-SSTSEGDVGTQGSAAAGEEILSQLYRP..LE--SNK-------Y---L--LK-G---C.-EQPR--.........SPKKVKAYSSAP-NKSL-.NGTRNSQPKEK..................................  111
MAC.US.x.251_1A11         -.............-TPLREQENSLESSSGH-SCTSEAAASTPELANLGEEILSQLYRP..LE--YNT-------Y---F--LK-G---C.-EQSRK-R-.......TPKKAKANTS-A-NK-I-.NRTRHCQPE-AKKETVEKAVATAPGLGR................  130
MAC.US.x.251_BK28         -.............-TPLREQENSLESSNER-SCILEADATTPESANLGEEILSQLYRP..LE--YNT-------Y---F--LK-G---C.-EQSRK-R-.......TPKKAKANTS-A-NKLIP.NRTRHCQPE-AKKETVEKAVATAPGLGR................  130
MAC.US.x.EMBL_3           ...........................#SNER-SCILEADATTPESANLGEEILSQLYRP..LE--YNT-------Y---F--LK-G---C.-EQSRK-R-.......TPKKAKANTS-A-NKLIP.NRTRHCQPE-AKKETVEKAVATAPGLGR................  116
SMM.US.x.H9               -.............-TPLKEQESSLESSREH-SSISEVDADTPESASLEXEILSQLYRX..VE--YNKX------Y---H--XK-G---C.-EQQR--.........TPKKTKANTF-A-DKSLX.XRARNCQPK--KKETVEXEVATDLGLXR................  128
SMM.US.x.PGM53            -.............-TPLKEQENSLGSSNER-SYTSEEDVLTPESAKVEEEILCQLYRP..LE--SNT-------Y---L--LK-G---C.-EQQR--R-.......TPKKAKANTF-APNESL-.TRARNSQPK--KKETVETEVATDLGLGR................  130
SMM.US.x.SIVsmH635F_L3    -.............-TPLKEQENSLESYREH-SSISEVDVPTPESANLEEEILSQLYRP..LEP-YNK----R--Y---H--LK-G---C.-EQHR--.........TPKKTK-NPLPA-NKSL-.TRTRNRQPK--KKEKVETEVAADLGLGR................  128
SMM.SL.92.SL92B           T.............GIQ*KGQGK*LRYYKRHCSYTLGMDAPIQGLDSQGDQIPWDLYRP..LET-YNS-F---------L--LR-G---T.-AKPR--.........VKKKAKANPFDT-N-SI-.SGTRNCQPK--QKEKVETEVDTDCGLGR................  126
STM.US.x.STM              -.............-TPLKEQESSLRSSSEP-SCTSEAVAATPGLANQEEEILWQLYRP..LEE-CNK-F-----Y---L--V--G---T.-E-SR--.........VKKKAK-YPI-A-NKSL-.TRARNSQPK--QKKEVETEVESTPGLGK................  128
MNE.US.x.MNE027           -.............-TPLREQENSLKSSNGR-SCTSEAAATTLESANLEEEILSQLYRP..LE--YNT-------Y---F--LK-G---C.-EQSR--R-.......TPKKVKANTS-A-NK-I-.DRTKHCQPK--QKETVETAVATVPGLGR................  130
GRV.ET.x.GRI_677          -.............D..K.EEE....PHPLLQDL......................HRP..LQP-TNK-------Y--EL--LQ-G--VR.-HVSRK-R.........KTST-DN-DPIRQ-SI-TV.QRNGQTTE-GKTEVEKAAAAN........................  94
VER.KE.x.TYO1_patent      -.............D..KGEAEQIVSHQDLSEDY......................QKP..LQT-KNK-F-----Y---L--LQ-G--VT.-HAPRT-R.........KKIASLDL-P-QH-SI-TKWGR-GQTTPTSQEKVETTAGSN......................  100
VER.KE.x.9063             -.............D..KGEEE-TVLHQDLIR-Y......................KKP..L-T-RNK-F-----Y---L--LQ-G--VT.-HAPRT-R.........KKSV-PNRL-QQD-SI-TR.GR-GQATQ-SQKKVERETTTAQILGRKDLERDKREAVGANA.  119
VER.KE.x.AGM155           -.............D..KGEEDQDVSHQDLIK-Y......................RKP..LET-TNK-F-----Y---F--LR-G---T.-HAFRT-R.........KKIASADRIPVPQ-SI-IR.GR-SQTTQ-SQKKVEEQAKANLRISRKNLGDETRGPVGAGN.  119
VER.DE.x.AGM3             -.............D..KGE-EQGAYHQDLIE-L......................KAP..LKR-TNK----C--Y---L--LQ-G--VT.-HAPRI-R.........KKIAPLDRFPEQ--SI-TR.GR-SQTTQKGQEKVETSARTAPSLGRKNLAQQSGRATGASD.  119
TAN.UG.x.TAN1             -.............-..S...EGDGMAESLLQDL......................HRP..L-P-TNK-F--R--Y------LQ-G---T.-HVSRI-RP........KK-HSN--NLV-Q-SI-AW.G-NSQTTQ-EKTKIPAAAETSRRPQ................  101
SAB.SN.x.SAB1C            -.............D.QEQEARPQVWEEL.QEEL......................HRP..LQ--DNT-F--V-----IL--HK-----R.-YVPRP-R........-SKK..IS-NQVSLHN..........................................  73
OLC.CI.97.97CI12          -.............D.....AQEWSLEETLTEDL-..........................E-KNK-------Y---L--LQ-G----.-STESP-K--.....KKVVSH-QDSRDT-HFQIP.SRTRNSQTTEKEKTASSSRAPSTSGTVPQSVQGH..........  107
DRL.x.x.FAO               -.............DARKVDLDQQDAG.THFEP.FQ....................GRT...NN-.NT-W--Y-AY---L--LQ-G---R.-NVSRP-K-T...ATVVSRVSETGDTPVARTHTA-R-PTT-DSQTEKKEKSKTEASGTSDQNPPSQDPPKLGAKKQ..  124
RCM.NG.x.NG411            -.............D.......................VKGAGLEEREEEILLDPFFTRE-S-.NS----R-VY---L--LQ-----N.-ASRR--RG-P...GNKKKN-INPFDT-E-SL-.AIGRNSQSK--KKEAVEKAVDTNITTGREDIFIS..........  115
RCM.GA.x.GAB1             -.............D.......................VQGVGLEHPEEVILYDPFRKRE-S-.NT-------Y---L--LQ-G---N.-ASRA--R-......SKEENKADKFPVPN-SI-.TT--NRKLQEKKEKTVEKKVATSTTIG................  105
MND_1.GA.x.MNDGB1         -.............-.......-SGKEDHNCPP-DSGQEEIDYKQLLEEYYQPLQ.........--ENK-W-------ML--QK-G---R.-HVYRK-VP.....GTNKKIPGSGEEAIRR-GYR.PVFSQDCQQNLYRKRPDNGEEESNSLELGEHNSENLKEESLNRS  135
MND_2.x.x.5440            -.............D...................VGEVASDKKEEDITHFDPFRART...-P-.NK-F--V--Y---L--LQ----VH.-HVYRQ-RP-R...GLLEKVPSNSQAALGRDIALFLMSSRETISGPTTSNGQTEKEEKSQT................  110
MND_2.CM.98.CM16          -.............D...................AGKAVSDKKEGDVTPYDPFRDRT...-P-.DT----R--Y---L--LQ----VH.-HAYRN-RS-Q...RLLEKISEDS-APVGRVIALFLMSSRETISVPTTSNSQAKKEKKSPTQKNRVPNQTAPK.....  122
MND_2.GA.x.M14            -.............D...................AGKAELDKKEANITPLDPFQGRT...-P-.NK-------Y---C--LQ-----H.-HVYRT-RT-K...RLLGKISQDSEAAVGRVIALFLMS........................................  87
GSN.CM.99.CN166           -.............PILSSLGDPQDIAWMNMNRDPEECLLPNWQQPGAAPA..........-P-.SA------AL---L--LR-G--L-.LHGRS-KRK-PK...-TEPDPSVSSNQD-A..........................................  92
GSN.CM.99.CN71            -.............N................RDPAECLLPKWQQPGAAPN..........-P-.SA----R-A---L--LR-G--L-.LHGRS-KRK-SE...-SEPAPSVPTNQDPI-IQ-LP.HSKWNAIRQTEQEKTVARKATSGPSDISQDSRVSCGTT....  119
DEB.CM.99.CM40            -.............-.......................SIDPFEERAN............-PH.PA-M--A-IL---L---Q-G-----.--PRR-KRKCRREYTPAAESN-PD-DFVPE--LP.-R-RKCFDSQKKEAEVEETAGPGGQHCREDSSVSSGRISNNC..  115
DEB.CM.99.CM5             -.............-.......................EIDPFKETAN............-PH.P--M--A-TM---L--MQ-G---Y.--SRR-KR--..GSSFAAN.DKD--EFIPE--LP.-R-WKCFNSE-KETKVEKASRAGGRHRPEDS.............  101
DEN.CD.x.CD1              -.............N.....................ADSIDPFAG.NK............-P-.NT----H-AY---L--LQ-G---V.---AP-KRP-......AR-SGTPD-EAV---SL-.D--RQQISQ-KEKEKVETATGSSGGNIREDS.............  99
TAL.CM.01.8023            -.............S...................SKEELPTTPTYDPFPAGG.........-P-.NK-------Y---L--LQ-G---R.-VPDR-IRK-RR..STSSSN-KDKDPV----L-.RAT-NQECQ-KKVQKVESPTGAGGEHRLKDSLPLSGRTSTASSA  122
TAL.CM.00.266             -.............S...................SKEELRTTPISDPFQEEGR........GP-.NK-------Y---L--LQ-G---R.-VPDR-FRK-RR..STPSSN-EDKDPV----L-.RAT-IQKRQE-TLQKVESKAATSGTDRKQDS.............  110
MON.CM.99.L1              -............M-...................PVDPDLPKEQHPPATPR.........-P-.---F-RV------I--L--G----------RK-R-....ATSPVPGLSSSKNPAR--SL-.H-TR-SQR-T-QAQAVATAATPDRQHI................  104
MON.NG.x.NG1              -...............QTQRGLCTRPDTTTMNSAXIVNPALPPEQHPSATPL.........-P-.N--F--I-AL-X----LQ-G---------GRK-R-G...P-HPTIGDRSY-DNPGVKSLP.KSL-NQERQ-KQEKAVESKAGTGXQSMREDF.............  126
WRC.CI.97.97CI14          -RKSLLCIISTGYCKEGSSCISPKGVLNNKRKGNNNFKEQNNCIETWREQQEIREKLSKETLSP-DHR-S-VA-VYS-TL--LQ-G----.-RSYSKKTKP...........DTTTAA-SRT--TTWKR-SKTA-KE............................  125
WRC.CI.98.98CI04          -KKVPLCIIFTGYYKEGSSCTLPKGVQNNKKRENNNFKEQNNSLETWKEQQEIRERLSKETLSP-NHK-G-Y--V-S--L--LQ-G----.-RSYSKKIGA...........STAKATTSRA--TTWKR-CQTE-KKQKKTTVL........................  132
WRC.GM.05.Pbt_05GM_X02    -KNRPLCTIFTGSFKEDYIYTSYKDVP-KKRTEKEIFKEQNNFIEIWKGLQEIREELSRETLSP-NHQ-R----I-S--L--LQ-G----.-HSYSKKTNK...........DKSDASTSRVI-......................................  112
SUN.GA.98.L14             -...................STQGHQQD.........QDQGKGTLEEAYKTNLE.........--DNK-W-RR------L--LQ-G---H.-YVYRG-.......QHQSI-KDK-EAVRRIGSRKPAPNPQ-KK-KKRQR....KKESTST.................  102
LST.CD.88.524             -...................QQPEQEQHT...-Q-QHLDQLEEIYKEAITDPLQ.........S-QNK-H--V-----IL--QQ-----R.-YVHRT-R.......RVTYILE-E-DPVDRTSSGRTE-AKNSQT-KERQRKT..EKTSTST.................  111
LST.CD.88.485             -...................QQQEQEQLT...RQ-QHQNPLKEIYKEAVTKPLS.........V-DNK-C--R-----IL--QQ-G---R.-YVHR--R.......RVT-VFEDK-DPVDGTTSGRKE-TQNSQT-KER.............................  100
LST.CD.88.447             -...................QQQEQEQLT...RQ-QHQDPLKETYKEAVTKPLS.........--NNK-C--------IL--QQ-G---R.-YVHR--K.......RVA-VFEDK-DPVDGTASGRKE-TKNSQT-KER.............................  100
LST.KE.x.lho7             -...................QQPEQEQPI...-Q-QHQDLMKETYNEAVTKALQ.........--QNT-C--------IL--QQ-G--LH.-YVHRK-R.......RVTSILE-Q-DPVDRTTSGRKK-TKNSQTEKKRKEQI..QRLSTST.................  111
SYK.KE.x.SYK173           -.................................S.STDQICQTQRVPPSFLEGTFLEKGPP-P-.NK-F--N--Y---L--LQ-G---T.-A-PRK-AA-.....SISEDDSAPTGT-PRAGRTQANPQTKKKAV-TEKDSTSLPSAENL.................  110
SYK.KE.x.KE51             -.................................S.TEDKVRETQGIPTSFLEGTFLSNGPQ-P-.SR-------Y--YR--LQ-G---T.-V-ERK-RHAR....TTAEDFAA-SSTT....R-RS-LPST-R-QQGQKKE*...........................  98
COL.CM.x.CGU1             -...................SLS.....................LYAKPLE.........--NDK-W--A------L--LS-G---R.-....-KNK..........SAT..QTQ-VA-TGSG....NSQ-TQ-KYTES....KKTAAADFGD*..............  79
```

H1B.FR.83.HXB2	.....*	86
H1A1.UG.85.U455	.....-	102
H1B.US.90.WEAU160	.....-	102
H1C.ET.86.ETH2220	.....-	101
H1D.CD.84.84ZR085	.....-	87
H1F1.BE.93.VI850	.....-	102
H1G.SE.93.SE6165	.....-	102
H1H.CF.90.056	.....-	100
H1J.SE.93.SE7887	.....-	102
H1K.CM.96.MP535	.....-	102
H101_AE.TH.90.CM240	.....-	102
H102_AG.NG.x.IBNG	.....-	101
H103_AB.RU.97.KAL153_2	.....-	87
H104_cpx.CY.94.CY032	.....-	103
H1N.CM.95.YBF30	.....-	102
H1O.BE.87.ANT70	.....-	115
H1O.CM.91.MVP5180	.....-	116
CPZ.CD.90.ANT	.....-	74
CPZ.CM.01.SIVcpzCAM13	.....-	100
CPZ.CM.05.SIVcpzEK505	.....-	94
CPZ.CM.05.SIVcpzMB66	.....-	102
CPZ.CM.98.CAM3	.....-	103
CPZ.GA.88.GAB2	.....-	100
CPZ.CM.98.CAM5	.....-	108
CPZ.CM.05.SIVcpzMT145	.....-	102
CPZ.TZ.00.TAN1	GYTRPFKTSSGSSGSACKH-	137
CPZ.CM.05.SIVcpzLB7	.....-	100
CPZ.US.85.CPZUS	.....-	103
CPZ.GA.88.GAB1	.....-	101
MUS_1.CM.01.1085	.....-	99
MUS_1.CM.01.CM1239	.....-	102
MUS_2.CM.01.CM1246	.....-	109
MUS_2.CM.01.CM2500	.....-	102
MAC.US.x.239	.....-	131
H2A.GW.x.ALI	.....-	137
H2A.DE.x.BEN	.....-	131
H2A.SN.x.ST	.....-	131
H2B.GH.86.D205	.....-	129
H2B.CI.x.EHO	.....-	139
H2G.CI.92.ABT96	.....-	113
H2U.FR.96.12034	.....-	112
MAC.US.x.251_1A11	.....-	131
MAC.US.x.251_BK28	.....-	131
MAC.US.x.EMBL_3	.....-	117
SMM.US.x.H9	.....-	129
SMM.US.x.PGM53	.....-	131
SMM.US.x.SIVsmH635F_L3	.....-	129
SMM.SL.92.SL92B	.....-	127
STM.US.x.STM	.....-	129
MNE.US.x.MNE027	.....-	131
GRV.ET.x.GRI_677	.....-	95
VER.KE.x.TYO1_patent	.....-	101
VER.KE.x.9063	.....-	120
VER.KE.x.AGM155	.....-	120
VER.DE.x.AGM3	.....-	120
TAN.UG.x.TAN1	.....-	102
SAB.SN.x.SAB1C	.....-	74
OLC.CI.97.97CI12	.....-	108
DRL.x.x.FAO	.....-	125
RCM.NG.x.NG411	.....-	116
RCM.GA.x.GAB1	.....-	106
MND_1.GA.x.MNDGB1	LIEDLTSFARE.....-	147
MND_2.x.x.5440	.....-	111
MND_2.CM.98.CM16	.....-	123
MND_2.GA.x.M14	.....-	88
GSN.CM.99.CN166	.....-	93
GSN.CM.99.CN71	.....-	120
DEB.CM.99.CM40	.....-	116
DEB.CM.99.CM5	.....-	102
DEN.CD.x.CD1	.....-	100
TAL.CM.01.8023	TVTRP.....-	128
TAL.CM.00.266	.....-	111
MON.CM.99.L1	.....-	105
MON.NG.x.NG1	.....-	127
WRC.CI.97.97CI14	.....-	126
WRC.CI.98.98CI04	.....-	133
WRC.GM.05.Pbt_05GM_X02	.....-	113
SUN.GA.98.L14	.....-	103
LST.CD.88.524	.....-	112
LST.CD.88.485	.....-	101
LST.CD.88.447	.....-	101
LST.KE.x.lho7	.....-	112
SYK.KE.x.SYK173	.....-	111
SYK.KE.x.KE51	.....	98
COL.CM.x.CGU1	.....	79

```
                        Rev start                     exon 1 end exon 2 start           NLS                              Leu-rich effector domain                          Rev end
H1B.FR.83.HXB2          M.AGRS.......GDSDE...ELIRTVRLIKLLYQSNPPP...........NPEGTRQARRNRRRRWRERQRQIHSISERIL.GTYLGRSAE.......PVPLQLPPLERLTLDCNEDCGTSGTQ..................GVGSPQILVESPTVLESGTKE*.....  116
H1A1.UG.85.U455         -.-R-........-NP--...D-LKA--I---------C-...........--R-S---K-------A----D-L-----.SDC---P--.......--------I---R---S-S-------QPQGTE...........T---G---S---SA--G----N-.....  124
H1B.US.90.WEAU160       -.----.........-------..D-LK------I--------...........S--------------------RK--GW--.N-----PT-.......----P---D-----K--------..................-----S---L---A---P---------.....  117
H1C.ET.86.ETH2220       -.----...........------...--LKA--I--I------Y-...........T---------------A-----TL-----.SNF---P--.......--------------N---S--S------QSQGTT...........E---N-*.......................  107
H1D.CD.84.84ZR085       -.----...........------..D-LTA-----I--------...........S---------------A---Y----G----.S------E-.......------------IN-N-S---------..................-------S----AI---R-E--.....  117
H1F1.BE.93.VI850        -.----.........----T...--LKA-KC--I------Y-...........K---------------A-----RAL-D---.SSC----E-.......------------HIN-S----QGPEE..................----S--SG--HA----------.....  117
H1G.SE.93.SE6165        -.----.........-ST--...--L-A-KA--I------Y-...........P---------------A-----SA------.TA----P--.......-----------H---S--S------QPQGTE...........T---RS*........................  107
H1H.CF.90.056           -.----.........-A--T...--LQVCKI--I------C-...........E-T--------------A-----RE------.TSC---PP-.......--T-----------N-S-------EK..................-E-----SL--S-I-GT---------.....  117
H1J.SE.93.SE7887        -.----...........----D...Q-LLA-----I------Y-...........K-N-S------------A--N--D-----P.SSC---P--.......---------I---R---S----N-----..................----D---SG-PCM--GA---------.....  117
H1K.CM.96.MP535         -.---R.........--P--...Q-LT---T--I--E---Y-...........-L--S--T----------A--K--S----L-.SAC-------.......---------I-K-N-N-D--P-K.--EG..................-L-----S--PC------------.....  117
H101_AE.TH.90.CM240     -.----.........-ST--...--L-A--I--I------F-...........SS----T-----------A-----SA------.S-C----T-.......---------V---N---S--G------QSQGTE...........T---R---SG--SVI-GP---N-.....  124
H102_AG.NG.x.IBNG       -.----.........--A--...--L-A--I--I------Y-...........P------T-K--------A-----RA------.S-C-------.......---------I---N---S---------LSQGTE...........T-------S---YII-G---------.....  124
H103_AB.RU.97.KAL153_2  -.----...........------...--LK-I---F--------...........S---------------------H------Q--.S-----PE-.......--L-H-----------S---------..................-------------------D----------.....  117
H104_cpx.CY.94.CY032    -.----.........-NI--...D-FKAA-A--I------Y-N...........--T--------------A--K----L-----.A-----P--.......----------K---N-S-------DK..................------VS--L-A--GT-A--------.....  118
H1N.CM.95.YBF30         -.----.........-VN--...--L-A--V--I------Y-...........-SK--------------A-----RA------.SSC--GPP-.......--D-P----D-----TE--S--P--E..................-Q-...........GTA-T---.....  105
H1O.BE.87.ANT70         -.----E........D-Q....-LQAIQI--I------Q-...........S-R-S-N--K-------R--A-VDTLAA-V-.A-VVHGPQN.......NNIVD-----Q-SIR....-PEGD.........Q...........LSEAWTVDPRAEDN*.............  102
H1O.CM.91.MVP5180       -.----E........E-QQ...-LQAIQI--I------C-...........T-A-S-N--K-------R--A-VD-LAT---.A-VVHG-QD.......NNLVD-----Q-NIR....-PEAD.........R...........LPGTGTVDPGTKDN*.............  103
CPZ.CD.90.ANT           -.---EEL.....EGT--Q...-LKA-KI--I------Y-...........K-A-S-A----K----KK--D-VEGLAA-V-.R-LV-GPD-.......HN-VD--D-QN-S-EPLGSPP-PLQNYFQTVQ...........PSGEQS-HPLDNNQN*.............  118
CPZ.CM.01.SIVcpzCAM13   -.----E.......HD-A...R-LQA-KI--I--E---Y-...........S-----N---------A-----RE--N-V-.AAH---PPQ.......-GH-E--E-DK-S-H-L-TV-KV--TVNQHPQ...........TDT-ETDSTEGDNPI-GKRI-N-.....  124
CPZ.CM.05.SIVcpzEK505   -.----.........-VN--...--L-AI-V--I------Y-...........-DS------K--------A-----RAL------.SSCV-GLQ-.......-ST-P----D--S-NPE--L---E-E..................H--...........GTA-T*......  103
CPZ.CM.05.SIVcpzMB66    -.--G-.........-N---...--L-A--I--I------Y-...........S---S-------------A-----N-L------.S-C-E-PTG.......--S-P---I-------SA--I--G--DPPQGTE...........T-T---NTPEGHS-I-GT-A-N-.....  124
CPZ.CM.98.CAM3          -.----E........--D-A...C-LQA--I--V--S---Y-...........DNK-------------A-------L------.SS-----P-.......-PA-L---I-----I-TG--I-KP--EPQQGTA...........TAEELQ--P-G-DNL-GK-ATN-.....  125
CPZ.GA.88.GAB2          -.----.........--E--...R-L-A-KA--I--D---Y-...........SNR---S----------A-----CE--Q--F.SSL---HP-.......-DD-P-------S-GPL-PA--I--TGPERTE...........TTQ-ASEPAERDR-I-AAR--N-.....  124
CPZ.CM.98.CAM5          -.----E........--E-T...R-LQA--I--I--D---Y-...........DNK-------------A------E-GQ-V-.SSC----S-.......-PD-L---ID----NPV--SAN---EHQQGTA...........TVE-L*......................  107
CPZ.CM.05.SIVcpzMT145   -.----E........--E....N-L-AI-I-RI--E---Y-...........SQS---R----K----------NQ-G--V-.ASC----S-.......-PQ-P---I---CISDL-TT-DT--IQPQGST...........TNRY--TPTERD-A--DNRNQ--.....  123
CPZ.TZ.00.TAN1          -.---EE..........-AN...-LY---I--I--D---Y-...........SGA-S-T-----------Q--H-VDALAS---.QYR--GPQ-.......-PH-DI-D-SK-H--PLDQPAS.........T...........ETGDNQLGTQP-NSA*............  105
CPZ.CM.05.SIVcpzLB7     -.----.........--R--...--L-A--I--I------F-...........H-----S----------A-H---SE-AG---.S-C-E-PEG.......----P---ID--S-NSE--I-DC--IHSEGTE...........T-----T-SAGNS---GK-A-N-.....  124
CPZ.US.85.CPZUS         -.----G........--A--...Q-L-A--I--I--A---F-...........-NS-------------KN-----NE--Q---.S-C----S-.......-PD-P---IGG--INSE--STNP--EHQQGTA...........TVE*.......................  105
CPZ.GA.88.GAB1          -.----E......PQD-A...R-LQA-KI--I------Y-...........S-----K-----------A--K--SE--G-V-.A-----PPK.......-GD-E--E-DK-S-Q-V-TTQDV--SNTSQPQ...........TAT-ETVPAGGNYSI-GK-A-N-.....  125
MUS_1.CM.01.1085        -.-...........-V-ENDR.----AC-A-QI--K---Y-...........TAD---S----------KRK----RE------.LSL---PPQ.......-SD-E--D-NK-S-H.......PLVATSES..............SPPDTEGVNKIVVHAQGTD*.......  110
MUS_1.CM.01.CM1239      -.-...........-A-ENVQ.D----C-I-RI------Y-...........TSR-S-A---------QQ--D--RA------.VAV---PPQ.......-SD-V--DISQ-SIN.......PSLDTPEG..............SSPDTEVVSTTVDRD*...........  106
MUS_2.CM.01.CM1246      -.-...........-A-GNDQ.----FC-Q-RI--D---Y-...........TVSR--AT-K--K---KK--.......AQ--.VAVFR-PPQ.......-ND-D--D-DQ-R-E.......PLVVDSKN..............LPPPDKGASK-VDSA*............   99
MUS_2.CM.01.CM2500      -.-...........-TGRDIE.--LKACQT--I--N---Y-...........S-G--KN----------Q--C--RAL-----.VS----PPQ.......-DSVE--D--K-S-K.......PLVADPAD..............SPSP-EVEHL-VAQG--NQHPPSGSP*  117
                        Rev start                      exon 1 end exon 2 start                                                                                                 Rev end
MAC.US.x.239            -.SNHE.........RE-....--RKRL---H--H-T--Y-...........TGP--ANQ--Q-K----R-WQ-LLALAD--Y.SFPD.PPTDT.......-LD-AIQQ-QN-AIESIP-PP...-NTPEALC...............D-TEDSR--QD*............  107
H2A.GW.x.ALI            -.TE-A.........GE-...D-Q-KL---R--H-T--Y-...........QGP--ANQ---------Q-WG--VALAD--F.TFPD.PP-SS.......-LDRAVQH-QG--IQDLP-PP...-DLPESSE...............-ADNNQGLAET*............  107
H2A.DE.x.BEN            -.SE-A.........DE-...G-QGKL--LR--H-T--Y-...........QGP--ASQ-------R-RQWLRLVALANKLC.AVPD.PPTDS.......-LDRAIQH-Q---IQELP-PP...-DLPES.................NSNQGLAET*............  103
H2A.SN.x.ST             -.NE-A.........EE-....--R-KL---R--H-T--Y-...........QGP--ASQ--------KQ-W--LVALADK-Y.TFPD.PP-DS.......-LEQTIQH-QG--IQELP-PP...-NLPESSE...............-IDSSQRLAEI*............  107
H2B.GH.86.D205          -.TA-...........EG...D-QKEL--LH--H-T--Y-...........QGP--ASQ--------KR-GL--LALAD--H.SFST.AP--E.......-LN-AIQR-QN--VEDLPNPPPNLN-SPTTQAPGCVPPVWDQLVPRSA-SGSKGYGRNSCECRRDLMGGSQ  136
H2B.CI.x.EHO            -.NA-...........ER...D-QKGL--LH--H-T--Y-...........QGP--ASQ--------KQ-GL--LALAD--H.PLPD.SPT-G.......-LD-AIQR-QN-IIKDLPNPP...-STPTAQASTCIPPIWDQLVPRSN-SSSQGCGRDSCERGEDLVGSPQ  133
H2G.CI.92.ABT96         -.SNL..........EE-....--RKRL---HF-H-T--Y-...........QGP-XANQ------K-KQ-WL--LALAD--Y.SFPX.PP-RE.......-LD-AIQQ-QS--IQDXP-SP...-AVPKNLQ...............D-PSN*..............  100
H2U.FR.96.12034         -.SNP...........E-...GVQ-RL---H--H-T--Y-...........TGP--ANQ---------Q-WL--LALADK-Y.TFPD.PPTDS.......--D-AIQQ-QG-AIQDLPEPP...AVVPEHQR...............TSAP*...............   98
MAC.US.x.251_1A11       -.SSHE.........RE-....--RKRL---H--H-T--Y-...........TGP--ANQ--Q------R-WQ-LLALAD--Y.SFPD.PPTDT.......-LD-AIQQ-QN-AIESIP-PP...-NTPEALC...............D-TKDSR--QD*............  107
MAC.US.x.251_BK28       -.SSHE.........RE-....--RKRL---H--H-T-SY-...........TGP--ANQ--Q------R-WQ-LLALAD--Y.SFPD.PPTDT.......-LD-AIQQ-QN-AIESIP-PP...-NTPEALC...............D-TKGSR--QD*............  107
MAC.US.x.EMBL_3         -.SSHE.........RE-....--RKRL---H--H-T-SY-...........TGP--ANQ--Q------R-WQ-LLALAD--Y.SFPD.PPTDT.......-LD-AIQQ-QN-AIESIP-PP...-NTPEALC...............D-TKDSR--QD*............  107
SMM.US.x.H9             -.SSN...........E-....--R-RL---HF-H-T--YX...........XGP--ANQ--R-----XQ-WQ--LAXAD--Y.SFPD.PP-DT.......-LD-AIQQ-QG-AIEXLPNPP...ASAPEPLK...............DAAESP*..............  100
SMM.US.x.PGM53          -.SSSA.........GE-....--R-RL---HF-H-T--Y-...........QGP--ANQ--R------Q-WQ--LALAD--Y.SFPD.PP-DS.......-LD-AVQQ-QD-S-ENIP-PP...-SVPEALT...............E...NPR-RQD*............  104
SMM.US.x.SIVsmH635F_L3  -.SST...........E-....--RKRL---HF-H-T--Y-...........QGP--ANQ--R------Q-WQ--LALAD--Y.SFPN.PPTDT.......-LD-AIQQ-QG-AIEDLP-PP...-SAPETLK...............DAAKSS*..............  100
SMM.SL.92.SL92B         -.QNP...........E-....--R-RL---H-IH-T--Y-...........AGP--ASQ--------KR-WI--VALAD--Y.SFPD.PP-DT......DID-AIQQFDS-SIQSLPEPP...-TVPETLR...............DQSAD*...............   99
STM.US.x.STM            -.SDQ..........EE-....-RKRL---QF-H-T--Y-...........QGP--ANQ---------Q-WN-LLALAN--Y.SFPD.PP-ST.......-LD-AVQQ-QG-SIQDLP-PP...PNLPKDLQ...............DTAEN*...............   99
MNE.US.x.MNE027         -.SSHA.........EE-....--R-RL---H--H-T--Y-...........TGPS-ANQ---------Q-WQ-FLALAD-VY.SFPD.PPTDT.......-LD-AIQQ-QN-AIESIP-PP...-NIPEILH...............D-TESPR--QD*............  107
GRV.ET.x.GRI_677        -..SLG.........KE-K.....QALKI--T--G---Y-Q...........FS--ARQ--RA-Q---KQ-Q--DK-AG-V-NTFED.QQLVA.QLQELQLENKDLV-QH-PDPPHIHQDS--IPAVWAPATPRGSNRACSSS-E-CEGS-GQTGCYCPIRLSGSHQQSK  141
VER.KE.x.TYO1_patent    -..PLG.........PE-R.....-LL---W---ST--Y-P...........SG---ARQ--RA-----QQ-D--RVLV--LQ.............EQVYA-DRLADEAQH-AIQQLP-PPH-A*..................................   84
VER.KE.x.9063           -..PLG.........PE-R.....KVFS--AY-NRT--Y-P...........V---ARQ--RA----------LRFLA---W.NE..TKE....RQLEQTLDQLALDVQH-AIQHLP-PPN-S*.................................   89
VER.KE.x.AGM155         -..PLG.........PE-R.....-LL---AF--R---Y-S...........V---ARQ--RA----KN--K--YALA---W.---.RQE....DQLVQAIDQLVLDTQH-VTQQLP-PPSQA*.................................   89
VER.DE.x.AGM3           -..PLG.........-E-R.....-LL---AF-NKN--Y-P...........V---ARQ--RA-----QA-E-LRALA---W.HS..RVE....EQLVQAIDQLVLDQQH-AIQQLP-PPS-S*.................................   89
TAN.UG.x.TAN1           -..SLG.........-EDLR.....-IIQI-RI--H--QY-P...........G---ARQ--R--Q-FQQQ---VAAL----F.IAQQR-DPSGGESLAAAFDQLVLDNQQ-VIETLP-PPQEPHDSSSTA*........................  103
SAB.SN.x.SAB1C          -..SLG.........QE-LL.....-RF-I--F--TT--Y-P...........GQ--ARQ--RA-Q--AKQRQ-VIHLA----.E-PVSQI....DHLAQEFDQLVLDNLQQPPSLPPGHP-ENQTANSSS*........................   99
OLC.CI.97.97CI12        LQSHLGR..............GER-WY-IN-I-GTPPIFRYP.........PGQ--ARQ--R--Q-VLREHL-RVELYH-VY.RDIESKLERLTITPVEGSSESHSLDPEELATRL-NLK-EE-TWDQFAAEETEKGKSSTEGSGAIGERRSRNNWS-RTSI-QIWQ*..  143
DRL.x.x.FAO             -..SAGPERE........PPPWFQEYL--VTR-W-GRI--.........RDDQLPQ-ARQ-KR--A-Q-RVEH--RTLQA---.QSLER-NSRD.....LVEGIERIH-AEERESSS*....................................   92
RCM.NG.x.NG411          -..LLGE.........EE-ADQ-IR-RI---H-IHL---Y-...........QSG--ANQ--RK--Q--R-WT--LQLA---F.LYPDTPPDRD.......-LQEA-SN-QQ---SDLPEPPVNPFSNPSSFA................VDRS*...................  107
RCM.GA.x.GAB1           -..LPGQ.........DE-...D-RKKI---NF--LI--Y-...........QHG--ANS--KK-----R-WQ-VQQLA----.LDSTDPPV-Q.......DLDAAIAD-QK-Q-NNLPEPPVD.FS*..................................   93
MND_1.GA.x.MNDGB1       -.STGNV............YQELIR-YLVVV-K--EGLSTCLFT...GLPAEPIPQ-ARQ--R-KQQL-T-RA-LRELEG---.KQI-D-GPDQ.....LCQGVTNLA-AEKSESS-*..................................   95
MND_2.x.x.5440          -STD-GDL..........DEGFWRKYQAIV-Q-WEGIS-SSLCLQGRLSPAPQPQ-ARQ--R--A-L-KTEH-VRELQT--*................................................................   71
MND_2.CM.98.CM16        -PTETGDL..........DRDFWKKYQKIVQR-WEGLS-SSLCLQGRLSPSPQPQ-ARQ-KR--A-L-KTEY--RQLQS--WEALEQRANQRLAEDFKGLRVAEESTSSN*..................................  100
MND_2.GA.x.M14          -STEPGEL..........GRDFWEKYH-IV-Q-WEGLS-SSLCLKGRLSPTSQPQ-ARR-KR--AQL-RTEH-VRELQ---WQTLEARANK-LVEGVNNLHLAEESAGSN*.................................  100
GSN.CM.99.CN166         -ADHAR........-NDQKLQ.N--LAC----T-HR-RLLTSLPA.DPYPTSSGTRSARRN-R--W-A-QG-VREI-NRILESLVGRPEEPGDLDLPDLGQLS-DS-WDREVPVGTAAESNPESNPASETAQGSSSQSDSLH*..........................  132
GSN.CM.99.CN71          -.-DPAN.......-RDQKLQ.N-LLAC----T--K---Y-...........TAS---S----KK----A----VRQ--H---.ESLV--PE-.......SGN-D--D-GH-S--SP--RQVPVGAAAEG................-PTSETAGG-GSQSS-LY*........  120
DEB.CM.99.CM40          -.-HAGGR.....-SAE-NTRQ-LKVIS---I------Y-...........KGG-SAST--R--QK--R--G-VD--A----.QSH--G-PT.......TADVA--D-SQ-H-AD*.................................   90
DEB.CM.99.CM5           -.-HAGGR......--A-Q...A-L-MI-I--S------Y-...........KGG-SAST--R--QK--R--E-VDD-AQ---.EAH---HLT.......ASDVA--N--E-R-AD*.................................   87
DEN.CD.x.CD1            -.--LQER.........-Q.....-VKA-----K------Y-...........T---SKS---R---K--RQ-A-VVA------.EYH----LT.......AADVE--D--K-SVSD*.................................   84
TAL.CM.01.8023          -.SLTGES......-KGED...P-LAAI-R-RI------Y-...........EQQ---S---R-Y---KA--E-VEN-AS---.CH---GPQQ.......-PQ----D--K-S-AELGPPEAP.....ESDH...........PTGSASH-RLPEN*..............  111
TAL.CM.00.266           -.SLTGGS......-RGED...P-LAAI-R-RI------Y-...........EQQ-S-S---R-Y---KA--L-VEQ-AS---.EH-----DR.......SDQIH--D-SQ-S-ADLGP-ESC.....ESQE...........PTGSASH-RLPTTPPDS-SSDSPVWVQ*  125
MON.CM.99.L1            -.--NG.R.....DEERR...L-SLALAAVRI-QE---Y-...........I-R--AN---------RQ-L--A------F.YH-----Q-.......-C--DI-D----SISDLP-PPE-VP...................EAAT-PAHTPA---GKP*.........  112
MON.NG.x.NG1            -.--AE.RG....AAVQD...I-RLAIGA-RIIRE---Y-...........S-S---S---------A--E-VX-LC---F.NH--D-LPR.......-NS-D--AI-G-R--DLPEPPE-A...................QPPD-SSPRPV--ASG-Q*.........  113
WRC.CI.97.97CI14        ..IVLTAR................KLSQIQQQ-L-AE--QQ...........LGR--ARQ--RN---LQQYRN-RLAVQ---C.NSL-EQL-N.......LTIQH--TPPSHCVESAFVPPVVDKRWVESSQRHQWKSIWSDTQEDCKE-EGEYAGRGISNWQQKDSSKQ*  133
WRC.CI.98.98CI04        ..IVPTAR................-L--VQQK-Q-AE--QQ...........LGR--ARQ--RS---QQYFR--RLAVQ---C.NNLVEQLDK.......LLICS--NPPSNSVESPCVPLLVDPAWVESSQKNQWSSACMHTQEDYKE-EGNNGERDIRK*.........  124
WRC.GM.05.Pbt_05GM_X02  ..IIHTAR................-LI-ISQT-Q-AESY-E...........PGR--ARQ-------FHSHR-IRLAVQ---A.D-L-EQLGN.......LTLQH--EPPAAK-QPS-LSVDRN-EWVESSQKTLWSSVCCATGEDCCASAEKKGEGDIHAWQQRDKQKQQ  134
SUN.GA.98.L14           -.STG..............DDSINQYL-IS-R--EGLA-G.........N..LPQ-HRQ--R--D-E-KNLH-LRAVQ---F.A-T-DSRLGR.....AFER-SVSDSSQVAESLGNSPS-KHLPPAKFLVAPTYDFLPSWATPLAD--R-AGFAPYSGYEQDQERVQNQ  138
LST.CD.88.524           -.STGP.............EG--PTYLK-SRI-WTGLQVG.........ERREPR-ARQ--RD-E-Q-KHLH-LRAVQ---F.QATVE-GL-R.....AFTG-TVCDSPEVAEGVGNTAPHPRVAIPHLAGTYSDPFLPPWATSLAN--QIPGGECSENCELDQEQEQKE  141
LST.CD.88.485           -.STGG.............EE--P-YLKIS-I-WTGLQVG.........ERREPR-ARQ--RD-E-Q-RHLH-LRTVQ---F.HSTVE-GL-R.....AFTG-TVCDSPEVDQGIWNSSPQPCVAVPFVARPLSDPFLPTWATSLE-Q-RMDGGKRSEDCALELEQDQKE  141
LST.CD.88.447           -.STGG.............EK-SP-YLKIS-I-WTGLQVG.........EKREPR-ARQ--RD-E-Q-KHLH-LRTVQ---F.HSTVE-GL-R.....AFTR-TVCDSPEVDQGLWNSSPQPCVAVPFVARPFSDPFLPTWATSSE-L-QTGGGKRSEDCELELEQDQKE  141
LST.KE.x.lho7           -.STGN.............GD--P-YL--SRI-WTGLQVG.........ERREPR-ARQ--RE-S-Y-DYLH-LRAVQ---F.QATVE-GL-R.....AFTR-AVSDSPEVAQGRGNTPPITSVAEPQLAVAFVDPFLPKWATPLADQ-QMDGGKRSED-ELAQGEMQKE  141
SYK.KE.x.SYK173         -.P..........DQGS-QLAVFL-MIAHL.....QE-Y-...........G---P--T--R---Q--Q-RT-RLYLQQ--F.EAIF-SRTA.......ALEDS-QQ-QISD*...................................   77
SYK.KE.x.KE51           -.S--ERE.....DTQEQLLRT-L-IAQQL#Q-EAAA-Y-...........L-Q-PSR---RN-N-Y-QL-A-RLYVQQ--F.E-IAR----.......TLEQSFGE-QITD*.................................   86
COL.CM.x.CGU1           -.TNAGVRPVAFTVSYAFFQKGWESAI-R-RV-HKH--W-R...........PGV--ARQ--RNTQ-A-KQ-Q--LA-DG-VV.EQLVASLQQWD.....-YCPASTSANN-RQEEECLADS-VSIGTQGSK................-SDS*..................  118
```


PLV Proteins

H1B.FR.83.HXB2	.....	116
H1A1.UG.85.U455	.....	124
H1B.US.90.WEAU160	.....	117
H1C.ET.86.ETH2220	.....	107
H1D.CD.84.84ZR085	.....	117
H1F1.BE.93.VI850	.....	117
H1G.SE.93.SE6165	.....	107
H1H.CF.90.056	.....	117
H1J.SE.93.SE7887	.....	117
H1K.CM.96.MP535	.....	117
H101_AE.TH.90.CM240	.....	124
H102_AG.NG.x.IBNG	.....	124
H103_AB.RU.97.KAL153_2	.....	117
H104_cpx.CY.94.CY032	.....	118
H1N.CM.95.YBF30	.....	105
H1O.BE.87.ANT70	.....	102
H1O.CM.91.MVP5180	.....	103
CPZ.CD.90.ANT	.....	118
CPZ.CM.01.SIVcpzCAM13	.....	124
CPZ.CM.05.SIVcpzEK505	.....	103
CPZ.CM.05.SIVcpzMB66	.....	124
CPZ.CM.98.CAM3	.....	125
CPZ.GA.88.GAB2	.....	124
CPZ.CM.98.CAM5	.....	107
CPZ.CM.05.SIVcpzMT145	.....	123
CPZ.TZ.00.TAN1	.....	105
CPZ.CM.05.SIVcpzLB7	.....	124
CPZ.US.85.CPZUS	.....	105
CPZ.GA.88.GAB1	.....	125
MUS_1.CM.01.1085	.....	110
MUS_1.CM.01.CM1239	.....	106
MUS_2.CM.01.CM1246	.....	99
MUS_2.CM.01.CM2500	.....	117
MAC.US.x.239	.....	107
H2A.GW.x.ALI	.....	107
H2A.DE.x.BEN	.....	103
H2A.SN.x.ST	.....	107
H2B.GH.86.D205	ESGESNHRDPQENQTRT*.....	153
H2B.CI.x.EHO	ESGRRDHCNTQEDQTRG*.....	150
H2G.CI.92.ABT96	.....	100
H2U.FR.96.12034	.....	98
MAC.US.x.251_1A11	.....	107
MAC.US.x.251_BK28	.....	107
MAC.US.x.EMBL_3	.....	107
SMM.US.x.H9	.....	100
SMM.US.x.PGM53	.....	104
SMM.US.x.SIVsmH635F_L3	.....	100
SMM.SL.92.SL92B	.....	99
STM.US.x.STM	.....	99
MNE.US.x.MNE027	.....	107
GRV.ET.x.GRI_677	KSAARP*.....	147
VER.KE.x.TYO1_patent	.....	84
VER.KE.x.9063	.....	89
VER.KE.x.AGM155	.....	89
VER.DE.x.AGM3	.....	89
TAN.UG.x.TAN1	.....	103
SAB.SN.x.SAB1C	.....	99
OLC.CI.97.97CI12	.....	143
DRL.x.x.FAO	.....	92
RCM.NG.x.NG411	.....	107
RCM.GA.x.GAB1	.....	93
MND_1.GA.x.MNDGB1	.....	95
MND_2.x.x.5440	.....	71
MND_2.CM.98.CM16	.....	100
MND_2.GA.x.M14	.....	100
GSN.CM.99.CN166	.....	132
GSN.CM.99.CN71	.....	120
DEB.CM.99.CM40	.....	90
DEB.CM.99.CM5	.....	87
DEN.CD.x.CD1	.....	84
TAL.CM.01.8023	.....	111
TAL.CM.00.266	.....	125
MON.CM.99.L1	.....	112
MON.NG.x.NG1	.....	113
WRC.CI.97.97CI14	.....	133
WRC.CI.98.98CI04	.....	124
WRC.GM.05.Pbt_05GM_X02	HQGNYCQTLHKP*.....	146
SUN.GA.98.L14	QGESIIVSEGKK*.....	150
LST.CD.88.524	QRRAIKN*.....	148
LST.CD.88.485	QKRHISSSCERHQTGTKEKMEIQIQREKWVSQ*	173
LST.CD.88.447	QKRHISSSCKKPQTGTKEEMEIQIQRKKWVSQ*	173
LST.KE.x.lho7	QRTVIEH*.....	148
SYK.KE.x.SYK173	.....	77
SYK.KE.x.KE51	.....	86
COL.CM.x.CGU1	.....	118

```
                                                                                   phos          phos
                                transmembrane domain                               cytoplasmic domain
                                Vpu start                                α-helix                                α-helix                Vpu end
H1B.FR.83.HXB2                  TQP...IPIVAIVALVVAIIIAIVVWSIVIIEY......RKILRQRKIDRLIDRLIERAEDSG.........NESE...GEISALVEMGVEMGHHAPW.DVDDL*   82
H1A1.UG.85.U455                 MT-...LE-W--TG-I--L-L-----T--G--*#.....K-L-K-K-----LN-IR-----------------D....GDTEELSLL----NYDLG.VDNN-X   81
H1B.US.90.WEAU160               M-S...LQ-L---------G----------L---.....-----------IRD-----------------....GDQEELSAL----------.-I----   82
H1C.ET.86.ETH2220               DLLAKVDYRIV---FI--L-------T-AY---.....--L----R----K-TR-----------------D....GDTEELSTM-D--NLRLL.--N---   85
H1D.CD.84.84ZR085               M-S...LQ-L------L-L-------T--F---.....-R-K-----W----IR--E-----------------....GDKEELSTL---------DVD----   83
H1F1.BE.93.VI850                MSY...LLAIG-A--I--L-------T--Y---.....K-LV-----NK-YK-IR-------------------....GDAEELAALG---PFI-G.-INN--   82
H1G.SE.93.SE6165                M-S...LV-L---G-I--F-A-----T--F---.....-E-RK-KR-GK-L--IR-----------------D....GDTEELVTL----DFD--.VG-N--   82
H1H.CF.90.056                   MYI...LGLGIGALV-TF--A.VI--T--Y---.....K-LV--K------E-IG-------NES......DG.D......TEELSKLM----LNLG.Y-A---   81
H1J.SE.93.SE7887                MI-...LQ-A----FI---FL--GM-T--Y---.....K-L-------K----IR-------NES......DG.D......TEELADL--R-P-DL-.N-N---   82
H1K.CM.96.MP535                 MVS...LA-SIVALV-AL-LA.-I--T--Y---.....--LVK-KR-NW----IR-------------------....GDAEELADIG-L--LILG.NI-N--   81
H101_AE.TH.90.CM240             MT-...LE-S---G-I--L-L----*T--A--V......K-----------VK-IR-------------------....GDTDELAKL----DFD--.VG-N--   81
H102_AG.NG.x.IBNG               M--...LT-T---G----F-A--------Y---.....---RK-K-----L--IR-----------------D....GDTEELSTLM---YEYIL.-N-N-X   82
H103_AB.RU.97.KAL153_2          M-S...LA-A-------VG------G---F---.....----------IR-------------------.....GDQ-ALM----LV--.-A----   79
H104_cpx.CY.94.CY032            MLF...WE-W---G----L--V----TL-F---.....K-LR---R--S-YN-IR-----------------D...GDAEELSTL-G--NFD--.VG-N--   82
H1N.CM.95.YBF30                 ML......SLGFI--GA-VS--VI--ALLYR-.......K--KL-E--KHIRQ-IR--E----------D.GDA-WLDGD-EWLVTLLSSSK.LDQGNW   82
H1O.BE.87.ANT70                 MHH...RDLL--IIISALLF-NVIL-GFILRK-..LEQKEQDRKE-E-LERLR-IR-IRD--DYES......-G.........EEEQ.E-MDLVLSHG.FDNPMF   83
H1O.CM.91.MVP5180               MHQ...ENLL-LI--SALCL-NVLI-LFNLRI-..LVQRKQDR-EQE-LERLR-IK-IRD--DYES.......-.........EEQQ.E-MELI-SHG.FANPMF   83
CPZ.CD.90.ANT                   MTNIFEYAFL-FSIVLWI-C-P..ILYKLYKI-.KQQQIDNKRN--I-EV-SR--SIDSAIEE.........D-.........EAD..TYYL-.SGF..ANPVYR   79
CPZ.CM.01.SIVcpzCAM13           MIL......-ALGCLAIAL-LLN-FI-RNLW......RLCKQNKLEVE-EN-AL--T----------------D...-GEEEKLQRLLHDHDNMHG.FANP-F   81
CPZ.CM.05.SIVcpzEK505           MLL...LIKLGFIG-AIETL-V----A--YRI-.......-EVKVEE--SQ-RQ-IRD-----------------.....DGDAEELANLLPPDR.IDQ-NW   78
CPZ.CM.05.SIVcpzMB66            MDI..VQQVGLL-V-IIELV-V--I-VK-Y......KLCKEDR--K------A-IR-------NES......DGDTEELQDLITEGDNLMHI-IRDNR.NN*.V-   88
CPZ.CM.98.CAM3                  MLT..W.EQIGLIGIGIE------A-G-AFK-W..RKGKEEEN-AQ--RT--E-I-S-Q----------------D.......GEDQRQLDREI-VYG.FDNPMF   83
CPZ.GA.88.GAB2                  MLS..MWVAIGLIGIGTLLV-N----G--G-SVYKRWKRH-EEQRI.--LI.IKIT------------------D....GEDKETL...ATLL-NNG.FDNPMF   84
CPZ.CM.98.CAM5                  MLI..W.EQIGLI--GIEL--V----G-AYKKW..REIKEERN-TQR-YN-AE-IRI-Q----------------N.......GEE-AQLDNLL--NG.FDNPMF   83
CPZ.CM.05.SIVcpzMT145           M-L....E--L-ILFIALMLV--FA-IAAY......KEYK-LQQV-R-E--Q--IRS------NES......DGD...EILLVEELMQVHQHQVNPD-.MDRI-F   85
CPZ.TZ.00.TAN1                  MIK..IVVGSVSTNVIGILC-LLILIGG.GLLIGIGIRRELERERQHQRV-ERLARRLSI---VEE......D-EF...............NWNNFDPHNYNPR   80
CPZ.CM.05.SIVcpzLB7             MDL....IELGLIG--IEL--V----LKAYQL-......KENI--KA-NK--E-IR--G-----------------.........GD-DELHAILRSG.-PELVL   76
CPZ.US.85.CPZUS                 MLN..WFE-GL-ALGI.EG-LVVII-GL-ARLW..RQIKI-ENT-QE-QN-LE-IRI-E----------------D....GEEEETL....AKLLSSLELDNX..   81
CPZ.GA.88.GAB1                  MTL..LV.GLVLILVGL.-AWN-CI-GYI-KWGYRRYKRHRLETE..-E--NLI-R------------------N...--EEERL-QLIHNYN-NNH.FANPMF   86
GSN.CM.99.CN166                 MH-.....AAVWWWGAAI-TFIYLCVALLALY.....LAWDKWVKG-PKPTQVAV-RLI-............D-ED.........SGIYDDASSELTGFNGFANP   74
GSN.CM.99.CN71                  MSA.....AALWWWGAAV-TFIYFCLA-FALY.....LAWDKWIKG-PK.IPVAV-RLV-............DDE-.........SGIFEDASSEPNAYGFANPG   73
MUS_1.CM.01.1085                MNY......WYLA-VI-TG-YFVIAIFAF..VLAYQRW...CKPKKVEVSV-RL-E-GDG---IFE......DAED..........D-AES.E---FANPAFEQ-   77
MUS_1.CM.01.CM1239              MNY.WYLAAALVTGIYFL-ALFAF-LAYQ...RWCRPK........VEVSV-RL-E-GDS---IFE......DAED...........PNGD----FANPAFEQ-   76
MUS_2.CM.01.CM1246              MHY.WYLGAAIVT-IYLV-ALVAFILAYQ...RWCQPK......PPIEVNVLRL-E-GDS---IFE......DACD........GEDEDSHRAFAN-SFEP*SNI   80
MUS_2.CM.01.CM2500              MNW.WWFAAAVVT-IYFV-ALVAF-LAYQ...RWCQPQ.....KGQVEVNV-RL-E-GDT---IFEDAE...DG-D..............D.QR-GFLNPAFE--   78
DEN.CD.x.CD1                    MLK.......IKLGPIEYYCLVFA-VITW-AAIGVGYLAYRAYKSYREELRYI--RLWSI---..........Y--S......QEDP....*GPPPY-V..DNWNN   76
MON.CM.99.L1                    MNY.WWSLVAITYS-ILIALPVAAWAW.W...RYYKITK-FKRIDQE-Q---QIHER-RH---VDTESESEQH-ET...................GFVNP-FNDD   82
MON.NG.x.NG1                    MNT.....YWIAAIS-WCV-VLLAPCTLY..LLYQSYKEH-RVSRFEQ-FQRLVQXYQEX---..........Y-D-.......DEDHNSFDNPLFDDG.-P-QWV   81
```

```
                                                                                                                                                         glycosylation NDT
                                                                                                                                                                V1
         Env start                signal peptide gp120 start                                                 glycosylation NVT                            glycosylation NSS
                                                 *                                 *                                                        *       *       *
H1B.FR.83.HXB2           MRVK...EKYQHLWRWGWRWGTMLLG.MLMICS..ATEKL.....WVTVYYGVPVWKEAT...TTLFCASDAKAYDTEVHNVWATHACVP.TDPNPQE...VVLVNVTENFNMW..KNDMVEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDLKNDTNTNSSSGR....  146
H1A1.UG.85.U455          ---M....GI-RNYPCL------I--.LII--...NAQQ-.....------------D-V...-------------A-------------.-------...ID------E----..--N--D-------------------------T-D-HNITINNTN-NTN......  142
H1B.US.90.WEAU160        -------...GIRKNYQHL-K--I----.I------..-A-N-.....---------------...---------------------------.-------...----E--------..--N-----------------------------T-N--NVNVTNLK-ETNTNSSSG  149
H1C.ET.86.ETH2220        -K-M....GI-RNCQQW-I--ILGFW.-----N..GMGN-.....------------D-S...P-----------------G-F----.---S----...LG-E----------..----------Q--------G------------T-N-NAI--N-KVTNN-......  143
H1D.CD.84.84ZR085        -------...GIKRNYQPL-K--I----MLM-TY-..-ADN-.....---------------...----------S-K--A--I---------.--------...IE-K----------..--N--D----------------------R--T-N---ASRNSTD-N-T......  144
H1F1.BE.93.VI850         ---R....GM-RN-QHLGK--LLF--.I-I--N..-ADN-.....---------------...------------ER-A-------------.--------...-F-K------D--..--N------T-------------------T-N--NAT-NSQEKPGA......  143
H1G.SE.93.SE6165         ---T....GI-RN-*HL-K---LI--.LVI---..-SNN-.....-----------ED-D...----------S-SA-S-----------.--------...IIME---Y----..--N-------------E------------T-N---VT-KG-KRNN-TDNSTE  148
H1H.CF.90.056            -E.T......-RNYPSL-----LI--.--L----..-AQN-.....--------------K...------------E--K------------.--------...M-ME---S----..E-N------T--------------------T-N--NVR-N-SNST--......  140
H1J.SE.93.SE7887         -E.T......-KN-QTL--G-L-IF-.-----K..-K-D-.....------------D-K...------------S--K-----------.---S----...MN-P------------..-----D--Q-----V--E-------I-----T-N-SNITSNS--T-N-S.....  141
H1K.CM.96.MP535          ---R....GM-RN-QTLGN--ILF--.I-I----.NAD--.....---------------...P-----------EK--------------.--------...-EME----------..--N------T-------E------E------T-N---Y-GTNS--NATST....  145
H101_AE.TH.90.CM240      -------...ET-MN-PNL-K---LI--.LVI----.-SDN-.....-----------RD-D...-----------HE--------------.--------...IH-E----------..--N-----Q--V------------------T-N--NANLTNGSSKTNVSN...  146
H102_AG.NG.x.IBNG        ---M....GI-KNYPLL-----NIFW.IMI--...NA-Q-.....------------T-E...------------------------------.--------...IH-E----K-----..--N----------------------------T-D-HNFN-SYSNS-NLTS....  144
H103_AB.RU.97.KAL153_2   ----------.EIRKHL-----LF--.----------N-.....---------------...------------SK-------Y-----.---S----...IP-K-------G..--N----------------------------T-N-----KEVTST.........  137
H104_cpx.CY.94.CY032     ---M....GM-RNYPHL-E---LI--.LVI---..-SNN-.....-----------RD-E...-------E---EK---I----------.--------...-A-I----------..------------------NEG----A---S---TFT-INATTTNS--GT.......  142
H1N.CM.95.YBF30          -K-M.GMQSGWMGMKS--LLFYL-VS.LIKVIG...S-QH.....------------R--E...-----------HS--A--I--Q------.--------...-L-P---K-----..E-K-AD--Q-------E--------------TML-N-SYGEERN-TN.......  144
H1O.BE.87.ANT70          -E.........KRNK.KLWTLYLAMA.LITP-LS..LRQ-.....YA---A-----ED--...PV------NLTS--K--I--SQ---.---T-Y-...YP-H---DD--I-..--Y-----Q--------------QM-F---QME--NIAGT--E..........  133
H1O.CM.91.MVP5180        -K.........KNN-.KSWSLYIAMA.L-IP-LS.YSKQ-.....YA---S-----E--A...PV------NLTS--Q--I--SQ---.------H-...FP-G---D--DI-..--Y--D---------E------E-M-F---QMN-V--QTNKTG..........  134
CPZ.CD.90.ANT            -----------.KPI..HII--.LA-..LIQFIEK.G-NED.....Y---F------RN--...P-----TN-SMTS---------TS---.I--D-I....--RL-..SVWFNAY.--Y---S-T--MXQ-FQ--H---------M-IKMN--GYNGTPT-P-TTTST...  137
CPZ.CM.01.SIVcpzCAM13    -K-M...-ERKRSFNGFYMTM-ICMA.L-IP-LT...SDE.....------------R-VN...-V---------HS--A--I-----------.--------...---H----D---N..-Q-----Q---S------------------TM--SNVTRQNS-SN-N......  143
CPZ.CM.05.SIVcpzEK505    -K-.........TEMQKN-LICCL-I-.LIK-IG....SE-.....------------RD-E...-V---------HS--A--I---Q------.--------...-LIP---R-D--K..-N--D--Q-------E--------------T-S-SSWRSVN-S..........  133
CPZ.CM.05.SIVcpzMB66     -ETQRNCRTLSLKEIILCTLVLGII-.IIK.-E....DNM.....------------R--D...-----------QNP-------SQ---S.-N---E-...IE-T----Y--A-E..-N---------VN-----V------I----T-N-SLF-CIKENGNTTNC....  146
CPZ.CM.98.CAM3           -K-M.......EKIKGSLIEKY-FM-.LIIP-LT.GSDQ-......-A------------D...-------E-S-LNK-A--I--SQ---.--------...-QIP---------..N-T-A---NT-------------------I--TME-RKVTFNSTS-RNKTST...  144
CPZ.GA.88.GAB2           -K-M....GKTRRRW...QPYCIIMA.LIIP-LKTESDQW.....----------------...-P---AN-T-LKE-P--I---Q----.---S-E-...-I-I----E--V-..D-A-----Q---T-------R-----N----Q-T--SINATG-ETNAT-......  143
CPZ.CM.98.CAM5           -KAM.......ERKKLN-IICY--M-.LITP-LT.GH-.W.....-A-----------E...-------E-S-LNK-A--I--SQ----.V--------...-YMP--S-S----..--N-A---Q---S------------------I--TMQ--T---IST--TN-......  140
CPZ.CM.05.SIVcpzMT145    -GM...KYYWWRNSCRSISIKLI-I-.WIAS-FG..E-NW.....-----------R--K...----------S-S--A--I---Q-----.---T-----...-L-P----E----E..-Y--D--Q-------E---------------T-T-NNPT-TSC-..........  140
CPZ.TZ.00.TAN1           -KNL...............IGIT-ILIIT-LGI.GFSTY.....YT--F---------Q...P---------DITSRDK--I----N---.L----Y-...-T-A--SIR---E..E-Y--QE-K---L--FQ--F---------F-IKMT-.TMT-T--KT.........  129
CPZ.CM.05.SIVcpzLB7      ------------.GT-TSCMPYYVV-.I-SL--L.IISSTKNENL--------------E...A-----------QKA-A--I--SQ----.---D-KA...IE-S----------T..-A--N--QQ-V-----------I--------T-H-SIP-F-NSS..........  140
CPZ.US.85.CPZUS          -K-M....EKKKRLW....LSYC--SSLII...P.GLSS-.....-A---------RDVE...-------------KQ-A--I---Q------.--------...-H-P----K-D---..E-N-A---Q--------------I--------TMT-LNPDSN............  131
CPZ.GA.88.GAB1           -K-M...--KKRD-NSLSIITIITII.L-TP-LT..S-.-.....------------HD-D...PV---------HS--A--I---Q-----.---S----...-F-P--I-S------..--N--D-------------------------T-Q-SKANFSQAK-LT........  142
GSN.CM.99.CN166          --KT.........VESMMMLAVS--AL-ALLIPG.LKCEN.....-T---------RD--...PP------PDIASN-PG-I-IST--L-.S--S-A-....-PL-I--K--IY..--Y--DEVRD-MV--FN-A----------M--GMY-NLTNTS-SEPTTT......  139
GSN.CM.99.CN71           -MR-.........VGFLKMLVVS-TLMA-L-PGS.RSKGN.....-T--F------RD-K...PP-------DITSN-PG-I-IST--L-.S--S-S-....-PL-I--E--IY..--Y--DEVRD-MV--FN-A-R---------M--RM--KLPNTS-T-TP-TTTA...  142
MUS_1.CM.01.1085         -PRM..........T-LRA.S-T--LTL-LSS-I.IWGRQ....----F-------D---...PP------NIANR-PG---I-T--L-.---A------...-QL-ISGEKFDIYK.-Y--D--TD-VR-IFN-A---------M--KM--AKVT-Y-RS..........  134
MUS_1.CM.01.CM1239       -PRM..........S.LMGM--T--LTL-LSS-I.IWGNQ....----F-------D---...PP------NIANR-PG---I-T--L-.---A------...-QL-ISGEYFNVYK.-Y--T--VD-VR-IFN-A---------M--KM--SRINVTRPS..........  134
MUS_2.CM.01.CM1246       -PAM........ERTRILTALSLT-ALS-SLILR.KHA-Q....----F-R---S-N---...PP-----N-NIANR-PG--LI-T--L-.---A------...-QL-ISNEHFNIYK.-Y--D--TD-VK-IFT-A---------M--KM--DL-NISMSS..........  137
MUS_2.CM.01.CM2500       -A-M..............T.SA-VFLTL-LS--L.IWG-Q....----F-------D---...PP------NIANR-PG---I-T--L-.---A------...-QL-I-GEYFDIYK.-Y--D--TD-VR-IFN-A---------M--RM---K-NSTRP-..........  130

                                        signal peptide end gp120 start
         Env start
                                                 *                                 *                                                        *       *       *
MAC.US.x.239             -GCL..........GNQL.....-IAIL-LSVYG.IYCT-.....Y---F----A-RN--...IP---.........-KNRDT-G-TQ-L-.DNGDYS-.....-AL----S-DA-N..-TVT--AI--VWQ-FET-I------S---ITMR-NKSET-RWGLTK-ITTTAS  132
H2A.GW.x.ALI             -MSS..........-NQL.....-VTIL-ASACL.VYC-Q....Y---F----A--N-S...IP---.........-KNRDT-G-IQ-L-.DNDDY--....IAL----A-DA-D..-TVT--AV--VWR-FET-I-----------IAM--SNISTEST-T-P-PGSTL.  131
H2A.DE.x.BEN             -EPG..........-NQL.....FVVIL-TSACL.VYCSQ....Y---F--I-A--N-S...IP---.........-KNRDT-G-IQ-L-.DNDDY--....IIL----A-DA-N..-TVT--AV--VWH-FET-I-----------AMN-SRVQGN-T-PNPRTSSSTT  132
H2A.SN.x.ST              KMCG..........-NQL.....FVASL-ASACL.IYCVQ....Y---F------RN-S...IP---.........-KNRDT-G-IQ-L-.DNDDY--....IAL----A-DA-N..-TVT--AV--VW--FET-I-----------AMR-NSTTAKNT-STPTTTTTA.  131
H2B.GH.86.D205           -AYF..........SSRL.....PIALL-IGI-G.FVC-Q....Y---F--I-A-RN--...VP-I-.........-TNRDT-G-VQ-L-.DNGDYT-....IRL-I--A-DA-D..-TVTQ-AVD-VWR-FET-I-----------AMN-SKTETNPGNA--TTTTKPT  132
H2B.CI.x.EHO             -AHV..........NNYL.....-VTLL-ISIYG.YMG-N....F---F--I-A--N-S...IP---.........-RNRDT-G-VQ-L-.DNDDYT-....IQL-I--A-DA-D..-TVTD-ATK-VW--FET-I-----------TM--NKTWSSASKETTTSSASLR  132
H2G.CI.92.ABT96          -AYL..........GNQL.....-IATX-LSTYG.FYCTK....Y---F--I-A-RN-S...VP---.........-KNRDT-G-TQ-L-.DNDDYS-....LAL----A-DA-D..-TVT--AV--VWN-FET-T----------IAMR-NKTETNKWGLTGTTTVTPA  132
H2U.FR.96.12034          -ACL..........GNQL.....-IALLCLNACV.IYCTQ....Y---F--I-A-RN--...IP---.........-NNRDT-G-TQ-L-.DNGDYS-....LAL----A-DA-N..-TVT--AV--VW--FET-I----------IAMR-SKTETNRWGLTGTAATTTS  132
MAC.US.x.251_1A11        -GCL..........GNQL.....-IAIL-LSVYG.IYCT-.....Y---F----A-RN--...IP---.........-KNRDT-G-TQ-L-.DNGDYS-....LAL----S-DA-E..-TVT--AI--VWQ-FET-I----------ITMR-NKSET-RWGLTK-STT.TT  131
MAC.US.x.251_BK28        -GCL..........GNQL.....-IAIL-LSVYG.IYCTQ....Y---F----A-RN--...IP---.........-KNRDT-G-TQ-L-.DNGDYS-....LAL----S-DA-E..-TVT--AI--VWQ-FET-I------S---ITMR-NKSET-RWGLTK-STTITT  132
MAC.US.x.EMBL_3          -GCL..........GNQL.....-IAIL-LSVYG.IYCTQ....Y---F----A-RN--...IP---.........-KNRDT-G-TQ-L-.DNGDYS-....LAL----S-DA-E..-TVT--AI--VWQ-FET-I------S---ITMR-NKSET-RWGLTK-STTITT  132
SMM.US.x.H9              -GCL..........GNQL.....-IALS-XSA-G.IYCVQ....Y---F----A-RN--...VP---.........-RNRDT-G-TQ-L-.DNGDYS-....LAI----A-DA-D..-TVT--AI--VWN-FET-I----------ITMR-NKSEXXRWGLTGXPAPTTT  132
SMM.US.x.PGM53           -GCL..........GNQL.....-IAIL-LSAYG.IYCTQ....Y---F----A--N--...IP---.........-KNRDT-G-TQ-L-.DNDDYS-....LAI----R-DA-N..-TVT--AI--VWN-FET-I----------ITMR-NKTET-RWGLTKNEIPTTT  132
SMM.US.x.SIVsmH635F_L3   -GCL..........GNQL.....-IALL-VSVLE.ICCVQ....Y---F----A--N--...IP---.........-KNRDT-G-TQ-L-.DNDDYS-....LA--I--A-DA-N..-TVT--AI--VWN-FET-I----------IAMR-NKTET-RWGLTGRAETTTT  132
SMM.SL.92.SL92B          -ACP..........GLHL.....-IDILFLSVLG.TWCAQ....Y--IF--I-A-RN--...IP---.........-QNRDT-G-VQ-L-.DNGDYS-....LAL----A-DA-D..-TVT--AI--VWN-FET-I----------IAM--NKNET-RWGLTRAATTTSS  132
STM.US.x.STM             -ACP..........GNQL.....-IAIL-LSACL.TYCTQ....Y---F----A-RN--...IP---.........-KNRDT-G-TQ-L-.DNGDYS-....LAI----A-DA-D..-TVT--AI--VWN-FET-I----------IAMR-NKNET-KWGLTGKTVTTVT  132
MNE.US.x.MNE027          -GCL..........GNQL.....-IAILFLSAYG.IYCIQ....Y---F----A-RN--...IP---V.......-RNRDT-G-TQ-L-.DNDDYS-....LAL-I--S-DA-E..-TVT--AI--VWH-FET-I----------ITM--NKSET-KWGLTK-STTTAA  132
GRV.ET.x.GRI_677         -GRL..........LIKI.......LIIAIGI-I.GIGN-....Y---F--I----NS-...VQA--M.......-PNT-M---TN-I-.D-HDNT-...-P-.-I--A-EA-D..-PL-K-AESN-HL-FE-TMR-----S-I-IKMS-VE-NGTAT-KATTTATTTM  130
VER.KE.x.TYO1_patent     --YT..........IITL.......-IIVIGIGI.VLS-Q....-I--F--I----NSS...VQA--M.......-PTTSL---TN-I-.D-HDYT-...-P-.-I--P-EA-GDR-PLIA-AASN-HL-FE-TM------S---IKMN-VE-NSTRERATTPTTTPKS  132
VER.KE.x.9063            -KPL..........LGIV....LIAIIGI.SSL..LS-RQ....----F--I----NSS...VQA--M.......-PTTRL---TN-I-.D-HDYT-...-P-.-I--P-EA-ADR-PL-A-AGSN-HL-FE-T------S---IKMS-VE-NSSRRNPTPTSASTTK  133
VER.KE.x.AGM155          -TKF..........LGIF....IV--IGI.GIGI.S-KQQ....-I--F------NSS...VQA--M.......-PTTRL---TN-I-.D-HDYT-...-P-.-I--P-EA-ADR-PL-A-AGSN-HL-FE-T------S---IKMN-VE--GSATSTPATSTTAGT  134
VER.DE.x.AGM3            -KLT..........LLIG.....I-LIGI.GVVL.N-RQQ....----F------NSS...VQA--M.......-PTTRL---TNSI-.D-HDYT-....-PL-I--P-EA-ADR-PL-A-AGSN-HL-FE-T-------S---IKMS-VE-NSSEP-TTPKSTTAST  133
TAN.UG.x.TAN1            -GPL..........-GKG...VL-VILG-SLIGL.LYGTQ....YI--F--I----NSS...VQA--M.......-PNT-L---TN-I-.D-HDYT-...-Q-.--S-K-EA-KDR-PL-A-AESN-HL-FEST---------M-IKMN--K-TSTAP-S-TPTSSSTT  136
SAB.SN.x.SAB1C           -KLL..........TVLL...WLSGCWS-VWLVQ...........Y---F--I----NSS...VQA--K.......-PNT-L--STN-I-.D-EPEGTIAE-PIP-I--K-DA-KNR-PL-G-AESN-HL-FEST-------S-M-IKMN-YR-EGGAA-TT-PSTSTAR  135
OLC.CI.97.97CI12         --K-..................LV-CLLWLSILGISIG.Q....Y---F--T-K-RDSN...D--M-V.......-ANKSL-V-T--L-.SFDEDEI...EQPIPGV-EVFYNVSD-EV-R-AWFG-A--L-AV---------F--KMQ-E-PSANFTEPITPISTAAP  132
DRL.x.x.FAO              -IRH..........I......IIG-CILGFLGIS.IG.-H....----F--T-K-RP-S...-H-I--.......-DN-SF-V-TS-I-.SLLHYE-...TEIN-IE---TVPMTE-EVIK-AWGALS-MI-AV------IN-Y--KME-S---IVNAASTTPTPTTVA  133
RCM.NG.x.NG411           -AKL...............ISIGIII-...-LGIII.VG--Q....Y---F-------NS-...VPM---.......-DNAQS-G-LN-I-.EGGISP-....-P---S-K-DA-E..-SLY--AKSNVWH-Y-ST-----R---M-ITMY--A-NGSWDGTATQSPTTTT  131
RCM.GA.x.GAB1            -DK-.............LVIVLIVVI-..IILVQG.SQKPQ....YI--F------RNS-...VPM--V.......-DNTQS-G-LN-I-.EGGISP-....-SI--S-R-DA-N..-SLY--AKDNVWN-Y-ST-----R-S---ITMN-SAINGSWDGIPT-APPTTT  133
MND_1.GA.x.MNDGB1        -KCP..........GLFL.....I-CLA#LV..G.IIG-Q....Y---F---------K...-H-I--.......-DNSSL-V-TN-I-..SLPDYDE..-EIPDIK---TGLIRE-QI-Y-AWHAMG-ML-TI------IN-Y--KMQ-QETE-VSA-TAKPITTPTT  132
MND_2.x.x.5440           -LRY...................LRYIVLGI-V-V.IVGEQ....-------T-K-HK-R...-H---.......-DNNSF-V-TS---.SLLHYE-...QHIP-I----TGPIEE-EI-T-AWGA-S-MI-AV-----R---Y--KM---EGQ-E-EQATAKTTTPVP  132
MND_2.CM.98.CM16         -LRY......................IILGI-VGL.GLGNQ....-------T-K-HK-E...-H---.......--NDSL-V-TS---SLLHYEEQ..PIPNITWNF.TGPMEE-EV-M-AWGA-S-MI-AV---------Y--KMQ--KG-DLATPIPTTSTTTTT  129
MND_2.GA.x.M14           -LRH.....................-ITAIV-IGI.GIG-Q....-------T-K-HP-R...-H---.......-DNNSL-V-TS---.SLLHYE-...QLIP-I----TVPIQE-EV-R-AWGA-S-MI-AV---------Y--KMQ--KGE-SSTSTTTPMPTTTT  130
DEB.CM.99.CM40           --LI..........QLLL...LI-GLIVIGTQGK.QKKQQ....YL---------VD-K...VD---TANSSE.....SG--VT--L-HALVREEVPMPN-TQ-.FNA-DNP.....IE--LWQ-MT--YK--F--------Y---MQ-IKTSTNPTPTNT-TTTTTI  136
DEB.CM.99.CM5            -LSL.................LIIIASLTLVGAA.P-KQQ....-----------HN--...VG---TANSSE.....SG--VT--L-HAMVRE-VPMPN-SQ.YFDS-SNV.....-EQ-LWD-MT--YR--F--------Y--RMN-SSTTTT-TAATTTI.....  127
DEN.CD.x.CD1             -SQV......RRTL....KVLLLI-W..TIGIIC.T---Q....Y---F--I-A-RNSA...AP---T.......AMDRG---AAS---.E--YQR-VWMENATDYFDIYHNP...IVDQVVQNMKD..-YK-TF--------V--RMN--RKEDNSTNAVTVSSTTTT  136
TAL.CM.01.8023           -KCL..........QVLCLAICVIWVANIGKSLT.SA-QQ....FL----------N---...PP---.......-MARDQ-S-SQ---.S--QNL-.....-P--L--Y-DVY..--Y--DHIAD-MDA-FLTAT---T----M--RMA---ITKNLTGTVAPPTTNST  137
TAL.CM.00.266            -KYL..........AVI........FCLAYLGRF.IIGNQ....Y--I--------N---...PP---.......-MARDQ-S-SQ---.S--QNL-.....-P--L--Y-DVY..--Y--DHIA--MD--FLTATR--T----I--RM--E-ITKNLTISANTTTPAPT  129
MON.CM.99.L1             --KL..................-A-STLCLMMTL.VNGSN.....-T---------P---...PP------PNYGSK-AG-N-LASS-L-.---T--S....LYL-I--E--AYQ..-Y---E-V--MK--FS-A---------M--RML-VEVNTVS-ASTTPAPSTPT  136
MON.NG.x.NG1             --MR.......MKTIIVLTILCLMM-.T-TSG-X.SGN-S....----F-----Y-P---...PP------PXYGSK-AG-N-LASS-L-.S--V-RA....LPL-I--S--AYN..-Y---E-V---K--FS-A---------M--RMR-QXINES-TVXPTTTXX.AT  145
SUN.GA.98.L14            --CP...................EILIGFSLLLG.SIAIQ....Y---F--T-K-EP-V...VP-I--.......SANNSL-V-TS-L-.DLQTYA-....-PITGLE---TEGISN-QI-Q-AWQAMT-MV-AIM-----IN-Y--KM--PPKPTTPSS--TVKSSCDY  131
WRC.CI.97.97CI14         -IF-....................-VLG-IGIAI.SLAVQ....Y--------Q-EDS-...DF-L--.......--NRSL-VSTR-I-.HLEHLE-...QYMPEIK---TRKVAE-QI-Q-AWTGVT-MI-TL------I--Y--KMQ--PGEGSETF.....DPCES  125
WRC.CI.98.98CI04         -LRF...................ILISIV-IGL.GIG-Q....YI-------Q-E-SK...DF-L--.......--NRSL-VSTH-I-.HIEHWE-...QY-P--K---TRKVAD-QIIQ-AWTGVT-MI-TI------I--Y--KMT--PGAP-S-NTKDF.NPCEI  129
WRC.GM.05.Pbt_05GM_X02   -NTF...................IFIAVVIGIGL.GLG-Q....YI-------Q-E--K...DFML--.......--NRSL-V-TH-L-.HIEEIT-...QYMEGIK---TKKVEE-EV-Q-AWTGVT-MI-TL------IN-Y--KMS--LGDEKS-YTEDIELTTVK  131
LST.CD.88.524            -ACP..........GNIL.....--LLC-GLTAG.EK--Q....Y---F----K-ED-V...VP-I--.......SANNSL-V-TS-L-.DLQSYAQ...IPIH-I-L--TQEIKD-QIIQ-AWGAMT-MV-AIM---T-IN-Y--RM---GEGLGKKRTTTTTPTTT  135
LST.CD.88.485            -DCL..........KLLL......FLLF-GAQGE.-.K-Q....Y---F----K-ED-V...VP-I--.......SANNSL-V-TS-L-.DLQSYAQ...IPIH-ISL--TQEIKD-QIIQ-AWSAMT-MV-TIM-----IN-Y--RM---GGGIGPK-TP-PIGKTTT  133
LST.CD.88.447            -NCP..........GLIL......-LII-RVQGE.K-K-Q....Y---F----K-E--V...VP-I--.......SANNSL-V-TS-I-.DLQSYAQ...IPIH-ISL--TKEIKD-QIIQ-AWSAMT-MV-TIM-----VN-Y--RM--EGEGIG-TKATKTKT.STT  133
LST.KE.x.lho7            -ACP..........GL-........ILL-LLGII..WG-Q....Y---F----N-DDNVS.VP-I--.......SANTSL-V-TS-L-.DLQSYA-...-PIY-IS---TIPVKD-QVIQ-AWSAMNAMV-SIM-----IN-Y--RMQ-GEV....TKTPTTTPKTTT  128
SYK.KE.x.SYK173          -AAF..........-TYI.....VCLFS-ISLGF.MEKQQ....Y---F--I-H-ED-Y...AP---T.......-SHKGG---KN---.SADQIEV...RVNITG-Y-PA-NSSHMIRQ-IL--MSA-FL-ANR-----A-M-IRML--LDNSPATS..........  124
SYK.KE.x.KE51            -KLF..........NIIS....VV--IV-SLGIV.-RLEQ....Y---F----N-ED-Q...VPM---.......-PHTGG---KN---.SSEQIEV...RVNISG-Y-TA-NSSHGIRQ-LLQ-MSD-FL-ADR-----S-M-IRTL-VED-DSSSG-LTTTI....  131
COL.CM.x.CGU1            -LRY..........LFIN.........L-VLGIV.L.SNK....-----Q---A-E--DVNDQQF--F-SSPE.......IQQVLG-L-PPPGK-V-...QNMP----A-DLF.--SFSGEVWIITQTTLE-R-R--A---AY-APMI--KVNRTE-GT-........  124
```

```
                                            glycosylation NIS
                                         glycosylation NCS                                                                          glycosylation NVS
                         V1                                          V2                                                          glycosylation NGT
                                                                glycosylation NDT           glycosylation NTS              glycosylation NKT                    glycosylation NGS
                                                   *                                                   *        *          *          *      *
H1B.FR.83.HXB2            ..................................M.IMEKGEIKNCSFNIS.TSIRGKVQKEYAFFYKLDIIPIDNDT............TSYKLTSCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNNKTFNGTGP..CTNVSTVQCTHGIRPVVSTQLLLNGS 264
H1A1.UG.85.U455           ...............................IT.DGVRE-M------MT.-EL-D-K--V-SL--R---VQ-NKTDN...........N--R-IN----T---------------------------KDPE---K--..-R-----------K------------ 262
H1B.US.90.WEAU160         G..............................E.K--E--M------VT.-L--N-RKT---L-----VM---H-N.............---T-IN-KS-T---------------------------D-K---K--..-K------------------------ 268
H1C.ET.86.ETH2220         ...............................I.NSAND-M-------T.-EL-D-KR-A--L------V-LN-GS.............-D-R-IN----T---------LD-----------Y-----RD---T----..-H-----------K------------ 261
H1D.CD.84.84ZR085         ..............................LP.TVKP--M-------T.-VVTD-RKQVH-L--R--VVQ---EGKNEIN.....DTYGT-R-IN----A-----------------------------D-R------..-K---S------------------- 271
H1F1.BE.93.VI850          ......................................MQ-----MT.-EV-D-KL-LS-L--R---V--G-NNS............SE-R-IN----T---------WD-----------Y------D-R------..-K-----------K------------ 255
H1G.SE.93.SE6165          TNN............................ST.VDNP---------VT.-E--D-KK-------R--VV--N-A..............-N-R--H--V-T-K------T-D----------T-------RD-E------..-K-----------K------------ 269
H1H.CF.90.056             .................................MEAG--LT-----VT.-VL-D-Q--VH-L--R--VV----NS.............-Q-R-IN------------------------------------------L..------------------------- 257
H1J.SE.93.SE7887          ................................VSSPDIMT------T.-E--N-RKQ---L--RQ-VV---SNN.............KN-I-IN------K--------Q-----------------D-N-----S..-K-----------K------------ 258
H1K.CM.96.MP535           ................................VVSPA---------T.-E-KD-KK--S-L--R--VL-LNGEGNNSS.........-E-R-IN----T---T----T----------------KD-R------..-K-----------K------------ 266
H101_AE.TH.90.CM240       ...............................II.GNITD-VR--T--MT.-ELTD-K--VH-L------VQ-EDKKTS..........SE-R-IN------K----I--D------T---Y------D-N------..-K---S-------K------------ 267
H102_AG.NG.x.IBNG         .................................DMN---------T.-EV-D-KK-MH-L--R--VVQ-NENNG............SQ-R-IN----A--------------------------KD-G------..-K-----------K------------ 261
H103_AB.RU.97.KAL153_2    ...............................NTS.SIKMM-M-------T.-DL-D--K----L-----VVQ-..............D--R-I------V-------I-------------------D-K-------..-----------------K------------ 255
H104_cpx.CY.94.CY032      .................................VI-EG------D-T.-E--D-KK----L--RI--V--NARVPING..SNRNNSTEE-M-IN--A-T-K---------------------------E-N-T-L--..-----S-R-----K------------ 269
H1N.CM.95.YBF30           ...........................MTTREP.DIGYKQM------AT.-ELTD-KKQV-SL--VE-VV--NA.............YNKT-R-IN---TAV------T-----------P----M---EGN-S-N-S..-------------K--I----I---- 267
H1O.BE.87.ANT70           ......................................NLM-K-E--VT.-V-KD-KE-KQ-L--VS-LMELNETSSTN.....KTNSKM-T--N--STT------------------Y--F---STE-----T..-R-ITV-T-------T-----I---T 253
H1O.CM.91.MVP5180         ...............................LL.NETIN-MR-----VT.-VLTD-KEQKQ-L--VS-LSKVNDSNAVN.........G-T-M--N--STI-K-----------------T-Y--F---DTD-----L..-H-I-V-T-----K-T-----I---T 257
CPZ.CD.90.ANT             ..........................VTPKTTTP.-VDGMKLQE-N--Q-.-GFKD-K--MK-I---G-LMKCQDNNETN..........C-Y-WH---TT---S-E-ST-------------Y---R-EDED-T-V-M..-K---V-H-----S-M-A-W-----T 264
CPZ.CM.01.SIVcpzCAM13     ....................NNPTLKNNTSGKN.ETGVLQM---T--TT.-EL-D-KKQV-SL--VD-LQSLGS............GTGDT-TMIN---TA-----------------------------DVN-S-K-K..-R-----H-----K---T---II--- 274
CPZ.CM.05.SIVcpzEK505     ...............................VN.QTNHVQMQ-----VT.-EL-D-KKQV-SL--MG----L-TNNSS........GNNSQ-R-IN---TAV------I-------Y----P----I---DQD-----E..-N-----------K--I----I---- 257
CPZ.CM.05.SIVcpzMB66      ...................TVQISTGNDSTANNI.TVGTIDMY-----AT.-EL-DRKKQV-SL--RQ-LE-LEGNKPPE......GDKNAL-R-YN---TAM----S---L-----------------L----D-N-T-I-Q..-K-----H-------------F----T 285
CPZ.CM.98.CAM3            .........................MTTNSPN.EKXDSTV---T--MT.-EV-D-EK-V-SL--VD-LVL-----D............T-R-IN---TA-------T-----------TP-Y--M---MPN-----TGR-N-I-------------T---I---- 270
CPZ.GA.88.GAB2            ............................NGIEK.DGLAQDMR--T--TT.-EL-D-K-QI-SL-W-N-LVGTN-.............TFR-IN---TA-------T-----------------L----D-DYP-K-K..-K-----H-----K-T-T---I---- 262
CPZ.CM.98.CAM5            ..............................TEL.KNI-EIM---T--MT.-EV-D-KKEV-SL--VD--V--GGNKSND...........T-R-IN---TTV------T-------------P----M---DQK-------..---I-------------T--VI---- 262
CPZ.CM.05.SIVcpzMT145     ..............................NST.DDRL-DMR-----VT.-EL-D-KRQV-SL--VE--TA-G-.............NST-R-IN---TA-------T-----------------L----DIDYK-NET..-K-----H-----K--AT---I---- 259
CPZ.TZ.00.TAN1            ...................LNSATTTLTPTVNLS.SIPNY-VY-----QT.-EF-D-KKQI-SL--RE--VKE-GN.............NN--Y-HN--------E-D-ST------R---------L---RDQN-T-K-Q..-S---V-H-----Y-MIA-A-H---- 261
CPZ.CM.05.SIVcpzLB7       ................................I.NSSNETM------TT.-EL-D-KKQT-SL--VD-LTQ-NK............TTE----IN---TA------I-----------------L-R-----Y--S--..-N-----H-----K--I----I---- 258
CPZ.US.85.CPZUS           ..............................SSA.VNTTDIMR------T.-EL-D-KKQV-SL--VD-LAH-N-N..............T-R-IN---TA-------T-----------P--------E-D-K-K-E..-K-----------K---T---II--- 249
CPZ.GA.88.GAB1            ..............................NQT.SSPPL-M------VT.-EL-D-KKQV-SL--VE-VVNLG-ENN...........T-RIIN---TA-------T-----------------------D-D-S-K-K..--------H-----K---T----I---- 262
GSN.CM.99.CN166           .................PRPPNVSTTTQWGSWG.GE.NGTGQPLY-----QT.-EF-DQKKQM-SL-WRE--MEETHG............NQSG-YIRN----Y---K-V-S--Q-V-------P-Y-M----DVN-T-V-T..-Y---A-T-----Q-L-A-W-H---T 275
GSN.CM.99.CN71            .................TTSPSSPTPTPWGNWG..G.NGTGQP-Y-----QT.-EF-D-KRQM-SL-W-D--MRAQDG............ND-D-YIIN----YV--K-V-T--Q-V-------P---M----DAN-T-V-K..-S---A-T-----Q-L-A-W-H---T 277
MUS_1.CM.01.1085          .............TGGNSTGTTAPPTTPWGSWGNGN.DSAGQPMF-----QT.-EF-D-KRQM-SL-FVD-LMRDSSESEND..............--YI-N----Y-----E-S--Q-V--Q------YSL----DAN-T-Q-E..-N--TARH-----L-L-A-W-Q---T 276
MUS_1.CM.01.CM1239        .............TTTTASTPTKTTPVTPWGKW.DA..NGTEQSMF-----QT.-EF-D-KKQM-SL--TD-LMRE----------------.T-YILN----Y---Q-E-S--Q-V--Q------YSL----DAN-T-Q-K..-N--TATH-----L-L-A-W-Q---T 271
MUS_2.CM.01.CM1246        .............TTAASATTSGHTTVTPWGRWADN..NETQLTMY-----QT.-EF-DVKKQM-SL-FVK-LMKGE-E..............T-YI-N----Y---Q-E-S--Q-V--Q------YSL----DAN-T-Q-I..-N--TARH---D-L-LIA-W-Q---T 275
MUS_2.CM.01.CM2500        ..................TKPTTPSTTPWGRWSDG..NDTMQVMY-----QT.-EF-D-KKQM-SL-FVD-LMRDGS-..............DG-YI-N----Y---Q-E-S--Q-V---------YSL----DAN-T-Q-K..-R--TATH-----L-L-A-W-Q---T 263

                                            *          *          *                      *        *          *          *      *
MAC.US.x.239              TTSTTA..........SAKVDMVNETSSCIAQDNCT.GL-QEQMIS-K--MT.GLK-D-KKEYNETW-SA-LVCEQ.........GNNTGNESRCYMNH------QES-D-HYWDA-RFR----P-Y-L-R--DTNYS-FM-K.-SK-VVSS--RMMETQT--WFGF--T 280
H2A.GW.x.ALI              ...................KPLINESDPCIKADNCPRGLGDE-MV--R--MT.GLQ-D-PKQYNETW-SK-VVCEP.........FNTTTNQ-RCYMNH-------ES-D-HYWDA-RFR----P-Y-L-R-DDINYS-FA-N.-SK-VAAT--RMMETQT--WFGF--T 271
H2A.DE.x.BEN              SRPPTS............AASIINETSNCIENNTCA.GLGYE-MMQ-E--MK.GLEQD-KRRYKDTW-LE-VVCDN.............TTAGTCYMRH----I-KES-D-HYWDAMRFR----P---L-R--DTNYS-FE-K.--K-VAAS--RMMETQT--WFGF--T 274
H2A.SN.x.ST               ...................NTTIGENSSCIRTDNCT.GLGEE-MVD-Q--MT.GLE-D-KKLYNETW-SK-VVCES..........NDTKKEKTCYMNH-------ES-D-HYWDTMRFR----P---L-R--DTNYS-FE-N.-SK-VAAT--RMMETQT--WFGF--T 269
H2B.GH.86.D205            TTS..............RGLKTINETDPCIKNDSCT.GLGEE--MQ-N-SMT.GLR-DELKQYKDTW-SE-LECNN...........TRKYTSRCYIRT---TI-QES-D-HYWDSLRFR----P--FL-R--DTNYS-FM-N.-SK-VASS--RMMETQS--WFGF--T 274
H2B.CI.x.EHO              SSTQ................TLLNEDSKCIQNDSCA.GIGLE-MID-Q-KMT.GLK-DESKQYKDTW--Q-LVCEK..........GTRSNESKCYIKT----I-QES-D-HYWDSLRFR----P---L-R--DTKYS-FM-N.-SK-VVSL.YRMMETQT--WFGF--T 272
H2G.CI.92.ABT96           TVPTTK..........MVTAELVNSTSQCLMYDNCT.GIQSESMVG-K--MT.GLK-DQRKEYNETW-SQ-LQXEQ.........SNKSENESRCYMRH------QES-D-HYWDA-RFR----P---L-R--DTNYS-FA-N.-SK-VVST--RMMETQT--WFGF--T 280
H2U.FR.96.12034           PAMQESS.............KVINETEPCIRNNSCS.GL-QEPLVS-R--MT.GLK-D-K-EYKETW-SH--VCEQ..........NTSESESKCYMNH------QES-D-HYWDA-RFR----P-Y-L-R--DTNYS-FM-N.-SK-VVST--RMRGTQT--WFGF--T 276
MAC.US.x.251_1A11         AAPKAM..........SEKINMINETSSCIVHDNCT.GL-QEQMIG-K--MT.GLK-D-KKEYNETW-SA-LVCEQ.........GNSTGNESRCYMNH------QES-D-HYWDA-RFR----P-Y-L-R--DTNYS-FM-K.-SK-VVSS--RMMETQT--WFGF--T 279
MAC.US.x.251_BK28         AAPTSA........PVSEKIDMVNETSSCIAQNNCT.GL-QEQMIS-K-TMT.GLK-D-TKEYNETW-ST-LVCEQ.........GNSTDNESRCYMNH------QES-D-HYWDT-RFR----P-Y-L-R--DTNYS-FM-K.-SK-VVSS--RMMETQT--WFGF--T 282
MAC.US.x.EMBL_3           AAPTSA........PVSEKIDMVNETSSCIAQNNCT.GL-QEQMIS-K-TMT.GLK-D-TKEYNETW-ST-LVCEQ.........GNSTDNESRCYMNH------QES-D-HYWDT-RFR----P-Y-L-R--DTNYS-FM-K.-SK-VVSS--RMMETQT--WFGF--T 282
SMM.US.x.H9               QTSTTPPS......PI..IAKVVNDSDPCIRSNNCT.GL-QEPMVS-K--MT.GLK-D-KREYNETW-SR-LVCEQ........NSNGNETDSKCYMNH------QES-D-HYWDA-RXR----P-Y-L-R--DTNYS-FA-N.--K-VVSS--RMMETQT--WFGF--T 283
SMM.US.x.PGM53            TTTTTKSPKA.....EAITAKVINESDPCISNNNCT.GL-QEPMVS-R--MT.GLKKD-KKEYNETW-SR--VCEQ.........STTSRNESKCYMNH------QES-D-HYWDA-RFR----P-Y-L-R--DTNYS-FA-N.-SK-VVSS--RMMETQT--WFGF--T 285
SMM.US.x.SIVsmH635F_L3    AKSTTSTTTTT......VTPKVINEGDPCIKNNSCA.GL-QEPMIG-K--MT.GLK-D-KIEYNETW-SR-L-CEQ..........PANGSESKCYMQH------QES-D-HYWDA-RFR----P-Y-L-R--DSNYS-FA-K.-SK-VVSS--RMMETQT--WFGF--T 283
SMM.SL.92.SL92B           PTTTSPLTAAS.....PSGEEIVNDTMSCTKNNNCS.GI-QEPMIG-Q--MT.GLK-DQKRQYNETW-SR-LVCEQ...........GGNESSRCYMNH------QES-D-HYWDA-RYR----P-Y-L-R--DTNYS-FA-N.-SK-VVSS--RMMETQT--WFGF--T 283
STM.US.x.STM              PTAAAAAT..........KPELVNETSSCVSNNNCT.GL-EESLVG-K--MT.GLK-D-KREYNETW-SS-L-CEQ..........NVTGEESRCYMRH------QES-D-HYWDALRFR----P-Y-L-R--DTNYT-FA-N.-SK-VVSS--RMMETQT--WFGF--T 279
MNE.US.x.MNE027           PTTKTTTT.........KEIEVVNENSTCVNRDNCT.GL-QEPMIS-K--MT.GLK-D-KREYNETW-SA-LVCEQ.........GNSTEDESRCYMNH------QES-D-HYWDA-RFR----P-Y-L-R--DTKYS-FM-N.-SK-VVSS--RMMETQT--WFGF--T 281
GRV.ET.x.GRI_677          T............TPCQN..........CSTEQIEG.E-AEEPAS--T-A-A.GYQ-DVKKNYSMTW-DQELVCNNKTGSEKGSK........DCYMIH--D---KE--D-TYWDTLRVR------Y-L----D-DYR-FA-K.-K---V-H--RL-NTTIT-GIG---- 267
VER.KE.x.TYO1_patent      TG...........LPCVGPTSG.ENLQSCNASIIER.E--DEPAS--T-AMA.GYV-DQKKNYSVVWNDAE-YCKNKTNSTSK..........ECYMIH--D---KE--D-TYWDQLRLR------Y-L----DEDY--YKQN.-S---V-H--GLMNTT-T-G------ 277
VER.KE.x.9063             APKTG........DPCIKSTNNNVNLQPCNASLIEE.EL-EEAAS--T-AMA.GY--DQKKNYSVVWNDAE-YCKNSSSSNSSTK........ECYMIH--D---KE--E-TYWDELRLR-------L---HDHDY--YKQN.-S---V-H--GLMNTT---G------ 284
VER.KE.x.AGM155           K............LPCVRNKTD.SNLQSCNDTIIEK.E-NDEAAS--T-AMA.GY--DQKKNYSVVWNDAE-FCKRSTSHNGTK.........ECYMIH--D---KE--D-TYWDELRLR------Y-L----DWDYA-FK-E.-S---V-H--TLMNTT-T-G------ 279
VER.DE.x.AGM3             TNITASTT....TLPCVQNK.TSTVLESCNETIIEK.ELNEEPAS--T-AMA.GYV-DQKK-YSVVWNDAE-MC...........KKGNNSNRECYMIH--D---KE--D-TYWDELRLR-------L----DYDYA-FKTN.-S---V-H--NL-NTT-T-G------ 284
TAN.UG.x.TAN1             .............DPCPNT.....DESSCNATLVTN.S-DYENSSI---AMA.GYR-DVKK-YNSTW-DQELVCEKENNTTGTR.........GCYMIH--D---KE--E-TYWDTLRLR------------KDTNYT-F-V..-R---V-S--GLMNTT--SAFGI--- 275
SAB.SN.x.SAB1C            .....................PEVVSVGFNDSVIEQ.E---EQAM----AMA.GYR-DVKKNYSTVWDDQEVVCEEGREKSNAT......HTVGCYMIH------KE--D-TYWDTFRLR------Y-L-R-ADTDYS-HKA..-R--TVSA--RL-NTT---GIGI--- 274
OLC.CI.97.97CI12          .............TTVTSSTGGQTTTIGLGVDNSTV.KP-TEKL-Q-RY--T.GLC-DCKIEMNQT-DLE-VACTN-................TCYMSH--------D-QYAVSDQQMVKL---P-RVLI----R-E-P......-R-I-A-T--PPLSAT--SFFGF--T 266
DRL.x.x.FAO               TTVNTS............VASTTTANDTLEIAVNNT.EPVMETNRV-KY-VT.GLC-DCKEEITEN-RYD-VVCKCGNETESPESNNCTNKKRTCYV----ST----D-N-A-TDMMKFRL---P-YIL-R--E-LNISSK...-K-ITA----RHMPATI-SHFGF--T 286
RCM.NG.x.NG411            RATTRTT..KEIGKDCYAN......QTECS.......SVPDADVM--E-TVA.GLK-DEKM-VNDTW-SR-LWCKA...........GNNNTRQCFIRH---TS-Q-F-EPKYW--FRLR----P-Y-----LDANYT-FDT..-Q--TATS---M-NTT---SFG---- 272
RCM.GA.x.GAB1             KTTTQRT..IGVEKECTAG......NETCE.......EVQDADVMS-E-AVA.GLK-DEKH-YNDTW-SR-LWCEK.......ETNSTNSTKKKCFVRH---TS-Q-F-EPKYW--FRLR----P---L-V-KD-NYT-FDT..-V--TATS---M-NTT-ASGFG---- 278
MND_1.GA.x.MNDGB1         TSTVASST.................EIYLDVDKNNT.EEKVERNHV-RY--T.GLC-DSKEEIVTN-RGD-VKCEN................NTCYMNH--E--N-ED-Q-GLLIR.C-LG-V-P-YVM-RY-E-LN-NKL...-S-I-A----QHLVAT--SFFGF--T 263
MND_2.x.x.5440            TTTTPSTT........TSSSTNKTTTPVLVVEKQNN.ETTTQQNRV-K--TT.GLC-DCKLEIEEN-RYE-VTCTKLNKTGS..ATNSTEPEYECYM----AT----DRN-A-TDRMTFRL---P--VL----E-LNKTKL...-G---A----APLPATI--MFGF--T 287
MND_2.CM.98.CM16          TTT.................KTVANTTELDIDTNNT.ETTTQQNRV-K--TT.GLC-DCKLEIEEN-RYE-VTC...........TNGTNNTY-CYM-Q--------D-N-A-TDE-KFRL---P-YVL-R-LERLNVSKK..----ITA----QPLPATT--MCGSS-T 266
MND_2.GA.x.M14            PTTP...................AVTGLEVTVQNN.ESTTQ-NRV-K--TT.GLC-DCKIEIKES-RYE-VTCT.........KKENDTETEECYM-H--------D-N-A-TDKMTFRL---P-YIL-R-RE-LNKTKL...-A---A----DPMPATI--MFGF--T 266
DEB.CM.99.CM40            ATTT...................KTTTD...WSGE.NITMTQYW-----V-.GPY-D-KE-SS-VWLED--QWA--KDGSGNR.........TGYMKH--D-----S-ETSR-K-FK-R------YGL-R-DD-N-----L..-N--TA-A--NL-HTMA--WVQF--- 270
DEB.CM.99.CM5             ........................KTTTD...WSGE.NKTVTQYW--T--VT.GPY-D-KEMQS-VWLED--QWETENQTNTNK......SIRHGYMNH--N-----S-ETSR-K-FK-R------YGL-R-DDRN-----L..----TA-A--NL-HTMA--WMQF--- 260
DEN.CD.x.CD1              TTTI..................KTPTTE...WYG..GNKTTIFLK-N--MT.-GKGFRDKRVMYEA.....HCRIE-VME.....DRETNQ-Y-YMRY--Q-T-----E--V---FK-RM-------L----DHPWS-Y-L..-K-M-ARV--DE-HTMA--W-IM-ST 269
TAL.CM.01.8023            TNSVTS.................KRTNDTNVWWG..NNATEPVY--TW-MT.GGF-D-KYTYFSHW-LA-LMKEPGNGN...........DSY-YA-G--L---A---E-THYQAF--Q--------LI---DNPWT-R-K..-H-I-A-H--DP-TTIAA-W-V---T 273
TAL.CM.00.266             .......................AMSNSNMPWWG..DNTTEPVY--TW-MT.GAF-D-RYTYFSHW-LA-LMKEPGNGS.............F-YA-G--V---A---E-THYQAF--Q--------LI---DSPWT-R-K..-H-I-A-H--DP-TTIAA-W-I---T 257
MON.CM.99.L1              P.............................WGNWGG.NGTGQPVY-----QT.-EF-D-KKQM-SL-W-E--MKEEGSNG............SH-YILN----Y-----E-SNY--V-L-----P-Y-L-R-DDPA-T-Q-S..-S---A-T---A-Q-I-A-WFQ--ST 261
MON.NG.x.NG1              TP............................WGSWGG.NGTGQPLY-----QT.-EF-D-XKQM-SL-X-E--MQDK...........DNTTDDY-Y-LN---XYX----E-SNY----------P-Y---R-DDPX-T-QNN..-X---A-T-----M-I---WFQ--AT 272
SUN.GA.98.L14             WTTTTAKT.TTQTTSSTSSTASTTTP.MPLDWNCTDTENIAESN-V-KY-VT.GLC-DCKTEVEQN-RDTEVTCNG--..............TCYM-H--D-I---D-H-GIMQNAYFRL-----YML-R-DE-LNATKK...-K-ITATP--NYMTST--SFFGF--T 280
WRC.CI.97.97CI14          NTTRA......TTSATTTVGTVSTTPVMQVENNCSQ.VNTSEINHV-R--TT.GMCKDCRIEIKES-RGD-VVCNN-S..............ECYINH--E---K-D-Q-GWQDEQY-RM---P-SVL----E-LNSSKL...-E---A-M--DMMTATI-SFFGF--T 269
WRC.CI.98.98CI04          STGAPTISTISAAPTTSTISAGTRTPAMPVDNNCTT.VNTSEINHV-R--TT.GMCKDCRIEIKES-RGD-VVCNK-G-.............ECYVNH--E---K-D-Q-GWQDEQYVRM---P-SVL----EPLNSSKL...-E---A-L--DMMTATI-SFFGF--T 280
WRC.GM.05.Pbt_05GM_X02    .............TTTPVTPPQNTTPSMDVENNSTS.LNLTDTNHV-R--VT.GLC-DCK-EIVES-RGD-VRCTQHG..........NGTQTCYINH--E---H-D-Q-AWHDEQY-RM---P-YVL-R--E-LNSSKL...-E---A-T--DLMTATI-SFFGF--T 272
LST.CD.88.524             PMPCFIT..ESTTTTTKSSGTEQPSLIEDTNCTRWL.NETTEINRE--Y-VT.GLC-DCKEEIKQN-RYDETTCNE-G..............TCYM-H--D-I-L-D-N-GVMSNAYFRL-----YML----E-LNFSAN..----ITATP--DYMISS--SFFGF--T 283
LST.CD.88.485             PMPCFIA..ET.TTTTTKGNTTMESLIEDTNCTRWL.NETTEINRE-KY-VT.GLC-DCKEEIKQS-RYDETTCDK-G..............TCYM-H--D-I-Q-D-N-GVISNAYFRL-----YML----E-LNFSAN..----ITATP--DYMIST--SFFGF--T 280
LST.CD.88.447             PMPCFIT..EKTTTTTKRSTTGGPELIEDINCIKWV.NETTEINRE-RY-VT.GLC-DCRTEIKQS-RYDETTCDK-G..............TCYM-H--D-I-Q-D-N-GVISNAYFRL-----YML----E-LNFSAN..----ITATP--DYMISS--SFFGF--T 281
LST.KE.x.lho7             QMPCFIN..EQV...TVKNPGNETRLEEDLNCTRGL.NETTERNAE-QY-VT.GLC-DCRTEIKQS-RYD-VTCSGER...........ENRTCYM-H--D-I---D-N-GVMQNAYFRL-----YML-R--EQLNFSKK...-E-ITATP--GYMLSS--SFFGF--T 276
SYK.KE.x.SYK173           ..................TPTTSPPTTPPNETWWG.DNSTEPRF-----LT.GGFKD-K-QYR-----D-LMKEEGN.............SSY-Y-LH------SA--E-QT-Q-F--Q----P-YSL----DTN-E-DDV..----TA-S--QEFNTLA--WFQ---T 258
SYK.KE.x.KE51             ..................SPTTTTNKAATPPAWWG.DNSTETRY-----MT.GGFKD-KHQYN-----S--MREEGN.............ESY-Y-QH-----MRN--E-QT-R-F--Q----P-YSL----DTE-E-DAE..-Q--TA-S---PFNTLA--WFQ---T 265
COL.CM.x.CGU1             ..................TVAPTTTNNSASDWDES.NWKEYPWY--RM-STAFLLKDRKEL-LG-SVEDLTVLGNKNDSNS..........IRATMKD-ANYTV--V-DMTIVD-VRTGF--AP-YML-R-DD-KWD---A..-N--TA-S---EFNIT-MSHV-V-A- 262
```

```
                                                                      glycosylation NNT
                                                              glycosylation NCT        V3 tip
                                            glycosylation NFT                               V3                                                                                                          CD4                      V4
                                                          glycosylation NTS                                          glycosylation NNT   glycosylation NKT                                          CD4          glycosylation NST
                                                                      *                                   *                                                                                                   *      *
H1B.FR.83.HXB2          LAEEEVVIR........SVNFTDNAKTIIVQLNTSVEINCTRPNNN.TRKRIRIQRGPGRAFV....TIGK.IGNMRQAHCNISRAKWNNTLKQIASKLREQFGNNKTII...............................FKQSSGGDPEIVTHSFNCGGEFFYCNST  388
H1A1.UG.85.U455         ---R-IR--.........-E---N---------VNP-K---S--Y-TRKNI-RYS.I-S-Q--Y....VT--I--DI------V--RD--R-IQ-V-EQ-KKK-N-KT.............................II-AS-----I--T-------------TS  386
H1B.US.90.WEAU160       ----DI--.........-E-------N------V-I------------.---K-TL.----VLY....-T-EI--DI-R----L--TS--------VE----IKQFKNKT-..............................V-----------M--------------  392
H1C.ET.86.ETH2220       I--G-TI--.........FE-L-N---I------E----T----S--.--ES---.----QT-Y....AT-DI--DI--------EE---K--QKVKE--QKH-P-.--..............................IE--P-----L--T-------------TS  383
H1D.CD.84.84ZR085       -----I---.........-E-L-N---I---H--Q---------YKK.E-Q-TP-..-Q-Q-LY....-TRYTTRIIG--Y----GV------R-V-R--GNLLNQT-..............................II--P--------T-------------TS  393
H1F1.BE.93.VI850        ----GI---.........-Q-ISN-------H--E--Q---------.---G-HL.----QT-Y....AT-AI--DI-K------GTQ-----EYVKAE-KSH-P--.-A.............................IK-N------L--TM-----R------DTS  378
H1G.SE.93.SE6165        ---GKIKV-.........-E-----T-V------KT-----------TMKRIRMG.I---QT-Y....AT-AI--DI------VTKR--KEA-QNV-AE-GKI-NKSSEN............................IT-NS-A---L--T----I-R-------TS  395
H1H.CF.90.056           ----QII--.........TK-IS--T-N-----K-P-N-T-------.--TS-HL.-------Y....AT-DI--DI---------TD--K--H-VVTQ-GIHLN-RT.............................IS--PN----M-VR------R-------TS  379
H1J.SE.93.SE7887        I--GDII--.........-E-IS----N------KT---V-Y------.---G-HM.----QVLY....AT-EI---I-ET-----ERD-S---RRV-T----H-NKT.............................IN-TSP----I------------L---TS  379
H1K.CM.96.MP535         -----II--.........-E-I---T-N------ET-Q----------.---S-HM.----K--Y....-T-DI--DI--------GE---M--SRVKE--K-H-K-GTI............................TFKPPNP------L--M---A-------T-  389
H101_AE.TH.90.CM240     -----II--.........-EDL-N-------H--K--------S---.--TS-T-.-----V-Y....RT-DI---I-K-Y-E-NGT---KV---VTE--K-H-NKT.............................II-QPP----L--TM-H---R-------T-  388
H102_AG.NG.x.IBNG       ---G-----.........-E-I-N---------ANP-K---------.---GVH-.----Q--Y....AT-DI--DI------V-KTE--K--H-VVTQ-KTY-K-TT.............................II-ANPL---V--T-------------TS  383
H103_AB.RU.97.KAL153_2  ----------.........-------T-------KEP----------.---G-H-.-------Y....AT-DIT-DI--------IT---------VI---K----KT.............................IV-N----------M-------------T-  377
H104_cpx.CY.94.CY032    --T-------.........-K-I---T-N-----AKA-K------G--.---SVH-.----LTWY....AT-EI--DI--------GND--D---V-SEE-KRL-P-KTIK.............................-APPV---L--T------K-------T-  391
H1N.CM.95.YBF30         -NTDGI---.............N-SHSNLL--W-ET-P------G--.-GGQVQ-..--AMT-Y....N-E-IV-DI---Y--V-KEL-EPMWNRTREEIKKIL-K-N...........................ITFRARERNE--L-VTHLM---R-------TS  387
H1O.BE.87.ANT70         -SKGKIRMM........AKDILEGG-N---T--STLNMT-E--QI..DIQEM--..--MAWYSMG..IG-TAGNSS-A-Y-KYNATD-GKI---T-ERYL-LVN-TG............................SINMT-NH-----L-VTHLH---H-------TA  379
H1O.CM.91.MVP5180       -SR-KIR-M........GK-I-ES--N---T---PINMT-I-EGIA.EVQD-YT..--M-WRSMTLKRSNNTSPRS-V-Y-TYNKTV-E-A-Q-T-IRYLNLVNQTEN.............................VTII-SRT----A-VSHLH---H-------TS  386
CPZ.CD.90.ANT           YQTNTS-VM........NGRKNESVLVRFGKEFENLT-T-I--G-RTV-.NLQ-.---MT-Y....NVEIAT-DT-K-F-TVNKTL-EQARNKTEHV-A-HWKKVDNKTN...........................AKTIWTFQD----VKV-W---Q------DI-  391
CPZ.CM.01.SIVcpzCAM13   --T-NITV-........VN-ASK-THDW----S-A-NLT-K-VG--.--GKVQ-.---MT-Y....NMDHIF-DT-K-F-ELNGTT--E--QKVRES-IKEIKA-ANGTY............................NIT-EP--------AN-M----------DTR  401
CPZ.CM.05.SIVcpzEK505   --TSNI---.........N-SK-....-LL----E-IP------G-K.--GQVQ-.---MT-Y....N-ENI--DT---Y-EVN-T.-EQIWNTTKQIIINN.......RK...........................NIT-IPNP---L-VTNLMI----------TS  372
CPZ.CM.05.SIVcpzMB66    -E-KVT-LD.........R-VSNDMD----K--ET-RL----TG--TIKG.-P-.---SQI-Y....G-ETV--DT---F-QLNKTV-T--F-KVRHA-N-TYKG.YLGNE............................TIT-GP-T---L-VTNLHLI---------TS  410
CPZ.CM.98.CAM3          V--NKTIA-...........R-GYXFLXXFQXX-S------G-X.S-GQ-Q-.----MT-Y....N-ENIV-DT-R-Y-Q-N-TV-DER-NETGQA---L-T-LTQV.............................N-TV-P-----VTNMM-----------T-  388
CPZ.GA.88.GAB2          -----E-VF.....RTKNMTAPGLSD-V----KRAIP---S--G--.-GRA-NL..S--TT-F....NTEAL---P-K-S-HLNGTL---I-NR-KQ-IKNSTTWHRGD-...........................T-TKHP-----V-NFM-----------TS  389
CPZ.CM.98.CAM5          ----GPAVAR.........SEDLRH-I--I-FKEGIN-I-M--G--.S-GQLQ-.----VS-Y....N-ENIV-DP------V-K-Q-MKK-NET-LAIAKYDKSIAKV............................T-TPNP------TNMM-----------T-  384
CPZ.CM.05.SIVcpzMT145   T-DNQT-A-...........IDPSENLA-I--KDP-K-T-R--G--.--GQ-Q-.--AMT-Y....N-ENVV-DT-K-Y-E-NGTQ-AKA-NETKEV--NILRK.............................NIS-MVP------VTN-H-----------TS  377
CPZ.TZ.00.TAN1          -E---TKAY.........FVN-SVNTPLL-KF-V-INLT-E-TG--.--GQVQ-.----MT-Y....N-ENVV-DT-K-Y-SVNATT-YRN-DWAMAAINTTMRARNETV...........................QQT-QWQRD----VTSFW---Q-------L-  386
CPZ.CM.05.SIVcpzLB7     ----KIM--...........YKA-...KVL------IH-----VG-KTIKG.-PL.----QL-Y....GTETVV-DT-E---E-NQTH-YKI-N-VKRE-TTV-NESNK............................TVH-AN-------VANLH-----------TS  377
CPZ.US.85.CPZUS         --TKN-TV-..........SK.NF-DI-L--FSEG-NMT-I--G--.-VGNVQL.----MT-Y....N-P-IV-DV-E------KLT-EKQR-YTLEIIKKEANLT-V.............................ELIPNA-----V-NMML-----------TI  368
CPZ.GA.88.GAB1          ---GNITV-........VE-KSK-TDVW----VEA-SL--H--G--.--GEVQ-.----MT-Y....N-ENVV-DT-S-Y-K-NGTT--R-VEEVKKA-ATSSNRTAAN-..............................TLNRA------VTH-M-----------TS  386
GSN.CM.99.CN166         YQPGNNTRIM.......MNGMKNESIV-GFGEDYHLNLT-I--G-K.-IRNLQ-..-A-MT-Y....SQVIVG--T-K-Y-RLDPDR--RAIREAMKAMN-HWE-KTG...........................RNDTQIRWTSEPK--L-VQ--W-Q-Q-------LS  404
GSN.CM.99.CN71          YQPGNNTRVM.......MNGKKNESIV-GFGEDYQLTLT-I--G-K.-L-NLQ-..-A-MT-Y....SQIIVG-DT-K-Y-KLNHT--DIAIR-AMKAMKNHWEKINNDTT.........................PNKTQIRWTSEPK--L-VQ--W-Q-Q-------LS  409
MUS_1.CM.01.1085        YL-GNNTAVM.......MNGDKNESIG-KFGE-FR-NLT-I--G-K.-IRNLQ-..-A-MT-Y....SQLIVD-DT-R-Y-K-NKTQ-DPA-R-AMLAM-KHWHGRLQKQNHSSINET.................KIKIRWRSEPK--Q-VQ--W-----------LS  413
MUS_1.CM.01.CM1239      YL-GNNTAVM.......MNGDKNESIG-KFGE-FR-NLT-L--G-K.-IRNLQ-..-A-MT-Y....SQLIVG-DT-R-Y-K-NKTQ-DVA-R-AMLAM-KHWHGRLEKQGE.NITEE.................KIKIRWTSQPK--N-VQ--W---A-------LS  407
MUS_2.CM.01.CM1246      YLPGNDTAVM.......MNGDKNESIA-KFGE-FR-NLT-I--G-K.-IRNLQ-..-A-MT-Y....SQLIVD-DT-R-Y-K-NKTQ-DPA-R-AMIAM-DHWH-RLQKANK.TINKT.................EIKIRWTSEPK--R-VQ--W---Q-------LS  411
MUS_2.CM.01.CM2500      YLPGNDTAVM.......MNGKKNESIA-QFGE-FR-NLT-L--G-K.-IRNLQ-..-A-MS-Y....SQLIVD-DT-R-Y-K-NKTQ-DPA-R-AMKAM-AHWA.KMD.....NRSES.................DIKIRWTSEPK--R-VQ--W-----------LS  394
                                                                                   V3
                                                                      *                                   *                                                                                                   *      *
MAC.US.x.239            R--NRTY-Y.........WHGR--RTI-SLNKYYNLTMK-R--G-K.-VLPVT-..MS-LV-.....HSQPINDRPK--W-WF.GG--KDAI-EVKQTIVKHPRYTG-NNT.........................DKINLTAPGG.----VTFMWT--R---L--KMN  405
H2A.GW.x.ALI            R--NRTY-Y.........WHGR--RTI-SLNKHYNLTMH-K--G-K.-VVP-TL..MS-LI-.....HSQPINKRP---W-WF.KGE-RKAMQEVKET-VKHPRYKG-NDT.........................NQINFT-PGR-S-A-V-YMWT--R---LH--M-  397
H2A.DE.x.BEN            R--NRTY-Y.........WHGR--RTI-SLNKYYNLTMR-K--G-K.-VLP-TL..MS-LV-.....HSQPINTRP---W-RF.GGR-REAMQEVKQT-VQHPRYKGINDT.........................GKINFT-PGA-S---VAFMWT--R---L---M-  400
H2A.SN.x.ST             R--NRTY-Y.........WHGR--RTI-SLNKFYNLTVH-K--G-K.-VVP-TL..MS-LV-.....HSQPINRRP---W-WF.KGE-KEAM-EVKLT-AKHPRYKG-NDT.........................EKIRFIAPGERS---VAYMWT--R---L---M-  395
H2B.GH.86.D205          R--NRTY-Y.........WHEK--RTI-SLNTYYNLS-H-K--G-K.-VVP--T..VS-LL-.....HSQPINKRP---W-WF.KGN-TEAI-EVKRTIIKHPRYKGGAKNIT.......................SVKLVSEHGK-S---TTYMWT--R---L---M-  402
H2B.CI.x.EHO            R--NRTY-Y.........WHGK--RTI-SLNSYYNLTMH-K--G-K.MVVP--T..VS-IL-.....HSQPINKRPK--W-WF.KGN-TEAIQEVKETIKNHPRYSG-TNI.........................SQIRLAEHARSS---VRYMWT--R---L---M-  398
H2G.CI.92.ABT96         R--NRTY-Y.........WHGR--RTI-SLNKYYNLT-X-R--G-K.-VLP-T-..MS-LV-.....HSQPINERP---W-WF.GG--KEAMQEVKETVVKHPSYKG-NDT.........................KXITFTTPGE-S---VKFMWT--R---L--KM-  406
H2U.FR.96.12034         RS-NRTY-Y.........WHGR--RTI-SLNKYYNLTMY-R--G-K.-VLPVT-..MS-LV-.....HSQAINN-PK--W-WF.GGN-KGAIQEVKETIANHPRYHG-KNI.........................SQINLAEPAG.--S-VKFMWT--R---L--KM-  401
MAC.US.x.251_1A11       R--NRTY-Y.........WHGR--RTI-SLNKYYNLT-K-R--G-K.-VLPVT-..MSQLV-.....HSQPINDRPK--W-WF.-G--KDAI-EVKQTIVKHPRYTG-NNT.........................DKINLTAPGG.----VTFMWT--R---L--KMN  404
MAC.US.x.251_BK28       R--NRTY-Y.........WHGR--RTI-SLNKYYNLTMK-R--G-K.-VLPVT-..MS-LV-.....HSQPINDRPK--W-WF.GG--KDAI-EVKQTIVKHPRYTG-NNT.........................DKINLTAPGG.----VTFMWT--R---L--KMN  407
MAC.US.x.EMBL_3         R--NRTY-Y.........WHGR--RTI-SLNKYYNLTMK-R--G-K.-VLPVT-..MS-LV-.....HSQPLTDRPK--W-WF.GG--KDAI-EVKQTIVKHPRYTG-NNT.........................DKINLTAPGG.----VTFMWT--R---L--KMN  407
SMM.US.x.H9             R--NRTY-Y.........WHXRS-RTI-SLNKYXNLTMR-R--G-K.-VLPVT-..MS-LV-.....HSQPINERPK--W-WF.GGE-KEAIREVKXX-VKHPRYKG-NXX.........................XQIKLTAPGG.----VTFMWT--R---L--KMN  408
SMM.US.x.PGM53          R--NRTY-Y.........WHSRS-RTI-SLNKYYNLT-R-R--G-K.-VLPVT-..MS-LI-.....HSQPINERPK--W-WF.GGS-REAIQEVKETIVKHPRYTG-NET.........................KKINLTAPGG.----VTFMWT--R---L--KMN  410
SMM.US.x.SIVsmH635F_L3  R--NRTY-Y.........WHGNS-RTI-SLNKYYNLTMK-R--G-K.-VLPVT-..MS-LV-.....HSQPINERPK--W-WF.GGN-SEAIQEVKET-VKHPRYTG-NDT.........................RKINLTAPAG.----VTFMWT--R---L--KMN  408
SMM.SL.92.SL92B         R--NRTY-Y.........WHGGS-RTI-SLNKYYNLT-K-R--G-K.-VLPVT-..MS-LV-.....HSQPINERP---W-WF.GG--REAMQEVKKTIVKHPRYTG-NDT.........................RKINLTAPGG.----VTFMWT--R---L--KMN  408
STM.US.x.STM            R--NRTY-Y.........WHGR--RTI-SLNKYYNLTMS-R--G-K.-VLPVT-..MS-LV-.....HSQPINERPK--W-WF.GGE-RGAI-EVKET-VKHPRYTG-NDT.........................AKIRIVAPGG.----VTFMWT--R---L--KMN  404
MNE.US.x.MNE027         R--NRTY-Y.........WHSK--RTI-SLNKYNNLTMK-R--G-K.-VLPVT-..MS-LV-.....HSQPINERPK--W-RF.EGN-KEAI-EVKQTIVKHPRYTG-NNT.........................DKINLTAPRG.----VTFMWT--R---L--KMN  406
GRV.ET.x.GRI_677        RS-NRTE-W.......QKGGNDNDTVI-KLNKFYNLTVR-R--G-K.-VLPVT-..MA-LV-.....HSQ-YNTRLK--W-HF.QGD-KGAW-EVREEVKKVKNLTEVS-E..........................NIHLRRIW..----SANFW---Q------KMD  392
VER.KE.x.TYO1_patent    YH-NRTQ-W.......QKHRVNN-TVL-LFNKHYNLSVT-R--G-K.-VLPVT-..MA-LV-.....HSQ-YNMKL---W-HF.EGN-RGAWREVKQ-IV-LPKDRYKGTN.........................NTEHIYL-RQW----ASNLW---Q------KMD  405
VER.KE.x.9063           YS-NRTQ-W.......QKHRVNNDSVIVLFNKYYNLTVT-K--G-K.-VLPVT-..MA-LV-.....HSQ-YNTKL---W-HF.QGR-REAW-EVKEEIIKLPKDRYRGTN.........................DTEKIYL-RQF----AANLW---Q------KMD  412
VER.KE.x.AGM155         YS-NRTQ-W.......QKHGVSNDSVL-LLNKHYNLTVT-K--G-K.-VLPVT-..MA-LV-.....HSQ-YNTRL---W-HF.QGN-KGAW-EVQEEIVKLPKERYQGTN.........................DTNKI-L-RQF----AANLW---Q------KMD  407
VER.DE.x.AGM3           YS-NRTQ-W.......QKHRVSNDSVLVLFNKHYNLTVT-K--G-K.-VLPVT-..MA-LV-.....HSQRYNTRL---W-HF.QGN-RGAW-EVKNEIVKLPKDRYQGTN.........................DTEEIYL-RLF----AANLW---Q------KMD  412
TAN.UG.x.TAN1           Q--NRTE-W.......QKHGVSN-SVI-KLNKHYKLK-V-R--G-K.-VLPVT-..MA-LV-.....HSQQYNTKL---W-HF.QGD-KGAWREVRKTIV-LPKEKYRGTN.........................NTRQIWLSRQW----AANIWL--Q------TPD  403
SAB.SN.x.SAB1C          YVANRTE-W.......QKNGNSNDSVI-RLNRYFNLT-R-R--G-K.-VLPVT-..MA-LV-.....HSQ-YNTRLK--W-WF.GGN-RGAW-EVKETIVRLPPKKYSGTN.........................DTNKI-L-RQW----SEFFF---Q------KMD  402
OLC.CI.97.97CI12        RHKASRL-PI.TGDLNHKPREGTGVYVMAIADKYKL--E-I-KG-R.SIVSVNS..AS-LVYY....AGLEPY--IDKGL-RF.KGQ-GWM-YALGKG-EKENNTWANYTGVNITGIGR.....NSTTFAKAIKR...FKFDE-LKR---DAAENMMMM---YMWF--I-  419
DRL.x.x.FAO             KHGQDEL-ET.....RGKPD-LNHKYVFR-NKKWNLK-I-R-KG-R.SIISTPS..AS-LL-Y....HGLEPGK-LKRGM-KF.EGQ-GKA-GSL-KE-KSVS-SIWRNTTNSTCISLKGKKI....QDRNVTGCMLKQLKVDNYTTK--LASENLMML----Y-F--VS  439
RCM.NG.x.NG411          IS-NRTW-Y.........QRKQS-RTV-GLNRHFNLTVT-R--S-R.-V-G-SL..AT-VF......ISLRVQKRPKG-W-RF.DGD-KGAW-EVKEAINKTK-YKG-RQP.........................ENITIRSVYG.--D-AKCFWL--D---L--KMN  396
RCM.GA.x.GAB1           INVN-TW-Y.........QRRQS-RTV-GLNSFYNLSVT-R--S-R.-V-G-SL..AT-VF......ISLRVEKRPKG-W-RF.EGN-TDAW-EVKERVKTTK-YRG-SNT.........................DKIKIRTVYG.--D-ARYFWL--N---L--KLN  402
MND_1.GA.x.MNDGB1       MHK.-GELIPIDDKYRGPEE-HQRKFVYK-PGKYGLK-E-H-KG-R.SVVSTPS..AT-LL-Y....HGLEPGK-LKKGM-TF.KGR-GLA-WSL-KE-NKLNDSI-VNQTCKNFTSTGEENKQNTDKQKEFAKCIKTLKIDNYTT-.--RAAEMMMMT-Q--M-F--V-  423
MND_2.x.x.5440          KHDYDEL-QT..NPRKGKDE-H-HKYVYR-DKKWGLQVR-R-KG-R.SIISTPS..ST-LL-Y....HGLEPGK-L-KGK-QL.EG--GQA-HSLSLE--KINDSIYKDNH.NMTCKSSNNK.......KNTTGCHLKTISISESTVK-E-GAE-IMLL----Y-F--W-  439
MND_2.CM.98.CM16        KHDYNEL-QT..NTKKGKEE-H-HKYVYR-DEKYNLQVV-R-KG-R.SVISTPS..AT-LI-Y....SGLEPGK-LKKGM-QL.KGQ-GIAMHDL-IE--KIDSSIWRNVT.KGYCKGLKRK.......ENRTGCALKTIKVSDYTTK-E-GAE-IMLL----Y-F--W-  418
MND_2.GA.x.M14          KHNYDEL-LV..NPQ...KE-H-HKYVYR-NKKWGLKVV-D-KG-R.SIVSVPS..AS-LI-Y....HGLEPGR-LKKGM-RL.IGQ-GRA-NALSRE-KKVNASIYRGFNASGPCRGINNR.....GVVNKTGCALKTIEVSNYTTI-E-GAE-IMIL----Y-F--W-  418
DEB.CM.99.CM40          DE-RAEELHI..IRKEVKGEVQNGSI--R-PAKYNLTLT-V--G-K.-YRA-HM..AT-LS-Y....-TFIQRLRIKR---RL.NGS-A-AT-EMRQ-IL-I--KANRTNN.............................LTIHYPK--R-VQSVW-Q-H-------IS  400
DEB.CM.99.CM5           DE-RAED-HI.....IRK.GLENKTI-VRIPPKYNLSLD-V--G-K.-YRAVHM..AT-LS-Y....-TFIPRLRIKR---RM.-GN-TGAV-E-R-R-K-IY...NVTK.............................ITVHYP---R-VQNTW-Q-H-------VS  382
DEN.CD.x.CD1            VELLDNHTQI.....IRNTSKDNQTIA-LFRKSSVIKLK----G-V.-L-GLNM..AA-IVMR....ATIIPRRAI-R-A-K-VGGN-TR-RQE-EKAVK-E-NI-G-VKW..............................DVPRV-----TSSAR-Q-R-------L-  396
TAL.CM.01.8023          K..-NQTQII.......HQAPKNESAI-LLGKEHNIT-Q-I--G-R.-I-DLQ-..AA-LM-H....SQIIAGKDLKR-Y-K-.EGN-TRAMEDVNE-IAKAYKDHINSSS........................GTAQIRWQGVRK--L-VSLFWHA-Q-------IS  401
TAL.CM.00.266           K..-NHTQVI.......HQAPKNESAI-LLGKEHNLT-T-I--G-R.-I-DLQ-..AA-LM-H....SQIIAGK-LNR-Y-KL.EGN-TKAMESANEEIAKRYKEHINNTD.........................KVNITWTAVRR--L-VQLFWHA-Q-------IS  384
MON.CM.99.L1            GNAPNTTVM........MNKQKNESIVVRLAKHLH-N-T-I--G-K.-IRNLQ-..-A-MT-Y....SQLIVG--T-K-Y-KVNKTQ-ETA-QAVHEAVKTEWEKKNNGTN.........................VTTISWRFQPQ--K-VQ--W---Q-------VS  390
MON.NG.x.NG1            S-TNXTTVMM........NGKRNESIV-RLXKTL--N-T-I--G-K.-LRNLQ-..-A-MT-Y....SQIIVG--X-K-Y-QVNKTQ-DAA-HAAHXAIXXEWESKNNNVT.........................VANITWRYQPK--N-VQ--W---Q-------LS  401
SUN.GA.98.L14           RHK-DEL-PI....NNKVGQGAEGEYVWK-AAKWGLV-Q-I-KG-R.SQVSTIS..ST-LL-Y....YGLEPGSKL-L-Q-KF.VGQ-GRAFATLGKL--QVEPLANMSNG....CTFDNITKT..CRFTNGTDFKKMIKFKPMEKH-A-AATEMLMMT--E-M-F--L-  432
WRC.CI.97.97CI14        RHK-GEL-EV.IGNINHKPKEGSG-FVYK-PDKYN-K-E-H-KG-R.SLVSVNS..AS-LIYY....AGLEPHR-I-KGL-KF.TGL-GKM-YDLGKN-EKYNSSWVNYTE..GDCNSKN....NSADFAKCIKR...FQIRNFTRQ-A-RATENIMMI----M-F--I-  421
WRC.CI.98.98CI04        RYKKGELVEI.IGNINHRPHEGTG-FVYK-PSRYK---E-N-KG-R.SVVSVNS..AS-LIYY....AGLEPHR-I-KGL-KF.KGL-GHM-YDLGKN-Q-YNNTWVNYTE...DCNQKN....DSREFAKCIKQ...FQIRNFTRK-A-RATENIMMI----M-F--I-  431
WRC.GM.05.Pbt_05GM_X02  KYDKGEL-EV.IGNINHKPKDGTG-FVYK-PAKYN-A-E-H-KG-R.SLVSVSS..AS-LIYY....AGLEPYK-I-K-L-KF.TGQ-GHM-YGLGKE-QAYNSSWVNYTE...ECNKKG....NSSAFATCIKQ...FQIRNFTKQ-A-RATENLMMI----M-F--I-  423
LST.CD.88.524           NHT-DEL-PL...TPKKMGDLNGAKFVYK-AGKWGLI-R-I-KG-R.SEVSTIS..ST-LL-Y....YGLEHGSRL-L-Q-KF.EGQ-GRMFNNLGRM-KKLNETAMNYKE..GNCSTPG.........KPCGRQLKGLPIANMTRR-V-LATEMLMHT--E-M-F--V-  431
LST.CD.88.485           NHT-DKL-PL...TPKKMG-LL-AKFVYK-SGKWGLI-R-I-KG-R.SEVSTIS..ST-LL-Y....YGLEHGSRL-L-Q-KF.EGQ-GRMFNNLGKM-QKLNATAMNYTN..GQCPEGK.........QPCGRHLKGLPIANMTRR-EGLATEMLMHT--E-M-F--I-  428
LST.CD.88.447           NHTDDEL-PL...TPKKMG-LLEAKFVYK-AGKWGLI-R-I-KG-R.SEVSTIS..ST-LL-Y....YGLEHGSRL-L-Q-KF.EGQ-GRMFHNLGNM-KRLNATAMNYTH..GNCIDNT.........KPCGRQLKGLPIANMTRR--NLATEMLMHT--E-M-F--V-  429
LST.KE.x.lho7           NHTRDEL-PL...TPNKMEDLNGAKFVYK-AGKWGLI-R-I-KG-R.SEVSTIS..ST-YL-Y....YGLEHGSRL-L-Q-KF.EGQ-GRMFNNLGKM-K-LNAEAMNYTEGTGTCDSKK.........TTCGRKLKGLPIANMTRH-A-LATEMLMHT--E-M-F--V-  426
SYK.KE.x.SYK173         YKAKDK-RFI.......KQKDKNESVI-L-PEALRLQ-I-E--G-E.SI-N-QL..AA-YFLPVI.QGKL-TGRDAKR-F-RV.TGN-TEFF--VHEQATKTWK-.........................VTNTTWRSQP---L-VR--W-Q---------VS  384
SYK.KE.x.KE51           YKAKDK-RFI.......RQKDKNESVI-L-PEKLGLQLV-E--G-E.SI-N-QL..AA-FFLPVI.QGRL-TGKAAKR-F-KV.KGE-GKFFE-VHNESIKVWK-.........................VTSTSWRSQPQ--L-VR--W-Q---------VS  391
COL.CM.x.CGU1           KELSDWAKD......REG-WKN-SGTIEYYWFPKDIALG-I-RG-S.SHRNLNT..AN-AK-Y....YELIPYSKGIYGR-QFVPMTGQ-KKNKTQFAIEIKKNLTAWLERISRKNI.....................TITPRNGNRTS---ATFTFVI-HRL-----AS  398
```

```
                                glycosylation NST
                             glycosylation NST
                                             V4                                    CD4                                                  glycosylation NES                                          fusion peptide
                                     glycosylation NNT                          CD4          CD4        glycosylation NIT              V5                                  gp120 end  gp41 start
                                                                             *                                 *
H1B.FR.83.HXB2               QLFNSTWFNSTWSTEGSNNT..........EGSDTITLPCRIKQIINMWQKVGKAMYAPPISG...QIRCSSNITGLLLTRDGGN......SNNESEIFRPGGGDMRDNWRSELYKY.....KVVKIEPLGVAPTKAKRRVVQ.....REKRAV..GIGALFLGFLGAA  526
H1A1.UG.85.U455              G----I-NG-MSNDM-P-G-..................---Q-----------R--Q-------Q-...V---E-------------TN......-TKN-T------------K------.....-------------R------E.....-------..-L--I---------  519
H1B.US.90.WEAU160            -------HANGTWKNTEGAD.............NN.-------------R--E----------E-...----L-------------SS.....EE-QT--------N-K----------.....----------------------------.....-----G.ML--M--------  529
H1C.ET.86.ETH2220            N-----KLELFN-STNL-....................---Q-----------G--R-------E-...I-M-R------------AK.....EPHSTK-----E--------------.....---E-K--------P-----E.....-----A...L-----------  514
H1D.CD.84.84ZR085            G----A-NI-GH--GLNDTI..................--I------------E----------E-...--N-----------------A......N-TQNDT------------------.....---------------------E.....-----I..-L--M--------  526
H1F1.BE.93.VI850             G---D-GS-NGT..........................------------V----G--R---TS--A-...N-T-N-----------------......ESNI-T---E--N-K----------.....---E------------Q---.....-----A..-L---------DS  502
H1G.SE.93.SE6165             G----SLLR-NS-EN-T.....................-----K----VR---R--Q-------A-...N-E-N-S----I-------N..NNTNTS----------------------.....-T---KS------R-R----E.....-------..-L--V--------  529
H1H.CF.90.056                G----S-EMH-NY-SNDTKG.............NEN----------V----R--R-------Q-...N-M-V------I--I-E--......ASA-NYT-------------------.....---------I----TR----E.....-------..-M--S--------  515
H1J.SE.93.SE7887             K----S-DKNSIEATNDTSX.............AT.--I--K----VR---RT-Q-I-----A-...N-T-T---------------R..GNG-E-GT-T---T--N-K----------.....---E-----------------E.....-------..----V-----T-  518
H1K.CM.96.MP535              K---E-GE-G-...........................------------------------I-----A-...S-N-------MI-------......N-THN-T------------------.....---Q-----I---R-R----.....-------..-L--V-F------  512
H101_AE.TH.90.CM240          K---N-CLGNETMAGCNDT...................-----K---------GA-Q------------...R-N-V-----I-------VN......-TDN-T------NIK----------.....---Q-----I---R------E.....-------..----MIF------  520
H102_AG.NG.x.IBNG            K------D--NSTANHTGSN.............DT.---Q------V-------Q-------Q-...I---D------------------......-STN-T-------------------.....-----------------R------E.....-------..-L--V--------  517
H103_AB.RU.97.KAL153_2       K------NGTEELNNTEGD...............IV-------------E-----R-----A-...--------------------------......QS-VT----T----------------.....-----------------R------------.....-------..----V--------  510
H104_cpx.CY.94.CY032         P-----HMQNGTNITSTDS-.............NS----Q--L--FVR---E--Q----S--A-...S-N---D---II------T.......-NT----------------------.....-------I----N--R-----------*-.....-----M--------  524
H1N.CM.95.YBF30              K---EELL-E-GEP........................-------R--V-L-TR---GI-----R-...VLN-T------V-EYS--P.......DTKET-VY-S--N-VNL--Q-------.....---S---I----G-----T-S........AF.-L-----------  514
H1O.BE.87.ANT70              KM--Y-F.SCNGT-CSVS-V..........SQGNNG----KLR-VVRS-IRGQSGL-----K-...NLT-M-----MI-QM-NTWNS....S-NNVT---I----K-I--T--FN-.....---RVK-FS----RIA-P-ISTRTH.-------..-L-M----V-S--  521
H1O.CM.91.MVP5180            GM--Y-FI-C-K-GCQEIKG.........SN-TNKNG-I--KLR-LVRS-M-GESRI-----P-...NLT-H-----MI-QL-QPW.....NSTGENTL--V----K-I--TK--N-.....---Q-K-FS-----MS-PIINIHTPH-------..-L-M----V-S--  531
CPZ.CD.90.ANT                PW--A-YT..GNLITNGAL...................IAH------V-H-GI-S-GI-LA-RR-...NVS-T-S---IM-.EGQIY.........--TVKVS-AARVA.-Q--A--SR-.....Q--E-X--S----TX--PEIKQHS..-Q--GI.----LF---L-S--  519
CPZ.CM.01.SIVcpzCAM13        KM--.........ES-PFHE..............NM-I----R--V-S-MR--RGI-----P-...H-T-N-L----I---.DHV.........-TNNT---I----KNI----------.....---R----S---------.HTVG..ER-Q---AF.-L-----------  527
CPZ.CM.05.SIVcpzEK505        ---TNQN.........-NTTG.............N---Q---R--V-L-TR---GI-----K-...P-N-L-----II-D........YTK-GT-KYTIY-T----TNL--Q-------.....---S---I----G----HT-T......-Q---AF.-L-----------  497
CPZ.CM.05.SIVcpzMB66         I---TSI.....IFNETKDD..............N--I----R--VRL--R--RGIFL---R-...T-N-I-----I-FA.....QQ.KTDRM-K-AM-T-V--E--N----------.....---R----------------T-H.........-AV.-L-----------  540
CPZ.CM.98.CAM3               T---Y--K-NNITKGDNTT.................FF----R--V-S-MR---GI-----R-...V-S-T--V--II-ETGH-I.....NNSITNITLY-T--N-V-L--L--H---.....---S---I----S----HT-S.........-AF.-L-----------  521
CPZ.GA.88.GAB2               R-ITCNSSDTSE........................YI---K-R-VV-S-MR---GIF---RR-...T-T-N-T------EVQN........GTGNNTEVYLS------I----------.....-I---------------YT-AKA.LD-S---AF.-L-----------  517
CPZ.CM.98.CAM5               P---KFNYTCDNQNNCTR-N......TNDTNSDNS-L-IS---R--V-S-MR---GI-----R-...N-T-I--------EAG-KWE.....N..NTITVY-S--E---L--L--NR-.....---S---I----------HT-A..........AF.-L-----------  527
CPZ.CM.05.SIVcpzMT145        EII-I-K.......INKTEN..............MTII----R--V-S-MR---GIF----R-...N-T-T-----M--EIHKNREDQGEDQDQNNTYVCLT--N-K-I-----------.....-I-E-Q--------SR-YA-EKQHH.-----L..-L-----------  514
CPZ.TZ.00.TAN1               NWT-TWTA.NRTNNTHGTL.................VA---LR--V-H-GI-S-GV-L--RR-...TVK-H------IM-AEKD-N.........NSYTPQFSAVVE-Y-KV--AR-.....---E-Q--S---RPG--PEIKANH.T-SR-D..V---L------S--  517
CPZ.CM.05.SIVcpzLB7          A---D-LLRNKTLLNETLHN..............NL------R--V-L-LR---GIF----R-...N---N-T----I-E.........KHT-TDNIT---I----T-I------N-.....-I-----------R-R--T-D.........-A..-L-VF--------  508
CPZ.US.85.CPZUS              P---M-YN....N-DNTTI.................--K---R--V-Q-MR---GIF----K-...VLS-N-----MI-DISISAV...NNDSRNITVM-T----TAL-KN--H---.....---S---I----G----HT-K.........-AF.-L-----------  498
CPZ.GA.88.GAB1               -I-TDNIT-GII........................I----R--VSS-MR--RGI-----R-...N-T-N---------S-TPVT....N-SGNLT---T--N-K-I----------.....---R----S------R-HT-ARQ.KD-Q---AF.-L-----------  516
GSN.CM.99.CN166              I--QLNNN.TIN-S.NIG-..........ITSKYKGQW-A-K-R-FVTQ-GY-S-SI-L--RQ-...H-N-T-------IDGAMYES..........S-NMTPSA-VA-A-NY--SR-.....---E-D--SM---P---KEHP....AV-----.LSL-IS--T--S--  538
GSN.CM.99.CN71               V--QFNNNITAVDDTNI--..........VTSKYKGQWMA---R-FVTQ-GY-SRSI-L--RQ-...H-N-T-------IDGAMYGH..........S-NMTPSA-VA-A-KY--SR-.....---E-D--SL---S-Q--PHPGV..HVK----.ISL-IS--T--S--  547
MUS_1.CM.01.1085             V--QFVNHTEINE.........SNIHTVTTKYHKDQWMV---R-FV-Q-GY-S-SI-L--RK-...LV--E-------IDGAMYRD..........S-NMTPSANVL-A--M-PK-VQ.....---E-D--SM---EVQ--P-P....HVHA---.ISL-IT-----S--  550
MUS_1.CM.01.CM1239           V--QFNNNITQINE........SNIHNVTTKYHQDQWMV---R-FVTQ-GY-S-SI-L--RR-...HVQ-T--V----IDGAMYKD..........TVNMTPSANVM-A--M--KN-.....---E-D--SM---EVQ--P-P....HVKA---.ISL-IT-----S--  544
MUS_2.CM.01.CM1246           V--QFENNTREIND........SNIHTVKTKYRHDQWMV---R-FVTQ-GY-S-SI-L--RK-...HVQ-T-HV----IDGAMYGN..........T-NMTPSANV...-TRASRLRD....T--E-D--S----EVR--P-P....HVKA---.ISL-IT-----S--  546
MUS_2.CM.01.CM2500           V--QFENDTRQIHN........SNIDTVTTKYHKDQWMV---R-FVTQ-GYTS-SI-L--RQ-...H-Q-T--V----IDGAMYGD..........TVNRLTPSANVL-A--M--KR-.....---E-D--S----EVQ--P-P....HVQA---.ITL-IT-----S--  532
                                                                                                                                                                   gp120 end  gp41 start
                                                                             *                                 *
MAC.US.x.239                 WFL-WVEDRN-AN.............QKPK-QHKRNYV--H-R----T-H----NV-L--RE-...DLT-N-TV-S-IANI-WIDG.........NQTNITMSAEVAELY-L--GD-......-L-E-T-I-L---DV--YTTGG..TS-N--G-.FVL-..------AT-  539
H2A.GW.x.ALI                 WFL-WVENKTGQEQ.....................HNYA--H------I-H-A--NV-L--RE-...ELT-N-TV-S-IANI-TDG..........NQTNITFSAEVAELY-L--GD-......-L-E-T-I-F---SER-YSST...PR-N--G-.FVL-..------AT-  522
H2A.DE.x.BEN                 WFL-WVEDKNQTRR.......................NY..-H------T-H----NV-L--RE-...ELA-E-TV-SIIANI-IDKN.........RTHTNITFSAEVAELY-L--GD-......-LIE-T-I-F---DQR-YSST...PV-N--G-.FVL-..------AT-  524
H2A.SN.x.ST                  WFL-WVENRTNQTQ.....................HNYV--H------T-H----NV-L--RE-...-LT-N-TV-SIIANI---E..........NQTNITFSAEVAELY-L--GD-......-LIEVT-I-F---PV--YSSA...PV-N--G-.FVL-..------TT-  520
H2B.GH.86.D205               WFL-WVENKTNTTR.....................RNYA--H-R----T-H----NI-L--RE-...ELS-N-TV-S-IANINSD..........N-TTNISVSAEVSELY-L--GD-......-L-E-T-I-F---DVR-YSSV...KP-N--G-.MVL-..------AM-  527
H2B.CI.x.EHO                 FFL-WVENRTGLKR.......................NYAS-H-R--V-T-H-I-RNV-L--RE-...ELS-N-TV-S-IANI-WIDK.........NLTNITVSAEVSELYKL--GD-......-L-E-T-I-F---SI--YSSV...TP-N--.GVLVL-..------AT-  523
H2G.CI.92.ABT96              WFL-WIEDRNMTLL..............KPQ-RQRRNYV--H-R----T-H----NV-L--RE-...DLT-N-TV-S-IANIEENRSN........NHTNIIFSAEVAELY-L--GD-......-L-E-T-I-F---NV--YSSV...TPKN--G-.FVL-..------AM-  540
H2U.FR.96.12034              WFL-WVENR-MEGQ..............TSR-RHKRNYV--H-R-V--T-H---RNV-L--RE-...DLT-N-TV-SMIANI-RI-K.........NQTNISMSAEVAELY-L--GD-......-L-E-R-I-F---PV--YSSV...TP-N--G-.FVL-..------AT-  534
MAC.US.x.251_1A11            WFL-WVEDRN-AN.............QKPK-QHKRNYV--H-R----T-H----NV-L--RE-...DLT-I-TV-S-IANI-WTDG.........NQTNITMSAEVAELY-L--GD-......-L-E-T-I-L---NVR-YTTGG..TS-N--G-.FVL-..------AT-  538
MAC.US.x.251_BK28            WFL-WVEDRDVTT.............QRPK-RHRRNYV--H-R----T-H----NV-L--RE-...DLT-N-TV-S-IANI-WTDG.........NQTSITMSAEVAELY-L--GD-......-L-E-T-I-L---DV--YTTGG..TS-N--G-.FVL-..------AT-  541
MAC.US.x.EMBL_3              WFL-WVEDRDVTTQ..............RPK-RHRRNYV--H-R----T-H----NV-L--RE-...DLT-N-TV-S-IANI-WTDG.........NQTSITMSAEVAELY-L--GD-......-L-E-T-I-L---DV--YTTGG..TS-N--G-.FVL-..------AT-  541
SMM.US.x.H9                  WFL-WVEXIKNG-.........RWTSQNQK-RYQKNYV--H-X----T-H----NVXL--RE-...DLT-N-TV-S-IAEI-WI-G.........NKTNITMSAEVAELY-L--GD-......-L-E-T-I-F---SV--YTTTG..AS-N--G-.FVL-..------AT-  546
SMM.US.x.PGM53               WFL-WVEDRN-SGS.........LWTQQTKADEKKRNYV--H-R----T-HR---NV-L--RE-...DLT-N-TV-S-IADI-WIDG.........NQTNITMSAEVAELY-L--GD-......-L-E-T-I-L---SVR-YTTGA...S-N--G-.FVL-..------AT-  548
SMM.US.x.SIVsmH635F_L3       WFL-WVEDRDQN-N.........RWKQQKKR-QQKRNYV--H-R----T-H----NV-L--RE-...DLT-N-TV-S-IAEI-WI-N.........N-TNITMSAEVAELY-L--GD-......-L-E-T-I-L---DVR-YTTTG..AS-N--G-.FVL-..------AT-  547
SMM.SL.92.SL92B              WFL-WVEDRN-S-P.........RWTTQTKK-QHKRNYV--H-R----T-HR---NV-L--RE-...DLT-N-TV-S-IANI-WIDN.........N-TNITMSAEVAELY-L--GD-......-L-E-T-I-M---HV--YTTS...TSKN--G-.FVL-..------AT-  546
STM.US.x.STM                 WFL-WIENR--SE.........MRDWNKNKK-QQKRNYV--H-R-V--T-H----NV-L--RQ-...DLT-N-TV-SIIANI-WT-N.........N-TNITASAEVAELY-L--GD-......-L-E-T-I-L---NV--YTTS...TS-T--G-.FVL-..------AT-  542
MNE.US.x.MNE027              WFL-WVEDKNLTG............TTQKPQ-QHKRNYV--H-R----T-H----NV-L--RE-...DLT-N-TV-S-IANI-WIDG.........NQTNITMSAEVAELY-L--GD-......-L-E-T-I-L---NV--YTTGG..TP-N--G-.FVL-..------AT-  542
GRV.ET.x.GRI_677             WFI-YLNNRTEDAE..GT-RTCDKGK.PGPGPCVQRTYVA-H-R-VV-D-YT-S-KV----RE-...HLE-N-SV-A-YVAI-YN-KS........GP-NVTLSPQV-SI-AY--GD-......-L-E-T-I-F---DVR-YTGPT...----VPF.VL-..---------  536
VER.KE.x.TYO1_patent         WFL-YLNNKTWDAD..H-FCSSKK.KGHAPGPCVQRTYVA-H-RSV--D-YTLS-KT----RE-...HLQ-R-TV--MTVELNYNSK........NRTNVTLSPQIESI-AA--GR-......-L-E-T-I-F---EVR-YTGG...HE-Q--V.PFVL-..---------  549
VER.KE.x.9063                WFL-YLNNQ-VDAN..H-QCSNETKKGNAPGPCVQRTYVA-H-RSV--D-YTIS-KV----RE-...HLE-T-TV--MTVELNYN-N........YRTNVTLSPQIEGI-AA--GR-......-L-E-T-I-F---NVR-YTGG...QD-S--V.PFVL-..---------  557
VER.KE.x.AGM155              WFL-YLNNLTVDAD..H-HCKNNAGKGRSPGPCVQRTYVA-H-RSV--D-YTIS-KT----RE-...HLQ-T-TV--MTVELNYN-Q........NRTNVTLSPQIETI-AA--GR-......-L-E-T-I-F---EVR-YTGG...QE-Q--V.PFVL-..---------  552
VER.DE.x.AGM3                WFL-YLNNRTVDPD..H-PCNGTKGKGKAPGPCAQRTYVA-H-RSV--D-YTLSRKT----RE-...HLQ-T-TV--MSVELNYNSK........NRTNVTLSPQIETI-AA--GR-......-L-E-T-I-F---EVR-YTGG...HD-T--V.PFVL-..---------  557
TAN.UG.x.TAN1                WFV-WLNNE-NSGR..NVDVEGNNCTTGKDKRCYKRTYV--H-RS-V-D-YTLS-KT----RE-...HLE-T-TV-SMMVSL-YNSK........ERTNVTLTANLENI-AY--GR-......-LIE---I-F---EIR-Y-GPT...----VPF.VL-..---------  547
SAB.SN.x.SAB1C               WFL-YLNNK-VDPD..H--CAKNN.....TKPCWQRTYV--H-R-VV-D-YTLS-KT----RE-...HLE-N-TA-A-YVELNYNSK........NRTNVTLSPQIESI-AN--GD-......-L-E-K-I-F----VR-YTG....PE-Q--V.PFVL-..---------  541
OLC.CI.97.97CI12             KILKTWNNA-DPKW.....................YPWSQ--LR-NVDD-M---RRI-L--QP-FNNR---TQRV-EMF-EIQKVEN........GTQVSLLPPS-NVNQFV-VASD-......-L-EL-LI-F---ETQ-Y......QP-N--G.ALVL-..L--V--L-  546
DRL.x.x.FAO                  KIWRTWNNRTSNVW.....................YPYAS-H-----DD-AR--RKI-M--V--FNNE---TQDV-EMFFEIQKLETE..ETGDSNYQ-KFIPQDEVQNQFTAVGAH-......-L--V--I-F---DVH-YNLPG...H-Q--GA.VVL-..I--L-SL-  575
RCM.NG.x.NG411               WFL-YLNNKTEGNK................N-ARQAMFV--IT-MVV-D-YT-SRKV-T--RPD...AL--NATV-Y--ADI-YTDN........MTTNVTLSAEVG-I-AA--GR-......---E-M-I-Y---NVR-YET......KQ--VP.LVL-..------S--  524
RCM.GA.x.GAB1                WFL-LLNNETVG................TTN-KRKAPFV--IT-M-V-D-YT-SRKV-T--RPD...ALK--AQVSY--ADI-YI-D.........--TNITLSAEVG-Y-AA--GR-......-AIE-R-I-F---EI--YQT......KQ--VP.LVL-..------S--  530
MND_1.GA.x.MNDGB1            RIMRAWNDPNEKKW.....................YPYAS-Q-R--VDD-MQ--RKI-L--T--FNNH---THRV-EMYFEMQKIDSNE.....TKMQ-KFLPPSETSNQFVAYGAH-......-L---M-I-I---DV--HTLPE..HHK---GA.VIL-..I--L-SL-  557
MND_2.x.x.5440               KIWKAWNSKQSSVW.....................YPYMSRN-R--VGD-H----KI-M--V--FNNE---TNDV-EMFFEVQKTDDD......-GYI-KFIPQDWIQNQYTAVGAH-......-L--VD-I-F---DIH-HHLPN...T-Q--GA.VLL-..M--L--L-  571
MND_2.CM.98.CM16             KIWRTWNDQNSSVW.....................YPWMS-N-R---DD-H----KI-M--V--FNNE----NDV-EMFFEVQKTEEG........YI-KFVPQDEVQNRFTAVGAH-.....QL--VD-I-F---EVA-YHLPE...A-Q--GA.VLL-..MF-L--L-  548
MND_2.GA.x.M14               RIWKTWNNQNSNVW.....................YPYMS-K-R--VDD-H---RKI-M--V--FNNH---TNDV-EMFFEVQLVEG.......KRYL-KFLPQDEVQNQFTAVGAH-......-L--VD-I-F---DVH-YHLPD...AKQ--GA.VLL-..M--L--L-  549
DEB.CM.99.CM40               KALDLLLLQNNT................RNSTW--KWLM----N-LV-T-YT--QHI-L--KE-...ELK---H-SAFVFDV-HY-G..........S-TLTPSA-I-AV--AD-F--......-IIEVK-I-F--SAVR-YEGPE..SV-H---AGIAF-..LVA--ST-  532
DEB.CM.99.CM5                KA-..NLLMNKT................NNITW-NNTLM--K-N-LV-T-F---QHV-L--KE-...ELK-H-SVSA-VFEVAHYGQ..........NMNITPSS-IKAV--AD------......-IIEVK-I-M---SIR-YEGPE..SA-Q---AGIAF-..LVA--ST-  512
DEN.CD.x.CD1                 RM---SQYPVNN................SNIINNQTYMV-N---LV-T-HT--MDI-L--AE-...E-T-T--VS-IIVDT-VHPN........RTH-TLTLSA-V-QV--A--ARW......-L-ELT-I-L---SVR-YEGPESAESVSR-RREVSLVLGLI---S--  533
TAL.CM.01.8023               SM-MK-NYTYNYTA-KGVLHQEKIGRQEHRYSDRTKFFMS-QLR-M--R-S--E-L--L--RE-...HLS-K--V--I-VDINHYKGN........DI-SVEPSSQI--A--AAVAR-......--IE-T-I-F---NIR-YDGP...ATK----APLAL-..-I---ST-  550
TAL.CM.00.266                SM-LN-SYEYNYTSNQGKLHQGVIGK.KPRSSDRTKFYMS-QLR-M--R-SR-E-L--L--RE-...HLQ-K-------VDINHHRGS.........-D-TVEPSSQI--S--AAVSR-......--IE-V-F-Y---NIR-YDGP...ANKQ---APLAL-..-I---ST-  532
MON.CM.99.L1                 A--INRRT-K-DG.......ISPFDVNNKPNTTYHGGW-A-T-R-MVTQ-GY-S-SI-L--RK-...HVQ-T----A--I-GELYQN..........NVTLVPSAQVS-S-----SR-......---E-D--SM---T-Q--TGV....H------.ITL-MA-----ST-  529
MON.NG.x.NG1                 R--X-SDPXKNKN....ETFESTFNVSTNASTAYNGEW-T--LR-LVTQ-GYIS-SX-L--RQ-...HVQ-T----A--VDGELYGS..........TVTLIPSARVS-A-----S--......---E-D--S----A-Q--HGL.....MRE-RA.ITL-MA-----ST-  543
SUN.GA.98.L14                RI-KVWNDTTSNKW.....................YPWAN-H--SV-DD-AS---KI-L--T--FNNR---ANRV-EAWF-LERVEDW.KVNGS-I-VVAFQPPTNTLNQFV-TGAH-......-L-R-R-I-F---DEH-YAP.....RK----APVAL-..A-AL-S--  568
WRC.CI.97.97CI14             RIADAWNNR-ETKW.....................YPWS--T-RN-VED-AH--IKV-L--V--FNNH---.....-QRVTEILFNWE........RNQVMFQAPSNIQNSFVAVGQ--......-L--LKMI-M--GDG--V-PPE..HR----G.AIIL-..I-----L-  548
WRC.CI.98.98CI04             KIADAWNNKNESKW.....................YPWS--T-RN-VDD-AH--VKV-L--V--FRNH---.....-QRVTEILFNWE........NKQVEFQAPSNIQNSFVAVGQ--......-L-RLKMI-L--GDGR-VTPKE..HH----G.AIVL-..I-----L-  558
WRC.GM.05.Pbt_05GM_X02       KIADWWNNKTEKRW.....................YPWA--T-RS-VDD-ASI--KI-L--V--FNNH---.....-QRVSEVLFNWE........NGQVEFQAPSH-QNSFVAVGQ--......-L--LKLI-----DG--VTPPD..HH----G.AIVL-..L-----L-  550
LST.CD.88.524                RI-QEWNNKNSDKW.....................YPWAN-H--S--DD-AT---KI-L--T--FNNR---THRV-EMWFEMEKWEPH..EDLGGNLS-KFLPPSWETNQFVATGAH-......-LIRLR-I-F---SEH-YAP....RG-QT--APLAL-..A--L-S--  567
LST.CD.88.485                RILQEWNNKNSDKW.....................YPWAN-H--SV-DD-AT---KI-I--T--FNNR---THRV-EMWFELEKWEPE..ENLGGNLSVRFLPPSWETNQFVATGAH-......-LIR-R-I-F---AEHQYAP....SSKQ---APLAL-..A--L-S--  564
LST.CD.88.447                RI-QEWNNKNSDKW.....................YPWAN-H--SV-DD-AS---KI-L--T--FNNR---THRV-EMWFQMEKFEPE..ENLGGNLSVRFLPPSWETNQFVATGAH-......-LIRLR-I-F---DVH-YAP....SSKQ---APLAL-..A--L-S--  565
LST.KE.x.lho7                RI-QEWNNKNSDKW.....................YPWAN-H--S--DD-ATI--KI-L--T--FNNR---THRV-EMFFEMEKWEPH..EDLGGNLS-KFLPPSWETNQFVA-GS--......-LI-LN-I-F---DEH-YAP....RG-QT--APLAL-..A--L-S--  562
SYK.KE.x.SYK173              K--ANITNGNASKN................NYASNLR-S-A-R----D-RY-R-LI-L--TA-...H-K-T--V-AV-TDIEYYPG..........-TLNFTPTANVE-V--AD-FN-......-LIQ-K-I-F---DQR-YELP...NT-----APLAL-..---L-S--  515
SYK.KE.x.KE51                KI-ANVSNGYANP................SNYAKNLY-S-A-R----Y-GY-T-L--L--TE-...H-K-T----AV-TDIEYYPN..........-QLNFAPTANVE-V--AD-FN-......-LIR-K-I-F---SQR-YELP....TKQ---APLAL-..---L-S--  521
COL.CM.x.CGU1                S-WKHDS...........................PVMN-T-RKLV-S-VTHARIL-G--PG-...HLQ-NWEKQPVIAFMGTIEG....DNDGNGCAYPAAPNFKHALSTL--GR-......-L--MRTTTYV--DI--S-NVNWHHG-Q--...GIFAFSI-AL-SG-  524
```

```
                              fusion peptide                                          glycosylation NAS          glycosylation NHT
                                                                    immunodominant region   glycosylation NKS              glycosylation NYT
H1B.FR.83.HXB2            GSTMGAASMTLTVQARQLLSGIVQQQNNLL.RAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKS..........LEQIWNHTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWAS.LWNWFNITNWLWYIKL  684
H1A1.UG.85.U455           --------I-----------------S---.----------K------------V-------Q------------------T----S------..........Q-D---NM--LQ-EK--SS--GI-YQ-------------LD--A-----N.-*-----S------R-  676
H1B.US.90.WEAU160         ----------------L-------------.---------FE------------V--------------------------T---------R-..........QDY---NM----E-------G--YN-------------------------.--T--D-S-------I  687
H1C.ET.86.ETH2220         --------I-----------------S---.K-------M------------T-V--I--H-R-----------------------S------..........Q-E--DNM---Q----S---DI-YN-L-V------D---KD--A----EN.---------------I  672
H1D.CD.84.84ZR085         -----------------V------------.---------------------------------------------H----T----S---R-..........V-Y--GNM---Q-E--D---G--YN-----I------K------------.-----S--Q------I  684
H1F1.BE.93.VI850          REH-----I---------------------.-----------------------V--------------------------N----S------..........Q-E---NM-----EK--S--SNI-YK-----------------A------.-----D-S-------I  660
H1G.SE.93.SE6165          --------I-----V-----------G---.-----------------------L--------------------------N----V------..........YNE--DNM--I--E-------YQ-Y--L------------D--A--Q---.-----G--R------I  687
H1H.CF.90.056             --------I-----------------S---.---Q-R--M--------------V-------R------------------N----S------..........QSE--DNM------KQ-S---EE-YR-L-V--T-------D--A------.--T--D-SH------I  673
H1J.SE.93.SE7887          --------I-----V-----------S---.K---------K------------V--------------------------N------------..........Y-D--ENM--IQ-E-------GI-Y-----A-----N--KD--A----TN.-------S-------I  676
H1K.CM.96.MP535           --------I-----------------S---.--------------------R-----------------------------N----S------..........W-E---NM-----EK--G--SDT-YK------T-------D--A------.-----D--K------I  670
H101_AE.TH.90.CM240       --------I-----------------S---.-----------------------V---------KF--L-------I---------ST---R-..........F-E---NM--I--E--S---NQ-YEILT------DR--KD----------.-----D---------I  678
H102_AG.NG.x.IBNG         ------R-I-----------------S---.K---------K------------V--L----R------------------T----S----T-..........FND--DNM--IQ-EK--S---DI-YN-------R------D--A------.-----D--------RI  675
H103_AB.RU.97.KAL153_2    --------I---------------------.-----------------------V--------------------------------T-----P..........-DE-*-NM----E-------G--YN----------------I-A------.-----D-SK------I  667
H104_cpx.CY.94.CY032      --------------------------S---.----------R------------V--L-S---------------------N----S------..........YND--DNM--LQ--K------QI-YG-L------------D--A------.-----S--K------I  682
H1N.CM.95.YBF30           --------I-------T----------I--.------------SI--------KV--I----R---I-SL------T--Y-T----ET---NTS.........YDT---NL--QQ--EKVR--SGV-FG---QA-E--NT--KS-----Q-D-.--S--G--K------I  673
H1O.BE.87.ANT70           -------AT--A--THT--K------D---.---Q---Q--R-SX---R--R--L--L-TL-QN----SL---K---V-Y-S-K--RT-IG..........N-S--DTL--Q----Q-S-IS-T-YEE-QKA-V---Q--KK-----E---.I---LD--K------I  677
H1O.CM.91.MVP5180         -------ATA---RTHSV-K------D---.---Q------R-S----R--R--LQ-L-TLIQN--R-NL---K-----Y-S-K--T---G.........RYNDDS--DNL--QQ--QH---VS-I-YDE-QAA-D-----VKA-----E---.-----D--K------I  690
CPZ.CD.90.ANT             --------IA--A-T-N-XH------A---.Q---T-------S---V------M----K--R-----SL---AD-VT-H-T----N--V-FTQTC.AKNSSDIQC--ENM--Q----LVQ-S-GQ-YNILQIAHE---R-KK--Y-----S-.-----D--Q------I  686
CPZ.CM.01.SIVcpzCAM13     -------AV---------------------.----------------V------L-------Q---I--L------S--Y-T----KT--G..........KSMSD--NL--QQ--KL-T---GT-FG-L--A-S------KD-----Q---.-----D---------I  685
CPZ.CM.05.SIVcpzEK505     --------I-------K-------------.------------S---------V--I----R---I--L------SV-Y-N----TT-............NNSYDT--GNM--QN--EQVR--SGV-FG-L-QA-E--SI--KS-----Q-S-.-----D--K------I  656
CPZ.CM.05.SIVcpzMB66      --------L-----------------S---.------------S---------V----------------L----------S----TT-T-.........KSYDD--YNM--Q--K-VS---DV-YN-L-KA-T---N--K-----------.-----D--S------I  698
CPZ.CM.98.CAM3            --------VV--------------------.------------S---------V-------R---I--L------A--Y-T----NT--AN.........TSFDE---NL--QD--KRVK--SGV-F----QA-E--NT--KS-----Q-S-.-----D--R-------  680
CPZ.GA.88.GAB2            --------VM------N-------------.------------S---------V-----------V-AL------TV-YST----T--NSNK.........SY-D--SNL--QQ--KLVE-H-GT-FQ-LQRAHE--NS--K------Q-S-.-----D---------M  676
CPZ.CM.98.CAM5            --------VV-----H--------------.------------S---------V-----------I-SL------A--Y-T-----T--AN.........TSYDE---NL--QD--KKVK--SGV-F----QA-E--NT--KD-----Q-S-.--S-----Q------I  686
CPZ.CM.05.SIVcpzMT145     --------VV---------T----------.------------S---------V-------R------L---T--T--P---R--KT-G-I.........SDYQV---NY--QQ----V----G--YT-L--ANT------K------S--N.--S--D---------M  673
CPZ.TZ.00.TAN1            --------IA--A---G---------Q---.Q------------S---------M----K-IR-----SL---AN--V-HSS----LT-AED-TKCNHSDAKYYDC---NL--Q----LVE-S-GT-Y--L-KA-T-----K---------S-.--D--D--Q------I  685
CPZ.CM.05.SIVcpzLB7       --------L-----------T-----S---.------------S---------V--I----------------------------S----RT-..........KTYNE--DNM-------VR---EI-YG---QA-D---N--KK-----H-T-.-----D-SH------I  666
CPZ.US.85.CPZUS           --------VV--------------------.------------S---------V-----------I--L------T--Y-T----DT---N.........LSYDA--GNL--Q----KVR--SGT-F----QA-E--NT--KS-----Q-S-.-----D---------I  657
CPZ.GA.88.GAB1            -------AV---------------------.K-----------SI--V------L-------Q---I--L------AV-Y-T----N--PGS.........NSTDD--GNL--QQ--KLVS---GK-FG-L--A-S------RD-----Q---.-----D--K------I  675
GSN.CM.99.CN166           -T-----ATA--E-S-S--A--M---E---.--V----S---PS--------T-LSSL-K--R--TI-QA---ANRP--H-I-----T--A-...........GSLPD-ENM--QK-SMLVE-D-YT-QQ-L-QANQ--AS-LN--MK-S--D-.--S--D-SD-QR---I  695
GSN.CM.99.CN71            -A-----ATA---S-S--A------E---.--V----S---S----------LSSL-K--R--TI-QA---ANQP--H-I-----D--AK..........NSTPD-E-M--Q--SKL-E-D-YT-QQ-L-NANH--S--MND--K-S--D-.--S--D-S-------I  704
MUS_1.CM.01.1085          -G-----ATA---S-S--A------A---.--V---------S----------LT-L-KFI---A--NL---ANRQ--H-R-----D--A-..........HTQPG-ENM--QQ-S-LVD-D-TT-QE-L-LA-R---E-QHK-QK-LE-D-.--E--D-SK------I  707
MUS_1.CM.01.CM1239        -G-----ATA---S-S--A------A---.--V---------S----------LT-L-KFI---A--NL---ANRQ--H-R-----T--A-..........HTLPG-ENM--QQ-SKLVD-D-NT-QG-L-AA-K---E-QHK-QK-LE-D-.--E--D-SK------I  701
MUS_2.CM.01.CM1246        -G-----ATA---S-S--A------A---.--V---------S----------LT-L-KFV---AI-NL---ANRQ--H-S-----D--A-..........HTQPG-DNM--QQ-S-LVD-D-AV--G-L-MA-Q---E-QHK-QK-LE-D-.--E--D-SK----V-I  703
MUS_2.CM.01.CM2500        -G-----ATA---S-S--A------A---.--V-T-------S----------LT-L-KFI---A--NL---ANRQ--H-R-----N--A-..........HTQPE-DNM--Q--S-LVD-D-AT-QG-L-LA-K---E-QHK-QK-LE-DA.--E--D-SK--F-V-I  689

                                                                                            *
MAC.US.x.239              --A-----L---A-S-T--A------QQ--.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NA---AFRQV-H-T---PNAS..............-TPK--NE--Q--E-KVDFLEEN-TA-L--A-I-----MY--QK-NS-DV.FG---DLAS-IK--QY  693
H2A.GW.x.ALI              --A--T-AL--SA-S-T--A------QQ--.DVVKR--EM-R-----T-N----VT-I-K-----AR-NS---AFRQV-H-T---VNNS..............-KPD-DNM--Q--EQQVRYLEAN-SEQL-RA-I-----TY--QK-NS-DV.FT--LDL-A-VK--QY  676
H2A.DE.x.BEN              --A---R-L--SA-S-T--A------QQ--.DVVKR--EM-R-----T-N----VT-I-K---H-AQ-NS---AFRQV-H-T---VNDS..............-SPD-KNM--Q--EKQVRYLEAN-SQSL--A-I-----MY--QK-NS-DI.-G---DL-S-VK--QY  678
H2A.SN.x.ST               -AA-----L--SA-S-T--A------QQ--.DVVKR--EM-R-----T-N----VT-I-K-----AQ-NS---AFRQV-H-T---VNDT..............-TPD--NM--Q--EQR-R-LEAN-SESL-QA-I-----MY--QK-NS-DV.FG---DL-S-IK--QY  674
H2B.GH.86.D205            --A---T-L--SA-S-T--A------QQPV.DVVKR--E--R-----T-N----VT-I-K-----AQ-NS---AFRQV-H-T---PNET..............-TPN--NM--QQ-EKQVHFLDAN-TA-L--A-I-----MY--QKINS-DV.FG---DL-S-IK--H-  681
H2B.CI.x.EHO              --A-----L--SA-S-T--A------QQ-V.DVVKR--E--R-----T-N----VT-I-K-----AQ-NS---AFRQV-H-T---VNES..............-KPD--NM--QQ-E-QVRFLDAN-TK-L--A-I-----MY--QK-NQ-DI.FS---DF-S-MA--R-  677
H2G.CI.92.ABT96           --A--T--L--SA-S-T--A------QQ--.DVVKR--EM-R-----T-N--T-VT-I-K-----AR-NS---AFRQV-H-T---D-LGA--T............--PQ--NM--Q--EKQ--FLEDN-TR-L--A-I-----MY--QK-NS-DV.FG---DL-S-VK-VY-  698
H2U.FR.96.12034           --A-----L--SA-S-T--A------QQ--.DVVKR--EM-R-----T-N--T-VT-I-K-----AS-NA---AFRQV-H-T---INDT..............-TPN-DNM--Q--EEKV-YLEEN-TQ-L-AA-I-----MY--QK-NN-DI.FG---DL-S-VK-VY-  688
MAC.US.x.251_1A11         --A-----L---A-S-T--A------QL--.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NA---AFRQV-H-T---PNAS..............-TPK--NE--Q--E-KVDFLEEN-TA-L--A-I-----MY--QK-NS-DV.FG---DLAS-IK--QY  692
MAC.US.x.251_BK28         --A-----L---A-S-T--A------QQ--.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NA---AFRQV-H-T---PNAS..............-TPD--ND--Q--E-KVDFLEEN-TA-L--A-I-----MY--QK-NS-DV.FG---DLAS-IK--QY  695
MAC.US.x.EMBL_3           --A-----FR--A-S-T--A------QQ--.GVVKR--E--R-----T-N--T-VT-I-K--E--AQ-NA---AFRQV-H-T---PNAS..............-TPD--ND--Q--E-KVDFLEEN-TA-L--A-I-----MY--QK-NS-DV.FG---DLAS-IK--QY  695
SMM.US.x.H9               --A-----L--SA-S-T--A------QQ--.DVVKR--E--R-----T-N--T-VT-I-K-----AXXNS---AFRQV-H-X---PNDT..............-TPX--NMX-Q--EKQV-FLEAN-TXXL--A-I---X-MY--QK-NXXDX.XG--XDL-X-IK--QY  700
SMM.US.x.PGM53            --A-----V-RSA-S-T--A------QQ--.DVVKR--E--R-----T-N--T-VT-I-K-R---AQ-NS---AFRQV-H-T---PNAS..............-VPN--NM--Q--E-QVDDLEAN-TQAL--A-I-----MY--QK-NS-DI.FG---DL-S-IK--QY  702
SMM.US.x.SIVsmH635F_L3    --A-----L--SA-S-T--A------QQ--.DVVKR-HE--R-----T-N--T-VT-I-K-----AQ-NS---AFRQV-H-T---PNDS..............-VPN-DNM--Q--E-KVDFLEAN-TQ-L--A-I-----MY--QK-NS-DI.FG---DL-S-IR--QY  701
SMM.SL.92.SL92B           --A-----L---A-S-T--A------QQ--.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NS---AFRQV-H-T-L-PNDS..............-VPD--NM--Q--EKKVEFLEAN-TQML--ARL-----MY--QK-NS-DV.FG---DL-S-VR--QY  700
STM.US.x.STM              --A-----L---A-S-T--T------QQ--.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NS---AFRQV-H-T---PNDS..............-VPD--NM--Q--E-KVDFLEAN-TQ-L--A-V-----MY--QK-NS-DV.FG---DL-S-VR--QY  696
MNE.US.x.MNE027           --A-----L---A-S-T--A------QQ--.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NA---AFRQV-H-T---PNAS..............-TPN--NE--Q--E-KVDFLEEN-TA-L--A-I-----MY--QK-NS-DV.FG---DLAS-IR--QY  696
GRV.ET.x.GRI_677          -TA----AT-----S-H--A--L---K---.A-V-Q--Q--K--I--V-N-N--VT-L-K--E--AR-NS---AW-QV-H-T---KYNN.............TPK-DNM--L--E-Q--ALEGN-TQ-L--A---ES--LDLYQK--D-SG.F-S--SLST--G-V-I  689
VER.KE.x.TYO1_patent      -TA----ATA----SQH--A--L---K---.A-V----QM-K--I--V-N-N--VT-L-K--E--AR-NS---AW-QV-H-T-E-PWTN.............RTPD-QNM--L--E-Q-ADLE-N-TRQLVKARE-E---LDAYQK-TS-SD.F-S--DFSK--NIL-M  703
VER.KE.x.9063             -TA----ATA----SQH--A--M---K---.A-V----QM-K--I--V-N-N--VT-L-K--E--AR-NV---AW-QV-H-T---QWQN.............MTPN-QNM--L--E-Q-GELEGN-TEQLVKARE-E---LDAYQR-TS-SN.F-S--DFSK--NIL-I  711
VER.KE.x.AGM155           -TA----ATA----SQH--A--L---K---.A-VG---QM-K--I--V-N-N--VT-L-K--A--AR-NA---AW-QV-H-T---TWNN..............TPE--NM--L--EKQ-EGLEGN-TKQL-QARE-E---LDAYQK-SD-S-.F-S--DFSK--NIL-I  705
VER.DE.x.AGM3             -TA----ATA----SQH--A--L---K---.A-V----QM-K--I--V-N-N--VT-L-K--E--AR-NA---AW-QV-H-T---QWNN.............RTPD--NM--L--E-Q-SYLEGN-TTQL--ARA-E---LDAYQK-SS-SD.F-S--DFSK--NIL-I  711
TAN.UG.x.TAN1             -AA---TATA----SQ---A--L---K---.A-V-Q--QM-K--I--V-N-N--VT-L-K--E--TR-NL---AF-QV-H-T---TFNN..............TPD-DNM--Q--ESQ-TALEGN-STTLVKAYE-EQ--MDTYQK-GD-T-.W--I-DVSS-F-W--W  700
SAB.SN.x.SAB1C            -AA----ATA----SQ---A--L---K---.A-V-Q--QM-K--I--V-N-N--VT-L-K--E--AR-N----AFRQV-H-T-L-KYNN..............TPD-ENM--Q--E-Q-EK-EAN-SRIL-QAHE-EQ--LDSYQK-VS-SD.F-S--DL-K-FGWM-I  694
OLC.CI.97.97CI12          --A--TV-TI----SSH--A---R--KHM-.EM--K--E--K--I--V-N----L-G--K-----A--KQ---EFMQV-H-T----YTQ..............SPD-GND--LK-Q-KVD-L-QE-DDRLTIAYQE-QR-QY--QRFQELSD.WTK---LA--MT-V-I  699
DRL.x.x.FAO               --A-CSVATAM---SQA--T-M-E--KQ--.-LV-Q--E--K--I--V-N----LT-L-E-IG--AM-SL----FAQV-H-N-V-PNES.............VTPN-TSE-----QKRVDSISNN-TLDLQKAYE-EQ--IF--QK-GDLT-.WA---DF-W-SK---I  729
RCM.NG.x.NG411            -TA----ATA----S-H--A--L---KK--.DIV-Q--E--K-----T-N----VT-I-K--A--S--NTF--AWRQV-H-T-E-IY-Q..............TPE--KQ--L--E-N-SRLEGN-SVALQDA-E-H-R-VHD-EK-NS-GD.MLS-L-MDW--K--RI  677
RCM.GA.x.GAB1             -TA----ATA----S-H--A--L---K---.DIVKR--N--K-----T-N----VT-I-K--A--S--NTF--AWRQV-H-V---TFNK..............TPE-QKES-LQ-E-N-SYLEAN-TIALQ-A-D-H---VH--EK-SN-GD.AFS-L-LDW-MQ---I  683
MND_1.GA.x.MNDGB1         --A--SV-VA----SQS-VT---E--KQ--.KL--Q-SE--K--I--V-N--T-LTSL-N-I---A--SQ----WAQV-H-S-E-TNTS.............ITPN-TSE--K--ETRTDYLQQN-TEMLKQAYDREQR-TY--QK-GDLT-.WAS--DF-W-VQ-L-W  711
MND_2.x.x.5440            --A--SVAVA----SQT--N---E--KV--.SL-DQHSE--K--I--V-N--V-LT-L-E-VA--SR-SV----FSQV-H-S-K-PNNS.............IVPN-TSE--L----RV-SIVTNMTIDLQRAYELEQR-IF--QK-GDLNFHGLTG-DL-W--K-V-I  726
MND_2.CM.98.CM16          -----SVAVA----SQA--N---E--KV--.SL-DQHSE--K--I--V-N----LT-L-D-VA--AR-SM----FAQV-H-H---PNDS.............ITPN-TSE--L---KRVTAL-DNMTVNLQKAYELEQ--IY--EK-GD-T-.WAS--DF-W--K-V-I  702
MND_2.GA.x.M14            --A--SVAVA----SQA--N---E--KI--.SL-DQHSE--K--I--V-N----LT-L-D-VA--SR-AV----FSQV-H-N---PNES.............ITPN-TSE--L----RVTAI-NNMTIDLQRAYELEQ--MY--QK-GDLT-.WAS--DL-W--K-V-I  703
DEB.CM.99.CM40            -AA-----TA----S-S---------QE--.K-V--HGQ--T--A--VRN-NT-LT-I-K-----AK-NE---AF-Q--H-T----N-L..............EDPD-DNM--Q--EMKVA---DEWEGALQRA-E---R-VHA-QS-QD-D-.-----DLSR-F-W-R-  686
DEB.CM.99.CM5             -AA-----TA----S-S---------QE--.K-V--HG---S--A--VRN-NT-LT-I-K-----SK-NE---AF-Q--H-T----HT-.............G-PD--NM--Q--E-KVA---DEWEGALQRA-E---R-VHA-QS-TD-D-.-----DLSR-F-W-R-  666
DEN.CD.x.CD1              -TV---VGTG-A--S-D--A------QE--.--V-GHSA----S-----N-N--LT-I-K-----A--NE---AW-Q--H-S-E-SNQS........IPGLALYQPDFQNM--QQ-ETD-AALFGNVSDALKKA-A---R-VH-VQS-YD-DN.-----DLSK-F-W-R-  693
TAL.CM.01.8023            -TA---VATA----S-S---------EH--.----H-------------N-N--LT-L-K--E--AR-NS---AW-Q--Y-S----KT-T-............STNPD-QNM--Q--EKLVD-ASDT-TV-LQ-A-E---R-VH--QK-ND-D-.--S---LSA-FRWLRI  707
TAL.CM.00.266             -TA---VATA----S-S---------EH--.----H-------------N-N--LT-L-K--E--AK-NS---AW-Q--Y-S----KT--.............YTDPQ-QNM--Q--EMKVD-H-G--SQ-LQ-A-E---R-VH--QK-ND-D-.--S---LSA-FRWLRI  689
MON.CM.99.L1              -G-----ATA---S-S--A------E---.--VT---S--------V------LT---KFI---T--NA---AN-AV-H-T----N--AK...........GHFPE-DNM--QQ-SELVD-D-MT-QQ-L-AA-E--G--QH--MKPGQ-DF.-----D-SK------I  686
MON.NG.x.NG1              -G-----ATA---S-S--A------E---.--VT---S----S----------LT---KFI---T--NS---ANRAX-H-Q-L--NT-AK...........GHFPE-DNM--QQ-SMLVD-D-A--QX-L--A-E--G--AH--MK-GQ-DW.-----D-SK------I  700
SUN.GA.98.L14             -TA--LV-TI-----QAV-Q--L---KQ--.VLV-K--E--R--I--V-N----LT-L-E-VQ--S--AS---QW-QV-H-N----YN..............ITPN-TKD-------QVKM-DDN-TA-LQ-AYVTELE-QNKFKQ-QEFN..F-S-LDLSQ-FL---Y  720
WRC.CI.97.97CI14          --A--TV-TV--I-SQH--A------K---.ELV-K--E--R--I--V-N----LS-I-GF-E--LK-KQ---ELTQV-H------Q--N.............ATPN-ENE--QH-YKRTDELQAN-SAILVIAHD-EQA-MF--QK-SD-TA.WGS-LQWF-IWNWL-W  703
WRC.CI.98.98CI04          --A--TV-TV--I-SQH--A--L---K---.ELV-K--E--R--I--V-N----LS-I-GF-E--LK-KQ---ELTQV-H-N----E--N.............ATPN--NE--QH-Y-RTDELQAN-SV-LV-AH-TEQ--MF--QK-QD-TT.WGS-LQWF-IWNWL-W  713
WRC.GM.05.Pbt_05GM_X02    --A--TV-TV--I-SQH--A------K---.ELV-K--E--R--I--V-N---LS---G--E--VK-KQ---ELTQV-H-----EEA-N.............MTPA--NE--QQ-HLRTDTLM-N-SDMLVQAHA-EQS-MF--QK-TD-TQ.WGS-LSWF-IWNWL-W  705
LST.CD.88.524             -TA--LV-TI-----QVVIQ--L---KQ--.VLV-K--E--R--I--V-N----LT-I-E-----A--AS---QW-QV-H-N----YN.............VTPN-TRD--I--E-QVGSLEAN-TT-LQ-AYTTELE-RNNFKK-QDFN..F-S-MDL-T-FQ---Y  719
LST.CD.88.485             -TA--LV-TI-----QVVIQ--L---KQ--.VLV-K--E--R--I--V-N----LT-I-E-----A--AS---QW-Q--H-N-E--YN.............ITPN-TRD--I----RVGVLEAN-ST-LQ-AYTTELE-RNAFKK-QEFN..F-S-LD-LS-FQ---Y  716
LST.CD.88.447             -TA--LV-TI-----QVVIQ--L---KQ--.VLV-K--E--R--I--V-N----LT-I-E-----T--AS---QW-QV-H-N-E--YN.............ITPN-TRD--I----QVGVLEAN-ST-LQ-AYTTELE-RNAFKK-QEFN..F---LD-LS-FQ---Y  717
LST.KE.x.lho7             -TA--LV-TI-----QAV-Q--L---KQ--.VLV-K--E--R--I--V-N----LT-L-E-V-H-A--AS---QW-QV-H-N-E-TYN.............ITPN-TKD--R--ESKVAI-DKN-T--LQ-AYTTELE-QNKFKK-QEFN..F-S-LD-SH-FT-V-Y  714
SYK.KE.x.SYK173           -TA--G-ATA--L-SQT--A------QK--.E-V-------G-----V-N-N--LT-L-T--R--AI-SN---AF-Q--H---T-EKACG-N-N.......FCPKPQ-KNM--HR-EQ-VD-L-DH-DG-LR-A-E---R-VHD-TK-QE-D-.--S--DLSK-FF-L-I  676
SYK.KE.x.KE51             -TA--S-ATA--L-SQT--A------QK--.E-V-------G-----V-N-N--LT-L-T--R--AIMSN---AF-Q--H---T-QQACG-N-........RCPTPQ-ENM--HT-E-QVD-L-DH-DN-LR-A-E-----VHD-TK-QE-D-.--S--DLSK-FQ-L-I  681
COL.CM.x.CGU1             -AA--S--VA--I--QS-NGRASASS-RM-LKLV-T-SA--------V-N--V-VATI-G--EE-AK-ASI--ANMQ--R-I----KT-.............GE-DP-QNM--KQ-HERVR---DI-EADLV-AYDL--E--KK-A--GD-TN.WFSG-GLF-IFK-VLY  680
```

```
                                                                                              glycosylation NGS                                                          glycosylation NAT
                                                                                                               *
H1B.FR.83.HXB2          FIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTHLPT.......PRGPDRPEGIEEEGGERDRD.RSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELLG.......RRGWEALKYW...........WNLLQYWSQELKNSAVSLLNATAIAVAEGTDR 828
H1A1.UG.85.U455         -VI-----I------T----I---------------LA-I.......-E-LG--GR-------QGK-.------S-F--IA-----N-----------FA---A-A-----RSSLKGL-L---G---L...........----L--GR---I--IT--D-V-V---GWI-- 827
H1B.US.90.WEAU160       --------I------T--------------------A.......-------------------.--G--D-F-T---V--------L---I-----AK-------.......----------C...........------------------------------ 831
H1C.ET.86.ETH2220       -------VI----I--------------------LI-H.......------LG--------QG--.--------F--IF-----------------I--AA-T-----RSSLKGLQ----T---L...........GS-V---GL---K--IN---T---V-G----- 823
H1D.CD.84.84ZR085       --------I-------------------------L--A.......-----------------Q--G.T-------FS--------N----------E-I--AA-------.......----------L...........--------R------I--VD----------- 828
H1F1.BE.93.VI850        --------I---------------K------L--LI-S.......------------G---QGK-.--V---T-F---A-----N------RH---FI--AA---DRGL.......---------L...........G--TR----------I--F-T---V------- 804
H1G.SE.93.SE6165        --------I-------------------------LTHH.......Q-E---------G---QG-G.--V---S-F-P-------------------SI---A-T-----RSSLKGL-L---G---L...........----L--GR------I---DTV-----NW--- 838
H1H.CF.90.056           --------I----I--------------------LV-N.......----------T--G---Q---.--V----F-PVV-------S----RL--------V-T-----.......---R-----L...........-------G-------ID---T---------G 817
H1J.SE.93.SE7887        ----I---I----I----A---------------LI-N.......-TEA---G----G---QG-T.--------F---A-----N-----------FV--AA-T-GT--.......L----I---L...........V--VW--G-------I----T----------- 820
H1K.CM.96.MP535         ----I---I----A-----V--------------LI--.......S--A------------Q-KN.--V---S-F---A-----N------RQ--N-I------L-RGL.......-G-------L...........---V-----------I----T------G---- 814
H101_AE.TH.90.CM240     --------I----I--------------------PSHH.......QKE---------G---QG--.--V---S-F---A-------------------T--AA-T-----HSSLKGL-----G---L...........G---L--G----I--I---D-----A-GW--- 829
H102_AG.NG.x.IBNG       --------I---------T-I-------------LTHH.......Q-E-----R---G---Q-K-.--V---S-F---A-------------------I--AA-T-----HNCLKGL-L--G----L...........---IS--V--------IN---TI--V--NW--- 826
H103_AB.RU.97.KAL153_2  -------------I---------G----------R--A.......Q-----------D-------.T------RF--------------F-I--H--------AA-------.......-----------------...........-------I----S--IN-IGTI-----GW--- 811
H104_cpx.CY.94.CY032    --------I----I-----------------L--LI--T.......Q--L---G-T------Q--S.--------F-P--------N------RH--N------A-T-----.......I--------L...........--F-L--G-------IN-F-T----------- 827
H1N.CM.95.YBF30         A----A-I--I--ISI-IT-IA-----------L--LI--.......A-------ET-GGV--Q--G.--V---S-FS--V-E---N-LI-L----T-S---LR-TL----QSLSRGLQLLN-LRTHL...........-GI-A--GK--RD--I----T---V------- 824
H1O.BE.87.ANT70         A-I---A---V-VIMI--N--KNI----Q---L-IPNHH.......QEEAGT-GRTGGG---EG-P.-W-PSPQ-F-P-LYT---TII-WT--L-SN-ASGIQKVISY-RLGLWILGQKIINVCRIC...........AAVT---L---Q---T---DTL-V---NW--G 828
H1O.CM.91.MVP5180       A-I---A-I-I-VIMII-NL-KNI----Q---L-IPV-H.......RQEAET-GRTG-----G--P.KWTA-PP-F-QQLYT---TII-WT--L-SN-ISGIR-LIDY--LGLWILGQKTI--CRLC...........GAVM---L-------TN--DTI-VS--NW--G 841
CPZ.CD.90.ANT           ------AI-----LLVLV-CLRK-----H-----IPTQN.......QQD-EQ--E-R----RK--I.-WRA-QH-FF--L-V--T-IIQWI-QIC-TC--NLWAVLQH-CR.......ITFRLCNHLE...........N--STL-.TIIRTEIIKNIDRL--W-G-K--S 830
CPZ.CM.01.SIVcpzCAM13   -L-A---II----IMS-V-VIR-----------L--LI-A.......Q-E-GTLGETGG----PGSG.--V---T-CWP--------N-VIW--RS-TS-AC--W-QL-K--QLLINIL-LIQ-K-TLL...........RGII---GR--RT--T---D------G----- 836
CPZ.CM.05.SIVcpzEK505   ---V-A-I--I--ISIIM-M-A-----------L--LI--.......T-------RT--DA--L-NG.--V---S-F---A-E-F-N-L--L----T-C-S-LR-TL---RQNIHKGLQLLN-LRI-L...........-GIIA--GR---I--IN--DT--V--------- 807
CPZ.CM.05.SIVcpzMB66    --I-----I-------L-------------------LI-A.......R-DR----E---G---P-NV.------S-F---A-N---D----L-----------L-TL--V-QTLLKGL---R---IHL...........RGI----G----T--I---DT----------- 849
CPZ.CM.98.CAM3          -----A----I---G-II-F-AK----------L--LI--.......T-E----G-T--DV--PGKG.------S-F-----E---N-L--C----------LG-TL-N--QSLNKGLQQLRNFSR-L...........-GVIT--GR--QT--I---D------------ 831
CPZ.GA.88.GAB2          --IL---VI-----IG-VN-IR-S-----------LI-A.......TG--EV----GGG---Q--G.--V---S-F---C---I-N-TI-L-RL---C-SVLWSLL-Q--QQVLRGL-LLR-L-SQL...........KGIG---L---RT--I---DT---I---R--T 827
CPZ.CM.98.CAM5          -LIV-A--I-F-LIGI-M-VIAK----------L--LI--.......T-E---RG-T--D---PG-G.--T---S-F---V-E---N-LI-F----I---S-LW-T-QI--QSINKGLQQLKDFST-L...........KGV-T--GR-IQI--T---DT------N---- 837
CPZ.CM.05.SIVcpzMT145   -LIV---II----C--IG-LI----K-------L--LI-A.......NQ---GLGET-KG---NV-G.------S-F-P-V-E---N-LS-L--Q---CASLIW-LL-I--QYSLRGVQQIGTL-HQL...........RGT----TT-I----I---DT-----G----- 824
CPZ.TZ.00.TAN1          A-I--A-------LMFIVNV-KQ-----T--FS-IPTQA.......EQD-EQ-G--AGG--G--NI.-WTPSPA-FFSIV-E---N-LIWI-QTFQNFIW-LWISLQA-KQGIISLAHSLVIVHRTIIVGVRQI....IEWSSNTYAS-RVLLIQAIDRL-NFTGWW--L 843
CPZ.CM.05.SIVcpzLB7     ----I---IVC--I----A----------------LS-I.......---LAQ-G-T------QGT-.--V--LD-F---A-N---D----L----T--V--AR-TL-IV-QYTLKGL-LV----L-L...........QGI----GK---T--T---DT----------- 817
CPZ.US.85.CPZUS         -LIV-AS---I---GVIF-L-AK----------L--LF--.......T-E------T--GA-KT-NV.--T---S-F---V-E---N-LI-L----E-----LR-T-QI--QNINKGLQLLN-LRARC...........-GVIA--AR---V--T---DT----------- 808
CPZ.GA.88.GAB1          -L-A---II----IMT-F-V-R-----------L--LI-V.......Q-EQG-LGE-D-G---Q--S.--V---E-C-P-------N-GIW--QS-TS-ACN-W-QLKT--HLILHSL-LLR-R-CLL...........GGII---GK---I--I---D------------ 826
GSN.CM.99.CN166         -VIV-AA-IA----MFI-NMLR-I---------P-ILI-Q.......AVERGP-DAT--GV-AQGKV.--V-YLT-FSS-L------N-VIWI-QI-AS-AWTLRG-GKF-WEQLQKAIQKAIQQARQLREVAARV....IAYIS-GI---QAA-TGI-DSL--FTWNW-EA 853
GSN.CM.99.CN71          ---V-AA--A---IMF--NMLR-----------P-ILT-Q.......QVER-P-D-TRKD--AQGNV.--V-YLT-FSS-L--------VIWI-QI-AT-VWTLQ--GKF-WEQ....LQKAFQQARQLGEVAARA....AAYIS-GI---QAAVSGI-DSL-VFTWNW-EP 858
MUS_1.CM.01.1085        -C-V-A---LF-L-MF--G-LR-----------P-ILIRP.......QT--EA---T-GD--SGS--.--V-YLS-FFS-F------N-VIW--QT-ISSASTLWWLLQRAVHLIQKGSI-L-FT-RALIEPLREALQRALAC--HGI---QTASTAV-DQL-TFTWNW-ET 869
MUS_1.CM.01.CM1239      -C-V-A---LF--FMF--G-LR-I---------P-I-TQP.......-EAAEP---T-GGV-SSG-T.--V-YL--F-S-F------N-VIWF-QT-VSSVSTLRW-LQRAAHLIQKWSV-L-IVIREQARRCSELLARLLAYI--GI---QAAGTAA-DQL--FTWNW-EP 863
MUS_2.CM.01.CM1246      -C-V-A--IA---MFI-GMLR-I---------P-I-TRQ.......SAEQEQQ-KT-KGA-RSG-H.-.........-S-FSE--HN-MIWT-QI-TSSASNLWWLTQRIYHHLTKGLVNP-TL--SFREALGRFIAH....A--GI-K-QAAVSGA-D-L--FTWNW-EP 854
MUS_2.CM.01.CM2500      -CLV-A--IA---IMFL-G-LRK---------PP-I-TRA.......-E--RT--ET--G--A-GSV.--V-YL--F-S-------N-IVW--QI--SS-SNLWWLIQ-IVHHLQKWSISL-LRARATSTHLLEVLRRLLAY---GL--FQEGVTAAIDGL--WT-NW--A 851

MAC.US.x.239            GVY--V-VIL----IYIVQMLAKL----R-VFSSPPSYFQQTHIQQ..D-AL-TREGK-RDGGEGG.GNSSWPWQIEYIHFLIRQLIR-LTWLFSNCRT-LSRVYQI-QPILQRLSATLQRIREVLRTE.........LTY---GWSYFHEAVQAVWRSATETL-GAWGD 851
H2A.GW.x.ALI            GVY-IV-I-A----IY-VQMLS-L-K--R-VFSSPPGYIQQIHIHK..DQEQ-TRG-T-EDVG-NV.GDRLWPWPIAY-HFLIHLLAR-LIGLYSICRD-LSRISPI-QPIFRSLQRALTTIRDWLRLK.........AAY---GCEWIQEAFRAFARIARETLTNTWRD 834
H2A.DE.x.BEN            GVH--V-IIA---AIY-VQLLS-F-K--R-VFSSPPGYLQQIHIHK..DRGQ-ANEGT-EDVGGDS.GYDLWPWPINYVQFLIHLLTR-LIGLYNICRD-LSKNSP....TRRLISQSLTAIRDWLRLK.........AAQ---GCEWIQEAFQAFART-RETL-GAWGW 832
H2A.SN.x.ST             GVY--V-IIV----IY-VQMLS-L-K--R-VFSSPPAYFQQIHIHK..DREQ-ARE-T-EDVGNSV.GDNWWPWPIRYIHFLIRQLIR-LNRLYNICRD-LSRSFQT-QLISQSLRRALTAVRDWLRFN.........TAY---GGEWIQEAFRAFAR--GETLTNAWRG 832
H2B.GH.86.D205          GLY--A---V----VYIVQMLA-L-K--R-VFSSPPSY-QQIPIRK..DRGQ-ANE-T-E-GGNDG.DYRSWPWQIEYIHFLLRQLRN-LIWLYNGCRT-LLKTFQ........ILHQISTNLQPLRLP.........VAY---GISWFQEALRAAAR--GETL-SAGET 831
H2B.CI.x.EHO            GLY--I-I-V---AIYIIQMLA-L-K--R-VFSSPPSY-QQIPIRK..DRGQ-ANE-T-E-GGNNE.GYRSWPWQIEYIHFPIRQLRD-LIWLYSGCRT-LSKTFQT-QPVLQPLRLPPA...................Y-R-GISWFQEAIQAAAR-AGETL-SAART 824
H2G.CI.92.ABT96         GLYV-A-VIV----IY-VQ-LG-L----R-VXSSPPSYVQQIHIRK..DQEX-TKE-I--ESGNKG.GYRSWPWQIEYIHFLXRQLGN-LTWLYSNCRA-LSRIXQT-QPLFQRISRTLQAIREHLRLE.........AAYFS-GFRW-QEACTAATR-AQETLTSTWRA 856
H2U.FR.96.12034         GLYV-A-III----IY-VQLLGSL-K--R-VFSSHPSYVEQIPIQR..DQEQ-TKG-I-E-AG-SG.GYRSWPWQIRYIHFLIHQLIR-LTWLYSSCRD-LSRICQS-QPLF.......QSIRERLHLE.........IAY---GW-YF-EAFQAFGK-ARETLSRTGRE 839
MAC.US.x.251_1A11       GVY--V-VIL----IYIVQMLAKL----R-VFSSPPSYFQ*THIQQ..DQAL-TREGK--DGGEGG.GNSSWPWQIEYIHFLIRQLIR-LTWLFSNCRT-LSRVYQI-QPILQRLSATLRRIREVLRTE.........LTY---GWNYFHEAVQVDWRSATETL-GAWGD 849
MAC.US.x.251_BK28       G-YV-V-VIL----IYIVQMLAKL----R-VFSSPPSYFQ*THTQQ..D-AL-TREGK--DGGEGG.GNSSWPWQIEYIHFLIRQLIR-LTWLFSNCRT-LSRAYQI-QPILQRLSATLRRVREVLRTE.........LTY---GWSYFHEAVQAGWRSATETL-GAWRD 852
MAC.US.x.EMBL_3         G-YV-V-VIL----IYIVQMLAKL----R-VFSSPPSYFQXTHTQQ..D-AL-TREGK--DGGEGG.GNSSWPWQIEYIHFLIRQLIR-LTWLFSNCRT-LSRAYQI-QPILQRLSATLRRIREVLRTELT.........Y---GWSYFHEAVQAGWRSATETL-GAWGD 853
SMM.US.x.H9             GVL--L-VI-----IY-VQMLA-L----R-VFSSPPAXVXQIPIXT..-QEL-TKEG---DGGX-G.GNRSWPXQIEYIHFLIRQLIR-LTWLFSSCRDWLLRXCQI-QPVLQSLSRTLQRAREVIRVEIT.........Y---GWRY-QEA-XXWWKFXRETL-SAWRD 858
SMM.US.x.PGM53          GVL--L-V------IY-VQMIA-L----R-VFSSPPVYVQQIPIHK..-QEQ-TKEG---DGG--G.GNRSWPWQIEYIHFLIRQLIR-LTWLFNNCRTYLLRIYQI-QPVFQRL...LQRIREVVRTE.........ISY---GCRWFQEAVQAAWR-RGETL-SAWRD 857
SMM.US.x.SIVsmH635F_L3  GVL--L-V------IY-VQMLA-L----R-VFSPPPAYVQQIPIHK..DQEP-TKEG---EGG--G.G-RSWPWQIEYIHFLIRQLIR-LTWLFSSCRDWLLRIYQI-QPVLQRLSRTLQRVREVIRIE.........ITY---GWSYFQEA-QAWWKVARETL-SAWRD 859
SMM.SL.92.SL92B         GVFL-I-I-L----IY-VQMLS-L----R-VFSSPPSYHQQIHIQR..DQEL-AKEGT--EGGNGG.GYRLWPWQIEYIHFLIRQLIRILTWLYNNLTR-ASRAYQN-QQLCQRLSEISQPIRELVRREAG.........YIR-GWNYFIEACQEAWRSAQE-IVGAWGL 858
STM.US.x.STM            GVYL-I---M--VAIYIMQLLA-L-K--R-VFSSPPSCRQQIPIHK..-QEQ-TKEGT-E-GG--G.GINSWPWQIEYTHFLIRQLVR-LTWLYNNFRACLSRIYQT-HPTFQRISRILQRIREVVRLG.........AAY---GCIWIQEA-QAAWR-AGETL-SAGRD 854
MNE.US.x.MNE027         GVY--V-VIL----IYIVQMLAKL----R-VFSSPPSYFQQTHIRQ..DQAL-TKEGT--DGG-SG.GNSSWPWQIEYIHFLIRQLIR-LTWLFSNCRT-LSRAYQI-QPIFQRFSTTLQRVREVLRTE.........LTY---GWSYFQEAVQVAWRSATETL-GAWGD 854
GRV.ET.x.GRI_677        GFLVIVIIL---FAWVLWGCIRNI----N--P.........QIHIHS..SAE--DNGGGQDRGGESSSSKLIRLQEESSTPSRINNWWLN-KSCS--IRTWCYNICLT--IFIRTAVG.......................Y---GL-Q-QEA-TG-AQ-L-R-AR-AWG- 826
VER.KE.x.TYO1_patent    GFLVIV-II---LLYT-YGCIV------V---P.......QIHIHQ.V-KG--DNAD-P-EGG-NSRIKLESW*KDSKSRCMQ-TAWLTRLN......TWLYNSCLT--IQLRKAFQ.......................Y---GLA---TG-QEI-QTL-GVAQNACHQ 835
VER.KE.x.9063           GFLV-V-II---LLYTIY-CIG----------P.......QIHIHP..WKG--GNAG-PEEGGE--ND-SDSWQKGSGTRQKRGSWSKRLTNWLSTFSTWLYNTFLT--IHLRGAWQ.......................YI--GLG---TA-QEARL-V-RFAQNAGHQ 849
VER.KE.x.AGM155         GFLA-I-VI---LLYTLYTCIA----------P.......QIHIHP..WKGQ-DNAG-PEEGGRTGKSKSTH*QKEFGGRDKRTSWCRQLTNWF-TLSIWLHNSCLT--LKLRSAWQ.......................Y---GLG---AA-QEAV-HL-SFARNAAHQ 842
VER.DE.x.AGM3           GFLD-L-II---LLYT-Y-CIA----------P.......QIHIHP..WKGQ-DNA-GP-EGG-KRKN-SEPWQKESGTAEWKSNWCKRLTNWCSISSIWLYNSCLT--VHLRSAFQ.......................YI--GLG---AA-QEAVV-L-RLAQNAGYQ 849
TAN.UG.x.TAN1           GFY--I--ILF-MAWLIWGCIA------F---P.......QINIRL..-REQ-DNAGG-DKDSSSS.-DKSPPSVKES-LPNRGGIQAEERAW-QHLTNWCL-ISSW--RLYQILRRSLTTLLQLLRQE.........CQYI--GW-QF-EG-ARSFE-L-S-AQSASRT 851
SAB.SN.x.SAB1C          A--VIA-IIVA-VLLVIIG-LRKF-K--A---SLPSSH*.QIHIHL..DKGQ-DKE--Q-SGGLNSDSRSYTWQREFL............RHLCH-LITWLRN.LTSWFSTIFSNLH-CLQDIQQRTRQLTA........H-E-GW----AA-A--TV-VVQ-ATSAS-S 839
OLC.CI.97.97CI12        GLLV-LAY---N--LGLVRKIRIM----VS-QLP.......FSDTLQ..DKEQ-DNG-G-DS-FFES..TFNEW-CTTECTGTLRVNWRDLLS-LWKGVVRVTR..................................WI.QKSLQHVWRTLKQKRQ......HGIN 814
DRL.x.x.FAO             GFF--MAII----LA-LW-T-G-F----R--.PYIFKDAYRPETTN..YHRQ-DREKG-EQD-GEW.NIRSEPSKPGSSKAWSEE..TVGTWLKESRGYIWLKNVKAV-EYGWSE.........................-QEAGRGIC.TLLQEVAQRIWREGRQLGLS 867
RCM.NG.x.NG411          G-F-IL-II----I-LLW-C-SKL-A--T--LSPPPCYHQQIPIRN..REEQ-IKEG-K--SGEGG.GHKYYNWQRG......YFYILILRPIEILCRR-YQICSNS-SVIYQSLQSIPSRIL-HLRWIGA.........K--DGW--F-GFSSW-AEMARTNAYYTWRG 829
RCM.GA.x.GAB1           GFF--I-II---VAWLLWNCLSNL----R--..SPPSYVQQIHIHN.T-EPQTP-EKR-D-GEEGG..NKYNNWLRE......YCWIQ-IHPLSRIWTQ-SQICRSCSSIIFQSLRWILA...................KI--GW--F-EFSSWFAEMALQNAYYTWRG 822
MND_1.GA.x.MNDGB1       GVFL-L-II----LL-LWNTIS-F----R-VFS.......QDCQQN.LYRK--DNG---SNSLELG.EHNSENL...KEESLNRSLIEDLTSFARE*PIWLWLKNLRAAIEYGF............LAVQEA...IRSLGS-LVSFAAISWR-GYTGWLQ-NRIFT-CRE 853
MND_2.x.x.5440          GLLV-VVII---MLACLW-VLGKF----R--PYVFKGDYLRPH.....NLK--DREGG-EPDLEKQ.NIKSESSRQESRKPWKPE...QVRSWLK-ST-Y-WLKNLQAVIEYGWQE.........................-KAAGAALYQVLQGFAQRLWSRGYQL-LSG 863
MND_2.CM.98.CM16        GLL--IVIIV---LACLW-VLGKF----R--PYVFKGDYLRPH.....NLKQ-GKER--EPDSEKQ.STKSDSSKVEFGKPWSKE...QIRDWLKTSRGYVWLKNLQAVIEYGWQE.........................-KTAGGKIF-VLQGYAQGLWSRGHQW-LST 839
MND_2.GA.x.M14          G-L-IMVVI----LACLW-TIG-F----R--PYVLKGDYLRPR.....NLRQ-DEER-GELNLEEQ.NIKSESSKKEFGRPWRPE...QIRNWLKESTIY-WLKNLQAVIEYGWQE.........................-KAAGARICQVLQSFAQRLWSGGHQL-LSV 840
DEB.CM.99.CM40          VVY-IAA-IL---AMFGVN-GSKLCR-----QTL-PK.........EEEVL-LAEEGGRS-GDG-A.-WTA-PR-FFS---E--QQ-LTWL-QTF-NCIWLIR-GF---QEGIYRLSFAL-QR-RLLGERFRVL....GGY---GWK-F-DTITWAGGELWQWLQGIAEM 842
DEB.CM.99.CM5           VVY-IAA-IL---AMFGVN-GTKLCR-----QTL-PK.........EVEVLQLG-EGDKS--GVKS.-WTTSPR-F-R-----ISQ-LTWL-QT--NCAWLIR-GF---QETIYRLSFAF-QR-RLLGEHLRVV....-GYC--GWK-F-DTIIWAGQEIWEWLQGIAQV 822
DEN.CD.x.CD1            VVYVIA---I---IMFIINVLS-LCR-----LQIPIR..........-QKAANQPEEG-GESGDGNRLKWWQYPR-F-SI--E--SQ-LTW--LT--N-VFLIRTACSRIWDFV..........HRHSLILLG.........YF--GL---QTGLRD-GTSAIQQGRATAEV 836
TAL.CM.01.8023          AVIV-AS-IL----MYIINVLKLL----L--P-K-PI-S.......N-E-GV--EEGT---KP--S.-WRTSPQ-FF-I--E--N--LSY--QT--N-VWLSWDLLKR-KATTQQAAQATYAY-RIKFTAAV.......ATI--GFS-IQAAWGDCVT-L-RFTWRW-SG 862
TAL.CM.00.266           AVIV-AS-IL----MYIVNVLKLL----L--P-K-PI-S.......N-D-EA-G-DAT---KQGSY.KWN-SQA-F-SI--E--TG-IRYI-QT-AN-AWLIWDLAKAAKAKSQQVAQATYAF-RLHLTAAAATLQY.......GFS-IQAAWRDGIT-L-RFTWWW-NG 844
MON.CM.99.L1            --IV-AA-I----LMFI-GVIS-LG----L--S-IPI-S.......HA-QPT-D-TGAG--DGSNS.--PAYLK-FFTI--E---N-V-WT-QI-K-SV-VIY--LQRVSQRLPPLLHIRLLQ-WESLRRLLAYCQY.......GI---QAAVT---D-L-RFTIVW--A 841
MON.NG.x.NG1            --IV-AA-----VLMFI-G-IR--G----L--P-...IPTQVPR....E-GA--ETG-G---QG-NRXTVYARGFLTIIWT-CRDLIVWTFQLLRDSG-T-YRSLQRVRSHLIPLLRDLCRQLREASSRLL........AY---GL--XQXACTGAID-L-RFTVIW--A 855
SUN.GA.98.L14           AVL-I-IIIAA--LSFIIQQIY-MC---RV--PSAYVEQDWLQETC.PK-TDK-EE--TEK--I..YINLEQSKKESLPPPWTVDWDEPLRDSL-VT--KWLKAWGIV-AQNIYHLLSFL-HL-TTSFH.HGQRLWQTLRGWLGSHLIQAT-RIRNACR-SRERVSSSQK 886
WRC.CI.97.97CI14        GLLVI-IV-----LACLMNSIRIA-----K-IS-......QNHLNN.LEEGQQDSE-GIEEDYNSIEINDWQYRRES-IVFLNSLLTS---ICQHLLHIVWSLLSCL-WWINDGW........SLLKDI..............SGRAYGAIRKKIAKSRKENTQ--VSAT 844
WRC.CI.98.98CI04        GLL-I-IV--I--LGWLMTSIRIA-----K-IS-......QSHLNN.LEEGL-DRE--AEEDNSTLDV-DWQFKKES-ITLLNSLINF-SAVCQTLLQIVWSLLACLFWWIQHGW........NLLKRI..............SGAAHACIRKKITKSRKETTEK-ISEN 854
WRC.GM.05.Pbt_05GM_X02  VLLVI-LV--I-VLGCLFTSLRIA-----K-VS-......QSHILN.LEEGL-DNE-T-EEDFIAIEE-DWQFKKESLILFWNSLATL-Y-ICLSLLQ-SYSLLSYLWI-IQNGW........SLLKRL..............CGAAYAALQEKIAARRQKRKEK-IYTP 846
LST.CD.88.524           AVL-II-III---LSFIIQS-VKMCR--RV-APSAYVEQDFKWENG.ESQEQ-DKEG-TEKD-ENIYINLEQCKKESFRPPWNVDWNEPLQDSLFVT--KWLKAWGI--LTLVWQFLTWLGHIVILSYH.HGQHLWQTLSRFLVE-AQKTA-WIRNKSRRNRGELSK-NG 887
LST.CD.88.485           AVL-II-IIV----SFIVQNIVKMCR--RV-APSAYVEQDYKWEKG.EN-EQ-DKEG-IEKDKEDIYINLEQCKKESSIPPWNVDWSEPLQDSLFVT--KWIKAFGI--LSLVWQFLSWLGHLVILFFQ.HGQHLWKASSGWMVENAQKIARWSWNKIRRNRRDISAAHA 884
LST.CD.88.447           AVL-II-IIV--V-SFIVQNIVKMCR--RV-APSAYVEQDCKWEKR.ENQEQ-DKEG-IEKDKENIYINLEQYKKESSTPPWNVDWSEPLQDSLFVT--KWIKAFGI--LSLVWQFLSWLGHLVILFFQ.HGQHLRKACSRRVVENAQKIA-WSWNKIRRNRRDISAAHA 885
LST.KE.x.lho7           AVL-ILVII---VLSFIIQN-VKMCR--RV--PSVYIEQDYKWEKE.ENQEQ-DRE--K-ADTETIYINLEQCKKESSRPLWNVDWNEPLQDSL-VT--KWLKEGGI--LSLVWQSLSWL-HL-ILFFQ.NGQRLWQTSSRWMVENAQKIQ-WLREKCRRNRGQLSS--- 882
SYK.KE.x.SYK173         GFYVI-A--L--L-SFSVG-IKNLLG--V-ILQNPT...........Q-RKD-GKPAD-EEGSGDR.EGLNVSTF-RESLRQS-EAGQQLWRTVCSSFRSLIRQLTITW-...........................FIS-GFN---IA-A--GREIRDW--AIWQA 804
SYK.KE.x.KE51           GFFAIAAI-I--VLSFAWG--RNMLG-----LQKPQP--.........LYHKA-A-P--GIETDTDN.YKPSVSTF-REFLRQS-AEAQKLWNRVSESCRSLIRGL-IAW-...........................FIS-GVT---EA-I--GREV-HQMVAIWQA 812
COL.CM.x.CGU1           AAYV----I----IMV-IACIRGAFRVKGFQQIGRTNVSS......QIRGQDR-WEQPDNA-EIHR.EQENSSSRFW...............RLMEEWWSSLL-HCSNGILT...........VLPQHPPTTS....DKRRSA-.-T-RYL..LVPRGPSLQIL-TLQ- 809
```

```
                                                   gp41 end
                                                    Env end
                                 *
H1B.FR.83.HXB2          VIEVV...QGACRAIRHIPRRIRQGLER..ILL*   856
H1A1.UG.85.U455         ---IG...TIG---LN------------..A---   856
H1B.US.90.WEAU160       ---I-...-RT----L------------..A---   860
H1C.ET.86.ETH2220       F--LI...-RIW--FCN----------A..A-Q-   852
H1D.CD.84.84ZR085       I-DI-...RR--K-VL---T--------..A---   857
H1F1.BE.93.VI850        I---L...-R-G--VLN--------A---..A---   833
H1G.SE.93.SE6165        ----A...-R-----LN--T--------..A---   867
H1H.CF.90.056           I-VI-...-R-W---L---------F--..S---   846
H1J.SE.93.SE7887        I--IA...-R-F---L------------..A---   849
H1K.CM.96.MP535         I--IG...-R-F--LL------------..A---   843
H101_AE.TH.90.CM240     ----A...---W---L------------..T---   858
H102_AG.NG.x.IBNG       A--IG...-RVG----N--------F--..A---   855
H103_AB.RU.97.KAL153_2  ---IG...-RF---M-N--------A-K..A-Q-   840
H104_cpx.CY.94.CY032    I--A-...-R-----CN-----------..A---   856
H1N.CM.95.YBF30         I--LA...-RIG-G-L------------..A-I-   853
H1O.BE.87.ANT70         I-AGI...-RIGTG--N-----------..S---   857
H1O.CM.91.MVP5180       I-LGL...-RIGQGFL--------A--..--V-   870
CPZ.CD.90.ANT           ILLAL...-TIV-I--EV---------I..A-N-   859
CPZ.CM.01.SIVcpzCAM13   I-RA-...-IVF-I-GN------------..T---   865
CPZ.CM.05.SIVcpzEK505   I--L-...-RIG-G-L------------..A---   836
CPZ.CM.05.SIVcpzMB66    I--IA...-RFG-G-LN-----------..A---   878
CPZ.CM.98.CAM3          IL--A...-IIG-G-L------------..S---   860
CPZ.GA.88.GAB2          I---A...TRIG-G-L------------..A--E   856
CPZ.CM.98.CAM5          I--F-...-IVG-G-L------------..S---   866
CPZ.CM.05.SIVcpzMT145   IL-AI...TRLG-G-L------------..A---   853
CPZ.TZ.00.TAN1          I--G-...VYIA-G--N---------L..A-N-   872
CPZ.CM.05.SIVcpzLB7     I--I-...-RIG-G-L---------F--..A---   846
CPZ.US.85.CPZUS         I--LT...RRLFLG-I------------..S--X   837
CPZ.GA.88.GAB1          I--AF...-VTL-I--N-----------..A---   855
GSN.CM.99.CN166         -LHAC...RRVW-EFLA--------A-I..LFN-   882
GSN.CM.99.CN71          LVQT-...GRVW-EFLA--------A-I..L-N-   887
MUS_1.CM.01.1085        -LQAG...RRVW-EFLA----V---A-I..L-N-   898
MUS_1.CM.01.CM1239      ILQIG...RRVW-EFLA--------A-I..L-N-   892
MUS_2.CM.01.CM1246      ILQFS...-KIW-KFLA--------A-I..L-N-   883
MUS_2.CM.01.CM2500      -LQ-S...KQL*K-FLA--K----A-I..L-N-   879
                                                    Env end
MAC.US.x.239            LW-TL...RRGG-W-LA----------L..T---   880
H2A.GW.x.ALI            LWGA-...-WVG-R-LAV-------A-I..A---   863
H2A.DE.x.BEN            LW-AA...RRIG-G-LAV-------A-L..A---   861
H2A.SN.x.ST             FWGTL...GQIG-G-LAV-------A-I..A---   861
H2B.GH.86.D205          LW-AL...RR-A---IA----------L..T---   860
H2B.CI.x.EHO            SWG-L...RR-AGE-IA--------A-L..A---   853
H2G.CI.92.ABT96         LWKTL...GRVG-G-LA----------L..T---   885
H2U.FR.96.12034         LW-TL...GRVG-WL-A--------F-L..A---   868
MAC.US.x.251_1A11       L*-AL...RRGG-W-LA----------L..T---   877
MAC.US.x.251_BK28       LW-TL...RRGG-W-LA----------L..T---   881
MAC.US.x.EMBL_3         LW-TL...RRGG-W-LA----------L..T---   882
SMM.US.x.H9             LW-TL...GRVG-G-LA-XX-------L..T---   887
SMM.US.x.PGM53          SW-TL...GRVG-WVLA----------L..A---   886
SMM.US.x.SIVsmH635F_L3  IW-TL...GRVG-G-LA-----V-----L..T---   888
SMM.SL.92.SL92B         IW-TL...GRVG-G-AA----------L..M-N-   887
STM.US.x.STM            LW-TL...GRVG-R-GA----------L..T---   883
MNE.US.x.MNE027         LW-TL...GRVG-W-LA-------E--L..T---   883
GRV.ET.x.GRI_677        LGAI-...RS-Y--VINS---V-----K..V-G-   855
VER.KE.x.TYO1_patent    IWLAC...RS-Y-N-VNS---V-----E..--N-   864
VER.KE.x.9063           IWLAC...RSTY-H-ISS---V-----E..--N-   878
VER.KE.x.AGM155         IWLAC...RS-Y---INS---V-----E..V-N-   871
VER.DE.x.AGM3           IWLAC...RS-Y---INS---V-----G..--N-   878
TAN.UG.x.TAN1           LWNAC...RS-Y---LEH---M--E---..WFN-   880
SAB.SN.x.SAB1C          -RHAC...RSIV--VIAH---M--E---..WFN-   868
OLC.CI.97.97CI12        LQQKK...-RRGSPLQKAQEQLGREGAE.TTGV-   844
DRL.x.x.FAO             SARAL...RALAQEVAA--------A-V..LFN-   896
RCM.NG.x.NG411          LCA-C...RDPAGWPATLC------F--..F-N-   858
RCM.GA.x.GAB1           LCA-A...RDFAGWPAMVC----------..LCN-   851
MND_1.GA.x.MNDGB1       A-IAA.....GTC-W--L-----SA---..P-N-   879
MND_2.x.x.5440          -RGAA....AFG-G-WN--------A-A..L-N-   891
MND_2.CM.98.CM16        AACFR....AIA-G-IN--------A-V..L-N-   867
MND_2.GA.x.M14          IRGAA....AIG---GN--------A-V..L-N-   868
DEB.CM.99.CM40          ALQGL...VR-GG--LRV-A-----A-L..--N-   871
DEB.CM.99.CM5           A-QGL...VWGG-NLLA--A-----A-L..L-N-   851
DEN.CD.x.CD1            -LAAL...TR-A-EVVA----------I..V-N-   865
TAL.CM.01.8023          -LAAA...RW-VQ--AA----T-----Y..C-A-   891
TAL.CM.00.266           -LAAA...RW-GE--AA----T-----Y..CFT-   873
MON.CM.99.L1            LLHAG...GRLW---VA--------A-I..F-N-   870
MON.NG.x.NG1            -LRLG...GRLW-GLVAV-------A-I..L-N-   884
SUN.GA.98.L14           ARSRTFSLGRKW-PKWNRT-GS-IPS-TTETT--   920
WRC.CI.97.97CI14        GSRKT...AASS-IRMLAT-LCASFDKCRRSFSE   875
WRC.CI.98.98CI04        SNRET...T-SS-VWKL-T-LCSSLNKSRRR-F-   885
WRC.GM.05.Pbt_05GM_X02  GSRET...SRSSSIRGITA-LCTNLNPCRRGQQ-   877
LST.CD.88.524           TNIRL...GQNR-RW.RFRF-SGLPS-TTETT--   917
LST.CD.88.485           RDIRL...GQKK-WRFRFRG-SGFPS-TTETA--   915
LST.CD.88.447           RNLRL...GQKK-WRFRFRG-SGFPS-TTETA--   916
LST.KE.x.lho7           KNIQL...GKKK-WRLRFGG-SGISS-ATETA--   913
SYK.KE.x.SYK173         IYAAT...RRVVE-VAAL---L-----I..Y-N-   833
SYK.KE.x.KE51           LLAYA...RRVAENVAAL---L-----I..Y-N-   841
COL.CM.x.CGU1           WLRSA...ARGW-RAPEYL-GWIYDRPQ..GPA-   838
```

```
                         myristoylation                                                                                                   phos
                          Nef start                                                                                         acidic cluster poly-P helix  phos                    HXB2 premature Nef end
H1B.FR.83.HXB2           MGGKWSKSSVIGWPTVRERMRR............AEPAADRVGAASRDLEKHGAITSS.......................NTAATNAACAWLEAQ.......EEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLWIYHTQGYFPD*QNY  126
H1A1.UG.85.U455          -------K-RVE--E--K---E............TPA--KG---V-Q--D-Y--V---.......................--SS---S-------.......--GD-----R-------------F---F--------D------K--E-----V-----F---W---  127
H1B.US.90.WEAU160        ---I---R-GS---AI----K-............-----EG---V----A--------.......................-E..N--D-V--K--.......-D-------R-------------H-----#-----------Y--K---------V-----F---W---  124
H1C.ET.86.ETH2220        ---TM--C-PV---AI---I--............-A---EG--------D-Y--L---.......................--P-N-PD-------E.......--------R-------------F---L------------Y-KK--E-----V-N---F---W---  128
H1D.CD.84.84ZR085        ---------IV---AI---I-K....TDPRERRRP-----G---V-----R--------.......................---T-------V---E.......-D-------R---------------------RK-------VY--K---------V-----F---W---  136
H1F1.BE.93.VI850         ---------IV---A-G----Q............TPT--EG---V----DRR------.......................--RT--PDL------.......----------R----V-------------------------Y-KK-G-T-----V-----------W---  127
H1G.SE.93.SE6165         ---------IV---E----I-N............TPT--EG---V-Q--DR-------.......................----N-PD-------E.......-DS------R---------F-G-F---F--------D---Y-KK--E-----V-N-------W---  128
H1H.CF.90.056            --------RMG--S-I------............---V-EG---V----DRR--V-IN.......................---S--RDA------E.......DG--------R------------G-F-------------D---Y-KQ---------V-N-------W---  128
H1J.SE.93.SE7887         --N-----.....--Q------............APA----G---V-Q--A--------.......................--------D--------.......T--------K--I--------G-----F--------D---Y-KK--E-----VHN-------W---  124
H1K.CM.96.MP535          ---------IV---AI-------...ARPAADRVGTQ----G---V-Q--AR---V---.......................--SHN-PD--------.......----------R-------------F--GF---------D---Y-KK--E-----V-----F---W---  136
H101_AE.TH.90.CM240      --S------IV---Q---KIKQ............TP--TEG---V-Q--D---------.......................--..ID--D-V--R--.......-D-------M------------G-F---F--------D---Y-KK--E-----V-N---F---W---  125
H102_AG.NG.x.IBNG        ---------IV---K-MK---Q............TPT--TG-----Q--DR-------.......................---Q--PD--------.......-D-N-----R------------G----------------D---Y-KK--E-----V-N---F---W---  127
H103_AB.RU.97.KAL153_2   ---------IV---Q----I--............APA---RG--PV-Q--D-Y--V---.......................----N--D--------.......----------R------------G-F-------------D---Y-KK--E-----V-----------W-..  126
H104_cpx.CY.94.CY032     ---------IV---EI------ARAEPERMRRAQ-----AG---V-Q--D------IN.......................-------PDKT-----E.......----------R----------F-G-L-------------D---Y-KK--E-----V-----F---WD--  140
H1N.CM.95.YBF30          --KI-----LV---EI---------......QTQEP-VEP-VGA----Q--ANR----IR.......................--RDN-ESI------E.......----------R-------I---Q-F---F---D-------VW-RK---------M-----IL--WH--  133
H1O.BE.87.ANT70          --NALR-GKFE--AA----------......TRTFPES--C-PG--QI--E-AAR-G-P--.......................H-PQN---L-F--SH.......Q---------A------------G-F---F-------------Y-HK-AE-----V-N---F---W---  133
H1O.CM.91.MVP5180        --NA----KFA--SE--D-------......SSSDPQQ-C-PG---V--E-ATR-G-S--.......................H-PQN---L-F-DSH.......KD-D-----R----------F---F---F--------D---Y-HK-AE------------F---W-C-  132
CPZ.CD.90.ANT            --SA---IKWV-A...-QAI-K............IH-TNP-DI-PCGNE-ASR--L---.......................TIGTEKDVITYS-DH.......T--GT----R----M----E-L-----X--------DX--XXLK-AA---M-MFN---I---W---  125
CPZ.CM.01.SIVcpzCAM13    --AQ-----LV---EI-K-IKE............-PI--EG--EV-K----Y-----R.......................H-PS--QTL---DEMKH.......H---------R------------G-F------RG-------VY--K--E-----V-----F---W---  129
CPZ.CM.05.SIVcpzEK505    --KV-----LV---Q-ID----............QQDP-EG---V-Q--ANR----IR.......................--KEN-Q-L----E-.......K--A-----C------------E-F---F--R--------VW-TK--E-----V-----I---W---  127
CPZ.CM.05.SIVcpzMB66     ---------LV---E--D-L--............TQT--EG--PV----AR------R.......................--SQ--ETL----EV.......QN--------R----------F-G-F---------------VY-KK--E-----V-----F---W---  127
CPZ.CM.98.CAM3           --NR-----LV---D----LSE............LNLL-PG--PV-Q---A---V-TR.......................H-PQN-QTL----EMQE.......HN--------R----------F---F-------K--------Y-IXXXXXX-HGY-----F---W---  129
CPZ.GA.88.GAB2           --N------IV---Q----L--............TQE--EG--EV-Q--ARR-----R.......................H-PQ--QTL----EM.........-Q-D-----R--------------F---------------VY-K---E-----V---------WN--  127
CPZ.CM.98.CAM5           --N------LV--AA--Q-I-Q............TP---SG--PV-Q---AR----TR.......................-H.TKYQTL----EMEN.......H---------R------------Q-F---------------VY-TK--E-----V-----I---W---  128
CPZ.CM.05.SIVcpzMT145    --A------LV---EI---IK-............-PT---G--EV-Q-IASR--V-IR.......................--PQ--ETL----EM.......QD--------K----------F-G-F---------------W-RK--E-----V-----------W---  127
CPZ.TZ.00.TAN1           --NIFGRWP..-ARKAI-DL..............HNTSSEP--Q--Q--QNK-GL-TN.......................TLGTSADVLEYSADH.......T--------R-A--M----E-L-I---W--------D--FF-PK-AA---T-M-N---V---W---  123
CPZ.CM.05.SIVcpzLB7      --S------LV---E--R-I-Q............ETV--EG--P--Q--AR-------.......................--PQ--ETI----E-.......K---IW---K----------F-E-F--GY-----------VY-KK--E----GV-----I---W---  127
CPZ.US.85.CPZUS          --N------IV---E--N-L-Q..........TQTTA--EG--PV-Q--AE-----TR.......................--PQN-QTL---DEMTN.......H-S-------R------------Q-F--GF-----------VY-R---E-----V-----I---W---  131
CPZ.GA.88.GAB1           --T------LV---E--R-I-E............-PT--EG--EV-K---R------R.......................--PE--QTL----EM.......DN--------R----T--------F---------------VY-R---E-----V-----F---W---  127
GSN.CM.99.CN166          --S-N--QQQQESS-ALLSS-GTGQRPYFTL...VDEYGENFWLSPDASD-GRRYYLT.......................EEPKPKRGS-..........--Y-PSC--R-R----DP---VM--------------MFYCED-HQK-EQYA-LEW-LV-GWLSF  131
GSN.CM.99.CN71           --S-N--QQSQESS-ALLSSHGTGQSPSFRL...LDEYGENSWLSPDASDRGRRYYLT.......................EESPSRQNCI..........DY-PSC--R------DP---LM--------------AMFYCED-HQK-ESYC-YEW-IV-GWLQW  130
MUS_1.CM.01.1085         --S-S--QPAQQSLISSPPSPGTGRKQYFKL...V-EYGENSWLSPDASGRGRRYSLT.......................EGASKRPVI.........E-H-.RTC--R-R--I-DP---LM--M-------------MFV-TE-LQK-ET-A-VEW-II-GWLM-  130
MUS_1.CM.01.CM1239       --S-N--QQEQQPL-SSPSSPGTGQSPYFRL...VDEYGENSWLSPDASDRGRKYSLT.......................EGS-ERPLI.........E-H-.RSC--R-R----DP---MMI----Y---------MFYCEE-HQK-ETYT-VDW-IV-GWLQF  130
MUS_2.CM.01.CM1246       --S-N--PQSQERS-PLRSSHGTGPNRYFSL...VKKYGENSWLSPDASARGRKYSLT.......................ENKGRRQII.........E-H-.CSC--R-K--I-DVN--LMI----Y--K-R----MFVCID-LQK-ETYA-VKSKIV-RSLI-  130
MUS_2.CM.01.CM2500       --S-N--KESQQPLMGSPSGPQTGQTPYFKL...VNNYKKHSWLSPNASARGRKYSLT.......................KGS-KKAII.........KKHK.RSC--R--H-IQNP---LIIN-------------MFY-EE-HQK-ETYA-VEW-LI-GWLM-  130
                          Nef start       R17Y mutation                                                    premature stop in Mac239                                                                 max HIV-1 similarity

MAC.US.x.239             ---AI-MRRSRPSGDL-Q-LL-ARGETYGRL...LGEVE-GYSQSPGG-D-GLSSL-C..EGQKYNQGQYMNTPWRNPAEEREKL-YRKQNMDDI.DE*.....DDDL--VS-R-K----T-S--L-I-M---I--------IYY-A--HR---IYLEKEE-II--W-D-  158
H2A.GW.x.ALI             --ASG--KRSGPLQGL---LLQTPGETCGGQ...CSGSGGGYSQSQGGSGRGQKLP-C..EGQRYQQGDFMNTPWRTPATEREKELYKQQNMDDV.DLDD....DDSL--VS---R-Q--T----L---M--LI--R-----MFY-E--HR---IYLEKEE-II--W---  160
H2A.DE.x.BEN             --ASG--KLSKHSRGL---LL-ARGDGYGKQ...RDASGGEYSQFQEESGREQNSP-C..EGQQYQQGEYMNSPWRNPATERQKDLYRQQNMDDV.DSD.....DDDLI-V----R--R-E----L-I-M---I------Q-MFY-R--HR---IYLEKEE-II--W---  159
H2A.SN.x.ST              --ASG--KRSEPSRGL---LLQTPGEASGGH...WDKLGGEYLQSQEGSGRGQKSP-C..EGRRYQQGDFMNTPWRAPA.EGEKGSYKQQNMDDV.DSD.....DDDL--V----R---E---RL-R-M--LI---------YY-D--RRV--IYLEKEE-IIG-W---  158
H2B.GH.86.D205           --SAG--KRSERQQGL--KLL-VPERPYGRL...SGERREQSSRSPGESD-DLNSP-C..................EGQNARG-EGGGQQDAD-.SDE.....DD-VGAICK--I---------L---M---I--Q-----MYY-E--HR---TYFENEE-IVSGW---  142
H2B.CI.x.EHO             --SAG--KQSKQQ-GL---LL-ARRGPRGES...SGERQE-SLQYPGGSD-GLNSP-C..................DGQKTLG-EGGGKQDSD..EDD.....-DN---VR-R-G-------F-L---M--------E---IFY-E--HK---TYLENEE-IVSGW---  141
H2G.CI.92.ABT96          --SAG--KRAPQQLGL-K-LLQARGEPYGKL...W-GLEEGYSQYREESG-GLSSL-C..EGQQYTQGQFMNTPWRNPATERAKL-YRXQNNDDV.DSD.....DNDL--V--Y-K----V-S--L-I-M---I--------IYY-E--HR---IYMEKE--II--W---  159
H2U.FR.96.12034          ---SI--KHSKRSGKL---LLAARGENYGRL...WGGLE-GFELYQGGSG-DLSSRFCEAQGQGYSEGEFMNTPWRTPATGKDKLQYKQQNMDDV.DEE.....DNDL--VA-R-R-A--T----L---M---I------D-IFY---HR---IYLENEE-II--W---  161
MAC.US.x.251_1A11        ---TI-MRRSRSTGDL-Q-LL-ARGETYERL...LGEVE-GSSQSLGE-D-GLSSL-C..EGQKYNQEQYMNTPWRNPAEEREKL-YRKQNMDDI.DEE.....DDDL--VS-R-K----T-S--L-I-M---I--------IYY-A--HR---IYLEKEE-II--W-D-  159
MAC.US.x.251_BK28        ---AI-MRRSKPAGDL-QKLL-ARGETYGRL...LGEVE-GSSQSLGG-G-GLSSR-C..EGQKYNQGQYMNTPWRNPAEEKEKL-YRKQNMDDI.DEE.....DDDL--VS-R-K----A----L-I-M---I--------IYY-A--HR---MYLEKEE-II--W-D-  159
MAC.US.x.EMBL_3          ---AI-MRRSKPAGDL-QKLL-ARGETYGRL...LGEVE-GSSQSLGG-G-DLSSR-C..KGQKYNQGQYMNTPWRNPAEEKEKL-YRKQNMDDI.DEE.....DDDL--VS-R-K----A----L-I-M---I--------IYY-A--HR---MYLE...-II--W-D-  156
SMM.US.x.H9              ---VT--KQRXAGGNLX--LLQARGETYGRL...W-GLEGEYSQSXDASG-GLSSL-C..EPQKYCEGQFMNTPWRNPATEGAKLGYRQQNMDDV.DNE.....D..L--C--S-R--V-I----L-I-M---I--------VYY-I--HK---XYLEXEE-II--W---  157
SMM.US.x.PGM53           --AAG--KQSRQRGGLG-KLLQARGETHGKL...W-GLE-GYSQSRGE-GRDWNLH-F..EGQGYSEGQFMNTPWRNPAREREKLKYRQQNMDDV.DDD.....DD-LI-VS-H-K----A-S--L-I-M---I--------IYY-E--HR---IYLEKEE-IT--W---  159
SMM.US.x.SIVsmH635F_L3   ---AI--KQHRRGGKL---LL-ARGETYGRL...W-GLEEGYSQSLGASG-GLSSL-C..EPQKYSEGQYMNTPWRNPTAEKAKLGYRQQNMDDV.DDE.....DDDL--VS-H-K----A----L-I-M---I--------IYYNEK-HR---MYMEKEE-II--W---  159
SMM.SL.92.SL92B          ---IT-SKPARKHGDL-KKLLSARGDSYGRL...WDGLEEGSQQSQGGSDRALNSCLT.RDGYKYNVGEFMNTPWRNPATERQKQQYRQQ-MDDI.DED.....DD-LA-VV-W-K----A-S--L-I-M---I------D-IYY-E--HR---IYLEKEE-II--W---  160
STM.US.x.STM             --ASG--KQRKQHGEL---LL-ARGETYGKL...L-GLGEGS-PSQGASD-GLNSH-C..EPQRYNEGQFMNTPWKNPAAESAKLEYRQQNMDDV.DEE.....DDNL--VA-H-R----E----L-I-----I-S------IYY-E--HR---MYLEKEE-IV--W---  159
MNE.US.x.MNE027          ---AT--KRSKSPGDL-Q-LL-ARGETYGRL...W-GLE-GYSQSLGGSD-SLSLL-C..EGQKYNQGQFMNTPWKNPAGEREKL-YRKQNMDDI.DEE.....DDGL--V--R-R----A----L---M---I--------IYY-E--HK---IYLEKEE-IV--W-D-  159
GRV.ET.x.GRI_677         --SSN--RQQQ-LLKLW.-GL-GKPGADWVLL..SD-LIGQSSTVQEECG-ALKKSWG.........KGKMTP.....DGRRLQEGDTF.DEW.DDD.....---.-----Q-R----Q----L---F-----S----D-IYY-E--EK--N-YALNEW-IID-W-A-  143
VER.KE.x.TYO1_patent     --SQN--PAHKKYSKLWQALHKTHVTRYGLL...-D-LIGTSSTVQEECD-ALRKSLI.......RKQNGNMTE.....EGRRLQEGDTW.EEW.SDD.....---.-----R-R----Q----L---F----------D-IYY-D--NK--N-YALNEW-IID-WNAW  146
VER.KE.x.9063            --LGN--PQHKKHVSLWHALHKTQVTKYGLL...-D-LIGTSS-VQEEYD-GLRKSLI.......RKRNGNMTP.....DGRRLQEGDEW.DEW.SDD.....-D-.-----K-R----Q----L---F----------D-IYY-E--NR--N-YALNEW-IID-WNAW  146
VER.KE.x.AGM155          --LGS--PQHKKQL-IWRALHATRHTRYGLL...-D-LIGQSSTLQEECD-GLRKSLI.......RKRNGNMTP.....EGRRLQDGDQW.DEW.SDE.....-D-.-----R-R----QI---L---F----------D-IYY-D--NK--N-YALNEW-IID-WNAW  146
VER.DE.x.AGM3            --LGN--PQHKKQLSLWHALHKTRATRYGLL...-D-LIGQSSTLQEECD-ALKESLI.......RKRNGKMTP.....EGRKLQEGDKW.DEW.SDE.....-D-.-----R-R----Q----L---F----------D-IYY-D--NQ--N-YALNEW-IID-WNAW  146
TAN.UG.x.TAN1            ---SN--REQQ-LLRLWRAL-KAPVVRYGML...-D-LIGQSSNIQEECD-NWNGG-T.........RRGKSTP.....EGRKL-ADDTW.DDW.EPE.....---.-----R-R----Q----L---F------------IYW-PK-EQ--N-YALNEW-IID-W-A-  144
SAB.SN.x.SAB1C           ----S--QQQRHSLWLWSKL-QAPVIQYDML...-D-LLGQSSHIQEECA-SLRDGLI.......RQGDSSRTE.....EGVKMKHQGRQ.PSWYDED.....---.-----R-CL---A----L-I-FG-----------IYY-E--KK----YALNEW-IVDGW---  147
OLC.CI.97.97CI12         --SIC-RRNRE-EVLY-RLRSNWGEKEQKQL..............ESEDQHQTNLAI......................................ESLMTGDTDTDDLE----K----V-QA---TLIE-------------MYW-P--ENLIY-YALNCW-IIGNF-AT  118
DRL.x.x.FAO              ---QN--KREEAYARFYK-L--GYGAKGGNLDYHQLEPSEPLLRKWQPFPGGLDKAQRSSSTS.............SLDTTHMQ--AAEKDVT-PTEEEQPSEE---------C--R---EP---DL--F----------K-IWW-L--ET----YAQNEW-FIKGW-S-  157
RCM.NG.x.NG411           ----S--A-AA-LRRW-GL-LTTPGEDYARFAETLQDGQP-CAEG-GRASRDFLTR..........................GGFTIETQQSVDAIDWYEDTDDTL-----R----I---S--L-I-------------IYF-E--RR----YMENEH-II--W---  142
RCM.GA.x.GAB1            ----S--N-AA-LLRW-FK-LTTPGEGYVRWHETLLDGQPWCAEG-GRASRDFV-R..........................GGIT-ETQASIDDIDWYEDTDDTL-----K--------S--L-I-M-----------IYW-I---R---MYLENEH-II--W---  142
MND_1.GA.x.MNDGB1        --SSQ--KRSEA-VRYSSAL-QLVGGPVTPDGYKQIESSQGAEKQ-LLRGRAYGTY-EGLDKVQNDPLTK......................DEKLDLTQQDPE---------CR--S--VPS--DLI-F---I------G-IYY-R--EE----YAENEW-FE-GW-Q-  148
MND_2.x.x.5440           ----S--QREQHYIKYYKAL--GYGAEGTNLDYQVLEEQQPLAG--GTSLAASDKVLRPYSTE.....................RDLRHQQDVTLPSEKEQPSD---------Y-TQ-VPEA---DLI-M------------IWW-K--EE----YAQNEW-LI-NW---  149
MND_2.CM.98.CM16         ----S--QQEEKYLKYYKA---GYGAEGTNGDYQQLH-SEPLLG-LSTSQEEFDREQKSSSTD...................................----T----Y--C-V-EP---DL--MP-----------IW--K--EE----YAQNEW-FI-TW-S-  129
MND_2.GA.x.M14           ----S--QREQ-YAKYYKAL--GYGAEGTNLDYQLLEEQRLLEGPLGTSLGASDKELKSCSTK..........................VDDEQK--R----Y-TQ-V-EA---NLI-M------------IWF-R--EE--N-YAQNEW-FI--W-E-  133
DEB.CM.99.CM40           ----N--IPSR-QVGSFGSGS-GLLRWRYRDLSEQ-EQFSECLLE-DREQSSSSTE..........................FWGSPRREIKCTNKQQQDLDQT-A-----R---SI-DP---LMI-Y-----------DVFY-A--HA--E-HAQNEW-II-GWLQ-  142
DEB.CM.99.CM5            ----N--TPSF-QGKKSGSGSKGLLRWRYKDLSGEAETFSPFLQE-DKEQSCYSTE........................EPLYADPHREIKCT-RAQECYD--HGI--A-R-R--I-EP---LMI-Y--Y---------DIFY-A--HA--EIHAH-EW-II-GWLK-  144
DEN.CD.x.CD1             --S-N--QASGI-GLAPSNKEGRQRKWYWRLLPEQQEKWWPSPDG-DRAL-LSSTR.........LLEHDEDEETFVDAEEIIQSSPALDPQHPRIHQYVVDHDD--------R-RR--HAP---DMI-M----------QDIFY-P--HA----YAQNEH-FITGW-T-  161
TAL.CM.01.8023           --SVK--LLGETASQRLQGSPGAGRLVYWRLRDGQCRQLQ-SRGERGKGLNIALLD..................AGWQASDANWMGK-LSEGEVWYEKEQEENDM-L-----R-------CN--LGI-F-F---------K-IFYNN--HA--N-YAHNEW-IL--W-S-  151
TAL.CM.00.266            --SVK--LLGETASQRLQGSHGGGRMVYWRLRDGQVKRSQLSPG-RGKDLNIALLD..................EEWQATGADYMGK-LTQGEVWYEKEQEENDM-L-----R-K----TC---LGI-F-F---------K-IFYNK--HA--N-YAHNEW-IL--W---  151
MON.CM.99.L1             --S-N--QQSPLSS-PLLGSQSSGRMRYFMLEDDYGEQSWLSPD--DRER-YSLTE..............................GRNGKQRRQPLDDDDDDDG--C--R-R----DP-W-LMM----Y-------GEMFYCED-HRKIEQYA-LEW-LI-GWLQ-  137
MON.NG.x.NG1             --S-X-XQPAQEPLMPSHGSQSSGQTRSFAWEDDYGEDSWLSPD--DRGRRFSLTE..........................GR-HRSNRRPTTVVDDADEDQDLP-C--R-M----DP-W-IMM------------DK-FYCXDXHQK-EQYA-VMW-LV-GWLQF  141
WRC.CI.97.97CI14         ---VF--T--EEHMERYA-RLQRAGRRIRRQRYQQLA-ERQQQ-VESGCLLPDSAQALIN..............................AE--SVKEDSSD-D-G----------I-QP-F-LL-----LF--------DIFW-A--EE-VN-YAHNEW-LIDGW-A-  138
WRC.CI.98.98CI04         ---IF--E--EQRMHAYA-RLQRAGRKQRRKRYQKIATERQQV-VESGNLLPDSAQALIR..............................AE-DSSKEDSS.DD-------H----I-QA-F-LL-----LF--------YIFW-P--EE-VN-YAFNEW-LIEGW-A-  137
WRC.GM.05.Pbt_05GM_X02   ---VF--D--EQRMLRYR-RLLRVGRKERRRRYTRLA-ERQAE--ASGELLPDSAQTLIR..............................AE-DSSKEDSS.DDDAI----R------QA-F-IL-----LF--------DIFW-P--EE-VN-YA--EW-IVEGW-A-  137
SUN.GA.98.L14            --NAFGRP-EV--VRTLF-L-AGSGTRAEPAGREYHRLRRQEVE...............................................PLVS-ENGGPNGI-Q---E---R--R--CKP---QLI-----I---------WY-RT-EE----YAENEW-FITGW-D-  123
LST.CD.88.524            --NIFG-P-ADS-WR-LRKL-AGSGTRAEGTEESYQKLMEQTLG..........................................WDKTEEDG.DSDS------R--R--CSS---TLI-----I---------YW----N---T-YCENEW-LIG-FM--  119
LST.CD.88.485            --NIFG-PAAD--WKTLR-L-AGAGTRSEGTEETYQQLMRETSD..........................................WDKRKDGDSDS-G------R--R--CSP---TLI-----I---------YW--------C-YCENEW-LIG-FMT-  120
LST.CD.88.447            --NIFG-PAAD--WKTLR-L-AGAGTRSEGTEETYQQLMQETSD..........................................WDKRRDGDSDS--------R--R--CSP---TLI-----I---------YW--------C-YCENEW-LIG-FMT-  120
LST.KE.x.lho7            --NAFGRPAAD--WKTLR-F-AGSGRNAEGTEDSYRALIEKTSN..........................................W-KRRDGDSDS--------K--R--CSP---TLI-----I---------YW--------I-YCENEW-LIG-FM--  120
SYK.KE.x.SYK173          --STS---QQLRSEGKYAIGW-LFGKQYTPLPDELSRPLQPCRGGFDKAWRSTLTE..................PIDPHGPDRDWGHSGGQKFSPGDIVQDEGDTGL-----C--T---TL---L-I-----I-N----Q-MNYCEK-DE--H-YLQNEH-II.-RI--  150
SYK.KE.x.KE51            --SQS--KQPSR-DEKWRTRWWPFGKPYSPMPDELLRMSQPYHEDFDRGWRSTLTE.....................PILDPKRDFDSGGKKWN-GDICHDEGD-DL-----L-------L---L-I-----I-R----Q-MFFC-K-HE--Q-YLKNEH-VI.-DIT-  147
COL.CM.x.CGU1            --SLL-CLNIRQQ-QT-GGVLGRLFGIYWYPGVQ-FRFLRHFRGGYDLRRVA-VERLNICEDGYMTVPKVLPN.........CPCPLCYRHPDVYYHNPEPEDV-DV-A-CY-MSRL-V-QA-I-LL--I-C----------I-RTPE-D-LIEQYA-IEW-CLKGWLE-  161
```

```
                                                                                                                        normal Nef end
H1B.FR.83.HXB2            TPGPG.......VRYPLTFGWCYKLVPVE.PDKIEEA.NKGENTSLLHPVSLHGMDD...PEREVLEWRFDSRLAFHH..............VARELHPEYFKNC..........................*   205
H1A1.UG.85.U455           -----.......I---------------D.-AEV---.TG---N-----ICQ--V--...E-K---M-K---T--LK-..............R-Y-----FY-D..........................-   206
H1B.US.90.WEAU160         -----.......T----C----F-------.-E-V---.-E---N-----M--------...H-K---M-K---K-----..............-----------D-..........................-   204
H1C.ET.86.ETH2220         -----.......----------F-----D.-SEV--I.-E---NC----A-----E-...ED----K-K---H--RR-..............M--------Y-D-..........................-   208
H1D.CD.84.84ZR085         -----.......I---------FE----D.-EVV-K-.TE--DNC----ICQ---E-...E-K---V---N-----E-..............K-K-KY-------..........................-   216
H1F1.BE.93.VI850          -----.......I-----L---F-----D.-EEV-K-.-E---NC----M-Q---E-...ED----R-K---S--LR-..............I---R---FYQD..........................-   206
H1G.SE.93.SE6165          -----.......T-F-------F----MD.-AEV---.-----N-----ICQ---E-...ED----V-----S--RR-..............I--------Y-D-..........................-   208
H1H.CF.90.056             -----.......E-F-------F-----N.-QEV-Q-.-E---N-----M-----E-...DG----M-K------LT-..............L--VK----.-D-..........................-   207
H1J.SE.93.SE7887          -----.......I---------------D.-SEV---.-E---NC----ICQ--IE-...E-----Q-K---S--RR-..............I------FY-D-..........................-   204
H1K.CM.96.MP535           -----.......I---------------D.-AEV--T.TE--DNC----INQ---E-...EH--I-M-K---S--RR-..............-------D-Y-D-..........................-   216
H101_AE.TH.90.CM240       -----.......I-F--C----F-----D.QREV--D.-----NC----M-Q--IE-...E-----M-K---A--RK-..............----Q----Y-D-..........................-   205
H102_AG.NG.x.IBNG         -----.......T-F-------F----MD.-AEV---.-E---N-----ICQ---E-...DD----I---------R-..............T--------Y-D-..........................-   207
H103_AB.RU.97.KAL153_2    .--IR........F-------------D.-AEV---.TE---N-----ICQ------...E-K---M-K------LT-..............R-------FY-D-..........................-   203
H104_cpx.CY.94.CY032      -----.......E-F--C----F-----D.-QEV---.TE----C----I-Q---E-...E-----K-K------YK-..............---------FY-D-..........................-   220
H1N.CM.95.YBF30           -----.......I---V-----F----LS.AEEV---.-E-D-NA----ICQ--A--...DHK---V-----S--RR-..............---------FY---..........................-   213
H1O.BE.87.ANT70           -----.......T-F------LF-----S.EEEA-RLG-TC-RAN----ACA--FE-...THK-I-M-K--RS-GNT-..............--MIT---L-QKD..........................-   214
H1O.CM.91.MVP5180         -----.......P-F------LF-----S.AEEA-RLG-TN-DA-----ACN--AE-...AHG-I-K-Q--RS-GLT-..............I-LQK---L-PK...........................-   212
CPZ.CD.90.ANT             -EE--.......-----CR---F------.SPPD......D-RNI----ACT--DG-...-HK-I-R-E--AS-MRR-..............I---R-----RD...........................-   199
CPZ.CM.01.SIVcpzCAM13     -----.......E---PC----F----LT.TEQV-A-.-E-D-NC----ICQ---E-...ESK---I--------LR-..............I---K----Y-D..........................-   208
CPZ.CM.05.SIVcpzEK505     -----.......----------F----LS.-EEV---.-Q-D-NV----MCQ---E-...-DK---V-----S--RV-..............R-------FYQ--..........................-   207
CPZ.CM.05.SIVcpzMB66      -----.......I---------F----L-.-EEV-R-.-E-D-NI----ICQ--QE-...EA----V-T------LK-..............R--------Y-D-..........................-   207
CPZ.CM.98.CAM3            -----.......------Y-G-F----LT.-EEV-Q-.-E-D-NI----ICQ---E-...AHG---V--------LR-..............I---Q----YQ-...........................-   208
CPZ.GA.88.GAB2            -----.......------Y---F-----S.-EAV---.-----DNL----LCT--FE-...E-K---I-K---Q--LR-..............L---K----YRD...........................-   206
CPZ.CM.98.CAM5            -----.......----------F----LT.-EEV-K-.-E-DTNI----ICQ---E-...EHG---V--------LK-..............L---Q----YGD...................HNPAPAQ-   214
CPZ.CM.05.SIVcpzMT145     -----.......----------F----LP.-EEV-K-.-E---NC----MCQ---E-...E-----V-T---K--LR-..............L---K----YRD...........................-   206
CPZ.TZ.00.TAN1            -----.......I----CR--LF-----D......PP.EDD-KNI----ACS--TT-...-DG-T-I-----S--RR-..............I---RY-----...........................-   196
CPZ.CM.05.SIVcpzLB7       -----.......I---I---G-F----LT.-EEV---.-----NL----ICQ---E-...EDK---I-KY--Q--LR-..............I-----S--Y---..........................-   207
CPZ.US.85.CPZUS           -----.......------Y---F----LT.EEEV-Q-.----TNI----MCQ---E-...EHG---I-Q--TE--RR-..............R-K-------R-...........................-   210
CPZ.GA.88.GAB1            -T---.......T-F--C----F----LT.EEQV-Q-.-E-D-NC----ICQ---E-...EDK---V--------LR-..............I---Q----Y-D...........................-   206
GSN.CM.99.CN166           -----.......T---TIP-F-IC-R--A..TTEDSE.PGDDEYL-T--AYQGRSE-...QHK-F-VFS-C-K--IKSG.............IQLDQLQQEERKMRLTANRFL...................-   219
GSN.CM.99.CN71            -----.......I---TMP-F-WC-R--A..MTEDSE.PGDDQYL-N--AYQGQQE-...HH--I-VFS-C----LKSG.............WQMNQLQQEERKKRLTANRFL...................-   218
MUS_1.CM.01.1085          -D---.......T---KIP-FLWC-RE-A..MTEGSE.EGDDQYL-D--AFQGREE-...HHKQF-VFS-C----LKSG.............LQLDQMQQEERKKRLATNHIL...................-   218
MUS_1.CM.01.CM1239        -E---.......T---TLP-FLWC-RE-A..ITEDSE.EGD-EYL-T-AAYQGKEE-...-HKQF-VFS-C----MKSG.............RQLDQIQQEERKRRLTANRIL...................-   218
MUS_2.CM.01.CM1246        -D---.......T---RQP-FLWC-R--A..MTE-SE.PGDDQYL-T--AYVGRDE-...HHK-F-VFS-C----LKSG.............PQLNQIQQKERKRRLTANRIL...................-   218
MUS_2.CM.01.CM2500        -E---.......-----MP-VLFC-R--H..ITENSE.LGDDEHL-C--AFIGREE-...-HK-F-VFS-C----LKSG.............RQLNQMQQEERKRRLAANRILSQSP...............-   222
                                   max HIV-1 similarity                                                                     normal Nef end
MAC.US.x.239              -S---.......I---K----LW-----N..VSD-AQ.ED.-EHY-M--AQTSQW--...-WG---A-K--PT--YTYE...AYVRYPEEFGSKSGLS-EEVRRRLTARGLL..NMADKKETR.........-   263
H2A.GW.x.ALI              -H---.......I---MF---LW-----D..VPQ-GE.DT.-THC-----QTSRH--...THG-T-V----PK--HDYK...AFILHPEEFGYKSGL--DEWKARLKARGIP..FSKNRNS...........-   263
H2A.DE.x.BEN              -H---.......----MY---LW---S--..LSQ-AE.ED.-ANC-V--AQTSRH--...EHG-T-V-Q---M--YNYK...AFTLYPEEFGHKSGL--KEWKAKLKARGIP..YSE...............-   258
H2A.SN.x.ST               -H---.......----KF---LW-----D..VPQ-GD.DS.-THC-V--AQTSRF--...-HG-T-V----PT---SYE...AFIRYPEEFGYKSGL--DEWKARLKARGIP..FS................-   256
H2B.GH.86.D205            -H---.......I---KY---LW------..VPAATR.EEE-THC-M--AQISSW--...IHG-T-I-Q---L--YDYV...AFNRFPEEFGYQSGL--EEWKARLKARGIP..TD................-   241
H2B.CI.x.EHO              -H---.......----KF---LW----IN..MIA-PE.DE.-THC-V--AQTSAW--...-HE-T-V-Q---L--YDYV...AFSRFPEEFGYQSGM--KEWKAKLRARGIP..TE................-   239
H2G.CI.92.ABT96           -A---.......I---X----LW-----T..VSD-AQ.ED.-THC-V--AQTSPW--...--T---A-K--PT--YDYR...AFILHPEEFRWKSGL--AVWKEKLKQRGLP..IE................-   257
H2U.FR.96.12034           -S---.......T-W-MC---LW--E--D..ANDKAQ.ED.-RLY-VGSAQTSCEE-...HWG-A-V-K---S--YSYQ...AFIKCPEEFGSKSGLS-EEVKRRLTARGLP..VKNC..............-   261
MAC.US.x.251_1A11         -S---.......I---K----LW-----N..VSD-AQ.ED.-EHY-M--AQTSQW--...-WG--PA-K--PT--YTYE...AYVRYPEEFGSKSGLS-EEVRRRLTARGLL..NMADKKETR.........-   264
MAC.US.x.251_BK28         -S---.......I---K----LW-----N..VSD-AQ.ED.-RHY-MQ-AQTSKW--...-WG---A-K--PT--YTYE...AYARYPEE#GSKSGLS-EEVRRRLTARGLL..NMADKRETR.........-   263
MAC.US.x.EMBL_3           -S---.......I---K----LW-----N..VSD-AQ.ED.-RHY-MQ-AQTSKW--...-W#---T-K--PT--YTYE...AYVRYPEEFGSKSGLS-EEVRRRLTARGLL..NMADKRETR.........-   260
SMM.US.x.H9               -A---.......I---MF---LW--X--N..VSD-AQ.ED.-THC-M--AQTSQW--...-WG---A-K--PX--YXYK...AFVEHPEEFGSQSGLSKEEVQRRLTARGLL..KMADKKKTS.........-   262
SMM.US.x.PGM53            -S---.......I---MF---LW-----S..ASD-AQ.ED.-THY-V--AQISQW--...-WG---A-K---Q--YRYE...AFIRHPEEFGSKSGLS-EEVKRRLTARGLL..KMADKKETS.........-   264
SMM.US.x.SIVsmH635F_L3    -L---.......T---MY---LW-----D..VSD-AQ.ED.-THC-V--AQT-QW--...-WG---A-K--PE--YSYK...AFIKYPEEFGSKSGLS-EEVKRRLTARGIY..KMADKRETS.........-   264
SMM.SL.92.SL92B           -S---.......I---KF---LWQ----D..VSD-AN.-D.-THC-V--AQTYQHS-...-WGK--A-K-NPH--YTYE...AFVRHPEEFGWKSGLT-EEVERRLANKPKPQKKMADKKETS.........-   267
STM.US.x.STM              -A---.......I---KQ---LW-----D..MSN-AQ.EDDGTHY-V--AQT-QW--...-WG---V-K--PL--HTYE...AFVRHPEEFGSKSGL-KEEVERRLTARGLL..KMADKKETS.........-   265
MNE.US.x.MNE027           -S---.......P---K----LW-----N..VSD-AQ.EG.-ENY----AQTSQW--...-WG---V-K--PT--YTYE...AYIRYPEEFGSKSGLS-EEVRRRLTARGLL..KMADKKETS.........-   264
GRV.ET.x.GRI_677          S----.......I---RV--F-F-----D..LHE-AR.-C.-RHC-M--AQMGEDP-GI.DHG---V-K--PK--VEY...........RPDMFK..DMHEHAKR.........................-   224
VER.KE.x.TYO1_patent      SK---.......I-F-KC--F-F-----D..LHE-AQ.TC.-RHC-V--AQMGEDP-GI.SHG-I-V-K--PM--IQY...........DPNREYFTDMHGLVKRK........................-   230
VER.KE.x.9063             ST---.......----RC--F-F------..MHE-AE.TC.-RHY-A--AQIKEDP-GI.SHG-T-V-K--PM--VQY...........DPNRQYFEDMHAIVKRK........................-   230
VER.KE.x.AGM155           SK---.......I---RC--F-F-----A..LHE-AE.TC.-RHC-V--AQ--EDP-GI.NHG-I-A-K--PM--VQY...........DPSREYFTDLYSTVGTGN.......................-   231
VER.DE.x.AGM3             SE---.......I---RC--F-F-----D..LHE-AE.TC.-RHC-V--AQVREDP-GI.NHG---V-K--PM--VQY...........DPNRKYLTDMHDLGKRK........................-   230
TAN.UG.x.TAN1             S----.......T-K-RC--F-FE----D..VSQ-AQ.D..-RHC----AQIEWES-...-WK-T-V-K-NPL--VQY...........NPD..SFKDMHGLVKRK........................-   223
SAB.SN.x.SAB1C            -D---.......T---KC----F-----D..LSE-AK.-S.--HC----AQVAYE--...AWK-T-V-K--PL--VDY.VAWRLHPEQVPSA.QG..................................-   227
OLC.CI.97.97CI12          EEVEE.....GTL-R-K-P--LWD-RR--.-LPETLEERG..........................................................................................   152
DRL.x.x.FAO               -K---.......I---KM--FLF-------..VDEDLV.-Q.QCNR--NSSQMGPI--....--G-R-M-Q--PG--YTF.QACIHHPEEFGHVKSIKSVE-PG......CCKRKWWQFKPTTEGCHGDNLQKC-   269
RCM.NG.x.NG411            -----.......I---TM---LWQ----D..VSD-AT.EN.-RHC----AETS------...-WG---A-K-KPL--VDY.AGYRLHPEFFGER.KNKTK.............................-   225
RCM.GA.x.GAB1             -----.......I---TL---LWQ----D..VSD-AR.ED.-EH-----AETS--E-...-WG---A-K-NPM--VDY.IGYRLHPEFFGERKNKTQ..............................-   225
MND_1.GA.x.MNDGB1         -T---.......T---K---FLF--E--S.RAIGD-Y.AA..-NH---SSQ-CPQE-...--G-T-M-SGTLI-PMTL.QH..............................................-   215
MND_2.x.x.5440            -K---.......I---K---FLF-----A..VSPPLE.ED.-CNR--NSSQMGIQE-....--G-R-M-K---G--YTF.YAPIIRPEEYKCV.TSLSYEA-K-EEKPDCCKRKWWQF...............-   251
MND_2.CM.98.CM16          -D---.......I---K--RFLF--C--A..VPPDQE.-N.-CNK--QSSQ-GIQEE..#-WG-R-V-K---G--YTF.YAPIKRPGEFKHV.QSIS-IA-A-EHKPECCKRKWWQF...............-   231
MND_2.GA.x.M14            -S---.......I---KR--FLF--I--Q..VPPDQE.-Q.-CNR--NSSQ-GIQEE..#-WG-R-M-K--PE--YTF.YAPIKCPKEYRHV.TSLSYTA-QE..KSDCCKRKWWQF...............-   233
DEB.CM.99.CM40            -K---.......----IY--FLF-------..IADPDY.END-RNI---DAHQGM-E-...-H--R-V-K---T--YCY.............K-GH.AEQR..EHTRRC.MFPKRK................-   228
DEB.CM.99.CM5             -E---.......P---TY--FLF-------..IADPDY.END-RNI---DAHQGQAE-...-HK-R-V-K---S--YCY.............K-GHEAHKP..EHNRRC.MFPKRK................-   231
DEN.CD.x.CD1              -D---.......I----E--FLW-----T..IEE-YD.--.-ENC---DRYEGQQA-...-W--T-V----PE--WCY.............K.AGHKKLQTSQHNNYCQRAKPRK................-   249
TAL.CM.01.8023            -E---.......T----C--ILW--C----..IHEDDT.ED.GH.L-M--AYDGQQE-...-WG-A-V-V--EK--YTP....................GAKMAEYDRLERKKKELLAPPQTAS..S......-   241
TAL.CM.00.266             -E---.......T----C--ILW--C----..IHEDDT.ED.GH.L-P--AYDGQAE-...-WG-S-V-V--EK--YTP....................GAKMAEWDRLEREKRMLLAPPQTASS........-   241
MON.CM.99.L1              -E---.......----TMP-FLWC-R--A..TTEDSE.EGD-DFL-T--AYQGR-E-...-H-QF-VFS-C-K--VKS....................GRQLAQLQQEERKKRLAANRIL............-   225
MON.NG.x.NG1              -X---.......----TIP-FLWC-R--A..IQEDSE.DGDXEFL-T--AYQGRDE-...-HKQF-VXS-C-K-GVKS....................GPQLDELQQEERKRRLTANRIL............-   229
WRC.CI.97.97CI14          -E---.......I-W-K-----W-----HFSEESDNTDPYY-RNK----ACA.EPE-...-WG-H-V-K--PK--CDFVAGRLPGQSIATGKE-I-SL......Y.K.NKKQCK..................-   234
WRC.CI.98.98CI04          -E---.......I-W-K-----W------FSEESNDSSQYY-KNK----GCQ.ESE-...-WG-H-V-K--P---CDFVASRLPGHIIAQGED-I-TL......H.KEKLKRHK.................-   234
WRC.GM.05.Pbt_05GM_X02    -D---.......----K-----W-----HFSEESPNSDQGF-KNL----AVA.ESE-...-WK-Y-V-K--P---CDFVAGRLADGQIATGIETL..Y......E.M.KKRNNK.................-   231
SUN.GA.98.L14             -K---.......----KA---LW--A--T..IDEDRD.PNHPCQA---SSQQGVNE-...-WG-R-I-T--PT--YDF.RAIQKHPEEFKHVTSLQWE......V..........................-   209
LST.CD.88.524             -D---.......T----A---LWQ-E--A..CDEFKD.PSD-RQC---SSQ-GVLE-...-WG-R-I-H-NPM--VDY.VALRKQP.................ASVQATAMRFNCERK.............-   207
LST.CD.88.485             -D---.......T--------LWQ-E--A..CDEYTD.PSDASQC---SSQ-GVQE-...-WG-R-I-H-SPM--VDF.VALRKQP.................KSIQASSLSLNCKRK.............-   208
LST.CD.88.447             -D---.......T-F--A---LWQ-E--A..CDEYTD.PSDYRQC---SSQ-GVQE-...-WG-R-I-H-NPM--VDY.VALRKQP.................KSIQASAIGL..QRK.............-   206
LST.KE.x.lho7             -D---.......T--------LWQ-E--A..CDEYKD.PSD-TQC---SSQ-GVLE-...-WG-R-I-H-NPM--VDF.IALKKQP.................AKIQNTAFAFDCKRK.............-   208
SYK.KE.x.SYK173           -S---.......T----I---LWE---N-..IEGCLE.YE.-H-L----A-GQ-SSSMGE-HV-LQPPPGYTPGWEMA.RLQLERQTGKPQE...LQSALSKN..IS.......................-   241
SYK.KE.x.KE51             -S---.......T----I---LWE-A-N-..IEGYLV.DE.-D-LMM--AAGV-ASED..-H--N-M-N-NPH--YTPGWEMARQQLERQTGKR.................................-   229
COL.CM.x.CGU1             EDEL-EDGALKED-K--VA--LW---YI-..QLG-Y-YSYDLSLLSVT...........................................................SRKKKKPQQVAIEMVD......-   224
```