V

HIV-1/SIVcpz Proteins

Contents

V-1	Introduction		**307**
V-2	Annotated features		**308**
V-3	Sequences		**310**
V-4	Alignments		**318**
	V-4.1	Gag	318
	V-4.2	Pol	326
	V-4.3	Vif	340
	V-4.4	Vpr	344
	V-4.5	Tat	346
	V-4.6	Rev	348
	V-4.7	Vpu	350
	V-4.8	Env	352
	V-4.9	Nef	366

V-1 Introduction

The HIV-1/SIVcpz protein alignments are based on the complete genome nucleotide alignment, but in some cases a few sequences were removed because they were too short to be informative (especially in Nef), had many stop codons or a problematic segment of amino acids. As with the other alignments in this compendium, they are intended to display the genetic variation of the world-wide HIV epidemic in a compact form. They are annotated in more detail than the complete genome nucleotide alignment.

V-2 Annotated features

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

Feature	Protein	Location	Page
Gag start, p17 start	Gag	1	318
membrane binding	Gag	1-30	318
phosphorylation site	Gag	112	318
p17 end	Gag	132	318
p24 start	Gag	133	318
CyPA binding	Gag	205-241	320
major homology region	Gag	285-304	320
p24 end	Gag	364	322
p2 start	Gag	365	322
p2 end	Gag	377	322
p7 start	Gag	378	322
Zn motif	Gag	392-404	322
Zn motif	Gag	413-425	322
p7 end	Gag	432	322
p1 start	Gag	433	322
p1 end	Gag	448	324
p6 start	Gag	449	324
Vpr binding	Gag	455-460	324
Vpr binding	Gag	489-494	324
p6 end, Gag end	Gag	501	324
Pol start	Pol	1	326
Gag-Pol TF start	Pol	1	326
Gag-Pol TF end	Pol	56	326
protease start	Pol	57	326
protease end	Pol	155	328
p66, p51 RT start	Pol	156	328
M41L	Pol	196	328
D67N	Pol	222	328
K70R	Pol	225	328
D110 catalytic site	Pol	265	328
polymerase motif	Pol	337-342	330
T215Y	Pol	370	330
K219Q	Pol	374	330
p51 RT end	Pol	595	332
p15 RNase H start	Pol	596	332
p66 RT, p15 Rnase H end	Pol	715	334
p31 Integrase start	Pol	716	334
p31 Integrase end	Pol	1004	338
Pol end	Pol	1004	338
Vif start	Vif	1	340
Vif end	Vif	193	342
Vpr start	Vpr	1	344
oligomerization	Vpr	1-41	344
amphipathic α-helix	Vpr	17-33	344
H(S/N)RIG motifs	Vpr	71-83	344
frameshift in HXB2	Vpr	72	344
Vpr end in HXB2	Vpr	79	344
Vpr end	Vpr	98	344

Feature	Protein	Location	Page
Tat start	Tat	1	346
C-rich region	Tat	22-37	346
nuclear localization	Tat	49-57	346
exon 1 end	Tat	72	346
exon 2 start	Tat	73	346
Tat end	Tat	102	346
Rev start	Rev	1	348
exon 1 end	Rev	25	348
exon 2 start	Rev	26	348
NLS	Rev	34-49	348
Leu-rich effector domain	Rev	75-83	348
Rev end	Rev	117	348
Vpu start	Vpu	1	350
transmembrane domain	Vpu	1-27	350
cytoplasmic domain	Vpu	28-82	350
α-helix	Vpu	43-51	350
phos	Vpu	53	350
phos	Vpu	57	350
α-helix	Vpu	58-70	350
Vpu end	Vpu	83	350
Env start	Env	1	352
signal peptide end	Env	30	352
gp120 start	Env	31	352
glycosylation NVT	Env	88-90	352
CD4 binding	Env	124	352
V1	Env	131-156	352
glycosylation NDT	Env	136-138	352
glycosylation NSS	Env	141-143	354
glycosylation NCS	Env	156-158	354
V2	Env	158-196	354
glycosylation NIS	Env	160-162	354
glycosylation NDT	Env	186-188	354
glycosylation NTS	Env	197-199	354
CD4 binding	Env	196	354
glycosylation NKT	Env	230-232	354
glycosylation NGT	Env	234-236	354
glycosylation NVS	Env	241-243	354
glycosylation NGS	Env	262-264	356
glycosylation NFT	Env	276-278	356
CD4 binding	Env	279	356
glycosylation NTS	Env	289-291	356
glycosylation NCT	Env	295-297	356
V3	Env	296-331	356
glycosylation NNT	Env	301-303	356
V3 tip	Env	312-315	356
glycosylation NNT	Env	339-341	356
glycosylation NKT	Env	356-358	356
CD4 binding	Env	365	356
V4	Env	385-418	356
glycosylation NST	Env	386-388	356
glycosylation NST	Env	392-394	358
glycosylation NST	Env	397-399	358
glycosylation NNT	Env	406-408	358

Feature	Protein	Location	Page
CD4 binding	Env	425	358
glycosylation NIT	Env	448-450	358
CD4 binding	Env	455	358
V5	Env	460-471	358
glycosylation NES	Env	463-465	358
CD4 binding	Env	469	358
fusion peptide	Env	512-527	360
gp120 end	Env	511	360
gp41 start	Env	512	360
immunodominant region	Env	588-607	360
glycosylation NAS	Env	611-613	360
glycosylation NKS	Env	616-618	360
glycosylation NHT	Env	624-626	360
glycosylation NYT	Env	637-639	360
transmembrane domain	Env	685-704	362
gp41 cytoplasmic tail start	Env	705	362
glycosylation NGS	Env	750-752	362
glycosylation NAT	Env	816-818	364
cytoplasmic tail end	Env	857	364
gp41 end	Env	857	364
Env end	Env	857	364
Nef start	Nef	1	366
myristoylation	Nef	2-7	366
acidic cluster	Nef	62-65	366
poly-P helix	Nef	69-78	366
phosphorylation	Nef	77-81	366
HXB2 premature Nef end	Nef	124	366
normal Nef end	Nef	207	368

V-3 Sequences

Sequences included in the HIV-1 protein alignments.

Name	Accession	Proteins	Author	Reference
B.FR.83.HXB2	K03455	All	Wong-Staal, F.	*Nature* **313**(6000):277-284 (1985)
A1.AU.03.PS1044_Day0	DQ676872	All	Li, B.	*J Virol* **81**(1):193-201 (2007)
A1.CH.03.HIV_CH_BID_V3538	JQ403028	All	Henn, M.R.	*PLoS Pathog* **8**(3):E1002529 (2012)
A1.CY.08.CY236	JF683783	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
A1.ES.05.X1608_8	FJ670519	All	Cuevas, M.T.	*ARHR* **26**(9); 1019-25 (2010)
A1.KE.11.DEMA11KE001	KF716475	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
A1.RU.11.11RU6950	JX500694	All	Baryshev, P.B.	*ARHR* **30**(6); 592-7 (2014)

Name	Accession	Proteins	Author	Reference
A1.RW.11.DEMA111RW002	KF716472	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
A1.SN.01.DDI579	AY521629	All	Meloni, S.T.	*J Virol* **78**(22):12438-12445 (2004)
A1.UG.11.DEMA110UG001	KF859745	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
A1.ZA.04.04ZASK162B1	DQ396400	All	Rousseau, C.M.	*J Virol Methods* **136**(1-2):118-125 (2006)
A2.CM.01.01CM_1445MV	GU201516	All	Carr, J.K.	*Retrovirology* **2010** Apr 28;7:39 doi: 101186/1742-4690-7-39
A2.CY.94.94CY017_41	AF286237	All	Gao, F.	*ARHR* **17**(8):675-688 (2001)
B.BR.10.10BR_MG029	KJ849799	All	Pessoa, R.	*Transfusion* **55**(5); 980-90 (2015)
B.CA.07.502_1191_03	JF320424	All	Rolland, M.	*Nat Med* **17**(3); 366-71 (2011)
B.CH.08.M2_0803101_NFLG8	KC797225	All	Castro, E.	*AIDS* **28**(12); 1840-4 (2014)
B.CN.10.DEMB10CN002	JX140658	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.ES.10.DEMB10ES002	KC473842	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.FR.11.DEMB11FR001	KF716496	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.GB.05.MM45d213_GN1	HM586212	All	Turnbull, E.L.	*J Immunol* **182**(11):7131-7145 (2009)
B.HK.06.HK003	FJ460500	All	Tsui, S.K.W.	*ARHR* **26**(1); 117-22 (2010)
B.HT.05.05HT_129389	EU839602	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Nadai, Y.	*PLoS ONE* **4**(3):E4814 (2009)
B.JP.12.DEMB12JP001	KF716498	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.KR.07.HP_18_07JHS10_3909	KJ140263	All	Kim, B.-R.	*Haemophilia* **21**(1); e1-11 (2015)
B.PE.07.502_0525_wg5	JF320191	All	Rolland, M.	*Nat Med* **17**(3); 366-71 (2011)
B.RU.11.11RU21n	JX500708	All	Baryshev	Unpublished
B.TH.08.MERLBDTRC10	JN860769	All	Rutvisuttinunt, W.	*ARHR* **28**(12); 1703-11 (2012)
B.US.11.ES38	JN397362	All	Buckheit, R.W.3.	*Nat Commun* **3**, 716 (2012)
C.AR.01.ARG4006	AY563170	All	Carrion, G.	*ARHR* **20**(9):1022-1025 (2004)
C.BR.07.DEMC07BR003	JX140663	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
C.BW.00.00BW5031_1	AF443115	All	Novitsky, V.	*J Virol* **76**(11):5435-5451 (2002)
C.CN.10.YNFL19	KC870038	All	Wei	Unpublished

Name	Accession	Proteins	Author	Reference
C.CY.09.CY260	JF683803	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
C.ES.08.X2363_2	EU786681	All	Fernandez-Garcia, A.	*ARHR* **25**(1); 93-102 (2009)
C.ET.02.02ET_288	AY713417	All	Brown, B.K.	*J Virol* **79**(10):6089-6101 (2005)
C.IN.09.T125_2139	KC156210	All	Parrish, N.F.	*PNAS USA* **110**(17); 6626-33 (2013)
C.KE.00.KER2010	AF457054	All	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
C.MW.09.703010256_CH256.w96	KC156214	All	Parrish, N.F.	*PNAS USA* **110**(17); 6626-33 (2013)
C.TZ.08.707010457_CH457.w8	KC156220	All	Parrish, N.F.	*PNAS USA* **110**(17); 6626-33 (2013)
C.YE.02.02YE511	AY795906	All	Saad, M.D.	*ARHR* **21**(7):644-648 (2005)
C.ZA.08.705010534_CH534.w12	KC156221	All	Parrish, N.F.	*PNAS USA* **110**(17); 6626-33 (2013)
C.ZA.10.DEMC10ZA001	JX140669	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
C.ZM.11.DEMC11ZM006	KF716467	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.CM.10.DEMD10CM009	JX140670	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.CY.06.CY163	FJ388945	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kousiappa, I.	*ARHR* **25**(8); 727-40 (2009)
D.KE.11.DEMD11KE003	KF716476	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.KR.04.04KBH8	DQ054367	All	Cho, Y.-K.	*ARHR* **29**(4); 738-43 (2013)
D.SN.90.SE365	AB485648	All	Takekawa	Unpublished
D.TZ.01.A280	AY253311	All	Arroyo, M.A.	*ARHR* **20**(8):895-901 (2004)
D.UG.10.DEMD10UG004	KF716479	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.UG.11.DEMD11UG003	KF716480	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
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)

Name	Accession	Proteins	Author	Reference
F1.AO.06.AO_06_ANG32	FJ900266	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Guimaraes, M.L.	*Retrovirology* **6**, 39 (2009)
F1.AR.02.ARE933	DQ189088	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Aulicino, P.C.	*ARHR* **21**(2):158-164 (2005)
F1.BR.10.10BR_RJ015	KJ849791	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Pessoa, R.	*Transfusion* **55**(5); 980-90 (2015)
F1.CY.08.CY222	JF683771	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
F1.ES.02.ES_X845_4	FJ670516	All	Fernandez-Garcia, A.	*ARHR* **25**(11):1187-1191 (2009)
F1.RO.96.BCI_R07	AB485658	All	Takekawa	Unpublished
F1.RU.08.D88_845	GQ290462	All	Fernandez-Garcia, A.	*ARHR* **25**(11):1187-1191 (2009)
F2.CM.02.02CM_0016BBY	AY371158	All	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
F2.CM.10.DEMF210CM001	JX140672	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
F2.CM.10.DEMF210CM007	JX140673	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
G.BE.96.DRCBL	AF084936	All	Oelrichs, R.B.	*ARHR* **15**(6):585-589 (1999)
G.CM.10.DEMG10CM008	JX140676	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
G.CN.08.GX_2084_08	JN106043	All	Liu, W.	*Zhonghua Liu Xing Bing Xue Za Zhi* **34**(1); 53-6 (2013)
G.CU.99.Cu74	AY586547	All	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
G.ES.09.X2634_2	GU362882	All	Cuevas, M.T.	*ARHR* **26**(9); 1019-25 (2010)
G.GH.03.03GH175G	AB287004	All	Takekawa	Unpublished
G.KE.09.DEMG09KE001	KF716477	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
G.NG.09.09NG_SC62	JN248593	All	Charurat, M.	*J Infect Dis* **205**(8); 1239-47 (2012)
G.ZA.01.TV546	KJ948662	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Wilkinson, E.	*ARHR* **31**(4); 412-20 (2015)
H.BE.93.VI991	AF190127	All	Janssens, W.	*AIDS* **14**(11):1533-1543 (2000)
H.CF.90.056	AF005496	All	Gao, F.	*J Virol* **72**(7):5680-5698 (1998)
H.GB.00.00GBAC4001	FJ711703	All	Holzmayer, V.	*ARHR* **25**(7):721-726 (2009)
J.CM.04.04CMU11421	GU237072	All	Yamaguchi, J.	*ARHR* **26**(6); 693-7 (2010)

Name	Accession	Proteins	Author	Reference
J.SE.93.SE9280_7887	AF082394	All	Laukkanen, T.	*ARHR* **15**(3):293-297 (1999)
J.SE.94.SE9173_7022	AF082395	All	Laukkanen, T.	*ARHR* **15**(3):293-297 (1999)
K.CD.97.97ZR_EQTB11	AJ249235	All	Triques, K.	*ARHR* **16**(2):139-151 (2000)
K.CM.96.96CM_MP535	AJ249239	All	Triques, K.	*ARHR* **16**(2):139-151 (2000)
U.CA.01.TV749	HM215251	All	Quesnel-Vallieres, M.	*Emerging Infect Dis* **17**(2):271-274 (2011)
U.CA.99.TV721	HM215249	All	Quesnel-Vallieres, M.	*Emerging Infect Dis* **17**(2):271-274 (2011)
U.CD.83.83CD003_Z3	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.CY.05.CY090	FJ388921	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kousiappa, I.	*ARHR* **25**(8); 727-40 (2009)
U.CY.08.CY223	JF683772	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
U.ES.10.DEURF10DZ001	JX140679	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
U.GR.99.99GR303	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)
01_AE.AF.07.569M	GQ477441	All	Sanders-Buell, E.	*ARHR* **26**(5):605-608 (2010)
01_AE.CF.90.90CF11697	AF197340	All	Anderson, J.P.	*J Virol* **74**(22):10752-10765 (2000)
01_AE.CN.10.YNFL03	KC870029	All	Wei	Unpublished
01_AE.HK.04.HK001	DQ234790	All	Tsui	Unpublished
01_AE.IR.10.10IR.THR48F	AB703616	All	Jahanbakhsh, F.	*ARHR* **29**(1); 198-203 (2013)
01_AE.JP.x.JRC77AE	AB565504	All	Umeki-Sakamoto	Unpublished
01_AE.TH.09.AA090a_WG11	JX447727	All	Rolland, M.	*Nature* **490**(7420); 417-20 (2012)
01_AE.TH.90.CM240	U54771	All	Carr, J.K.	*J Virol* **70**(9):5935-5943 (1996)
01_AE.US.05.306163_FL	JX863920	All	Heipertz, R.A. Jr.	*ARHR* **29**(10):1310-1320 (2013)
01_AE.VN.98.98VNND15	FJ185235	All	Liao, H.	*Virology* **391**(1):51-56 (2009)
02_AG.CM.10.DE00210CM013	KF859739	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
02_AG.CY.09.CY256	JF683799	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
02_AG.ES.06.P1423	EU884501	All	Fernandez-Garcia, A.	*ARHR* **25**(1); 93-102 (2009)
02_AG.GW.05.CC_0048	FJ694792	All	Vinner, L.	*APMIS* **119**(8); 487-97 (2011)
02_AG.KR.12.12MHR9	KF561435	All	Cho	Unpublished
02_AG.LR.x.POC44951	AB485636	All	Takekawa	Unpublished
02_AG.NG.09.09NG_SC61	JN248592	All	Charurat, M.	*J Infect Dis* **205**(8); 1239-47 (2012)
02_AG.NG.x.IBNG	L39106	All	Howard, T.M.	*ARHR* **10**(12):1755-1757 (1994)
02_AG.SE.94.SE7812	AF107770	All	Carr, J.K.	*AIDS* **13**(14); 1819-26 (1999)
02_AG.SN.98.98SE_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)

Name	Accession	Proteins	Author	Reference
04_cpx.CY.94.94CY032_3	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)
07_BC.CN.98.98CN009	AF286230	All	Rodenburg, C.M.	*ARHR* **17**(2):161-168 (2001)
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.CM.95.95CM_1816	AF492624	All	Wilbe, K.	*ARHR* **18**(12):849-56 (2002)
12_BF.AR.99.ARMA159	AF385936	All	Carr, J.K.	*AIDS* **15**(15):F41-F47 (2001)
13_cpx.CM.96.96CM_1849	AF460972	All	Wilbe, K.	*ARHR* **18**(12):849-56 (2002)
14_BG.ES.05.X1870	FJ670522	All	Cuevas, M.T.	*ARHR* **26**(9); 1019-25 (2010)
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.AR.99.ARMA038	AY037281	All	Carr, J.K.	*AIDS* **15**(15):F41-F47 (2001)
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.99.KER2003	AF457051	All	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
22_01A1.CM.01.01CM_0001BBY	AY371159	All	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.ES.08.X2456_2	FJ670526	All	Cuevas, M.T.	*ARHR* **26**(9); 1019-25 (2010)
25_cpx.CM.02.1918LE	AY371169	All	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
26_AU.CD.02.02CD_MBTB047	FM877782	All	Vidal, N.	*ARHR* **25**(8):823-832 (2009)
27_cpx.FR.04.04CD_FR_KZS	AM851091	All	Vidal, N.	*ARHR* **24**(2):315-321 (2008)
28_BF.BR.99.BREPM12609	DQ085873	All	Sa Filho, D.J.	*ARHR* **22**(1):1-13 (2006)
29_BF.BR.01.BREPM16704	DQ085876	All	Sa Filho, D.J.	*ARHR* **22**(1):1-13 (2006)
31_BC.BR.04.04BR142	AY727527	All	Sanabani, S.	*ARHR* **22**(2):171-176 (2006)
32_06A1.EE.01.EE0369	AY535660	All	Adojaan, M.	*JAIDS* **39**(5):598-605 (2005)
33_01B.ID.07.JKT189_C	AB547463	All	SahBandar, I.N.	*ARHR* **27**(1); 97-102 (2011)
34_01B.TH.99.OUR2478P	EF165541	All	Tovanabutra, S.	*ARHR* **23**(6):829-833 (2007)
35_AD.AF.07.169H	GQ477446	All	Sanders-Buell, E.	*ARHR* **26**(5):605-608 (2010)
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.03.UY03_3389	FJ213783	All	Ruchansky, D.	*ARHR* **25**(3); 351-6 (2009)
39_BF.BR.04.04BRRJ179	EU735535	All	Guimaraes, M.L.	*AIDS* **22**(3):433-435 (2008)
40_BF.BR.05.05BRRJ055	EU735537	All	Guimaraes, M.L.	*AIDS* **22**(3):433-435 (2008)
42_BF.LU.06.luBF_18_06	EU170139	All	Struck, D.	*ARHR* **31**(5); 554-8 (2015)
43_02G.SA.03.J11223	EU697904	All	Badreddine, S.	*ARHR* **23**(5):667-674 (2007)
44_BF.CL.00.CH80	FJ358521	All	Delgado, E.	*ARHR* **26**(7); 821-6 (2010)
45_cpx.FR.04.04FR_AUK	EU448295	All	Frange, P.	*Retrovirology* **2008** Aug 1;5:69 doi: 101186/1742-4690-5-69
46_BF.BR.07.07BR_FPS625	HM026456	All	Sanabani, S.S.	*Virol J* **2010** Apr 16;7:74 doi: 101186/1743-422X-7-74
47_BF.ES.08.P1942	GQ372987	All	Fernandez-Garcia, A.	*ARHR* **26**(7); 827-32 (2010)
48_01B.MY.07.07MYKT021	GQ175883	All	Li, Y.	*JAIDS* **54**(2):129-136 (2010)
49_cpx.GM.03.N26677	HQ385479	All	de Silva, T.I.	*Retrovirology* **7**(1):82 (2010)
50_A1D.GB.10.12792	JN417240	All	Foster, G.M.	*PLoS One* **9**(1); e83337 (2014)
51_01B.SG.11.11SG_HM021	JN029801	All	Ng, O.T.	*ARHR* **28**(5); 527-30 (2012)
52_01B.MY.03.03MYKL018_1	DQ366664	All	Tee, K.K.	*JAIDS* **43**(5):523-529 (2006)
53_01B.MY.11.11FIR164	JX390610	All	Chow, W.Z.	*J Virol* **86**(20):11398-11399 (2012)
54_01B.MY.09.09MYSB023	JX390976	All	Ng, K.T.	*J Virol* **86**(20):11405-11406 (2012)

Name	Accession	Proteins	Author	Reference
55_01B.CN.10.HNCS102056	JX574661	All	Han, X.	*Genome Announc* **1**(1):E00050-12 (2013)
56_cpx.FR.10.URF5_patient_A	JN882655	All	Leoz, M.	*AIDS* **25**(11):1371-1377 (2011)
57_BC.CN.09.09YNLX19sg	KC899008	All	Han, X.	*PLoS ONE* **8**(5):E65337 (2013)
58_01B.MY.09.09MYPR37	KC522031	All	Chow, W.Z.	*PLoS ONE* **9**(1):E85250 (2014)
59_01B.CN.09.09LNA423	JX960635	All	An, M.	*J Virol* **86**(22); 12402-6 (2012)
60_BC.IT.11.BAV499	KC899079	All	Simonetti, F.R.	*Infect Genet Evol* **2014** Apr;23:176-81 doi: 101016/jmeegid201402007 Epub 2014 Mar 3
61_BC.CN.10.JL100010	KC990124	All	Li, X.	*Genome Announc* **2013** Jun 27;1(3) pii: e00326-13 doi: 101128/genomeA00326-13
62_BC.CN.10.YNFL13	KC870034	All	Wei, H.	*ARHR* **30**(4):380-383 (2014)
63_02A1.RU.10.10RU6637	JN230353	All	Baryshev, P.B.	*Arch Virol* **157**(12); 2335-41 (2012)
64_BC.CN.09.YNFL31	KC870042	All	Hsi, J.	*ARHR* **30**(4); 389-93 (2014)
65_cpx.CN.10.YNFL01	KC870027	All	Feng, Y.	*ARHR* **30**(6); 598-602 (2014)
67_01B.CN.11.2011.ANHUI.HF115	KC183779	All	Wu, J.	*PLoS ONE* **8**(1):E54322 (2013)
68_01B.CN.11.2011.ANHUI.WH73	KC183782	All	Wu, J.	*PLoS ONE* **8**(1):E54322 (2013)
70_BF1.BR.10.10BR_PE004	KJ849758	All	Pessoa, R.	*PLoS One* **9**(11); e112674 (2014)
71_BF1.BR.10.10BR_PE008	KJ849759	All	Pessoa, R.	*PLoS One* **9**(11); e112674 (2014)
72_BF1.BR.10.10BR_MG002	KJ671534	All	Pessoa, R.	*Genome Announc* **2**(3):e00386-14 (2014)
O.BE.87.ANT70	L20587	All	Vanden Haesevelde, M.	*J Virol* **68**(3):1586-1596 (1994)
O.CM.98.98CMA104	AY169802	All	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.98.98CMABB141	AY169807	All	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.98.98CMABB212	AY169804	All	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.98.98CMU5337	AY169808	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)
O.GA.11.11Gab6352	JX245015	All	Liegeois, F.	*ARHR* **29**(7); 1085-90 (2013)
O.SN.99.99SE_MP1299	AJ302646	All	Toure-Kane, C.	*ARHR* **17**(12):1211-1216 (2001)
O.US.10.LTNP	JN571034	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Buckheit, R.W.3.	*ARHR* **30**(6); 511-513 (2014)
N.CM.02.DJO0131	AY532635	All	Bodelle, P.	*ARHR* **20**(8):902-908 (2004)
N.CM.02.SJGddd	GQ324959	All	Vallari, A.	*ARHR* **26**(1):109-115 (2010)
N.CM.04.04CM_1015_04	DQ017382	All	Yamaguchi, J.	*ARHR* **22**(1):83-92 (2006)
N.CM.06.U14296	GQ324962	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Vallari, A.	*ARHR* **26**(1):109-115 (2010)
N.CM.06.U14842	GQ324958	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Vallari, A.	*ARHR* **26**(1):109-115 (2010)
N.FR.11.N1_FR_2011	JN572926	All	Delaugerre, C.	*Lancet* **378**(9806); 1894 (2011)
P.CM.06.U14788	HQ179987	All	Vallari, A.	*J Virol* **85**(3); 1403-7 (2011)
P.FR.09.RBF168	GU111555	All	Plantier, J.-C.	*Nat Med* **15**(8); 871-2 (2009)
CPZ.CD.06.BF1167	JQ866001	All	Li, Y.	*J Virol* **86**(19):10776-10791 (2012)

Name	Accession	Proteins	Author	Reference
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.06.SIVcpzTAN13	JQ768416	All	Takehisa	Unpublished
CPZ.US.85.US_Marilyn	AF103818	All	Gao, F.	*Nature* **397**(6718):436-441 (1999)
GOR.CM.07.SIVgor2139_287	FJ424866	All	Takehisa, J.	*J Virol* **83**(4):1635-1648 (2009)
GOR.CM.07.SIVgorCP2135con	FJ424863	All	Takehisa, J.	*J Virol* **83**(4):1635-1648 (2009)

```
                          membrane binding                                                                                     phosphorylation site
                          Gag start, p17 start                                                                                                                                       p17 end p24 start
B.FR.83.HXB2                MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKI...EEEQNKS.........KKKAQQ.....AAAD..TG.............HSN...QVSQNYP 133
A1.AU.03.PS1044_Day0        ------I----R--A--------------R---L-----------L--X---SA---Q--ME----A-K--X--IK--F-------------DV-----------.---IK---..........-QRT--.........----.--..............N-G...K------ 133
A1.CH.03.HIV_CH_BID_V3538   ------I----K--A------------A----L-----------I-------A---Q---E----A-K-----YK---------------GV-----------.---I----..........-QRT--......TTN..--..............S-S...S------ 133
A1.CY.08.CY236              -----------K--A--R-----------RM--L-----------L-------A---Q--ME----A-N-----K--F-------------DV-----------.---IK---..........-Q-T--.........----.--....NSSTG......N-S...---R--- 138
A1.ES.05.X1608_8            -----------K--A--------------RM--L-----------L--D----A---Q--IE-I--A-K--T-----F-------------NV-----------.---L----..........-Q........A----.--..............N-S...K------ 130
A1.KE.11.DEMA11KE001        -----------K--A--R---------A-RM--L-----------L--S----A---Q--IE----A-N--T------------------DV-----------.---I----..........---T--......T---.--..............NRS...N------ 133
A1.RU.11.11RU6950           -----------K--A--------------RI--L-----------L--S---------Q---E----T-K-----FK--------------------------.---I----..........-Q-T--.......--TG.--.................------ 129
A1.RW.11.DEMA111RW002       -----------K--A------------Q-R---L-----------L------SA---Q---E----A----T--VK-----I----------AVR----------.---I-K--..........-Q-TT-A.........----.--..............N-S...N------ 134
A1.SN.01.DDI579             .#---------K-NK------------Q-RI--L-----------L--S----T---Q--MQ----A-K--T-------------------DV-----------.---I---N..........-Q-T--........-T-.--..............N-R..------- 131
A1.UG.11.DEMA110UG001       -----------K--A--------------R---L-----------L-S-----A---Q--MA---SA-G--T-G-K--F-----------Q--V-----------.---I-K--..........-Q-T--.......-E-G.--..............T-.........---Y-- 131
A1.ZA.04.04ZASK162B1        -----------K--S---T----------RI--L-----------L-------A---Q--IE----A-K--T--IK---------------DV-----------.---I----..........-Q-T--AAAAQ---N.--..............N-S..------- 138
A2.CM.01.01CM_1445MV        ..---------K--A----------R---RM--L--------KYSI-----------K--IR--HSA-PV-T---K-----I-V--Y---K--V---------L........Y..........-Q-T--.........---A.--..............N-.....S----- 129
A2.CY.94.94CY017_41         ------I----K--A--------------R---L--------K-SI-------P------IR----A----T---K------VV--W----VDV-----------..........-Q-T-H.........----.--..............N-.....S----- 129
B.BR.10.10BR_MG029          -----------Q--------------Q-------------------------A------E----A----T--IK--------------K-D-R------E-V..........-----------.....--.-..S-..............N-S...K------ 132
B.CA.07.502_1191_03         --------------------------------------------------A-----I---------------F----V------------------E--.......S--.........---V--.........---TAD--.............N-S...T------ 135
B.CH.08.M2_0803101_NFLG8    -----------------------------R---------------------G--K---I------K------K--------------K-DV----------.-------C..........--RE--.........-----.IK.............DNS....-I----- 133
B.CN.10.DEMB10CN002         -----------Q--K-------------Q-----------------SS---------K--Q-------------K---------------DVR-----------.--------..........-----------.E---..A-..............N-S.KAS------ 135
B.ES.10.DEMB10ES002         -----------K--K---------------------------------------AG----IE----A----T----------------RG--V------E-V.......G..........--E--P.........---A.--..............N-S.....---H-F- 132
B.FR.11.DEMB11FR001         ----------------------------------------------------G------E-------V-----------I-V------K--V------E--.........--..........-----------.....--..............N-G...-A----- 133
B.GB.05.MM45d213_GN1        -----------------------------R------------------------------A------------------------------DV------E--.......G..........---T--.......-T-N...--NQAAXASG......S-S...------- 141
B.HK.06.HK003               -----------K--K--------------Q--------------------------------SA-------K------V-----G-D--------------.--------..........R--V--.........----.--..............N-S...------- 133
B.HT.05.05HT_129389         ..-----------K-------------------------------------A-----------A-----------------------K--V-------R-........C..........-----------.....--..............N-S...------- 131
B.JP.12.DEMB12JP001         -------IN-------------S---------------L---------K--M----A----------F-V--V--------------------------.......S--.........-----------.....--..............NNS...------- 133
B.KR.07.HP_18_07JHS10_3909  -----------K--Q--------------R----------------------------------------K--F----------------------.--------..........---------.-T-A.--..............N-S...------- 133
B.PE.07.502_0525_wg5        ------I----K--S--------------R---L------D-----------A------Q-----------IK--------------K--VR-------V........E-..........---------.--TN.--..............K-S...------- 133
B.RU.11.11RU21n             ------I--------------------R---------------------G--K---E---------------K--F---V-------DV-----------.-----------.........--R--.........-----.--NSSPG.........N-S...P------ 138
B.TH.08.MERLBDTRC10         -----------K--Q-----------------------------------A--------------------K--F-----------------------R-.--------..........-Q---P.........----.--..............NNS...------- 133
B.US.11.ES38                -----------D-----#----------Q-------------------------K------E----A-------K--F----------G-DV-----------.--------..........-----P........-T--.--..............N-S...------- 132
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--......-E-A...................DKG..K------ 129
C.BR.07.DEMC07BR003         ------I-R-EK--T--R---------T-MM--L-----------LD------R--K--IQ----A----TK--V--H----V------EK--VR----------...K------.............R--.......-E-A...................DKR..K------ 128
C.BW.00.00BW5031_1          ------I-R--K-----------------R---L-----------L--S----AD--K--IK----A-K--T------------------RN--VQ--------V.......NC..........QQ-T--.......-E-A...................GG...KI----- 130
C.CN.10.YNFL19              ------I-R-RK--A-------------R-M--------------L----------K--MK----A----T-----F----------EG--VR--------V.......I..........QQ-T-E......-KEA...................DE...K--R--- 130
C.CY.09.CY260               ------I-R-EK--K------------T-M---L-----------L----------K--MS------K--T-----F----------AG--VR------------.--------..........QQ-T--KTQQ.VE-A..G-......................K------ 134
C.ES.08.X2363_2             ------I-R--K--A------------H-M---L-----------L----------K--IK----A-K--T--K---------------EK--VR------------........A..........QQ--------.KTQQ........................DDK...KI----- 131
C.ET.02.02ET_288            ..------R-EK--K------------H-M---L-----------L--------T--K--IK----A-K--T-----F----------EG-D-R-------R-........V..........QQ-T--......-G....E..............ADKG..K------ 129
C.IN.09.T125_2139           ------I-R-EK--K------------H-M---L-------DK--L---------------MK----A----T-----F----------EG--VR------------........C..........QQ-T--......-K....E..............ADG...K------ 130
C.KE.00.KER2010             ..----I-R--K--A------------H-R---L-----------L----------K--IK----A----T------------------NK--VR------------.--------..........QQ-T--......GK....A..............ADG...K------ 128
C.MW.09.703010256_CH256.w96 ------I-R--K--K---------------Q---L-------K--L--D-------K--IR--H-A----T---K--------------KN-DVR-------R-........C..........QQ-T--.......--....A..............ADG.....---R-F- 130
C.TZ.08.707010457_CH457.w8  ---S--I-R-EK--K------------S-M---------------L-------E---K--MK----A----T-----------V-----K---VR-------R-.--------..........QQ-T--......-E....A..............AD....K------ 129
C.YE.02.02YE511             ..----I-R--K--T--------------M---L---------------------K--IR----A----TA--K--------------KD-AV-------------.--------..........QQ-T--......-T....A..............EDK...K------ 128
C.ZA.08.705010534_CH534.w12 ------I-R--K---------------H-MI--L-----------L--------A----MK----A----T-----F----------EE-DVR-------R-.--------..........QQA-..............A..............ADKG..KI----- 127
C.ZA.10.DEMC10ZA001         ------I-R--K--K------------C-Q---L-----------L--------K----IQ----A----T---K--F----------EG-DVR------------.--------..........QQ-T--......-K....A..............TEGK..A------ 131
C.ZM.11.DEMC11ZM006         ------I-R-EK--K--R-K-------H-M---L-----------L--------D----IN----A----T------------------A--DVR-------------........-R-.........-Q-T--......-E....A..............ADKG..K------ 131
D.CM.10.DEMD10CM009         -----------K--Q--S-------N-R------------------L-------AD--K---S--E-AIK------I--H----------K--DV-------E--.......K-I..........--RE.A.........----.--..............N-S...------- 132
D.CY.06.CY163               ---------R--K-----T-------H---Q---------------I-------D--K---------A---S---F-----I------QAG----------E--.......-N-..........-----------.....--..............N-S...------- 133
D.KE.11.DEMD11KE003         -----------Q--T--A-------S--------------------L----------K--M-----AV------K--F-----W---EG-KV------------.......-T--.........-----------.....-R..............N-S...------- 133
D.KR.04.04KBH8              -----------K--T*--------R--M-------------K--L-------K----I-----I----K-K--F----------E--NV------------.--------..........-----------.-T--.-K..............NNS...------- 132
D.SN.90.SE365               -----------Q-------------------------------L--------------I----AI-------------I------EK--V-----VE-M.....R--..........---V--......-T--.L-..............NNS...------- 133
D.TZ.01.A280                ..---------Q--A--------------Q---------------L-----------K-#...---AI-------K--F----------RK--V-------E-L......T--..........-----------.-T--.--..............S-S...------- 127
D.UG.10.DEMD10UG004         ---------R--K--E-------S--R---R---L-----------I-----------K-VM-L----IK------K--------I----AG-KVT-----------.......-T--.........-E-V--......-TT-..-R..............N-S...------- 133
D.UG.11.DEMD11UG003         -----------K--D--R-Q-----N-R------------------I--C-----------I-----AI------IK-------------E--KV------------.......-T--.........-----------.....-K..............N-S...------- 132
D.YE.02.02YE516             ..---------K--E----------S---R---------------L-------D--K---------A-K---------F----------E--GV--------------........I..........----P-.........----.--..............N-S...------- 131
D.ZA.90.R1                  -----------R--A--R--------Q-R-----------------L----------K--MA-----I-------------------------E--DV--Q----------.--------..........-----------.....--..............N-S...------- 133
F1.AO.06.AO_06_ANG32        -----------K--D-------------Q-----L-----------L------------K-I----------------K------VV--Y----EK-S--------L...K-----..........QQ-T--......VT-V..-A...................A....A----- 130
F1.AR.02.ARE933             -----------K--T--------------R---L-----------ID------A----K-I--H--I-------K-------V--S------V-------E-L...-D---..........QQ-R--.........----.--.................------ 129
F1.BR.10.10BR_RJ015         ------I----K--Q--------------RM--LI-----------D------P---KK-IA-----------------F--I-V--F---KV-V-------E-L.---------..........-Q-T--......-T-A.--..............AAA...G------ 132
F1.CY.08.CY222              -----------K--A-------------Q--M--L-----------LD-----------KK-I---------------------IVV--Y---K--V-------E-L...-A--T--..........QQ-T--.........----.K-..................------ 129
F1.ES.02.ES_X845_4          -----------K--D-------------Q-RM--L-----------ID-----------K--I-----A-K------K--F-L--V--Y---KEK-----D----L.---------..........QQ-T--.........---AA.K-..................------ 130
F1.RO.96.BCI_R07            -----------K--A--------------RM--L--------------------P---K-I-----------------------I-V------V-V-------E-L......R--..........QQ-T--.........----.KV..................------ 129
F1.RU.08.D88_845            -----------K--D----------R---R-------------------D----------KK-IA------P-----I-------V--Y----PV-------E-L.--------..........QR-P--.........---A.--..............AAAADQG------ 135
F2.CM.02.02CM_0016BBY       ..---------K--D--------------R--------K------L-------T---K--I----S---------IK-------V--Y---K-Q-R---------L...Q---D-Y..........QQ-T-P.........----.K-..................------ 127
F2.CM.10.DEMF210CM001       ------I----K--S--------------R---L-----------L-------T---KK-IA-------------K----A-VV--Y-----------------L...Q------..........QQ-T--.........----.K-..................------ 129
F2.CM.10.DEMF210CM007       -----------K--A-------------Q-R---L-----------L-------TD--K--IL-------------K--F-A-VV--Y------V---------L...Q---D--..........Q--T--.........----.K-..................------ 129
G.BE.96.DRCBL               -----------K--A-------------R-RM--L-------D---L-------A---QK-MA----A----T--IK--F-----------K--VR------EEV...-KI-K--..........QQ-E...............N..............S-S...------- 126
G.CM.10.DEMG10CM008         -----------K--A--------------QM--------------L-------T---Q--MT----------T---K--------------R---V--------E-...-KI-Q--..........QQ-T--........--GG..E-..............N--...------- 133
G.CN.08.GX_2084_08          -----------K--A------------------L-----------L--D----AD--WK-MD----A----T--IK--F-A----------K--V--Q--IKEV...-KIXQ--..........QQX--K.........--.....A..............N-S...------- 130
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...------- 133
G.ES.09.X2634_2             -----------K--K----------R-Q-----------------L-------A---Q--MV----A----T--IK--------------G--V-------EV...-KI-K-G..........Q--.-........--M-..E-..............DNS....-T----- 132
G.GH.03.03GH175G            -----------K--A--------------RM--L------M----L-S-----A---Q--MM----A----T------F---------------V-------E-V...-KL-KN-..........QQ-.-........--KA..E-..............NN-...------- 132
G.KE.09.DEMG09KE001         ------I-R--K--A--R-----------R---L-----------L--D----E---Q--MR-----I---T--VK---------W---RK--VR-----EEV...-KA-K--..........QQNV--........--R-..G-..............NNS...------- 133
G.NG.09.09NG_SC62           .#---------K--A------------R-RM--L-----------L-------SA---Q--MM-----I---T--IK--F-----------K--V------EEV...-KI-K--..........QQEK-K.......--...................NS...------- 126
G.ZA.01.TV546               ..---------K--A-------------Q-R---L------M----L-------X---Q---V-----X--AT---------------K---V-------EEV...-KI-KN-..........QQ-T--........--MX..-X..............X-S...------- 131
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--....AP-----.KE..............KDS...KI----- 135
H.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------ 133
H.GB.00.00GBAC4001          -----------K--A--------------R---L------D---L--D----AD--LK----I--A----T--IK--F-L--V------K--VQ--S---N-V...K-I---N..........QQTT--........-T-X.KE..............KDS...KI----- 133
J.CM.04.04CMU11421          -----------K--T------------R-R---L------D---L--X-----K--Q---V-----Y---T--IK--------------EX--V----X--------.....-L-K-N..........-QQ--K......-EX--.KR..............XXS...------- 133
J.SE.93.SE9280_7887         ------I----K--D--------------RI--L------D---L------SAK--Q---V----A----TQ-IK--------------------M---E--.....-I---N..........-QQ--K......-ET-..KK..............DNS...------- 133
J.SE.94.SE9173_7022         ------I----K--D-------------Q-RI--L------D---L------SAK--Q---V----A----T--IK--------------------------------.....-I---N..........-QQT-K......-ET-..KK..............DNS...------- 133
K.CD.97.97ZR_EQTB11         -----------K--K----Q---------R---L-----------L--N----V------IR---------------F-------W---S-QVR--------L........-RT..........QQ-T--......GK--..K-..................------ 129
K.CM.96.96CM_MP535          -----------K--A------------------L-----------L-------T------IT-I---I------IK-----I-V--F---K--V--------L.----------..........QR-T--......E---..K-..................------ 129
U.CA.01.TV749               -----------K--------------R---L-----------L-S-----A----K-I-----A----T---K---------------G--V-------R-.........K-C..............K-.......--DA..--..............SGR...E----F- 129
U.CA.99.TV721               -----------Q--------------Q-R---L-----------L--SF---A----KLI-----A----T---K--F---V---------V-------R-.........--C..............K-.......-TGA..--..............NGR...E------ 129
U.CD.83.83CD003_Z3          -----------K--A----------R---R---L------K--I-------A---Q--I-----A-------K-I--L--------------V------------.....V---G..........-Q---------.........----.K-..............SNS...------- 133
U.CD.90.90CD121E12          ...K-------K--A----------R---R---L-----------L-------A---Q--IV----AF------K-----------------------E-----.....A---G..........-Q---------.........----.K-..............DNS...K------ 130
U.CY.05.CY090               -----------K--K--------------R---L-----------L-------A-------I----SAIN-----K--------------G--VR--------L..----------..........-R--------.......ET--.K..........................E----F- 129
U.CY.08.CY223               ------I----K--A--------------R---L-----------L-------A---Q---E----ST-R------K--F--I-------------V-----------.....LK---..........-Q-T--.........---AD.--..............S-S...K------ 134
U.ES.10.DEURF10DZ001        -----------K--D--------------R---L-----------L-------A---K--IE--EST-K------K--F------F------KVG-----------.....MKQ--..........-Q-T--.........---A.--..............GC.........---F- 130
U.GR.99.99GR303             -----------K--A-------------Q-R---L-----------L-------------Q--ME----A-K------K-----------------DV-----------.....L-KR-..........-Q-T--.........----.--..............GTS...K------ 133
U.NL.95.U_NL_95_H10986_D1   ---------R--Q--T------------------L-------------------A---Q--IM----A----T-------------------CK------------------.....I---N..........---T.........E--.--..............KTS...------- 130
01_AE.AF.07.569M            -----------K--A---------------M--L-----------L-------A---Q--IE---ST-K------K------V--W-------V-----------------.......K--..........RQ-T--.........---G.--..............N-S....-A----- 133
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.--..............SNS...K------ 133
01_AE.CN.10.YNFL03          ------I----K--A------------------LI-----------L-------A---Q-LIE---ST-K------K------V--W------QVN---------L.....V-KR-..........QQ-T--........--TG.--..............S-S...K------ 133
01_AE.HK.04.HK001           ------I-T-EK--A--------------MI--L-----------L-------A---Q--IE---ST-K------K--F------W-------V---------L.....I---N..........Q--T--.........---G.--..............SNS...K------ 133
01_AE.IR.10.10IR.THR48F     -----------K--A--------------R---L-----------L-------A---Q-LIE---ST-K------K--F------W-------V---------L.....A-T--..........QQ-T--......T--G.--..............S-S...K------ 133
01_AE.JP.x.JRC77AE          -----------R--A--------------RI--L-----------L-------AI--Q--IE---ST-K------K--F-A-VV-W-------V---------L.....V-K-E..........QQ-T--.........---G.--..............N-S...K------ 133
01_AE.TH.09.AA090a_WG11     ---------T--K--A--------------RM--L-----------L-------A---Q-L-E---ST-K---------F-AI-V-----K------------------.....A-K--..........QQ-T--.........---S.--..............SG.....------ 131
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-S...K------ 133
01_AE.US.05.306163_FL       -----------K--A--------------RM--------------L-------A---Q-LIE---ST-K---------F-AI---W---K---V----------L.....-....E-..........QQ-T--.........----.--..............S-S...K------ 129
01_AE.VN.98.98VNND15        -----------K--A--------------RM--L-----------L-------A---Q-LIE---ST-K------K--F--I------R---V-----------------.....V----..........QQ-T--.........---G..P-..............S-S...K------ 133
```

membrane binding

Gag start, p17 start

phosphorylation site

p17 end p24 start

```
B.FR.83.HXB2                   MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKI...EEEQNKS..........KKKAQQ.....AAAD..TG.............HSN...QVSQNYP 133
02_AG.CM.10.DE00210CM013       ------I----K--Q--------------R---L-----------L-------A---Q--ME---SA-G--T---K--F--I---W------D----------V....-V---RQKTQ.....QQ-T--.....---A..--.............S-......---F- 135
02_AG.CY.09.CY256              -----------K--A------------Q-R---L-----------L-------A---Q-LIE---ST-K------K--F--I---W---S--T-----------....-V----..........-Q-T--.....---A..--.............S-....------- 130
02_AG.ES.06.P1423              -----------K--A--------------R---LI----------L------SA---Q--ME---ST-R------K-------V------F--R----------....-IK---..........-Q-T--.....---A..ST.............GC....------- 130
02_AG.GW.05.CC_0048            -----------K--A--R-----------R---L-----------L------SA---Q-LME---ST-R------K-----I---W------D-R--------V....-V---KQKAQ.....QQ-T--.....---A..A-.............S-....------- 135
02_AG.KR.12.12MHR9             -----------K--A--------------R---L-----------L-------A---Q-LIE-I-SA-R-----FK-----IV--W------D----------L....-V----..........-Q-T--.....---A..--.............I-....------- 130
02_AG.LR.x.POC44951            -----------K--A--R-----------R---L-----------I-------A---Q--ME---ST-S------K--F--I---W---R-LD-T---------....-M----..........-Q-T--.....---A..--.............S-....------- 130
02_AG.NG.09.09NG_SC61          .#---------K--A------------Q-R---L-----------L-------A---Q-LME---ST-K------K--F--L---W------D-R--------V....-M-K--..........-Q-T--.....---A..A-.............S-....------- 128
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-....------- 130
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--.....---A..A-.............S-....------- 130
02_AG.SN.98.98SE_MP1211        #----------K--A--R-----------R---L-----------L-------A---Q--IE---ST-K------K--F--------------V---------L....-I-E--..........-Q-T--.....---A..--.............S-......---F- 129
03_AB.RU.97.KAL153_2           -----------K--A-----------E--RI--L-----------L--S--------Q---E---T-K-------K----------------------------....-I----..........-Q-T--.....--TG..--.............S-S...K------ 133
04_cpx.CY.94.94CY032_3         -----------K--A--R-----------R---L-----------L-------A---Q-LME---ST-K------------IT--W------DVQ---------....-I-S--..........-Q-T--.....---A..A-.............G-S...N------ 133
05_DF.BE.x.VI1310              -----------K--A--------------R---L-----------L-----------K--IS---AI----T---K-----------------E---V-------E--.........K.........R-----.....-E-G..A-.............N-S...-A----- 134
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-S...NL----- 132
07_BC.CN.98.98CN009            ------I-R--K--K------------H-M---L-----------L--------------MK----A----T------F----------TE-DVR---------.........I..........QQ-T--......-K....E.............ADG...K------ 130
08_BC.CN.97.97CNGX_6F          .#----I-R--K--K------------H-M---L-----------L-----------K--IK----A----T------F----------EE--VR---------.........I..........QQ-T--......-K....K.............AEE...K------ 128
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..--.............NNS...K------ 133
10_CD.TZ.96.96TZ_BF061         -----------K--E------------Q-----L-------D---I-----------K--I----AI-------IK-------------E--KVA---------....-----T..........-------.....-T--..--.............N-S...------- 133
11_cpx.CM.95.95CM_1816         -----------K--K---------R-Q-R---L-----------L--S----A---Q--MV-I--A-N--T------------------Q--VR---------....-I----..........-Q-T--.....----..--.............N-S...K------ 133
12_BF.AR.99.ARMA159            -----------------------------R----------------------------K-I-----------------F----V------K--V-------E-L....-------..........QQ-T--.....----..K-...................------- 129
13_cpx.CM.96.96CM_1849         -----------K--K------------R-RM--L-----------L-------T---Q---R-------------K-----------------V-------EEV...-KV--S-..........QQRT--.....--KN..E-.............N--...------- 133
14_BG.ES.05.X1870              -----------K--S--------------R---L-----------L--D----A---Q--M-----A----T--I---F----------------------EEV...-KA-K--..........Q--.--.....--K-..E-.............NNS...-A----- 132
15_01B.TH.99.99TH_MU2079       -----------K--A------------------L-----------I--S---SA---Q---E---ST-K------K---------W--QN---V----------....-V--N..........QQ-T--.....---G..--.............S-S...T--H--- 133
16_A2D.KR.97.97KR004           -----------K--A--------------R---L-----------I-----------K--I---E--IK-----IK--F-A-V----------VR--------L.........H..........-Q-T-P.....----..--.............S-....G------ 131
17_BF.AR.99.ARMA038            ..-----------------------R---R----------------------AD--K-I----------------------I-V--Y---KVDV-------E-L....-------..........RQ-T--.AAADKTQQ..AA.............ADK...G------ 135
18_cpx.CU.99.CU76              -----------K--K--------------R---L-----------L-------A---Q--IE---S-IK-----IK--------------G-DV---------....-I----..........-Q-T--.....---A..A-.............S-S...------- 133
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--..-R.............N-S...H------ 133
20_BG.CU.99.Cu103              -----------K--E--------------RM--L-----------L-------V---Q--IN----A-K--T---K--F-------------GV--------EV...-KL-K--..........QQ-M--.....--TG..E-.............N-S...------- 133
21_A2D.KE.99.KER2003           ..---------K--A------------Q-R---L-----------I-----------K--I------I-------K--F----------A--DV---------....------..........-N---P.....---N..--.............N-S...------- 131
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-S...N------ 131
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--.....--TG..E-.............N-G...------- 133
24_BG.ES.08.X2456_2            -----------K--A--------------RM--L-------D---L------AG---Q---N----A----T--IK-------------K--VR--------EV...-KL-K--..........QQ-V--.....--TG..E-.............S-S...------- 133
25_cpx.CM.02.1918LE            ..----I----R--A---------R-R-R---L-----------L--S---NG---Q-------SA-K------K-------V--------DV----------....-I--N..........-Q-K--.....ET--..-S.............N-S...------- 132
26_AU.CD.02.02CD_MBTB047       ------I----Q-----------------R---L-------K-SI-------A---Q---E---AVK-------K--F---V--------NV---------V....-M--..........------.....--DN..--.............S-S...K------ 130
27_cpx.FR.04.04CD_FR_KZS       -----------K--A----------R---RM--L-----------I------AAD--LD--R----A--------K--F-L-----------DV-------EEV...-KV-KNR..........QQ-T--.....E.....E-.............S-S...-A----- 130
28_BF.BR.99.BREPMI2609         -----------------------------R------------------------------I----------------------V------K-------------....-------..........------.....----..A-.............N-S...------- 133
29_BF.BR.01.BREPM16704         -------------------------N-----------------------------------S-------------K--F----V--------D--------E--....-------..........-T----.....---G.--NSRQQAAAGTG..N-S...------- 144
31_BC.BR.04.04BR142            ------I-R-EK--T--R---------Q-MM--L-----------LD-------G--K--IQ----A-K--TK--V-------------EK--VR---------...K..........QQ-T--.....-E....A.............AGKG..N------ 132
32_06A1.EE.01.EE0369           -----------K--E----------R-------------------L-------A---Q--VE-V--A-R------K--F-------------KVA---------....-I----..........-Q------.....--TA..--.............N-....------- 130
33_01B.ID.07.JKT189_C          -----------K--A--------------R---L-----------I-------A---Q--ID---ST-K------Q--H--IV--W-------V---------M....-T--R..........QQ-T--.....---G..--.............SNS...T----F- 132
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..--.............SGX...K------ 131
35_AD.AF.07.169H               ------I-R--K--T--------------R---L-----------L--S----T---Q--ME----A-K---------F-----------Q-A-----------....-I----..........-Q--K-.....T---..--.............SG....K------ 132
36_cpx.CM.00.00CMNYU830        ..---------K--A--------------RI--L------------D-----SA---Q--IE----T-K------K----------------DV-----------...-KL-SN-..........-Q-T--.....------..--.............S-S...T--R--- 131
37_cpx.CM.00.00CMNYU926        ..------T--K--A------------Q-RM--L-----------L-------A---Q--IE---ST-K------K--F------W------DVR-------L....-V----..........-Q-T--.....---G..--.............S-C...KI----- 129
38_BF1.UY.03.UY03_3389         -----------K--A------------------------------------S----K-MAH-------------K-------V--F----VDV-------E-L....-------K--..........QQ-T--.....----..K-...................------- 129
39_BF.BR.04.04BRRJ179          -----------------------------R------------------------------IA-------------------I-V--Y---KV-V---------L....-------..........------.....----..--.............NNS...------- 133
40_BF.BR.05.05BRRJ055          --------------K--------------Q-----------------S----A----------H----------K-----I--------K--V-----------....-------..........--R---....A----..A-.............DNS...K------ 134
42_BF.LU.06.luBF_18_06         #----------------------------------------------------------M-----A--------K--------------K-DVR---------....-------..........------.....----..--.............N--...------- 132
43_02G.SA.03.J11223            -----------K--A--------------RM--L-----------I------SA---Q--ME-X-ST-R------K--F--I---W------D----------V....-M----..........-Q-T--.....---A..--.............S-....------- 130
44_BF.CL.00.CH80               ------I----K--A--------------RM--LI-----------D------P---K-I---------------------VV--F---V-V-------E-L....-------..........QQ----.KTQQ---AADK-...................------- 135
45_cpx.FR.04.04FR_AUK          -----------K--M------------------L-------D---L-------A---Q--IE----N-K-----IK------V---------DV-----------....-I----..........-QNT--.....---A..--.............S-S...S------ 133
46_BF.BR.07.07BR_FPS625        -----------K--A--------------RM--LI----------LDS-----P----K-IA---S------------F--VV---F---KV-V-------E-L....-------..........QQ-T--.....----..K-...................------- 129
47_BF.ES.08.P1942              ---------------------------T-----------------L--------------------A-----------F-A---------K-DV-----------....-------..........---V--.....----..--.............NN-...------- 132
48_01B.MY.07.07MYKT021         ------I----K--A--------------R---L-----------L-------A---Q-LIE---ST-K----------F--IV--L-------V----------....-M-K--..........QQ-T--.....-E-G..--.............S-S...K------ 133
49_cpx.GM.03.N26677            ------I----K--K------------Q-R---L-------------------A---Q--IE----A-K--------------------------V----------....-I-K--..........-QGT--.....---A..--.............NIS...SI----- 133
50_A1D.GB.10.12792             ------I----K--T----------R---R---L-----------L--D----A---L--MD-I--A-R--T------F--------------V----------....-V--R..........-Q-TP-.....----..AA.............SCS...-A----- 133
51_01B.SG.11.11SG_HM021        .#---------K--K--R-----------R------------------------------------------X-----F----V------M-DV----------....-------..........--Q---.....---N..--.............N-S...------- 131
52_01B.MY.03.03MYKL018_1       ------X----K--A--------------QM--L-----------L-------A---Q--IE---ST-K------K------V--W------DV---------M....-A----..........-Y-T--.....---G..--.............NXS...K------ 133
53_01B.MY.11.11FIR164          --------T--K--A--R-----------XI---------------L-------A---Q--IE-I--X-K------K--H-L--V-W------VX---------M....-X-K-G..........QQ-T--.....----..--.............S-S...K------ 133
54_01B.MY.09.09MYSB023         ------I----Q-----R-----------R------------Y---------S----K---E----A------GF---F----V---------V----------....-A-K-G..........QQ-K--.....----..--.............S-S...------- 133
55_01B.CN.10.HNCS102056        -----------K--A--------------R---L-----------L-------T---Q-LIE---ST-K----------F--IV--W------DV--------L....-V-K--..........QQ-T--.....--TG..A-.............S-S...------- 133
56_cpx.FR.10.URF5_patient_A    -----------K--A--------------R---L-------D---LD-----SA---Q-LIE----A-K------K--F--I---L---L--N----------V....-------..........------.....---A..--.............S-S...------- 133
57_BC.CN.09.09YNLX19sg         ------I-R--K--K------------R-M---L-----------L--E----A---K-LMK----A----T-GFK--F----------AE-DVR-------R-....-K---I..........QQ-T--......-Q....E.............ADG...K------ 130
58_01B.MY.09.09MYPR37          --------T--K--A----------------M--------------L-------A---Q--IE---ST-K------K--F------W-------V--------RL....-A--G..........-Q-T--.....---G..A-.............N-S...K------ 132
59_01B.CN.09.09LNA423          ------I-T--K--A--------------RI--L------------D------K---Q--IE----A-K---------F------W------DV---------M....-A-..........................--T--.....G--G..--.............S-S...KI----- 129
60_BC.IT.11.BAV499             ------I-R-EK--T--R-K-K-----H-M---L-----------LD-------G--K--MN----A-----K--V--H-A--------EK--VR---------...K---R..........QQTTR-......-E....A.............ADKG..K------ 129
61_BC.CN.10.JL100010           ------I-R--K--K------------H-M---L-----------L-----------K--IK----A----T------F----------SE-DVR---------........-N-..........QQR---......-K....E.............ADG...K------ 130
62_BC.CN.10.YNFL13             ------I-R--K--K------------H-M---L-----------L--------D--K-LMK----A----T------F--L-------KG-DVQ---------....--......................-T--......-E....-.............ADG...KG----- 126
63_02A1.RU.10.10RU6637         -----------K--A--------------R---L-----------L-------A---Q-LME---ST-K------K--F--I---W------D-R---------....-V----..........-Q----.....---A..--.............S-....------- 130
64_BC.CN.09.YNFL31             ------I-R--K--K------------H-M---L-----------L-----------K--MK----A----T------F----------EE-DVR--------V.........I..........QQ-T--.....EK.....-.............ADG...K------ 130
65_cpx.CN.10.YNFL01            --------T--K--A--R----------RM--L-----------I-------A---Q--IE---TVK-------K--F-A-VV--------D----------L....-------..........QR-T--.....---S..--.............SGG...K------ 133
67_01B.CN.11.2011.ANHUI.HF115 -----------K--A------------Q--M--L-----------L-------A---Q---E---ST-K------K--F--I-V-W------A-----------....-V----..........QQ-T--.....---G..A-.............S-S...K------ 133
68_01B.CN.11.2011.ANHUI.WH73 -----------K--A--------------RM--L-----------L-----------Q---E---ST-K------K--F------W------------------....-V-KR-..........QQ-T--.....---G..A-.............S-S...K------ 133
70_BF1.BR.10.10BR_PE004        I-----I--------------------Q-R---V---------------------------A--------R------F-------------DV-------E--....-------..........------.....---G..--.............N-SQSS-SG---- 136
71_BF1.BR.10.10BR_PE008        -----------K--K----------------R----------------------------------A--------------------------M-DV----------....-------..........-R----.....---G..AD.............N-S...------- 133
72_BF1.BR.10.10BR_MG002        -----------------R-------N-------------------------------I--------------K--F--I-V--------DV-------E--....-------..........------.....E---..--.............KGS...P------ 133
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-L...K-VMGSRK...................S-D-AKEDT.............SAR...-AG---- 131
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-L...R--KGGRK...................S-G-AKEDT.............SAR...-TG---- 131
O.CM.98.98CMABB141             ---S----R-SK--T--Q---K--C----R---L----K------C--E----A--NE-L-Q--E-A-K----S-Q--W--I-V-W---N-FKVG--QQ-IQ-L...R-VMGNRK...................S-S-AKEDT.............SAK...-TG---- 131
O.CM.98.98CMABB212             --------T-SK--A------K--S----R---LI---K-----EC------YA--NE-L-Q--E-A-A----S-K--W-LL-V-W---S--D-Q--QQ-VQ-V...K-KMANRK...................S-D-AKKET.............SPK...-T----- 131
O.CM.98.98CMU5337              ---S----T-SK--A--H-H-K--C----R---L-----------C--E----A--NEKL-Q--E-A-K----S-Q--W--I-V-W---N-YKVG--QQ-IQ-L...K-VMGSRK...................S-G-AKEDA.............SAR...-TG---- 131
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---N-YR-E--QQ-IK-L...R-VMGSXK...................S-SDAKEDT.............SAR...-TG---- 131
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-L...K--MK-RK...................PVSSTKENT.............S-K...-TG---- 130
O.GA.11.11Gab6352              ---S----R-SK--S--Q---K--S----R---L-----------C--E----A--NEKL-Q--E-A-K---DS-Q--W--I-V-W---NGFR-N--QQ-IQ-L...K-VM.NKK...................S-S-AKEDT.............SER...-TG---- 130
O.SN.99.99SE_MP1299            --------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-L...K-VMGSRK...................S-GTAKEDT.............SAR...-TG---- 131
O.US.10.LTNP                   ---S----T-SR--A--Q---K--S----R---L-----------CS-E----A--NEKL-Q--E-A-K----S-Q--W--I-V-W---N-YK-G--QQ-IQ-L...K-VMGSRK...................P-G-AKEDA.............SAR...-TG---- 131

N.CM.02.DJ00131                --------T--K--Q--A-Y---------RM--L-----------C----MD-AS--A-L-N--E-A-A----G------AL-V-----SN-QVHN-Q------KEIK-KKERHKS........EP-NPE.....-G-A..AA.............AD-...NI-R--- 138
N.CM.02.SJGddd                 --------T--K--Q--S-Y---------R---L-------D---C----MD-AV--A-LIN--E-A-K----G-------L-V-----S--QVHN-Q-------..K-K-EQHK.......PPEP-NPE.....-G-A..AA.............VXS...NI-R--- 136
N.CM.04.04CM_1015_04           --------T--K--K--A-Y----------M--L-----------C----MD-AH--N-L-N--E-A-N----G-------I-V-----ST-PVLN-Q-------..K-KKEQHKS........EA-NPE.....-G-A..AA.............AD-...NI-R--- 135
N.CM.06.U14296                 --------T--K--Q--A-Y---------RM--L-----------C----MD-AN--A-L-N--E-A-K----G-------L-V-----S--QVQN-Q----XX...KXK-XQHKP........EP-NPE.....TE-A..AA.............ADS...NI-R--- 135
N.CM.06.U14842                 --------T--K--Q--G-Y---------RM--L-----------C----MD-AA--A-L-N--E-A-N----G------XL-V-----NKRQ-HN-Q----H-...K-VKEQHKP........-P-EPE.....-G-A..AA.............VD-...XI-R--- 135
N.FR.11.N1_FR_2011             --------T--K--Q--A-Y---------RM--L-----------C----MD-A--XV-L-N--E-A-K----G-------L-V-----S--S-Q--QA-------..K-KKEQHK.......QPEP-NPE.....-G-A..AA.............PD-...NI-R--- 136

P.CM.06.U14788                 --------T--R--A--R---------------V-----------C----M--A---EKL-E--E-A-R---DG-Q--W--LVV-W---K----S--QQ-IT-W...K--MQ-R..........--T.SEG..........................-T...GT----- 126
P.FR.09.RBF168                 --------T--R--A--R---K--S----M---------------C--E-M--A---EKL-E--E-A-R---DG-Q--W--LVV-W---R-VD-S--QQ-IT-W...K--MQRR..........--A.LEGS.....................TGT...GT----- 128

CPZ.CD.06.BF1167               --------R-EK-----SV----------LI--L---GN--Q---L----M--AD---K-IM--S--VEI--P-II--F--I-V-H-I-EGEK-Q--EA-VKIV...K-KLKITSSETTGNEETA-GKKET....-VTSSGQTEKIYPTLP......S-G...-TGR--- 154
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-L...KAAV-PE..........QGETPSA.............................-A-F- 123
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--EQL...KRHHGEQQS.........-TESNSG...SREGGASQGAS..........A-A...GI-G--- 141
CPZ.TZ.06.SIVcpzTAN13          --------R-EK--T--S-----------LI--L---GS-----M----M-NK---WK-IL-----IDI--P-IV------C--F--AGEKVR--E--VKIV...KVKLTTV..........-DD-AATSSGQGQNAGGR-TAVTPSGKSENASRVTVPPSEGGRL-- 157
CPZ.US.85.US_Marilyn           --------T--R--A---------------MM--L-----------C----M--AD--L-L-K--E-A-K----G----F--L-V-W---S-VTVE--QQ--V-L...K-VVQLQK.......QQEE-E--QQ........EAS.............G--...IG-S--- 135

GOR.CM.07.SIVgor2139_287       --------T--R--A--R---K--S----M---V-----------C--E-M--A---EKL----ESA-K---DG----W--LVV-W---R--DVS--QQ--E-F...K-VMS-RK..............SD...........................SRP...ET----- 124
GOR.CM.07.SIVgorCP2135con      --------T--R--A--R---K--S----M---V-C---------C--E-M--A---EKL----E-A-K---DG----W--LVV-W---R--DVS--QQ--E-F...K-AMS-RK..............SD...........................SRP...ET----- 124
```

```
                                                                                                                CyPA binding                                                     major homology region
B.FR.83.HXB2                  IVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTN.NPPI...PVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDR 299
A1.AU.03.PS1044_Day0          ----A----I--NL-----------I--------------------------V---I------------D-----------L--------P---I----------A---P---LQ---G.----...---D-------------------------------S------- 299
A1.CH.03.HIV_CH_BID_V3538     ----A----I--NL-----------I-------------T------------M---I------------D-----------T--------Q---L------------------M----S.----...-----------------------V------------------- 299
A1.CY.08.CY236                ----A----I--NL-----------I--------------------------V---I------------D-----------L--P-----P---I--------------P---LQ---G.----...---D-------------------------K------------- 304
A1.ES.05.X1608_8              ----A---WT--NL-----------I-------------T--------H---M---I------------D-----------L--------P--------------A---P---LQ---S.---V...---D-------------------V------------------- 296
A1.KE.11.DEMA11KE001          ----A----I--NL-----------I--------I--------D----Y---M---I------------D-----------L---Q----P------------------P--------S.----...---D-------------------V-----K------------- 299
A1.RU.11.11RU6950             ----A----I--SM-----------I-------------T------------M---I------------D-----------L--AQ---FP---------------------------S.----...---D-------------------------K------------- 295
A1.RW.11.DEMA111RW002         ----A-------SL-----------I--------------------------M---I------------D-----------L---P---VP---I-D------------S----R---S.----...---D-------------------V-----K------------- 300
A1.SN.01.DDI579               ----A-------SL-----------I-------------------T------M---I------------D-----------M--------P---------------------------S.---V...------R----------------V------------------- 297
A1.UG.11.DEMA110UG001         ----A--------------------I-----------------D--------M---I------------D-----------L--AQ----P---I-D------------I----T---G.----...---D-------------------V-----K------------- 297
A1.ZA.04.04ZASK162B1          ----AH------PL-----------I---G-N--------------------M---I------------D------M-------------P-----------------N---------S.----...---D-------------------V------------------- 304
A2.CM.01.01CM_1445MV          ----A----------------------------------------------------------------D--------------------P---------------------------S.----...-----------------------V------------------- 295
A2.CY.94.94CY017_41           ----A----------------------------------T-----------------------------D--------------------P---------------------------S.D---...-----------------------V------------------- 295
B.BR.10.10BR_MG029            ----A-----Y-S---------------------------------P-----I---A------------------------L---Q---V-----------------------VA----.--A-...------------------------------------------- 298
B.CA.07.502_1191_03           ----L--------L-----------I-------------------------------------------------------L---Q---V------D---------------------H.----...-----------------------V------------------- 301
B.CH.08.M2_0803101_NFLG8      ----L--------L-------------------------------------------------------------------L------------------------------------H.----...-----------------------V------------------- 299
B.CN.10.DEMB10CN002           ----L-------P-----------------------------A------------------------------D-------L--QQ---------------------------------.----...-----------------------------K------------- 301
B.ES.10.DEMB10ES002           ----L--------M-------------------------------------------------------------------M------------------------------------H.----...-----------------------S------------------- 298
B.FR.11.DEMB11FR001           ---------------------------------------------------------------------------------L-------------------------------------.----...-----------M-----------------K------------- 299
B.GB.05.MM45d213_GN1          -------------L-------------------------------------------------------------------L---Q--------------------------------H.----...-----------------------------K------------- 307
B.HK.06.HK003                 ----M-------------------------------------A--------------------------------------L-------V----I-------------------A---H.----...-----------M-----------------K------------- 299
B.HT.05.05HT_129389           ----M----------------------------------T-----------------------------------------------------------------------A---S.----...------------------------------------------- 297
B.JP.12.DEMB12JP001           ----L----------------------------------------------------------------------------L-------NP-----------------------A----.----...------------------------------------------- 299
B.KR.07.HP_18_07JHS10_3909    ----L----------------------------------------------------------------------------L-------------------------------------.----...------------------------------------------- 299
B.PE.07.502_0525_wg5          ----LX-------------------------------X-------------------------------------------L-------------------------------------.----...------------------------------------------- 299
B.RU.11.11RU21n               ----L--------L-------------------------T-----------------------------------------L-------V-----------------------------.----...-----------M------------------------------- 304
B.TH.08.MERLBDTRC10           ----L-------P-----------------------------------------------------------------------------------------------------A----.----...---D-------------------------K------------- 299
B.US.11.ES38                  ----L--------L-------------------------T-----------------------------------------L--AQ----------------------------A---S.---V...-----------------------------K------------- 298
C.AR.01.ARG4006               ----L-------PL-----------I-------------T-----------------------------D-----------L-------V------------------N-----A---S.----...---D-------------------V-----K------------- 295
C.BR.07.DEMC07BR003           ----L-------P-AA-----------------------T-----------------------------D-----------L-------V----V-----------------------G.----...-----------------------V-----K------------- 294
C.BW.00.00BW5031_1            ----L--------------------IG----N--I--I-T--------R--------------------D-----------L-------V------------------------A---S.---V...---D-------------------V-----K------------- 296
C.CN.10.YNFL19                ----L-------P-----------------------------------------------------I--DV----------L--------------------------------A---G.---V...---D-------------------S-----K------------- 296
C.CY.09.CY260                 ----L--------------------I----C----VR-VT-V------S---S----------------D-----------L-------V------------------------A----.----...---D-------------------V------------------- 300
C.ES.08.X2363_2               ----L--------------------I-------------T-----------------------------D-----------L--------------------------S-----A---S.---V...---D-------------------V-----K------------- 297
C.ET.02.02ET_288              ----L-------SL-----------I--------I----T-----------------------------D-----------I--------Q-----------------------A---S.----...---D-------------------V------------------- 295
C.IN.09.T125_2139             ----L-------S------------I-------------T-----------------------------D-----------L------------S-------------------A---G.---V...---D-------------------------K------------- 296
C.KE.00.KER2010               ----L--------L-----------I-------------T-----------------------------D-----------L-------NP-----------------------A---S.----...-----------------------V------------------- 294
C.MW.09.703010256_CH256.w96   ----L-------PL---------------G-N-------T----------M------------------D-----------L---Q---V----I-------------------A---S.---V...---D-------------------V-----K------------- 296
C.TZ.08.707010457_CH457.w8    ----L--------L-----------I-------------T-----------------------------D---------------P---V----L-D-----------------A---A.--A-...---D-------------------V-----K------------- 295
C.YE.02.02YE511               ------------SL-----------I-------------T-----------------------------D-----------L-------V------------------S-----A---G.---V...---D-------M-----------V-----K------------- 294
C.ZA.08.705010534_CH534.w12   ----L-------P--------------------------T---------------------------------D-------L-------V----------------S-------A---S.----...---D-------------------V-----K------------- 293
C.ZA.10.DEMC10ZA001           ----L-------SL-----------I-------------T-----------------------------D-------------------V------------------------T---S.---V...-----------------------V------------------- 297
C.ZM.11.DEMC11ZM006           ----L--------L-----------I-------------T-----------------------------D-----------L---Q---V----V-------------------A---S.----...---D-------M-----------V-----K------------- 297
D.CM.10.DEMD10CM009           ----L--------L-----------ID---------------------A----------------------T---------L-------V------------------------A---S.----...---D-------------------V------------------- 298
D.CY.06.CY163                 ----L------------I-S-----I----------------------H-------------------------------L--------------------------------D---S.---V...-----------------------V------------------- 299
D.KE.11.DEMD11KE003           ----L--------------------I-----N--I---------------------A------------------D-------L---Q---HP----------------------------.----...---D-------------------V------------------- 299
D.KR.04.04KBH8                ----L-------PL-----------------N-------T------------S---A-------------------S----L------------L----------------D------S.----...-----------------------VG----K------------- 298
D.SN.90.SE365                 ----L--------L-----------I-------------T-----------------------------------------L-------------------------------------.----...-----------------------V------------------- 299
D.TZ.01.A280                  ----L--------------------I-------------------------------------------------------L--------------------------------A----.---V...-----------------------V------------------- 293
D.UG.10.DEMD10UG004           ----L--------------------I-------------------------------------------------------L-------V-----------------------------.----...-----------------------V-----K------------- 299
D.UG.11.DEMD11UG003           ----L--------------------I-------------------------------------------D-----------L---Q---V------------------S---------S.----...-----------------------V-----K------------- 298
D.YE.02.02YE516               ----L--------------------I-----N------------------------------------------G----L--AQ--------I--------V------E---R---S.----...-----------------------V------------------- 297
D.ZA.90.R1                    ----L-------P------------I-------------------------------------------------------L--------------------------------A----.----...---D-------------------V------------------- 299
F1.AO.06.AO_06_ANG32          ----F--------------------I-------------------------------------------D--------------PQ--------I-------------------Q---S.---V...--------L--------------V-----K------------- 296
F1.AR.02.ARE933               V---P-----Y-SL-------------------------------------------------------D-----------L--A-----P------------------P----Q---S.---V...-----------------------V-----K------------- 295
F1.BR.10.10BR_RJ015           ----L-------SL-----------I-------------------------------------------D-----------I--PQ----P---I-------------------Q---S.---V...-----------------------V-----K------------- 298
F1.CY.08.CY222                ----L-------S------------I-------------------------------------------D-------D---L--------P-----------------------Q---S.---V...---D-------------------V-----K------------- 295
F1.ES.02.ES_X845_4            ----L--------M-----------I-------------V------------S----------------D-----------L--QQ---FP---L-------------------Q---G.---V...----------V------------V-----K------------- 296
F1.RO.96.BCI_R07              ----L--------------------I-------------------------------------------D-----------L--------P-----------------------Q---S.T--V...-------G---------------V------------------- 295
F1.RU.08.D88_845              ----H-------S------------I-------------------------------------------D--------------------P-----------------------Q---S.---V...-----------------------V-----K------------- 301
F2.CM.02.02CM_0016BBY         ----L--------------------I-------------------------------------------D-----------L---Q----P---I-------------------A---S.---V...-----------------------V------------------- 293
F2.CM.10.DEMF210CM001         ----L--------------------I-------------------------------------------D--------------------P-----------------------A---S.---V...---D-------------------V-----K------------- 295
F2.CM.10.DEMF210CM007         ----L--------------------I--------------------------V----------------D---------------Q----P---I-------------------A---S.---V...---D-------------------V-----K------------- 295
G.BE.96.DRCBL                 ----A----------------------------------T---------------------------------D-------L--QQ--------I-D-T------A--------R---S.---V...-----------------------V------------------- 292
G.CM.10.DEMG10CM008           ----A---P----L-------------------------T-----------------------------D-----------T--QQ--------I-------------N----VA---G.----...-----------------------V-----K------------- 299
G.CN.08.GX_2084_08            ----A-------SL-----------I-------------T-----------------------------D-----------L--PQ---FP---I-D-----------N-----A---S.----...-----------------------V-----K------------- 296
G.CU.99.Cu74                  ----A---P------------------------------------------------------------D-----------L--PQ----P---I------------------------.----...-----------------------V-----K------------- 299
G.ES.09.X2634_2               ----A----------------------------------------------------------------DA----------T--PQ----P---I----------------D--R---S.----...-----------------------V-----K------------- 298
G.GH.03.03GH175G              ----A-------P--------------------------------------------------------D-----------L--QQ--------I-D-T---------------R---S.----...-----------------------V------------------- 298
G.KE.09.DEMG09KE001           V---A-------------------A--------------V-----------------------------D-----------M--QQ--------I-------------------Q---S.---V...-----------------------V------------------- 299
G.NG.09.09NG_SC62             ----A-----------K--------I-------------T-----------------------------D---D-------L--QQ---V----I-------------------T----.----...-----------------------V-----K------------- 292
G.ZA.01.TV546                 ----A-------P------------I-------------------------------------------D-----------I--PQ----P---I-------------------R---S.----...---D--------------------------------------- 297
H.BE.93.VI991                 ----A-----------------------------------------------A----------------D-----------L--------P----------------------VA---G.----...---D-------------------V-----K------------- 301
H.CF.90.056                   ----A-----------------------------------------------A----------------D--------------------P-----------------------A---G.--A-...---D-------------------V-----K------------- 299
H.GB.00.00GBAC4001            ----A---------------------------------------------------S---------X--D-----------T--------P-----------------N-----A---X.---X...---X-------X-----------V-----K------------- 299
J.CM.04.04CMU11421            ----M---P----L-----------X-------------------------------------------D-----------X-------P----A-------------------A---G.----...-----------------------V------------------- 299
J.SE.93.SE9280_7887           ----L---P----L-----------I-------------------------------I-----------D------------------------V-----------------------G.----...-----------------------V------------------- 299
J.SE.94.SE9173_7022           ----L---P----L-----------I-------------------------------I-----------D-------------------V----V-------------N---------G.----...-----------------------V------------------- 299
K.CD.97.97ZR_EQTB11           ----L--------L-----------I-------------------------------------------D-----------M---Q----P---I-------------------T---S.----...-----------------------V------------------- 295
K.CM.96.96CM_MP535            ----L--------L-----------I-------------T-----------------------------D---D-------L--------P-----------------------A---S.---V...-----------------------V------------------- 295
U.CA.01.TV749                 -M--A----TY-SLT----------I---N------------------------M-----------L--D-----------L--------P----------------------VA---G.---V...---D--R----------------V------------------- 295
U.CA.99.TV721                 -M--A----TY-SLT----------I---N------------------S-----M-----------I--D-----------M-------V-----------------------V----S.----...---D--R----------------V------------------- 295
U.CD.83.83CD003_Z3            ----M--------------------I-------------------------------------------D-------D---L--------P---------------------------S.----...---D-------------------V-----K------------- 299
U.CD.90.90CD121E12            ----M-------P------------I-------------------------------------------D-------D---L--------P---------------------------S.----...---D-------------------V------------------- 296
U.CY.05.CY090                 ----V--------L-----------I---S---------V-----------------------------D--------------------P-----------------------A----.---V...-------K--V------------V-----K------------- 295
U.CY.08.CY223                 ----A-------SM---------------G----------------------M---I----------------D-------L-------V-----------------------V----S.----...---D-------------------VG----K------------- 300
U.ES.10.DEURF10DZ001          ----A---P----------------I-------------T------------M---I------------D---------------Q----P---I-------------------K----.----...----------V------------V------------------- 296
U.GR.99.99GR303               ----A----I--PM-----------I-------------T---X--------M---II-----------D-----------L--TQ---FP---------------------------S.----...----..-----------------V------------------- 297
U.NL.95.U_NL_95_H10986_D1     --R-L--------L-------------------------------------------------------D--------------------P-----------------------A---S.---V...-----------------------------K------------- 296
01_AE.AF.07.569M              ----A-------PL---------------G-N----S---------------M---I------------D-----------T--------P----------------------------.T---...-----------------------V------------------- 299
01_AE.CF.90.90CF11697         ----A-------PL---------------G-N--------------------M---I------------D-----------L--------P------------X----N---------S.----...---D-------------------V------------------- 299
01_AE.CN.10.YNFL03            -I--A-------PL-----------I---G-N--------------------M---I------------------------L--------P----------------------------.----...---D-------------------V------------------- 299
01_AE.HK.04.HK001             ----A----I--SL---------------G-N--------------A-----M---I------------------------------------P---------------------------S.--X-...---D-------------------V------------------- 299
01_AE.IR.10.10IR.THR48F       ----A-------SL---------------G-N-----------D--------M---I----------------------------Q----P---V-------------------A----.----...---D-------------------I------------------- 299
01_AE.JP.x.JRC77AE            ----A-------SL-----------I---G-N--------------------M---I----------------D---D---L--AQ---FP---------------------------S.----...-----------------------VG----K------------- 299
01_AE.TH.09.AA090a_WG11       ----A-------PL---------------G-N--------------------M---I------------------------L--------P-----------------N-----A---G.--T-...---D-------------------V------------------- 297
01_AE.TH.90.CM240             ----A----A--PL---------------G-N--------------------M---I------------------P--------------P----------------------------.----...---D-------------------V------------------- 299
01_AE.US.05.306163_FL         V---A--H----PL---------------G-N--------------------M---I------------------------I--------Q---L--------------I---------.----...-----------------------V------------------- 295
01_AE.VN.98.98VNND15          ----A-------PL---------------G-N--------------------M---I------------------------------------P-----------------N----------.----...---D-------------------V------------------- 299
```

```
                                                                                                   CyPA binding                                                 major homology region
B.FR.83.HXB2                  IVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTN.NPPI...PVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDR 299
02_AG.CM.10.DE00210CM013      ----A-------SM-----------I-------------T------------M---I------------D--------------------P---------------------------S.---V...----------V------------V------------------- 301
02_AG.CY.09.CY256             ----A--------V----------AI--------------------------M---I------------D-------D---T--------P---------------------------S.----...-----------------------V------------------- 296
02_AG.ES.06.P1423             ----A----T--NL-----------I---G----------------------M---I------------D-----------L---Q----P----------------------M----S.----...---D-------------------V-----K------------- 296
02_AG.GW.05.CC_0048           ----A----I--SM----------AI-------------T------------M---I------------D---D-------T--------P---------------------------S.----...----------V------------V------------------- 301
02_AG.KR.12.12MHR9            ----A----T--PM-----------I--------------------------M---I------------D-----------T--------P---------------------------S.--A-...-----------------------V------------------- 296
02_AG.LR.x.POC44951           ----A----T--S-----------AI-------------T------------M---I------------D---------------Q----P---------------------------S.----...----------V------------V------------------- 296
02_AG.NG.09.09NG_SC61         ----A-------PM-----------I---G---------T------------M---I------------D-----------T---Q----P---L-----------------------G.----...----------V------------V------------------- 294
02_AG.NG.x.IBNG               ----AK---T--SM-----------I---G----------------------M---I------------D--------------------P---------------------------S.----...----------V------------V------------------- 296
02_AG.SE.94.SE7812            ----A----T--S---------------------------------------M---I------------D--------------------P---------------------------S.----...----------V------------V------------------- 296
02_AG.SN.98.98SE_MP1211       ----A-------S------------I--------------------------M---I------------D-------D---T--------P---------------------------S.----...---D-------------------V------------------- 295
03_AB.RU.97.KAL153_2          ----A----T--SM-----------I--------------------------M---I------------D-----------L--AQ---FP--------------SS-----------S.----...---D-------------------V------------------- 299
04_cpx.CY.94.94CY032_3        ----A-------S------------I--------------------------M---I------------D-------D---T--------P---------------------------S.---V...-------------------T---I------------------- 299
05_DF.BE.x.VI1310             -------------L-----------I-------------------------------------------------------L---Q---V------D-----------------A----.----...-----------------------V------------------- 300
06_cpx.AU.96.BFP90            ----A--------M-----------I-D-----------T------------M---I------------D--------------------P---I-----------------------S.----...-----------------------VG----K------------- 298
07_BC.CN.98.98CN009           ----L--------L----------------------------------------------------I--D-----------L--------------------------N-----A---S.---V...---D-------------------------K------------- 296
08_BC.CN.97.97CNGX_6F         ----P-------PL-------------------------T-----------------------------D-----------L-------V-----------------------------.----...-----------------------------K------------- 294
09_cpx.GH.96.96GH2911         ----A--------------------I-------------T------------M---I------------D-------------------NP------------------V----A---G.----...----------V------------V------------------- 299
10_CD.TZ.96.96TZ_BF061        ----L-------PLT------------------------------------------------------------------L---Q---V----I-------------------R---S.----...-----------------------V------------------- 299
11_cpx.CM.95.95CM_1816        ----A----I---L-----------I--------------------------M---I------------D-----------M---Q----P---L--------------P--------G.---V...------R----------------V------------------- 299
12_BF.AR.99.ARMA159           ----L-------PL-------------------------------------------------------D-------D---L--------P-----------------------Q---S.---V...------------------------------------------- 295
13_cpx.CM.96.96CM_1849        ----A-------S--------------------------------------------------------D-----------M---QR--FP---I-------------------T----.----...-----------------------V------------------- 299
14_BG.ES.05.X1870             ----A-------P---------------------------------------L----------------DA----------M---QQ---P---I-------------------T----.----...-----------------------V------------------- 298
15_01B.TH.99.99TH_MU2079      ----A-------PV---------------G-N--------------------M---I--------------------K------------P----------------------V----S.--A-...---D-------------------V------------------- 299
16_A2D.KR.97.97KR004          ----A---TY--NL-------------------------------------------------------D-----------L--------P---------------------------S.---V...-----------------------V------------------- 297
17_BF.AR.99.ARMA038           ----L-------T--------------------------------------------------------D-----------L---Q----P---I----------A--N-----Q---S.---V...-----------------------V-----K------------- 301
18_cpx.CU.99.CU76             ----A-------S-----------AI--------------------------M---I------------D-----------I---Q--------I-D-----------------R----.----...---D-------------------V------------------- 299
19_cpx.CU.99.CU7              --HKL--------L-----------I-----N-------------------------------------------------L--------------------------------A---S.----...-----------------------V------------------- 299
20_BG.CU.99.Cu103             ----A---PI-----------------------------------------------------------D-----------T---QQ---P---I----------A--------T----.----...-----------------------V-----K------------- 299
21_A2D.KE.99.KER2003          ----A---W----------------------N----------------S---S----------------------------L---Q----------------------------A---G.----...-----------------------V------------------- 297
22_01A1.CM.01.01CM_0001BBY    ----A----I---L---------------G-N-------------T--T---M---II-----------D-----------M---Q---YP-----------------------A---S.---V...-----------------------V------------------- 297
23_BG.CU.03.CB118             ----A---P------------------------------------------------------------D-----------L---QQ-------I-------------------R----.----...-----------------------V-----K------------- 299
24_BG.ES.08.X2456_2           ----A---PI-----------------------------------------------------------D-----------M---QQ---P---I-------------------T----.----...-----------------------V-----K------------- 299
25_cpx.CM.02.1918LE           ----A---PI-----------------------------------------------------------D-----------L---Q----P---I-------------N-----R----.----...-----------------------V------------------- 298
26_AU.CD.02.02CD_MBTB047      ----A--------L-------------------------V--------------M------------I-D-----------L---Q--------I-------------------A---S.----...-----------------------V------------------- 296
27_cpx.FR.04.04CD_FR_KZS      --H-A--------L-----------I-------------------------------------------D-----------L---QQ---------------------------A----.----...-----------------------V------------------- 296
28_BF.BR.99.BREPM12609        ----L--------L-----------I-------------------------------------------D-----------L--------P-----------------N-----A---S.---V...-----------------------V------------------- 299
29_BF.BR.01.BREPM16704        ----L-------P------------I-------------------------------------------D-----------L--------P-----------------------Q---S.---V...-----------------------V------------------- 310
31_BC.BR.04.04BR142           ----L-------P-A------------------------T-----------------------------D-----------L-------V---------------S--------A---S.---V...---D-------------------------K------------- 298
32_06A1.EE.01.EE0369          ----A----T--S------------I--------------------------M---I------------D-----------LY-TQ---FQ---------------------------S.----...---D-------------------V------------------- 296
33_01B.ID.07.JKT189_C         ----A-------PL---------------G-N--------------------M---I--------------------D---T--------P---------------------------S.----...---D-------------------V------------------- 298
34_01B.TH.99.OUR2478P         ----A-------PV---------------G-N--------------------M---I--------------------D-------Q----P---------------------------S.----...---D-------------------V------------------- 297
35_AD.AF.07.169H              ----A----I--SL-----------I----------------------G---M---V------------D-----------L---Q-----------------------P----A----.----...---D--R----------------V-----K------------- 298
36_cpx.CM.00.00CMNYU830       ----A--------L---------------G-N--------------------M---I------------D-------------------HP---I---------E---------A----.----..S-----------------------V------------------- 298
37_cpx.CM.00.00CMNYU926       ----A----I--SL-----------I-------------T---D----H---A---V------------D-------D-----------HP---L------------------VA---S.----...---D-------------------V------------------- 295
38_BF1.UY.03.UY03_3389        ----L-------PL-----------I-------------------------------------------D-------D---L--------P-----------------------Q---S.---V...-----------------------V----LK------------- 295
39_BF.BR.04.04BRRJ179         ----L-------SL-----------I-------------------------------------------------------L-------------------------------VA---S.---V...-----------------------V------------------- 299
40_BF.BR.05.05BRRJ055         ----L--------------------I----------------A--------------------------D-----------L--------P-----------------------Q---S.---V...-----------------------V------------------- 300
42_BF.LU.06.luBF_18_06        ----L-------------------------------------A--------------------------D-----------L--AQ----P---I-------------------Q---S.----...-----------------------V------------------- 298
43_02G.SA.03.J11223           ----A-------SL-----------I-------------T-----------------------------D-----------L--PQ----P---I-------------------T---S.----...----L------------------V-----K------------- 296
44_BF.CL.00.CH80              ----L--------L-------------------------T-----------------------------D-------------------AP---------------------------S.---V...------------------------------------------- 301
45_cpx.FR.04.04FR_AUK         ----A----I---------------I--------------------------M---I------------D-------------------AP----------------------V----G.---V...------R----------------V------------------- 299
46_BF.BR.07.07BR_FPS625       ----L--------L-------------------------------------------------P-----GG--------Y-F-T------P-----------------------Q---S.---V...-----------------------V------------------- 295
47_BF.ES.08.P1942             ----A--H-I--P--------------------------------C-----------------------D-----------L--------P-----------------------Q---S.---V...---D-------------------V-----K------------- 298
48_01B.MY.07.07MYKT021        ----A-------PV---------------G-N--------------------M---I--------------------D---T--------P-----------------N---------G.----...---D-------------------V------------------- 299
49_cpx.GM.03.N26677           ----T-------PM-----------I--------------------------M---I------------D--------------------P-------------------------I-G.----...-----------------------V------------------- 299
50_A1D.GB.10.12792            ----L--------------------I-------------------------------------------D-----------L------#------------------------------.----...-----------------------V-----K------------- 298
51_01B.SG.11.11SG_HM021       ----L--------------------I-------------------------------------------------------L-------T------------------N---------S.----...-----------------------V------------------- 297
52_01B.MY.03.03MYKL018_1      ----A-------PV---------------G-N--------------------M---I---------------------------------X---------------------------X.----...---D-------------------V------------------- 299
53_01B.MY.11.11FIR164         -M--A-------PL-----------X---G-N--------------------M---I------------D-------D---T--------P----------------------------.----...---D-------------------V------------------- 299
54_01B.MY.09.09MYSB023        ----L-------PL-----------------N----------A------------------------------D-------L-------V----------------------------S.----...-----------------------V------------------- 299
55_01B.CN.10.HNCS102056       ----A-------PL---------------G-N--------------------M---I------------------------A--------P----------------------------.----...---D-------------------V------------------- 299
56_cpx.FR.10.URF5_patient_A   ----A-------SM-----------I-------------T------------M---I------------D-----------T--------------------------N----------.----...----------V------------V------------------- 299
57_BC.CN.09.09YNLX19sg        ----L-------P----------------------------------------------------------------D---L-------V-----------------------------.----...-----------------------S-----------S------- 296
58_01B.MY.09.09MYPR37         ----A-------SL---------------G-N--------------A-----M---I------------------------I--------P---------------------------S.----...---D-------------------V------------------- 298
59_01B.CN.09.09LNA423         ----A----T--SL---------------G-N--------------------M---I------------------------L--------P-----------------N---------S.----...---D-------------------V-----K------------- 295
60_BC.IT.11.BAV499            ------------P------------I-------------T-----------------------------D-----------L---Q---VQ---L-------------------A---A.---V...---DL------------------V-----K------------- 295
61_BC.CN.10.JL100010          ----L-------P--------------------------T-----------------------------D-------D---L--PQ---L---LI-----------S-------T---G.---V...---D-------------------S-----K------------- 296
62_BC.CN.10.YNFL13            ----L--------L-----------I-------------T---------------------------I-D------------------------I-------------------A---G.---V...---D-------------------VG----K------------- 292
63_02A1.RU.10.10RU6637        ----A----T--SM-----------I-------------T------------M---I------------D-----------T-------NP----------------------------.----...----------VM-----------V-----K------------- 296
64_BC.CN.09.YNFL31            ----L-------P--------------------------------------------------------D-----------L-------V------------------------A---G.---V...---D-------------------------K------------- 296
65_cpx.CN.10.YNFL01           ----L-------P--------------------------------------------------------D-----------L--------------------------------A---G.---V...---D-------M-----------------K------------- 299
67_01B.CN.11.2011.ANHUI.HF115 ----A--------V---------------G-N--------------------M---I--------------------D---X---Q----P---I-------------------A----.----...---D-------------------V------------------- 299
68_01B.CN.11.2011.ANHUI.WH73  ----A-------PV---------------G-N--------------------M---I--------------------D---T---Q----P---I-------------N-----A----.----...---D-------------------V------------------- 299
70_BF1.BR.10.10BR_PE004       ----M---I----------------I-K---------------D----S---A---I------------D-----------T---Q----PA--IK--------R---N-----H---S.----...-----------------------V-------R----------- 302
71_BF1.BR.10.10BR_PE008       ----H-------P--------------------------T--------S--------------------D-----------L--------P-----------------------Q---G.---V...---D-------------------VG----K------------- 299
72_BF1.BR.10.10BR_MG002       ----L-------P-T---------------------------A--------------------------D---------------Q----P---I-------------------Q---G.----...-----------------------V------------------- 299
O.BE.87.ANT70                 --S-A-------------------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------------- 298
O.CM.98.98CMA104              V---A--------L----------A------N--I----M-------SY-I-----AI----G-L-V---V----------T--IAV--LP---I---T--------------VH-T-RA-N--...---D--RK--V-----M-K----V-----K------------- 298
O.CM.98.98CMABB141            V-P-A----I--------------A------N--I----M------I-Y-I-----AI----G-L-V---V----------T--IPV--LP---I-D-T----------Q---VH-I-RA-N-L...---D--RK--V-----M-K----V------------------- 298
O.CM.98.98CMABB212            V-A-A--------X----------A------N--I----M------I-Y-I-----A-----G-L-V---V-----G----A--PAM--LP---L-D-T----------Q----A-T-RA-N--...---D--RK--V-----M-K----V-----K------------- 298
O.CM.98.98CMU5337             V-P-A--H-----L----------A------N--I----M------I-Y-I-----AI----G-L-V---V--D---D---T--PAV--LP---I---T----------Q---VH-I-RA-N--...---D--RK--V-----M-K----V-----K------------- 298
O.CM.99.99CMU4122             -IP-A-------P-----------A------N--I----M------I-Y-I-----AI----G-L-V---V-----S----T--PAM--LP-------T----------Q---VH-T-RA-N--...---D--RK--V-----M-K----V------------------- 298
O.FR.92.VAU                   V-A-A--------L----------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 297
O.GA.11.11Gab6352             -IP-A-------------------A------N--I----M------I-Y-I-----AI----G-L-V---V-----SD---T--PA---YPA--I-D-T----------Q----H-T-RA-N--...---D--RK--V-----M-K----V-----K------------- 297
O.SN.99.99SE_MP1299           --T-A-------SL----------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------------------- 298
O.US.10.LTNP                  V---A----T-VPL----------A------N--I----M------V-Y-I-----AI----G-L-V---V------D---T--PPV--LP---I---T----------Q---VH-T-RP-Q--...---D--RK--V-----M-K----V------------------- 298

N.CM.02.DJO0131               L---A----I--PLT----------I-----N-------M--------A---------------------V----------T--XPX--LP---L---------------X--VT---S.---V...-------K--V----R-------V---E-K------------- 304
N.CM.02.SJGddd                L--TA----I--PLT----------I-----N-------M--------S---------------------V----------T---PV--LP---I---------------A--VA---A.---V...---D--R---V----R-------V---E-K------------- 302
N.CM.04.04CM_1015_04          L---A------HPLT----------I-----N-------M--------S---S-----------------V----------T---PV--LP---L---------------P---T---S.----...---D--R---V----R-------V---E-K------------- 301
N.CM.06.U14296                L---A--H----PLT----------I-----N-------M--------G---------------------V----------T---PV--LP---L---------------A--VA---X.---V...---D--R---V----R--K--X-V---E-K------------- 301
N.CM.06.U14842                LI--A---L---PLT----------I--------I----M--------S--------------------DV----------T---PV---P---I--------------NVA-VA---A.---V...-X-D--R---V----R-------V-----K------------- 301
N.FR.11.N1_FR_2011            L---A----I--PLT----------I-----N-------M--------A---------------------V------D---T---PV--MP-------------------A---A---A.---V...------R---V----R-------V---E-K------------- 302

P.CM.06.U14788                ----A----T-MPL----------A------N--I----M------I-D-I-----A-----G-L-V---V----------T---PV--LP---L---T-G--------K----T---R.-N-V...---D--RK--V-----V-K--C-V-----K------------- 292
P.FR.09.RBF168                ----A----T-LPL----------A------N--I----M-----SI-D-I-----A-----G-L-V---V----------T---PV--LP---L-D-T-G--------NK---T-T-RN-N--...---D--RK--V-----V-K----I-----K------------- 295

CPZ.CD.06.BF1167              --FDAAQ-PR--P-----------T----N-K----------A---I-H-------AI-D--G---V---V------D---L--T----LV--TL---T----------IV---Q-V-T.PQNQGGV---D------------L-K----V------------------- 323
CPZ.CM.05.SIVcpzMT145         V-R-A---LT-NP-----------A------N-------M-----------------I--------I---V--D-------T--T-----------------------S-E--V----A.----...---D--R---V-------K----V-----K------------- 289
CPZ.GA.88.GAB1                L---A-----------------------------------------L---V-----A-----G---V---V----------L--T---------L---------------------T-A.----...---DV-R--V------V----C-V------------------- 307
CPZ.TZ.06.SIVcpzTAN13         --TDL--TAR--P-----------T----G------------------Y---S---AI-E----------V----------L--A----QQA-ML---T-A---------T---Q---T.P-AQGGV---D------------L----N-V-----K-----A------- 326
CPZ.US.85.US_Marilyn          VI--A--------M----------A------N-------M----------V-----AI----G---V---A------D*--T-L-----------------------------V----A.----...---D--R---VV----V-K--C-VG----K------------- 300

GOR.CM.07.SIVgor2139_287      V-A-A-------PL----------A------N--I----M-----SI-Y-I-----AI----G-L-V---V----------M--AVV--LPV--L---S-G--------Q---VA-I-RN-N-V...---D--RK--V-----V-K--C-V-----K------------- 291
GOR.CM.07.SIVgorCP2135con     V-A-A-------PL----------A------N--I----M-----SI-Y-I-----AI----G-L-V---V----------L--AVV--LPV--L---S-G--------Q---VA-I-RN-N-V...---D--RK--V-----V-K--C-V-----K------------- 291
```

```
                                        major homology region                                           p24 end p2 start        p2 end p7 start          Zn motif          Zn motif          p7 end p1 start
B.FR.83.HXB2                   FYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVT......N..SAT.I.MM......Q..RGNFR.NQRK..IVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCT....E.RQANFLGKIWPS.YK 442
A1.AU.03.PS1044_Day0           -F-A------T----S----------------S------SG----------------S------------AQ......Q...TS.-.--......-..-----GG-KR...I----------L--------------------------....-.------------.S- 441
A1.CH.03.HIV_CH_BID_V3538      -F--------T-D-------------------S--R---AG------------------------------Q......H...TN.-.--......-..----K.G---...I----------L--------------------------....-.-------R----.N- 440
A1.CY.08.CY236                 -F-V------T--------D-------------------SG-S--------------A------------AQ......H...-N.-.--......-..-----.G-K-...I----------L--------------------------....-.------------.S- 445
A1.ES.05.X1608_8               -F-C--------D--T-------I--------S--R---AG----------------S------------AQ......Q..HTN.-.--......-..K---K.G---...I----------L--------------------------....-.-------R----.S- 438
A1.KE.11.DEMA11KE001           -F--------T----G-------I--------S--R---TG------------------------------Q......H...TN.-.--......-..----K.G-KR...I----------L----------------R---------....-.------------.S- 440
A1.RU.11.11RU6950              -F--------T-D------D---------------R---SG------------------------------Q......-...-N.-.--......-..KS---.GPKR...I----------L-------------R------------....-.-------R----.S- 436
A1.RW.11.DEMA111RW002          -F-------------G---D---I--------S--R----G-S--------------S------------AQ......Q...TS.-.--......-..----K.G-KR...I----------L--------------------------....-.-------R----.N- 441
A1.SN.01.DDI579                -F-C------T-D-------------------S--R---TG------------------------------Q......H...TN.V.--......-..-----.G-KR...I----------L--------------------------....-.------------.N- 438
A1.UG.11.DEMA110UG001          -F--------T----G-------I--------S--R---AG----------------S------------AQ......Q...TN.-.--......-..----K.G-KR...I----------L--------------------------....-.-------R----.S- 438
A1.ZA.04.04ZASK162B1           -F-A---G--T--------G------------Q------TG-S---V----------S-------------Q......Q...PN.T.--......-..----K.G-KR..NI----------L--------------------------....-.------------.S- 446
A2.CM.01.01CM_1445MV           -F--------T-D------D---------------R------------------------------------IH......S.TNQN.V.--......-..-----.GPKR...I----------L--------------------------....-.-----------P.N- 438
A2.CY.94.94CY017_41            -F--------T--------D-----------RS--R----G-S--------------S-----------H-Q......S.TNTN.-.--......-..-----.G-KR...I----------L--------------------------....-.------------.N- 438
B.BR.10.10BR_MG029             ------------D-------------S-------------G-------------------------------......-..TGN.-.--......-..-----.T--R..T-----------I-K-----------------------V....-.------------.H- 441
B.CA.07.502_1191_03            ----------------------------------------G-------------------------------......S..PGN.-.--......-..-----.----..T-----------I-K------------------------....-.------------.H- 444
B.CH.08.M2_0803101_NFLG8       ------------D---------------------------G----------------S------------AA.........P--.-.--......-..G-K--.---R..T-----------I-K-----------------------A....-.------------.H- 441
B.CN.10.DEMB10CN002            ----------T-----------------------------------------S--------S-------------Q......-..TT-.-.--......-..-----.----..T----------D--I-K------------------------....-.------------.H- 444
B.ES.10.DEMB10ES002            ------------D----------------------------------------------------------N......T...TS.V.FV......-..-----.--K-..T-----------I-K--------------R---------....-.------------.H- 440
B.FR.11.DEMB11FR001            ----------------------------------------------------------------------------......-..---.-.--......-..K----.--KR..SI----------L---------R----------------....-.------------.-- 442
B.GB.05.MM45d213_GN1           --R-------T----------------------------TG-------------------------------......-..--A.-.--......-..K----.--KR.RP--------V--I-K--K----R---------------N....-.------------.H- 451
B.HK.06.HK003                  --R----------------------------------------------------A--------------......-..---.V.--......-..-----.----..-----------I-K--------------Q---------....-.------------.H- 442
B.HT.05.05HT_129389            ---------------G------------------------G-------------------------------......-..TNA.-.-I......-..GSH--.----..F-----------I-K------#-----------------....-.------------.-- 439
B.JP.12.DEMB12JP001            ----------------------------------------------------------------------------......-..PT-.-.--......-..-----.--KR..TI------------K------------------------....-.------------.H- 442
B.KR.07.HP_18_07JHS10_3909     ----------------------------------------G-----------------------------M-......-..PT-.-.--......-..-----.----..T-----------I-K------------------------....-.------------.H- 442
B.PE.07.502_0525_wg5           ------------D-----------------------------------------------------------A-......S.SA-A.-.--......-..K----.----..-I----------L-K-------R----------------....-.------------.P- 443
B.RU.11.11RU21n                ----------------------------------------------------------------------------......S..---.M.--......-..---YK.----..T-----------I-K------------------------....-.------------.H- 447
B.TH.08.MERLBDTRC10            --R---------D-------------S------------------------------S------------A-......-.SA-A.-.--QRNNFK-..----K.G--R..V-----------I-K--------------Q---------....-.------------.H- 449
B.US.11.ES38                   ------------D------------------R-----M--G-----------------------------M-......-..-T-.V.--......-..K----.----..TI----------L--------------------------....-.------------.H- 441
C.AR.01.ARG4006                -F--------T------------------------R----G-S----------------------------N......-...TN.-.--......-..----K.GPKR..TI-----R---L--------------------------....-.------------.N- 437
C.BR.07.DEMC07BR003            -F--------T--------D------------S--R----G------------------------------N......-...TN.-.--......-..-S--K.GPKR..TI----------L----------------Q---------....-.------------.H- 436
C.BW.00.00BW5031_1             -F--------T-D------D------G-----N--R----G-----------------------------AN......-...IN.-.--......-..-N--K.GPKR..-------------I-K------------------------....-.------------.N- 438
C.CN.10.YNFL19                 -F--------T-D------D---------------R----G-SI--------------------------AN......S....N.-.L-......-..-S--K.GPKR..-------------I-K------------------------....-.------------.H- 437
C.CY.09.CY260                  -F--------T------------------------R----G----------------S------------AN......H..---.-.--......-..KS--K.GSKR..-------------I---------R----------------...V-.------------.H- 443
C.ES.08.X2363_2                -F--------T----G-------------------R----G----------------S------------TN......-....A.-.--......-..-S--K.GPKR..-------------I-K-----------R-----------S....-.------------.H- 438
C.ET.02.02ET_288               -F--------T-D----------------------R----G-S---------------------------AN......Q...-N.-.--......-..K---K.GP-R..TI----------L--------------------------....-.--------L---.H- 437
C.IN.09.T125_2139              -F--------T-D------D---------------R----G----------------S------------TN......-....A.-.--......-..-S--K.GA-R..-------------I----K---------------------....-.------------.H- 437
C.KE.00.KER2010                -F--------T--------D---------------R----G-------S---------------------AN......-...-N.-.-I......-..-S--K.SPKR..-------------I--------------------------....-.------------.H- 436
C.MW.09.703010256_CH256.w96    -F--------T-D------D---------------R----G----------------S------------TN..........TN.-.L-......-..-S--K.GT-R..-------------I-K------------------------....-.------------.H- 437
C.TZ.08.707010457_CH457.w8     -F--------T--------D---I----------------G----------------S------------AN......-...-N.-.--......-..-S--K.GTKR..-I-----------I--------------------------....-.------------.H- 437
C.YE.02.02YE511                -F--------T-D------D---------------R----G-S----------------------------N......-...PN.-.--......-..-N--K.GPKR..-------------I-K-----------------------N....-.------------.H- 436
C.ZA.08.705010534_CH534.w12    -F-------ST---------------------G--R---H---I-------------S-------------N......-...TN.-.--......-..KS--K.GPKR..-I-----------I-K-------R----------------....-.------------.H- 435
C.ZA.10.DEMC10ZA001            -F--------T------------------------RG---G-----------------------------AN......S....N.-.--......-..KS--K.GP-R..-------------I-K------------------------....-.-------R----.H- 438
C.ZM.11.DEMC11ZM006            -F-C-----CT-D------D--------------------G-S--------------S------------AN......H...-N.-.--......-..KS--K.GAKR..----------V--I-K--------------Q---------....-.--------L---.H- 439
D.CM.10.DEMD10CM009            ------------D---------------------------G----------------S---------I--AN......-......V.--......-..----K.GSKR..T-----------I---------R----------------....-.-------R----.N- 440
D.CY.06.CY163                  ------------D---------------------------G----------------------------KAA......T..NTA.-.--......-..----K.SPKR..-------------I-K------------------------....-.----------S-.H- 442
D.KE.11.DEMD11KE003            ------------D--R----------------N----M--G----------------S------------A-......S.VN-A.-.--......-..-N--K.GPKR..-------------V-K-------R----------------....-.------------.Q- 443
D.KR.04.04KBH8                 ---------------------------------------AG-S--------------S------------A-......S.SN-A.-.--......-..----K.GP--..-------------I-K-------R----------------....-.------------.HN 442
D.SN.90.SE365                  -------------------D------------N-------G----------------S------------A-......S.SA-A.-.--......-..----K.G---..-------------I----K-----------Q---------....G.------------.S- 443
D.TZ.01.A280                   ------------D---------------------------------------------S------------A-......-.VN-A.-.--......-..----K.GP--..-I-----------I-K------------------------....-.------------.H- 437
D.UG.10.DEMD10UG004            ------------D---------------------------------------------S------------A-......-.VN-A.-.--......-..--KYT.GP--..SI----------L----K---------------------....-.-------RF---.Q- 443
D.UG.11.DEMD11UG003            ---------G--D------------------------M--.----------------S------------A-......-.AN-A.-.--......-..----K.GPKR..-I-----------I-K-------R---R------------....-.-------N----.R- 441
D.YE.02.02YE516                ----------------------------------------G-S--------------S-----------IA-......S.NN-A.-.--......-..----K.-P--..-------------I-K------------------------....-.-------R----.S- 441
D.ZA.90.R1                     ------------D---------------------------Q-----------------------------A-......-.LATA.V.--......-..----K.G---..-I-----------I-K--------------R---------....-.------------.H- 443
F1.AO.06.AO_06_ANG32           -F-C------T----G---D-------------------SGGSI---------------------------Q......-...-S.-.--......-..KS--K.G---..TI----------L----------------R---------....-.-------R----.H- 438
F1.AR.02.ARE933                -F-A------T-D------D----------------G--SG----------------S------------A-......S...T-.V.--......-..KS--K.G-KR..----------V--I-K--------------R---------....-.------------.N- 437
F1.BR.10.10BR_RJ015            -F--------T----G---D----------------------S-----------------------------......-...P-.V.--......-..KS--K.G--R..P-----------I----------------R---------....-.------------.H- 440
F1.CY.08.CY222                 -F--------T----G---D--------------------G---------------------I-------A.......-...TA.-.--......-..KS--K.G---..T-----------I--------------------------....-.------------.N- 436
F1.ES.02.ES_X845_4             -F--------T--------D--------------------G------------------------------Q......-...TS.-.--......-..K---K.G---..-I----------L----------------R---------....-.-------R----.H- 438
F1.RO.96.BCI_R07               -F--------T----G---D-----H--------------G----------------S----I-------T-......S...-A.-.--......-..KS--K.G---..-I-----------I-K------------------------....-.------------.H- 437
F1.RU.08.D88_845               -F-------CT----G---D-----------R--------G----------------S----I-------A-......S...TA.-.--......-..KS--K.G-K-..-------------I----------------R---------....-.-------R----.S- 443
F2.CM.02.02CM_0016BBY          -F--------T----G---D--------------------G---------------------I---------......A...TS.V.L-......-..KS--K.G-KR..-------------I-K-------R----------------....-.------------.N- 435
F2.CM.10.DEMF210CM001          -F-I------T--------D---------------------------------------S----I-------A-......AVTA-S.M.--......-..KS--K.G--R..TI----------L---------R------Q---------....-.------------.N- 440
F2.CM.10.DEMF210CM007          -F--------T----G---D--------------------G-------------------------------......-...T-.-.--......-..KS--K.G--R..----------V--I-K------------------------....-.--------L--P.S. 436
G.BE.96.DRCBL                  -F--------T----S---D---I--------I---G--QG----------------S------------AS......G.AA-A.-.--......-..KS--K.GP-R..TI----------L-----------------------E--....-.------------.N- 436
G.CM.10.DEMG10CM008            -F--------T-D------D---------------R----G-------------------------------......-.TGTA.-.--......-..----K.GPKR..NI----------L--------------------------....-.------------.S- 442
G.CN.08.GX_2084_08             -F--------T-D----------------------RG--TG----------------S-----------NAS......G..A-A.-.--......-..KS--K.GP-R..XI----------L----------------R---------....-.------------.N- 439
G.CU.99.Cu74                   -F--------T----G---D------------A--R----G----------------S------------AS......G.AG-A.V.L-......-..-S--K.GP-R..NI----------L--------------------------....-.--------L---.N- 443
G.ES.09.X2634_2                -F--------T--------D---------------R----G-------S---------------------A-......G...-A.-.--......-..KS--K.GPKR..MI-----R---L--------------------------....-.-------R----.N- 440
G.GH.03.03GH175G               -F-A------T----H-------I------------S--TG------------------------------S......G..ATS.-.--......-..NS---.GPKR..NI----------L-----------------------E--....-.------------.N- 441
G.KE.09.DEMG09KE001            -F-V--------D--S---D-------------------QG----------------S------------AS......L...PS.-.--......-..KS--K.GP-R..TI----------L----------------R---------....-.------------.N- 441
G.NG.09.09NG_SC62              -F--------T----G---D---------------R----G----------------S------------A-......-.A--A.-.--......-..KS--K.GP--..NI----------L----------------R---------....-.-----------P.HS 436
G.ZA.01.TV546                  -F-C------T--------D----------Y--------QG----------------S------------AS......G..ATA.-.--......-..KS--K.GP-R..MI----------L--------------------------....-.------------.H- 440
H.BE.93.VI991                  -FRV------T-D------D-----------R-------RG-SI-------------S--------------......-.A--A.-.--......-..K---K.GP-R..T---S-------I----------------Q---------....G.------------.S- 445
H.CF.90.056                    -F--------T-D----------------------R---QG-SI-------------S--------------......-.TNTA.-.--......-..K---K.G---..F-----------I----------------R---------....-.------------.S- 443
H.GB.00.00GBAC4001             -F--------T-D------D-------------------QG-SI--------------------------M-......-.MNTA.-.-V......-..K---K.G--R..T-----------I----------------R---------....-.------------.S- 443
J.CM.04.04CMU11421             -F--------T--------D---I-------R--------G-----------------------------M-......-..T-N.-.--......-..X---K.G-KR..MI-----P--L----------------Q---------....-.------------.N- 442
J.SE.93.SE9280_7887            -F-A------T-D------D-------------------SG-------------------------------......-...TN.-.--......-..-----.DHKR..-----------Q--I-K------------------------....-.------------.S- 441
J.SE.94.SE9173_7022            -F-A------T-D------D-------------------SG-------------------------------......-...TN.-.--......-..-----.DHKR..-----------Q--I-K------------------------....-.------------.S- 441
K.CD.97.97ZR_EQTB11            -FRV------T--------------------R-------SG-------------------------------......-...SA.V.--......-..----K.G--R..-I----------L-------------------------S....-.--------F--L.N- 437
K.CM.96.96CM_MP535             -F--------T--------D--------------------G-S--------------S----I---------......-...PV.V.--......-..K---K.GH--..-------------I--------------------------....-.------------.H- 437
U.CA.01.TV749                  -F----------D------D------------S------TG----------------S----I-------AQ......-...TN.-.--......-..-----.G-KR...I----------L----K---R-----------------....-.-------R----.H- 436
U.CA.99.TV721                  -F----------D------D-----------RS------AG----------------S----I-------AQ......-...TN.-.--......-..-----.G-K-...I----------L--------------------------....-.-------R----.H- 436
U.CD.83.83CD003_Z3             -F-V-----S---------D---I---------------TG----------------S------------T.......-...TA.-.--......-..----K.GP-R..-------------I-K--------------R---------....-.------------.N- 440
U.CD.90.90CD121E12             -F-------S---------D---T---------------TG-----------------------------M-......-...T-.V.--......-..----K.G---..-------------I-K--------------R---------....-.------------.N- 438
U.CY.05.CY090                  -F--------T--------D------------H-------G----------------T----I------KAA......T..--N.-.-I......-..----K.GP--..M-------------K-------R----------------....-.------------.QS 438
U.CY.08.CY223                  -F--------T--------D---------------R----G----------------S------------AG......G..A-A.-.--......-..K---K.GP-R..NI----------L-K--------------R---------....-.------------.N- 443
U.ES.10.DEURF10DZ001           -F--------T------------------------R----G----------------S------------AQ......Q...HN.-.-V......-..-----.GP-T...I----------L----K----R----------------....-.------------.S- 437
U.GR.99.99GR303                -F----GK--T----------------------------SG------------------------------Q......S...PN.-.--......-..K----.G-KR...I----------L---------X---------------P....-.-------R----.S- 438
U.NL.95.U_NL_95_H10986_D1      -F-I--------D------D----------------------S-------------------I-------A-......K...TV.V.--......-..----K.GTKR..-I-----------I--------------------------....-.-------RV---.N- 438
01_AE.AF.07.569M               -F--------T--------------------RS------SG----------------S------------AQ......H...--.-.--......-..----K.G-KR...I----------L--------------------------....-.------------TN- 441
01_AE.CF.90.90CF11697          -F--------T---------------------S------TG------------------------------Q......H...-N.-.--......-..----K.G-KR...I----------L---------R------Q---------....-.-----------L.N- 440
01_AE.CN.10.YNFL03             ----------T---------------------S------TG-------V--------------------H-Q......Q...-S.-.--......-..----K.GPK-...I----------L--------------------------....-.-------R----..- 439
01_AE.HK.04.HK001              ----------T---------------------S------AG----------------S----------N-AQ......Q...T-.V.--......-..---XK.G-KR...I-----R---L--------------------------....-.-----------P.N- 440
01_AE.IR.10.10IR.THR48F        ----------T---------------------S------SG----------------S-----------HTQ......Q...-N.-L--......-..----K.G-KR...I----------L--------------------------....-.------------..- 440
01_AE.JP.x.JRC77AE             --------------------------------S------TG----------------S------------AQ......H...TN.-.L-......-..----K.G-KT...I----------L---------------------L----....-.-------R----.H- 440
01_AE.TH.09.AA090a_WG11        ----------T---------------------S------TG----------------S-------------Q......P...TN.-.-I......-..G-K-N.GP-R...I----------L--------------------------....-.------------.H- 438
01_AE.TH.90.CM240              ----------T---------------------S------TG----------------S-----------HAQ......H...--.-.--......-..----K.G-KR...I-----R---L---------Q----------------....-.------------.N- 440
01_AE.US.05.306163_FL          ----------T---------------------S--RG--SG----------------S------------AQ.....AQ...VN.-.--......-..----K.G-KR...I----------L--------------------------....-.-----------P.N- 437
01_AE.VN.98.98VNND15           ----------T---------------------S------TG----------------S-----------H-Q......H...TN.V.--......-..----K.G-KR...I----------L--------------------------....-.------------..- 439
```

```
                                   major homology region                                            p24 end p2 start        p2 end p7 start           Zn motif          Zn motif            p7 end p1 start
B.FR.83.HXB2                  FYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVT......N..SAT.I.MM......Q..RGNFR.NQRK..IVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCT....E.RQANFLGKIWPS.YK 442
02_AG.CM.10.DE00210CM013      -F--------T---------------------S--R----G------------------------------Q......Q...TN.-.-..........-S---.GR-T...I----------L----K---------------------....-.------------.N- 442
02_AG.CY.09.CY256             -F--------T------------------------R----G----------------S-------------Q......Q...PN.-.-..........-----.G-KT...I----------L----K----R----------------....-.-------R----.S- 437
02_AG.ES.06.P1423             -F--------T---------------------S------TG-X----------------------------Q......Q...-G.V.-..........-----.G--T...I----------L----K----R----------------....-.-------RX---.S- 437
02_AG.GW.05.CC_0048           -F--------------------M-----------RG----G----------------S------------AQ......Q...SN.-.-..........-----.G--T...I----------L----K----R----------------....-.------------.N- 442
02_AG.KR.12.12MHR9            -FR-------T---------------S-----A--R----G----------------S-------------Q......Q...PN.-.-..........-----.G--T...I----------L----K----R---------------P....-.------------.S- 437
02_AG.LR.x.POC44951           -F--------T---------------------S------SG------------------------------Q......Q...-N.-.-..........-----.GR-T...I----------L----K---------------------....-.------------.S- 437
02_AG.NG.09.09NG_SC61         -F--------T---------------------N--R---QG--------------------K---------Q......Q...TN.-.-..........-----.G--T...I----------L----K----R----------------....-.------------.S- 435
02_AG.NG.x.IBNG               -F--------T---------------------S--R---TG------------------------------Q......Q...-N.V.-..........-----.G--T...I----------L----K----R----------------....-.------------.S- 437
02_AG.SE.94.SE7812            -F--------T-D-------------------S--R----G----------------S-------------Q......Q...PN.-.-..........-----.G--P...I----------L----K---------------------....-.------------.S- 437
02_AG.SN.98.98SE_MP1211       -F--------T---------------S----R--------G-------S---------------------AQ......Q...SN.-.-..........-----.G--T...I----------L----K----R----------------....-.------------.S- 436
03_AB.RU.97.KAL153_2          -F--------T-D----------------------R---SG--------V---------------------Q......-...-N.-.-..........KS---.GPKR...I----------L--------------------------....-.-------R----.S- 440
04_cpx.CY.94.94CY032_3        -F-C------T---------------------S------TG----------------S------------AS........AA-A.-.-..........KSK-K.G--R..TI----------L--------------------------....-.-------RM---.S- 443
05_DF.BE.x.VI1310             ----------T----------------------------QQ----------------S------------A-........SA--.A.--.........----K.GP--..-I----------I-K-----G--------R---------....-.--------V---.H- 444
06_cpx.AU.96.BFP90            -F--------T--------D---------------R----G-----------------------------AS......V...GA.-.-..........KS--K.GPKR..SI----------L--------------------------....-.------------.N- 440
07_BC.CN.98.98CN009           -F--------T-D------D---------------R----G-SI-------------S------------TN......S.......-.L-.........-S--K.GSKR..------------I--------------------------....-.------------.H- 437
08_BC.CN.97.97CNGX_6F         -F--------T-D------D---------------R---SG-S--------------S------------TN.............-.L-.........-S--K.GSKR..------------I-K------------------------....-.--------F---.H- 435
09_cpx.GH.96.96GH2911         -F--------T-D---*------------------R----G------------------------------Q......-...TN.-.L-.........----K.G-KR...I---------XL--------------------------....-.------------.N- 439
10_CD.TZ.96.96TZ_BF061        ------------D--------------------------------------------S------------A-......G..GN-.-.-..........----K.GPK-..SI------------K------RR------R----L---P....-R------------.N- 443
11_cpx.CM.95.95CM_1816        -F-C------T----Q----S-----------S--R---TG------------------------------Q......Q...TN.-.-..........----K.G-KR...I----------L--------------------------....-.------------.S- 440
12_BF.AR.99.ARMA159           -F--------T--------D-------------------SG-----------------------------A-......S...TA.V.L..........KSS-N.G--R..---------T--I-K--K--------------------P....-.------------.N- 437
13_cpx.CM.96.96CM_1849        -F--------T-D------D---------------R----G-----------------------------AS......SAAA-A.-.-..........KS--K.GPKR..-I------R---L-------------------------M....-.------------.N- 444
14_BG.ES.05.X1870             -F--------T----G---D---------------R----G------------------------------S......G...-S.V.-..........K---K.G--R..NI----------L--------------------------....-.-------R----.S- 440
15_01B.TH.99.99TH_MU2079      ----------T------------A--------S------SG-S--------------S-----------H-Q......Q...TN.V.-..........----K.G-KR...I----------L----------------E---------....-.-------RL---..- 439
16_A2D.KR.97.97KR004          -F--------T---------------------S--R----G-S--------------S-------------Q........TNSNIM.-..........K----.G-K-...#----------L--------------------------....-.------------.HS 440
17_BF.AR.99.ARMA038           -F-------------G---D--------------------G-Q-----------------------------..........P-.V.L..........KS--K.GP--..N-----------I----------------R---------....-.------------.N- 443
18_cpx.CU.99.CU76             -F--------T---------------------N--R----G----------------S-----------H-Q......-...TN.-.-..........----K.G-KR...I----------I--------------------------....-.------------.N- 440
19_cpx.CU.99.CU7              ------------D---------------------------G----------------S------------A-......-.SH--.-.-..........K---K.GL--..V-----------I--------------------------....-.------------.H- 443
20_BG.CU.99.Cu103             -F--------T----G---D---------------R----G------------------------------S......G.AG-S.-.L-.........KS--K.GPKR..NI----------L--------------------------....-.--------L---.H- 443
21_A2D.KE.99.KER2003          -------------------D------------S-------G----------------S------------A-......-.SAV-.V.-..........----K.GPKR..-I----------I-K--------------------R---....-.------------.S- 441
22_01A1.CM.01.01CM_0001BBY    -F--------T-D-------------------S--R---TG------------------------------Q......H...TG.-.-..........----K.G-KR...I----------LT-------R-----------------L.TL-.------------.S- 441
23_BG.CU.03.CB118             -F--------T----G---D---------------R----G-S---------------------------IP......G.AG-S.-.-..........-S--K.GPKR..NI-------D--L-----------------D--------....G.--------L---.H- 444
24_BG.ES.08.X2456_2           -F--------T----G---D---------------R----G-S---------------------------IS......G.AG-S.-.-..........KS---.GP-R..NI----------L-K-----------------------S....-.--------L---.H- 443
25_cpx.CM.02.1918LE           -F--------T----G---D---------------R----G-S---------------------------AS......G.TA-A.-.-..........----K.GP-R..NI----------L-----------------D--------....-.-------R----.S- 442
26_AU.CD.02.02CD_MBTB047      -F--------T----A---D--------------RG---QG-S----------------------------Q......-...-N.-.LV.........G---K.G--G.K-I----------L----K-----------R---------....-.-------RF--P.N- 439
27_cpx.FR.04.04CD_FR_KZS      -F--------T--------S---------------R----G----------------S------------AS......G.AVTA.-.-..........K---K.G-KR..-I----------L----------------R---------....-.----------SP.N- 440
28_BF.BR.99.BREPM12609        -F-------CT----G---D--------------------G----------------S------------A-......-...T-.V.--.........----K.G---..------------I----------------R---------....-.------------.H- 441
29_BF.BR.01.BREPM16704        -F-------CT----G---D--------------------G----------------S------------A-......-...T-.V.--.........KS--K.G--R..------------I-----------------------E--....-.------------.H- 452
31_BC.BR.04.04BR142           -F--------T--------D------------S--R----G------------------------------N......-...TN.-.-..........-S--K.GPKR..TI----------L-------------------------N...N-.-------R----.H- 441
32_06A1.EE.01.EE0369          -F--------THD----------------------R----G-----------------------------AQ......-...-N.-.-..........KS---.GPKR...I----------L----------------Q---------....-.------------.S- 437
33_01B.ID.07.JKT189_C         ----------T---------------------S------SG----------------S-------------Q......Q...TG.-.-..........----K.G-KR...I--------.-L----------------Q---------....-.--------L---.H- 438
34_01B.TH.99.OUR2478P         ----------T---------------------S------TG----------------S------------TQ......Q...-NMM.-..........----K.G-KR...I----------L--------------------------....-.K-I-----F---.N- 439
35_AD.AF.07.169H              -F-V------T----G----------------S------TG----------------S------------AQ......H...TN.-.-..........----K.G-KR...I----------L--------------------------....-.------------.NN 439
36_cpx.CM.00.00CMNYU830       -F--------T-D-------------------S------TG------------------------------Q......H...-G.-.-I.........-S---.G--T...I----------L--------------------------...N-.------------.S- 440
37_cpx.CM.00.00CMNYU926       -F-V-----ST---------------------S------TG----------------------------TAQ......Q...TS.-.-..........-S--K.GPKR..NI----------L--------------------------....-.-------R----.S- 437
38_BF1.UY.03.UY03_3389        -F-VV----ST--------D------------S-------G-------S----------------A------......-...SA.V.--.........K---K.GP-R..T-----------I-K--------------R---------....-.------------.H- 437
39_BF.BR.04.04BRRJ179         ---------------A---D--------------------G----------------S------------AN......Q...TN.-.-V.........----K..P-R..T-----------I-K--------------------L---....-.------------.N- 440
40_BF.BR.05.05BRRJ055         -F--------T--------D--------------------G----------------S----I-----N---......-...TA.V.-..........KS--K.G--R..------------I-K--------------R---------....-.-----------P.-- 442
42_BF.LU.06.luBF_18_06        -F--------T----G---D--------------------G-S---------------------------A-......G...TA.-.L-.........-S--K.GP--..------------I-K--------------R---------....-.------------.N- 440
43_02G.SA.03.J11223           -F--------T-D------D------------N--R---SG----------------S------------AS......G.AA-A.-.-..........K---K.GP-R..NI----------L--------------------------....-.------------.H- 440
44_BF.CL.00.CH80              -F--------T--------D---------------R----G-------------------------------......-...TA.V.-..........KS--K.G--R..------------I--------------------------....-.------------.N- 443
45_cpx.FR.04.04FR_AUK         -F--------T------------------------R----G----------------S---KI-------AQ......S.TNLM.M.-..........----K.G-K-...I----------L--------------------------....-.------------.SR 442
46_BF.BR.07.07BR_FPS625       -F--------T----G---D--------------------G----------------S----I---------......H...TN.V.L..........KS--K.G--R..------------I----------------R---------....-.------------.N- 437
47_BF.ES.08.P1942             -F-V------T-D-------------S-------------Q-----------------------------A-......A..-NV.-.-..........K--Y-.G---..---------R--I-K------------------------....-.------------.S- 441
48_01B.MY.07.07MYKT021        ----------T----S----------------S------TG----------------S------------AQ......Q...GG.-.-..........----K.G-KR...I----------L--------------------------....-.--------L---.H- 440
49_cpx.GM.03.N26677           -F--------T-----------------------RG----G-S----------------------------Q......-...SN.-.-..........K---K.G---...I----------L-K-----------R------------....A.G------R----.H- 440
50_A1D.GB.10.12792            --R---------D----------------------------------------------------------Q......S...SN.-.-..........----K.G-KR...I----------L--------------------------...N-.------------.H- 440
51_01B.SG.11.11SG_HM021       ------------D--------------------------------------------S------------A-......S..ATS.-.-..........-----.---R..T-----------I-K------------------------....-.------------.H- 440
52_01B.MY.03.03MYKL018_1      ----------T---------------------S------SG----------------S------------AQ......Q..HXN.-.-..........----K.GPKR...I----------L----------------Q---------....-.------------.N- 441
53_01B.MY.11.11FIR164         ----------T---------------------S------TG----------------S-----------NAQ......H......-.-..........----K.G--R...I----------L--------------------------....-.--------L--X.H- 440
54_01B.MY.09.09MYSB023        ------------D---------------------------G----------------S----I---------.........PP-.-.I..........-----.--GR..T-----------V--------------------------...N-.------------.H- 443
55_01B.CN.10.HNCS102056       ----------T---------------------S------TG----------------S-----------HMQ......S...TN.-.-..........----K.G---...I------R---L--------------------------....-.-------RL---.N- 440
56_cpx.FR.10.URF5_patient_A   -F--------T-----------------------RG----G-----------------------------AQ......Q...-N.-.-I.........-----.G--S...I----------L----K----R----------------....-.------------.N- 440
57_BC.CN.09.09YNLX19sg        ------------D--------------------------------------------S------------TN......K...................-S--K.GPKR..------------I--------------------------....-.--------L---.H- 437
58_01B.MY.09.09MYPR37         ----------T---------------------S------AG----------------S-----------HAQ......Q...TG.-.-..........----K.GNXX...I----------L-K------------------------....-.--------F--P.N- 439
59_01B.CN.09.09LNA423         ----------T---------------------S------TG----------------S------------AQ......Q...T-.-.-..........----K.G-KR...I------R---L--------------------------....-.------------..- 435
60_BC.IT.11.BAV499            -F--------T---------S-----S--------R----GX----------------------------LN..........-N.-.-..........-S--K.GPKR..TI----------L--------------------------....-.-------R----.H- 437
61_BC.CN.10.JL100010          -F--------T-D------D---------------R----G----------------S------------TN......S....A.-.L-.........-S--K.GPKR..------------I--------------------------....-.------------.H- 437
62_BC.CN.10.YNFL13            -F--------T-D------D---------------R----G-S--------------S------------TN......S....A.-.-..........-S--K.GPKR..------------I-------------------------A....-.------------.H- 433
63_02A1.RU.10.10RU6637        -F--------T---------------------A--R----G------------------------------Q......Q...SN.-.-..........-----.G-KT...I----------L--------------------------....-.-------RL---.NN 437
64_BC.CN.09.YNFL31            -F--------T-D------D---------------R----G-SI-------------S----------G-AN......-....V.-.L-.........----K.----..------------I-K------------------------....-.------------.H- 437
65_cpx.CN.10.YNFL01           -F--------T-D------D---------------R----G-SI-------------S------------AN......-....-.-.-..........----K.GPKR..------------I-K------------------------....-.------------.H- 440
67_01B.CN.11.2011.ANHUI.HF115 ----------T---------------------S------TG----------------S-----------X-Q......H...TN.-.-..........K---K.GPKR...I----------L--------------------------....-.------------.N- 440
68_01B.CN.11.2011.ANHUI.WH73  ----------T---------------------S------TG----------------S-----------X-Q......Q...-S.V.-..........K---K.GPKR...I----------L--------------------------....-.------------.N- 440
70_BF1.BR.10.10BR_PE004       -F-A-K----T----G---D---------------R----E--I-K------------------------AN..........................KS--K.G-GR..---------QR-I-K--K-------------R-L-----....-.------R-----.H- 443
71_BF1.BR.10.10BR_PE008       -F--------T---------------------A--R----G------------------------------N..........T-.-.-..........KS--K.G---..-I------R---L----K-----------R---------....-.-------RM---.S- 440
72_BF1.BR.10.10BR_MG002       -F--------T----G---D---------------R----G----------------S------------A-......K...TA.V.-..........KS--K.G---..V-----------I----------------R---------....-.------------.H- 441
O.BE.87.ANT70                 ----------T---------------------Q---S---G-------V--------T----------ATAQ...QDLKGGYTA.V.F-.........--QNP.IRKG..TI----------I----------------Q--------R...NG.K-------Y--P.GG 447
O.CM.98.98CMA104              --------------------------------Q-------G-------V--------T----------ASAQ...QDLKGGYTA.V.F-.........--QNP.IR-G..P-----------I-------------N--QD-------R...NG.K-------Y--P.GG 447
O.CM.98.98CMABB141            --------------------------------Q---S---G-------V--------T---KI-----ASAQ...QDLKGGYTA.V.F-.........--QNP.IRKG..PI----------L-K-------R------Q-----R--K...NG.K-------Y--P.GG 447
O.CM.98.98CMABB212            ---------------S----------S-----Q-------G-------V--------T---KI-----VSAQ...QDLKGGYTS.-.F-.........--QSP.-K-R..TI----------V-K-------R------Q--------N...NG.--------Y--P.GG 447
O.CM.98.98CMU5337             ----------T---------------------Q---S---G-------V--------T----------ATAQ...QDLKGGYTA.V.F-.........--QNP.IRKG..PI----------L----------------Q--------R...NA.K-------Y--P.GG 447
O.CM.99.99CMU4122             --------------------------------Q---S---G-------V--------T----------ATAQ...QDLKGGYTA.V.F-.........--QNP.GRKG..PI----------L----------------Q--------R...NG.K-------Y--P.GG 447
O.FR.92.VAU                   ----------T---------------------Q-------G-S-----V--------T----------ASAQ...QDLKGGYTS.V.F-.........--QKP.IRKG..PI----------VTK--------------Q--------K...NR.--T---K-Y--P.RG 446
O.GA.11.11Gab6352             ----------T---------------------Q-------G-S-----V--------T----------ATAQ...QDLKGGYTS.V.F-.........--QNP.VK-G..P-----------I-K---S--R-------Q--------K...NG.--------Y--P.GG 446
O.SN.99.99SE_MP1299           ----------T---------------------Q---S---G-------I--------T----------AAAQ....DLKGGYTA.V.F-.........--QNP.SRKG..PI----------L----------------Q--------K...NG.--------Y--P.GG 446
O.US.10.LTNP                  --------------------------------Q---S---G-------V--------T----I-----ATAQ...QDLRGGYTA.V.F-.........--QNP.-R-G..AI----------I-K--------------Q---L----R...NG.K-------Y--P.GG 447

N.CM.02.DJO0131               ------------D-------------------Q------------------------A-------------Q......T..ATS.V.FA.........GR--K.GI-R..TIR---------L----K---RG------Q--------K..NEG.--------G-SP.F- 449
N.CM.02.SJGddd                ----------T---------------------Q---S----------------------------------Q......A..TTS.V.FV.........----K.GI--..TI----------L----K---R-------Q--------N..NEG.-------RG-SP.F- 447
N.CM.04.04CM_1015_04          ----------T-D-------------------Q-------S-S-------------------------A-AQ......A..ATS.V.FV.........----K.GT-R..TI----------L----K---RX------Q--------K..NEG.--------G-SP.F- 446
N.CM.06.U14296                --R---------X------D------------Q-------G----------------A-------------Q......A..TTS.V.FA.........----K.GI-R..PI----------L----K---RG------Q--------K..NEG.-------RG-SP.F- 446
N.CM.06.U14842                ----------X--------D------------QL------G----------------A-------------Q......G..ATS.V.FV.........----K.GM--..TI----------V----K---RG------Q--------K..NEG.--------G-SP.F- 446
N.FR.11.N1_FR_2011            ----------T--------D------------QL------G----------------A-------------Q......P..TTN.-.FA.........----K.GI--..PI----------L----K---R-------Q--------K..SEG.-------RG-SP.F- 447

P.CM.06.U14788                ---------------S---D------------Q-------G-------N--------T----------AAANQASQELKGGYT-.V.F-......-S.......G---..P--------V--I-K--K---RR------Q--------K...SG.-----------P.GG 440
P.FR.09.RBF168                ---------------S---D--I---------Q-------G-------N--------T----------AAANQANQDLKGGYT-.V.F-......-S.......G--X..P--------I--I-K--K---R-------Q--------K...SG.-----------P.GG 443

CPZ.CD.06.BF1167              ----I-------P---------------------------G------L---------A----------Q-TQ......Q..NVN...-V......-G.-AAPQ.GK-GAGNI-------I-----------R-------Q----L----..QSK.SGV-----RT-W.WG 470
CPZ.CM.05.SIVcpzMT145         ------------D------D--I---------N--R----G----------------S-------------N......K...TE.-.F-......-R.G-QN-.PPN-..KI----------L--------R-------Q--------N..QNN.GRI-----N--P.WG 434
CPZ.GA.88.GAB1                -------------------D------------Q-------G----------------S-----------M-Q......-QGR-D.V.FF.........K-QGA.GPKR..KI----------L----K---R----R--Q---------....G.--V-----G---.RS 452
CPZ.TZ.06.SIVcpzTAN13         ----I-------P--T------I---------H------QG------L---------T---KI-----ATA-......-TVGVN...-L.........G-KKP.LRKG..QLQ-------L---K-----------R--Q---------ARTNN.TGV-----RT-L.WG 474
CPZ.US.85.US_Marilyn          -----------P---T-------------------R----G------V---------A----------C-MK.........PSS.V.FL.........K--AG.KPGR..KI----------L--------R-------Q-----------..AGN.--------H-SP.SW 445

GOR.CM.07.SIVgor2139_287      ------------D---------I---------Q-------G-------N--------T----------AAAH...SDLKGGYTA.V.F-......-KGGN-QG.IKKG..P--------V--I-K------R-------QD-------R...SG.--------Y--P.GG 442
GOR.CM.07.SIVgorCP2135con     ------------D---------I---------Q-------G-------N--------T----------AAAH...SDLKGGYTA.V.F-......-KGGX-QG.IKKG..P--------V--I-K------R-------Q--------R...SG.--------Y--P.GG 442
```

```
                                                                                               Vpr binding
                               p1 end p6 start       Vpr binding                               p6 end, Gag end
B.FR.83.HXB2                 ....GRPGNFLQSRP............EPT..........APPEESFRS.....GVETTT..PP......Q.....KQEPI....D.....KE........LY.PLTSLRSLFGNDPSSQ*   500
A1.AU.03.PS1044_Day0         ....------P----............---..........---A-L-GM.....-E-I.-..S-......P.....---QK....-.....--........QVP--V--K----------    500
A1.CH.03.HIV_CH_BID_V3538    ....------P-N-L............---..........---A--GV.....RE--.-..SS......M.....---QR....-.....R-........QPP--V--K-------L---   499
A1.CY.08.CY236               ....------P----............--S..........---A-I-GM.....-E-I.-..S-......P.....---QK....-.....S-..........P--V--K----S------  501
A1.ES.05.X1608_8             ....------P----............--S..........---A-X-GM.....-E-M.-..S-......P.....---QKQEQGN....R-........QAP--I--K-------L---   501
A1.KE.11.DEMA11KE001         ....------P-N-L............---..........---A-ICGM.....-DGI.A..F-......P.....---QK....-.....R-........QGP-SI--K------QL---   499
A1.RU.11.11RU6950            ....------P----............--S..........---A-N-GM.....-E-I.-..-S......L.....---QK....-.....R-........QHP-SI--K-------L---   495
A1.RW.11.DEMA111RW002        ....------P----............--S..........---V-LLGM.....-E-I.A..S-......P.....R--QR....-.....R-........QGP--V--K-------L---   500
A1.SN.01.DDI579              ....------P----............---..........---ADLLGM.....-E-I.-..SS......P.....---QK....-.....R-........P.P--V----------L---   496
A1.UG.11.DEMA110UG001        ....------P----............---..........---A-L-GM.....-E-I.-..S-......P.....N--QK....-.....R-........QAP-FA--K-------*---   496
A1.ZA.04.04ZASK162B1         ....------P-N-L............---..........---AG-YGM.....-E-I.-..S-......-.....-K-QT....-.....R-........Q.P--A---------L-P-   504
A2.CM.01.01CM_1445MV         ....------P---T............---..........---A-NVGM.....-E-I.A..SS......S.....---LR....N.....R-........QHT-TI---------L---   497
A2.CY.94.94CY017_41          ....------P---T............---..........---A-NL-M.....-E-I.-..SS......L.....---LE....T.....R-........P-N-AI--K-------LL--   497
B.BR.10.10BR_MG029           ....-----------............---..........--------F.....-E----..--......-.....-----....-.....--........----A--K-------L---   500
B.CA.07.502_1191_03          ....--------N--............--S..........-------KF.....-E----..--......-.....----T....-.....--........M-.-FS--K-------L---   503
B.CH.08.M2_0803101_NFLG8     ....------F----............---..........--------F.....-E----.-S..QKQE-.......---QK....T.....E-........K-.--A--K-------Y---   504
B.CN.10.DEMB10CN002          ....--------N-T............--S..........--------F.....-E----..-A......-.....----T....-.....--........----S--K-------L---   503
B.ES.10.DEMB10ES002          ....-----------............---..........--------F.....-E-K-A..--......-.....---Q-....-.....--........------------------   499
B.FR.11.DEMB11FR001          ....-----------............---..........--------F.....-E----..-S......-.....---QT....-.....E-........K-------------*---   499
B.GB.05.MM45d213_GN1         ....-----------............---..........--------F.....-E.---..-S......P.....----T....-.....-G........----A-------------   509
B.HK.06.HK003                ....-----------............---..........--------Y.....-E----..--......-KQG..L-G--....-.....-D........------------------   504
B.HT.05.05HT_129389          ....--------N-T............--S..........--------F.....-E-.--..--......-.....---SR....-.....--........----A--K-------*---   496
B.JP.12.DEMB12JP001          ....-----------............---APP.......--------F.....-EK-E-..-S......-.....-----....-.....--........Q---A--K-------*---   503
B.KR.07.HP_18_07JHS10_3909   ....-----------............--S..........--------F.....-E----..-S......-.....-----....-.....--........----A-------------   501
B.PE.07.502_0525_wg5         ....----------TQSRP........---..........--------F.....VE----..-S......-.....----T....-.....--........----A--K----------   506
B.RU.11.11RU21n              ....-----------............---..........--------F.....-E----..-S......-.....---QK....-RDRD.--........----A-------------   510
B.TH.08.MERLBDTRC10          ....--------N--............---..........--------F.....-E----..--......-.....-----....-.....--........----A---------L---   508
B.US.11.ES38                 ....----------ESRP.........--S..........--------F.....-EG-A-..-S......-.....-----....-.....-D........----A-------------   504
C.AR.01.ARG4006              ....-----------............---..........---A---F......EE.-.-..-S......P.....----K....-.....R-........-------K----S--L---   492
C.BR.07.DEMC07BR003          ....------P---L............---..........---A---F......EE.I.-..-S......P.....----K....-.....R-........-------K----S--L---   491
C.BW.00.00BW5031_1           ....-----------............---APP...MPT---A---F......-E.-.-..-A......P.....---MK....E.....R-........---I--K----S--L---   499
C.CN.10.YNFL19               ....--------N-T............---..........---A---F......-E----..-S......-.....-S---....-.....--........-------K-------L---   496
C.CY.09.CY260                ....--------N--............---APPTV..VPT---A--KF......EER-.-..-A......P.....----T....-.....RD..........-I--K----S--W---   505
C.ES.08.X2363_2              ....-----------............---APPA...EPT---A---F......EG.-.A..-A......P.....---LK....-.....R-..........T---K-------W---   500
C.ET.02.02ET_288             ....--------N--............A--APL....EPT---A---F......EE.A.-..--......P.....---LK....-.....R-..........A---K-------L---   498
C.IN.09.T125_2139            ....-----------............---..........---A---F......EE.--P..-A......P.....----K....-.....R-..........T---------S--L---   493
C.KE.00.KER2010              ....--------N--............---..........---A---F......EE.-A...LA......P.....---LK....E.....R-........-------K----S--L---   491
C.MW.09.703010256_CH256.w96  ....--------N--............---APPL...EPT---A--ST..............-A......L.....---QK....-.....R-..........-SI--K-------L---   494
C.TZ.08.707010457_CH457.w8   ....--------N-T............---..........---A---EF.....EE.-.-..H-......P.....---QK....N.....R-..........T---K----S--L---   492
C.YE.02.02YE511              ....-----------............---..........---L---F......EE.-.-..-A......P.....---AK....-.....--........-------K----S--LLP-   489
C.ZA.08.705010534_CH534.w12  ....-----------............---APPA...EPT---A---F......EG.-.-..-A......P.....R--TK....-.....R-........-------K-------L---   497
C.ZA.10.DEMC10ZA001          ....-----------............---..........---A---F......EE.-.-..-A......P.....---SK....-.....R-..........---A-K----S-LL---   493
C.ZM.11.DEMC11ZM006          ....-----------............---APL...QPT---A---F......EE.-.-..-A......P.....----K....-.....R-..........AI---------L---   499
D.CM.10.DEMD10CM009          ....----------EPTAPL.......---..........---A--GF.....-E-I.-..LS......-.....---QR....-.....-N........KD.--A--K-------L---   504
D.CY.06.CY163                ....--------N--............---..........---A--GF.....-E-I.-..-S......-.....---QR....-.....--........-------------L---   500
D.KE.11.DEMD11KE003          ....--------G--............---..........---A--LGF.....-E-I.-..-S......-.....---QK....G.....--........QC.----K-------W---   501
D.KR.04.04KBH8               ....------P---T............---..........---A--GF.....-E-I.-..-S......-.....-R-QK....-.....N-........----D--K-------L-K-   500
D.SN.90.SE365                ....------P----............---..........---T--GF.....-E-M.-..-SQKQEQA-.......---QK....-.....--........----A--K-------L---   507
D.TZ.01.A280                 ....-----------............---..........---A-I-GF.....-E-I.K..-S......-.....---QKDK..-.....--........---SA--K-------L---   497
D.UG.10.DEMD10UG004          ....-----------............---..........---A-C-GF.....-E--.N..-S......-.....Q--QK....-.....--........----A-------------   501
D.UG.11.DEMD11UG003          ....--------R--............---..........---A--GF.....-E-I.-..--......-.....---QK....-.....EK.DE......-------K----S--L---   501
D.YE.02.02YE516              ....-----------............---..........---A--GF.....-E-I.-..-S......-.....---QKDK..-.....-D........--------------L---   501
D.ZA.90.R1                   ....-----------............---..........---A--GF.....-E-I.-..-S......-.....---QK....-.....--........-------K----S--L---   501
F1.AO.06.AO_06_ANG32         ....-----------............---..........---A--GF.....RE-M.P..--......P.....---QK....E.....EGQGQG......--P--A--K------Q*---   500
F1.AR.02.ARE933              ....-----------............---..........---A--GF.....RE-I.-..-S......P.....---QK....-.....EG........--P--A--K-------*L--   495
F1.BR.10.10BR_RJ015          ....------I-N--............--S..........---A--GF.....RE-I.P..S-...RQE-.....---QK....E.....EG........--P--A--K------QWL--   502
F1.CY.08.CY222               ....-----------............--S..........---A--GF.....KE-I.A..-S......P.....E--QK....-.....-G........--P--A--K-------*L*-   493
F1.ES.02.ES_X845_4           ....-----------............---..........---A--GF.....RE--.-..-S......P.....---QK....E.....EGQG......--P--A--K------Q*---   498
F1.RO.96.BCI_R07             ....--------N--............---..........---A--GF.....RE--.-..-S......P.....---QK....-.....EG........M-P--A--K-------*---   495
F1.RU.08.D88_845             ....------P---L............--S..........---A--GF.....-E-I.-..-S......P.....R--QK....-.....-D........QCP----K------Q*---   501
F2.CM.02.02CM_0016BBY        ....------I----............---..........---A-G-GF.....-E-I.-..-S......P.....---QK....-.....-G........--P----K-------*---   493
F2.CM.10.DEMF210CM001        ....-----------............-QSRP....EPS---A--GF.....-E-I.-..-S......L.....---QK....-.....--........QLP-----------Q*---   503
F2.CM.10.DEMF210CM007        ....--------N--............---..........---A--GF.....-E-I.-..-S......P.....---QK....-.....--........PAY---A-K----S--*---   494
G.BE.96.DRCBL                ....--------N--............---..........---A-N-GF.....-E-I.A..-S......P.....---QK....E.....--........----S--K------Q*-K-   493
G.CM.10.DEMG10CM008          ....--------N--............---..........---A--GF.....-E-I.A..-N......-.....---QK....E.....--........----A--K----S--*-M-   499
G.CN.08.GX_2084_08           ....--------N-L............---..........---A--GF.....-EGI.-..-S......P.....----R....EQEPRE-G........-P.--A--K-------*---   501
G.CU.99.Cu74                 ....--------N--............--S..........---A--GF.....-Q-I.A..-S......P.....---QK....E.....-D...........-I--K----S--*---   497
G.ES.09.X2634_2              ....--------N--............---..........---A--GF.....RE-I.A..-S......L.....-P--R....EPEQRE--........----A--K----S--*---   502
G.GH.03.03GH175G             ....--------N--............--S..........---A--GF.....-R--.A..-S......P.....----R....E.......--E......-------K-------*---   499
G.KE.09.DEMG09KE001          ....--------N--............---..........---A--GL.....-E-I.A..-S......P.....--DQK....G.....R-..........P----K----S--*---   497
G.NG.09.09NG_SC62            ....------P-K-L............---..........---A--GF.....-E-I.A..-S......P.....----R....E.....--........SP.--A--K----S--*---   493
G.ZA.01.TV546                ....-----------............---..........---A--GF.....-E-I.A..-S......P.....R---K....E.....--........TP.--A--K----S--*---   497
H.BE.93.VI991                ....------P-K-L............---..........---A--GF.....-E-I.-..-S......P.....R--LK....E.....Q-..........P-----------Q*---   501
H.CF.90.056                  ....-----------............---..........---A--GF.....-E-M.-..-S......P.....---QLK...-.....--..........P--A------S--LL--   501
H.GB.00.00GBAC4001           ....-----------............---..........---A-C-GF.....-E-M.-..--......P.....---MK....E.....--..........P---------S--L---   500
J.CM.04.04CMU11421           ....-----------............---..........---A--GF.....-E--....XX......P.....---AK....-.....--........-------K---X--X---   497
J.SE.93.SE9280_7887          ....-----------............---..........---A--LGL.....-E-I....-S......P.....----K....-.....--........-------K----S--L---   498
J.SE.94.SE9173_7022          ....-----------............---..........---A--LGF.....-E-I....-S......P.....----K....-.....--........------------S--L---   498
K.CD.97.97ZR_EQTB11          ....E-------N--............---..........---A--GF.....-EKI.-..-S......L.....R--MK....-.....Q-........QGP----K----S--L---   496
K.CM.96.96CM_MP535           ....-----------............---..........---A--GF.....-E-I.-..-S......P.....R--TK....-.....--........QSP----K-------L---   496
U.CA.01.TV749                ....------P----EPTAPPFPQSSL---..........---A--LGK.....AEGI.-..SS......P.....L-D......-.....N.........---V----------L---   500
U.CA.99.TV721                ....------P----............A--..........---A--LQ-.....RP-P.-..A-......PAENLW-G-E-....I.....RQ........EQ.--V----------L---   499
U.CD.83.83CD003_Z3           ....--------N--............---..........---A--GF.....-E-.-...-S......P.....----R....-.....--...S.....--P----K----S---L--   500
U.CD.90.90CD121E12           ....--------N-T............---..........---A--GF.....-E-V.-T.-S......P.....---LR....-.....--........SP.--A--K----S--LL--   497
U.CY.05.CY090                ....-----------............--S..........---A--GF.....-E-I.-..SL......P.....R--QK....-.....--........QGP-----------S---L--   497
U.CY.08.CY223                ....------P----............---..........---A--LGF.....-E-I.AA.-S......P.....-P-QK....V.....--........ELY-----K----S-Q*-E-   502
U.ES.10.DEURF10DZ001         ....------P----............---..........---M--LGM.....-E-.-...SS......L.....---QG....-.....-G........--P----K----S--*---   495
U.GR.99.99GR303              ....------P----............---..........---A--GF.....-E-I.V..-S......P.....--G-K....-.....--........----A--K----S--*---   495
U.NL.95.U_NL_95_H10986_D1    ....-----------............---..........---A--GF.....-E-I.-..-S......P.....---MK....-.....--........QDP----K----S--W---   497
01_AE.AF.07.569M             ....------P----............---..........---AD.WGM.....-E-I.-..SL......L.....---QK....-.....--........HLP-TV--K-------L---   499
01_AE.CF.90.90CF11697        ....------P---L............---..........---A--LGM.....-E-I.-..SF......P.....---QK....-.....--........HPS--V--K-------L---   499
01_AE.CN.10.YNFL03           ....------P----............---..........---A-NWGM.....-E-I.P..SL......L.....---QT....-.....--........SPP----K-------LL--   498
01_AE.HK.04.HK001            ....------P----............--S..........---A-NW-E.......I.-..SL......L.....---QK....-.....--........HPS-SI--K-------L---   496
01_AE.IR.10.10IR.THR48F      ....------P----............--S..........---A-NWGM.....-E-I.A..SS.......................T-........HPP-SI--K-------L---   492
01_AE.JP.x.JRC77AE           ....------P----............--S..........---A-NWGM.....-E-M...................---QK....-.....--.........HP--V--K-------L---   494
01_AE.TH.09.AA090a_WG11      ....------P----............---..........---A-NWGM.....-E-V.-..S-......L.....R--QK....-.....--........HPP--V--K-------L---   497
01_AE.TH.90.CM240            ....------P----............---..........---A-NWGMGEEIT-E-I.-..SL......P.....---QK....-.....--........HPP--V--K-------L---   504
01_AE.US.05.306163_FL        ....------P----............---..........---A-NWGM.....-E-.-...SL......L.....---QK....-.....--.........RP-VV--K-------L---   494
01_AE.VN.98.98VNND15         ....------P---T............---..........---A-NWGM.....-E-I.-..SL......L.....---QK....-.....--........HPP-SI--K-------L---   498
```

```
                             p1 end  p6 start          Vpr binding                                                                   Vpr binding
                                                                                                                                   p6 end, Gag end
B.FR.83.HXB2                ....GRPGNFLQSRP............EPT..........APPEESFRS.....GVETTT..PP......Q.....KQEPI....D.....KE........LY.PLTSLRSLFGNDPSSQ*   500
02_AG.CM.10.DE00210CM013    ...------P----.............---...........--LA--LGM.....-E-R.P..SS......P........---K......-.....-G.........--P--A--K-------*L--   500
02_AG.CY.09.CY256           ...------P----.............---...........---A--GM......-E-.--..--......L......---TG......-.....-G.........--P-----K-------*---   495
02_AG.ES.06.P1423           ...------P----.............---...........---A--LGM.....-E-X.P..-S......P.....R---K......-.....--..........--P--A--T------Q*---   495
02_AG.GW.05.CC_0048         ...------P----.............---...........---A--..M.....-E-M.-..S-......P.....T---G......-.....-G.........--P-----K----S--*---   498
02_AG.KR.12.12MHR9          ...------P----.............---...........---A--LGM.....-E-I.-..S-......PKQEQ.RE--R......-.....-G.........--P-----K-------*---   499
02_AG.LR.x.POC44951         ...------P----.............---...........---A---GM.....-E-I.-..S-......P.........---R......-.....-G.........--P-----K-------*---   495
02_AG.NG.09.09NG_SC61       ...------P-N-L.............---...........---A--VGM.....RE-A.-..SS......-......-P--G......-.....-D.........--P-----K-------*---   493
02_AG.NG.x.IBNG             ...------P----.............---...........---A---GM.....-E-I.P..-S......P.....Q---R......-.....-G.........--P-----K-------*L--   495
02_AG.SE.94.SE7812          ...------P----.............---...........---A--LGI.....-E-I.-..SS......-.........---G......-.....-G.........--P--A--K-------*---   495
02_AG.SN.98.98SE_MP1211     ...------P----.............---...........---A---GM.....-E-I.-..SS......P.........---G......-.....-G.........--P-----K-------Y---   495
03_AB.RU.97.KAL153_2        ...------P----.............--S...........---A-N-GM.....-E-I.-..-S......L.........---QK.....-.....R-.........QHP-SI--K----D--L---   499
04_cpx.CY.94.94CY032_3      ...---------N--............---...........---A-CLER.....KE-.-..SS......L.........---R......-.....--.........--.-----K----S--L---   501
05_DF.BE.x.VI1310           ...------------............--S...........---A----F.....-E-I.A..SS......P.........---QK.....-.....EG.........--P--A--K-------L---   503
06_cpx.AU.96.BFP90          ...---------N--............---...........---I---GF.....-E-I.A..-S......P.........---SK....E.....--EKG.......--.--A--K----S--*---   500
07_BC.CN.98.98CN009         ...------------............---...........--------F.....-E----..-S......-.........---T-......-.....--.........--.-----K----------   496
08_BC.CN.97.97CNGX_6F       ...------------............---...........---A----F.....EE.-.-..-A......P.........---K......-.....R-.........-----------S--L---   490
09_cpx.GH.96.96GH2911       ...------P----.............--S...........---A---GM.....RE-.-..SS......X.........--XKD......-.....G-.........XPP--X--K-------L---   497
10_CD.TZ.96.96TZ_BF061      ...------------............---...........---A---GF.....-E-I.-..-S......-.........---QK.....-.....--.........PH.--A--K-------L---   501
11_cpx.CM.95.95CM_1816      ...------------............---...........---A---GF.....-E-V.A..-S......-.........----K.....-.....E-.........--.--A--K----S------   497
12_BF.AR.99.ARMA159         ...---------N--............---...........---A---GF.....-E-I.-..SS......P.........-R-QK.....-.....EG.........Q-P--A--K-------*-*-   494
13_cpx.CM.96.96CM_1849      ...-----------GPTAPPESRP...---...........---A---GF.....-E-I.A..-S......-.........----K....E.....--.........T-.--AA-K----S--*-*-   508
14_BG.ES.05.X1870           ...------P-N-L.............---...........---A---GF.....-E-I.A..-S......P.........----K....E.....--.........M-.--A--K----S--*---   497
15_01B.TH.99.99TH_MU2079    ...------P----.............--S...........---A-NWGM.....-E--.-..SL......L.........---QK.....-.....R-.........HTP--V--K-------LL--   498
16_A2D.KR.97.97KR004        ...------P---T.............---...........---A-D-GM.....-E--.-..-L......-.........---LK....N.....R-.........QHT-AI--K-------LL--   499
17_BF.AR.99.ARMA038         ...---------N--............---...........---L---GF.....-ETI.-..-Q......P.........----K.....-.....EG.........--P--A--K-------*-K-   501
18_cpx.CU.99.CU76           ...---------K--............---...........---A--LGF.....-E-I.A..-S......P.........---SK....E.....EK.........Y-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-I.A..-S......P.........--GQK....E.....-D.........--.--I--K----S--*---   497
21_A2D.KE.99.KER2003        ...-----------S............---...........---A-M-GF.....-E-I.-..-S......-.........---QK.....-.....--.........--.-SV--K-------LL--   499
22_01A1.CM.01.01CM_0001BBY  ...------P-G--.............---...........---A--VGM.....-E-I.-..SS......-.........---QR....E.....--.........QPP--V--K-------L---   500
23_BG.CU.03.CB118           ...------P-N-L.............---...........---A---GF.....-E-I.A..-S......P.........--G-K....E.....QD.........---I--K----S--*---   498
24_BG.ES.08.X2456_2         ...------P-N-L.............---...........---A--LGF.....-E-I.A..-S......P.........--GQP....E.....-D.........---I--K----S--*---   497
25_cpx.CM.02.1918LE         ...---------N--............---...........---A--LGF.....-E-I.-..-S......P.........---QK.....-.....--.........--.-MA--K----S--*---   499
26_AU.CD.02.02CD_MBTB047    ...------P---L.........V--..............---A-NLGM.....EEG-.-..-S......P.........--KEG....E.....--.........QSP--I--K----------   496
27_cpx.FR.04.04CD_FR_KZS    ...---------N--............---...........---A--LGI.....ME-M.-..SY......P.........---LK.....-.....--.........-P--V--K-------L---   497
28_BF.BR.99.BREPM12609      ...---------N--............---...........---A----F.....-E--.-..-S......P.........---QK....E.....EG.........--P--A--K-------*---   499
29_BF.BR.01.BREPM16704      ...---------N--............--S...........---A--LGF.....-E-R.-..-S......P.........---QK.....-.....EG.........K-P--A---------*---   510
31_BC.BR.04.04BR142         ...------------............---APPASRLEPT---A----F.....EE.-.-..-S......P.........---QR.....-.....R-.........--.-----K----S--L---   506
32_06A1.EE.01.EE0369        ...---------N--............---...........---A----F.....EE--AA..-S......S.........----K....K.....--.........--.-SA---------*---   495
33_01B.ID.07.JKT189_C       ...------------............---...........---A----F.....-EQ---..-S......-.........---Q-......-.....--..........LY-VA--K----------   497
34_01B.TH.99.OUR2478P       ...------P----.............---...........--------F.....-EG---..-S......-.........------.....-.....--..........LY-SA--K-------LL--   498
35_AD.AF.07.169H            ...------P-G--.............---...........---A---GF.....-E-I.K..-S......-.........---QK.....-.....G-.........--.-----K----------   497
36_cpx.CM.00.00CMNYU830     ...------P----.............---...........---A-N-GI.....-E-M.S..SS......-.........---SR.....-.....-G.........--P-----K-------*L--   498
37_cpx.CM.00.00CMNYU926     ...------P----.............---...........---A---EV.....KE-V.A..SS......L.........-----...G.....--.........--P--A--K----S--*P--   493
38_BF1.UY.03.UY03_3389      ...------------............--S...........---A---GF.....-E-I.-..----....PKQEQ.---QK.....-.....EG.........M-P--A--K-------*---   499
39_BF.BR.04.04BRRJ179       ...-----------T............---APT........---A--LMF.....-E--.-A.-S......-.........---R......-.....--.........--.---------------   502
40_BF.BR.05.05BRRJ055       ...---------N-L............---...........---A---GF.....-E-I.-..--......P.........---QK....N.....EG.........--P-----K-------...-   498
42_BF.LU.06.luBF_18_06      ...---------N--............--RT......EPT---A---GF.....-E-I.-..SS......L.........---QK.....-.....EG.........--P---------------   503
43_02G.SA.03.J11223         ...------------............---...........---A---GF.....-E-I.A..-S......-.........----R....E.....--.........SP.--A--K-------W-K-   498
44_BF.CL.00.CH80            ...------P-N-L.............---...........---A---GF.....-E-I.-..-S......P.........---QK.....-.....EG.........K-P-FA--K-------*-K-   501
45_cpx.FR.04.04FR_AUK       ...------P----.............--S...........---A---GM.....-E-I.-..SC......P.........---QK.....-.....R-.........QV..YP--K-------L---   499
46_BF.BR.07.07BR_FPS625     ...---------N--............---...........---A---EI.....RE-I.-..-S......-.........---QK.....-.....EG.........--P--A--K-------W---   496
47_BF.ES.08.P1942           ...---------N--............---...........---A--LGF.....-E--.-..-S......P.........---QK.....-.....EG.........--P--A---------*---   498
48_01B.MY.07.07MYKT021      ...------I----.............---...........---A----F.....-E----..-S......-.........------.....-.....--..........LY--A------S------   499
49_cpx.GM.03.N26677         ...---------N--............---...........---A---GF.....-E-I.A..-S......P.........---QK.....-.....E-.........--.-SA--K----S--*LP-   496
50_A1D.GB.10.12792          ...------------............---...........---A---GF.....-E-V.I..-S......-.........---QK.....-.....--.........--.-SA------------   498
51_01B.SG.11.11SG_HM021     ...------------............---...........---A----F.....-E----..-S......-.........------.....-.....--.........Q-.--A--K-----LL---   499
52_01B.MY.03.03MYKL018_1    ...------P---X.............---...........---A-NWGM.....-E-.X-..SS......L.........---QX.....-.....-D.........PPP-SV--K----S--LL--   500
53_01B.MY.11.11FIR164       ...---------N--............---...........--------F.....-EG---..-S......-.........------.....-.....--..........LY-----K----------   499
54_01B.MY.09.09MYSB023      ...---------N--............---...........--------L.....-E----..--......-.........------.....-.....--.........--.-VSA-K-------*---   501
55_01B.CN.10.HNCS102056     ...------P----.............---...........---A-DWGM.....-E-I.-..SS......P.........---QK.....-.....-G.........HRP-VV--K-------L---   499
56_cpx.FR.10.URF5_patient_A ...------P----.............---...........---A--VGM.....-E-A.-..YS......P.....R--SRDK......-.....--.........--P-----K----S--*---   500
57_BC.CN.09.09YNLX19sg      ...------------............---...........--------F.....EE.-.-..-A......P.........---LK.....-.....R-.........--.-----K-------L---   492
58_01B.MY.09.09MYPR37       ...------------............--X...........--------F.....-E----..-X......-........-G--......-.....--..........LY-----K----S------   498
59_01B.CN.09.09LNA423       ...------P----.............---...........---A-NWGM.....-E-S.....L......L.........---QK.....-.....--.........QPH-SI--K-------L---   492
60_BC.IT.11.BAV499          ...------------............---...........---A----F.....EE.-.-..--......P.........---QK.....-.....R-.........--.-----K----S--L---   491
61_BC.CN.10.JL100010        ...------------............---...........---A----F.....EE.-.A..SA......P.........---MK.....-.....R-.........-----------S--L---   492
62_BC.CN.10.YNFL13          ...------------............---...........---A----F.....EE.-.-..-A......P.........---SK.....-.....R-.........-----------S--L---   488
63_02A1.RU.10.10RU6637      ...------P----.............---...........---A--VGM.....-E-I.-..SS......P.....Q---R......-.....-G.........--P--A--K-------*---   495
64_BC.CN.09.YNFL31          ...------------............--S...........---A----F.....-E----..-S......-.........-P-GT.....-.....--.........--.--A--K-------LC--   496
65_cpx.CN.10.YNFL01         ...-----------L............---...........---A----F.....EE.-.-..-A......P.....R--SK......-.....RG.........--.-----K-------L---   495
67_01B.CN.11.2011.ANHUI.HF115..------P----.............---...........---A-DWGM.....-G-I.-..FS......P.........---QK.....-.....--..........H-SV--K-------L---   497
68_01B.CN.11.2011.ANHUI.WH73...------P----.............---...........---A-DWGM.....-G-I.-..SS......P.........---QR.....-.....R-..........Q--V--K-------L---   497
70_BF1.BR.10.10BR_PE004     .......---R--I-N--.........---...........---A----F.....-E-M.-.......--.PK.......--EQK.....-.....EG.........R-P-----K-------W---   504
71_BF1.BR.10.10BR_PE008     ...---------N-T............---...........---A---GF.....-E-M.-..-S......-.........---QK....S.....EG.........--P--A--K-------W-K-   499
72_BF1.BR.10.10BR_MG002     ...G--------N-T............---...........---A---GF.....-E-K.-..-S......P.........----K....E.....EG.........--P--A--K-------WL--   501
O.BE.87.ANT70               ....T----YV-RPA............H-S...........---M-EEVK.....-Q-NQ..............E-KGG....P.....N-..........--.-FA--K----T-Q...-   499
O.CM.98.98CMA104            ....T-----M-KQV............A-S...........---M-EVVK.....-Q-NQ..............E-KED....S.....N-..........--.-FA--K----T-Q...-   499
O.CM.98.98CMABB141          ....A----YM-KQV............P-S...........---M-EMMK.....-Q-NQ..............E-KGD....T.....T-..........--.-FA--K----T-Q...-   499
O.CM.98.98CMABB212          ....T----Y--KQV............S--...........---X-EIVK.....-Q-NQ..............S-GXX....Q.....X-..........--.-FA--K----T-Q...-   499
O.CM.98.98CMU5337           ....T----YA-RQV............S--...........---M-EEVK.....RQ-NQ..............E-KGD....Q.....N-..........--.-FA--K----T-Q...-   499
O.CM.99.99CMU4122           ....T----YA-RQV............S-S...........---M-EE-K.....-P-SQ..............E-KGD....P.....N-..........--.-FA------T-Q...-   499
O.FR.92.VAU                 ....T----YV-EQV............S-S...........---M-EMMK.....-Q-EDQ.............N-KGE....Q.....N-..........--.-FA------T-Q...-   499
O.GA.11.11Gab6352           ....T----YV-KQV............H-S...........---M-EEVQ.....KQ-NQ..............E-KGD....P.....N-..........--.-FA--K----T-Q...-   498
O.SN.99.99SE_MP1299         ....T----YA-RQV............S-S...........---MTEEMK.....-Q-NQ..............E-KED....Q.....N-..........--.-FA------T-Q...-   498
O.US.10.LTNP                ....T----YV-KQV............A-S...........---M-EEVN.....RQ-NQ..............E-KGS....P.....T-..........--.-FA------T-Q...-   499

N.CM.02.DJ00131             ...------P-TTTTRK..........---...........---L-NYG-.....QE-R.S..TQ......G.....-ERQE....N..........TENS--P-----------L----   513
N.CM.02.SJGddd              ...------P-TTARK...........---...........---L-..........E-R.S..TQ......G.....-EMQE....N.....Q-...KTENS--P-----------L----   508
N.CM.04.04CM_1015_04        ...------P-TATRK...........---...........---L--YG-.....QE-G.S..TQ......G.....-EMQE....N.....Q-...KTENS--P-----------L----   512
N.CM.06.U14296              ...------P-TTTKK...........---...........---L-NYGF.....QE-R.S..TQ......G.....-EMQE....N.....Q-...RTENS--P------------L---   512
N.CM.06.U14842              ...------P-TTTTRK..........---...........---L-DYGC.....QE-R.N..TQ......R.....-EMQE....-.....Q-...KTENS--P----------------   513
N.FR.11.N1_FR_2011          ...------P-TTTRK...........---...........---L-NYGF.....QE-K.S..TQ......G.....-EMQE....-.....Q-...RTENS--P----------------   513

P.CM.06.U14788              ....K----YV-KQV............Q--...........---M-EEEM.........-..........-NKQ..XE-KE....-.....EK...E.....--.--A--K----T-Q*Y--   495
P.FR.09.RBF168              ....K----YV-KQV............Q--...........---M-EDEM.........-..........RDKQ..REXEE....X.....E-.........E-.---------T-Q*Y--   497

CPZ.CD.06.BF1167            ....C-----V-NQSQTQNQEQ...................SQTQ-KLVE.....KA-..P..SAP.....P.....ME-VNHW..NS....-D..........--GF--K-I--S-Q*WK-   531
CPZ.CM.05.SIVcpzMT145       ...-----YV-G-K.............---...........---M---GF.....QE-K.K..VK......E.....EVREE....S.....-P.........PS.-----K----------   492
CPZ.GA.88.GAB1              ...------V-N-T.............---...........---I--YGY.....QE-E.K..SQ......-.....EKKEG....E.....SS.........--.-P--K----S------   509
CPZ.TZ.06.SIVcpzTAN13       ....C-----V-NN-............-KE...........K-Q-QEG.......VQ.-.VPTA-......PM....EMG-M....K.....EG.........-..F--K-I--S-Q...-   525
CPZ.US.85.US_Marilyn        SGGSK------EN-K............---...........---I-D-GY.....QE--.V..TQ......E.....--GKE....-.....--.........PF.Q---K----S----E-   506

GOR.CM.07.SIVgor2139_287    ....E----YV-KQV............Q--...........---M-EGEK.......................E.....MRK-E....K.....E-.........--.-----K----T-Q*YL-   492
GOR.CM.07.SIVgorCP2135con   ....E----YV-KQV............Q--...........---M-EGEK.......................E.....MRK-E....K.....E-.........--.-----K----T-Q*YL-   492
```

```
                             Pol start
                             Gag-Pol TF start                                                                                                   Gag-Pol TF end  protease start
B.FR.83.HXB2                 FFREDLAF.LQ....GKAREFSSEQT............RAN..........SPTRRELQV.....WGRDNN..SP......S.....EAGAD....R.....QG........TV.SFNFPQVTLWQRPLVTIKIGGQLKEALLDTGADDTVL.EEMSLPGRWKPKMIGGI  106
A1.AU.03.PS1044_Day0         ----N---.Q-....-E--K-P----............---..........---SGA-WD.....G---.-..L-......-.....----E....G.....--........-GPTLS---I------I--VR-E---R-------------.-DID---K---R-----  106
A1.CH.03.HIV_CH_BID_V3538    ----N---.Q-....-E---------............G--..........---S---WG.....E--N.-..LL......D.....----E....G.....--........-AP--S---I------V--V--------------------.--IN---K---------  106
A1.CY.08.CY236               ----N---.Q-....-E--K--P---............--I..........---S-D-WD.....G---.-..L-......-.....----E....-.....--..........P-LS---I---------VE-----R-------------.-DID-------R-----  103
A1.ES.05.X1608_8             ----N---.Q-....-E--K-PT---............--I..........---S-X-WD.....G---.D..LT......-.....----EAGAGK.....--........-GPTL----I------I--VR-------------------.-DIN---K---------  110
A1.KE.11.DEMA11KE001         ----N---.Q-....RE-----T---............G--..........---S-D-WD.....G---.S..L-......-.....-T--E....G.....--........-GSTL----I------V--VR-------------------.-DIE---K---R-----  106
A1.RU.11.11RU6950            ----N---.Q-....RE--K------............--I..........---SGK-WD.....G---.-..PL......P.....-T--E....G.....--........-AS------I--------AVR-------------------.-DID---K---------  106
A1.RW.11.DEMA111RW002        ----N---.Q-....-E--K--P---............--I..........---SGT-GD.....G---.S..L-......-.....-T--E....-.....--........-GPT-S---I---------V--E---R-------------.-NIN-T-K---------  106
A1.SN.01.DDI579              ----N---.Q-....-E--K------............---..........---S-SPGD.....GR--.-..LL......P.....----E....G.....--........-.A--S---I------I--VR-------------------.-DIN---K---------  105
A1.UG.11.DEMA110UG001        ----N---.Q-....RE--K---K--............---..........---G-A-WD.....G---.-..L-......-.....-S--E....-.....--........-GPTLC---I---------V--------------------.-DIN---K---------  106
A1.ZA.04.04ZASK162B1         ----N---.Q-....-E---------............G--..........---S---WD.....G---.-..L-......T.....-E---....-.....--........A.A-LS---I-----------V------------------.-DIN---K---------  105
A2.CM.01.01CM_1445MV         ----N--S.Q-....RE--KL-P--N............---..........--AS--RWN.....G---.S..FL......V.....---TE....K.....--........-TH-Y----I---------V--E-----------------.-DIN---K-R-------  106
A2.CY.94.94CY017_41          ----N---.Q-....RE--K-----N............---..........---S---EN.....G---.-..LL......P.....---TG....D.....--........-IQ-C----I---------V--E-----------------.--IN---K---------  106
B.BR.10.10BR_MG029           --------.P-....-----L-P---............---..........---------.....--G---..--......-.....-----....-.....--........--.-LS---I---------V--------------------.---N-------------  106
B.CA.07.502_1191_03          ----N---.P-....-E--K------............--I..........---------.....--G---..--......-.....-----....-.....--........N-.-LL---I------V-----------------------.-DIN---K---------  106
B.CH.08.M2_0803101_NFLG8     ----N---.P-....------FP-P-............---..........----G----.....---N--..PQ..SEAGA.....---TE....N.....R-........E-.--S---I-----VL--V----VR------------II.-D-T---K----IV--V  110
B.CN.10.DEMB10CN002          ----N---.P-....----------D............--I..........---------.....--G--D..-S......-.....-----....-.....--........--.PLI---I------I--V------VK------------.-DIN-------------  106
B.ES.10.DEMB10ES002          ----N---.P-....-E--K------............---..........---------.....--GE-S..-T......-.....-----....-.....--........--.-L----I------I--V-V------------------.---N-------------  106
B.FR.11.DEMB11FR001          ----N---.-.....---G--P----............---..........---------.....--G---..-L......-.....--R--....-.....R-........K-.---L--I------------------------------.---N-------------  105
B.GB.05.MM45d213_GN1         --------.P-....-----------............-T-..........----G----.....------..-L......P.....-----....-.....--........--.--S---I-----------V------------------.---N---K---------  105
B.HK.06.HK003                ----N--L.P-....-----------............---..........----G----.....------..--......-EAG..T--T-....-.....--........P-.---L--I---------V-V------------------E--IN---K---------  110
B.HT.05.05HT_129389          --------.-.....-----LP---D............--I..........----G----.....----.D..--......-.....---VE....G.....--........--.--S---I------------------------------.---N---K---------  105
B.JP.12.DEMB12JP001          --------.P-....-----LP---A............---SPT.......---------.....--K-RD..-L......-.....-T---....-.....--........--.--S---I-----------V------------------.-D-N---K---------  109
B.KR.07.HP_18_07JHS10_3909   ----N---.P-....-----------............--I..........--S-G----.....--G---..-L......-.....---T-....-.....--........--.--S---I------I-----------------------.-----------------  106
B.PE.07.502_0525_wg5         ----N---.P-....-E--KL----DSEQA........---..........----G--I-.....C-G---..PL......-.....-----....-.....--........--.--S---I------------------------------.---N-------------  110
B.RU.11.11RU21n              --------.P-....-E----P----............---..........----G-F--.....--G---..-L......-.....----E....QGQG...--........A-.-LS---I----------R------------------V.-D-Y-------------  110
B.TH.08.MERLBDTRC10          ----N--L.P-....-----KL----............---..........--S------.....------..--......-.....-T-TN....-.....--........-I.-LS---I--------A-------R-------------.---N---K---------  106
B.US.11.ES38                 ----N---.P-....-E--KLP----REQT........--I..........----G----.....----S-..-L......-.....-T---....-.....--........P-.--S---I---------V-------------------.---N-------------  110
C.AR.01.ARG4006              ----N---.Q-....----K-XP---............---..........---S-----.....R-.-.-..PL......-.....----E....-.....--...........TL----I-----------V------------------.---K---N---------  102
C.BR.07.DEMC07BR003          ----N---.P-....-E--K-PP---............---..........---S-----.....R-.-.-..PL......-.....----E....G.....--...........TL-L--I------V--V-V------------------.--TK---N---------  102
C.BW.00.00BW5031_1           ----N---.Q-....-E---LPP---............-T-SPT....NAN---S-----.....R-.-.-..PR......-.....---DE....-.....E-...........TL----I--------S--V---I--------------.-DIN---K---------  108
C.CN.10.YNFL19               ----N---.P-....---G------N............---..........-S-S-----.....------..-I......-.....-V---....-.....--........-I.PL----I---------V-V---I--------------.--I----K---------  106
C.CY.09.CY260                ----N---.P-....-E----P---A............---SPTDS..GAN---S-----.....R-E-.-..PR......-.....-----....-.....--............L----I--------S--V---I--------------.-DIN---K---------  109
C.ES.08.X2363_2              ----N---.P-....-E-G-LP----............---SPTS...RTN---S-----.....R-.--S..-R......-.....---VE....-.....-R...........DL----I----------RV---I----X---------.-DIN---K---------  109
C.ET.02.02ET_288             ----T---.P-....-E---P-----............S--SPT...RANG---S-----.....R-.G.-..PS......-.....----E....-.....--...........-L----I----------R-------------------.--IN---K---------  108
C.IN.09.T125_2139            ----N---.P-....-E---PP----............---..........---S-----.....R-.--P..PS......-.....----E....-.....--...........NL----I--------S-RV---I--------------.--IN---K---------  103
C.KE.00.KER2010              ----N---.P-....-E----P---A............---..........---S-----.....R-.--S..PC......-.....----E....-.....E-...........TL----I---------V-----IR------------.--IN---K---------  102
C.MW.09.703010256_CH256.w96  ----N---.P-....-E----P---A............---SPTT...RTN---S---H..............PR......-.....----E....-.....--...........TL----I----------RV---I--------------.--L----K---------  104
C.TZ.08.707010457_CH457.w8   ----N---.P-....-E-GK-P---N............---..........---S---R-.....R-.-.-..P-......-.....----E....K.....--...........NL----I--------S--V---I--------------.-DIN---K---------  102
C.YE.02.02YE511              ----N---.P-....------P----............---..........---I-----.....R-.N.-..PR......-.....K--GE....-.....--...........TL----I------V----V---I--------------.--IN---K---------  100
C.ZA.08.705010534_CH534.w12  ----N---.P-....-E----P---A............-T-SPTS...RAN---S--F--.....R-.-.-..PC......-.....---DE....G.....--...........T-----I--------S------I--------------.--IN---K---------  109
C.ZA.10.DEMC10ZA001          ----N---.P-....------P----............---..........---S-----.....R-.-.-..PC......-.....---VE....-.....--...........TL-C--I--------S------I--------------.--I----K---------  102
C.ZM.11.DEMC11ZM006          ----T---.P-....-E---LPP---............---SSP....SAN---S-----.....R-.-.-..PR......-.....----E....-.....--...........-Y----I--------A--V---T--------------.--IN---K---R-----  107
D.CM.10.DEMD10CM009          --------.Q-....---GKL-----RANSPT......---..........---S---R-.....------..PL......-.....-T-TE....G.....-K........QG.-LS---I------F--V--------------------.-DID---K---------  111
D.CY.06.CY163                ------V-.P-....-----------............---..........---S---R-.....--G-.-..PL......-.....-T--E....G.....-R........--.-L----I--------S--V------------------.-DIN---K---------  105
D.KE.11.DEMD11KE003          --------.PE....---G-LP-G-A............-TH..........---S--PW-.....--G-.-..TL......-.....-T--E....-.....--........A-.-S----I------------E-----------------.--IN---K---------  105
D.KR.04.04KBH8               --------.P-....W---KLP---N............---..........R--S---R-.....--G-.-..PL......-.....-T--E....-.....-R........--.-LG---I--------KV-----I--------------.-DIE---K---------  105
D.SN.90.SE365                ----N---.Q-....-E---L-----............---..........---N---R-.....--G-.D..PLSETGASA.....-T--E....-.....--........--.--S---I------V-AV--------------------.-DIN-S-K---------  111
D.TZ.01.A280                 ----N---.P-....R----LP----............---..........---S-D-R-.....--G-.K..TL......-.....-T--ERQ..G.....--........--.--S---I------------------------------.-------K---------  107
D.UG.10.DEMD10UG004          ----I---.PE....-----LP----............---..........---S-V-R-.....--G-.K..PL......-.....AT--E....G.....--........A-.-LSL--I---------V--------------------.--IN---K---------  105
D.UG.11.DEMD11UG003          ----Y---.P-....-----LP--KA............---..........---S---R-.....------..T-......-.....-T--E....-.....RE.RR.....A-.-L-L--I---------V--------------------.---N---K---------  107
D.YE.02.02YE516              ----N---.Q-....-E-GKL-----............---..........C--S---R-.....------..PL......-.....-T--EGQ..-.....--........S-.-S----I----------N----Q--------------.-DID---K---------  107
D.ZA.90.R1                   ----N---.P-....---G-------............---..........---S---R-.....--G-.-..PL......-.....-T-TE....-.....--........A-.--DL--I---E--------------------------.---N---K---------  105
F1.AO.06.AO_06_ANG32         ----N---.P-....-E-GK------............---..........---S---R-.....Q-G-.A..--......-.....----E....G.....R-TGTG....--S-LS---I------V--V-VE-----------------.-DIN---K---------  110
F1.AR.02.ARE933              ----N---.Q-....-E--K-P----............---..........--AS---R-.....QRG-.-..PL......-.....----E....-.....R-........--P-LS---I--------A-N-------------------.-DIN---K---------  106
F1.BR.10.10BR_RJ015          ----N---.P-....---WKLHP--A............-TI..........--AS---R-.....Q-G-.P..IS..EAGA......---TE....-.....R-........--P-LS---I------V--V-V--E-R-------------.-DIN---K---------  109
F1.CY.08.CY222               ----N---.Q-....-E--KL-----............--I..........--AS---R-.....QRG-.S..PL......-.....G---E....G.....--........A-S-LSL--I------V-VV-V-----------------.-DIN---K-R-------  106
F1.ES.02.ES_X845_4           ----N---.P-....-E-GK------............---..........---S---RI.....QRGN.D..PL......-.....----E....G.....G-TG......--S-LSL--I------V----VE-----------------.-DIN-S-K---------  108
F1.RO.96.BCI_R07             ----N---.P-....-E--KL-----............---..........-SA----W-.....Q-GN.-..-L......-.....----E....G.....R-........N-S-LSL--I--------S--V------------------.-DIN---K---------  106
F1.RU.08.D88_845             ----N---.Q-....RE------X--............--I..........--AGG--R-.....R-G-.-..PL......-.....----E....-.....--........S-P-LD---I------V--V-V----R-------------.-DIN---K---------  106
F2.CM.02.02CM_0016BBY        ----N---.Q-....RE-W--H----............---..........G-AS-G-R-.....RR--.-..-L......P.....----E....-.....--........--S-LD---I------V----VE---R-------------.-DIN-S-K---R-----  106
F2.CM.10.DEMF210CM001        ----N---.Q-....-E-G-------............--EQT.....RAF--AS---W-.....R-G-.-..PL......P.....----E....-.....--........-AS-L----I------I--V--E-H-R------------.-DIN-T-K---R-----  111
F2.CM.10.DEMF210CM007        ----T--S.Q-....RE---------............---..........--AS---R-.....-RG-.-..-L......P.....----E....-.....--........AGV-L-C--I-----------V----R-------------.-DIN---K---------  105
G.BE.96.DRCBL                ----N---.Q-....-E----P---A............---..........-------R-.....R-G-.S..PL......P.....----E....G.....K-........-I.-SI---I------I-KVR-----I-------------.--ID---K---------  105
G.CM.10.DEMG10CM008          ----N---.Q-....-----------............---..........-------W-.....R-G-.S..PQ......P.....-T--E....G.....K-........A-.-LSL--I--------NVRV----I-------------.-DIN---K---------  105
G.CN.08.GX_2084_08           ----N---.Q-....-E-G--P---A............---..........-------R-.....RRG-.-..PL......P.....-T--E....GTGAEGE-........VT.--SL--I---------VR-----I-------------.-QID---K---------  110
G.CU.99.Cu74                 ----T---.Q-....RE-G----K-A............--I..........---C---R-.....RTG-.S..PL......P.....--R-E....G.....E-............L-L--I---------VR-----I-------------.-DIN---K---------  102
G.ES.09.X2634_2              ----N---.Q-....-------P--A............---..........---S---R-.....QRG-.S..PL......P.....----E....GAGAEGK-........VI.PLSL--I---------V------I-------------.-DIN---K---------  110
G.GH.03.03GH175G             ----N---.Q-....-E--------A............--I..........-S-S---R-.....REG-.S..PL......P.....----E....G.....K-G.......-I.-L----I---------VR-----I-------------.--IN---K---------  106
G.KE.09.DEMG09KE001          ----N---.Q-....R------P--A............---..........---C---RA.....RRG-.S..PL......P.....--RPE....G.....E-..........A-L----I---------V------I-------------.--I----K---------  104
G.NG.09.09NG_SC62            ----N--S.P-....WE-------EA............---..........---S---R-.....R-G-.S..PL......P.....----E....G.....K-........VT.-LSL--I---------VR-----I-------------.-QID---K---------  105
G.ZA.01.TV546                ----N---.P-....-E--------A............---..........---S---R-.....R-G-.S..PL......P.....---TE....G.....K-........DT.--S---I--------AVRVE---I-------------.-DLN---K---------  105
H.BE.93.VI991                ----N---.Q-....------PP-EA............---..........---S---R-.....RRG-.H..PL......-.....----E....-.....T-..........T------I------I--V--E-----------------.-DIN---K---------  104
H.CF.90.056                  ----N---.Q-....RE--K--P--A............-T-..........---S---R-.....RRG-.D..PL......-.....---AE....G.....--..........T-LS---I---------V--E---R-------------.--IN---K---------  105
H.GB.00.00GBAC4001           ----N---.Q-....RE---L-P--A............-T-..........---S-M-R-.....R-G-.D..T-......-.....---DE....G.....K-..........T-L----I---------V--E---R-------------.--IN---K---------  104
J.CM.04.04CMU11421           ----X---.Q-....XE---L-X--X............---..........---S---R-.....R-----..XX......P.....-TRSE.....................-.-S----I----------------R------------I.--ID---X---------  102
J.SE.93.SE9280_7887          --------.Q-....RE--L-P----............---..........---S--PRA.....RRG--...PL......P.....-T--E....G.....--.........-.-S----I----------R-----R-------------.-DID--RK---------  104
J.SE.94.SE9173_7022          --------.Q-....RE----P----............---..........---S--PR-.....RRG--...PL......P.....-T--E....G.....--.........-.-S----I----------R-----R-------------.--ID---K---------  104
K.CD.97.97ZR_EQTB11          ----V--S.Q-....RE--K------............---..........---S---W-.....R-E-.-..PL......-.....-T-NE....-.....S-........-GS------I------V--V-V----R-------------.--IN---K---------  106
K.CM.96.96CM_MP535           ----N---.P-....-E---------............---..........---S---R-.....R-G-.-..PL......-.....---DQ....-.....--........-EP------I------I----V----R-------------.--IN---K---------  106
U.CA.01.TV749                ----N---.P-....-E---LPT---RANSPTIPSEQP---..........---S--PGE.....G---.-..LL......P.....AT-.............Q...........--S---IS--K-----VS-----I-------------.-NIN---K---------  111
U.CA.99.TV721                ----N---.P-....-----L-----............S--..........---S--PSE.....QT-A.-..--......TSREPL-G-R-....N.....-A........GT.A-S---IS-----I--VN-----I-------------.-NIN---K-R-------  110
U.CD.83.83CD003_Z3           ----N---.Q-....-E---------............---..........---S---R-.....RRG-.-..PL......A.....----E....G.....--..V.....PLP------I---------V-------------------V.---T---K---------  107
U.CD.90.90CD121E12           ----N---.Q-....----------N............---..........-S-S---R-.....RRGS.H..PL......A.....----E....G.....--........VS.--S---I---------V-------------------V.---T---K---------  106
U.CY.05.CY090                ----N---.PE....WE--KLPP---............--I..........---S---RI.....RRG-.-..L-......-.....-T--E....-.....--........-GS------I-----------VE---R-------------.KNIN-S-K---------  106
U.CY.08.CY223                ----N---.Q-....-E---L-----............-T-..........---S--SG-.....RRG-.SR.PL......P.....-T-TE....G.....--........RII-LD---I------I-RVT-E---I-------------.-NIN---K---------  107
U.ES.10.DEURF10DZ001         ----N---.Q-....RE--K------............-T-..........---NG--GD.....G---.-..LL......-.....---TG....G.....--........-IP-L----I---------VR-E---I-------------.-DIN---K---------  106
U.GR.99.99GR303              --------.Q-....-----------............--D..........---S---R-.....RRG-.S..PL......P.....----K....G.....--........AI.--S---I---------VR-----I-------------.-GI----K------R--  105
U.NL.95.U_NL_95_H10986_D1    ----S---.Q-....RE----P----............---..........---SG--R-.....RRG-.-..PL......-.....---DE....-.....--........-GS------I------M--VN----V--------------.-DIN---K-R-R-----  106
01_AE.AF.07.569M             ----N---HQ-....---GK------............---..........--AS-.-GD.....G---.-..LI......T.....-T-TE....-.....--........-SS-YS---I------I-PV-----I--------------.-DIN---K---------  106
01_AE.CF.90.90CF11697        ----N--S.Q-....-E---------............---..........---S---GD.....G---.-..LL......P.....----E....-.....--........-PF--S---I-F-------V--E-----------------.-DIN-------------  106
01_AE.CN.10.YNFL03           ----N---.-.....----K------............---..........---S-K-GD.....G---.T..IL......T.....-----....-.....--........KSS------I------------------------------.-DIN---K---------  105
01_AE.HK.04.HK001            ----N--S.Q-....-E---------............--F..........---SGK-EG..........-..LL......T.....----E....G.....--........-PF---L--I---------V-------------------.-DIN---K---------  103
01_AE.IR.10.10IR.THR48F      ----N---.-.....R----------............G-I..........--AS-K-GD.....G---.S..LL...........................D-........ASS------I---------V--E---R-------------.-DIN---K---------   98
01_AE.JP.x.JRC77AE           ---Q----.P-....---GKL-----............-TI..........---S-K-GN.....G--...................----E....-.....--.........TS--S---I--------AV------------------V.-DIN---K---------  101
01_AE.TH.09.AA090a_WG11      ----NM--.P-....---G-------............---..........---S---GD.....G--S.-..L-......T.....----E....-.....--........-PS-LS---I---------V------R-------------.-DIN---K---------  106
01_AE.TH.90.CM240            ----N---.Q-....---G-------............---..........---S-K-GDGGRDNG---.-..LL......T.....----E....-.....--........-SS--S---I---------V-------------------.-DIN---K---------  111
01_AE.US.05.306163_FL        -------S.Q-....-E--------A............---..........---S---GD.....G---.-..LL......T.....----E....-.....--.........TS-CS---I---------VR-------------------.-DIN---K---------  104
01_AE.VN.98.98VNND15         ----N---.-.....---G------N............---..........---S-K-GD.....G---.-..LL......T.....----E....-.....--........-SS------I---------V-------------------.-DIN---K---------  105
```

```
                              Pol start
                              Gag-Pol TF start                                                                                         Gag-Pol TF end protease start
B.FR.83.HXB2                  FFREDLAF.LQ....GKAREFSSEQT............RAN..........SPTRRELQV.....WGRDNN..SP......S.....EAGAD....R.....QG........TV.SFNFPQVTLWQRPLVTIKIGGQLKEALLDTGADDTVL.EEMSLPGRWKPKMIGGI 106
02_AG.CM.10.DE00210CM013      ----N---.Q-....-E--------A............-T-..........---S--PWD.....G--E.T..LL......-.....----E....G.....--........IIS--S---I---------VR-----I-------------.--IN---K--------- 106
02_AG.CY.09.CY256             ----N---.Q-....-E--KL-P---............-T-..........---SG--WD.....G-----..--......-.....---DG....G.....--........-IS-L----I------II-VR----PI-------------.-DID---K--------- 106
02_AG.ES.06.P1423             ----NX--.Q-....-E--Q------............GT-..........--AS---GD.....G--X.T..PL......P.....---TK....G.....--........-IS-LS--HI------V--V------X-------------.--ID------------- 106
02_AG.GW.05.CC_0048           ----N---.Q-....-E---------............GT-..........---S-K..H.....G---.D..L-......-.....NT-TG....G.....--........--S------I---------VR-E---I-------------.--IN------------- 104
02_AG.KR.12.12MHR9            ----N---.Q-....-E---------............GT-..........---S---RD.....G---.-..L-......-EAGA.-G--E....G.....--........-IS-L----I---------VR-----I-------------.--IN---K-R------- 110
02_AG.LR.x.POC44951           ----N---.Q-....RE--KL-----............GT-..........---G---WD.....G---.-..L-......-.....---TE....G.....--........-IS------I---------VR-----I-------------.-DIE------------- 106
02_AG.NG.09.09NG_SC61         ----N---.Q-....-E--K------............GT-..........---S--CWD.....E--G.-..LL......-.....---TR....G.....--........SLP---L--I---------VN-----I-------------.--I-------------- 106
02_AG.NG.x.IBNG               ----N---.Q-....-E--K------............GT-..........-S-S---WD.....G---.T..-L......-.....T--TE....G.....--........AIS------I---------VR-E---I-------------.-DIN---K--------- 106
02_AG.SE.94.SE7812            ----N---.Q-....-E--K------............GT-..........--AS--PWD.....R---.-..LL......-.....---TG....G.....--........-IS-LS---I---------V-----I-------------.--IN---K--------- 106
02_AG.SN.98.98SE_MP1211       ----N---.Q-....-E--KL-----............GT-..........---S---WD.....G---.-..LL......P.....---TG....G.....--........-IP------I-------F-VR-E---IK-------N----.-KIN---K--------- 106
03_AB.RU.97.KAL153_2          ----N---.Q-....RE--K------............--I..........---S-K-WD.....G---.-..PL......P.....-T-TE....G.....--........-AS------I---R-----VR--------------------.-DIN---K--------- 106
04_cpx.CY.94.94CY032_3        ----NV--.Q-....RE--K-----A............---..........--A-GM-RE.....ERG-.-..LL......-.....---TE....G.....--........-I.------I-----------L---IR-------------.--IN---K--------- 105
05_DF.BE.x.VI1310             ----S---.P-....-E---LPP---............G-L..........--AS-----.....--G-.S..LL......-.....----E....G.....R-........--P-LS---I------V---R-------------------.-DIN---K--------- 106
06_cpx.AU.96.BFP90            ----N---.Q-....-E-------A............---..........---H---RF.....RRG-.S..PL......P.....-T-VE....G.....E-GKG.....AI.-LSL--I---------VR-----I-------------.-DIN---K--------- 108
07_BC.CN.98.98CN009           ----N--L.P-....-----------............---..........----G----.....------..-I......-.....---DN....-.....--........-I.------I------------------------------.-D-N---K--------- 106
08_BC.CN.97.97CNGX_6F         ----I---.P-....-E----PP---............---..........---S-----.....R-.-.-..PS......-.....---TE....-.....--........ ...TL----I--------S--V---I--------------.--VN---K--------- 102
09_cpx.GH.96.96GH2911         ----N---.Q-....RE---L-X---............G-I..........G-XS---WD.....E---.X..LL......-.....XXXEG..........R-........-AP--X---I------I--V--------------------.--IN---K--------- 105
10_CD.TZ.96.96TZ_BF061        ----N---.Q-....R----LP----............---..........-S-S---R-.....--G-.-..TL......-.....-T--E....-.....--........AA.-LS---I------------------------------.---N---K-----M--- 105
11_cpx.CM.95.95CM_1816        ----N---.Q-....-E----P--A............---..........---S---R-.....RRGG.S..PL......P.....-T--E....-.....G-........AI.--S---I------I--V--A---R-------------.--IN---K--------- 104
12_BF.AR.99.ARMA159           ----N---.Q-....-E--K-P---A............---..........--AS---W-.....RRG-.-..LL......-.....----E....-.....R-........--P-LS---I--------I--V------------------.-DIN---K--------- 106
13_cpx.CM.96.96CM_1849        ----N---.Q-....RE---------RTNSPTREQT..---..........---S---R-.....RRG-.S..PL......P.....---TE....G...............DI.-LSC--I------I-IV---E--I------------.K-IN------------- 113
14_BG.ES.05.X1870             ----N---.Q-....-E-----P--A............---..........-------W-.....RRG-.S..PL......P.....--R-E....G.....K-........D-.PLSL--I---------V-----I-------------.-DIN---K--------- 105
15_01B.TH.99.99TH_MU2079      ----T---.--....-----------............--I..........---S---GD.....G----...LL......T.....----E....-.....--........-YS--S---I---------------------P----P--.-DINF--K-N------- 105
16_A2D.KR.97.97KR004          ----N---.P-....RE--------N............-T-..........---S-G-WN.....G-G-.-..PL......A.....----E....K.....--........-TH-C----I---------V--E---R-------------.-DIN---K--------- 106
17_BF.AR.99.ARMA038           ----N---.Q-....-E--K-P---A............---..........--AI---R-.....R-D-.-..P-......A.....----E....-.....R-........--P-LS---I--------E--V------------------.-DIN---K-R------- 106
18_cpx.CU.99.CU76             ----N---.Q-....-E-----P-KA............---..........---SG--G-.....RRG-.S..PL......P.....--RVE....G.....GE..........I--L---I---------V-----I-------------.-DIN---K--------- 104
19_cpx.CU.99.CU7              ----N---.P-....-E--KLP----............---..........--------R-.....--GN.-..PL......-.....-T-SE....-.....--........ ...-LSL--I------------------------------.-DIN---K-R------- 105
20_BG.CU.99.Cu103             ----T---.P-....------P--A............---..........---C---R-.....RRG-.S..-L......P.....--R-E....G.....E-........ ....L-L--I---------VR-----I-------------.-DIN---K--------- 102
21_A2D.KE.99.KER2003          ----N---.--....-E-G-L----V............---..........---S-DVR-.....R----...PL......-.....-T--E....-.....--........ ...-LS---I------------------------------.---N---K--------- 105
22_01A1.CM.01.01CM_0001BBY    ----N---.Q-....-E---L--G--............---..........---S--RGD.....G----...LL......P.....----E....-.....E-........-AS--S---I------------------------------.-DID---K--------- 106
23_BG.CU.03.CB118             ----T---.P-....------L---A............---..........---C---RI.....RRG-.S..PL......P.....--R-E....G.....A-........ ....L-L--I---------VR-----I-------------.-DIN---K--------- 102
24_BG.ES.08.X2456_2           ----T---.P-....----G-L---A............---..........---C--PR-.....RRG-.S..PL......P.....--R-A....G.....ER........ ....L-L--I---------V-V----I-------------.-DIN---K--------- 102
25_cpx.CM.02.1918LE           --------.Q-....RE---L-P--A............--D..........---S--PR-.....--G-.-..PL......P.....----E....-.....--........VI.-YG---I---------V-----II------------.-DIN---K--------- 105
26_AU.CD.02.02CD_MBTB047      ----I--S.Q-....-E--------V............S--..........-S-S---GN.....GRG-.-..PL......P.....-TERG....-.....--........--S-S----I------------E---R-------------.-DIN-S-K--------- 104
27_cpx.FR.04.04CD_FR_KZS      ----N-VS.Q-....RE-----P--A............---..........---S--PRD.....H-G-.D..LL......-.....-T--E....G.....--........ ..T--S---I---------V--E---R-------------.--IN-Q-K--------- 104
28_BF.BR.99.BREPM12609        ----N---.P-....-E--K-P---A............---..........--AS-----.....RRG-.-..PL......-.....----E....-.....G-........--P-LSL--I-----------V---R-------------.-DIN---K--------- 106
29_BF.BR.01.BREPM16704        ----N---.P-....-E---P---A............--I..........--AS---R-.....-R-E.-..PL......-.....----E....-.....R-........E-P-LS---I-----------V------------------.-DIN---K--------- 106
31_BC.BR.04.04BR142           ----N---.P-....-E----PP---............---SPASEQTRAD--AS------.....R-.-.-..PL......-.....----E....-.....--........ ...TL----I-----------V------------------.--IK---N--------- 112
32_06A1.EE.01.EE0369          ----N---.Q-....-E---L----A............---..........---C-----.....R-G-SS..PL......L.....----E....E.....K-........AI.-LS---I------I--VR--E--I-------------.-DIN---K--------- 106
33_01B.ID.07.JKT189_C         ----A---.P-....-E---LP-KP-............---..........---S-----.....--T--...-I......-.....-----....-.....--.........AI-CS---I----------------R-------------.--IN---K--------- 106
34_01B.TH.99.OUR2478P         ----I---.Q-....---G--P----............---..........----G----.....------..-I......-.....-T---....-.....--.........TI--S---I---------V--------------------.-DIN---K--------- 106
35_AD.AF.07.169H              ----N---.Q-....RE--K---G--............---..........---S---R-.....-----K..AL......-.....-T--K....-.....R-........V-.-S----I---------V-L---QR------------.-DIN---K--------- 105
36_cpx.CM.00.00CMNYU830       ----N---.Q-....-E---------............-T-..........---S---WD.....R---.V..LF......-.....---IE....G.....--........-IS-L----I---------VR-----I-------------.-DIN---K--------- 106
37_cpx.CM.00.00CMNYU926       ----N---.Q-....RE--K------............G--..........---S---RG.....ERGS.S..LL......-.....-----....G.....--........-IP--S---I------IT-V------I-------------.-DIN---K--------- 104
38_BF1.UY.03.UY03_3389        ----N---.P-....-E-GK-P---A............--I..........---S-KFW-.....RRG-.-..P-......-EAGA.----E....G.....R-........N-P-LS---I------------------------------.-DIN---K--------- 110
39_BF.BR.04.04BRRJ179         ----N---.Q-....-E--------N............---SPN.......---S--PN-.....--G-.S..-L......P.....-T--E....G.....--........--.---L--I------I----V---R-------------.-DIN---K-T------- 109
40_BF.BR.05.05BRRJ055         -------S.--....-E----P---A............---..........--AS---W-.....RRG-.-..P-......-.....----E....K.....R-........--S-L----I------V--V-----------------N--.-DIN---K--------- 106
42_BF.LU.06.luBF_18_06        ----N---.Q-....-E--K-PP--A............--KD......RAD--AS---W-.....R-G-.-..LL......-.....----E....-.....R-........S-P-L----I-----------V-E-I-------------.---N------------- 110
43_02G.SA.03.J11223           ----N---.P-....-E---LP---A............---..........---C---R-.....RRG-.S..-L......P.....----E....G.....E-........IT.-ISL--I--------KV---E--I------------.-QID------------- 105
44_BF.CL.00.CH80              ----N---.Q-....-E--K-P----............---..........---S---W-.....RRG-.-..PF......-.....----E....G.....R-........K-P-LC---I--------K--V---I-------------.-DIN-----T------- 106
45_cpx.FR.04.04FR_AUK         ----N---.Q-....RE---L-----............G-I..........--SS---WD.....G----...LL......-.....----E....G.....--........AG..LS---I------V--V-V-----------------.--IN---K--------- 104
46_BF.BR.07.07BR_FPS625       ----N---.Q-....-E--K-P---A............---..........--AS---RD.....QRG-.-..PL......-.....----E....-.....R-........-LP-LS---I-----------V---R-------------.-DIN---K--------- 106
47_BF.ES.08.P1942             ----N---.Q-....-E-----P--A............---..........--AS---G-.....-RG-.-..PL......-.....---TE....-.....R-........--S-LS---I---------V-V--M-------------.-DIN---K--------- 105
48_01B.MY.07.07MYKT021        ---KA---.P-....-E---LY----............---..........---SG----.....------..-I......-.....-----....-.....--.........TI--SL--I------------------------------.---N----R-------- 106
49_cpx.GM.03.N26677           ----N---.P-....-E----P----............---..........---SG--R-.....RR--.S..PL......P.....----E....-.....G-........--.--S---I---------R--------------------.--ID---K--------- 104
50_A1D.GB.10.12792            ----N---.P-....R-----PP---............---..........---S---R-.....---S.-..TF......P.....-T--E....-.....--........A-.--S---I---------VRV------------------.---N-Q-K--------- 105
51_01B.SG.11.11SG_HM021       --------.P-....---G-------............---..........---S-----.....--G--...-L......-.....-----....-.....--........--.-LSL--I------------------------------.--ID------------- 106
52_01B.MY.03.03MYKL018_1      ----N---.Q-....R--G------A............---..........---S-K-GD.....G--.D-..LL......T.....----X....G.....--........SSS--S---I---------V-V---R-------------.-DID---K--------- 106
53_01B.MY.11.11FIR164         ----A--X.P-....-E-X-LP-XP-............---..........----G--P-.....------..-I......-.....-----....-.....--.........AI------I-----------VA----------------.---N------------- 106
54_01B.MY.09.09MYSB023        ----N--L.P-....------L----............---..........----G----.....G-----..-T......-.....-----....-.....--........-I.-CVC--I------I-------E--------------.-DIN------------- 106
55_01B.CN.10.HNCS102056       ----AV--.Q-....---GK------............---..........---S-G-GD.....G----...LL......T.....K---E....-.....--........-SP-CS---I------------E-----------------.-DIN---K--------- 106
56_cpx.FR.10.URF5_patient_A   ----N---.Q-....-E--K------............GT-..........---S--RWD.....G--S.-..LL......-.....---VKGQ..G.....--........-IP-L----I---------VR-----I-------------.-DIN---K--------- 108
57_BC.CN.09.09YNLX19sg        ----AV--.P-....-E---P-----............---..........---------.....R-.-.-..PS......A.....--RTE....-.....--........ ...TL----I--------S-RV---I-------------.--IN---K--------- 102
58_01B.MY.09.09MYPR37         ----I--S.Q-....RE---LP-XP-............--X..........----G----.....--X---..-X......-.....-----....-.....--.........TIP-D---I------------------------------.-DID---K--------- 106
59_01B.CN.09.09LNA423         ----N---.--....R---K--PK--............---..........---S-K-GD.....G-GE.....L......T.....---TE....-.....--........-AS-L----I---------V--------------------.-DIN---K---R----- 103
60_BC.IT.11.BAV499            --------.P-....-E--K-PP---............GT-..........-T-S-----.....R-.-.-..--......A.....----E....-.....--........ ...T-----I------V----V-------P----------.--TK---N--------- 101
61_BC.CN.10.JL100010          --------.P-....-E----PP---............---..........--AS-----.....R-.-.S..LS......P.....---NE....-.....--........ ...T-----I--------S--V---I-------------.--IN---K--------- 102
62_BC.CN.10.YNFL13            ----N---.P-....-E--K--PK-A............---..........---S-----.....RR.-.-..PS......-.....---IE....-.....--........ ...T-----I--------P-IV--M-------------.--IN---K--------- 102
63_02A1.RU.10.10RU6637        ----T---.Q-....RE--K------............---..........---S--RGD.....G----...LL......-.....A---E....G.....--........AIS--S---I---------VR-E---I-------------.-DID---K--------- 106
64_BC.CN.09.YNFL31            -------L.P-....-----------............--I..........---S--F--.....------..-I......-.....---G-....-.....--........-I.PLS---I--------S--V---I-------------.---N---K--------- 106
65_cpx.CN.10.YNFL01           ----N---.P-....-E----PP---............--D..........---S-----.....RR.-.-..PS......-.....---VE....-.....--........ ...TL----I--------S-RV---I-------------.--IN---K--------- 102
67_01B.CN.11.2011.ANHUI.HF115 ----N---.Q-....---X---X---............---..........---S-G-GD.....GR--.-..LL......T.....--R-E....-.....--........ ..T-LS---I---------V--------------------.-DIN---K--------- 104
68_01B.CN.11.2011.ANHUI.WH73  ----N---.Q-....X----------............---..........---S-G-GD.....GR--.-..LL......T.....--R-E....-.....--........ ..TALS---I---------V-V------------------.-DIN---K--------- 104
70_BF1.BR.10.10BR_PE004       --K-N---.P-....-E-KK-Y---A............---..........---S-----.....RR--.D..P-......-E.......---GAE...-.....RR........--P-L----I-----------V---I-------------.XDIN--XKX---X-X-- 108
71_BF1.BR.10.10BR_PE008       ----NV--.Q-....-E--K-P---D............---..........--AS---R-.....R---.D..PL......-.....----E....K.....R-........--P-LS---I--------K--------------------.-DID---K--------- 106
72_BF1.BR.10.10BR_MG002       ----N---.P-...R-E---LP--N............G--..........GSAS---R-.....RRGE.-..PL......-.....----E....G.....R-........--P-LS---I------M---RV------------------.-DIN---K--------- 107
O.BE.87.ANT70                 ---QI--S.GG...HE--QLCA-TS............TPI..........---DGGGSE.....GTGES.................GTERG....P.....ER........AL.-VCL--IP--D--I--ARV--H-C-V-----------.NNIQ-E-K--------- 102
O.CM.98.98CMA104              ---QI--S.GG...HE--Q-YA-TS............GPI..........--SDGGSSE.....GTGEP.................GTERG....L.....KR........-L.-VCL--IP--D--V--A-V--H-C-V-----------.NDIQ-E-K--------- 102
O.CM.98.98CMABB141            ---QI--S.GG...RE--QLYA-TS............API..........---NGGNDE.....GT-ES.................GTKGG....H.....NR........-L.--C---IP--D--V-IARVE-H-C-V-----------.NNIQ-E-K-T------- 102
O.CM.98.98CMABB212            ----V--S.GG...HE--QLXA-TS............VP-..........---XGGDSK.....GAGES.................-SRXX....X.....GR........--.PVCL--IP--D--I--A----H-C-------------.NNIQ-E---T------- 102
O.CM.98.98CMU5337             ----I--S.GG...HE--QLCT-TS............VT-..........---NGGGSE.....ET-ES.................GTEGR....P.....ER........AL.-VCL--IP--D--V--ARV--H-C-V-----------.NNIQ-E-K-T------- 102
O.CM.99.99CMU4122             ----I--S.GG...HE--QLCT-TS............VPI..........-SSDGGGKE.....GT-ES.................GTEGG....P.....ER........-L.-VCL--IP--D--V--ARV--H-C-V-----------.NNIQ-E-K-T------- 102
O.FR.92.VAU                   --K-I--S.ER...HE--QLCAGTS............VPI..........---NGRDDE.....GA-GGS................-SEGG....A.....ER........AL.-VCL--IP--D--V--ARV----C-V-----------.NNIQ-E-K-A------- 103
O.GA.11.11Gab6352             ---QI--S.GG...HE--QLCA-TS............APL..........---DGGGSA.....ET-ES.................GTEGG....P.....KR........AL.-VCL--IP--D--V--ARV--H-C-V-----------.NNIQ-E-E-T------- 102
O.SN.99.99SE_MP1299           ----I--S.GG...HE--QLCA-TS............VPI..........---DDGGNE.....GT-ES.................GTEGG....P.....ER........-L.-VCL--IP--D--V--ARV--H-C-V-----------.TNIQ-E-K-T------- 102
O.US.10.LTNP                  ---QV--S.GG...HE--QLCT-TS............GPI..........-S-DGGGGK.....QTGES.................GTERE....P.....NR........-L.-VCL--IP--D--II-A-V--H-C-V-----------.NNIQ-E---T------- 102

N.CM.02.DJ00131               ---KG-VS.--....RET-QLPPDNNNKE.........--H..........--AT---W-.....S-GE.E..HT......G.....-GE-G....E..........HRELSIPTL----I------VI-V---KEVR------------I.--IQ-E-K--------- 111
N.CM.02.SJGddd                ---QG-VS.--....RET-K-PPDNSKE..........G-H..........---T-..........-GE.E..HT......G.....-RD-G....E.....P-...EDRELS-PTL----I------VI-V---EEVR------------I.--IQ-E---------- 108
N.CM.04.04CM_1015_04          ---KG-VS.--....RET--LPPDSNKE..........--H..........--AT---W-.....S-GG.E..HT......G.....KGD-G....E.....P-...EDRELS-PTL----I------VISV---NEVR------------I.--IQ-E---------- 113
N.CM.06.U14296                ---KG-VS.--....RET--LPPDNNKE..........--H..........--AT---W-.....S-GE.E..HT......G.....KGD-R....E.....P-...EDRELS-PTL----I------VI-V---KEVR------------I.--LQ-E-K--------- 113
N.CM.06.U14842                ---QG-VS.--....RET--LXPDNNNKE.........--H..........--AT-G-W-.....S-GE.E..HT......E.....KGD-G....G.....P-...EDRELS-PTL-L--I------II-V---KEIR------------I.--LQ-E-K--------- 114
N.FR.11.N1_FR_2011            ---KG-VS.--....RET-KLPPDNNKE..........GTH..........--AT---W-.....S-GE.E..HT......G.....-GD-G....G.....P-...EDRELS-PTL----I------VI-V---EEEI------------I.-DLQ-E---R------- 113

P.CM.06.U14788                -------S.GG....QE--QLCT-TS............TTD..........---YGGGGD..........-.............-EQA..-G-KG....G.....RE..R....V-.P-SL--I---D--VIPV----HIC------------V.DNLP-E----------- 103
P.FR.09.RBF168                ---K---S.WG....QE--QLCT-TS............TT-..........---YGGGRD..........-............PGQA..-GXGG....G.....X-........GI.P--L--I---D--VIPVRV--HIC------------V.DKLP-E-----N----- 102

CPZ.CD.06.BF1167              ----THPL.VG....VQT-K-CA-SESDSESGT.................ESDTG-TSG....ES-..T..LSS......-.....NG-GESL..EQ.....--................WL--EIH-----MMEVTVQ--VCT-----------F.ADLK-K-SCN-RT---- 109
CPZ.CM.05.SIVcpzMT145         ---KK--S.-G....-E---LCAR-E............GT-..........-S-YGI-RI.....S-GE.E..-Q......G.....GGEGG....E.....-A........A-.------I-------IPV--E---I------------I.-DLN---K-S------- 105
CPZ.GA.88.GAB1                ----R---.P-....RE--QLCA--N............-T-..........G--D---W-.....P-GR.E..E-............GEERG....-.....EQ........SI.-T-L--I-------IPV-VE---C------------I.-RIQ-Q-L--------- 104
CPZ.TZ.06.SIVcpzTAN13         ----THPL.VG....VQT---CAK-P............-EG..........EAPGARG.....SA.-.SAYC-........-N....GD-PH....E.....RR..............VL--IN-----MMAVRVQ-H-CQ---------S-F.CNLQ-K-Q----T---- 101
CPZ.US.85.US_Marilyn          ----T-VP.IVERGIKET--LPGK-E............G-H..........-S-N-GFR-.....SR--.S..DT......G.....-T-KG..........K-........AL.-A----I-------LR--VA--IV--------------.DNIQIE-T-R---M--- 108

GOR.CM.07.SIVgor2139_287      ----I--S.RG....RE--QLCA-AS............AT-..........---NGGGRK.........................G.....DEETR....K.....G-........-L.-S----I---D--VIPV----H-C------------V.-HL--E-----R----- 98
GOR.CM.07.SIVgorCP2135con     ----I--S.RG....RE--QLCA-AS............AT-..........---NGGGRE.........................R.....NEETR....E.....G-........-L.-S----I---D--IIPV----H-C------------V.-HL--E-----R----- 98
```

```
                                                                     protease end p66, p51 RT start                           M41L         K70R
                                                                                                                                          D67N                                    D110 catalytic site
B.FR.83.HXB2                GGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD 276
A1.AU.03.PS1044_Day0        ------K-----S-----KR------------------MX------------------------------------A------T----D-------------------I------------------------------------------------------------- 276
A1.CH.03.HIV_CH_BID_V3538   ------------------K-------------------M--------------------T-----------------------T---N---------------------------------------------------------------------------------H 276
A1.CY.08.CY236              ------K-----P-----K-------------------M-----------------------------R--------------T------------------------I------------------------------------------------------------H 273
A1.ES.05.X1608_8            ------K-----P-----KR------------------M--------------------------------------------T----D-------------------I------------------------------------------------------------- 280
A1.KE.11.DEMA11KE001        ------K--E--V-----K-------------------M--------------------------------------------T------------------------I------------------------------------------------------------- 276
A1.RU.11.11RU6950           ------------------K-------------------M---L----------------T-----------------------T---K---------------------------------------------------------------------------------- 276
A1.RW.11.DEMA111RW002       ------K-----------K-------------------M--------------------------------------------TD--A---------------------------------------------------------------------------------- 276
A1.SN.01.DDI579             ------------------K-------------------M---------------K----T-----------------------TA--K---E------------------------------------------------------------------------------ 275
A1.UG.11.DEMA110UG001       ------K-----V-----K-------------------M---------------D----T-----------------------T----D--------R----------I------------------------------------------------------------- 276
A1.ZA.04.04ZASK162B1        ------------------K-------------------M---------------K------R---------------------T-------------------------------------------------------------------------------------- 275
A2.CM.01.01CM_1445MV        ------------V----EKR------------------I-V-L----------------------------------------T---KD-------------------------------------------------------------------------------Y 276
A2.CY.94.94CY017_41         ------------A-----KR------------------M-V-L----------------------------------------T---K-----------------------------------------------------------------A---------------H 276
B.BR.10.10BR_MG029          -----------VP---------------------------------------------------------------------------------------------------------------------------------------------------------- 276
B.CA.07.502_I191_03         ------K-----P-------T---------------------L-----------------------------------------------------------------------------------------------------------Q------------------Y 276
B.CH.08.M2_0803101_NFLG8    ----R-K-----P-------V-----I-----------VM--L----------------------------------------------L------------------I-------G-----IM----------------------G----N------------------ 280
B.CN.10.DEMB10CN002         ------K-----PV------T-----I------------M--L-----------------------------------------------------------------------------K------------------------------------------------- 276
B.ES.10.DEMB10ES002         -----------VSV------------I------------M-----------------------------------------------------------------------------------------------------------------------------------H 276
B.FR.11.DEMB11FR001         -----------VH----------------------------------------------R------------------------K------------------------------------------------------------------------------------- 275
B.GB.05.MM45d213_GN1        ------------P----------------------------E-----------------------------------------T------------------------I------------------------------------------------------------- 275
B.HK.06.HK003               ------K----VS-------------I--------------------------------------------------------I------------------------I---R------------------K-------------------------------------- 280
B.HT.05.05HT_129389         -----------VH-------T------------------------------------------------------------------------------------------------------------------------------X----------------------- 275
B.JP.12.DEMB12JP001         ------K-----A-----------------------------L----------------------------------------I-------------------------------------------------------------------------------------- 279
B.KR.07.HP_18_07JHS10_3909  -----------VA--------T---------------------------------------------------------------------------------------------N-NR-------------------------------R------------------N 276
B.PE.07.502_0525_wg5        ----------EVSL-------T---------I----------L-------------------------R----------------------------------------------------------------------X----------QR------------------ 280
B.RU.11.11RU21n             -----------VP-----Q-------------------------------------------------R--------------I-------------------------------------------------------------------N------------------ 280
B.TH.08.MERLBDTRC10         ---------E--P-------------I---------------L-----------------------------------------------------------------I-----------------------------------------Q------------------- 276
B.US.11.ES38                ------K--E--P-------------I-----------------------------I-------------------------------------------------------------------------------------------------------------C 280
C.AR.01.ARG4006             ------------VM----K-------------------M---L----------------------------------------T---E--------T------------------------------------------------------------------------- 272
C.BR.07.DEMC07BR003         ------------P-----K-------------------M---L----------------------------------------TA--D---R-----------------------------------------------------------------------------H 272
C.BW.00.00BW5031_1          ------------M-----Q-------------------M---L----------------------------------------TA--E--------T------------S------------------------------------V----------------------- 278
C.CN.10.YNFL19              ------------P-----K-----------------------L----------------------------------------T---N--------T-----------I----------------------------------------Q-------------------H 276
C.CY.09.CY260               -----------XX-----K----S--------------M---L---------S------------------------------XD--E--------X------------------------------------------------------------------------- 279
C.ES.08.X2363_2             -----------VT-D---K----S--------------M---L----------------------------------------TA--E---------------------------------------------------------------------------------- 279
C.ET.02.02ET_288            ------------H-----K-----------------------L----------------------------------------T---E---------------------------------------------------------------------------------- 278
C.IN.09.T125_2139           ------------P-----K-------------------M---L-------------I--------------------------TA--E--------T-----------I------------------------------------------------------------Y 273
C.KE.00.KER2010             -----------VI-----K-T-----------------I---L----------------------------------------T---E--------E------------------------------------------------------------------------- 272
C.MW.09.703010256_CH256.w96 ------------S-----K-------------------M---L----------------------------------------TA--E--------T------------------------------------------------------------------------- 274
C.TZ.08.707010457_CH457.w8  ------------V-----K----S--------------M---L----------------------------------------TA--ED-------T------------------------------------------------------------------------- 272
C.YE.02.02YE511             ------------------KR------------------M---L-------------------------R--------------T---ED-------T------------------------------------------------------------------------- 270
C.ZA.08.705010534_CH534.w12 ---------E-V------K----A-------I------M---L----------------------------------------TA--E--------T------------------------------------------------------------------------- 279
C.ZA.10.DEMC10ZA001         ------------P---Y-K-------------------M---L----------------Q-----------------------TA--E--------T-----------I------------------------------------------------------------- 272
C.ZM.11.DEMC11ZM006         ---------E--P-----K-----------------------L----------------------------------------TA--D--------T-----------I------------------------------------------------------------- 277
D.CM.10.DEMD10CM009         ------------P-------------------------------------------I--------------------------T-------------R----------I----------------------------------------------------I---------- 281
D.CY.06.CY163               --------------------------------------M--------------------------------------------T----D--------R------------------------------------------------------------------------ 275
D.KE.11.DEMD11KE003         ------------G-----Y-------------------------------------------------------------------------TQ------I-----R----------I------------------------------------------------------------Y 275
D.KR.04.04KBH8              ---------------------------------------------------------------V----R--------------T-------------R------------------------------------------------------------------------ 275
D.SN.90.SE365               ------K---N---------------------------M-----------------I--------------------------T------------------------I------------------------------------------------------------H 281
D.TZ.01.A280                ------------------Q------------------------------------A----------------------------T-------------R----------I------------------------------------------------------------- 277
D.UG.10.DEMD10UG004         ------------TVD----------------------------------------------------------------V---T---R---E-----R----------I------------------------------------------------------------Y 275
D.UG.11.DEMD11UG003         ------K---H-SVD-----------I--------------------------------------------------------T----D--------R----------I------------------------------------------------------------- 277
D.YE.02.02YE516             ---------E-V-V---------------------------------------------------------------------I-------------R----------I------------------------------------------------------------- 277
D.ZA.90.R1                  ------------PL------------I---------------L----------------------------------------T------------------------I------------------------------------------------------------C 275
F1.AO.06.AO_06_ANG32        ------------T-D---Y--T----I-----------M-----------------I--------------------------T---E---------------------------------------------------------------------------------- 280
F1.AR.02.ARE933             ------K---N----------TC-----------F-I-M--------H------K------------------------L---S---I-----------------X---------------------------------------------------------------- 276
F1.BR.10.10BR_RJ015         ------K---N-I-D------T----------------M--------------------------------------------T---L---------------------------------------------------------------------------------- 279
F1.CY.08.CY222              ------K-----T-D------T----------------M--------------------------------------------T-------------------------------------------------------------------------------------- 276
F1.ES.02.ES_X845_4          ------K-----T-D----R------------------M--------------------------------------------T-------------------------------------------------------------------------------------- 278
F1.RO.96.BCI_R07            ------K-----T-D------T----------------M--------------------------------------------T---AD--------------------------------------------------------------------------------- 276
F1.RU.08.D88_845            ------K-----V-------------------------M-----------------------------R-----------M--K---------------------------------S------------------C--------------------------------- 276
F2.CM.02.02CM_0016BBY       -----------VP-----Q-------------------M--------------------------------------------T-------------------------------------------------------------------------------------- 276
F2.CM.10.DEMF210CM001       -----------VH-----Q-------------------M--------------------------------------------T-------------------------------------------------------------------------------------- 281
F2.CM.10.DEMF210CM007       ------------------Q-------------------M-----------------------------R--------------T-------------R------------------------------------------------------------------------ 275
G.BE.96.DRCBL               ----------------S-KR-----------I------M-----------------------------R--------------T---N--------------------I---------R--------------------P------G---Q-R----------------- 275
G.CM.10.DEMG10CM008         ------K--E--P---E-K-------------------M--------------------------------------------T-------------R----------I-----------------------------------------Q------------------- 275
G.CN.08.GX_2084_08          ------K----VP---E-K------------I------M--------------------------------------------T---------------------------------------------------------------------R----------------- 280
G.CU.99.Cu74                ------------------K------------I------M------------------------------I-------------TD--AD-------------------I------------------------------------------------------------- 272
G.ES.09.X2634_2             ------K-----------KR-----------I------M--------------------------------------------T------------------------I------------------------------------------------------------- 280
G.GH.03.03GH175G            ------------------K------------I------M-----------------------------RI-------------T---KD---------------------------------------------------------------R----------------- 276
G.KE.09.DEMG09KE001         ----------------G-K------------I------M--------------------------------------------T-------------E----------I------------------------------------------------------------- 274
G.NG.09.09NG_SC62           ------------PLK-E-K--------R---I------M--------------------------------------K---E--T---M--------------------I---------------------------------------------I--------------- 275
G.ZA.01.TV546               ------------SL----KR-----------I------I--------------------------------------------T------------------------I------------------------------------------------------------- 275
H.BE.93.VI991               ---------E-VA---F-K-------------------I---M-----L----------T-----------------------T---L--------------------I------N--R------------------------------------S-----G-------H 274
H.CF.90.056                 ---------E-VA-----K-------------------I--------------------------------------------T-------------R-------S--I----------------------------------------------S-------------- 275
H.GB.00.00GBAC4001          ---------E-VA-----K-------------------I--------------------T-----------------------T-------------R----------I------------------------------------------N-----------------Y 274
J.CM.04.04CMU11421          ---------SN-P---E-K--X----------------M---L----------------------------------------T---L---------------------------------------------------------------------------------- 272
J.SE.93.SE9280_7887         ---------NEVP---E-K-------I-----------M---L--------------------------I-------------TQ--A---E-----RV----------------------------------------------------------------------Y 274
J.SE.94.SE9173_7022         ---------NEVP---E-K-------I-----------M---L--------------------------I-------------TQ--A-L-E-----R-----------------------------------------------------------------------Y 274
K.CD.97.97ZR_EQTB11         -----------VCM----Q-------------------M-------------------------------------------------------------------------------------I------------P------------------------------------ 276
K.CM.96.96CM_MP535          -----------V------Q-----------------------------------------------------------------T-------------------------------------------------------------------------------------- 276
U.CA.01.TV749               ------K-----SM----KR------------------M-----------------------------R--------------L-------------R------------------------------------------------------------------------ 281
U.CA.99.TV721               ------------SM----KR------------------M-------------H---------------R--------------I----D--------R------------------------------------------------------------------------ 280
U.CD.83.83CD003_Z3          ------------------K-------------------M--------------------------------------------T---L---------------------------------------------------------------------------------- 277
U.CD.90.90CD121E12          ------------------KR------------------M--------------------------------------------T-------------------------------------------------------------------------------------- 276
U.CY.05.CY090               ---------EK-----E-K--------R----------M------------------------------I----------R--T--------KR--------------I------N----K-------------N--------------------------N-------- 276
U.CY.08.CY223               ------K-----------KR------------------M--------------------------------------------T---N---E-----------------------------------------------------------------------------H 277
U.ES.10.DEURF10DZ001        ------------------K-------------------M------------------------------I----------R--T---AD-------------------I------------------------------------------------------------- 276
U.GR.99.99GR303             ------------------K---------S---------V--------------------A---------I-----X-------T----------------------------------R--------------------------------------------------- 275
U.NL.95.U_NL_95_H10986_D1   ---------E-VC-----Q-------------------M-----------------------------------------R--TD-------------------------------------------------------------------------G----------- 276
01_AE.AF.07.569M            ----------------W-K-------------------M---------------N----T-----------------------T---K-L-A----------------I------------------------------------------------------------- 276
01_AE.CF.90.90CF11697       ------------------KR------------------M---------------D----T--------------*-----*--T---K---E---------------------E----------------------------V--------------------------Y 274
01_AE.CN.10.YNFL03          ------------------K-------------------M---L-----------D----------------------------T---K---E------------------------------------------------------------------------------ 275
01_AE.HK.04.HK001           ------------------K-------------------M---------------D----------------------------T---K---E----------------------------------------------I------------------------------- 273
01_AE.IR.10.10IR.THR48F     ------------------K--V----------------M---------------D----T-----------------------T---K---------------------------------------------------------------------------------- 268
01_AE.JP.x.JRC77AE          ------K-----A-----K-------------------M---------------D----T-----------------------T---K---E------------------------------------------------------------------------------ 271
01_AE.TH.09.AA090a_WG11     ------------S-----K-------------------M---------------D-I--T-----------------------T---K---E------------------------------------------------------------------------------ 276
01_AE.TH.90.CM240           ------K-----------K-------------------M---------------D----T-----------------------T---K---E------------------------------------------------------------------------------ 281
01_AE.US.05.306163_FL       ------K-----------K-------------------M---------------D----T-----------------------T---K---E-------------------------------------------------------------------I---------H 274
01_AE.VN.98.98VNND15        ------------------K-------------------M---------------D----T-----------------------T---K----------------------------G----------------------------------------------------- 275
```

```
                                                           protease end p66, p51 RT start              M41L          K70R                       D110 catalytic site
                                                                                                                    D67N
B.FR.83.HXB2                  GGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD 276
02_AG.CM.10.DE00210CM013      -------------M----K-------------------M-----------------------------------------------------------------------------------------------------------------R----------------- 276
02_AG.CY.09.CY256             ------------X-----K-------------------X-----------------------------R--------------TD--L---------------------------------------------------------------------------------- 276
02_AG.ES.06.P1423             ------------------K---------------X---M----------------X-----X---------R--------X--TD---D--------R--X--------X------------------------------------------------------------ 276
02_AG.GW.05.CC_0048           ------------------K-------------------M-------------S------------------------------TD--K---------------------------------------------------------------------------------- 274
02_AG.KR.12.12MHR9            ------------------K-------------------M--------------------------------------------TD-----------------------------------------------------------------E------------------- 280
02_AG.LR.x.POC44951           ------------P-----K-------------------M---------L-----K----------------------------TD--A--------------------------------------------------------------E------------------- 276
02_AG.NG.09.09NG_SC61         ------------S-----K-------------------M--------------------R-----------------------TD------------------------------------------------------------------------------------- 276
02_AG.NG.x.IBNG               ------------------K-------------------M--------------------------------------------TD------------------------------------------------------------------------------------- 276
02_AG.SE.94.SE7812            ------------------K-------------------M--------------------------------------------TD------------R----------I------N--R-------------------I------------------------------- 276
02_AG.SN.98.98SE_MP1211       -----L------------K--M-S------------K-I-------------------L--------------------L-H-TD--A-------------------------R-------------------------------------------------------- 276
03_AB.RU.97.KAL153_2          ------------------K-------------------M---L----------------T-----------------------TD--K------------------------------------G--------------------------------------------- 276
04_cpx.CY.94.94CY032_3        ------------P-----K-------------------M---L----------------------------------------T----D-------------------I------------------------------------------------------------- 275
05_DF.BE.x.VI1310             -------------V----Q----------------------------------------------------------------T----D--------R----------I---------R-----------------------------------I--------------- 276
06_cpx.AU.96.BFP90            ------K-----------K------------I------M--------------------------------------------T---K--------------------------------------------------------------------------I------- 278
07_BC.CN.98.98CN009           ---------E--P-----------------------------L----------------------------------------TA--D--------T-----------I-----------------------------------------R------------------Y 276
08_BC.CN.97.97CNGX_6F         ---------E--P-----K-------------------M---L----------------------------------------TA--D--------T----D------I---R----S---------------------------------------------------- 272
09_cpx.GH.96.96GH2911         ------------------Q-------------------M------------------------------I-------------T---A---------------------------------------------------------------------------------- 275
10_CD.TZ.96.96TZ_BF061        ------------------Y----------------------------------------------------------------T---K---------R----------I------------------------------------------------------------Y 275
11_cpx.CM.95.95CM_1816        ---------EE-----E-K-------------------M--------------VD----------------------------T---------------------------------------------------------------------------------------- 274
12_BF.AR.99.ARMA159           ------K---NV--------------I---------------L----------------------------------------T---------------------------------------------------------------------------------------- 276
13_cpx.CM.96.96CM_1849        ----------NMS---E-K-------I-----------M--------------------------------------------T---K---------------------------------------------------------------------------------- 283
14_BG.ES.05.X1870             ------------------K------------I------M------------------------------I-------------I-------R----------------I-------------------------------------S----------------------- 275
15_01B.TH.99.99TH_MU2079      ------------P-----K-------------------M---L----------------------------------------T---K---E----------------I-------------------------------------------------V----------- 275
16_A2D.KR.97.97KR004          ------------T---EKR-------------------M-V-L----------------------------------------T---K--------------------------------------------------------------M---------R--------H 276
17_BF.AR.99.ARMA038           ------K---D-------------------------------L------------------------------------------------------------------------------------------------------------------------------- 276
18_cpx.CU.99.CU76             ------------------KR------------------M--------------------------------------------T------------------------I------------------------------------------------------------- 274
19_cpx.CU.99.CU7              ------------P-----Q----------------------------------------------------------------T----D-------------------I------------------------------------------------------------- 275
20_BG.CU.99.Cu103             ------------------K------------I------M---------------D------------------------------------------------------------------------------------------------------------------- 272
21_A2D.KE.99.KER2003          ------------PL---------------------------------------------Q-----------------------T---MD--------R----------I--------------------------------------X---------------------S 275
22_01A1.CM.01.01CM_0001BBY    ------------------K-------------------M--------------------A---------I-------------T---K--------------------I------------------------------------------------------------Y 276
23_BG.CU.03.CB118             ------------P-----K------------T------M--------------------------------------------------L-------------------------N------------------------------------------------------ 272
24_BG.ES.08.X2456_2           -------------M----K------------I------M-----------------------------------------------------------------------------------------------------------S----------------------- 272
25_cpx.CM.02.1918LE           ----------E-VL--S-K------------I------I--------------------------------------------T---ND--------R----------I----------------------------------------------S-------------- 275
26_AU.CD.02.02CD_MBTB047      -----------VT-D---K-------------------M--------------V----------------------------IT---N----------V---------I------------------------------------------------------------- 274
27_cpx.FR.04.04CD_FR_KZS      ---------EEVTVD---K------------------------------------------------------------------------------R--------------R-----------------------------------------I--------------- 274
28_BF.BR.99.BREPM12609        ------K---D---D------T----------------M------------------------------------------------------------------------------------------------------------------------------------ 276
29_BF.BR.01.BREPM16704        ------K---D---------------------------M--------------------------------------------M-----------------------------------------------K-------------------------------------- 276
31_BC.BR.04.04BR142           ------------------K-------------------M---L----------------------------------------T---N---R-----------------------------------------------------------N------------------ 282
32_06A1.EE.01.EE0369          ------------H-----K-------------------M---------L-------------------R-----------R--T------------------------I------------------------------------------------------------- 276
33_01B.ID.07.JKT189_C         ------------------K-------------------M--------------------------------------------I-------------------------------------------------------------------------------------- 276
34_01B.TH.99.OUR2478P         ------------P-----------I-----------------L----------------------------------------A---A--------------------I------------------------------------------R------------------ 276
35_AD.AF.07.169H              ------K-----A-----K-------------------M---------------D----R-----------------------T------------------------I------------------------------------------------------------- 275
36_cpx.CM.00.00CMNYU830       ------------P-----K------------I--V---M--------------------A-----------------------T---A------------------------R---------------------------------------R----------------- 276
37_cpx.CM.00.00CMNYU926       ------------P-----K-------------------M---------------N----S---------------S-------T---N---------RV---------I-------------------------------------------R----------------- 274
38_BF1.UY.03.UY03_3389        ------K---N-C----D--VT----I---------------L----------------------------------------M---AD------------------------------------------K------------------R------------------- 280
39_BF.BR.04.04BRRJ179         --IV-----EN-P------------------A------------------------------------R--------------L---A-L-EA------------------------------M-L-------------------------N----------I------H 279
40_BF.BR.05.05BRRJ055         ----R-K--ED-P------------------------D---------------------------------------------T-------------------------------N--R--------------------------------------------------- 276
42_BF.LU.06.luBF_18_06        ------K---Y-S-D------T----I---------------L----------------------------------------T---M---------------------------------------------------------------------------------- 280
43_02G.SA.03.J11223           ------K-----P---E-K------------I------M-------------------------------------A------T---L--------------------I------------------------------------------------------------- 275
44_BF.CL.00.CH80              ------K---NDT---------------------------------------------------------------G------I-------------------------------------------------------------------------------------- 276
45_cpx.FR.04.04FR_AUK         ------------V-----K----A--------------M--------------------------------------------T------------------------I------N------------------------------------------------------ 274
46_BF.BR.07.07BR_FPS625       ------K---D---------------------------M---------------D-------------R--------K-----T---A----------------------------------------------------------------R----------------- 276
47_BF.ES.08.P1942             ------------I-----K----------------------------------------------------------------I-------------------------------------------------------------------------------------Y 275
48_01B.MY.07.07MYKT021        ------------P----------------------------------------------------------------------M------------------------------------------------A-------------------R--S-------------- 276
49_cpx.GM.03.N26677           ---------ND-----E-K--V----------------M---L----------------------------------------T---------------------------------------------------------------------------------------- 274
50_A1D.GB.10.12792            ------------P-------------I--------------------------------------------------Q-----T---AD--E----T-----------I-----------------#------------------------------------------H 274
51_01B.SG.11.11SG_HM021       -----------VSVD---------------------------------------------------------------------------------------------I------------------------------------------------------------- 276
52_01B.MY.03.03MYKL018_1      ------------------K-------------------M---------------D----------------------------I-------------------------------------------------H------------------------------------ 276
53_01B.MY.11.11FIR164         ---------E---L----K-------------------M---L-----------N----------------------------T---K---E----------------I-------G----------------------------------------------------- 276
54_01B.MY.09.09MYSB023        ------------P-----------------------------L--------------------N----R--------------I-----L-------------------------N--R------------------------------E----M--------------- 276
55_01B.CN.10.HNCS102056       ------------------KR------------------M---------------D----------------------------T---K---E----------------------------------------------------------R------------------- 276
56_cpx.FR.10.URF5_patient_A   ------------SV----KR------------------M-------------H------S-----------------------T---KDV-------R------------------------------------------------------------------------ 278
57_BC.CN.09.09YNLX19sg        ------------P-----KR----------------------L----------------------------------------TA--D--------T-----------I------------------------------------------N-----------------Y 272
58_01B.MY.09.09MYPR37         ------------P-----------------------------X----------------------------------------I-------------------------------------------------------------------------------------- 276
59_01B.CN.09.09LNA423         ------------C-----K-------------------M----------------------------------------------------------------------------------------------------------------------------------- 273
60_BC.IT.11.BAV499            --S---------A-----K-------------------I---L----------------------------------------TA--D--------T------------------------------------------------------------------------Y 271
61_BC.CN.10.JL100010          ------------P-----K-----------------------L----------------------------------------TA--D--------T-----------I------------------------------------------------------------- 272
62_BC.CN.10.YNFL13            ------------P-----KR------------------M---L----------------------------------------I-------------------------------------------------------------------------------------- 272
63_02A1.RU.10.10RU6637        -------------M--Y-K-------------------M--------------------------------------------TA-----------------------I------------------------------------------------------------- 276
64_BC.CN.09.YNFL31            ------K-----P-------------I---------------L----------------------------------------TA--D--------T-----------I------------------------------------------R-----------------Y 276
65_cpx.CN.10.YNFL01           ------------A-----K-------------------I---L-----------D----------------------------TA--D--------T-----------I------------------------------------------------------------Y 272
67_01B.CN.11.2011.ANHUI.HF115 ------------------K-------------------M--------C------D----------------------------T---------------------------------------------------------------------------------------- 274
68_01B.CN.11.2011.ANHUI.WH73  ------------------K-------------------M---------------D----------------------------T------------------------------------------------------------------------------I------- 274
70_BF1.BR.10.10BR_PE004       XX--------N------X---T----I-----------M-------------S------------------------------T---A---------------------------------------------------------------------------------- 278
71_BF1.BR.10.10BR_PE008       ------K--ENM--------------------------M--------------------------------------------T----D-----------------------------------------------------------------I-------I------- 276
72_BF1.BR.10.10BR_MG002       ------K-----------Q----------------------------------------------------------------T---M---------------------------------------------------------------------------------- 277
O.BE.87.ANT70                 ------KE--NVTV--E-REVQ----------------I--GL-----------AP-------------------SK---E--TA--Q---Q-----R----------I-------G-----------------E-----------G---Q-Q---------C------- 272
O.CM.98.98CMA104              --C-R-KEFNNVTV--E-REVQ----------------I--GL---------S-AP---R---------------SK---E--TA--Q---Q-----RV---------I----------------------K--------------G-----Q---------C------- 272
O.CM.98.98CMABB141            ------KE-TNVT--VE-R-VQ----------------I--GL---------H-AP-------------I-----SK---E--TA--Q---Q-----R----------I-------------------------------------G---Q-Q---------C------- 272
O.CM.98.98CMABB212            ------KE-NNVTV-VQ-KEVQ------------L---I--GL-----------AP-------------------SR---E--TA--Q---Q-----R----------I-------------------------------------G---QRQ---------C------- 272
O.CM.98.98CMU5337             ------KE-NKVAV--E-REVL----------------I--GL-----------AP-------------------SK---E--TA--Q---Q-----R------------------------------------E-----------G---Q-----------C------- 272
O.CM.99.99CMU4122             ------KE-YXVKV--E-REVQ----------------I--GL-----------AP-------------I-----SK---E--TA--Q---Q-----RV---------I--------S----------------------------G---Q-Q---------C------- 272
O.FR.92.VAU                   ------KE-YNVAV-LE-REVR------------L---I--GL-----------TP-------------------SR---E--T---K-L-Q-----R----------I-------G-----------------------------G---Q-Q-------K-C------- 273
O.GA.11.11Gab6352             ------KE-NNVTV--E-R-VQ-V--------------I--GL------------P-------------------SK---E--TA--Q-L-Q-----R----------I-------------------------------------G--QQRQ---------C------- 272
O.SN.99.99SE_MP1299           ------KE-N-VPV--E-REVL----------------I--GL-----------AP-------------I-----SK---E--TA--Q---Q-----R----------I-------------------------------------G---Q-Q---------C------- 272
O.US.10.LTNP                  ------KE--NVAV--E-REVQ----------------I--GL-----------AP-------------I-----SR---E--TA--R---Q-----R----------I-------------------------------------G---Q-Q---------C------- 272
N.CM.02.DJ00131               ----------NVT-D-Q-R--V----------------F-----------------------------R-------A---E--RQ--A---Q-----R----------I----------------------------------------Q------------C------- 281
N.CM.02.SJGddd                ----------NVTV--Q-R--V------------------------------------------------------A---E--R-------------R----------I---R-------------------------------------Q-----------C------- 278
N.CM.04.04CM_1015_04          ----------N-T-D-Q-R--V----------------F---------------K---------------------A---E--R-------------R----------I---R-------------------------------------Q-----------C------- 283
N.CM.06.U14296                ----------NVTVD-Q-R--V----------------------------------------------R-------A---E--R-------------R----------I---R-------------------------------------Q-----------C------- 283
N.CM.06.U14842                ----------NXT-D-Q-R--V----------------F-----------------------------------------------A---E--R-------------R----------I-----------------------------------------Q-----------C------- 284
N.FR.11.N1_FR_2011            ----------N-SVD-Q-R--V----------------------------------------------R-------A---E--R-------------R----------I---R-----------------------------------------I-------C------- 283
P.CM.06.U14788                ------KEFEEVKX--E-RQVY----------------I--L-----S-------I------A-----R------SK---E--RA--Q---Q----T-----------I-------GS----------------------------G--Q-R----I-----C------- 273
P.FR.09.RBF168                ----T-KXFEXVKV--E-REVY----------------I--L-----S----S--I---X--A-X------X---SK---E--RX--Q---Q----T-----------I--------S------X---------------X-----G----R----------C------- 272
CPZ.CD.06.BF1167              ---VP-QE-YNVP-T-A-KTVKTK--I-----------V-R-L---------TV-I-K----E------------SR---E--T---ENLRA----IEV--D------I------------------------------------------R-VI-------I------- 279
CPZ.CM.05.SIVcpzMT145         ------K-FERVN---E-K-ITS---------------V--KL-----------D--K-F---------------SK---E--T---K--------------------I-------------------------E---------------------------C------- 275
CPZ.GA.88.GAB1                ------K-F-NVH---E-R-VV----------------I---L----V----S----------------------SA------T---Q--------------------I-----------------------------------------------------C------- 274
CPZ.TZ.06.SIVcpzTAN13         ----P-SE-YN-PVQ-GNKEVRA-----D--L------I-Q-L---------KV-V-K-E--Q------------SK---E--T---STL-------A----------I--------S--------------------L------------R-V--------I------- 271
CPZ.US.85.US_Marilyn          ------K---HVN---E-R--Q-S--------------I-------------S-K----R--------R-------A------T---Q--------TRV---------I-------------------------------------------R---------C------- 278
GOR.CM.07.SIVgor2139_287      ------KE-EDVE---EHR-VV----------------I--K----------S-DV---S---------------SK---E--TA--Q---Q----------------I-------------------------------------G---------------C------- 268
GOR.CM.07.SIVgorCP2135con     ------KE-EDVE---EHR-VV----------------I--K----------S-DV---S---------------SK---E--TA--Q---Q----------------I-------------------------------------G-----R---------C------- 268
```

```
                                                                                                        K219Q
                                                               polymerase motif               T215Y
B.FR.83.HXB2                EDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTE 446
A1.AU.03.PS1044_Day0        -N-----------T-----------------------------------Y-AK----I--------------------A------A---S--F------------------------------N----E---------------------A----K--------A----D 446
A1.CH.03.HIV_CH_BID_V3538   -S--------------------------------------A----------T---E---------------W--E---A------A---S--F------------------------------Q----------------------------------------A----D 446
A1.CY.08.CY236              -------------T--------------------------C------A---SK-------------------------A------A---K--F------------------------------Q----E---------------------A----K--------A----D 443
A1.ES.05.X1608_8            -S-----------T-------------------------------------SK--E-I--------------------A------A---S--F------------------------------E----E---------------------A----K--------A----D 450
A1.KE.11.DEMA11KE001        -S-----------T--A----------------------------------SK--E-I--------------------A------A---N--F---------------------------------------------------------A----K---R----A----- 446
A1.RU.11.11RU6950           -S-----------K-------------------------------------LK---------------------------V----A---S--F------------------------------M--D----------------------------K--------A----D 446
A1.RW.11.DEMA111RW002       -G------------------V------------------------------L---E-I---------------------------A---S--F------------------------------Q--------------------------A--Q-K-------------- 446
A1.SN.01.DDI579             -S-----------T------V-------------------Y----------S---E-I------------N--------------A---K--F------------------------------Q--------------------------A-------------A----D 445
A1.UG.11.DEMA110UG001       -S-----------T-------------------------------------SK--EMI--------------------A------A---S--F------------------------------Q--D----------------------------K---R---------D 446
A1.ZA.04.04ZASK162B1        -------------T------V------------------------------S----------------------------V--------S---------------------------------Q--D-----------------------S-------------A----D 445
A2.CM.01.01CM_1445MV        --------------------V----------------------I-------TK--E-I--------------------A------A------F------------------------------K--------------------------A--Q-K---R----A----D 446
A2.CY.94.94CY017_41         -------------T------V------------------------------SK-TELI----------------S---V------A---K--FY-----------------------------K--------------------------A----K--------A----D 446

B.BR.10.10BR_MG029          -------------K---------------------------------------H---I----------------E---I-------------F------------------------------E--------------------------A------------------- 446
B.CA.07.502_I191_03         -------------A--------------------------Y----------------------------------E-------------K--F--------------------------------------------------------------T-------------- 446
B.CH.08.M2_0803101_NFLG8    -------------T--------------------------A--------------E------V---------------A----------K---Y-----L------------I----------E-------------------------------K--------A----- 450
B.CN.10.DEMB10CN002         K-------------------------------------------------K---------------L--------E-------------K--FC---------------------N----------------------------------S----K-------------- 446
B.ES.10.DEMB10ES002         -------------T--------------------------C--------------E------------------------------------F------------------------------K--------------------------A----K----------S--- 446
B.FR.11.DEMB11FR001         K-------------------------------------------------------M-----------------------------------F---------------------------------------------------------A----K----------S--- 445
B.GB.05.MM45d213_GN1        KE----------------------------------------------------------------------------I---Q-------------------------------------------------------------------A-------------A----- 445
B.HK.06.HK003               -------------T-------------------------------------R-H------------------------I-------------F------------------------------E--D----------------------------K-------------- 450
B.HT.05.05HT_129389         K--------------------------------------------------X--------------------------X------E------F------------------------------X--N-X---------------------S----K-------------- 445
B.JP.12.DEMB12JP001         K------------L--------------------------C------------Y--------------------E----------R---G--F-----------------------------------------------------------------------A-G--- 449
B.KR.07.HP_18_07JHS10_3909  KN----------------------------------------------------------------------------I-V-----------FY---Q-------------------------M--------------------------A----K-------------- 446
B.PE.07.502_0525_wg5        KE-----------------------------------------------------E------------------KE--E----------K---------------------------------I--------------------------A----K-------------- 450
B.RU.11.11RU21n             P------------T---------------------------------D-------E-----------------------------K----------------------------------------------------------------A-L-----------A----- 450
B.TH.08.MERLBDTRC10         KE-----------V------V-----------------------R-----------------------------E-----------------F------------------------------R------------------------------IK-------------- 446
B.US.11.ES38                -S-----------------------------------------------------E----------------------A----------K--F------------------------------M--D-E--------------------------KH-------A----- 450

C.AR.01.ARG4006             ------------------------------------S--------------A---E----------------------A---K--E---K---------------------------------Q-------------------------------K-------------D 442
C.BR.07.DEMC07BR003         -S-------------------------------------------------A---E----------------------A------E---K--F----------------------------S-Q-----------------------------R----------A----D 442
C.BW.00.00BW5031_1          -N-----------V-------------------------------------AK--E------------------V---A-V----K------F------------------------------Q----------------------------------------A----D 448
C.CN.10.YNFL19              -E-----------R-----------------------------I-------A---E----------------------A------E------F------------------------------Q---------------------------------------------D 446
C.CY.09.CY260               -X--------------A-----------------------C----------A--------------------------X------X---K--F------------------------------Q----E----------------------------------------D 449
C.ES.08.X2363_2             -N----------------------------------S---A------D---AK--E----------------------A------N---K--F----------------------------A-K---------------------------------------------D 449
C.ET.02.02ET_288            -S-----------T-------------------------------------AK--E-----------------------------E---K--F------------------------------Q----E--------------------------K--------A----D 448
C.IN.09.T125_2139           -------------T------------------------------R------AR--E----------------------E------E---K---------------------------------Q--QQ---------------------------K--------A----D 443
C.KE.00.KER2010             KN-----------R-------------------------------------AK--E----------------------A------E------F------------------------------Q--D--------------------------R--H------------D 442
C.MW.09.703010256_CH256.w96 K--------------------------------------------------AK--E----------------------A------K---K--F----------------------------S-Q-------------------------------K-M------A----D 444
C.TZ.08.707010457_CH457.w8  ---------------------------------------------------VR-------------------------A------E---K--F------------------------------K----N---------------------S----K-------------D 442
C.YE.02.02YE511             KE-------------------------------------------------A---E------------------K---A-V----E---K--F------------------------------Q--D-------------------------------------A----D 440
C.ZA.08.705010534_CH534.w12 -N------------------------------------------R------MK--EM---------------------A------E---K--F------------------------------L---------------------------------------------D 449
C.ZA.10.DEMC10ZA001         -G-------------------------------------------------A---E----------------------A------E---K--F------------------------------Q--N-E---------------------A----K-------------D 442
C.ZM.11.DEMC11ZM006         -S-------------------------------------------------A---E----------------------E---K--E---K--F----------------------------A-Q----E--------------------------K-------------- 447

D.CM.10.DEMD10CM009         KE-----------T--------------------------C---------K----EL----------------------------E---K---------------------------------T--------------------------A--Q-KH-------A----- 451
D.CY.06.CY163               -------------T--------------------------C---------K----EL---------------------A---K--E-------------------------------------T----EN----------------------------------A----- 445
D.KE.11.DEMD11KE003         -------------------------------------------------------EM----------------------------E---K--F------------------------------K--D-ED-----------------------------RCI--A-T--- 445
D.KR.04.04KBH8              -------------------------------------------------------E----------------------A---K--E---K---------------------------------T----E---------------------A----K--------A----- 445
D.SN.90.SE365               -------------T---------------------Q---------------NK--E-----------------------------E------F-----------------------------VE----E---------------------A----KE-------A----- 451
D.TZ.01.A280                K------------T-----------------------------------------E-----------------------------D---K--F------------------------------K----E---------------------S----K--------A----- 447
D.UG.10.DEMD10UG004         -G---------------A-------------------------------------EM---------------------I------G---K--F------------------------------K----E--------------------------K---RC-K-A-T--- 445
D.UG.11.DEMD11UG003         -N-----------------------------------------------------EM---------------------I------E---K--F----------------------------T-T----E-------------------------------CI----T--- 447
D.YE.02.02YE516             -S-----------T--------------------------C--------------EL-------------------------K--E-------------------------------------T--Q-EN-------------------------K---R----A----D 447
D.ZA.90.R1                  ----------------------------------------C-------------------------------------A------E------F------------------------------K----E----------------------------------------- 445

F1.AO.06.AO_06_ANG32        KE-----------V--------------------------C----------TR--E----------------------A------E---K--F------------------------------Q--D--------------------------Q----------A-G--D 450
F1.AR.02.ARE933             -------------V--------------------------C----------TK---M----------------------------K---T--F------------------------------Q-------------------K-----------K--------A----D 446
F1.BR.10.10BR_RJ015         KE-----------T--------------------------C------D---T--------------------------A-V----K---S--F------------------------------Q----------------------------------------A----D 449
F1.CY.08.CY222              KE-----------L-------------------------------------VK-------------------------I------A---K--F------------------------------Q--D---------------------------I---------A----D 446
F1.ES.02.ES_X845_4          KE---------------K----------------------C----------IK-------------------------E------N---K--F------------------------------Q--D----------------------------K--------A----D 448
F1.RO.96.BCI_R07            KE-----------L--------------------------Y----------TR----I---------------L----A------T---K--F------------------------------Q--N----------------------------K--------V----D 446
F1.RU.08.D88_845            KE---------------------------------------------K---TK-------------------K-------V----D---K---------------------------------Q--D-E--------------------------K--------A----D 446
F2.CM.02.02CM_0016BBY       KE-------------------------------------------------AR--E--V-------------------A------E------F----------------------------A-Q--N-S------------------------R-KH-------A----D 446
F2.CM.10.DEMF210CM001       KE------------------------------------------R------TK-SE-----------------L----A-V----E---K--F----------------------------T-Q--D-S------------------------RIK-----I--A----D 451
F2.CM.10.DEMF210CM007       KE------------------------------------------H------AK--E-----------------------------G------F------------------------------Q----S------------------------R-KHI------A-T--D 445

G.BE.96.DRCBL               -N-----------T-----------#..-----------------------T---E----------------------A------E------F------------------------------Q--N-EN-------------------------K-----I--A----D 442
G.CM.10.DEMG10CM008         -S-----------V--------------------------C---R-------K--EM---------------------A------E---K--F------------------------------Q--D-E--------------------------K-----I-------- 445
G.CN.08.GX_2084_08          -N------------------------------------------X------IK--EM----------------T----A------D------F------------------------------K--D-E-------------------------IK--------A----D 450
G.CU.99.Cu74                K------------T-----------------------------------------E-E--------------------A----------S--F------------------------------Q--D-E--------------------------K--------A----D 442
G.ES.09.X2634_2             -S--------------A-----------------------C------K---IK----E--------------K-----A------K---S--F------------------------------Q--D----------------------------K---R----A----D 450
G.GH.03.03GH175G            -S-----------------------------------------I-------TK-------------------------A-V----E---K--F------------------------------Q--D-E--------------------------KH--R----A----- 446
G.KE.09.DEMG09KE001         -S-----------V------K------------------------------IK---M---------------------E------K---S--F------------------------------Q----E--------------------------K--------A----D 444
G.NG.09.09NG_SC62           -N-------------------------------------------------TK--EL---------------------A----------K--F--------------------------I---K--D-ED--------------------S----K--------A----D 445
G.ZA.01.TV546               K--------------------------------------------------TK--EM-----------------E---A------G---K--F------------------------------Q--D-EI-------------------------K---R---------D 445

H.BE.93.VI991               -------------T-----------------------------------------EVI--------------------E------A------F----Q------------------------VK-------------------------------K---X----A----- 444
H.CF.90.056                 KE---------------------------------------------A---E---EM---------------------A------A---K--F----------------------------TVK---------------------------N---K--------A----D 445
H.GB.00.00GBAC4001          -------------T--------------------------N--------------EM---------------------A------A---K--F----------------------------TVK----------------------------------------A----D 444

J.CM.04.04CMU11421          -X------------------X-------------------C----------TK--XL-----------------E---A--X---E---K--F------------------------------Q----EX-------------------X---R-K-M------A-S--D 442
J.SE.93.SE9280_7887         ----------------------------------------C------K---ER--E------------------E---R--K---E---K--F------------------------------Q----ED-------------------------K------K-A----D 444
J.SE.94.SE9173_7022         ----------------------------------------C------K---ER--E------------------E---R--K---E---K--FY-----------------------------Q----ED------------------------IKE----I--A----D 444

K.CD.97.97ZR_EQTB11         K---------------------------------------C----------RK---M-L-------------------A------E------F------------------------------Q--D----------------------F-----K--------V----D 446
K.CM.96.96CM_MP535          K-------------------V-------------------H----------IK--EM-------------------P--------E---K--F------------------------------Q--D----------------------------K--------V----D 446

U.CA.01.TV749               -N--------------------------------------C----------R---EL---------------------V------E---K--F----------------------------T-K----------------------------VR-KH--R----A----D 451
U.CA.99.TV721               ----------------------------------------C--------------EL---------------------V------E---K--F-----------------------------VT-------------------------------K---R----A----D 450
U.CD.83.83CD003_Z3          -N--------------------------------------------------E--E----------------------A------A------F------------------------------Q--N--D-----------------------Q---------------D 447
U.CD.90.90CD121E12          -N----------------------------------S--------------TK--E-A--------------------A------A------F------------------------------Q--D--D-----------------------Q----------A----D 446
U.CY.05.CY090               KE-----T-----T--A-----------------------C----------TK--E-------N------N-K-----A--K---T---K--F------------------------------Q--D-------N------------------RIKH--R----A----D 446
U.CY.08.CY223               -------------T---------------------------------D---AK--EM---------------------A-V----N---K--F------------------------------Q--D-----------------------A----K-----I-------- 447
U.ES.10.DEURF10DZ001        -E-----------V--------------------------A----------TK--EV---------------------A------E---K--F------------------------------Q--D--------------------------R-------I-------D 446
U.GR.99.99GR303             K------------T-------------------------------------IK--E-I--------------------A------E---K--FF-----------------------------Q----VN------TEVS------N---A--RD---------A----- 445
U.NL.95.U_NL_95_H10986_D1   PG--------------------------------------C--------Y-AK--EVI----------------E---A---G--D---K--F------------------------------Q--D-E--------------------F---R-KN-------A----D 446

01_AE.AF.07.569M            -S-----------T------------------------------------KIK--E-----------------------------A---S--F-----------------------R------E--------------------------A----X--------A----D 446
01_AE.CF.90.90CF11697       -G-----------------------------------------------Y-AK--E-----------------------------D---K--F-----------------------R------E--------------------------A----K--------A----- 444
01_AE.CN.10.YNFL03          -S-----------------------------------------I-------IK--EM---------------------V------A---S--F------------------------------T--------------------------A----K-------------- 445
01_AE.HK.04.HK001           KX-------------------------------------------------AN--E-I--------------------I------A---S--F-----------------------R------E--------------------------A---IK---R----A----D 443
01_AE.IR.10.10IR.THR48F     -S-----------T---K--------------------------------KIK--EMI--H------------------------A---N---------------------------------K-----H-------------------------K--------A----D 438
01_AE.JP.x.JRC77AE          -S-------------------------------------------------IK--EV----------------------------A---S--F------------------------------E--------------------------A----K--------A----D 441
01_AE.TH.09.AA090a_WG11     -S-------------------------------------------------LK-S-----------------------A------A---S--F-----------------------R------E-------------------------------K--------A----D 446
01_AE.TH.90.CM240           -S-------------------------------------------------IK--EM-----K----------------------A---S--F-----------------------R------E--------------------------A----K--------A----D 451
01_AE.US.05.306163_FL       -S-------------------------------------------------TK--E-I---------------------------A---S--F-----------------------R------E-------------------------------K-----I-------D 444
01_AE.VN.98.98VNND15        -S-------------------------------------------------IK--E-----------------------------A---S--F-----------------------R------E---R----------------------A----K--------A----D 445
```

polymerase motif   T215Y   K219Q

```
B.FR.83.HXB2                    EDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTE 446
02_AG.CM.10.DE00210CM013        K------------V--------------------------A----------AK--E----------------------A------E---K--F----------------------------H-Q--------------------------A----K---R----A----D 446
02_AG.CY.09.CY256               KE-----------V--------------------------A--------Y-TK--E----------------------A--X---X---X--F----------------------------H-Q--------------------------A----K--------A----D 446
02_AG.ES.06.P1423               -E------------------------X-------------A------X---TK--E------------------EK--A------G---K--F------------------------------Q--D-----------------------A----K--------A----D 446
02_AG.GW.05.CC_0048             KE------------------V-------------------A----------TK--E----------------------A------E---K---------------------------------Q--------------------------A----K---R----A----- 444
02_AG.KR.12.12MHR9              K------------V--------------------------A----------AE--EV---------------------A--DK--E------F------------------------------Q-------------------------------K--------A----D 450
02_AG.LR.x.POC44951             ----------------------------------------A----------TK--EM---------------------A---Q--D---K--F------------------------------Q--------------------------A--R-K--------A----D 446
02_AG.NG.09.09NG_SC61           P------------V--------------------------A----------AK--EV-V-------------------A------E-------------------------------------Q--D-E---------------------A-------------A----D 446
02_AG.NG.x.IBNG                 K------------V--------------------------A----------TK--E----------------------A------G---K--F-----------------------------VE--------------------------A---IK---R----A----D 446
02_AG.SE.94.SE7812              K---------------------------------------A----------TK--E----------------------A------E------F----E--------------------------------------------V--QVR--A----K--------A-T--D 446
02_AG.SN.98.98SE_MP1211         -------------V--------------------------A----------TK--EL---------------------A------E---K--F------------------------------Q-------------------------------K-------------D 446
03_AB.RU.97.KAL153_2            Q------------T-----------------------------------------E-----------------------E-----E------F---------------------------------------------------------A-------------A----- 446
04_cpx.CY.94.94CY032_3          PE-----------T--------------------------C----------FK--E----------------------A------E------F------------------------------QPA-----------------------------K--------A----D 445
05_DF.BE.x.VI1310               KE--------------------------------------C--------------EV---------------------A--K---E---A--F------------------------------Q--D----------------------------K-------------- 446
06_cpx.AU.96.BFP90              -K-----------T-----------------------------I-------IK--E----------------------A------E---K---------------------------------Q--D-E-----------------------------------A----D 448
07_BC.CN.98.98CN009             -------------T------------------------------------------------------------------------------F------------------------------Q----------------------------------------A----D 446
08_BC.CN.97.97CNGX_6F           K------------V--------------------------C----------------------------------K---------E---K--F------------------------------Q----------------------------------------A----D 442
09_cpx.GH.96.96GH2911           K------------V-------------------------------------IK--E----------------------A--K---E---K--F------------------------------Q--D-E---------------------A----K--------AR---D 445
10_CD.TZ.96.96TZ_BF061          -------------------------------------------------------QM---------------Q--H-TI------G---K--F-A------------------------L---Q--K-N-----------------------------------A-V--D 445
11_cpx.CM.95.95CM_1816          -S-------------------------------------------------T---E----------------------E-V----K------F------------------------------E--D-EC------------------------I--------------D 444
12_BF.AR.99.ARMA159             -A----------CV------T-----------------------------------------------------------------------F--------------------------------------------------------------K---R---------- 446
13_cpx.CM.96.96CM_1849          K------------V--------------------------C----------T---E----------------------E-V----R------F----------------------------A-Q--D----------------------------K--------A----D 453
14_BG.ES.05.X1870               -S-------------------------------------------------RK--EM---------------------A------K---N--F------------------------------Q--D-E--------------------------K--------A----D 445
15_01B.TH.99.99TH_MU2079        -S--------------------------------------C-----------K--EM----------------------------A---S--F-----------------------R------E--------------------------A----K--------A----D 445
16_A2D.KR.97.97KR004            -------------T------V----------------------------------E----R------------------------N---K--FY---------------E-------------K--------------------------AE---K--------A----D 446
17_BF.AR.99.ARMA038             KE------------------------------------------------------L--------------------------------K---------------------------------I-------------------------------K-------------- 446
18_cpx.CU.99.CU76               PE--------------------------------------------------E--E---------------------------------K--F------------------------------Q-------------------------------K-----I-------D 444
19_cpx.CU.99.CU7                -S---------------------------------------------D-------E---------------------------------K--F------------------------------S----E-----------------------------------A----- 445
20_BG.CU.99.Cu103               -------------------W-----------------------------------G---F----------------------------WK--FY-----------------------------Q--D------------------------------------------- 442
21_A2D.KE.99.KER2003            -------------------------------------------------------EM----------------------------E-------------------------------------K----E----------------------------------------- 445
22_01A1.CM.01.01CM_0001BBY      -G-----------T------------------------------R--D---AK--E------------------E---I--D---A---S--F------------------------------Q--N-----------------------A-------------A----- 446
23_BG.CU.03.CB118               -E--------------------------------------C--------------EL---------------------A----------K--F--------------------------------------------------------------K-------------- 442
24_BG.ES.08.X2456_2             K------------L--------------------------C------X-------EL---------------------A----------K--F-------------Q---------------------E------L----------------------------A----- 442
25_cpx.CM.02.1918LE             KE-----------T---------------------------T-----GT--AK--E-----------------------------E---K--F------------------------------K--------------------------A--Q----------A----D 445
26_AU.CD.02.02CD_MBTB047        KE------------------V------------------------------A---E----------------------A------A---K--F------------------------------K----EN--------------------A----K--------A----D 444
27_cpx.FR.04.04CD_FR_KZS        K---------------------------------------C----------T---E----------------------A------A---K--F------------------------------Q----E--------------------------K---R----A----D 444
28_BF.BR.99.BREPM12609          KE------------------------------------------R-----------------------------------V--------E--F-----------------------------V------I-------------------------K-------------- 446
29_BF.BR.01.BREPM16704          KE-----------V------------------------------------------------------------EK----------------F---------------------------------------------------------A----K-------------- 446
31_BC.BR.04.04BR142             -S---------------------------------------------------------------------------------------K--FD-----------------------------E--D-E----------------------------------------D 452
32_06A1.EE.01.EE0369            K---R--------------------------------------------------EM---------------------V------E---K--F------------------------------E--D-E-----------------------------------AR---D 446
33_01B.ID.07.JKT189_C           -S--------------------------------------R---R------IK-------------------D------------A---N--F------------------------------E-------------------------------K--------A----D 446
34_01B.TH.99.OUR2478P           K---------------------------------------C---------------M----------------------------D------F------------------------------E--------------------------A----K--------A----D 446
35_AD.AF.07.169H                K------------------------------------------I-----------E-I--------------------I------A---K--F------------------------------T----E----------------------------------------D 445
36_cpx.CM.00.00CMNYU830         K---------------------------------------A----------TK--E----------------------A------K---N--F------------------------------Q-----N--G-----------------A----K--------A----- 446
37_cpx.CM.00.00CMNYU926         -S-----------T--------------------------C----------A---E----------------------A------D---K--F------------------------------Q--D--I--------------------A--------R---------D 444
38_BF1.UY.03.UY03_3389          K--------------------------------------------------Q----------------------E-----------------F--------------------------------------------------------------K-------------- 450
39_BF.BR.04.04BRRJ179           -E-----------V----------C------------------------------E------V-----A-----K---A---------W---FY------------H----------------M-------------------------------K-------------- 449
40_BF.BR.05.05BRRJ055           KE--------------------------------------C------D---AK--EL----------------------------E---K--FF--EQ-------------------------M--D-------------------------------------A----D 446
42_BF.LU.06.luBF_18_06          K------------T-----------------------------------------E----------------------I--------------------------------------------Q--N-----------------------A----K-------------- 450
43_02G.SA.03.J11223             -S-----------T---------C---------------------------TK--E------V----------L----A------E---X--F------------------------------Q--D-E--------------------------K---R----A----D 445
44_BF.CL.00.CH80                KE-----------T------V--------------------------------------------------------------------K-----------------------------------------------------------------K---R----P----- 446
45_cpx.FR.04.04FR_AUK           K--------------------------------------------------TR--EM---------------------A------E---K--F----Q-------------------------Q--D-E---------------------S----K--------A----D 444
46_BF.BR.07.07BR_FPS625         Q------------V--------------------------C----------AK--E-----------------------------N---K--F------------------------------Q--N-E------------------------R-K--------A----D 446
47_BF.ES.08.P1942               -E-----------T---------------------------------------------------------------------------K-----------------------------------Q---N-----------------------------R---------D 445
48_01B.MY.07.07MYKT021          K------------V---A----------------------C--------------E-I---------------------------A---N--F------------------------------Q-------------------D------A--Q-K-------------D 446
49_cpx.GM.03.N26677             -N-------------------------------------------------T#--E-----------------------------A---K--F------------------------------Q--D-EA-------------------------K-----I--A----D 443
50_A1D.GB.10.12792              ----------------------------------------A----------SK--E-I--------------------A------A---S--F-----------------------------------E---------------------A----K-----I-------- 444
51_01B.SG.11.11SG_HM021         K----------------------------------------------D---------I----------------E-----V-----------F------------------------------E--------------------------A----K-------------- 446
52_01B.MY.03.03MYKL018_1        -S-----------V---A----------------------C---------------------------------------------------F-----------------------R------E-------------------X------A---XK--------A----D 446
53_01B.MY.11.11FIR164           -S-------------------------------------------------IK--E-----------------------------A---S--F-----------------------R------E---R----------------------A----K--------A----D 446
54_01B.MY.09.09MYSB023          -------------V--------------------------C---------------------V-----A--------------------S--FF---Q-------L----------R------E--------------------------A----K--------V----D 446
55_01B.CN.10.HNCS102056         -S-----------R-------------------------------------------E-------------------------------K---------------------------------A--------------------------A----K--------A----D 446
56_cpx.FR.10.URF5_patient_A     KE-----------T-----------------------------I--------H---------------------R-----V--------K---------------------------------Q--D------------------------------------------- 448
57_BC.CN.09.09YNLX19sg          ----------------------------------------H---R------A----------------------E---A------E---K--F------------------------------M---------------------------------------------D 442
58_01B.MY.09.09MYPR37           -S-----------V-------------------------------------IK--E------------------------V----A---S--F------------------------------M-----N--------------------A-----E-------A----- 446
59_01B.CN.09.09LNA423           -S-----------R-------------------------------------IK--E-I--------------------I-V----A---S--F-----------------------------VE--------------------------A----K---RV---A----D 443
60_BC.IT.11.BAV499              ------X------------------------------V--G----------------I--------------------A------K---Q---A--------------T--------------Q--D-------------------------------------A----D 441
61_BC.CN.10.JL100010            K-----------------------------------------------------------------------------A------E---K--F------------------------------Q----------------------------------------A----D 442
62_BC.CN.10.YNFL13              K------------L--------------------------C---------------------------------A---A----------K--F---------Q--------------------Q----------------------------------------A----D 442
63_02A1.RU.10.10RU6637          K------------L--------------------------A----------RK--E----------------------A------A---N--F------------------------------Q--D-------------------------------------A----D 446
64_BC.CN.09.YNFL31              ---------------------------------------------------A---E----------------------A------E---K--F------------------------------Q--------I-------------------------------A----D 446
65_cpx.CN.10.YNFL01             ----------------------------------------C----------------D--------------------A------E---T--F----------------------N-------Q-------------------------------------I--S----D 442
67_01B.CN.11.2011.ANHUI.HF115   -S-------------------------------------------------NK-------------------------I----------K--------------------------R------E-------------------------------K-------------- 444
68_01B.CN.11.2011.ANHUI.WH73    -N------------------------------------------R---------------------------------I----------K--------------------------R------E-------------------------------K-------------- 444
70_BF1.BR.10.10BR_PE004         K------------V--------------------------C----------AK---M---------------------A------E---K--F------------------------------Q--N-------------------------------------A----D 448
71_BF1.BR.10.10BR_PE008         KE-----------V--------------------------C------D---AK-------------------------A------E---K---------------------------------Q-------------------------------K--------A----D 446
72_BF1.BR.10.10BR_MG002         K---------------------------------------C----------AR--------------------------------E---K--F------------------------------Q--------------------------G----K---R---------D 447
O.BE.87.ANT70                   P------------V---------------------------------D---RD--ELE-C------------PLTE--KR--L--E--YQ--F----------------------------S-Q--N--V---------I----------Q--R--E----I----S--- 442
O.CM.98.98CMA104                P------------V------V--------------------------D----D--ELE-C------------PLTE--KRV-L--E--YK--F------------------------------Q--D-EV---------I----------Q--R--E----I----S--- 442
O.CM.98.98CMABB141              P---P--------V--A------------------------T-I---D----D--ELE-C------------PLTE--KRV-L--E--YQ--F------------------------------Q--S--V---------I----------Q--R-KE----I----S--- 442
O.CM.98.98CMABB212              S------------V---------------------------P-----D----N--EVE-C------------PLAE--KRV-L--E--YQ--F------------------------------Q--N-EV--------------------Q--R-KE----I----S--- 442
O.CM.98.98CMU5337               P---X--------V---------------------------------D----D--ELE-C------------PLTE--NRV-K--E--YQ--F------------------------------Q--N-EV-----------------------R--E----I----S--- 442
O.CM.99.99CMU4122               P------------V-----------------------------V---D----D--ELE-C------------PLTE--KRV-L--E--YQ--F------------------------------Q--N-EI-----------------------R--E----I----S--- 442
O.FR.92.VAU                     PE-----------V------V--------------------------D----N--ELE--------------PLAE--KRV-L--K--YK--FY------------Q----------------Q--D-EV--------------------H--R-KE----I----S--- 443
O.GA.11.11Gab6352               PN-----------L------V----------------------V---D----D--ELE-C--V---------PLTE--KRV-L--E--YK--FY-----------------------------Q--D-EV---------I----------Q--R-KE----I----S--- 442
O.SN.99.99SE_MP1299             P------------V---------------------------------D----D--ELE-C------------PLTE--KRV-L--E--YQ--F------------------------------Q--N-EE--------------------Q--R-KE----I--A-S--- 442
O.US.10.LTNP                    P------------V------V--------------------------D----D--ELE-C------------PLTE--KRV-S--E--YQ--F------------------------------Q--N--V---------I----------Q--R-KE----I----S--- 442
N.CM.02.DJ00131                 -------------T--------------------------------------KH-E-I-------------S-LA---ETV----S---K--F----------------------------T-K---Q-V-------------------------K---R-I-------- 451
N.CM.02.SJGddd                  K---------------------------------------A-----------KH-E-I----------------A---ETV----R---E--F------------------------------K-----V-------------------------K-----I--A----- 448
N.CM.04.04CM_1015_04            -------------V------------------------------T-------KH-E-I--------------DLA---ETV----G---K--F------------------------------K-----T-------------------------------I-------- 453
N.CM.06.U14296                  -N---------------------------------------------D----KH-E-I---------------LA---ETV-K--E---Q--F------------------------------K-----V-------------------------K---R-I-------- 453
N.CM.06.U14842                  P------------V------V-------------------------------KH-E-I---------------LA---EAVQ---D---K--F-----------------------------VK----EV-------------------------K---R-I-------- 454
N.FR.11.N1_FR_2011              KE------------------V--------------------T----------KHSE-I---------------LT---EAV----G---K--F------------------------------K-----V-------------------------KH--R-I-------- 453
P.CM.06.U14788                  P------------V--D---L--V----------------H-----------S--EVE--------------PLSE--QRV-K--E--YV--F------------------------------K----E------------------------R-KE----I----S--- 443
P.FR.09.RBF168                  P------------V--------------------------Y----------ES--EVE--------------PLSE--QRV-K--E--YV--F-----------------------------------EA-----------------------R-KE----I----S--- 442
CPZ.CD.06.BF1167                PE-Q---------V---R--K--C-T--------------Y------D----KH-E-DV-------------TAEK-KEVVQ---N--KA---E-----Y-EK------------E-----Q-K---PET-----------------------RT-E----I--V-G--D 449
CPZ.CM.05.SIVcpzMT145           P----F------------------------------S--------------TKH-E-I--------------P-EE--KRV----A---K--F----------------------------Q-K---QEV------------------------IK-----I--A-G--- 445
CPZ.GA.88.GAB1                  K-------------------V---------------S--------------EK----T----------------D---K-V--------K--F------------------------------Q----EV---------I--------------IK-----I----K--D 444
CPZ.TZ.06.SIVcpzTAN13           P------------L--N---K-F------------------------D----KH-EVDV------------YSEEE--KLMK-------A---E-----Y-EK------------------N-T---PEQ----H---------------A---TKE----I--V-G--D 441
CPZ.US.85.US_Marilyn            KE-----------V---------------------------------D-----H--VI--------------NLEK--E-V-L---Y--T--F--------E---------------------Q--Q-EI------------------------IK-----IK-A----- 448
GOR.CM.07.SIVgor2139_287        P------------V------V-------------------H------D----K--E-E--------------P-TE--Q-V-K--E--YA--F------------------------------K----E--------R---------------R-KE----I----S--- 438
GOR.CM.07.SIVgorCP2135con       P------------V------V-------------------H------D----K--E-E--------------P-TE--Q-V-K--E--YA--F------------------------------K----E------------------------R-KE----I----S--- 438
```

```
                                                                                                                                      p51 RT end  p15 RNase H start
B.FR.83.HXB2                 VIPLTEEAELELAENREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPF.KNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTN 615
A1.AU.03.PS1044_Day0         -VT-----------------------------E----------D-----------.----------K-S-X----R--A-V---VVM---------------------D---MD---------------------------D---------------------------D 615
A1.CH.03.HIV_CH_BID_V3538    IVT-----------------D-----------E-V-------PD-----------.---------KKGS------R---AV---VA-------------R----R----A--M----------------------------D--I----------------------I-D 615
A1.CY.08.CY236               IVT-----K----------R------------E-V--------E-----------.----------K-S---------A-V----AM----------------------L--MD---------------------------D--A------------------------D 612
A1.ES.05.X1608_8             -VT-----------------D----------------------D------F----.---------KKGS------R--ANV---VVM------------R------------MD---------------------------D--A------------------------D 619
A1.KE.11.DEMA11KE001         -VT-----------------D----T------E----------D-----V-----.---------KKS------------V---VVM--------I------------D---MD---------------------------D--E------------------------D 615
A1.RU.11.11RU6950            IVT----------------R-----L--------V--------D-----------.---------KKGS----------AV---VA-------------R---------A--M-----------------------------------------------RI-------D 615
A1.RW.11.DEMA111RW002        -VT-------------------------------V--------D-----------.----------K-S-----------V---V-M-------------------------MD---------------------------D----------------G----------D 615
A1.SN.01.DDI579              IVT-----------------D----A-----------------N-----------.----------K-SS-----R--V-V---VV-------------R------------M------Y---------------------D--I------------------------D 614
A1.UG.11.DEMA110UG001        IVT-----------------D----------------------D-----------.----------K-S------R----V---VV---------------------------D---------------------------D--------------S------------D 615
A1.ZA.04.04ZASK162B1         -V------------------N---------A------------D-----------.---------KK-S-----------V---VV--------------------------M-C--------------------------D--L------------------------D 614
A2.CM.01.01CM_1445MV         IV---R------E-------N-------------T--------D-----------.---------KK-TT---------------VM------------R----------------------------------------Q----------------------------D 615
A2.CY.94.94CY017_41          IVT--K------E-------T----------------------D-----------.---------KR-ST----I-----------M-------------------------A-------------------------------A------------------------D 615
B.BR.10.10BR_MG029           -----K-------------------------------------------------.----------T-----------A------V-------------R-------------------------------------------------------------S-------D 615
B.CA.07.502_1191_03          -V---T-------------R----------T------L-----------------.-----------------------------A-----------------------------------------------------------------------------------D 615
B.CH.08.M2_0803101_NFLG8     -V--SR--------------T-----F----------------------------.-------------T---------------A-------------R----R---D--------------------------------------------------S---------D 619
B.CN.10.DEMB10CN002          -V---R---I--------------------------------E-----------Y.-------H---------------------A-Q-----------R---------A------------------------------------------------D----------D 615
B.ES.10.DEMB10ES002          -----------------------------------------------------Q-.---------K---------R---------AM------E-I---R------------I----------------------------D----------------D----------D 615
B.FR.11.DEMB11FR001          -----Q-----------------------------------EH------------.-----------------------------AQ------------R---------A-----------------------------------------------------------D 614
B.GB.05.MM45d213_GN1         -----------------------------------------EC-----------Y.------------S-----------V----A----------------------QA----------------------------------I------------------------D 614
B.HK.06.HK003                -----K-------------------------------L-----------------.------------------I----------A-------------R--------X------------------------------------------------------------- 619
B.HT.05.05HT_129389          -V----------------------------------------H------V-----.-----------------------------A--G------I---R--------DI-----------------------------------------------------------D 614
B.JP.12.DEMB12JP001          -V---A-------------R--------------V-------L------------.-----------K------I----------A--------R----R---------A--M--------------------------------------------------------D 618
B.KR.07.HP_18_07JHS10_3909   -V------------------V----------------------------------.---------K-------------------AI----------------------A---------*-----L--------------------------------------------- 614
B.PE.07.502_0525_wg5         --T---------------------------------------L---S--------.---------K-------------------A-----------------------A----------------------------------L------------------------D 619
B.RU.11.11RU21n              -----A----------------------------V--L-----------------.-----------------------------A---------------------------------------------------------------------------------I-D 619
B.TH.08.MERLBDTRC10          IVS----------------R-----------------------------------.-------------T---------------A--------------------A--A---D------------------------------E----------------I-------- 615
B.US.11.ES38                 -V--SK------------I------------------------------------.---------K--S--------------------------I---R--------------------------------------------L----------S-------------D 619
C.AR.01.ARG4006              IV---A-----------------------------------E-------------.---------K--T------R---------AQ---I--------R---------S---D------------------------------T----------------M-------D 611
C.BR.07.DEMC07BR003          IV----------------------------------------E------------.---------KT-T------R---------AQ------------R---------S---D-----------------------------------------------M-------D 611
C.BW.00.00BW5031_1           IV-----------------R----------------------DD-----------.---------KI-T------R---------AL--------I---R-------------D-----------------------------VA-V--------------I-------D 617
C.CN.10.YNFL19               IV-------------------T---A----------------LD-----------.---------K--T---------A-------M------------R------------AD---------------------------D--A-V--------------K-------D 615
C.CY.09.CY260                IV----------------------------------------ED-----------.---------KV-T----------------AK-C----------R-------------D------------------------------A-------------D--I-----I-D 618
C.ES.08.X2363_2              IV----------------------------------------E------------.---------KRKA----------------SM------------R--------D----D------------------------------I------------------R-----D 618
C.ET.02.02ET_288             IV----------------------------------------DD-----------.---------KR-T-------L--------AL--------I---R-------------D------------------------------A----------------I-------D 617
C.IN.09.T125_2139            IV-----------------------A-----------------------------.---------K--T----------------AL--------------------------D---------------------------D--A-V--------------I-------D 612
C.KE.00.KER2010              IV--------------A-------------------------HD-----------.---------K--T----------------A--------R----R-------------D---------------------------D--A----------------V-------D 611
C.MW.09.703010256_CH256.w96  -V----------------------------------------DD-----------.---------KQ-T----------------CM--------I---R-------------D------------I-----------------A----------------K-------D 613
C.TZ.08.707010457_CH457.w8   IV----------E------R-----A----------------HD-----------.---------K--S----------------AM------------R-------------D-----------------------------------------------I-------D 611
C.YE.02.02YE511              IV----------------------------------------YD-----------.---------KI-T----------------AM--------I---R-------------D-----------------------------M--V--------------M-------D 609
C.ZA.08.705010534_CH534.w12  IV----------------------------------------LD-----------.---------K--T---------A------AM------------R----R----I---D------------------------------A------------------------D 618
C.ZA.10.DEMC10ZA001          IV----------------------------------------ND-----------.---------KR-T---------------VAQ-C----------R-------------D------------------------------I-----------S-D--I-------D 611
C.ZM.11.DEMC11ZM006          -VT---------------------------------------HD-----------.---------K--T------R---------AM-C----------R-------------D-------------------------------------------------------D 616
D.CM.10.DEMD10CM009          -V-----------------------------------------S---W-------.---------K---------R---------AM------------R---------A--R----------------------------D--A------------------------D 620
D.CY.06.CY163                IV----------------------------------K------S---W-------.---------K-------------------AI------------R---------A--M-------------------------------I----------S-I-----------D 614
D.KE.11.DEMD11KE003          IV---------------------------Q-------------D---------QY.---------K--T-----------------H-C----------R--------D------------------------------------------------------------D 614
D.KR.04.04KBH8               -----K-------------------------------L-----------------.---------K--T----------------AQ-------R----R---------A------------------------------R---I---------------R--R-----D 614
D.SN.90.SE365                IV----A------------------------------------D-----------.---------KR-S----------K------Q------------R---------A--I--------------------------------------------------------D 620
D.TZ.01.A280                 -----K-------------R-T---A---------------E-------------.---------K--T----------------AL-G------I---R--------------------------------------------I---------------R--------D 616
D.UG.10.DEMD10UG004          I------------------R----------A------------D---------QY.----------K--T---------------AQ-C----------R-------------D------------------------------T------------------------D 614
D.UG.11.DEMD11UG003          -V-----------------------------------V----PD---------QY.---------KL--T--------AA-----SQ-C----------R---------I---D------------------------------I------------------------D 616
D.YE.02.02YE516              -----R-------------R-----------------V---------W-------.------------------------V----AQ------------R---------A--M-------------------------------I------------------------D 616
D.ZA.90.R1                   -V----------------------------------------D------------.----------T-----------A----NLS-------------R-------------------------------------------------------------V-------D 614
F1.AO.06.AO_06_ANG32         IV---A------------------------------------HE-----------.---------KR-S---------------VAM-----------------------------------------------------T---A------------K---K-------D 619
F1.AR.02.ARE933              IV---A------------------------------------------------L.------------S----------------AQ----------------L----D---ISI-------------------------T---A----------S-----K-------D 615
F1.BR.10.10BR_RJ015          -V---T------E------------A-----------L-----------------.---------K--S---------------VSL--------I---R---L----D-------------------------------T--------------S--G--E-------D 618
F1.CY.08.CY222               IV---A----K-----------------------V-------------------C.---------K--S----------------VQ--------I---------K---A-*ID--------------------------T---T----------S-----I-------D 614
F1.ES.02.ES_X845_4           IV---A-------------------A---------V-------------------.---------K-KS----------------AL----------------------A------------------------------T---A-------------D--K-------D 617
F1.RO.96.BCI_R07             IVT--A-----------------------------------------------T-.---------K--S----------------A--C--------------------A----------------A-------------T---A----------S-----Q-------D 615
F1.RU.08.D88_845             IVT--A------------------------------------------------Y.---------K--S------R---------AV-------------------------A---------------------------T---PE---------S-----K-------D 615
F2.CM.02.02CM_0016BBY        -V---A------------------------------------HD----------H.R--------KR-S-----------V----A--G------V---R----------------------------------------T---I-----------S----I-------D 615
F2.CM.10.DEMF210CM001        -V---A--------------------------E----------D---------TH.---------KR-S-----------V---VA---------I---R------------A---------------------------T----------------------------D 620
F2.CM.10.DEMF210CM007        -VS--A-----M------------------------------HD----------H.---------KR-S----------DV----A--G-----RV---R----------------------------------------TD----V----------------------D 614
G.BE.96.DRCBL                IVSM-A---M---------------------------V----L------V----Y.---------KGGS-----------V----A--G-I----I-------K-----V------------------------------T---P-V--Y-------------------D 611
G.CM.10.DEMG10CM008          IV---A--------------D----------------V----VE----------Y.---------KRGS-----------V----A----I----I---R---R----DA---D--------D--------------R--T---P----Y-------------------D 614
G.CN.08.GX_2084_08           IV----------T------------------------V----PD----------Y.---------KRGS----------------A----I----V-------R-----V---D--------D-----------------T---X----Y-------------------D 619
G.CU.99.Cu74                 IV------------------------------EI---V----ND----------Y.---------KRGS---------V-V----S-----------------R-----V---D--------D--------------R--T---L----Y-------------------D 611
G.ES.09.X2634_2              IV------------------------------E----V----LD----------Y.---------KRGS-----------V---VS----I--------R---RR-----------------D--------------R--T---A-----F-----S------------D 619
G.GH.03.03GH175G             IVS--A---M----------------------E----V----LD----------Y.---------QRGST----------V----A--G------V-------R-----V------------------------------T---P----Y-------------------D 615
G.KE.09.DEMG09KE001          IV---T-------------------------------V----LD----------Y.---------KRGS-----------V----A-----------------R-----V------------D-----------------A---P----Y-----------V-----I-D 613
G.NG.09.09NG_SC62            IV-----------------R---------S--E----V----EP----------Y.-----E---KRGS-------E---V----SI--------I---R---K----DV------------D--------------R--T---A----Y--------K----------- 614
G.ZA.01.TV546                IV---A--------------------------E----V----LE----------Y.---------KKGS-----------V----AV--------I-------R-----V------------D--------------R--T---P----Y-------------------D 614
H.BE.93.VI991                IV---K-------------------A------E---------PD-----------.---------K--S-----------V----A---------I---R--------------H---------------H---------T---E----Y-----------M-------D 613
H.CF.90.056                  I----K-------------R--I---------------R----------------.---------K--T-----I----------S---------I---R------------------------------H---------T---A----Y-I-----------------D 614
H.GB.00.00GBAC4001           I----T------E------R-----------------------E-----------.---------K--T-----I----------C---------I---R------------------------------H---------T--------Y-----------I-R-----D 613
J.CM.04.04CMU11421           IV---K---------K---X-----X----A-E----V----ED-----------.---------KR------------------AL------------R-------------D------------------------------M---------------XK-----I-D 611
J.SE.93.SE9280_7887          IV---R---------K-------------SA-E----V----LD-----------.---------KR-S---------A-V----AL-A----------R----R--------D------------------------------M----------S-----T-------D 613
J.SE.94.SE9173_7022          IV---R---------K--------------ARE----V----LD-----------.---------KR-S---------SQV----AL-A----------R-------------D------------------------------M----------S-----V-------D 613
K.CD.97.97ZR_EQTB11          IV---A------------------------------------H-----------Y.----------I-S-----------V---VAM------------R--------G-------------------------------T--------------------Q-------D 615
K.CM.96.96CM_MP535           IV---A------------------------------------ND----------H.------------S----------------A--G----------R----------------------------------------T---------------H----K-R-----D 615
U.CA.01.TV749                -VT---------------------------------------ND-----------.---------KRKS----------N-----AM-C-----R----R--------DA--M----------------------------D--M------------------------D 620
U.CA.99.TV721                IVT---------------------------------------ND-----------.---------KSKST---------A-----AM-C----------R--------DA--M----------------------------D---------------------------D 619
U.CD.83.83CD003_Z3           IV-M-K----------------------------V--------D--S--------.---------KR-S----------------SM--------I-------------S---D---------------------------D--M----------------K-------D 616
U.CD.90.90CD121E12           IV-M-K-------------E-----------------------D--S--------.---------KR-S----------------SM-------------------------AD------------------------------A-----------S----K-------D 615
U.CY.05.CY090                IVT-----K--------------------------------H-------------.---------KI-S--P--I----------S-------------R--------DI-G----------------------------T---A------------I---E-------D 615
U.CY.08.CY223                -VT----------------------A----A---V--------D----------H.---------KR-S-----IR--V------AQ-G------V---R---------A--M---------------------------R---A------------------------D 616
U.ES.10.DEURF10DZ001         IV---A---------------A---A----------------E-----------H.----------IKA----------------AL------------R---------I---D------------------------------I------------------------D 615
U.GR.99.99GR303              IV---A----------D-------------------------H-----------Y.----------V-S-----I-----V------------------R-----G-------D----P-------.------------D-G-------------------R-------D 613
U.NL.95.U_NL_95_H10986_D1    IV---A-------------------------------------D----------Y.----------RKS----------------AM------------R-----------------------------------------D--E----------------M-------D 615
01_AE.AF.07.569M             IV---X-------X------T-------------V--V-----D-----------.----------K-S-----------V----A-------------R------------M----------------------------D--------------S------------D 615
01_AE.CF.90.90CF11697        -V------------------N-----------E-V--------D-----------.----------K-S------R--A-V---VA-------------R----R-------A----------------------------D--A----L------S--N---------D 613
01_AE.CN.10.YNFL03           -V------------------S----A-----------V-----D---------Q-.----------K-S------R--A-V----A-------------R----R------------------------------------D--------------S------------D 614
01_AE.HK.04.HK001            -V------------------T-------------V--V-----D-----------.----------K-S------R--A-V----V---------A--------R-------M----------------------------D--------------S------------D 612
01_AE.IR.10.10IR.THR48F      IV------------------T-------------V--V-----D-----------.----------K-S------R----V----A-------------R------------M----------------------------D--------------S------------D 607
01_AE.JP.x.JRC77AE           I-------------------T-------------V--V-----D-----------.----------K-S------R----V----A-------------R----R------------------------------------D--------------S--N-I-------D 610
01_AE.TH.09.AA090a_WG11      IV------------------T-------------L--------D-----------.----------K-S------R----V---VAM------------R----R-------M----------------------------D--------------S------------D 615
01_AE.TH.90.CM240            IV------------------T-------------V--V-----D-----------.----------RGS------R----V---VA-------------R----R-------M----------------------------D--------------S------------D 620
01_AE.US.05.306163_FL        -V------------------T----------------V-----D-----------.----------K-S-----------V----V-------------R----R-------M----------------------------D--------------S------------D 613
01_AE.VN.98.98VNND15         IV------------------T-------------V--V-----D-----------.----------K-S------R----V----A-------------R------------M----------------------------D--------------S------------D 614
```

```
                                                                                                                                                 p51 RT end  p15 RNase H start
B.FR.83.HXB2                  VIPLTEEAELELAENREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPF.KNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTN 615
02_AG.CM.10.DE00210CM013      IV----------------------------T---V--V-----D------F----.---------KK-S-----------V---VA-------------R----R----A--M----------------------------D--A------------------------D 615
02_AG.CY.09.CY256             IVT-S-------------------------T---V--------D-----------.---------KK-S-----------V---VAL------------R----R----A--M----------------------------D--P------------------------D 615
02_AG.ES.06.P1423             -V----------------------------A---V--------D-----------.---------KK-S---------V-V---VV-----------R-R----#----A--M----------------------------D------------------S--------D 614
02_AG.GW.05.CC_0048           IVT---------------------------T---V--------D-----------.---------KKKS-----------V----VM------------R--------DA--M---------D------------------D-------------------I-------D 613
02_AG.KR.12.12MHR9            IV----------------------------T---V--------D-----------.---------KK-S-----------V---VA-------------R---------A--M----------------------------D--A------------------------D 619
02_AG.LR.x.POC44951           IVT-----------------K---------T---V--------D-----------.---------KK-S-----I-----V---VA-------------R----R---DA--M----------------------------D---------------------------D 615
02_AG.NG.09.09NG_SC61         IV----------------------------T---V--------D-----------.---------KK-S-----------V---VSM------------R--------DA--M---------D---------------------A------------------------D 615
02_AG.NG.x.IBNG               IVA---------------------------T---V--L-----D-----------.---------KK-S-----------V---VAM------------R----R-------M----------------------------D-------------------I-------D 615
02_AG.SE.94.SE7812            IVT---------------------------T---V--------D-----------.---------KK-S-----------V---VA-------------R----R----A--M----------------------------D-------------------I-------D 615
02_AG.SN.98.98SE_MP1211       IV----------------------------T------------D-----------.---------KK-S-----------V---VA-----V-------R----R----A--M----------------------------D--I------------------R-----D 615
03_AB.RU.97.KAL153_2          -----A----------------------------V--------------------.----------L------------------A---------------------------------------------------------------------------S-------D 615
04_cpx.CY.94.94CY032_3        IV---T-------------------A----------------------------H.----------T-S------R---------AM-C----------R--------D-------------------------------TD--A----------------Q-------D 614
05_DF.BE.x.VI1310             -------------------------------------V----L------------.----------A--V--------A------A-------------R---------I----------------------------------L----------------------I-D 615
06_cpx.AU.96.BFP90            IV---A-------------------------------V----Y-----------H.----------IKS----------------AL------------R----------------------------------------T--------------------Q-------D 617
07_BC.CN.98.98CN009           IV-----------------------I-----------------D-----------.---------K--T----------------AM------------R---------I---D---------------------------D----V--------------I-------D 615
08_BC.CN.97.97CNGX_6F         IV-----------------------A------E----------D-----------.---------K--T----------------AM--------I---R-------------D---------------------------D--A-V--------------I-------D 611
09_cpx.GH.96.96GH2911         IVT-------------------------------V--------D-----------.----------K-S------R----V---V---------------------------A----------------------------D--I------------------------- 614
10_CD.TZ.96.96TZ_BF061        IV-----------------------------------------D----------H.---------KR-T----------------AQ-C--------------------I---D--------D----------------------------------------------D 614
11_cpx.CM.95.95CM_1816        IV---A-----------------S--S----V----LD-----------------.---------KR-T------R----VI---C---------I---R----R--------D------------------------------I----------------V-------D 613
12_BF.AR.99.ARMA159           -----K-------------------------------V------------F----.-----------------------------A-----------------L----D-------------------------------T---AE---------S-K---K-------D 615
13_cpx.CM.96.96CM_1849        IVS--T------------------------E-E--V-------E----------Y.---------KRGS-----------V---VA---------L-------R-----V---D--------D--------------R--T---L----Y-------------------D 622
14_BG.ES.05.X1870             IV---V----------------------E---E----V----PD----------H.---------KRGS-----------V----A----X--------R---R-----V------------D-----------------T--------Y-------------------- 614
15_01B.TH.99.99TH_MU2079      IV------------------A----A--------V--V-----D-----------.----------K-S------R----V----A-------------R------------M----------------------------D--------------S------------D 614
16_A2D.KR.97.97KR004          IV---K------E-------D----------------V----PD-----------.---------KR-ST---------------VM--------I---R----------------------------------------T---T------------------------D 615
17_BF.AR.99.ARMA038           IV---K-------------------------------------------------.-----------------------------A-------------R---L----D-------------------------------T---AE---------S-----K-------D 615
18_cpx.CU.99.CU76             IV---V----------------------------V-------YD-----------.---------KR-S---------V-V---VA-----------------------------------------S--F---------T---A----------------I-------D 613
19_cpx.CU.99.CU7              -------------------------------------V-----D-----------.---------K-------------------SL--------I---R---------A--R-------------------------------M------------------------D 614
20_BG.CU.99.Cu103             -V---T-------------------------------------------------.-----------KS---------A------AQ-------------------------M--------------------------------------------------------D 611
21_A2D.KE.99.KER2003          -------------------------------------------D-----------.---------KK-ST---------------AM-A----------R------------M--------------------------------------------------------D 614
22_01A1.CM.01.01CM_0001BBY    -V-----------------R-----------------------D-----------.----------K-S-----------V---VA--G----------R---------A--M----------------------------D--I-----------S------------D 615
23_BG.CU.03.CB118             -V---A-------------------------------------L-----------.-----------KST----------V----AQ------------R---------A--M--------------------------------------------------------D 611
24_BG.ES.08.X2456_2           -V---T-------------------------------L-----------------.---R-------KN----------------AQ--------I---R---------A--M-H--------------------------------------------D---------D 611
25_cpx.CM.02.1918LE           IV---V--------------N-----------E----V-----D----------Y.---------KGGS---------AA-----A--G------I---R-------------D------------------------------I------------------R-----D 614
26_AU.CD.02.02CD_MBTB047      IVT--K--------------D----------------------------------.---------KRKS---------A-V---V-M------------R--------D--------------------------------D--M------------------------D 613
27_cpx.FR.04.04CD_FR_KZS      IVS----------------R-----------------V-----D-----------.---------KR-C-----I---A-VM---A-----------R-R-------------D--------------------------Q--------------------I-------D 613
28_BF.BR.99.BREPM12609        -V---------------------------------V-------D-----------.----------T-----------------VA---------I---T------------I-------------------------------I----------------Q-------D 615
29_BF.BR.01.BREPM16704        -V---K------------I-----------A------------N-----------.-------------T----------------------------------------------------------------------T---E----------S-----K-------D 615
31_BC.BR.04.04BR142           IV-----------------------A-----------------D-----------.---------K--T------R----------L------------R----R--------D-----------------------------------------------M-------D 621
32_06A1.EE.01.EE0369          IV---A-------------------------------V----ND----------H.---------KIKS----------------AM----V-------R----------------------D-----------------T-----V--------------K-------D 615
33_01B.ID.07.JKT189_C         IV-------------K----T-------------VV-V-----D-----------.----------KSS------R----V----A--------------------------M----------------------------D--M-----------S------------D 615
34_01B.TH.99.OUR2478P         IV------------------T-------------V--V-----D-----------.----------K-S------R----V----A-------------R------------M----------------------------D--------------S----R-----I-- 615
35_AD.AF.07.169H              IVS--------------L-------A-----------------D-----------.---------KK-S-----------V---VAM------------R---------A--MD---------------------------D---------------------------D 614
36_cpx.CM.00.00CMNYU830       -V------------------D---------------------ED-----------.----------K-S-----I----------A---------I-------------I--M----------------------------D---------------------------D 615
37_cpx.CM.00.00CMNYU926       -VTK-K---------IL-V-------L-V-E-E---#L-----D------F--S-#---------KK-S------R--A-----VA----I--------R----R----A--RD-----------------------------------------G-----Q-------D 612
38_BF1.UY.03.UY03_3389        -V-----------------R------------E-V--V-----C-----------.------------ST----------V---SV-----------R-R---------A--M---------------------------T---AE---------S-----T-------D 619
39_BF.BR.04.04BRRJ179         -V---A-------------------------------------S-----------.----------T-S---------V-V---VSL---------------------D----------------------------E--T---I------------------------D 618
40_BF.BR.05.05BRRJ055         IV---T------------------------------------H---S--------.---------KLK-T---------------A-------------R---LR---D-------------------------------T---A----------------------I-D 615
42_BF.LU.06.luBF_18_06        -V-----------------------------------------------------.--------G-------------------VXL-A-----RI----------------I----------------------------D---------------------------D 619
43_02G.SA.03.J11223           -VT---------------------------T-E----------D-----------.---------KRGS-------E--QV---VA-------------R---------------------------------R-------D----V------------S-Q-------D 614
44_BF.CL.00.CH80              -V---R---I----------------------E----------C-----------.------------------I------------------------R---L----D-------------------------------T---A----------S---S-R-------D 615
45_cpx.FR.04.04FR_AUK         IVL--A------E------RR---------T---------------S-------Y.----------IKS---------------VAL------------R----------------------------------------T---A-V--------------K-------D 613
46_BF.BR.07.07BR_FPS625       IVS--A-------------------------------L-----------------.---------K--S----------------SL--------I-------L----D-------------------------------T---A----------S-----K-------D 615
47_BF.ES.08.P1942             IV---K-------------------------------------------------.----------S--------R-----------------------R----R--------D------------------------------I------------------------- 614
48_01B.MY.07.07MYKT021        IV------------------N-------------V--V-----D-----------.----------K-S------R----V---VA-------------R----R--------------R---------------------D--I-----------S----R-------D 615
49_cpx.GM.03.N26677           IV---A------------------------A-E----V----LD-----------.---------KR-------I---A-L--RVAI-G----------R-------------D-----------------M---------D---------------------------D 612
50_A1D.GB.10.12792            IVT--A--------------D----------------------D-----------.----------K-S------R--V-V---V-R------------R------------MD---------------------------D--E------------------------D 613
51_01B.SG.11.11SG_HM021       -V-----------------------A--------V--------------------.----------V-------I----------A---------I---R------------M---------------------------R---G------------------------D 615
52_01B.MY.03.03MYKL018_1      -V------------------T-------------V--V-----D---------Q-.----------K-SS-----R----V----A-------------R--------D--------------------------------D---------------------------- 615
53_01B.MY.11.11FIR164         IVQX-X--------------T-------------V--V-----D-----------.----------K-S------R----V----A-------------R-----------------------------------------D-------------------------I-D 615
54_01B.MY.09.09MYSB023        IV---A--------------T----------------------L----------Y.-----------------------------A------------------R---------------------------------------EE------------------------ 615
55_01B.CN.10.HNCS102056       IV------------------T-------------V--V-----D-----------.----------K-S------R----V----A-----------R-R---------------------------------------------------------------------- 615
56_cpx.FR.10.URF5_patient_A   -----A------------------------T------------------------.---------KT--------R--------------I--------------------------------------------------D-------------------I-------D 617
57_BC.CN.09.09YNLX19sg        -V------------------------------E--------E-------------.---R-----K--T-----I----------AR-C----------R-------------D---------------------------D--A------------------------D 611
58_01B.MY.09.09MYPR37         -V---R----------------------------V--V-----X-----------.----------K-S------R--A-V----A-------------R----R-------M----------------------------D--------------SK-----------D 615
59_01B.CN.09.09LNA423         IVT-----------------N------------IV--V-----D-----------.----------K-S-----------V----A-------------R------------M----------------------------D--------------SK------------ 612
60_BC.IT.11.BAV499            IV-----------------------------------------D-----------.---------K--T----------------AQ------------R--------GA---D------------------------------G-V--------------M-------D 610
61_BC.CN.10.JL100010          IV-----------------R-----I----------------HD-----------.---------K--T----------------AM------------R-------------D---------------------------D--A----------------I-------D 611
62_BC.CN.10.YNFL13            IV---T------E------R-----------------------D------F----.---------K--T----------------AM------------R-------------D---------------------------D--A-V--------------I-------D 611
63_02A1.RU.10.10RU6637        IVT-------------------------------V--------D-----------.---------KRGS----------AV---VA-------------R----R---DA---------------------------------------------------I-------D 615
64_BC.CN.09.YNFL31            IV-------------------T---------------------D-----------.---------K--T----------------AM------------R-------------D---------------------------D--A-V--------------I-------D 615
65_cpx.CN.10.YNFL01           IV-------------------T---A----A------V-----D-----------.---------K--T----------------SM--------I---R--------N----D------------------------------E----Y-------------------- 611
67_01B.CN.11.2011.ANHUI.HF115 -V------------------------------E----L-----------------.---------K-----------------------------I----------------I----------------------------------------------D---------D 613
68_01B.CN.11.2011.ANHUI.WH73  -V---------------------C--------E----L-----------------.------------------------------X--------I----------------I----------------------------------------------D---------D 613
70_BF1.BR.10.10BR_PE004       IVT--T------E---------------------------------------D--.---------K--S----------------A-----------R-----L------------------------------------T---T----------S-----V-------D 617
71_BF1.BR.10.10BR_PE008       IV---A-----------------N-------------L----------------C.---------K--S------R--A------AQ----------------L----D-------------------------------T---A----------SS----K-------D 615
72_BF1.BR.10.10BR_MG002       I----A------------------------T-----------------------Y.---------K--ST---------------AL------------R----------------------D-----------------T---A----------------K-------D 616
O.BE.87.ANT70                 -V--SR------E----R--Q-------Q-D---WVN-----GE---------EH.--------T-QKAS----IR--A-VI--VSQ---I----L------VTR-------AD-------------S----I----R--S---M----Y-------------------E 611
O.CM.98.98CMA104              -V--SK------E----R----------Q-D---WVSX-----E--------DEH.--------T-QKAS----IR--A-VL--VSQ--------L------VTR-------AD-------------S----I----R--S---M--D-----------D---------E 611
O.CM.98.98CMABB141            -V--SK------E----K---S------Q-D---WVD------E--------DKY.----------QKAS----IR--A-VL--VSQ-A------L------VTR----A--AD-------------S----I-------S---I----Y---------D---------E 611
O.CM.98.98CMABB212            -VA-SK------E----R----------Q-D---WVN------E--------DEH.----------QKAS----IR--A-VL--VGQ-A------L------VTR----A--A--------------S----I-------T------Q---------S-D---------E 611
O.CM.98.98CMU5337             -V--SK------E----R----------Q-E---WVN------K---------EY.--------T-QKTS----IRL---VI--VSQ-A------L------VTR-------AD-------------S----I-------N---M----Y--------KN---------E 611
O.CM.99.99CMU4122             IV--SK------E----R----------Q-D---WVN------E-----V---EH.--------T-Q-AS----IRI-A-VI--VSQ-A------L-------TR-------AD-------------S----I----R--S---M----Y---------N---------E 611
O.FR.92.VAU                   -V--SK------E----K----------Q-D---WVN------E--------DEH.----------QKAS----IR--ANVI--VSQ-A-I----L------VTR-V-----AD-------------S----I----K--N---I--------------D-------I-E 612
O.GA.11.11Gab6352             -V--SK------E----R----------Q-D---WVN---------S-----DEH.----------QKSS----IR--A-VL--VSQ-A-I----L------VTR----S--AD-------------S----I----R--S---P----Y---------D---------E 611
O.SN.99.99SE_MP1299           IV--SK------E----K----------Q-D---WVN------W--------DEH.--------T-QKAS----IR--A-VL-RVSQ-A-I----L-------TR-------AD-------------S----I-------S---M----Y---------D---------E 611
O.US.10.LTNP                  IV--S-------E----K----------Q-D--IWVN------E---------EH.--------N-QKAS-----R--A-VL--VSQ--------L---R---TR----S--AD----N--------S----I-------S---I----Y--------KT---------E 611
N.CM.02.DJO0131               -VNF-K-------------R-L--------G-E----------D----------Y.---------KR-S-----I---V-V---VA--G------I------V---V--A--DN--------------------------T---T----Y-------K--------F--D 620
N.CM.02.SJGddd                -VNF-K---------K-----TL---------E---D------H---------LY.---------K--S-----IR--V-V-R-VA-------------R--V---V--A--DH-------------S------------T---S-V--Y---------D------F--D 617
N.CM.04.04CM_1015_04          -VTF------------------L-------E-E----------S---------SX.---------K-KS-----I-----V---VA----------------V---V--A---N--------------------------T---S----Y-------K--------F--D 622
N.CM.06.U14296                -VTF-K----------------L-------G-E--------------------AH.---------K--S-----I---A-V---VS----X-----------V---M--A---H--------------------------T---S----Y----------------F--D 622
N.CM.06.U14842                -VNF-Q----------------L-------E-E-V-------HD---------LY.---------RKA-----EI---A-V---VA----------------V---M------H----S---D-----------------T---C----Y----------R-----F--D 623
N.FR.11.N1_FR_2011            -VNF-K---------------TL-------G-E--------------------LH.---------K--S-----I---V-V-R-VA-------------R--V---V--A---H--------D-----------------T---A----Y----------------F--D 622
P.CM.06.U14788                -VAF-R------E--K------------Q-E-E-VD-------A-----VF--EH.----------QKAT----IR--A-VI--VSQ--------L---R--VNRNV-----SD-------------S----I--------D--P-T----------------------D 612
P.FR.09.RBF168                -VA--R------E--K------------Q-E-E-VD-------E---------EH.----------QKAT----IR--A-VI--TSQ--------L---R--VNRSM-----SD-------------S----I--------D--P-I----------------------D 611
CPZ.CD.06.BF1167              -VQ--A------E--KQ----T-Q-T----KEP-K-AV-----K-----V--KEN.QI--------NH-T---EL---AGV-------G-I---RV-V-Q--VT--V-QQ--SD---VS-------------I----N-LT--MPE---Y--------DS---------D 618
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.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
CPZ.TZ.06.SIVcpzTAN13         PVE--R------E--K-----K-Q-A----KLP-Q-AV---------------EG.---------KSP-T---EIR--AGLI--GN--------TV---L--VT----SQ---D---V--V-----I-----IR---N-LSD--PE-------------S-K-------D 610
CPZ.US.85.US_Marilyn          -VNF-H---M--E-----------------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 617
GOR.CM.07.SIVgor2139_287      -VT-SK------E---------------Q-E-E-VNV------E------F-DEY.----------QKAT----IR--A-VI--VSQ--------L---R---CRS--DAF--D----S-------IS----I-----------P--------------S-Q-------D 607
GOR.CM.07.SIVgorCP2135con     -VT-SK------E---------------Q-E-E-VNV------E------F-DEY.----------QKAT----IR--A-VI--VSQ--------L---R---CRS--DAF--D----S--------S----I-----------P--------------S-Q-------D 607
```

```
                               p66 RT, p15 Rnase H end  p31 Integrase start
B.FR.83.HXB2                 RGRQKVVTLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQD.EHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLE 784
A1.AU.03.PS1044_Day0         ------IS--E--------H----------S--------------------R-----------K--G-D----S-------------------S---R-----------E.D--R------T---------II------C------------------------------ 784
A1.CH.03.HIV_CH_BID_V3538    K-----IS--E--------H--H-------S--------------------R-----------E---------S-------------------K---R-----------E.---R------------H---I-------------------------------------- 784
A1.CY.08.CY236               -------S--E--------H----------S--------------------K-----------K--G-D-I--S-------------------T---R-----------E.D---------T----------I--------N------------------M--------- 781
A1.ES.05.X1608_8             -------S-AE--------H----------S--------------------R-----------K--E-D----S-------------------S---------------E.---R----------------I------------------T---------M--------- 788
A1.KE.11.DEMA11KE001         -------S--E--------H--L-------S----------------S---R----I------K--G-D----S-------------------T---------------E.---R------------------------------------------------------- 784
A1.RU.11.11RU6950            -------P-NE--------H--H-------S--------------------R----I--K---K--E--R---S-------------------N---R----E------E.---------K----------I-------------------------------------- 784
A1.RW.11.DEMA111RW002        --K----S--E--------H--L-------S--------------------R-DA-V------K--E-D----S-----------------------------------E.---R----------------I-------------------------------------- 784
A1.SN.01.DDI579              -------S-ME--------H--H-------S----------------------D--I------K--E------S-------------------S---------------E.---R------T---------I-------------------------------------- 783
A1.UG.11.DEMA110UG001        -------S--E--------Y----------A-------------V------K-D---------K--G-D----S-------------------T---------------E.D--R------T----------I------------------------------------- 784
A1.ZA.04.04ZASK162B1         -------P--E--------H----------P--------------------R--A-I------K--E--R---S-------------------S---R-----------E.D-D-------T---------I-------------------------------------- 783
A2.CM.01.01CM_1445MV         -----I-S--E----R---H----------S--------------------R-----------E-----RI--S-------------------S---------------E.---R---------H------I------------------I------------------- 784
A2.CY.94.94CY017_41          -----I-S--E--------H------------------------------ER----I------K--E--R---S-------------------S---------------E.---R---------H--------------------------------------------- 784
B.BR.10.10BR_MG029           ------IH----------------------S---M----------------K------S----E-----------------------------X---XX----------E.---X--X----------X--X-------------X--------------X--------- 784
B.CA.07.502_1191_03          K-----IN--------------H----------------------------K------S----------------------------------N---RI----------E.D-D----------N--------------------------------------------- 784
B.CH.08.M2_0803101_NFLG8     -----AIP-------Q--------------S---------------H----K------S----------------------------------T---R-----------E.---R-------------I--II---------------------------V--------- 788
B.CN.10.DEMB10CN002          -------S-E-----------------------------------L-----K--------------G------------------------------------------E.D--R--T---------------------------------------------------- 784
B.ES.10.DEMB10ES002          -------PM-------------H----------------------L-----R--------------N-------------------K------T---R-----------E.---R-------------I----------C------------------------------ 784
B.FR.11.DEMB11FR001          K-----IS--------------H-----------L----------------K-----------------------------------------T---------------E.----------------------------------------------------------- 783
B.GB.05.MM45d213_GN1         K------SI----------------------K-------------------K------S--------------------------------------------------E.------N-------------I-------------------------------------- 783
B.HK.06.HK003                -------S--------------H------------------------G---K---------------------------------------------------------E.------N---------------------------I------------------------ 788
B.HT.05.05HT_129389          -------S--------------H----------------------------K-----------E------I---------------------NT---R-----------E.----------------S------R----------------------------------- 783
B.JP.12.DEMB12JP001          --K----S--------------H----------------------------K------S------------------------------R---T-V--------------.----------------------------C------------------------------ 787
B.KR.07.HP_18_07JHS10_3909   -------S---------------------ES-----------I--------K-----------------------------------------N---R-------E---E.D--------K---N------I-------------------------------------- 783
B.PE.07.502_0525_wg5         K-K----SI------------------------------------------K---D------------KR------------------------------------R--E.----------------------------------------------------------- 788
B.RU.11.11RU21n              K------S--------------H----------------------------K------S----------------------------I----------------------.----------------------------------I------------------------ 788
B.TH.08.MERLBDTRC10          K--------P------------N----------------------------K------S----------------------------I---------------------E.------N-------------I-------------------------------------- 784
B.US.11.ES38                 -------N--------------N-------S--------------------K------SK------R------------------------------------------E.----------------------------------------------------------- 788
C.AR.01.ARG4006              -----I-A--------------Q----------------------------K--------------Q------S-------------------S---R-----------E.D------------NE-----I------------------T------------------- 780
C.BR.07.DEMC07BR003          -----I-P--E-----------Q-------S--------------------K-----------------R---S-------------------S---------------E.D------------NE-----I------------------T------------------- 780
C.BW.00.00BW5031_1           -----I-S--E-----------Q-------S--------------------K-----------------RI--S-------------I-----S---------------E.----------S---------I----------------------------V--------- 786
C.CN.10.YNFL19               ---K-I-S--E-------------------S--------------------K---------------------S-----------------------------------E.-------------------------------Q--------------------------- 784
C.CY.09.CY260                -----I-S-NE-------------------P--------------------K---G--S-------Q--R---S-------------I-----S---------------E.D-----------V-E-------------------Q----I------------------- 787
C.ES.08.X2363_2              -----I----E-----------Q-------P---------------H----K-----------K--N--R---S-------------------K---------------E.D-----T-------E-----I-----I-N----------L------------------- 787
C.ET.02.02ET_288             -----I-S--E-----------Q-------S--------------------K----V------E-----R---S-------------------S---------------E.-------------NE-----I-------------------------------------- 786
C.IN.09.T125_2139            ---K-I-S--E-----------C-------S--------------------K------S----------R---S-------------------S---------------E.--------------------I----------Q--------------------------- 781
C.KE.00.KER2010              -------SISE-----------Q-------S--------------------K--------------R------S-------------------S---------------E.--------------E-----I------------------T------------------- 780
C.MW.09.703010256_CH256.w96  -------S--G-----------Q-------S--------------------K-----------------RI--S-------------------S---------------E.D--R----------E-----I------------------I---------V--------- 782
C.TZ.08.707010457_CH457.w8   -----I-S--E-----------Q-------S-----------------H--K----I---------N--R---S-------------------N---------------E.--------------E-----I-------------Q----I------------------- 780
C.YE.02.02YE511              -----I-S-NE-----A-----H-------S--------------------K---------------------S-------------------S---------------E.------N-------------I------------------I------------------- 778
C.ZA.08.705010534_CH534.w12  -----I-P--E-----A-----Q-------S--------------------K----I---------N--RT--S-------------------K------------R--E.------N-------------I-------------------------------------- 787
C.ZA.10.DEMC10ZA001          K---------E-----------QI------P--------------------K-----------E-----R---S-------------------S---------------E.------N---------------------------Q----T------------------- 780
C.ZM.11.DEMC11ZM006          -----I-S-SE-----------Q-------S--------------------K--------------R--R---S-------------------S---------------E.---R--N-------E-----I-------------------------------------- 785
D.CM.10.DEMD10CM009          -----A-SF-------------N-------P--------------------K------S----------------------------------N----I------S----.--DR-----------------I------------------------------------- 789
D.CY.06.CY163                -------PF-------------N-------P--------------------K------SL---------------------------------R-----------S----.--D---N--------------I-----------------------W------------- 783
D.KE.11.DEMD11KE003          -------S--------------NM---------------------------K------S--------------S-------------------N----I----------E.------N---------------------------------------------------- 783
D.KR.04.04KBH8               ------IP--------------N----------------------------K------S----------------------------------N------------E--E.D-------------E-------------C-----Q------------------------ 783
D.SN.90.SE365                -----AIP--------------N-------S--------------------K----V-S----------------------------------K---------------E.------N--------------I-----------------I------------------- 789
D.TZ.01.A280                 K------S--------------N------------I---------------K------S----------------------------------S---------------E.D--R---------------S--------C------------------------------ 785
D.UG.10.DEMD10UG004          -------P--------------N-------S----------------------D-------------------S-------------------N----I----------E.------N--------------------------------L------------------- 783
D.UG.11.DEMD11UG003          -------S-A------------N----------------------------K---------------------S-------------------S---------------E.------N-------------I------------------L------------------- 785
D.YE.02.02YE516              -------PF-------------N-------S--------------------K----V-S----------------------------------S-----------NQ--E.D-D--------------I----------------Q------------------------ 785
D.ZA.90.R1                   -------PF-------------N----------------------------K------S----------R-----------------------Q---------------E.------N---------------------------------------------------- 783
F1.AO.06.AO_06_ANG32         --KE-A-P--E-----A-----H-------S--------------------K--------------Q--R---S------------------------I----------E.------N---------------------------------------------------- 788
F1.AR.02.ARE933              -----A-P--E-----A--K-VQ-------A--------------------K----------------------------------------------I----------E.D-----N---------------------------------------------------- 784
F1.BR.10.10BR_RJ015          -----A-F--E-----A--E--Q-------S--------------------K---------------------S------------------------I----------E.D-----N----------I----------------------------------------- 787
F1.CY.08.CY222               -------S--E-----A-----H-------S--------------------K------S----R--Q------S--------R----------N----I----------E.---------------------I--K------R--------------------------- 783
F1.ES.02.ES_X845_4           -----A-S--E-----A-----H-------P--------------------K--------------Q-TRI--S------------------------I----------E.D-------------------I-----I-------------------------------- 786
F1.RO.96.BCI_R07             -----A-S--E-----A-----H-------S--------------------K--------------Q--RI--S------------------------I----------E.------N--K----------------------------T-------------------- 784
F1.RU.08.D88_845             K----A----E-----A-------------S--------------------K--------------Q--R---S-------------------S---------------E.------N---------------------------------------------------- 784
F2.CM.02.02CM_0016BBY        -------S--E-----------H-------S-----------------H--K--------------Q------S-------------------T---------------E.------N-------------I------------------I------------------- 784
F2.CM.10.DEMF210CM001        K------S--E-----------C-------S-----------------H--K--------------Q------S-------------------S---------------E.---------------------------------------I------------------- 789
F2.CM.10.DEMF210CM007        -------S--E-----------Q-------P-----------------L--K--------------Q------S-------------------T---------------E.----F-N---------------------------------Q------------------ 783
G.BE.96.DRCBL                K---III---E-----A--H--Q------QS--------------------R----I----------------S-------------------S---------------E.---R-----K----------I-------------------------------------- 780
G.CM.10.DEMG10CM008          --K--IIS--E-----------N----------------------------K----I---------R------S-------------I-----T---------------E.---R--N------N------I------------------I------------------- 783
G.CN.08.GX_2084_08           K----II-IQE-----A--H--Q-------S--------------------R-------X------X---I----------------------S---------------E.---R----------------I-------------------------------------- 788
G.CU.99.Cu74                 K-K---I---------------Q-------S--------------------R---------------------S-------------------T---------------E.---R----------------II------------------------------------- 780
G.ES.09.X2634_2              --K---I--SE-----A-----QI------S--------------------R---------------------S-------------------S---------------E.---R--N---------------------------------------------------- 788
G.GH.03.03GH175G             K-K--II---E-----A-----Q------KS--------------------R---------------------S-------------------N---------------E.---R----------------I------------------L------------------- 784
G.KE.09.DEMG09KE001          K-K--II---E-----S--H--Q-------S--------------------R-D-----K-------------S-------------------N---------------E.---R----------------I-------------------------------------- 782
G.NG.09.09NG_SC62            -EK--I--IQE--------H--Q-------P--------------------R----I----------------T------------K------S---------------E.D--R----------------I-------------------------------------- 783
G.ZA.01.TV546                --K--I----E-----------Q-------S-----------------I--R--------------N------S-------------------N---------------E.---R----------------I-------------------------------------- 783
H.BE.93.VI991                --K--I-S--E----------------E--P--------------------K-----------E------F--S-------------------S---------------V.Q-------------------I-------------------------------------- 782
H.CF.90.056                  --K----S--E----------------------------------------K-----------E---------S-------------------S-V-------------E.---R--N----V--------I-------------------------------------- 783
H.GB.00.00GBAC4001           --K--I-S--E----------------------------------------K-----------E---------S-------------------S---------------E.---R--N--K---N------I------------------I------------------- 782
J.CM.04.04CMU11421           K--------------Q------L---X---X-A------------------K-----------E--X--R---S-------------------S---------------E.---R----------------I-------------------------------------- 780
J.SE.93.SE9280_7887          K------------------H------R------------------------K-----------E---------S-------------------S---------------E.D---------------------------------------------------------- 782
J.SE.94.SE9173_7022          K-----I-----------------------I--------------------K-----------E---------S-------------------S---------------E.----------------------------------------------------------- 782
K.CD.97.97ZR_EQTB11          K-----ISI-E-----------H-------S--------------------K----------------DR---S-------------------S---------------E#------N--------------------------------I------------------- 784
K.CM.96.96CM_MP535           -------SI-E-----A-----C-------S--------------------K---D-------------RI--S-----------------------------------E.------N-------------I------------------I------------------- 784
U.CA.01.TV749                -------S--E--------H----------S-----------I--------R-----------K--E---I-MS-------------------S---------------E.-------------N--------------------I--------------V--------- 789
U.CA.99.TV721                -----I-S-NG--------H----------S-----------I--------R----I--W---K--E---I--S-------------------S---------------E.-------------N-------------------------I---------V--------- 788
U.CD.83.83CD003_Z3           ------IS--E-----A----L----------A------------------S---------------------S-------------A---------------------E.---R----------------I-------------------------------------- 785
U.CD.90.90CD121E12           Q----TIS--E-----A----L----------A------------------S---------------------S-----------------------------------E.--------------------I------------------I------------------- 784
U.CY.05.CY090                -------S--E---P---------------S-----------------H--K----------------D----S-------------I-------V-------------E.-------T-------------E------------V------------------------ 784
U.CY.08.CY223                K------S-NE--------H----------S--------------------R------S--------------S-------------------N---------------E.------N-------------I----------------------------V--------- 785
U.ES.10.DEURF10DZ001         -------S--E--------H--H-------------------I--------K---A--S----K--Q-N----S-------------------N---------------E.---R---------G--------------------------------G------------ 784
U.GR.99.99GR303              ---X---A--E-----------H-------S--------------------T-----------K--G--R---S-------------I-----N---------------E.D-XR---------N-------I--------AGIR------------------------- 782
U.NL.95.U_NL_95_H10986_D1    -------S--E-----------Q-------SA-------------------N---D----------Q-DQ---S-------------------N----I----------E.D-----T------------------------------------------M--------- 784
01_AE.AF.07.569M             ------I---E--------H--H-------S--------------------R----V------E---------S-------------------S---------------E.---R-----K------------------N------------------------------ 784
01_AE.CF.90.90CF11697        -------S--E--------H--H-------SD--------H----------R----I------K--E--R---S-------------------S---------------E.---R------T---------I-------N------------------------------ 782
01_AE.CN.10.YNFL03           -------S-AE--------H--Q-------S--------------------R----V------E---------S-------------------S---------------E.---R------T-V---------------N------------------------------ 783
01_AE.HK.04.HK001            -------S--E--------H----------S--------------------R----I-S----E---------S-----------NK------S--------N------E.---R------T---------I-------N------------------------------ 781
01_AE.IR.10.10IR.THR48F      ---KR--S--E--------H--H-------S--------------L-----R----V------E---------S-------------------S---------------E.---R------T------------------------------------------------ 776
01_AE.JP.x.JRC77AE           -------S--E--------H--Q-------P--------------------R----V------E--R------S-------------------S---------------E.---R------T-----------------N------------------------------ 779
01_AE.TH.09.AA090a_WG11      -------S--EA-------H--Q-----X----------------------R----V------E----D----S-------------------S---------------E.---R------T-----------------N------------------------------ 784
01_AE.TH.90.CM240            -------S--E--------H--H-------S--------------------R----V------E---------S-------------------S---------------E.---R------T---------I------TN------------------------------ 789
01_AE.US.05.306163_FL        -------S--E--------H--Q-------S--------------------R-D--V------E---------S-----------------------------------E.---R------T---------I-------N------------------------------ 782
01_AE.VN.98.98VNND15         -------S--E--------H--H-------S--------------------R----V------E---------S-------------------S---------------E.---R------T---------I-------N------------------------------ 783
```

```
p66 RT, p15 Rnase H end p31 Integrase start
B.FR.83.HXB2                  RGRQKVVTLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQD.EHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLE 784
02_AG.CM.10.DE00210CM013      ----RIIS--E--------H--H-------S--------------------R-----------K--E------S-------------------S---------------E.---R----------------I-------C------------------------------ 784
02_AG.CY.09.CY256             K---------E--------H--H-------S--------------------R--------------G-D----S-------------------S---------------E.D--R--------G------I-------------------------N------------- 784
02_AG.ES.06.P1423             ------XS--E--------H----------S--------------------R-D-----L---E----D----S-------------------T---------------E.------N-------------I-------------------------------I------- 783
02_AG.GW.05.CC_0048           ----------E--------H--H-------S--------------------R----I------E--R-D-I--S-------------------S---R-----------E.---R----------------I------------Q------------------------- 782
02_AG.KR.12.12MHR9            -------S--E--------H----------S--------------------R-------L---K--E-D----S-------------I-----S---------------E.---R----------------I--------------------------------------- 788
02_AG.LR.x.POC44951           K------S--E--------H--H-------S--------------------R-----------E--G-D----S-------------------N---------------E.D--R----------------I--------------------------------------- 784
02_AG.NG.09.09NG_SC61         -------P--E-----------H-----A-S--------------------R-----------E----DR---S-------------------S---------------E.---R----------------X--------------------------------------- 784
02_AG.NG.x.IBNG               -------S--E--------H--H-------S--------------------R-----------K----D----S-------------------N---------------E.---R-----K----------I------------M-----------G------------- 784
02_AG.SE.94.SE7812            -------S-AE--------H--H-------S--------------------R------T----E----DR---S-------------------S---------------E.---R------------------------------------------------------- 784
02_AG.SN.98.98SE_MP1211       ----N--S--E--------H--L----E--S--------------------R-----------K--E-D----S-------------------K---------------E.---R---------E------I-------------------------------I------- 784
03_AB.RU.97.KAL153_2          -------S--------------H----------------------------K------S--------------------------------------E-----------E.------G----------------------------------------------------- 784
04_cpx.CY.94.94CY032_3        -------S-SE-------------------S-----------I--------R---D----------R-D----S-------------------N---------------E.------N-------------S----------N---------------------------- 783
05_DF.BE.x.VI1310             K----A-S--E-----A-----H----------------------------K--------------E-K----S------------------------------------.----------------------------------------------------------- 784
06_cpx.AU.96.BFP90            ----------E-----------N-------P--------------------K------------------I--S-------------------N---------------E.D--R--------N------I-----------------I------------------- 786
07_BC.CN.98.98CN009           ---K---S--E--------H--CI------S--------------------K-----------------R---S-------------------S---------------E.---R---------------------------Q--------------------------- 784
08_BC.CN.97.97CNGX_6F         ---K-I-S---------------I------S--------------------K-----------------R---S-------------------N---------------E.-------------------I----------Q--------------------------- 780
09_cpx.GH.96.96GH2911         -----I-S--E--------H--H-------S--------------------R-----------K--E------S-------------------S---------------E.---R------------------------------------------------------- 783
10_CD.TZ.96.96TZ_BF061        ----RAISI-------------N-------S--------------------K---------------------P-------------------R---------------E.------N-----------------------------L------------------- 783
11_cpx.CM.95.95CM_1816        ----------E--------E--H----------------------------K------S--------------S-------------------H---R-----------E.---R----------------I-----------------I------------------- 782
12_BF.AR.99.ARMA159           K------S--E-----A--H----------S--------------------K------S--------------S---------------------V--I----------E.------N--------------------------------------------------- 784
13_cpx.CM.96.96CM_1849        --K--II---E-----------Q-------S--------------------R---------------------S-------------------N---------------E.D--R------T-----------------------------T------------------- 791
14_BG.ES.05.X1870             K-K--II---E-----A-----QI------S--------------------R-----------------------------------------S---------------E.---R--N----------I----------------------------------------- 783
15_01B.TH.99.99TH_MU2079      -------S--E-----------Q---R---S--------------------R----V------E---------S-------------------S---------------E.---R------T---------I--------------------------------------- 783
16_A2D.KR.97.97KR004          K----I-S-----------H-------------------------------R----I--L---K--E--R---S---------E---------S--------------------.-----R-------H--------------C-----V------------------------- 784
17_BF.AR.99.ARMA038           -----A-S--E-----A--H--K-------S--------------------K----I----------------S-------------------------I---------E.------N------------------------------------I--------------- 784
18_cpx.CU.99.CU76             K-----ISI-E--------E--H-------Q--------------------K----I------E---------S-------------------S---R-------E---E.--D---N-----------------------------------------------R---- 782
19_cpx.CU.99.CU7              -------S--------------N----------------------------K------S----------R-----------------------S---------------E.-------------------I--------------------------------------- 783
20_BG.CU.99.Cu103             --K----S--------------H----------------------------K------S----------------------------------S---------------E.---R--------N------I--------------------------------------- 780
21_A2D.KE.99.KER2003          -----I-S--E----R---H-------------------------------R--------------G-KR---S-------------------S---------------E.-----------H------I--------------------------------------- 783
22_01A1.CM.01.01CM_0001BBY    -------S--E-I------H--H-------S--------------------R-------L---K--E------S-------------------S----I----------E.D--R------T-V------I---------N--------I------------------- 784
23_BG.CU.03.CB118             --K----S--------------H-------S--------------------K------S----------------------------------S---------------E.---------K------------------------------------------------- 780
24_BG.ES.08.X2456_2           -------S--------------H----------------------------K----V-S----------------------------------S---------------E.Q--R------------------------------------I------------------- 780
25_cpx.CM.02.1918LE           -------S--E--------H--H-------S--------------------K----I----------------S-------------------N---------------E.---R----------------I-----------------L------------------- 783
26_AU.CD.02.02CD_MBTB047      K---------E--------H-------------------------------R----I------K--E------S-------------I-----S---------------E.D-------------N------I-----I-----------I---------V--------- 782
27_cpx.FR.04.04CD_FR_KZS      K---------E-----------Q-------S-----------------R--K---------------------S-------------------S---------------E.---R----------------I--------------------------------------- 782
28_BF.BR.99.BREPM12609        -------P--------------H----------------------------K------S----------------------------------------I---------.-----------S-D-------I--------------------------------------- 784
29_BF.BR.01.BREPM16704        -----A-S--E-----A-----Q-------S--------------------K--------------N------S-------------------------I---------E.------N-----------------------V---------------------------- 784
31_BC.BR.04.04BR142           -----C-S--E-----------Q-------S--------------------K-----------L--N--R---S-------------------N---------------E.------------E------I-----------------T------------------- 790
32_06A1.EE.01.EE0369          ---------SE-----------N-------S--------------------K---------------------S-------------------S---------------E.---R--------N------I-------------------------------I------- 784
33_01B.ID.07.JKT189_C         -------S--E--------H--H-------S--------------------R----V------E----D----S-------X-----------S---------------E.---R------T---------I-------N-----------------------V------- 784
34_01B.TH.99.OUR2478P         -------S--E--------H--H-------S--------------------K----V------E---------S-------------------S---------------E.---R------T---------I-------N------------------------------ 784
35_AD.AF.07.169H              -------P--E--------H--H-------S--------------------K-------K---K-VE-D----S-------------------N----I------E---E.---R----------------I----------------------------M--------- 783
36_cpx.CM.00.00CMNYU830       -------S--E--------H--H-------SK-------------------R-----------K--E-D----S-------------------K----I----------E.D--R----------------I--------------------------------------- 784
37_cpx.CM.00.00CMNYU926       K----I-SF-E--------H--H------ES--------------------R-D--I------KI-E--R---S-------------I-----S----GV---------E.D-KR--------N------I--------------------------------------- 781
38_BF1.UY.03.UY03_3389        -----A-S--E-----A--H--S-------S--------------------K---------------------S-------------------------I---------E.------N--------------------------------------------------- 788
39_BF.BR.04.04BRRJ179         -------S-E--------------------S-----------------K--K------S----E------I--T------------------------------------E.------N-------------I--------------------------------------- 787
40_BF.BR.05.05BRRJ055         -------P--------------H----------------------------K------S----------------------------------------I----------.----------------------------------------------------------- 784
42_BF.LU.06.luBF_18_06        -----A-S--E-----A-----Q----------------------------K----I----------------S-------------------T----I----------E.D-----------------I------I-C------------------------------ 788
43_02G.SA.03.J11223           -------S--E--------H----------S--------------------R-----------K--E-D----S-------------I-----S---------------E.---R----------D-----------------Q------------------------- 783
44_BF.CL.00.CH80              -----A-S--E-----A--H--Q-------P--------------------K---G----------E--R---S-------------------------------------E.D-------------------II-----I-------------------------------- 784
45_cpx.FR.04.04FR_AUK         -------S--E-----------H-------S--------------------K---D----------M-D----S-------------------S---------------E.----------------------------------------------------------- 782
46_BF.BR.07.07BR_FPS625       -----A-S--E-----S-----Q-------S------------------------D-------L----------------------------------------------.----------------------------A------------------------------ 784
47_BF.ES.08.P1942             K------S--------------L----------------------------K--------------N------S-----------------------------------E.------N---------I-----------------Q------------------------- 783
48_01B.MY.07.07MYKT021        -------S--E--------H--H-------S--------------------R----V------E---------S-------------------S---R-------R--E.D------------------II-------------------------------------- 784
49_cpx.GM.03.N26677           K------S--E--------H----------S--------------------K-----------------R---S-------------------K---------------E.------------E------I-----------------I------------------- 781
50_A1D.GB.10.12792            -------S-NE----M---H--H------R------------------L--K---G-------K--E-D--------------------------S---------------E.D--R------T------------------------------------------------ 782
51_01B.SG.11.11SG_HM021       -------S--------------H------------------------------------K------Q--------------------------S---R-----------E.D-----------------------------------L------------------- 784
52_01B.MY.03.03MYKL018_1      -----I-S--E----Q---H--H-------S--------------------R----V------E---------S-------------------S---------------E.---R-----------------I-------------------------------------- 784
53_01B.MY.11.11FIR164         -------S--E--------H--H-------S--------------------K----V------E---------S-------------------K---------------E.---R------T---------I-------N------------------------------ 784
54_01B.MY.09.09MYSB023        K--------P------A-----H----------------------------K------S-------Q--R-----------------I-----S---------------E.------N--------------------------------------------------- 784
55_01B.CN.10.HNCS102056       K------S----------------------S--------------------K-----------M-S--R-------------------------S---------------E.D--R------T------------------------------------------------ 784
56_cpx.FR.10.URF5_patient_A   -------S--E--------H----------S--------------------K----V-S--------------S-------------I-----S---R-----------E.D---------T-------------------N-------------G--------------- 786
57_BC.CN.09.09YNLX19sg        ---R-I-S--E------------I------S--------------------K-----------------R---S-------------------S---------------E.D--R----------------I----------Q--Q------------------------ 780
58_01B.MY.09.09MYPR37         -------S--E--------H--H-------S--------------------R----V------E---------S-------------------S---------------E.D--R------T---------I-------N------------------------------ 784
59_01B.CN.09.09LNA423         -------S--E--------H--H-------S--------------------R----V------E----D----S-------------------S---------------E.---R----------------II-------------------------------------- 781
60_BC.IT.11.BAV499            ------IS--E-----------Q-------S-----------------K--K---------V-------R---S-------------------K---R-----------E.D-----------NE--L--I-----------------T------------------- 779
61_BC.CN.10.JL100010          ---K-I-S--E-----------CI------A--------------------K---------------------T-------------------S---------------E.---R----------------I---------NQ--Q----R------------------- 780
62_BC.CN.10.YNFL13            ---N-I-S-NE-----A------I------S--------------------K-----------------R---S-------------------H---------------E.---R-------------I--I----------Q--------------------------- 780
63_02A1.RU.10.10RU6637        ------IS--E--------H--H------RS--------------------R----I--K---K--E------S-------------------S---------------E.D--R--------N------I-----------------I------------------- 784
64_BC.CN.09.YNFL31            ---K-I-S--E------------I------S--------------------K--------------R--R---S------------------------------------E.-------------------I----------Q--------------------------- 784
65_cpx.CN.10.YNFL01           K---------------A-----H-------V--------------------K-----------------R---S-------------------S---------------E.------N-------------I--------------------------------------- 780
67_01B.CN.11.2011.ANHUI.HF115 -------S--------------H-----------------------H----K------S----------------------------------T---------------E.------------V---------------------I----L------------------- 782
68_01B.CN.11.2011.ANHUI.WH73  -------P--------------H-----------------------H----R---------------------------------------------------------E.------------V---------------------I----------------------------- 782
70_BF1.BR.10.10BR_PE004       -------SI-E-----A-----Q----------------------------K------S----------------------------------N-----------E---E.-----------------------------N------------------------------ 786
71_BF1.BR.10.10BR_PE008       -------S--E-----A-----Q-------S--------------------K-----------------RI--T-------------------T---------------E.------N--------------------------------------------------- 784
72_BF1.BR.10.10BR_MG002       K----A-S-IE-----A-----Q-------ST-------------------K----I------E---------S-------------------------I------R--E.------N--------------------------------------------------- 785
O.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--E.D----------K-L--E-G---------I---P--HI----I--------EV--I----M- 780
O.CM.98.98CMA104              Q-K--IIK-DE-----A--M-VL------KEK---------V----SS--T----PI-Q----E-T---Q---T-------------I-----KD--R----E---Q--E.D------X---L----G---------I-N-P--HI----I--------EV--M----T- 780
O.CM.98.98CMABB141            Q-K--IIK-EE-----A--M-VL------KEKI-------------SS--T----PI-Q----E-T-------T------------KI-----KD--R----E---Q--E.D----------L----G---------I---P--HIR-------------EV--I----V- 780
O.CM.98.98CMABB212            Q-K--IIE-KE-----A--M-VL------KEK---------V----SS--T--D-PI-Q----E-T---Q---T------------KI-----KD--R----E---Q--E.D----------L----G---------I---P--HI----T--------E--M------ 780
O.CM.98.98CMU5337             Q-K--IIK-EE-----A--M-VL------KEX--------------SS--T----PI-Q----E-TR--R----------------KI-----KD--R----E---Q--E.D----------L----G---------I-N-P--HI----T-----Y--EV--I------ 780
O.CM.99.99CMU4122             Q-K---IK-EE-----A--M--L---K--KET--------F-----SS--T--D-PI-Q----E-T---Q----------------KI-----KD--R----E---Q--E.D----------L----G---------I-N-P--HI----I--------EV--I----T- 780
O.FR.92.VAU                   Q-K--IIK-DE-----A--M-VL------KEK---------V----SS--T--D-PI-Q----E-T-R-Q---T------------KI-----KD--R----E---Q--E.D----------L----G---------I---P--HIRR-------------E--M------ 781
O.GA.11.11Gab6352             Q-K--I-K-EE-----A--M-ML------KET---------T----SS--T----PI-Q----E-T--D----T------------KI-----KD--R----E---Q--E.D----------L----GI-A------I---P---I----I-----Y------M------ 780
O.SN.99.99SE_MP1299           K----IIK-EE----RA--M-VL------KET----Q----V----SS--T----S--Q----E-T---Q---T------------KI-----KD--R----E---Q--E.D----------L----G---------I-N-PQ-HI----I--------EV--M------ 780
O.US.10.LTNP                  Q-K--IIN-DE-----A--M-VL------KEK---------V----SS--T----PI-Q----E-TR--Q---T------------KI-----KD--R----E---Q--E.D----------K-L--E-G---------I---P--HIR---I--------EV--M----V- 780
N.CM.02.DJO0131               --K----SIE------------L-------Q---------------HSH--K------S----E-----R---S-------------------S-----------E---E.D--R-----------------S-------------------I----N----V--------- 789
N.CM.02.SJGddd                K------SIA------A-----L-------Q-----------M---HS---K------S----E-----R---S-------------I-----S-----------E----.D--R-----------------I------------------------N----V--------- 786
N.CM.04.04CM_1015_04          -------S-EN-----------LM----X-Q-----------M---HS---K------S----E-----R---S-------------------S-----------E---E.---R-----K-----------I------------------------N----V--------- 791
N.CM.06.U14296                -------SIX------A-----LM---E-EK-----------M---H----K----V-S----E-----R---S-------------------S---------------E.D--R-----------------I------------------------N----V--------- 791
N.CM.06.U14842                K------SIA------------LM------QD----------M---HS---K------S----E-----R---S-------------------S------------X--E.D--------K-----------I------------------------N----V--------- 792
N.FR.11.N1_FR_2011            ------ISIEN-----A-----LM---E--R-A---------M---HS---K------S----E-----R---S-------------------S---------------E.D--R-----------------I------------------------N----V--------- 791
P.CM.06.U14788                -----IIK-EE-----A--E-VL---KE--EQA--------V----S-T----D-P--QK---EMT-------S------------NI-----KD--R----E---Q--E.D----------L--E-G---I-----INN-P--HV-----------------------M- 781
P.FR.09.RBF168                -----IIK-EG-----A--E-VL---KE--KQA--------V----S-T----D-P--QK---EMT-------S------------DI-----KD--R----E---Q--E.D--------K-L----G---I-----INN-P--HV-----------------------T- 780
CPZ.CD.06.BF1167              --KSR-IE--E----QA-----L---K-AKEK---------V---L-GT----D-P--EE--QE-MN--A---N------------E------KN-------E---Q--E.D--RF----KML-DEY----I-----I-Q-N--HR----I--------E--------I- 787
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------ 783
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------ 782
CPZ.TZ.06.SIVcpzTAN13         ---Y-TRE-EN----QA--W-VD---K---SQ---------VM-VL-GM----D-PI-EE--QK-MQ-NA----------------E------KN-------E--TE--E.D-D------K---DEY----I-----I-Q-S--HK----I--------EV--------- 779
CPZ.US.85.US_Marilyn          -----II--EN-----A--T-V----K--ENT--V--G---V----HS---------------E------S-IS-------------I-----S---------------E.D-D------T-----------------------------P----I------------------- 786
GOR.CM.07.SIVgor2139_287      K----IRK-EG-----A--E-VL---EE--KN---------V----A-S----D-P--QK---EMTT-------------------EI-----KD--R----E---Q--E.D---------TL----G---------IN--P--HV----R------------------- 776
GOR.CM.07.SIVgorCP2135con     -----IRK-EG-----A--E-VL---EE--KN---------V----A-S----D-P--QK---EMTT-------------------EI-----KD--R----E---Q--E.D----------L----G---------IN--P--HV----R------------------- 776
```

```
B.FR.83.HXB2                 GKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTGATVRAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRNPLWKGPA 954
A1.AU.03.PS1044_Day0         ------------------------------------------VV---------S-A-K------N-Q------------------------------E---------------------------------I--------------H---------------D-I----- 954
A1.CH.03.HIV_CH_BID_V3538    ----I------------------------------------RV----------S-A-K------N-Q------------------------------E---------------------------------I------------------------------D-I----- 954
A1.CY.08.CY236               --I---------------------------------A-----VV-----P---S-A-K------D-Q------------------------------E---------------------------------I---S------------I-------------D-I----- 951
A1.ES.05.X1608_8             --I----------------------------M----------VV---------S-AFK--------Q------------------------------E---------------------------------I------------------------------D-I----- 958
A1.KE.11.DEMA11KE001         -------------------------------I----------VV---------S-A-K--------Q------------------------------E---------------------------------I------------------------------D------- 954
A1.RU.11.11RU6950            ----I-------------------------------------V------P---SNA-K------N-Q------------------------------E---------------------------------I----------------I-------------D-I----- 954
A1.RW.11.DEMA111RW002        ------------------------------------------VV---------S-A-K------D--------------------------------E---------------------------------I------------------------------D-I----- 954
A1.SN.01.DDI579              ----I----------------------A--------------V----------S-AMK------N--------------------------------E---------------------------------I------------------------------D-I----- 953
A1.UG.11.DEMA110UG001        ------------------------------------------MV---------S-AFK-------VQ------------------------------E---------------------------------I----------------I-------------D-I----- 954
A1.ZA.04.04ZASK162B1         --I---------------------------------A----RV----------S-A-K------NVQ----------------------------------------------------------------I------------------------------DSV----- 953
A2.CM.01.01CM_1445MV         -------------------------------I----------V------P---S---K-------VQ--------------------------------------------------------D-------I-------------R----------------E-I----- 954
A2.CY.94.94CY017_41          ---------------------T----D----I----------V------P--IS---K--------Q----------------------------------------------------------------I-------------R----------------D-I----- 954

B.BR.10.10BR_MG029           --X-----------------------X----XX---------X------X--ISXA-K--X----X--------------------X------------------------X-------------------I-----N------------------------D------- 954
B.CA.07.502_1191_03          --I----------------------------I-R---------------P--IS---K-------------------------------------------------------------------------I------------------------------D------- 954
B.CH.08.M2_0803101_NFLG8     --I---------------I------------------------------P---SN--K-------------------------------------------------------------------------I---------T--------N-----------D------- 958
B.CN.10.DEMB10CN002          --I----------------------------I---------Q-----------STA-K-------V--------------------N--------------Q-----------------------------I---S--L---A-------------------D------- 954
B.ES.10.DEMB10ES002          ---------A---------------------------------V-----P---SN--K-------------------------------------------------------------------------I------------------------------D------- 954
B.FR.11.DEMB11FR001          --I--------------------------------------------------ST--K----------D-------------------------------------------------R------------I------------------------------D------- 953
B.GB.05.MM45d213_GN1         -------------------------------I----------A----------S-A-K--------------------------L----------------------------------------------------EM--R--------------------D------- 953
B.HK.06.HK003                --I----------------------------I-----------V---------SNA-K-----------------------------------E-------------------------------------I------------------------------D-I----- 958
B.HT.05.05HT_129389          -------------------------------I---------R-----------S-V-K-----------------------------------E---E----------------------------------------------------------------D------- 953
B.JP.12.DEMB12JP001          --I----------------------------I---------------------ST--K----------------------------N----------------------------------------------M----------------------------D------- 957
B.KR.07.HP_18_07JHS10_3909   --I----------------------------I---------T-----------STA-K-------V------------------------T----------------------------------------I-----N---Q---R---Q----------N-D-I----- 953
B.PE.07.502_0525_wg5         --I----------------------------I---------------------S---K-------------V------------------------------------------------------------------------------------------D------- 958
B.RU.11.11RU21n              --I----------------------------I----------V----------STA-K-----------------------------Q--Q--------------------------------------------------Q------------------N-D-I----- 958
B.TH.08.MERLBDTRC10          --I--------------------------------------------------SN--K-----------------------------------E----------------------------------------------------H---M---------N-D------- 954
B.US.11.ES38                 --I--------------------------------------------------ST--K-----------------------------------E--------------------------------------------M---------------------N-D------- 958

C.AR.01.ARG4006              --I----------------------------I----------V----------SNA-K--------Q----------------------------------------------------------------I----------------I-------------D-I----- 950
C.BR.07.DEMC07BR003          --I----------------------------I-----------------P--ISN--K--------Q------------------------------------------------------------------------------N--I-------------D-I----- 950
C.BW.00.00BW5031_1           --I-------------------------V--I----------A----------S-A-K--------N------------------------------E----------------------E----------I------------------------------D-I----- 956
C.CN.10.YNFL19               --I-----------M----------------I----------V----------S-A-K--------------------------L------------E---------------------------------I---------R------I-------------D-I----- 954
C.CY.09.CY260                --I----------------------------I----------V----------S---K--------Q------------------------------E---------------------------------I---------R------I-------------D-I----- 957
C.ES.08.X2363_2              -R----------------------------YI----------V------P---S-A-K--------Q----------------------------I-----------------------------------I----------------I-------------D-I----- 957
C.ET.02.02ET_288             --I---------------------------YI----------V----------SSA-K--------T--------------------------V------------------------R----D-C-----I-------------N--L-------------D-I----- 956
C.IN.09.T125_2139            --I-----------M----------------I----------V----------S-A-K--------Q----------------------------------------------------------------I------------------V-----------D-I----- 951
C.KE.00.KER2010              --I---------------------------YI----------V----------S-A-K--------Q----------------------------------------------------------------I---------R------IQ------------D-I----- 950
C.MW.09.703010256_CH256.w96  --I------------------T--------YI----------V------P---S-A-K--------Q----------------------------------------------------------T-----I----------Q-----I-------------D-I----- 952
C.TZ.08.707010457_CH457.w8   --I------------------E--------YI----------V------R---SN--K---------------------------------------------------------------D-T-----I----------------I-------------D-I----- 950
C.YE.02.02YE511              --I---------------------------YI----------V----------S-A----------Q------------------------------E---------------------------------I----------------I-------------D-I----- 948
C.ZA.08.705010534_CH534.w12  --------------M-----------D----I----------V----------S-A-K--------Q----------------------------------------------------------------I----------------I-------------D-I----- 957
C.ZA.10.DEMC10ZA001          --I---------------------------YI----------V------T---S-A-K--------Q----------------------------------------------------------------I----------------I-------------D-I----- 950
C.ZM.11.DEMC11ZM006          ------------------------------YI----------V----------S-A-K--------Q----------------------------------------------------------------I----------------I-------------D-I----- 955

D.CM.10.DEMD10CM009          --I----------------------------I----------VV---------S---K-------------------------------------------------------------------------I------------------------------D-I----- 959
D.CY.06.CY163                -----------------K------------------------VV---------S-A-K-------V-------------------------------------R---------L-----------------I---------R--------------------D-I----- 953
D.KE.11.DEMD11KE003          ------------------------------------------VV---------S---K-------------------------------------------Q-----------------------------I------------------------------D-I----- 953
D.KR.04.04KBH8               ------------------------------------------V----------S-V-K-----------------------------D---------E---------------------------------I---------R--------------------D-I----- 953
D.SN.90.SE365                ------------------------------------------VV---------S---K-------------------------------------------------------------------------I---------R------I-------------D-I----- 959
D.TZ.01.A280                 -------------------------------I----------V----------S---K----------------------------R--------------------------------------------I---S-L--T--------------------D-I----- 955
D.UG.10.DEMD10UG004          --I---------------------------------------VV-----R---STA-K---------------------------------------------R---------------------------I---------------V--------------D------- 953
D.UG.11.DEMD11UG003          -------------------------------I----------VV--------IS-A-K---------------------------------------E---------------------------------I------------------------------D-I----- 955
D.YE.02.02YE516              --I---------------------------------------VV---------SGA-K-------------------------------------------------------------------------I------------------------------D-I----- 955
D.ZA.90.R1                   ----I-------------------------------------VV---------S-A-K------------------------------------------------L------------------------I------------------------------D-I----- 953

F1.AO.06.AO_06_ANG32         ----M--------------------------I----------I----------S-A-K--------Q--------------------------------------------------------------------S------------------------N-D-V----- 958
F1.AR.02.ARE933              --------------L----------------I--------A--------P---S-A-K-S------Q--------------------------------------------------------------------S-----R------I-------------D-V----- 954
F1.BR.10.10BR_RJ015          --I-----------L----------------I-----------------T---S---K--------Q----------------------------I-----------------------------------I---S--------------------------D-V----- 957
F1.CY.08.CY222               E-I----------------------K-----I----------I----------SGA-K--------Q---E-----------K-I--K--------K----------------------------------I------------------------------D-I----- 953
F1.ES.02.ES_X845_4           --I---------------------------------------M----------S-A-K--------Q----------------------------------------------------------------I----------------I-------------D-V----- 956
F1.RO.96.BCI_R07             -------------------------------I----------R----------SGA-K--------Q--------------------------------------------------------D-------I---S------------I-------------D-V----- 954
F1.RU.08.D88_845             --I----------------------------I----------I----------ST--K--------Q------------------------------------R---------------------------I---S--------------------------D-V----- 954
F2.CM.02.02CM_0016BBY        --I----------------------------I----------I----------STV-K--------Q----------------------------------------------------------------I----------------S--------F----D-I----- 954
F2.CM.10.DEMF210CM001        --I----------------------------I---------------------STV-K-------VN---------------------------------------------------------C------I---S-L---A-----S---K----F----D-I----- 959
F2.CM.10.DEMF210CM007        --I---------------------------------------I----------S-A-K--------Q----------------------------------------------------------------I---S--------------------------D-I----- 953

G.BE.96.DRCBL                --I----------------------------I----------I----------S-A-K------S-T----------------------------------------------------------------I---S--------------------------D-I----- 950
G.CM.10.DEMG10CM008          -QI------------------------A---I----------V------R---SNA-K------N-T------------------------------E---------------------------------I---S--------------------------D-I----- 953
G.CN.08.GX_2084_08           --I----------------------------I----------I----------SN--K------N-T--------------------------E-------------------------------------I---S------------SR------------D-I----- 958
G.CU.99.Cu74                 --I----------------------------I----------V----------S-A-K------N-T----------------------------------------------------------------I---S---------------------F----D-I----- 950
G.ES.09.X2634_2              --I----------------------------I----------I----------S---K------N-T--------------------------M---------R---------------------------I---S-L-----------------------D-I----- 958
G.GH.03.03GH175G             --I-I--------------------------I----------VL---------S-A-K------N-T----------------------------------------------------------------I---S---------------------F----D-I----- 954
G.KE.09.DEMG09KE001          ----M----------------------A---I----------V---------IS-A-K------N-T------------------------------------R--------------------------MI---S--------------------------D-I----- 952
G.NG.09.09NG_SC62            --I------------------T---------I----------V----------SN--K------N-T--------------------K-----------T-------------------------------I---S--------------------------D-I----- 953
G.ZA.01.TV546                --I-I-------------------------------------V----------SQA-K------S-T----------------------------------------------------------------I------------------------------D-I----- 953

H.BE.93.VI991                --I----------------------------I----------M----------S-A-K------D-H------------------------------------R-----------------------R---I-------P--------SQ--K---------D-I----- 952
H.CF.90.056                  -Q-----------------------K----------S-----V----------S-A-K------D-Q----------------------------------------------------------------I----------------SN--K---------D-I----- 953
H.GB.00.00GBAC4001           -------------------------------I----------I----------SGA-K-------VQ----------------------------------------------------------------I----------------L---K---------D-I----- 952

J.CM.04.04CMU11421           --I--------------------------F-I----------V----------S---K--------X------------------------------E-----R-------------------D-T-----I---------R--------------------D-I----- 950
J.SE.93.SE9280_7887          ---------------------------A-F-I----------V----------SGA-K------D--------------------------------E---------------------------------I---------R--------------------D-I----- 952
J.SE.94.SE9173_7022          ---------------------------A-F-I-----G----A----------SGA-K------D--------------------------------E---------------------------------I------------------------------D-I----- 952

K.CD.97.97ZR_EQTB11          --I----------------------------I---------RV----------S-V-K------D--------------------------------E---------------------------------I------------------------------E-I----- 954
K.CM.96.96CM_MP535           --I----------------------------I----------V------T---STV-K-------V----------------------------------------------------------------------------------LN--K---------E-I----- 954

U.CA.01.TV749                --I----------------------------I----------V------G--IS-A-K------N--------------------------------E---------------------------------I---------RA-------------------D-I----- 959
U.CA.99.TV721                --I----------------------------I----------IL-----P--IS-A-K------N--------------------------------E-----------------------------------M----L---I-------------------D-I----- 958
U.CD.83.83CD003_Z3           --------------L----------------V---------TV----------S-A-K--------Q----------------------------------------------------------------I------------------------------D-I----- 955
U.CD.90.90CD121E12           --I-----------------------D----V---------T-----------S-A-K--------Q------------------------------E---------------------------------I----------------IQ------------D-I----- 954
U.CY.05.CY090                --------------------S----------I----------V----------S-V-K---------------------------------------E---------------------------------I---------RL-----LQ------------D-I----- 954
U.CY.08.CY223                ---------------------------A--------------V----------SSALK------DV-----------------------------------------------------------------I----------Q-----S-------------D-I----- 955
U.ES.10.DEURF10DZ001         --I----------------------------I----------V------P---S-A-K------NVT----------------------------------------------------------------I---S---------------R----------D-V----- 954
U.GR.99.99GR303              ----P------N------------------------------V----------S-AMK------NVT--------S-------------------------------------------------------I---------R------I-------------D-I----- 952
U.NL.95.U_NL_95_H10986_D1    --I----A----------------------YI----------V------P---S-A-K-------------------------------------------------------------------------I-M-------T--------------------D-I----- 954

01_AE.AF.07.569M             --I-M-------------------------------------V----------S-A-K------NVR------------------------------E---------------------------------I-M--------------IN------------D-I----- 954
01_AE.CF.90.90CF11697        --------------------------------------*---V----------S-AMK------NVQ------------------------V-----E---------------------------------I-----E------------------------D-I----- 951
01_AE.CN.10.YNFL03           ------------------------------------------V----------S-A-K------NVR----------------------------------------------------------------I---------T--------------------D-I----- 953
01_AE.HK.04.HK001            ------------------------------------------V----------S-A-K------NVR--------------------D---------E---------------------------------I------------------------------D-I----- 951
01_AE.IR.10.10IR.THR48F      --I---------------------------------------V----------S-A-K------NVR------------------------------E---------------------------------I---------R--------------------D-I----- 946
01_AE.JP.x.JRC77AE           ------------------------------------------V----------S-A-K------NVR----------------------------I-E---------------------------------I---------R---NR---------------D-I----- 949
01_AE.TH.09.AA090a_WG11      -------------------------------I----------V----------S-A-K------NVQ------------------------------E---------------------------------I------------------------------D-I----- 954
01_AE.TH.90.CM240            ------------------------------------------V----------S-A-K------NVQ------------------------------E---------------------------------I------------------------------D-I----- 959
01_AE.US.05.306163_FL        --I-I-------------------------------------V----------STA-K------NVR------------------------------E---------------------------------I------------------------------D-I----- 952
01_AE.VN.98.98VNND15         ------------------------------------------V----------S-A-K------NVR------------------------------E-------------------------E-------I------------------------------D-I----- 953
```

```
B.FR.83.HXB2                    GKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTGATVRAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRNPLWKGPA 954
02_AG.CM.10.DE00210CM013        -----------------------------F-I----------V----------S-A-K------NVT----------------------------------------------------------------I----S---I---------------------D-I----- 954
02_AG.CY.09.CY256               ---------------------TDP-------I----------I----------S-ALK----R-NVT---------------K------------------------------------------------I---SS---------------H---------D-I----- 954
02_AG.ES.06.P1423               --I-I--------------------------I----------I----------S-A-K------NVT----------------------------------------------------------------I----S-----------I-------------D-I----- 953
02_AG.GW.05.CC_0048             -------------------------------I----------V----------S-A-K------NVT---------------------------------------------------------------------S-------------------------D-I----- 952
02_AG.KR.12.12MHR9              --I----------------------------I----------V----------S-A-K------NVT-----------------------N------E--------------------------------MI--L-S---------H-I---------G---H-I----- 958
02_AG.LR.x.POC44951             --I----------------------------I----------V----------S-A--------NVTH-------------------------------------------------------R-------I----S------------Q------------D-I----- 954
02_AG.NG.09.09NG_SC61           --I----------------------------I----------V----------S-A-K------DVT----------------------------------------------------------------I----S------------Q------------D-I----- 954
02_AG.NG.x.IBNG                 --I----------------------------I----------V----------S-A-K------NVT----------------------------------------------------------------I----S-----------I-------------D-I----- 954
02_AG.SE.94.SE7812              --I----------------------------I----------V----------S-A-K------NVT--------------------------------------------L------R------------I----S-------------------------D-I----- 954
02_AG.SN.98.98SE_MP1211         --I-I--------------------------I-----KMAS-VV---------S-A-K------KVTPGI-NS---------------------------------------------R------------I----S------------N--K---------D-I----- 954
03_AB.RU.97.KAL153_2            --I----------------------------V----------I----------STA-K-----------------------------Q--QT---------------------------------------I----------------I-------------D-I----- 954
04_cpx.CY.94.94CY032_3          ----M--------------------------I----------M--A---P---S-A-K------D-N----------------------------------------------------------------I----S-------------------------E-I----- 953
05_DF.BE.x.VI1310               --------------------------D----I----------M------P---SGA-K--------Q------------------------------------------T---------------G-----I---T------------I-------------D-V----- 954
06_cpx.AU.96.BFP90              --I----------------------------I----------V----------S-A-K------N-T------------------------------E---------------------------------I----S-----------I-------------D-I----- 956
07_BC.CN.98.98CN009             --I-----------M----------------I----------V----------STA-K--------Q----------------------------------------------------------------I------------------------------D-S----- 954
08_BC.CN.97.97CNGX_6F           --I----------------------------I----------V----------S-A-K--------Q------------------------L-------------------------------------------------R------I-------------D-I----- 950
09_cpx.GH.96.96GH2911           --I----------------------------I----------V----------S-A-K------NVT----------------------------------------------------R-----------I----S-------------------------D-I----- 953
10_CD.TZ.96.96TZ_BF061          ------------------------------------------VV---------S-A-K--------------------------------------------Q----------------------------I----------------I-------------G-I----- 953
11_cpx.CM.95.95CM_1816          -------------------------------I----------V----------S-I-K------D-Q------------------------------E----------------------------------------L-------------K---------D------- 952
12_BF.AR.99.ARMA159             --T-----------L----------------I-----------------P--IS-M-K--------Q--------------------------S-------------------------------------I---S-----R------I-------------D-V----- 954
13_cpx.CM.96.96CM_1849          --I----------------------------I----------V----------S-A-K------N-T------------------------------E---------------------------------I------L---------I-------------D-I----- 961
14_BG.ES.05.X1870               --I----------------------------I-----------------T---S-A-K------N-T----------------------------------------------------------------I----S--------------------F----D-I----- 953
15_01B.TH.99.99TH_MU2079        --------------V---------------------------V----------S-AMK------NVR---WD------------I----------I-E---------P-----------------------I-L-----P---PN-S.--------------D-I----- 952
16_A2D.KR.97.97KR004            --I----------------------------I----------V------P--IS-P-K-------VQ----------------------------I-------------------S---------------I-------------R----------------D-I----- 954
17_BF.AR.99.ARMA038             --I-----------L----------------I-----------------P--SSNA-K--------Q--------------------------R-I-----------------------------------I---S-----R------I-------------D-V----- 954
18_cpx.CU.99.CU76               --I-I-------------------------L-----------V------P---S-A-K------N-T----------------------------I----------------------------------------SE-----F------------------D-I----- 952
19_cpx.CU.99.CU7                ------------------------------------A-----V------P---S-A-K------NVQ----------------------------------------------------------------I----S--------------S----------D-I----- 953
20_BG.CU.99.Cu103               --I------------------S---------I----------V------P--IS-A-K------N-T----------------------------------------------------------------I----S-------------------------D-I----- 950
21_A2D.KE.99.KER2003            ------------------------------------------V------P--IS-A-K-----------------------------------------------------------------D-------I-------------R-----------F----D-I----- 953
22_01A1.CM.01.01CM_0001BBY      -------------------------------I----------V----------S-A-K------E-Q----------------------------I-E---------------------------------I------------------------------D-I----- 954
23_BG.CU.03.CB118               --I----------------------------I----------V------P--IS-A-K------D-T----------------------------------------------------------------I------------------------------D-I----- 950
24_BG.ES.08.X2456_2             -------------------------------I----------V------P--IS-A-K------D-T----------------------------------------------------------------I----S--------------------F----D-I----- 950
25_cpx.CM.02.1918LE             --I---------------------------------------VV-----P---S-A-K-----------------------------------T---E---------------------------------I----S-----------I-------------D-I----- 953
26_AU.CD.02.02CD_MBTB047        --------------M----------------I---------RVV---------S-A-K------N-Q------------------------------E---------------------------------I---------R------I-------------D-I----- 952
27_cpx.FR.04.04CD_FR_KZS        --I---------------------------------------V----------S-AMK-------V-----------------------------------------------------------------I----S----R--------------------D-V----- 952
28_BF.BR.99.BREPM12609          --I----------------------------------------V---------SNA-K------E----------------I-------------------------------------------------I---S-----R--------------------D-V----- 954
29_BF.BR.01.BREPM16704          --------------L----------------I-----------------T---S---K--------Q----V-----------------------------------------------------------I----------------I-------------D-V----- 954
31_BC.BR.04.04BR142             --I----------------------------I----------V----------SN--K--------Q----------------------------------------------------------------I------------------------------D-I----- 960
32_06A1.EE.01.EE0369            --I-I--------------------------I----------V----------S-A-K------N-T----------------------------------------------------------------I----S-------------------------D-I----- 954
33_01B.ID.07.JKT189_C           ------------------------------------------V----------S-A-K------NVQ-------------I----------------E---------------------------------I------------------------------D-I----- 954
34_01B.TH.99.OUR2478P           ----------------T-------------------------V----------S-A-K------NVR----------------------------------------------------------------I------------------------------D-I----- 954
35_AD.AF.07.169H                ------------------------------------------VV---------S-AFK------S-Q------------------------------E---------------------------------I------------------------------D------- 953
36_cpx.CM.00.00CMNYU830         --I-I---------------S---------------------V----------S-A-K------D-Q----------------------------------------------------------------I----S-------------------------D-I----- 954
37_cpx.CM.00.00CMNYU926         -----------------------------------------RV----------S-A-K------NV-----------------------------------------------------------------I----S-------------------------D-I----- 951
38_BF1.UY.03.UY03_3389          --------------L----------------I-----------------P--IS---K--------Q--------------------------N------------------F------------------I---SS----R-F----I-------FS----D-V----- 958
39_BF.BR.04.04BRRJ179           -------------------------------I---------------------ST--K-------------------------------------------------L----V-------------------------------------------------D------- 957
40_BF.BR.05.05BRRJ055           -------------------------------I-----------------P--IS-A-K--------R----------------------------I-----------------------------------I---S-----R---------------F----D-V----- 954
42_BF.LU.06.luBF_18_06          --I-----------L----------------I---------------------S-V-K-----------------------------------E-------------------------------------I------------------------------D-I----- 958
43_02G.SA.03.J11223             ----M---------------------D----I----------V----------S-A--------SVT----------------------------------------------------------------I----S-L-----------------------D-I----- 953
44_BF.CL.00.CH80                ----I---------L----------------------------------P---S-A-K------D-Q--------------------------R-------------------------------------I---------R------I-------------D-V----- 954
45_cpx.FR.04.04FR_AUK           --I-I--------------------------I----------V----------S-A-K------DV-------------------------------E--------------------R------------I------------------------------D-I----- 952
46_BF.BR.07.07BR_FPS625         --I-----------L----------------I-----------------T---S-A-K--------Q--------------------------K-------------------------------------I---S-----R------I-------------D-V----- 954
47_BF.ES.08.P1942               -----------------------------------------------------ST--K--------------------------I----------I----------------------R-------------------L-----------------------D-V----- 953
48_01B.MY.07.07MYKT021          --------------------------G----P----------V------------A-K------NVR------------------------------E---------------------------------I-M----------------------------D-I----- 954
49_cpx.GM.03.N26677             --I-----------M---------------YI----------V----------S-A-K--------Q--------------------------E-------------------------------------I------------------------------D-I----- 951
50_A1D.GB.10.12792              ------------------------------------------VV---------S-A-K------N-Q--------------------------T-------------------------------------I---SS-------------------------D------- 952
51_01B.SG.11.11SG_HM021         --I----------------------------I---------T-------P---STAAK-------------------------------------I-----------------------------T--------V-------------------------N-D-I----- 954
52_01B.MY.03.03MYKL018_1        -------------------------------I----------V----------S-A-K------NVQ------------------------------E---------------------------------I------------------------------D-I----- 954
53_01B.MY.11.11FIR164           ------------------------------------------V----------S-A-K------NVQ------------------------------E---------------------------------I-M----------------------------D-I----- 954
54_01B.MY.09.09MYSB023          D-I----------------------------I-----------V-----R---SNA-K-------V---------------------------E-----------------------------------------------R--------------------D------- 954
55_01B.CN.10.HNCS102056         --I-----------V----------------I---------------------S-A-K------NVQ------------------------------E---------------------------------I------------------------------D-I----- 954
56_cpx.FR.10.URF5_patient_A     -------------------------------I--------------------IS---K-------------------------------------------------------------------------I----S---S---------L-----------D------- 956
57_BC.CN.09.09YNLX19sg          --I----------------------------I----------VV---Y-----S-A-K--------T----------------------------------------------------------------I----S-----------I-V-----------D-I----- 950
58_01B.MY.09.09MYPR37           --------------------------D---------------V----------S-A-K------NVR------------------------------E---------------------------------I------------------------------D-I----- 954
59_01B.CN.09.09LNA423           ------------------I------------I----------V----------S-A-K------NVQ------------------------------E---------------------------------I------------------------------D-I----- 951
60_BC.IT.11.BAV499              --I----------------------------I----------V----------SN--K--------Q----------------------------------------------------------------I----------------I-------------DAI----- 949
61_BC.CN.10.JL100010            --I-----------M---------------------------V----------S-A-K-------V------------------------T----------------------------------------I------------------------------D-I----- 950
62_BC.CN.10.YNFL13              --I----------------------------I----------V----------SSA-K--------Q-----------------------R----------------------------------------I----------------I-V-----------D-I----- 950
63_02A1.RU.10.10RU6637          --I--------------------------------------RVV-----P---SSA-K------N-Q----------------------------------------------------------------I------------------------------D-I----- 954
64_BC.CN.09.YNFL31              --I-----------M----------------I----------V----------S-A-K------D--------------------------------E---------------------------------I---------R--------------------D-S----- 954
65_cpx.CN.10.YNFL01             --I-----------M----------------I----------V----------STA-K--------Q----------------------------------------------------------------I------------------------------D-I----- 950
67_01B.CN.11.2011.ANHUI.HF115   -------------------------------I---------------------SS--K----------------------------N----------E---------------------------------I------------------------------D-I----- 952
68_01B.CN.11.2011.ANHUI.WH73    -------------------------------I---------------------SGA-K----------------------------N--------I-E---------------------------------I------------------------------D-I----- 952
70_BF1.BR.10.10BR_PE004         --I--------------------------------------------------SNA-K-------------------------------------I---------------L-------------------I------------------------------D-I----- 956
71_BF1.BR.10.10BR_PE008         -------------------------------I---------------------S---K------------------------------------------------------------------------------------------I----------E--D-V----- 954
72_BF1.BR.10.10BR_MG002         --------------L----------------I----------I----------S---K-------------------------------------I-----------------------------------I------------------------------D-V----- 955
O.BE.87.ANT70                   --I-I--------F----------------------A-----V------P---ST-MK------N-QH---------------A------S--Q---------R--------V------------T-----I--L-SQ---T------L-XHK---------D-I----- 950
O.CM.98.98CMA104                ----I--------F----------------------A-----V------P---S--MK------N-XH---------------A------S--E---------R---------------------T----LI--L-SQ---T------L-------------D-I----- 950
O.CM.98.98CMABB141              ----I--------F----------------------A-----V------P---S--MK-----TN-QH---------------A------S--Q------------------V------------T-----I--L-SH---T------S-------------D-I----- 950
O.CM.98.98CMABB212              --I-I--------F---------S------------A-----V------P---S-AMK-----T--QH-------------------N--S--Q------------------V------------T----LI--L-SQ---T------L-------------D-I----- 950
O.CM.98.98CMU5337               --I-I--------F---------S------------A-----I------P---S-IMK-----TN--H---------------A------S--Q---------R--------V--------L---T-----I--L-SQ---T------F---K-Q-------D-I----- 950
O.CM.99.99CMU4122               ----I--------F----------------------A-----V------P---S--MK-----TN--H---------------A------S--Q---------R--------V------------T-----I--L-SQ---T------F---K-Q-------D-T----- 950
O.FR.92.VAU                     --I-I--------F----------------------A---I-IL-----P---S-AMK------N-QH---------------------ES-MQ------------------V------------T----LI--L-SQ---T------L-------------D-I----- 951
O.GA.11.11Gab6352               ----I--------F----------------------S-----V------P---ST-MK-----TN--H---------------A------S--Q------------------V------------T-----I--L-SH---T------L-------------D-I----- 950
O.SN.99.99SE_MP1299             --I-I--------F----------------------A-----V------P---S-AMK------N--H---------------A------S--Q------------------V--Y---------T-----I--L-SQ---T------F---K-Q-------D-I----- 950
O.US.10.LTNP                    --I-I--------F----------------------A-----I------P---S--MK------N-QH---------------A------S--Q---------R--------V------------T-----I--L-SQ---T------L-------------D-I----- 950
N.CM.02.DJO0131                 --I-----------L----------------I----------V------P---S---K------N-T----------------------------I---------------------------E-T-----I---------TK--T--L-------------D-I----- 959
N.CM.02.SJGddd                  --I-----------L----------------I----------V------P---S---K------N-Q-------------A----------------------R---------------------T-----I---------TK--T--LQV-----------D-I----- 956
N.CM.04.04CM_1015_04            --I-----------L----------X-----I----------V------P---S---K------S---------------A--------------I-----------------------------T-----I---------TK--T--L-V-----------D-I----- 961
N.CM.06.U14296                  --I-----------L------------A---I----------V------X---S---K------N---------------X--------------I-----------------------------T-----I----------N--T--L-V-----------D-I----- 961
N.CM.06.U14842                  --I-----------L----------------I----A-----V----------S---K----------------------A--------------I-----------------------------T-----I-------------T--L-V------F----D-I----- 962
N.FR.11.N1_FR_2011              ----I---------L----------------M----------V------P---S---K------N---------------A--------------I-----------------------------T-----I---------TK--T--L-V-----------D-I----- 961
P.CM.06.U14788                  --I----------FM------V-----A---V----S-----V------P---S-A-K-----LN-TH-------------------------H------------------V------------T--D--I--L--Q---T------L-----Q-------D-I----- 951
P.FR.09.RBF168                  --I----------FL----------------V----A-----V------P---S-A-K-----LN-TH-------------------------L------------------V------------T--D--I--M--Q---T------L-------------D-I----- 950
CPZ.CD.06.BF1167                ---V---------F-----LSS---K---F-I----------K------A---S-A-K------N-Q--------------------Q--I-VD-I-E-----R---A---Y-------------TP-D--L--LT------------F-L-----------D-V----- 957
CPZ.CM.05.SIVcpzMT145           --I-I--------------------------V-----------------P---SNA-K--------R--------------------------E-------------------------------------I----------------L---Q--------KD-I----- 953
CPZ.GA.88.GAB1                  ----I--------------------------------------------P---S-A-K------D-------------------L----------------------------------------T-----I---------S------L-V-K---------D-I----- 952
CPZ.TZ.06.SIVcpzTAN13           ----I--------F-------E---K-----I----------R------P---STA-K------Q-QH-------------------Q--Q--E-I------Q----V---Y-----------E-T----LL--LT-N---QQ-----L-V---------A-D-I----- 949
CPZ.US.85.US_Marilyn            --I-I---------L----------------I----------V----------SS-----------Q----------------------------I-----------------------------T-----I----SEL--DL-----L-V-----------D-I----- 956
GOR.CM.07.SIVgor2139_287        --I----------F----------------------A-----I------P---S---K-----LNVTH-------------------------Q------------------V------------T--D--I-ML--Q---T------L-------------D-I----- 946
GOR.CM.07.SIVgorCP2135con       --I----------F----------------------A-----I------P---S---K-----LNVTH-------------------------Q------------------V------------T--D--I-ML--Q---T------L-------------D-I----- 946
```

```
                                                                                p31 Integrase end
                                                                                         Pol end
B.FR.83.HXB2                KLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED..........*  1003
A1.AU.03.PS1044_Day0        ---------------T--------------------------G--------------  1004
A1.CH.03.HIV_CH_BID_V3538   ----------------------------------------G--------------  1004
A1.CY.08.CY236              ----------------------------------------G--------------  1001
A1.ES.05.X1608_8            ----------------------------------------G--------------  1008
A1.KE.11.DEMA11KE001        ---------------------------K------------G--------------  1004
A1.RU.11.11RU6950           ---------------N----------------------------------------  1004
A1.RW.11.DEMA111RW002       ------------------------V---------------G--------------  1004
A1.SN.01.DDI579             -------------------I--------------------G--------------  1003
A1.UG.11.DEMA110UG001       ----------------E---------------------G--G--------------  1004
A1.ZA.04.04ZASK162B1        R---------------E----------------------------------------  1003
A2.CM.01.01CM_1445MV        ------------------------V---------------G--------------  1004
A2.CY.94.94CY017_41         ----------------------------------------G--------------  1004
B.BR.10.10BR_MG029          ----------------E-----------------------------------------  1004
B.CA.07.502_1191_03         ----------------E-----------------------------------------  1004
B.CH.08.M2_0803101_NFLG8    -------------V--------------------------------------------  1008
B.CN.10.DEMB10CN002         ----------------E-----------------------------------------  1004
B.ES.10.DEMB10ES002         ------------------I---------------------G--------------  1004
B.FR.11.DEMB11FR001         ------------------------V---------------------------------  1003
B.GB.05.MM45d213_GN1        ----------------------------------------------------------  1003
B.HK.06.HK003               ----------------------------------------------------------  1008
B.HT.05.05HT_129389         ----------------------------------------------------------  1003
B.JP.12.DEMB12JP001         ----------------------------------------------------------  1007
B.KR.07.HP_18_07JHS10_3909  ------------------------V---------------M----------------  1003
B.PE.07.502_0525_wg5        ----------------------------------------------------------  1008
B.RU.11.11RU21n             ------------------------V---------------------------------  1008
B.TH.08.MERLBDTRC10         -------------E--------------------------------------------  1004
B.US.11.ES38                ----------------------------------------------------------  1008
C.AR.01.ARG4006             -----------L-------------V---K--------A--------------------  1000
C.BR.07.DEMC07BR003         -----------L----------V---K--------A--------------------  1000
C.BW.00.00BW5031_1          ------------------A---------K---Q----A-------G---N---------  1006
C.CN.10.YNFL19              ------------------------------XK-------A----G-----QNME-----  1004
C.CY.09.CY260               ----------------------------E-------A----G-----QNME-----  1011
C.ES.08.X2363_2             ------------------------------------A----G--------------  1007
C.ET.02.02ET_288            ------------------------------------A----G----N---------  1006
C.IN.09.T125_2139           ---------------------------K-------A----G--------------  1001
C.KE.00.KER2010             ---------------------------K-------A----G--------------  1000
C.MW.09.703010256_CH256.w96 ---------------------------K-------A----G--------------  1002
C.TZ.08.707010457_CH457.w8  -------------K---------V---KH------A----D-----QNME-----  1004
C.YE.02.02YE511             ----------------------V---K-------A----G--------------  998
C.ZA.08.705010534_CH534.w12 ---------I------------------K-------A----D--------------  1007
C.ZA.10.DEMC10ZA001         --------------K--------------------A----G--------------  1000
C.ZM.11.DEMC11ZM006         ----------------E-------------------A----G-----QNME-----  1009
D.CM.10.DEMD10CM009         ----------------------------------------G--------------  1009
D.CY.06.CY163               ------------------------V------------G----G--------------  1003
D.KE.11.DEMD11KE003         ------------------------V-----------------------------------  1003
D.KR.04.04KBH8              --------------S-------------------------------------------  1003
D.SN.90.SE365               ----------------E-----------------------------------------  1009
D.TZ.01.A280                -------------H-----------------V-------------------------  1005
D.UG.10.DEMD10UG004         ----------------E-----------------------------------------  1003
D.UG.11.DEMD11UG003         ------------------------V-----------------------------------  1005
D.YE.02.02YE516             ------------------------V---------------G--------------  1005
D.ZA.90.R1                  ------------------------V-----------------------------------  1003
F1.AO.06.AO_06_ANG32        ----------------E-------------------------G--------------  1008
F1.AR.02.ARE933             ----------------E-----------------------M-G--------------  1004
F1.BR.10.10BR_RJ015         ----------------E---L---------------------G--------------  1007
F1.CY.08.CY222              ----------------E---I-----------------------G--------------  1003
F1.ES.02.ES_X845_4          ---------------RE-----------------------G--------------  1006
F1.RO.96.BCI_R07            ----------------E--------------V---------G--------------  1004
F1.RU.08.D88_845            -------------E--E-----------------------G--------------  1004
F2.CM.02.02CM_0016BBY       ---------------NE--------V----------------G--------------  1004
F2.CM.10.DEMF210CM001       ---------------NE---------------X--------G--------------  1009
F2.CM.10.DEMF210CM007       ---------------NE-----------------------------A----------  1003
G.BE.96.DRCBL               ---------------NE-----------------------G--------------  1000
G.CM.10.DEMG10CM008         ---------------NE-----------------------G--------------  1003
G.CN.08.GX_2084_08          ---------------NE-----------------------G--------------  1008
G.CU.99.Cu74                ---------------NE--------------------A---G--------------  1000
G.ES.09.X2634_2             ---------------NE-----------------------G--------------  1008
G.GH.03.03GH175G            ---------------NE------------K----------G--------------  1004
G.KE.09.DEMG09KE001         ---------------N------------------------G--------------  1002
G.NG.09.09NG_SC62           ----------------E-----------------------G--------------  1003
G.ZA.01.TV546               ---------------NE----------VL-------------G--------------  1003
H.BE.93.VI991               ----------------E-----------------------G--------------  1002
H.CF.90.056                 ----------------E------E----------------------------------  1003
H.GB.00.00GBAC4001          ----------------E-----------------------G--------------  1002
J.CM.04.04CMU11421          ------------------------V---XX-----------G--X-----------  1000
J.SE.93.SE9280_7887         --P-------------E-----------------------G--------------  1002
J.SE.94.SE9173_7022         ----------------E-----------------------G--------------  1002
K.CD.97.97ZR_EQTB11         -----------#.Y--E-----------------------G--------------  1002
K.CM.96.96CM_MP535          ----------------E-----------------------G--------------  1004
U.CA.01.TV749               -------------E--E----------------------D--------------  1009
U.CA.99.TV721               ----------------------------------------G--------------  1008
U.CD.83.83CD003_Z3          ----------------E-----------------------------N----------  1005
U.CD.90.90CD121E12          ----------------E-----------------------------N----------  1004
U.CY.05.CY090               -------R--------E-----------------------G--------------  1004
U.CY.08.CY223               ----------------E-----------------------G--------------  1005
U.ES.10.DEURF10DZ001        ----------------E-----------------------G--------------  1004
U.GR.99.99GR303             ----------------E-----------------------G--------------  1002
U.NL.95.U_NL_95_H10986_D1   ----------------E-----------------V-----------------------  1004
01_AE.AF.07.569M            ----------------E-----------------------G--------------  1004
01_AE.CF.90.90CF11697       ----------------------------------------G---N----------  1001
01_AE.CN.10.YNFL03          ---------------------------K------------G--N-----------  1003
01_AE.HK.04.HK001           ----------------------------------------G--------------  1001
01_AE.IR.10.10IR.THR48F     ----------------------------------------G--------------  996
01_AE.JP.x.JRC77AE          ----------------------------------------G--------------  999
01_AE.TH.09.AA090a_WG11     ----------------------------------------G--------------  1004
01_AE.TH.90.CM240           ----------------------------------------G--------------  1009
01_AE.US.05.306163_FL       ----------------------------------------G--------------  1002
01_AE.VN.98.98VNND15        ----------------------------------------G--------------  1003
```

```
                                                                       p31 Integrase end
                                                                                 Pol end
B.FR.83.HXB2                 KLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED..........*  1003
02_AG.CM.10.DE00210CM013     -------------------------------------------GG-T--..........-  1004
02_AG.CY.09.CY256            -----------------------------K-------------G--E--..........-  1004
02_AG.ES.06.P1423            -------------------------V------------N----G-----..........-  1003
02_AG.GW.05.CC_0048          -----------------------------K-------------G----N..........-  1002
02_AG.KR.12.12MHR9           --------------------------N----------------G-----..........-  1008
02_AG.LR.x.POC44951          -------------------------V--L--------------G-----..........-  1004
02_AG.NG.09.09NG_SC61        ---------------N---------------------------G-----..........-  1004
02_AG.NG.x.IBNG              -------------------------------------------G-----..........-  1004
02_AG.SE.94.SE7812           -------------------------V--V--------------G-----..........-  1004
02_AG.SN.98.98SE_MP1211      -------------------------T--L--------------GG-N--..........-  1004
03_AB.RU.97.KAL153_2         ---------------N---------------------------------..........-  1004
04_cpx.CY.94.94CY032_3       --------------------------------------N----G-----..........-  1003
05_DF.BE.x.VI1310            ----------------E--------------------------G-----..........-  1004
06_cpx.AU.96.BFP90           ----------------E--------------------------G-----..........-  1006
07_BC.CN.98.98CN009          -----------------------------K--------A----G-----..........-  1004
08_BC.CN.97.97CNGX_6F        -----------------------------K--------A----G-----..........-  1000
09_cpx.GH.96.96GH2911        ---------------N--------------H------------G-----..........-  1003
10_CD.TZ.96.96TZ_BF061       -------------E-----------V---K--------A----------QNME......-  1007
11_cpx.CM.95.95CM_1816       --------------------I----------------------G-----..........-  1002
12_BF.AR.99.ARMA159          --------------D-E--------------------------G-----..........-  1004
13_cpx.CM.96.96CM_1849       ----------------E--------------------------G-----..........-  1011
14_BG.ES.05.X1870            ---------------NE--------------------------G-----..........-  1003
15_01B.TH.99.99TH_MU2079     -------------------------------------------G-----..........-  1002
16_A2D.KR.97.97KR004         ------------------------R------------------G-----..........-  1004
17_BF.AR.99.ARMA038          ----------------E--------------------------G-----..........-  1004
18_cpx.CU.99.CU76            ----------------E--------------------------G-----..........-  1002
19_cpx.CU.99.CU7             -------------------------------------------G-----..........-  1003
20_BG.CU.99.Cu103            ---------------NE--------------------------G-----..........-  1000
21_A2D.KE.99.KER2003         -------------------------V-----------------------..........-  1003
22_01A1.CM.01.01CM_0001BBY   ------------------------------E------------G----N..........-  1004
23_BG.CU.03.CB118            --------------SNE--------------------------G-----..........-  1000
24_BG.ES.08.X2456_2          ---------------NE-------------H------------G-----..........-  1000
25_cpx.CM.02.1918LE          ----------------E--------------------------G-----..........-  1003
26_AU.CD.02.02CD_MBTB047     -------------------------------------------G-----..........-  1002
27_cpx.FR.04.04CD_FR_KZS     ----------------E-----------LKH------------G-----..........-  1002
28_BF.BR.99.BREPM12609       -------------------------------------------------..........-  1004
29_BF.BR.01.BREPM16704       ----------------E--------------------------G-----..........-  1004
31_BC.BR.04.04BR142          -----------L-------------V---K--------A--M-------QNME......-  1014
32_06A1.EE.01.EE0369         ----------------E-------------E--------E---G-----..........-  1004
33_01B.ID.07.JKT189_C        -----------------------------------------G-G-----..........-  1004
34_01B.TH.99.OUR2478P        -------------------------------------------G-----..........-  1004
35_AD.AF.07.169H             -------------------------------------------G-----..........-  1003
36_cpx.CM.00.00CMNYU830      ------------------------------N------------G-----..........-  1004
37_cpx.CM.00.00CMNYU926      -----------------------------------------M-G-----..........-  1001
38_BF1.UY.03.UY03_3389       ----------------E--------------------------G-----..........-  1008
39_BF.BR.04.04BRRJ179        ----------------E--------------------------G-----..........-  1007
40_BF.BR.05.05BRRJ055        ----------------E--------V---K-------------------..........-  1004
42_BF.LU.06.luBF_18_06       -------------------------------------------------..........-  1008
43_02G.SA.03.J11223          ----------------E---------------------N----G-----..........-  1003
44_BF.CL.00.CH80             ----------------E--------------------------G-----..........-  1004
45_cpx.FR.04.04FR_AUK        -------------------------V-----------------G-----..........-  1002
46_BF.BR.07.07BR_FPS625      ----------------E--------------------------G-----..........-  1004
47_BF.ES.08.P1942            ----------------E--------------------------G-----..........-  1003
48_01B.MY.07.07MYKT021       -----------------------------------------M-G-----..........-  1004
49_cpx.GM.03.N26677          --------------K-----------------------A----G-----..........-  1001
50_A1D.GB.10.12792           --------------------L-----------------S--M-G-----..........-  1002
51_01B.SG.11.11SG_HM021      -------------------------------------------G-----..........-  1004
52_01B.MY.03.03MYKL018_1     -------------------------------------------G-----..........-  1004
53_01B.MY.11.11FIR164        -------------------------------------------G-----..........-  1004
54_01B.MY.09.09MYSB023       -------------------------------------------------..........-  1004
55_01B.CN.10.HNCS102056      -------------------------------------------G-----..........-  1004
56_cpx.FR.10.URF5_patient_A  -------------------------------------------G-----..........-  1006
57_BC.CN.09.09YNLX19sg       -----------------------------K-------------G-----..........-  1000
58_01B.MY.09.09MYPR37        -------------------------------------------G-----..........-  1004
59_01B.CN.09.09LNA423        -------------------------------------------G-----..........-  1001
60_BC.IT.11.BAV499           -----------L-------------V---K--------A--M-------..........-   999
61_BC.CN.10.JL100010         -----------------------------K--------A----G-----..........-  1000
62_BC.CN.10.YNFL13           --------------K--------------K--------A----G-----..........-  1000
63_02A1.RU.10.10RU6637       ---------------N---------------------------------..........-  1004
64_BC.CN.09.YNFL31           -----------------------------K--------X----G-----..........-  1004
65_cpx.CN.10.YNFL01          -----------------------------K-------------G-----..........-  1000
67_01B.CN.11.2011.ANHUI.HF115-------------------------------------------G-----..........-  1002
68_01B.CN.11.2011.ANHUI.WH73 -------------------------------------------G-----..........-  1002
70_BF1.BR.10.10BR_PE004      ----------------EL-------------------------G-----..........-  1006
71_BF1.BR.10.10BR_PE008      ----------------E--------V-----------------G-----..........-  1004
72_BF1.BR.10.10BR_MG002      ---------------------------V---------------------..........-  1005
O.BE.87.ANT70                Q-------------KG--------------E-------T-SM--G-T-SESVEQPSEIP-  1010
O.CM.98.98CMA104             Q-------------KG--------------H-------A-SM--G-T-SESVEQPGEIS-  1010
O.CM.98.98CMABB141           Q-------------QG----------------------T-SM----T-SESVEQPGEIP-  1010
O.CM.98.98CMABB212           Q-------------KG----------------------N-SM----T-SENMDQLGETP-  1010
O.CM.98.98CMU5337            Q-------------KG--------------H--X----P-SM--G-T-SESVEQPGEIP-  1010
O.CM.99.99CMU4122            Q-------------KG--------------N-------T-SM--G-X-SENVEQLGEIP-  1010
O.FR.92.VAU                  Q-------------KG----------------------A-S---G-T-GESVEQPSEIP-  1011
O.GA.11.11Gab6352            Q-------------KG----------------------S-S---G-T-SENMEQPGEIS-  1010
O.SN.99.99SE_MP1299          Q-------------KG--------------H-------T-SM--G-T-SENVEQPGEIP-  1010
O.US.10.LTNP                 Q-------------KG----------------------T-SM--G-T-SESVEQPSEIP-  1010

N.CM.02.DJO0131              ---------------G-----------------------G----G-A-NQDMA......-  1013
N.CM.02.SJGddd               --------------KG-----------------------G----G---NQEME......-  1010
N.CM.04.04CM_1015_04         ---------------G-----------------------G----G---NQDME......-  1015
N.CM.06.U14296               ---------------G-----------------------G----G---NQDMD......-  1015
N.CM.06.U14842               ---------------G-----------------------G----G---NQDME......-  1016
N.FR.11.N1_FR_2011           ---------------G-----------------------G----G---NQEME......-  1015

P.CM.06.U14788               T-------------KG--------------N------------DT-R-SESLEQSG...-  1008
P.FR.09.RBF168               T-------------KG---------------------------DI-R-SESLE......-  1004

CPZ.CD.06.BF1167             R-----------KEGEE------------K-----V--AGG--D--N--..........-  1007
CPZ.CM.05.SIVcpzMT145        --------------QDE---------------------A-S-------S..........-  1003
CPZ.GA.88.GAB1               T-------------QGEL----------------------------N--..........-  1002
CPZ.TZ.06.SIVcpzTAN13        R-----------KEGE-------------KE-------AGS-DD--N-A..........-   999
CPZ.US.85.US_Marilyn         ------------KE-EEV-----------K--------A-SM-G----S..........-  1006

GOR.CM.07.SIVgor2139_287     T-------------KG---------------------------DT-G-SESLEQSS...-  1003
GOR.CM.07.SIVgorCP2135con    T-------------KG---------------------------DT-G-SESLEQSS...-  1003
```

Vif start

```
B.FR.83.HXB2                 MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHYESPHPRISSEVHIPL.GD.AR....LVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPELADQLIHLYYFDCFSDSAIRKALLGHIVSPRCEYQAGH.NKVGSLQYLALAALIT.....PKK  158
A1.AU.03.PS1044_Day0         ---------------------N-------HI-K-TKK-S----F--RN-KV--------.--.-K....I-VR-----Q---K--Q--H-------L------I--D--------H--N----------I--QV-----------.-----------R--VE.....---  158
A1.CH.03.HIV_CH_BID_V3538    ----------------K----H---------I-K--KE-C-------R--KV--------.-E.--....--VR---------K-----H-------Q------I--D--------IH------E------I--E--R---------.-----------K--VA.....--Q  158
A1.CY.08.CY236               ------------------K--H------------KR-KN-R----F--R--KV--------.R-.-T....--VK-----Q---K--Q--H-------L------I--DM--H-------N----------I--QV---S-D-----.-----------R--VA.....---  158
A1.ES.05.X1608_8             ---------------------N----------K--K-------F--R--KV--------.--.-K....--VI---------K-----H-------L------I--D-------MH--N----------I--Q----S------.S----------K--V-.....QPR  158
A1.KE.11.DEMA11KE001         ---------------------H----Y-----K--KN-C----F-CRN--A--------.--.--....--VR---------KE----H------KLR-----I--DQ----------N-----E------I--QV-------HT--.-----------K--V-.....-T-  158
A1.RU.11.11RU6950            ---------------------N----------K---E-R-------G---V--------.--.--....--VK------A--K-----H-------QG-----I--D---------H-----E------IV---GR---N-P---.-----------K--V-.....-TR  158
A1.RW.11.DEMA111RW002        ---------------------H--------I-R--KD-V-K----CRN--E--------.--.-T....IIVK---------KE-Q--H-------L-S----I--DM-------H--------------I---V---S---LS--.-----------K--VA.....-AR  158
A1.SN.01.DDI579              ---------------------N--------I-K--KD-------DCR---A--------.--.-M....--VR------P--Q-----H-------Q------I--D--------H-------------R-I--Q--R-------T--.-----------R--VA.....S--  158
A1.UG.11.DEMA110UG001        --------V------------N--------I-K-VN--C-------R--KV--------.--.-K....--VR---------K-----H-------L---------D--------H------E------I--QV---S-------.-----------K----.....-T-  158
A1.ZA.04.04ZASK162B1         ---------------------N------I-I-K--K----------R--KV--------.--.--....--VR-----QP--K-----H-A-----Q------I--D------------------------VI--QV-R-----S---.-----------K--V-.....--R  158
A2.CM.01.01CM_1445MV         ---------------------N--------I-R--KE------F--R--KV--------.-E.--....IVR------P--KE----H-------Q-K----I--D---H-----------E------I--EL---K--------.-----------K--VA.....STR  158
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-VVA.....STR  158
B.BR.10.10BR_MG029           ---------------------N--------I-R-----V-------TN-----------.--.-K....-------------------------------------------G-------M---------------I---R------------.-------------------L-.....---  158
B.CA.07.502_1191_03          ------------------------------I-K--N--------D----K---------.--.-K....---A-----------------------R---------NI-------------------------N-----R-R---------.-Q---------V--TA.....T--  158
B.CH.08.M2_0803101_NFLG8     ----------------------------IHI-K--K--I-K-----TN-----------.--.-K....--V-----------------------ERK----I--G-------Q-H--------E----N-I--NR-R---------.-E---------T----.....--R  158
B.CN.10.DEMB10CN002          ----------------K-N-------Y---I-----K-S-------S-----------Y---.--.-K....--V-----------A----H-A------G--------G-------------------------S-IV-R--------------.----------------V--V-.....---  158
B.ES.10.DEMB10ES002          ------------------------------I-----K-I-K-----SN---V--------.--.-K....---K-----------------------RS---N-----D-----------------E----N-I--R--------------.S------H---T------.....--R  158
B.FR.11.DEMB11FR001          --------------------------------K--KK-I-------NN-----------.-E.-N....-----------------------------ER---------T-----------------------KN-I--R---S---------.-------------------K-.....---  158
B.GB.05.MM45d213_GN1         ------A-------------------------K--Q----------T------------.-E.-K....---K----------E------------I-----------N------------------E----NTI---R------------.------------------T--A.....-R-  158
B.HK.06.HK003                -A-------------------H----Y--HI-K--K--S-K-----I--K---------.--.ET....---------Q------M-H----------G-R-----G----K----H---------V---I---R---S--------.-Q---------K-----.....---  158
B.HT.05.05HT_129389          -------------------A------------KR------------T------------.--.--....-------------------------------R---------G-------------------E------I--YR------------.-----------V-----.....---  158
B.JP.12.DEMB12JP001          ------------------NA--------I-I-K--ER-V-------SN-----------.-G.-K....---I-----------------------Q---------N-------------------E----Q-------R---S---R---.S----------T----....PT-R  159
B.KR.07.HP_18_07JHS10_3909   ------------------------------I-K--KE-V-------T------------.--.-K....--------------E---------------N-----D----K----H-------------H-I--RL-R-E---P---.-T----K----T-----.....---  158
B.PE.07.502_0525_wg5         ------------------N----I-----H-----KR-S------GT------------.-E.EKLEEK--------------------------DK---------N------------------E----N-IV-R---------H---.------------------T--V-.....---  162
B.RU.11.11RU21n              ---------------------N------IH--K--S--V-K-----RN-----------.--.-K....---K---------K-------A-----E-----------NI------------------E----N-I--RR---S----------.-------------------R-.....T--  158
B.TH.08.MERLBDTRC10          ------------------N-------Y--H--K-------------T------------.--.-E....--VI-----N----E------A---------------------NV------------------E------I--SR-R---------.-----------T--L-.....---  158
B.US.11.ES38                 --------------------------------R-------------T------------.--.-K....---K---------------------K-R-----------Q------------------E----Q-I---------------------.-------------------VK.....---  158
C.AR.01.ARG4006              -------L-----------A-N--------I-K-----Y-------R--KV--------.--.--....---I-----Q---------H-------LXK-------G-------MH-----T-----------R-------------.-----------T--LK.....---  158
C.BR.07.DEMC07BR003          -------L--------K----N--------I-KR-K--Y-------R--K--------V.-E.-K....---------Q-----A----H---V---LR--------D-------IH-----A-----------R-----D------.-----------T---K.....--R  158
C.BW.00.00BW5031_1           -D----GL-------------N----------RR-N-------H--R--KV--------.--.--....---K----------------H-------LR--N-----G-------MH-----A-------I--R--I---D------.-Q---------T---K.....--R  158
C.CN.10.YNFL19               -------L--------K----N----------KR-N------D-R--KV--------.-E.--....---K-----Q---K-----H-------LRK-N--IE-G-------M------A-------I-----I---D-----N-----------T---K.....--R  159
C.CY.09.CY260                -------L--------K----N----------K--N----------R---V--------.-E.-S....---R-----Q----E----H-------W---------G-------MH-----A-----Q-I-----I---------.K----------T---K.....---  158
C.ES.08.X2363_2              -------L-----------A-N----------KR-N----------R--K---I-----.--.--....-I-K-----Q----E----H-------LG--------G-------MH-----T-------I---R-T---D------.-----------T--VK.....---  158
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.....---  158
C.IN.09.T125_2139            -------L-------------H----------RR-K--V-------R--KV--------.-E.-K....---K-----QP--------H-------LRS-N-----G-------MH--N-AE-------I---L-I---D-P---.T----------T---R.....--Q  158
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-  158
C.MW.09.703010256_CH256.w96  -------L--------K----N----------RR-D----------R--K---------.--.--....---K-----Q----A----H-------L---------D-------MH-----A-----Q-I---V-----D------.-----------T---K.....---  158
C.TZ.08.707010457_CH457.w8   -------L----I--------N----------RR-N--Y-------R--KV--------.--.--....---K-----Q----E----H-------LR--------G-------MH-----A-------I-----T---D------.-Q---------T--VK.....---  158
C.YE.02.02YE511              ------GL-------------N--------I-RR-S----------R--KV--------.-E.--....---K-------------------H-------LR--------G-------MH-----T-------I-----I---D------.T----------T---K.....---  158
C.ZA.08.705010534_CH534.w12  -------L----I---K----N--------I-K--S---K------R--KV--------.--.-K....--VK-----Q----A----H-A-----LRE-------G-------MH-----A---------------------------.-----------T---K.....---  158
C.ZA.10.DEMC10ZA001          ------GL-------------N--------I-KR-K-------F--R--K-----Y---.--.--....I-K-----Q----E----H-------LRK-------G-------MH-----A-----Q-I-----I---D------.-----------T---K.....--G  158
C.ZM.11.DEMC11ZM006          ------GL-------------N--------I-R--S----------R--KV--------.-E.--....---K-----Q----E----H-------LR--------G-------MH-----A-------I---------D-P---.-----------T---K.....---  158
D.CM.10.DEMD10CM009          -------------------I----------I-K-TA--------D----K---------.--.-K....--VI----------------------------R---------G-------M--------E------I---R------------.-----------T---K.....--Q  158
D.CY.06.CY163                --------V-----------------Y---K-K--A--------DNH--K---------.--.-K....--V------N------------------R---------G-------M-------------------I---R------------.-----------T---A.....--Q  158
D.KE.11.DEMD11KE003          --------------------------------KQ-KR-W------C-N-----------.-E.-K....--VK----------E-------------R---------G-------TH-----A-----Q-I---R-IS--N-P---.-----------T--LA.....---  158
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  158
D.SN.90.SE365                ------------------K-------------K--Q--C-K---D----K---------.-E.--....--VK----------E-------------E---------G-------I--------E----Q-I---------D------.-----------T--RA.....-R-  158
D.TZ.01.A280                 --------------------------------R--Q--L-----D-H--K---------.--.--....--V-----------E-------------R---------G-------M------T-------I---------D------.-----------I---K.....---  158
D.UG.10.DEMD10UG004          ------------------K-------------K--K--W--------------------.-E.--....--VK---------K--------------G---------G-------T-----AE-----Q-I-----I---N------.-----------T---K.....--Q  158
D.UG.11.DEMD11UG003          ------------------S---G-----I---KQ-N--L-----DC-N-K---------.-E.--....--VR-----Q----E-------------R---------G-------I-----AE-------I-----F---N------.-----------T--V-.....--R  158
D.YE.02.02YE516              ------------------N-----------I-XE-K--------D----K---------.-E.-K....--VI-----N------------------QR--------S-------T--------E------I--LR-----D------.-----------T---A.....-RR  158
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---A.....S--  159
F1.AO.06.AO_06_ANG32         ------------------N-------Y-IH--K--K--V----F--R---V--------.RK.DK....---------N----E-------------LGK----I--SM--------------------E----R-I--YR-----N------.-----------T---H.....---  158
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-  158
F1.BR.10.10BR_RJ015          ------------X-----N-------Y--HI-R--K--C-------RN-----------.DK.-N....I-V-------------------------QG--R--I--G-------I--------E------I--YR-----N-----ST----------T---A.....--G  159
F1.CY.08.CY222               ------------------N-------Y--HI-K--KN------F--R--K--------I.-E.-I....I-V----------E------A-----QG-----I--G-------I--------E------I---KI----N------.-----------T---P.....---  158
F1.ES.02.ES_X845_4           ------------------N-------Y--H--K--KK---------R------------.D-.-K....V---------------------------QG--R--I--GT------IH------------R-I--YK-----N------.-----------T---A.....T--  158
F1.RO.96.BCI_R07             -*----------------N-------Y--HI-K--KK-S----F--R---V--------.EE.-K....--V-----------E-------------QG-----I--G-------I------------------I--YR-----N------.-----------T---A.....---  157
F1.RU.08.D88_845             ------------------N-------Y-----K--K-------F--R------------.EE.-K....I-V-------P-----------------LG--R--I--SM------IH-------E----------RI----N------.-----------T--VA.....---  158
F2.CM.02.02CM_0016BBY        ---------------------N----------RR-Q----------R---V--------.ENDSK....---V---------K-----H-------Q---R-----G-------K--------E------I---R-----N------.-----------T--V-.....---  159
F2.CM.10.DEMF210CM001        -X-------------------N----------R--A--------Q-R---V--------.A-DST....---V---------------------------R---R-----G--------H-------E-T--N-I---R-----N------.-----------T---N.....--R  159
F2.CM.10.DEMF210CM007        ---------------L-----N----------K--K----------R---V--------.K-DSK....---V--------------------H-------Q-X-R-----G--------H-------E----N-----YR-----N------.-----------T-----.....---  159
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.....-TR  158
G.CM.10.DEMG10CM008          ---------------------N----------K--K----------R--KV--------.--.--....IVR-----N---KA----H------SLR-----I--D-------SH---------------I--Q--R-------------.-----------K--V-.....-TR  158
G.CN.08.GX_2084_08           ---------------------N----------K--K----------R---V--------.--.-K....IVR-----------K--Q--H-------QR--H-----D--------H---------------I--QL-R-K------------.-----------KV-V-.....-T-  158
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-  158
G.ES.09.X2634_2              ---------------------N----------K--K----------R--KV------L-.--.-K....--VR----------------H-------QG-----I--D--N------Q--E--N-------I--R-----------P---.S----------K--V-.....-TR  158
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.....-AR  158
G.KE.09.DEMG09KE001          ---------------------N----------K--K-------F--R--KV--------.-E.-T....---K-----N---K--Q--H-------QR--I--I--D--------H--N----------II-E--R-K---------.-----------K--VR.....TP-  158
G.NG.09.09NG_SC62            ------------------K--N-------HI-K---K-S-------K---V--------.E-.-K....--VR---------K--Q--H-------QR-----I--D--------H-----------------QV-R------------.-----------K--VQ.....-T-  158
G.ZA.01.TV546                ------A--------------N--------X-RX-KD------X--R---V--------.-E.--....I-VR------K--KE-Q--H-------LR-----I--D--------H-------E------I--Q------D------.-----------K--VK.....-VR  158
H.BE.93.VI991                ------------------K--N----------K--KK-V-------TN-KT-------V.--.--....-------------------H-------QE-----I--D--------H---------------I---R---I-D------.R----------T---S.....--R  158
H.CF.90.056                  ------------------N-------Y--HI-R----------F--T------------.-E.--....---------N----E-------------L-------E-G-------MH-------E------I--RV-R---N-P---.KQ--T------T--VA.....---  158
H.GB.00.00GBAC4001           ------------------N-------Y--HI-K--KD------F--T--K---------.-E.--....---I-----N------------------QR--------D-----------N---E------I---R-Y-S----T--.-----------T--V-.....---  158
J.CM.04.04CMU11421           ------------------N-------Y-XHX-K---R-L-----D-N--K---V-Y--X.-E.-E....X-------Q------------------L-X-R-----G-------MH-------E------X---X-----D-P---.-Q---------T---K.....-XX  158
J.SE.93.SE9280_7887          ------------------N-------Y--N--K---K-L-----D-N--K---------.-E.-I....---------Q------------------QR--R-----G-------MC----------------I--Q------D------.-----------T---R.....--R  158
J.SE.94.SE9173_7022          ------------------N-------Y--N--K---Q-L-----D-R--K---------.-E.--....--V------Q------------------R---R-----G-------MH----------------I--Q------D------.-----------T---K.....--R  158
K.CD.97.97ZR_EQTB11          ------------------N-------Y-----K--NR-R-----D-N--K---------.--.-E....--V-----------E-------------L-K-R-----G-------I--------E-----------R------------.TQ---------T---A.....---  158
K.CM.96.96CM_MP535           ------------------N-------Y--HI-K--NR-Y-------R--K---------.--.-E....--V------L------------------L---R---E-D-------I--------E--V---I---R-----C------.-----------T--VA.....-RR  158
U.CA.01.TV749                ------------I--------N----------K-------------R---V--------.-E.--....--V-----------E-------------LG--R--I--G-------TH-------E------I--Q-------------.-----------K--VX.....--G  158
U.CA.99.TV721                -------------------A-N----------K-----------Q-R--KV---I----.-E.--....IV------------E----S------ELR--R--I--G-------TH-------E------I--Q-------------.-----------K--VA.....--G  158
U.CD.83.83CD003_Z3           ------------------K--N--------I-K--K--V-K-----TN-----------.-N.--....I-V-A----------------------QG-----I--D--------H---------------I---------------T--.-----------TT-VA.....-T-  158
U.CD.90.90CD121E12           ---------------------N----------K--NK-V-K-----IN-----------.EN.--....I--A---------------------QG-----I--D-------MH----------------I---------D--T--.-----------T--VP.....-I-  158
U.CY.05.CY090                ----------------K-K--N--------I-K---N-R-K-----RN--V--K-----.KE.-K....-------------------------QG--I-----G-------MH-----T-----------LR-----D-PE--.I---------IT---A.....-E-  158
U.CY.08.CY223                ---------------------N--------I-K--K----K-----MN--V--------.-E.-K....IIVR----------K--Q--H-------Q-S----I--D---K----Q-------E------I--EV-R---AFS---.-----------K--V-.....-RR  158
U.ES.10.DEURF10DZ001         ------------------N-------Y--HI-K--KK-M-----D-R--K--------I.-E.-V....---------N----A-------------L---R--I--S-------VH----------------I--D--Y------.S-------------VK.....T--  158
U.GR.99.99GR303              ------------------N-------Y--H--K--KR-V-----D-N------------.RE.-K....--V-----------E-------------M-S-R-----S-------MH-------E------I--Q------D------.-----------T---N.....--R  158
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  158
01_AE.AF.07.569M             ---------------------N------K-I-K--KK-L-------QN-KV----Q---.-E.G-....I-R-----Q---K--Q--H-------LRK-N--I--D--------H---------------I--QV-RH--D-PS--.-----------K---A.....---  158
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-.....---  158
01_AE.CN.10.YNFL03           ---------------------N----------K--K----------R--KV--------.-E.-K....--I----PQ---K-----H-------Q------I--D----R----Q-------E----R-I--QV-RH----PS--.----P------K--T-.....--R  158
01_AE.HK.04.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.....--G  158
01_AE.IR.10.10IR.THR48F      ---------------------N----------K--KK---------W--KV-------I.-E.-K....--R-----Q---K-----H-------QGK----I--D----R----Q-----ET----G-I--KV-R---X-PS--.-----------K--T-.....---  158
01_AE.JP.x.JRC77AE           ---------------------N----------K--KK-V-------Q--KV--------.-E.--....--R-----Q---K-----H-------QGK-N--I--D----K----Q---------------I--QV-RH----PS--.-----------K--T-.....---  158
01_AE.TH.09.AA090a_WG11      ---------------------N----------K--KK-C----------KV--------.-E.-K....I--R-----Q---K--Q--H-------QGQ-N-----D----R----H--------G----S-I--QV-RH--D-PS--.-----------K--T-.....--R  158
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  158
01_AE.US.05.306163_FL        ------------------K--N-I----I---K--KK-S-------Q--KV----Q-I-.-E.--....--R-----Q---K--Q--H-------QRE-N--I--D----R----Q-----A-------I--QV-RR----PS--.-----------K--T-.....--R  158
01_AE.VN.98.98VNND15         ---------------------N--------I-K--KK-C-------Q--KV--------.-E.-I....--R-----Q---K--Q--H-------QRE-N--I--D----R----Q--E--------R-I--QV-RR----PS--.-----------K--T-.....--R  158
```

```
                               Vif start
B.FR.83.HXB2                   MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHYESPHPRISSEVHIPL.GD.AR....LVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPELADQLIHLYYFDCFSDSAIRKALLGHIVSPRCEYQAGH.NKVGSLQYLALAALIT.....PKK  158
02_AG.CM.10.DE00210CM013       ---------------L--K--H--------I-R--KD-S-------R---V--------.--.-K....-IVR-----------Q--H------KQES----I--D-------MH------G------I--QV-----------.-----------K--VP.....-V-  158
02_AG.CY.09.CY256              ---------------R-----H-I--Y--HI-R--KD-C----F--R---V--------.--.--....--VK----------------H------EQR-----I--D-------H----------------QV-R----------.K----------K--V-......-I-  158
02_AG.ES.06.P1423              -------------------N----------K--GK---------R--K---------.ES.-T...---------------------H------M------I--D-------HH-N---E------I--QV-R----------.-Q---------K--V-......-TR  158
02_AG.GW.05.CC_0048            -----------------A-N--------R-K-TKD--------TR--KV--------.--.--....--VK-----N-----------H------KQ-----I--D-------H-----------Q-I--QV-R-S--------.-----------K--V-......-TR  158
02_AG.KR.12.12MHR9             -------------------N-I------I-K--KK-V-------N--KV--------.--.-K....---R-----------------H------KQ-----I--D-------H--------------I--QV-R-I--------.S----------KV-V-......-TR  158
02_AG.LR.x.POC44951            -------------------N----------R--KD---------R--KV--------.--.-T....---R-----------------H-------Q-----I--D-------H--K--T-------I---V---I----E--.-----------K--V-......-TR  158
02_AG.NG.09.09NG_SC61          -------------------N--------I-K-TQ----------R--KA--------.--.--....-IVR-----------------H------EQ-----I--D-------H-----T----------QE-R----------.R----------K--V-......-TR  158
02_AG.NG.x.IBNG                -------------------N----------K--K-------F--R--KVC-------.--.--....--VR-----------------H------KQ-----I--D-------------N----------I--E--R----------.-----------N--VA.....-T-  158
02_AG.SE.94.SE7812             -------------------N----Y---K-R--KD---------S---V--------.--.--....--VR-----------------H------KQR----I--D-------H--------------I--Q--R----------.T----------K--V-......-TR  158
02_AG.SN.98.98SE_MP1211        -------------------N----------K--HR---------R--KV--------.--.-T....-IVR-----QP----------H----V-QQ-----I--D-------H------E----------QV-R-K--------.-----------K--V-......-TR  158
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-.....---  158
04_cpx.CY.94.94CY032_3         -A--------------K----N----------K--K-*--------R--KV--------.-E.--.....-VR-----QP--Q-----H------LR----------D-------MH------E------I---R-------------.---------------S......---  157
05_DF.BE.x.VI1310              -----------------N-------Y--H--K--NR-C----F--RN----------.--.-K....--V---------------------------R---------S-------V-------E----N-I--R-------F---.-----------T----------.---  158
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-  158
07_BC.CN.98.98CN009            -------L--------K----N-----------RR-N---------D-R--KV-------V.-E.-K....---K-----Q---------H--------LR--T--IE-G---------------A-------I-----I---D-----N-----------T---K......---  159
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  159
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-  158
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------.KQ----------T------------  158
11_cpx.CM.95.95CM_1816         ----------------------N--------I-K-TNR-R--------R---V--------.-P.-T...--VI----------E------------L---R------S-------M------TE------I---R-R-----P---.-----------K--V-......-T-  158
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..........  158
13_cpx.CM.96.96CM_1849         -----------------N-------Y--H--K--GR------D-N--K---------.--.-K....--V-----------K-------------L---R-----G--------TH-------------I---R-----D-----.-----------T---H..........  158
14_BG.ES.05.X1870              ---------------------N-I-------I-K--K--S-K------R--KV--------.--.-K...X-VR---------K-----H------QR-----I--DX--X---QH------E------I--G-----K---P---.S----------K-VV-........Q  158
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  158
16_A2D.KR.97.97KR004           ---------------------N----------K--KD-C-------IN--V--------.-G.-K...IIVK------P--K-------------QE-----I--D--GH----H--------------I--K--R-----------.-----------K--VG.....ETR  158
17_BF.AR.99.ARMA038            -----------------N-------Y-IH--KQTKR-V----F--RN--V--------.E-.-K....---A------P---------------QG--R--I--G------I-------E------I--NKI----N------.-----------T---A..........  158
18_cpx.CU.99.CU76              ---------------------N-I---------K--KN---------R--KV--------.--.-T....--V----------KE----H-------LR--R-----D----------------E----N-I---R-----------.----------------.........  158
19_cpx.CU.99.CU7               ---------------------N--------I-K--NK-V-------RN-KV--------.--.-K....--VK----------K-----H-------QR-----I--D--------H------------------QV-R----------.-----------K--V-......-TR  158
20_BG.CU.99.Cu103              ---------------------H----------K--K--W----F--R--KV--------.--.-K....-IVR----------K-----H------KQRE----I--D--HM------N-----------I--E--R----FH---.-----------K--V-......-T-  158
21_A2D.KE.99.KER2003           ---------------------N--------I-K---R---------R---V--------.-E.-T....-IVR------S--K-----H-------Q------I--D---H----H--------T------I--EL------------.-----------KV-VA.....A-R  158
22_01A1.CM.01.01CM_0001BBY     ---------------------N----------K---------------R--K---------.--.--....--VR----------KE----H-------Q---------D--------M--------------I--Q-----S----T--.-----------R--VA..........  158
23_BG.CU.03.CB118              ---------------------H------K-I-K--K--C-K-----TN-----------.-E.-K....-I-K------------------------------G-------Q-------E-----S-I---R---S----------.-----------K--V-......-T-  158
24_BG.ES.08.X2456_2            ---K-----------------N--------I-K--KD-W----F--R--KV--------.--.-N....-IVR----------K-----H------KQRK----I--D---R---QH--K---------------D-F------P----.-----------K--V-......-T-  158
25_cpx.CM.02.1918LE            -----------------N-------Y--H--K--KK---------H------------.--.-E....--V-----------E----------------L---R-----G---R---MH------E------I------K-D-----.K----------R--VA..........  158
26_AU.CD.02.02CD_MBTB047       ---------------------H-----Y---R-K--K--L-------R--KV--------.--.--.....---R-----Q----E----H------LR-----------G--------Q--E-------------IV-Q-IIHK--------.K----------K--VA.....T-R  158
27_cpx.FR.04.04CD_FR_KZS       ------G-----------NA-------Y--HK-K--K--H-----D-H--K---------.-E.....--V-----Q-------------------L---R-----G-------IH------E------I--YR-----D-P---.-Q----------T-------------  158
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..........  158
29_BF.BR.01.BREPM16704         -----------------N-------Y--HI-K--KR------F--R---V--------.EE.DEL...V-------------------------------GK-R--I--G-------IN------E------I--QKI-------------.----------------.......-R  159
31_BC.BR.04.04BR142            -------L--------K----N--------I-KR-S--S-K-----RN-K---------.-E.--.....---K-----Q---------H---L---LRG-------G---R---MH-----AE------I---R-----D------.-Q----------T---K..........  158
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..........  158
33_01B.ID.07.JKT189_C          ----------G-----K-K--H------V-I-K--K----------Q--KV--------.-E.-K...I--R-----Q---K-----H------QRD----I--D-------N----------------I--QV-RH----PS--.-----------K--T-.....Q-R  158
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  158
35_AD.AF.07.169H               -----------------K--N--------I-K--K--V-------RN--V----Q---.--.--....I-VR----------K-----H------LQ-----I--DV-------H--N---------------I--E-----S-------.-----------K-----.....-T-  158
36_cpx.CM.00.00CMNYU830        ---------------------N-------HI-KT-KD---------R--KV--------.--.--....---R-----------------H------KQRK----I--D--------H--------------I--Q-----D-P---.-----------K--V-......-T-  158
37_cpx.CM.00.00CMNYU926        ---------------------N-------HI-K--K-----------RD--V--------.-E.-W....--VR-----QP----------H------LRK-R--I--D--------H--------------I--Q-----D-P---.-----------K--V-......-T-  158
38_BF1.UY.03.UY03_3389         -----------------N-------Y--H--K--KR-C-K--F--RN--V--------.QE.DK....-------------------------------QGS-R--I--D-------I-------E------N-I---RI----N------.-----------T---A.......-Q-  158
39_BF.BR.04.04BRRJ179          -----------------N-------Y--H--K--KE---------T-------------.-E.--.....--V------------------------------R--R-----D---------------AE----H-I--R--R---N------.-----------T---A..........  158
40_BF.BR.05.05BRRJ055          ----------------------------I-R--NK-------K-T--K---------.-E.-A....--V-----------E---------K---RK-----N-D-----------N----E---N----R-----S-N------.---K-------T---K..........  158
42_BF.LU.06.luBF_18_06         ---------------------N------I---K--K--H-------I--K----------.--.S-......--K--------------------------R---------D-----------------E-----NTI---R-----------.-Q---------T-------------  158
43_02G.SA.03.J11223            ---------------------N----------K--K----------RN-KV-------X.X-.-T....-XXX-----Q-------Q--H------XQ-K-------D--------H------E------I--Q--R-----------.-----------K--VK.....-TR  158
44_BF.CL.00.CH80               -----------------N-------Y--H--K--KR-L----F--R---V--------.E-.-X...--------------------A-----QG--R--I--G-------IH-----T----------YK-----N-P---.-----------T--VA..........  158
45_cpx.FR.04.04FR_AUK          ---------------------H----------K--KR--------TR--KVAA------.--.-T....--VR-----Q---K-----H-------L---------D---------H-------------I--QV-R-------T--.-----------K--V-......-TR  158
46_BF.BR.07.07BR_FPS625        -----------------N-------Y-----K--KR--------TR---V--------.DK.DT...------------------------------QG--R-----D--------I--------------I--QK-----N------.-----------T--VA..........  158
47_BF.ES.08.P1942              -----------------SK------Y-IHT-K--KK------F--R--K---K-----.EEENK....--V------N----------------------QGK-R--I--G-------I-------E------I---R-----N------.-----------T---A.....T--  159
48_01B.MY.07.07MYKT021         ---------------------N----------K--KK-C-------QN-KV--------.EG.-K.......---K-----Q---K--Q--H------QRD----I--D--------HH-------------I--QV-RR----PT--.-----------K--T-.....Q-R  158
49_cpx.GM.03.N26677            -------L--------------N-------HI-R--S----------R--KV--------.-E.--.....--VR-----Q----E----H-------LR-----------G-------MH-----T----------I------R---V-P---.-----------T---K.....T-R  158
50_A1D.GB.10.12792             -------V--------------H-I----I-I-K--KN-V----F--RN-KT----------.--.-K....--VK----------K-----H-------LR-----I---Q--------H-------E------I--QV---S--------.-Q----------K--V-......-TR  158
51_01B.SG.11.11SG_HM021        ---------------------N----------K--KK---------Q--KV--------.-K.--.......---R-----Q---Q--Q--H------QRD-N--I--D--------H--E-----------I--Q--IH----PS--.-----------KT-T-.......-R  158
52_01B.MY.03.03MYKL018_1       ---------------------H-V--------K--KK-V-------Q--KV--------.-E.-K....--VR-----Q---K-----H-------QGL----I--D--------Q-----A-------I--QV-RH----PS--.-----------K--T-.......-R  158
53_01B.MY.11.11FIR164          ---K-----------------N------I---K--K.-V-------QN-KV--------.--.-K...I--R-----Q---Q-----H-------QGN-N--I--D--------HH----------NTI--QRIRS----PS--.-R----------T--T-.....Q-R  157
54_01B.MY.09.09MYSB023         -----------------K----------I-I-K--K--V-K---D-M--K---------.-G.-K....-------------------------D--RES---------D-------------------E----N-I--RV---S-A------.-----------K--T-..........  158
55_01B.CN.10.HNCS102056        ---------------------N--------I-K--K----------Q--KV--------I.-E.--.....---R-----Q---K--Q--H------QRT-R--I--D---R-----H--N------------R-I--QV-RH----PS--.-----------K--T-..........  158
56_cpx.FR.10.URF5_patient_A    ----------------K----N--------I-K-TK----------R--KV--------.--.--....-IVR-----Q----------H-------Q------I--D--------H-------E------I--QV-R-S--------.S----------K--V-......-T-  158
57_BC.CN.09.09YNLX19sg         ----------------K----N-----------RR-N----K---D-R--KV--------V.-E.-T....-I-K-----Q----------H---V---LR--T--IE-D---H--------------A-------I-----I---D-----K-----H----T---K..........  159
58_01B.MY.09.09MYPR37          ---------------------H-------HI-K--KK-S-------Q--KV--------.--.-K...I--R-----Q---K-----H------QGN-N--I--D--------N--------------I--QV-RR----PS--.-----------K--T-..........  158
59_01B.CN.09.09LNA423          ---------------------N----------K--KN---------Q--K--------I.-E.-K...I--K-----Q---K-----H---V---LR--N--I--D---R----Q--------------I--QV-RR----PS--.-----------K--T-.......-G  158
60_BC.IT.11.BAV499             -------L--------K----N-----------KR-S--Y-------R------------.EE.T-.........-S-Q----------H-------LR--R------D--------VH-----A-------I---R-----D------.-----------T---K..........  158
61_BC.CN.10.JL100010           -------L--------K----N-----------KR-N---------D-R--KV--------.-E.--.....---K-----Q---------H---V---LRK-----IE-D--------------A-------I-----I---D-----N-----------T---K..........  159
62_BC.CN.10.YNFL13             -------L--------K-K--N-----------RR-N-----------RN-KV-------V.-E.--.....---K-----Q---------H-------M------IE-GM-------H-----A-------I-----I---D-----..----------T---R..........  157
63_02A1.RU.10.10RU6637         ---------------------N------I---K---E---------R--KV--------.--.--....--VR-----QK----------H-------RE--------D---------H--E---ET-----IV-E---T--N-S---.-----------K--V-......-TR  158
64_BC.CN.09.YNFL31             -------X--------K----N-----------KR-K-----------R--K---------.-E.--.....---K-----QP----------H------WR--N--IE-G-------M-------A-------I-----I---D-----T-----------T---K..........  159
65_cpx.CN.10.YNFL01            ----------------K----N--------I-RR-K-----------R--K---------.--.--....--MK-----Q----E----H------MRG-N--IE-D-------------------A-----R-I--Q--L---N------.-----------T---K..........  158
67_01B.CN.11.2011.ANHUI.HF115  ---------------------H-I------I-K---K---K-----Q--KV--------.-E.-I.....---R-----Q---K-----H-------Q-N-R--I--G--------H--------------I--QV-RR----PS--.-----------K--T-........Q  158
68_01B.CN.11.2011.ANHUI.WH73   ---------------------H----------K---K---K-----Q--KV--------.-E.-M....--VR-----Q---K--Q--H--M---Q-N----I--D--------H--------------I--QV-RR----PS--.-----------K--T-........Q  158
70_BF1.BR.10.10BR_PE004        -----------------N-------Y--HI-K--KN------F--R--KV--------.-K.-E.......-I-----N-----------------------QG--R--I--G-------I-------E----N-I--Q-------------.S----------TXXRX.......-X  158
71_BF1.BR.10.10BR_PE008        -----------------N-------Y--H--K--KR------F--R---VG-------.EE.-T......-V--------------------------------RG--R--I--GT------I-------E-----------R---------------.---------------TP..........  158
72_BF1.BR.10.10BR_MG002        -----------------N-------Y--HI-K--KR---------T-------------.--.-E......-V---K-------------------------R--H-----N--------H------E----S-I--R--R-----------.----------------A.......-R  158
O.BE.87.ANT70                  -------L-------QKVKA-N----Y-K-R-R-TEN-W-------RN--V--S-Y--V.-V.-H...V-V------MP----E---H------Y-K-K--I---T--RM---H--T--TA--V---I--QR-LTK---PT--.SQ--T--L---R-VVK.....ARS  158
O.CM.98.98CMA104               -------L-------QKVKV-N----Y-K-R-R--KD-Y-----X-RN--V--G-Y--V.-Q.-C...I-V------MP---EDQ--H------QY-K-R--I---T--RM------T--TE------T--QR-LTK-D-PE--.S---T--L---RVVVK.....T-R  158
O.CM.98.98CMABB141             ---K---L-------QKVKA-N----Y-K-R-R--KD-Y-------RN-----S-Y---.-P.-C...I-VS-----MP---EE-M-H------QY-K-T--I---T--RM------T--TE----ETI--QR-LTK---S---.SQ--T--L---R-VVK.....E-R  158
O.CM.98.98CMABB212             ---------------QKV---T------K-I-K--KN-H-K---FRS---G-G-Y--V.AE.-K...IIV------MP---EE---H------QY-N-T--I---T--RM---H-----T--------I--QR-LTK---P---.SQ--T--F---K-IEK.....V-R  158
O.CM.98.98CMU5337              -X-----L-------QKVKA-N----Y-K-RTR---D-S-K--F-XRN--V--S-Y--V.-M.--......VKV------MP----E-M-H------QY-X-T--I---T--RI------T--TE------I--QR-LTK---P---.SQ--T--L---RVVVK.....A-R  158
O.CM.99.99CMU4122              -------L-------QKVK--N----Y-KHR-R-TKE-H-K--F-FRN--A--G-Y--V.-M.-C...V-V------MP----E---H------QY-N-T--I---T--RI--QH--T--TE------II-QRILTK---S---.SQ--T--L---RVVVK.....E-R  158
O.FR.92.VAU                    -------L-------QKVKA-N----Y-K-R-R--TD-Y-K-----IN--V--R-Y--V.-P.-C...IMV------MP---EEQ--H------QCRK-I--I---T--K-----H--T--TE----R-T--QR-L-Q---PE--.SQ--T--L---R-VVK.....A-R  158
O.GA.11.11Gab6352              ------AV-------QKVK--N----Y-K-R-R--KE-Y----F-HRN-----S-Y--V.-E.-W...V-V------MP----E---H------QY-K-K--I---T--RM--I---T--TE------I--QR-LVK--FP---.SQ--T--L---RV-EK.....A-R  158
O.SN.99.99SE_MP1299            -------L-------Q-VK--N----Y-K-R---T-D-Y----F--RN--V--G-Y--V.-G.PW...I-V------MP----E---H---------Y-K-K--I---T--RM--I---A--TE------I--QR-LT----P---.SQ--T--L---RVVVK.....E-R  158
O.US.10.LTNP                   -------L-------QKVKA-N----Y-K-R-R-TKD-H------TRN-----S-Y--V.-M.-C...V-V------MP---EDQ--H------QY-K-K--I---T--RI--Q---T---E----R-I--QR-LTK---P---.SQ--T--L---R-VVK.....A-R  158
N.CM.02.DJ00131                --------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-WVG.....V--  158
N.CM.02.SJGddd                 -A------V-------K-KK-N----Y------K--Q--Y------TR--K--------V.-Q.--....----V------T---QS----H-------LRK-K------M------H---N--AA----Q-I--RP-V-K---P---.QQ--T-------R-WVG.....V--  158
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-WVG.....I--  158
N.CM.06.U14296                 --------V-------K----T-----------KR-K--Y------IN---K-------V.-Q.-K....--VI-----T---QP----H-------LGK-K------M--R---------N--TA----Q-A--RP-L-K---P---.KQ--T-------T-WVG.....VQ-  158
N.CM.06.U14842                 --------V-------K----N-----------K--T--H------TX--KK-------V.-Q.-K.....---I------T---QP----H-------LGK-K------M--K---------N--TA----Q-V--KP-L------P---.QQ--T-------T-WVG.....V--  158
N.FR.11.N1_FR_2011             -A------V----------K-N-----------K--K--Y------TN--K-G-----------.-P.-------....-I------T---QP----H-------LRK-K------M--K---------N--TA----Q-V--RQ-I-K---P---.KQ--T---I--T-WVG.....A--  158
P.CM.06.U14788                 ------------IP-EKVKA-N----Y-K-R-K-T-K--------TT------A-Y--V.-T.-T....IIV------MP---EEQ--H-A-V---QGK-T--I---T--R----H--Q------V-R-I--DR-LNX---S----.SQ----------KVVVG.....KV-  158
P.FR.09.RBF168                 ------------IS-EKVKA-N----Y-K-R-K-TKN--------TV------G-Y--V.-A.-T....I-V------MP---EEQ--H-A-V---QGK-T--I---T--R----H--Q------V-R-I--DR-LNK---S----.SQ-----F---KVVVG.....KV-  158
CPZ.CD.06.BF1167               -------RV---I-------A-N------I-K-KHCKE-K-----DID---KGG-----ITEK.-K.........-I-----C--KA----H-A-----QG--W-E-----T---I--SR--P--TER---Q-I--EKILQH-HFPT--.S-------F--FKKVLE.QNK.Q-P  162
CPZ.CM.05.SIVcpzMT145          -------Q----------VK--N---Y---R-K--KN-------DGTN--MA-QIY--I.NQ.-K....-IV------MP---------H--A----ERG---------T--R----H-----TAT---N-I-EKA-L-K--FPE--.S-------W--KQ-VK.....T--  158
CPZ.GA.88.GAB1                 -----------------K--N----Y-I-R-K-----------DH-N-KVA--I---FRDY.SK.....-IV----A-SP---A---H----Q--LGS-V-----FT--R---SQ-----AET---R-I--QL-A------KE--.RQ------F---K---S.....ERR  159
CPZ.TZ.06.SIVcpzTAN13          -------QV--MI----L---N------IFTTKYCKE-K------TDV-KRAG-I----E-K.-------...I-VLH---AC---P---H--G---QG-----I---T--------TK--A--TAG-V-Q-I--ERILTF-H--T--.RQ--T------FLKVVEAQYKTA--  164
CPZ.US.85.US_Marilyn           -------L----------VKA-N----Y---R-K--K-----------TN--VN------I.DK.-K...I--I-----K---QA----H------QGT-K--I---M--K----------TA-------V--RP-I-K---PT--.-------F---K--VG.....QS-  158
GOR.CM.07.SIVgor2139_287       ------------IP-E-VKA-N----Y-K-RTKQ-KQ---------AVN--V--G-Y--V.-P.-E....IIV------MP--KEEQ--H--G----QGK-I--L---T--RI------Q--TE--V---IV-EKLLYK-N-P---.SQ--T-------R--VG.....KVR  158
GOR.CM.07.SIVgorCP2135con      ------------IP-E-VKA-N----Y-K-RTKQ-KQ---------AVN--V--G-Y--V.-P.-E....IIV------MP--KEEQ--H--G----QGK-I--L---T--RI------Q--TE--V---IV-EKLLCK-N-P---.SQ--T-------R--VG.....KVR  158
```

```
                                                Vif end
B.FR.83.HXB2                IKPPLPSVTKLTEDRWNKPQKTKGHRGSHTMNGH...*  192
A1.AU.03.PS1044_Day0        -------R--A-----------P----------...-  193
A1.CH.03.HIV_CH_BID_V3538   T------IR------------------------C...-  193
A1.CY.08.CY236              -------K-------------D----------...-  193
A1.ES.05.X1608_8            -------R--------------L---R-------...-  193
A1.KE.11.DEMA11KE001        -------RI----------R---P-E--I---C...-  193
A1.RU.11.11RU6950           ER------R------------R-R--N----C...-  193
A1.RW.11.DEMA111RW002       TN------R--A--------R----P--N-I--C...-  193
A1.SN.01.DDI579             TR------R-----------R---------C...-  193
A1.UG.11.DEMA110UG001       T-------R------------R-PKE----------  193
A1.ZA.04.04ZASK162B1        P-------K----------------------C...-  193
A2.CM.01.01CM_1445MV        T-------R--V----------R----N-----C...-  193
A2.CY.94.94CY017_41         T-------R--V--------------------C...-  193
B.BR.10.10BR_MG029          R-------A----------R--T-------------...-  193
B.CA.07.502_1191_03         T-------R--------------------I------...-  193
B.CH.08.M2_0803101_NFLG8    -------K----------------....V-------...-  189
B.CN.10.DEMB10CN002         R-------R-----------------N--TS-----...-  193
B.ES.10.DEMB10ES002         -------R--------H---R--------------...-  193
B.FR.11.DEMB11FR001         -------K--------V--------------Q...-  193
B.GB.05.MM45d213_GN1        ------------------R---------------...-  193
B.HK.06.HK003               --------------------T--------H-----...-  193
B.HT.05.05HT_129389         -------Q------------I----E-P--------...-  193
B.JP.12.DEMB12JP001         R--------------------------------...-  194
B.KR.07.HP_18_07JHS10_3909  -------R----------------E----------...-  193
B.PE.07.502_0525_wg5        -------A----------R----------------...-  197
B.RU.11.11RU21n             ------------------R--D------IS-----...-  193
B.TH.08.MERLBDTRC10         -------R----------R-------Q-T------...-  193
B.US.11.ES38                V-----------------------E----------...-  193
C.AR.01.ARG4006             R-------R--V----------RDR--N-I-------...-  193
C.BR.07.DEMC07BR003         R-------K--V-------*--RD---N---------...-  192
C.BW.00.00BW5031_1          P-------R--A-------R--R-R--N---------...-  193
C.CN.10.YNFL19              -------IQR-V-----N---M-R--N---------...-  194
C.CY.09.CY260               T-------R--V--------------N---------...-  193
C.ES.08.X2363_2             R-------R--V------S--ARDC--N-S-------...-  193
C.ET.02.02ET_288            R-------S--V----------R-R--N--V------...-  193
C.IN.09.T125_2139           -------IK--V---------I--R--N-IL------...-  193
C.KE.00.KER2010             -------Q--V-----N----R-R--N---------...-  193
C.MW.09.703010256_CH256.w96 -------S--V--K-------R-R--N---------...-  193
C.TZ.08.707010457_CH457.w8  -------R--V-----N----R-R--N-I-------...-  193
C.YE.02.02YE511             -------R--V----------M-R--N-------...-  193
C.ZA.08.705010534_CH534.w12 T-------S--V----------R-R--N---------...-  193
C.ZA.10.DEMC10ZA001         R-------R-----------EI-----N---------...-  193
C.ZM.11.DEMC11ZM006         R-------R--V--------------N-----L...-  193
D.CM.10.DEMD10CM009         R-------R----------Q---------T------...-  193
D.CY.06.CY163               R-------R----------Q----K-----S-----...-  193
D.KE.11.DEMD11KE003         -------K----------R-----E----------...-  193
D.KR.04.04KBH8              -------R----------E---R---P-----WSF-  196
D.SN.90.SE365               -------R-----------------E----------...-  193
D.TZ.01.A280                -------K----------R-------P----------...-  193
D.UG.10.DEMD10UG004         T-------R----------R----------------...-  193
D.UG.11.DEMD11UG003         -------K----------R-R--------S------...-  193
D.YE.02.02YE516             R-------K----------Q----------------...-  193
D.ZA.90.R1                  T-------R------------R-------------...-  194
F1.AO.06.AO_06_ANG32        V-------R--V--------E--------------...-  193
F1.AR.02.ARE933             T-------Q--V----------R-------------...-  193
F1.BR.10.10BR_RJ015         T-------Q--V--------------#---D--WNFW  196
F1.CY.08.CY222              T------IK--V--------E-R----N---------...-  193
F1.ES.02.ES_X845_4          T-------R--V--------E-R---E----------...-  193
F1.RO.96.BCI_R07            T-------K--V-------REIR-------------...-  192
F1.RU.08.D88_845            T-------K--V--------E---------------...-  193
F2.CM.02.02CM_0016BBY       V-------K--V--K--N------------------...-  194
F2.CM.10.DEMF210CM001       TR------S--V-----N--R--------------...-  194
F2.CM.10.DEMF210CM007       --------R--V-----N----R-------------...-  194
G.BE.96.DRCBL               RR------R--------E----R---ENP--------...-  193
G.CM.10.DEMG10CM008         R-------Q---------S--IRD-KE-P--S-----...-  193
G.CN.08.GX_2084_08          -------K------------------NPS-S-----...-  193
G.CU.99.Cu74                -------K-----------R-R---ENP--------...-  193
G.ES.09.X2634_2             A-------R----------R-R----NP--T-----...-  193
G.GH.03.03GH175G            K-------R----N------------ENP-----WNC-  196
G.KE.09.DEMG09KE001         -------R---------S--IR-R-ENL--T-----...-  193
G.NG.09.09NG_SC62           -R------R------*----------ENP--S-Y...-  192
G.ZA.01.TV546               RR------K--------E----R---XNP-----WSC-  196
H.BE.93.VI991               T-------R--V----------R-------------...-  193
H.CF.90.056                 -------R--V----------R-------------...-  193
H.GB.00.00GBAC4001          T-------R------------R-------------...-  193
J.CM.04.04CMU11421          R-------Q--V-----N----R-------------...-  193
J.SE.93.SE9280_7887         R-------Q--V----------T---E----------...-  193
J.SE.94.SE9173_7022         R-------Q--V----------RD--E----------...-  193
K.CD.97.97ZR_EQTB11         T---V---Q--V----------R--------S-Q...-  193
K.CM.96.96CM_MP535          P---V---K--V----------R-----Q-------...-  193
U.CA.01.TV749               T-------Q--A-G---------A---N---------...-  193
U.CA.99.TV721               T-------Q--A----------RD------------...-  193
U.CD.83.83CD003_Z3          R-------R--V------------K-----H-----...-  193
U.CD.90.90CD121E12          R-------R--V---------I---K----------...-  193
U.CY.05.CY090               R-------Q--V-----R-R--R----N---------...-  193
U.CY.08.CY223               T-------R------------I----ENQ--------...-  193
U.ES.10.DEURF10DZ001        R-------Q--V----------RD-----S-------...-  193
U.GR.99.99GR303             T-------K--V-------------E-Q--------...-  193
U.NL.95.U_NL_95_H10986_D1   -------R--V----------R---E---I------...-  193
01_AE.AF.07.569M            T-------K------------.G--GENP--------...-  192
01_AE.CF.90.90CF11697       TR------R------------R---ENP--------...-  193
01_AE.CN.10.YNFL03          -R------K-----------IR---ENP--------...-  193
01_AE.HK.04.HK001           SR------K-------SE----R---ENP--------...-  193
01_AE.IR.10.10IR.THR48F     -------K------------I----ENL--------...-  193
01_AE.JP.x.JRC77AE          TR------KI----------IR---ENP--------...-  193
01_AE.TH.09.AA090a_WG11     -R------K------------IR---ENP----C...-  193
01_AE.TH.90.CM240           -R------.EI----------.R---ENP--------...-  191
01_AE.US.05.306163_FL       -R------K------------IR---ENS--------...-  193
01_AE.VN.98.98VNND15        -R------K------------IR---ENP--------...-  193
```

```
                                                Vif end
B.FR.83.HXB2                IKPPLPSVTKLTEDRWNKPQKTKGHRGSHTMNGH...*   192
02_AG.CM.10.DE00210CM013    T-------R-------KE-----------S--------   193
02_AG.CY.09.CY256           TR------K--A----------R-----RS--------   193
02_AG.ES.06.P1423           TR------K------------R------S--------   193
02_AG.GW.05.CC_0048         K-------K-------E----------RS--------   193
02_AG.KR.12.12MHR9          R-------K--A-----E----RD-----S--------   193
02_AG.LR.x.POC44951         T-------K--A-N------IR-----RS--------   193
02_AG.NG.09.09NG_SC61       R-------K---------S--R------S--------   193
02_AG.NG.x.IBNG             T-------R--A----KE----R-----RP--------   193
02_AG.SE.94.SE7812          T-------K--A-----E----R-----RS--------   193
02_AG.SN.98.98SE_MP1211     R-------K--A-----E----R----NRS--------   193
03_AB.RU.97.KAL153_2        -------------------R--D-------S-----.   193
04_cpx.CY.94.94CY032_3      T-------K--V---------R-R-ENQI--------   192
05_DF.BE.x.VI1310           T-------R--------------R--N-----Y...-   193
06_cpx.AU.96.BFP90          R-------Q--V---------R---E---------   193
07_BC.CN.98.98CN009         -------IK--V-----N---IR-R--N--------   194
08_BC.CN.97.97CNGX_6F       -------IK--V-----N---IR-R--N--------   194
09_cpx.GH.96.96GH2911       T-------K--V--------R-R---ENQ-------   193
10_CD.TZ.96.96TZ_BF061      T-------R----------R---------------   193
11_cpx.CM.95.95CM_1816      V-------K------------R---EN-----CWNC-   196
12_BF.AR.99.ARMA159         T-------K--V-----N----R-------------   193
13_cpx.CM.96.96CM_1849      -------K--V----------RD--ENQ--------   193
14_BG.ES.05.X1870           T-------K----------R-R----NP--------   193
15_01B.TH.99.99TH_MU2079    -R----R-K-------D----IR---ENP--D-----   193
16_A2D.KR.97.97KR004        T-------R----------------------------   193
17_BF.AR.99.ARMA038         T-------K--V--------E-R-------------   193
18_cpx.CU.99.CU76           T-------R--A----------R----NQ--------   193
19_cpx.CU.99.CU7            SR------R------------R---ENP--------   193
20_BG.CU.99.Cu103           --------R----------R-R---ENP--------   193
21_A2D.KE.99.KER2003        T-------R------------I----E--M-------   193
22_01A1.CM.01.01CM_0001BBY  --------R------------R---EYP-----WKC-   196
23_BG.CU.03.CB118           --------R----------R-R---DN---------   193
24_BG.ES.08.X2456_2         --------R----------R-R---EN---------   193
25_cpx.CM.02.1918LE         --------R--------E--------ENQ-------   193
26_AU.CD.02.02CD_MBTB047    T-------K----------------E----------   193
27_cpx.FR.04.04CD_FR_KZS    ---S----R------------R---ENQ----Q...   193
28_BF.BR.99.BREPM12609      T-------R----------I----N----------   193
29_BF.BR.01.BREPM16704      --------R----------R----------------   194
31_BC.BR.04.04BR142         --------R--V-----R----RDR--NQ-------   193
32_06A1.EE.01.EE0369        R-------Q--V-----E----R---E------R...   193
33_01B.ID.07.JKT189_C       -R------K------------IR---ENP--S-----   193
34_01B.TH.99.OUR2478P       -R------K----------QIR---ENP-T-------   193
35_AD.AF.07.169H            TR-----IR--A----------R-P-E------C...   193
36_cpx.CM.00.00CMNYU830     R-------K------------R-----RST------   193
37_cpx.CM.00.00CMNYU926     KR------K------------R---ENP-------   193
38_BF1.UY.03.UY03_3389      T-------K--V----------R-------------   193
39_BF.BR.04.04BRRJ179       -------LK-------------T-------------   193
40_BF.BR.05.05BRRJ055       -------LK----NK-----R---------------   193
42_BF.LU.06.luBF_18_06      --------K----------------------------   193
43_02G.SA.03.J11223         K-------A---------S---R-----RX-------   193
44_BF.CL.00.CH80            T-------Q--V---------E-----E---------   193
45_cpx.FR.04.04FR_AUK       T-------K--A----------RD--------C...   193
46_BF.BR.07.07BR_FPS625     T-------K--V--------E-R-------------   193
47_BF.ES.08.P1942           A-------Q--VX-------------E---------   194
48_01B.MY.07.07MYKT021      --------K------------IR---ENP--S-----   193
49_cpx.GM.03.N26677         -------TR-------------R-R--N---------   193
50_A1D.GB.10.12792          --------R--A-----E----R-P-----I------   193
51_01B.SG.11.11SG_HM021     -R------K------------I----ENPI-H-----   193
52_01B.MY.03.03MYKL018_1    -R------K------------I-D--EH---------   193
53_01B.MY.11.11FIR164       -R------K------------I--R-ENPI-------   192
54_01B.MY.09.09MYSB023      --------K------------IR---ENL--S-----   193
55_01B.CN.10.HNCS102056     -R------K-------------M---EN---------   193
56_cpx.FR.10.URF5_patient_A K-------A-----I-------R-----RP-------   193
57_BC_CN.09_09YNLX19sg      ------IKQ-V-----N----R-R--N---------   194
58_01B.MY.09.09MYPR37       -R------K------------IR---ENP--S-----   193
59_01B.CN.09.09LNA423       -R------K------------IM---E-P-------   193
60_BC.IT.11.BAV499          R-------R--V----------RDR--N---H-----   193
61_BC.CN.10.JL100010        -R------K--V---------IR-R--N---------   194
62_BC.CN.10.YNFL13          --------K--V-----N----R-R--N--X-------   192
63_02A1.RU.10.10RU6637      ER------K-------------------RS-------   193
64_BC.CN.09.YNFL31          R------IK--V-----D---IM-R--N---------   194
65_cpx.CN.10.YNFL01         -------IK--V-----N---I--R-..D--D-----   191
67_01B.CN.11.2011.ANHUI.HF115-R------K------------I----ENP-------   193
68_01B.CN.11.2011.ANHUI.WH73-R------K------------I----ENP-------   193
70_BF1.BR.10.10BR_PE004     -------XK------------I---K----------   193
71_BF1.BR.10.10BR_PE008     ---------Q--------------KE----------   193
72_BF1.BR.10.10BR_MG002     R----------------E-------K-N--------   193
O.BE.87.ANT70               R-------Q--------HLRIRDQLK-PS-------   193
O.CM.98.98CMA104            N-------Q--------RHL-IRDQLE--S-------   193
O.CM.98.98CMABB141          N-------Q--------YLRIRDQPE--S-------   193
O.CM.98.98CMABB212          N-------Q---------H---RDQL---S-S-----   193
O.CM.98.98CMU5337           N-------Q-----Q-KGHLRITDQLE--S---Y...   193
O.CM.99.99CMU4122           N-------Q---------HLRIMDQL-N-S-------   193
O.FR.92.VAU                 NR------Q-------RHLRI-DQP--QL-H-----   193
O.GA.11.11Gab6352           N-------Q-----K-S-HLMIRD-LE--S-------   193
O.SN.99.99SE_MP1299         N-------Q-------SRHLRIRDQLE--S-------   193
O.US.10.LTNP                S-------Q-------RHLR-RDQPE--S-------   193
N.CM.02.DJO0131             R----------------H---Q----NPI--------   193
N.CM.02.SJGddd              R---------------RH---Q----NPI--------   193
N.CM.04.04CM_1015_04        R-------A-------RH---Q----NPI--------   193
N.CM.06.U14296              R-------A-------SER---Q-R--NPI--------   193
N.CM.06.U14842              R---------------EH---Q----N-I--------   193
N.FR.11.N1_FR_2011          R---------------EH---Q----NP--------   193
P.CM.06.U14788              R-------QI--Q-I-SN--R---XQE--S-S-CYKP-   196
P.FR.09.RBF168              R-------QI--Q-I-SX--R-X-RQE--S-S-CYKP-   196
CPZ.CD.06.BF1167            HR-------V-A--------RI-A-QENRGI--Y...   197
CPZ.CM.05.SIVcpzMT145       NR------A--A--------R-RD-QE-PMSS-C...   193
CPZ.GA.88.GAB1              HR------A---------H-R--V-QENL-R------   194
CPZ.TZ.06.SIVcpzTAN13       PGR---A-V---------H--IR----N-II--Y...   199
CPZ.US.85.US_Marilyn        RR----------------C-----QPEN---S-----   193
GOR.CM.07.SIVgor2139_287    VR--F---Q--AA-T--RS-QMR--KEN-S-------   193
GOR.CM.07.SIVgorCP2135con   VR--F---Q--AA-T-SRS-QMR--KEN-ST------   193
```

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

```
B.FR.83.HXB2                 MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIW.LHGLGQHIYETYGDTWAGVEAIIRILQQLLFIH#FRIGCRHSRIGVTR....QRRA..RNGASRS..*   96
A1.AU.03.PS1044_Day0         --------------Y------------Q--------P-.------Y--N------E-------T------V-.-----Q-----IX-....G--G.---SG-X..-   97
A1.CH.03.HIV_CH_BID_V3538    --------------Y---M---------H--------Q-.-Q-------D------E----V--------V-.-----H-----LM-G...G--G..-D--G---..-   98
A1.CY.08.CY236               --------------Y------------H--------P-.------Y--N------E--Q----M------V-.-----Q-----I--....G--G..-D-SG-P..-   97
A1.ES.05.X1608_8             ---------------------------H--------P-.------Y--N------E--V---K-------V-.-----Q-----IIP....G--G.---S----..-   97
A1.KE.11.DEMA11KE001         --------------Y---M-D---D-----------M-.---------N------E----V--------V-.-----Q-----II-G...G--G.---S----..-   98
A1.RU.11.11RU6950            --------------Y---M-----DI-H--------P-.------Q--N------E-------T------V-.-----Q-----I-Q....R--V.---SR---..-   97
A1.RW.11.DEMA111RW002        --------------Y---M-----D--H--------P-.---------D-H----E---------------V-.-----Q-----II-....G--V..-D-PG---..-   97
A1.SN.01.DDI579              --------------Y---M--------Q--------P-.---------N------E-------M------V-.-----Q-----II-R...G--V..-D----P..-   98
A1.UG.11.DEMA110UG001        --------------Y---------D-----------P-.---------N------E------------I-V-.-----Q-----II-....G--G.---S----..-   97
A1.ZA.04.04ZASK162B1         --------------Y---M--------H-----Y--P-.--A------N------T---V----------V-.-----Q-----IIP....G--G..-D--G---..-   97
A2.CM.01.01CM_1445MV         --------------Y---M--------Q--------P-.--D------DI-----E-------Y------V-.-----Q-----II-....R--V..-D-T--P..-   97
A2.CY.94.94CY017_41          --------------Y---M--------Q--------H-.------Y--N------E---V---Y------V-.-----Q-----II-....R--V..-D----P..-   97

B.BR.10.10BR_MG029           -------H------Y------------S--------A-.--S-----T-----------LL----------.-----Q-----I--.........---.---T---..-   97
B.CA.07.502_1191_03          --------------Y-Q----------A--------P-.------Y-----------------T--------.-----Q-----ILQ.......---.R----N---..-   98
B.CH.08.M2_0803101_NFLG8     ----------....-S-----------S--------P-.--D---Y---------R------T---RM----.-----Q-----I--......---T.--------..-   93
B.CN.10.DEMB10CN002          ---------------------------S--------P-.--S---Y---------T-------------M----.Y----Q-----II-.........--.#D------..-   96
B.ES.10.DEMB10ES002          --------------Y---------------------P-.------Y---------T--L----M-------V-.-------------I-.......---T.---------..-   96
B.FR.11.DEMB11FR001          --R-----------Y----I-------R--------T-.--S---Y--Q------------------------.-----H-----IN-.....-T-T.--------..-   97
B.GB.05.MM45d213_GN1         --------------Y---------------------P-.------Y---------T----L-K-----------.-----Q------IQ....R--.---------..-   97
B.HK.06.HK003                -------H--P---Y-A-------------------T-.------Y---------E---------------V-.Y----Q-----I--.......--T.--------..-   97
B.HT.05.05HT_129389          --------------Y---A--------S--------V-.-------------------------------.-----Q-----I-PTR..-----.---TG---..-   99
B.JP.12.DEMB12JP001          --------------Y---------------------P-.------L---N----------------------.Y-L--Q-----I--............---G---..-   97
B.KR.07.HP_18_07JHS10_3909   --------------Y---------D--S--------P-.--S------------------A-----------.-----Q-----II-.............-------..-   97
B.PE.07.502_0525_wg5         --------------Y---------------------V-.------Y----------------------H------.-----Q-----I--............---------..-   97
B.RU.11.11RU21n              --------------Y-Q-------------------A-.------Y--------------L--T---------.-----Q-----II-....RS--..-H------..-   97
B.TH.08.MERLBDTRC10          --------------N------------S--------P-.--S---Y--R------T----L--M---------.-----Q-----I-P....T---.----------..-   97
B.US.11.ES38                 --------------Y----------------------.--S---Y---------E-----------------.-----Q-----II-.......----.---S----..-   97

C.AR.01.ARG4006              --------------Y------------Q--------P-.--S---Y---------M---------------V-.-------------II-.......----.--------..-   97
C.BR.07.DEMC07BR003          ----L---------Y------------Q--------P-.--S---Y---------T---------------V-.-----Q-----IL-.......----.---PN---..-   97
C.BW.00.00BW5031_1           --------------Y------------Q--------P-.--S------------M----L-----.....-.-----Q-----IIL.....-T.----------..-   92
C.CN.10.YNFL19               ---S---H------Y------------Q--------P-.-----------------LL-T------------.-----Q-----IV-.......-G..-T--N---..-   97
C.CY.09.CY260                --------------Y---A--------Q--------V-.--S-------------T----L--------------.-----Q-----IL-.......--------N---..-   97
C.ES.08.X2363_2              ---V--G--L----F------------Q--------P-.------Y---------T---V-------------.--N--Q-----II-.......-T.----G---..-   97
C.ET.02.02ET_288             --------------YS--------D--Q--------P-.--S---Y--N------R----L-TF---S----.-----H-----I--...........--------..-   97
C.IN.09.T125_2139            --------------YT--A--------Q--------P-.------Y---------T----L---------T-.-----Q-----IR-............--------..-   97
C.KE.00.KER2010              ---P----------Y------I-----Q--------P-.--SI--Y---------T-----------------.-----H-----IL-..........----------..-   97
C.MW.09.703010256_CH256.w96  --------------Y------------Q--------P-.--S-------------T---------------V-.-----Q-----II-.......----.----N---..-   97
C.TZ.08.707010457_CH457.w8   ---P--------------A--I-----Q--------P-.--A---Q--N------T-----------------.-----Q-----IV-...........---------..-   97
C.YE.02.02YE511              -------H------Y------------Q--------P-.--S---Y---------T-----L--------V-.-----Q-----TI-....H---.----N---..-   97
C.ZA.08.705010534_CH534.w12  ---------------------------Q--------P-.--S---Y--N-----T---VL------------.-----Q-----IL-...........----N---..-   97
C.ZA.10.DEMC10ZA001          -----G--------Y------------Q--------P-.--S---Y-------------------------------.-----Q-----II-.......-T.---------..-   97
C.ZM.11.DEMC11ZM006          --------------Y---------I-Q--------P-.--S---Y---------T-----------------.-----H-----IL-.......-T.---S----..-   97

D.CM.10.DEMD10CM009          -----A--------Y---A--------SK---------.---------D------T----L----------------.-----Q-----IIP.........----.---S----..-   97
D.CY.06.CY163                -----A--------Y------------S--------T-.------------------L--T---I-----.-----Q-----IIP....P---.---S----..-   97
D.KE.11.DEMD11KE003          --------------Y---------------------V-.------Y--------------------------.-----Q-----IL-.......-I.---SN---..-   97
D.KR.04.04KBH8               -----G--------Y---A--------S--------P-.--D-----A------V----LL-V------V-.-----Q-----II-.........---.---S----..-   97
D.SN.90.SE365                ---------------------------S--------T-.--S---Y-------------------------.-----Q-----I-P......-I.---S----..-   97
D.TZ.01.A280                 --------------Y------------S--------L-.--N---F-------------------------.-----Q-----I-L......-T.---S--P..-   97
D.UG.10.DEMD10UG004          --------------Y------------S--------P-.------Y-------------V------------.-----Q-----IM-.......-L.---S----..-   97
D.UG.11.DEMD11UG003          --------------Y---A--------R----------.--S------------T----LT-S-----V-.-----Q-----INP.......-T.----R---..-   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.AO.06.AO_06_ANG32         -----G--------Y---------------------P-.--S---Y--Q------L----L--M---M----.--V--P-----I-P.......-V.---N-----..-   97
F1.AR.02.ARE933              --------------Y---A-D---------------P-.--S------N----------LL-----------.-----H-----I--............---------..-   97
F1.BR.10.10BR_RJ015          ------------#---G----------H--------A-.------Y--N-.........................----H---M-I--.......-V..G---G---..-   73
F1.CY.08.CY222               -----G--------Y---A-----------------P-.---------N------E-----------------.-----Q------M-.......-V.---------..-   97
F1.ES.02.ES_X845_4           -----G--------Y---------------------P-.------Y--N------V----LV----------.--M--Q-----I--.......-V.---S----..-   97
F1.RO.96.BCI_R07             -----G--------Y---------------K----P-.---------N------E-----------------.-----Q-----I--.....-A-V.--------..-   97
F1.RU.08.D88_845             -----G--------Y---------------------P-.--S------D------E-----------------.-----Q-----I--.......-V.--------..-   97
F2.CM.02.02CM_0016BBY        ---P----------Y------------H--------E-.---------N------E-------T---------.-----H-----II-.......-I.---NG---..-   97
F2.CM.10.DEMF210CM001        ---P-----------------------Q--------E-.---------S------E-------T--------.-------------IIQ.......-I.---SG-P..-   97
F2.CM.10.DEMF210CM007        ---P----------Y------------Q--------E-.---M--Y--N------E-------M---------.-----H-----II-.......-I.---N----..-   97

G.BE.96.DRCBL                --R---------------A--------H--------P-.------L--N------E-----------------.-----Q-----IIPQ...R--V..-D-SG---..-   98
G.CM.10.DEMG10CM008          ---V--N-------Y---A--------H--------P-.------Y--N------E---------------V-.-----Q-----IIP....R--V..-D-PG---..-   97
G.CN.08.GX_2084_08           --------------F------------T--------P-.---------N------E-----------------.-----Q-----INLR...R--V..-D-SG---..-   98
G.CU.99.Cu74                 ------------------------------------P-.------YV-N------E----------------.-------------ISL....R--V..-D-PG---..-   97
G.ES.09.X2634_2              --------------Y-D-A--V--D--R--------P-.------Y--N------E---L-------------.-----Q-----IL-....R--V.M--P----..-   97
G.GH.03.03GH175G             ------N-------Y---A-----------------P-.------Y--N------E-------M--------.Y----Q-----IIP.......-V..-D-P--P..-   97
G.KE.09.DEMG09KE001          ---V----------Y-D-------------------P-.---------N------------------------.-----Q-----IIP.......-V..-D-PG---..-   97
G.NG.09.09NG_SC62            -K------------Y---I-G------V--------P-.---------N------E----TK----------.-----Q-----I-P.......-V..-D-SG---..-   97
G.ZA.01.TV546                --R-----------Y---A-----D-----------P-.-----X---X------E---X--------------.-----Q-----I-P.......-V.---P----..-   97

H.BE.93.VI991                ---------------------------S-----Y--P-.--D------N------E-------T-----A-.-----Q-----I--.......-V.---P----..-   97
H.CF.90.056                  ------------------------I-----------V-.--Q------N------V----L--T---------.-----Q-----I--.......-V.---P----..-   97
H.GB.00.00GBAC4001           --------------Y------------S--------A-.--Q---Y--N------R----LT--------T-.-----Q-----I--.......-V.---P----..-   97

J.CM.04.04CMU11421           ---P----------X---------------------X-.---------X------E--RX-------------.-----Q-----IMP.......-G.---------..-   97
J.SE.93.SE9280_7887          -------H------YH--------------------P-.--S---Y--S------E-----------------.-----H-----IIP.......-G.---------..-   97
J.SE.94.SE9173_7022          --------------YH--------------------P-.------YV-S------E----V------------.-----H-----IIP.......-G.---------..-   97

K.CD.97.97ZR_EQTB11          --------------Y----I-I-----R--------P-.--D------NN-----E-----------------.-----H-----IL-.......-G..G-------..-   97
K.CM.96.96CM_MP535           --------------N------I-----R--------P-.--N------T------E-L---------------.-----H-----IIP.......-G.---S----..-   97

U.CA.01.TV749                --------------Y------------S--------P-.--D------N------E-----------------.-----H-----II-.......-G.---SN---..-   97
U.CA.99.TV721                --------------Y------------S----------.--D------N------E-----------------.-----H-----II-.......-G.---S----..-   97
U.CD.83.83CD003_Z3           --------------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.CY.05.CY090                ---T----------Y------I-----R--------P-.--N------N------E---L--M---------.-----Q-----II-.......-R..-D------..-   97
U.CY.08.CY223                --------------N------------H--------P-.---------N------E--R--------------.-----Q-----IVP.......-G.---------..-   97
U.ES.10.DEURF10DZ001         --------------F------------Q--------P-.----------------E--R----------V-.-----Q-----IIP....G--G.----G---..-   97
U.GR.99.99GR303              --------------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

01_AE.AF.07.569M             ------.R--R---N------------I--------P-.------Y--N------E-------M------V-.-----Q-----IVP....G--G.----G---..-   96
01_AE.CF.90.90CF11697        --------------Y------------I--------P-.------Y--N------E-------T------V-.-----Q-----II-....G--G.----G---..-   97
01_AE.CN.10.YNFL03           ------N-------Y---------------------P-.------YV-N------E-----------------.-----Q-----IIP....G--G.----G---..-   97
01_AE.HK.04.HK001            --R-----------Y---------------X----XP-.--S---Y--D------E----T-T------V-.-----Q-----IIP....G--G.---------..-   97
01_AE.IR.10.10IR.THR48F      --------------------D---------------H-.------Y--NQ-----E---------------V-.-----Q-----IIP....G--G.---------..-   97
01_AE.JP.x.JRC77AE           ------------------------------------P-.---------N------E---------------V-.-----Q-----ILP....G--G.----G---..-   97
01_AE.TH.09.AA090a_WG11      ------------------M--I-----I--------P-.--S---Y--D-----LE--Q------------V-.Y----Q-----IIP....G--G.----G---..-   97
01_AE.TH.90.CM240            -------.E------Y--------------------P-.-Q----Y--NN-----E-------M------V-.-----Q-----IMP....G--G.---TG---..-   96
01_AE.US.05.306163_FL        -------------LN---------------------P-.------Y--T------E-----------------.-----Q-----ILP....G--G.----G---..-   97
01_AE.VN.98.98VNND15         ---------S----Y---------D-----------P-.------Y--NN-----E-------T------V-.-----Q-----IIP....G--G..G-E-G---..-   97
```

```
                                         amphipathic α-helix
                                      oligomerization
                          Vpr start                                           H(S/N)RIG motifs
                                                                    Vpr end in HXB2
                                                              frameshift in HXB2                          Vpr end
B.FR.83.HXB2                  MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIW.LHGLGQHIYETYGDTWAGVEAIIRILQQLLFIH#FRIGCRHSRIGVTR....QRRA..RNGASRS..*  96
02_AG.CM.10.DE00210CM013      --R-----------F------------H--------P-.----E-----------Q--Q-----------V-.-----Q-----IIP....G--G..--------..-  97
02_AG.CY.09.CY256             --------------F--A------------------P-.--E------N------E--R-L-----S---V-.-----Q-----II-....G--G..----G--..-  97
02_AG.ES.06.P1423             --------------F----------I-Q--------P-.---------N------E--V-------------.-----Q-----IIP....RXGT..GK.-G--..-  96
02_AG.GW.05.CC_0048           --R-----------F------------H--------P-.----------------E--G----L------V-.-----Q-----II-....G--G..--------..-  97
02_AG.KR.12.12MHR9            --R-----------F-------V-------------P-.---------N------E--G---T--H----V-.-----Q-----II-....G--G..----G--..-  97
02_AG.LR.x.POC44951           --------------F------------Q--------P-.-----EY--N------E-------T------V-.-----Q-----II-....G--G..---S---..-  97
02_AG.NG.09.09NG_SC61         ---V----------F--A---------H--------P-.---------N------E--G-----------V-.-----Q-----II-....G--G..----G--..-  97
02_AG.NG.x.IBNG               --R-----------S------------H--------P-.----------------E--K-----------V-.-----Q-----IIQ....G--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.98.98SE_MP1211       --R-----------F------------H--------P-.----------------E--G-----------V-.-----Q-----IIP....G--G..----G--..-  97
03_AB.RU.97.KAL153_2          --------------Y------------S--------V-.--S---Y--------------------------.-----Q-----IIQ....R---..--------..-  97
04_cpx.CY.94.94CY032_3        --------------N---------------------P-.---------N------E----------------.-----Q-----I-P.......---R#AGD-----..-  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
07_BC.CN.98.98CN009           ---S----------Y------------Q--------P-.--S--------------------T---------.-----Q-----IL-.......---T.--------..-  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-G...R--G..-D--G--..-  98
10_CD.TZ.96.96TZ_BF061        --------------Y----------------------.--N------------T-----------------.-----P-----IV-.......----.----T-T-..-  97
11_cpx.CM.95.95CM_1816        --------------Y--M---------H--------P-.---------N-----V---------------V-.-----Q-----IIP....G--G..--------..-  97
12_BF.AR.99.ARMA159           ---P----------Y---------------------P-.---------N------E----------------.-----Q-----IM-.......--V.--------..-  97
13_cpx.CM.96.96CM_1849        --------------N------------H---K----P-.---------N------E-----T-M--------.-----Q------V-.......--G.---S---..-  97
14_BG.ES.05.X1870             --------------Y---------------------P-.------Y--N------E---XL-X---------.-----Q-----I-P.......--V..-D-S---..-  97
15_01B.TH.99.99TH_MU2079      -G------------Y-G-P-----------K.-S-P-.------S--TN-----E-----V---H----V-.-----H----AIIP....G---.R---PG--..-  97
16_A2D.KR.97.97KR004          --------------Y--------D--R---------P*.--S---Y--A------T-----------M----.-----Q-----IM.....P-G-..---S---..-  95
17_BF.AR.99.ARMA038           -----G--------Y---------------------P-.--S------N------E----------------.-----H-----II-.......--V.---D---..-  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-P....R--V..-D-PG--..-  97
20_BG.CU.99.Cu103             ------------------------------------P-.-----V---N------E---L------------.-----Q-----ISL....R--V.---P----..-  97
21_A2D.KE.99.KER2003          ---------------D-----------S--------P-.--S------A-----------M-----------.-----H------I-.......----.---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-----ISP....R--G..---PG--..-  97
24_BG.ES.08.X2456_2           ------------------------------------P-.-----A---N------E----------------.-----Q-----ISP....R--R..-D-SG--..-  97
25_cpx.CM.02.1918LE           --RT----------N---------------------P-.------Y--N------E--V-------------.-----Q-----IIP.......--G..---P---..-  97
26_AU.CD.02.02CD_MBTB047      --------------Y------------Q--------P-.----------------G--------------V-.-----S-----IF-....S--G..----G-P..-  97
27_cpx.FR.04.04CD_FR_KZS      --------------N----I----------------P-.------YV-N------E--V-------------.-----Q-----IIP.......--G.---P--P..-  97
28_BF.BR.99.BREPM12609        --------------Y------------S---------.--------------------L-------M-----.---------------IN-.......----.--------..-  97
29_BF.BR.01.BREPM16704        --------------Y------------S---------.----------------------------------.---------------INP.......----.--------..-  97
31_BC.BR.04.04BR142           --------------N------------Q--------P-.------Y--------T--------------V-.-----Q-----II-.......----.---S---..-  97
32_06A1.EE.01.EE0369          --R------------------------S---K----P-.--S---Y--N------E----------------.-----N-----IIP.......--G.---T---..-  97
33_01B.ID.07.JKT189_C         -------------------------------K----P-.------X--N------X--------------X-.-----Q-----I-P....G--G..----G--..-  97
34_01B.TH.99.OUR2478P         -----A------------------------------P-.------Y--DN-----Q--------------V-.-----Q-----ILP....G--G..--------..-  97
35_AD.AF.07.169H              --------------Y--M----D--H----------P-.-Q---Y--N------------------------.------L----I--....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-#FMA*-NISMKPM-ILGKELKP**KYYSNYCLSI.-----Q-----I-P....R--V..---P---..-  94
38_BF1.UY.03.UY03_3389        ---------------------------A--------P-.--S-----N------E-----T-T---------.-----H-----IH-.......--V..---N---..-  97
39_BF.BR.04.04BRRJ179         -------H------Y------------T--------V-.------Y---------------T-T--------.-----Q-----I--.......----.---S---..-  97
40_BF.BR.05.05BRRJ055         --------------Y------------S--------P-.--N--Y---I----T------------------.-----H-----II-....L--S..---S---..-  97
42_BF.LU.06.luBF_18_06        --------------Y---------------------P-.------Y--------E---------L---T-.-------------L-L....R---..--------..-  97
43_02G.SA.03.J11223           ---V----------X------------H--------P-.---------N------E-------X-----V-.-----Q-----II-....G--G..-H--G--..-  97
44_BF.CL.00.CH80              -----G--------Y------------S--------P-.--S---Y--N------E----------------.-----H-----II-.......--V.----N---..-  97
45_cpx.FR.04.04FR_AUK         --------------Y--M------------------P-.--------------------V------------.-----Q-----IIP....G--V..-D-PG--..-  97
46_BF.BR.07.07BR_FPS625       -----G--------Y---------------------P-.--S------N------E----------------.-----Q-----II-.......--V.----G--..-  97
47_BF.ES.08.P1942             --------------Y---------------------P-.---------N------E-------M--H-----.-----N--K--II-.......--T.--------..-  97
48_01B.MY.07.07MYKT021        ------------------------------------T-.------Y--N------E--------------V-.-----Q-----IIP....G--G..----R---..-  97
49_cpx.GM.03.N26677           --------------Y-----I----Q----------P-.--S---Y--N------E--V---K---------.-----H-----II-.......--G.-------P..-  97
50_A1D.GB.10.12792            --R-----------Y-Q-A-DM--D--H----------.--E-----N------E-------T-----V-.-----Q-----II-....G--G..---S---..-  97
51_01B.SG.11.11SG_HM021       --------------Y-A-------------------P-.------Y--D------E-------T--------.Y----Q-----IIP....G--G..-------P..-  97
52_01B.MY.03.03MYKL018_1      ------------A-Y--A------------------P-.-Y---Y--TN-----E----------------V-.-----Q-----I-P....G--G..--------..-  97
53_01B.MY.11.11FIR164         --------------N------------H--------Q-.------X--N------E--Q-----------X-.-----Q-----IIP....G--G..--------..-  97
54_01B.MY.09.09MYSB023        -----------------------------------S-P-.-L-------N-----RE--------------V-.-----Q-----IIP....G--G..----G--..-  97
55_01B.CN.10.HNCS102056       -------H------Y---------------------P-.---------N------E----------I-----.-----Q-----ILP....G--G..----G--..-  97
56_cpx.FR.10.URF5_patient_A   --------------S------------H--------P-.---------------------------------.-----Q-----ILP....-G-V..-D-SD--..-  97
57_BC.CN.09.09YNLX19sg        ---P-----------------------Q--------P-.------YV-------T------K----------.-----Q-----IL-.......----.--------..-  97
58_01B.MY.09.09MYPR37         ---------------------------T--------P-.------Y--N------E--------------V-.-----Q-----I-P....G--G..----G--..-  97
59_01B.CN.09.09LNA423         -------H------Y---------------------P-.------Y--H------E-------T-----V-.-----Q-----IIP....G--G..----G--..-  97
60_BC.IT.11.BAV499            --------------Y-A----------Q--------P-.------Y--------T--------------V-.--V-GQ-----II-.....-N--.-------N..-  97
61_BC.CN.10.JL100010          ---------------------------Q--------P-.--S---YV-DN-----T----------------.Y----H-----IV-.......----.--------..-  97
62_BC.CN.10.YNFL13            ---P----------YX-----------Q--------T-.------Y----------------T-----T-.-----Q-----IL-.......----.--------..-  97
63_02A1.RU.10.10RU6637        --------------F------------H--------P-.------Y----------E--R-----------V-.-----Q-----I--....G--G..---S---..-  97
64_BC.CN.09.YNFL31            --RP---H------Y------------Q--------P-.------YV--N-----T---T------------.-----Q-----IM-.......----.--------..-  97
65_cpx.CN.10.YNFL01           ---S--------..R--G---------Q--------P-.---------N------E--------------V-.-----Q-----ILP....G--G..--------..-  95
67_01B.CN.11.2011.ANHUI.HF115 ------N-------Y---------------------P-.------Y--N------E-------T-----V-.-----Q-----ILP....G--G..----G--..-  97
68_01B.CN.11.2011.ANHUI.WH73  --------------Y---------------------P-.------Y--S------E-------T-----V-.-----Q-----ILP....G--G..----G--..-  97
70_BF1.BR.10.10BR_PE004       --------------Y--A---------S---------.---------D------------------------.-----Q-----IIQ....R---..---TG--..-  97
71_BF1.BR.10.10BR_PE008       ------------------------------------P-.-----------R--------LVK----------.---R-Q-----ILL....R--T..----G--..-  97
72_BF1.BR.10.10BR_MG002       --R-----------Y---------------------V-.------Y---N------------T--------.-----Q-----I-P....S---..--------..-  97
O.BE.87.ANT70                 ------N---AK--F--A---------A--------P-.--A---Y---------V--M-----------T-.Y----Q-----IN...PRG-GR..---S---..-  98
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.98.98CMABB141            ---V------A---F--A---------A--------P-.-MS---Y--D------V--M-------L-I-A-.-----Q-----INPSNTRG-GR..---S--P..- 101
O.CM.98.98CMABB212            ------N---A---F------------E--------P-.-QAC--Y--D------E--M-----------T-.Y----Q-----IIPSNTRG-GR..---S---..- 101
O.CM.98.98CMU5337             --R----H--A---F--V--T------A--------H-.-QA--NY---------V--M--T--------T-.-----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.GA.11.11Gab6352             -----N----P---F--A---------A--------P-.--A---Y---------V--M-----------T-.-----Q-----INPTNPRG-GR..---S---..- 101
O.SN.99.99SE_MP1299           ----------A---F--S---------A--------P-.-QA---Y--D------V--M-----------A-.-----Q-----I-PSNTRG-GR..---S---..- 101
O.US.10.LTNP                  ----------A---F--A---------T--------H-.------Y---N-----R--M--------M--T-.Y----Q-----ID...PRG-GR..---S---..-  98

N.CM.02.DJ00131               -------A------Y---------------------T-.---------N------E----------------.Y----Q-----I-P....RG-...--------..-  96
N.CM.02.SJGddd                -------A------Y--A-------------------.---------D------E-------M--------.Y----Q-----IIP.......---.--------..-  96
N.CM.04.04CM_1015_04          -------A------Y--A----D--------------.---------D------E-------M--------.Y----Q-----I-P.......---.--------..-  96
N.CM.06.U14296                --R----A------Y--A------------------M-.---------N------E-------M--------.Y----Q-----I-P.......---.--------..-  96
N.CM.06.U14842                --R----A------Y--X------------------X-.---------N------E----------------.Y----Q-----I-P.......---.--------..-  96
N.FR.11.N1_FR_2011            --R----T------Y----------------------.---------N------E----------------.-----Q-----I-P.......---.-------I-..-  96

P.CM.06.U14788                ---P------P---F--M-QT-----E---------P-.--S---Y--N------E--T---------I---.Y----Q-----ILTPS..R-GR..-H-P---..-  99
P.FR.09.RBF168                ----------P---F--M-QT-----A----P--LQY.--S---Y--N------E--T-----X--I---.Y----H-----ILPPS..R-GR..-H-P---..-  99

CPZ.CD.06.BF1167              ------N---P---W-Q-L-DT---I----M-----NT.-QNI-NWVFQN---S-E--QVL-TL--KA--T-.Y-Y--A-----I.......GPR..-------RS-  96
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.......---R..S--SN--..-  97
CPZ.TZ.06.SIVcpzTAN13         --------------Y-Q-I-DV-----Q--------PV.-Q-V-NWVFTV---S-E-AQEL-K---RA--S-.Y-H--V-----......................  84
CPZ.US.85.US_Marilyn          ---V------A---Y---A-DV-----E--L-----P-.-Q----Y--S------E----------------.Y----N-----ISLARRTPQGR..---S---..- 101

GOR.CM.07.SIVgor2139_287      ---I-T-E------F------M-----Q--------P-.--Q----V-D------E--T--------VI-V-.Y----H-----ILPPS..R-GR..---PN--..-  99
GOR.CM.07.SIVgorCP2135con     ---I-T-E------F------V-----Q--------P-.--Q----V-D------E--T--------VI-V-.Y----H-----ILPPS..R-GR..---PN--..-  99
```

Tat start · C-rich region · nuclear localization · exon 1 end · exon 2 start · Tat end

B.FR.83.HXB2	MEPVDPRLEPWKHPGSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRQ.RRRAHQNSQTHQASLSKQPTSQPR..GD.PTGPKE*KKKVERETETDPFD...............*	100
A1.AU.03.PS1044_Day0	-D----N----N------T---NK-------Y--PS--LN-G------------.--GPP-GNKNY-XPVP---IP-TQ..R-.S---E-P-----SKA---R--...............-	102
A1.CH.03.HIV_CH_BID_V3538	------N-----------T-P-S--------W------LK-G----------KH.--GT-PS-KD--NP-----IP-TQ..R-.S--QE-S-----SKA-A-R--...............-	102
A1.CY.08.CY236	-D----N----N------T---SK-------Y---D--LK-G------------.---PP-S-ED--NPIPE--IP-TQ..-V.S---E-S-E----SK--A-R--...............-	102
A1.ES.05.X1608_8	-D----N----N------T-P-SK-------Y------LN-G------------.--GTP-S-KD--NPIP---L--AQ..-I.----E-S--E--SK--A-R-A...............-	102
A1.KE.11.DEMA11KE001	-D----N----N------T-P-NK----R--Y--PA--LN-G------------.---TPHS-KD--NPVP------TQ..--.S---E-S-----SKA-A-R--...............-	102
A1.RU.11.11RU6950	-D----N----N----------S--------W---L--LK-G-----------H.--GTPHS-KD--NPI----LPHTQ..R-.Q---E-SA----SKA---R-A...............-	102
A1.RW.11.DEMA111RW002	-D----N----N------A---S--------Y--P---LS-G--V-------KR.--GTP-SNKD--NPVP--SIP-TK..-N.----E-S--E--SKA---R--...............-	102
A1.SN.01.DDI579	------N----H----------S-----A--W------LN-G-----------K-.--GTP-SNKD--NPVP---LP-TP..--.------S-----SK----R-A...............-	102
A1.UG.11.DEMA110UG001	-D----N----N------A---SP-------Y------LN-G------------.---TSHS-KD--DLVP--SIP-TQ..-V.---T--S-----SK---RL-Q.....RADS.....-	107
A1.ZA.04.04ZASK162B1	------N-D--N------S-P-S-----Y--W------LN-G-----------R.--GSPPS-KD--NPIQ---I--T-..R-.S---E-S-----SKA---R--...............-	102
A2.CM.01.01CM_1445MV	------K----N------R---N-----V--W------LN-G-----------P.---PS-S-KD--NPV-E-SLPHAQ..RV.S--Q--S--E--SK---HQ--...............-	102
A2.CY.94.94CY017_41	------K----N-----------K----R--Y---L---N-G-----------P.--KPSPSNKD--NPIP--SLP-AQ..RV.----E-P--E--SKA---R--...............-	102
B.BR.10.10BR_MG029	-----------Q------R----T-------Y---A--L--G------------.---TP-G-K--D--P---A--Q..-E.----E-S-----K------GY...............-	102
B.CA.07.502_1191_03	---I--S----------------N---------------K-G------------.---SP-S-KID--P-----A--H-..--.-------P--E------HQVG...............-	102
B.CH.08.M2_0803101_NFLG8	-------------------R---N-----G--------TK-G------------.---S-PD-----------------S-..--.------P--------IH-S-...............-	102
B.CN.10.DEMB10CN002	---------------------------------L-----TK--------------.---PP-D-K--NP-P---A----..--.------S-----------VH...............-	102
B.ES.10.DEMB10ES002	---------------------------------------TQ-G------------.----P-D----V---------------..--.S-----S----T-----DP...............-	102
B.FR.11.DEMB11FR001	----------------------NK----R--Y---L-----G------------.---TP-D-P-------------------..--.------S-E----KA--H-DA...............-	102
B.GB.05.MM45d213_GN1	---------------------------R---------K-G---N--------.---PSEG-----D----------------..--.Q-----S-E-----------G-W.....STDS.....-	107
B.HK.06.HK003	--------------------------------L------MQ-G------------.----PPE--N--V--P------C-..--.----E-S-----T--A---VP...............-	102
B.HT.05.05HT_129389	--L------------R---X-----X--Y-----L--G------------.---XS-X------A-----A-----..--.-K---Q-----------Q-W...............-	103
B.JP.12.DEMB12JP001	------S--------------N----------P---TK-G------------.---PP-G-----VP-P---A--L-..--.------S-E---------QV-...............-	102
B.KR.07.HP_18_07JHS10_3909	--------------------P----------L-------K-G------------.----P-DNKN--V----------A-..--.--Q--S-E---K--M--VH...............-	102
B.PE.07.502_0525_wg5	------N-----------R---N---------W------S--G------------.----PKD-------PI-----------..--.-A---S-----T---H-H-...............-	102
B.RU.11.11RU21n	-------------------------------L------MK-G------------.---SP--N--D----P---A----..--.------S-----------H--...............-	102
B.TH.08.MERLBDTRC10	------------------R-T------------------K-G------------.--GPP-D-------------A----..--.-E---S-------------GH...............-	102
B.US.11.ES38	-D----N-----------R---------------------G------------.----P-G-PNN-VP--E--A---S..--.------P---------K---H...............-	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.07.DEMC07BR003	-D-I-HN----N--------P-NK----R-SY---I--L--G------------.--S-PPS-ED--DLI----L--A-..--.Q---EK------SK--A----...............-	102
C.BW.00.00BW5031_1	------N----N--------P-NK----Q-SY--P---L--G---Y--------.--S-PSS-EN--NPV----LP-A-..-N.S--SE-S-----SK-------W..........RTDS-	107
C.CN.10.YNFL19	L--I--N----N------D---N-----R-SY--L---QK-G------------.--S-P-S-ED--NPI----L-RTK..--.S--SE-S-----SK-K----A...............-	102
C.CY.09.CY260	---I--N----N-----------NK----Y--Y--LA--Q--G------------.--S-PR--ED--DLI----L--T-..--.------E-S*--K-SNK-ETQLV...............-	101
C.ES.08.X2363_2	--L---K----N--------P-NT----S-SY--IA--QK-G------------.--S-PSS-KS--DLIP---LP-T-..--.Q--SE-S-----SK--A--S-...............-	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.IN.09.T125_2139	------N-D--N------D---N--F--S-SY--L---QK-G------------.--S-P-S-ED--NPI----L-RAQ..--.---S--S-----SK------A...............-	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.MW.09.703010256_CH256.w96	---I--K----N-----------NK------SY--L---Q--G------------.--S-PPS-KD--NP-P---L--T-..-N.---SE-S-----SQ-------W...............-	103
C.TZ.08.707010457_CH457.w8	--Q---K----N-----------NT-F--R-SY--L---Q--G-----------K.--S-PPS-ED--NPI----LP-T-..--.---SE-S--E--SK-T---L-...............-	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.....RTDS.....-	107
C.ZA.08.705010534_CH534.w12	---I--N----N-----------NK------SY--L---Q--G------------.--STPSS-KD--NP-----LP-T-..-N.S--S--S--E--SK--------...............-	102
C.ZA.10.DEMC10ZA001	--Q---K------------P-NS------SY--L---QK-G------------.--S-PPS-ED--NPI----ST-T-..--.---S--S-----SK--A-QSG...............-	102
C.ZM.11.DEMC11ZM006	-D----K----N-----------NK----R--Y--L---Q--G------------.--S-PPG-KD--NPI----SP-A-..-N.---SE-S-----SK-KA--YA...............-	102
D.CM.10.DEMD10CM009	-D----N----N------S---NS----R--Y---R---S-G----------G-.--KPPKDNT---VPVPD--S--H-..--.-----------E--SK--A----W...............-	103
D.CY.06.CY163	-D----N----N------S---NS----R--Y---C-----G----------G-.---PP-GNK---GPVPE--S--HP..-P.-P--NPS--E-ASK--A----W...............-	103
D.KE.11.DEMD11KE003	-D-I--S----N------R-P-NK-------Y---Q--VS-G------------.----P-GG-V--DPIP---S----..--.---------------SKA-----AW...............-	103
D.KR.04.04KBH8	-D----N-----------R---N--------L------LK-G------------.---PP-GD-A--VP-----P----..--.----------Q--E--SK--ANQ--W...............-	103
D.SN.90.SE365	-D----N----N--------P-N--------Y---N-----G------------.---TP-GN----V-VP---S--H-..--.---------------SKA---Q--W...............-	103
D.TZ.01.A280	-D----N----N------R-P-N--------Y---N-----G------------.---TP-GD-A--DPIP---S----..--.Q---------------SKA-A-R--W...............-	103
D.UG.10.DEMD10UG004	-D----NI---N--------P-NK----N--Y---------G------------.---PP-SG-A--DPIP---S----..--.---------------SKA---Q--C............WE.	104
D.UG.11.DEMD11UG003	--L---N-D--N------R-P-NK-------Y---N-----G------------.--GTP-GG-A--VPVP---S--L-..--.---------------S-A----Y-W...............-	103
D.YE.02.02YE516	-D----N---------R-S-P-SP----------------G----------G-.---SP-SN----VPVP---S----..--.---------------SQAKAP---W...............-	103
D.ZA.90.R1	-D----N----N--------P-NK-H-----Y---L---S-G------------.--KPPHSD-D--VPIPE--S----..--.---------------S-A--NQ--W...............-	103
F1.AO.06.AO_06_ANG32	--I---N-D--N------T-P--K----A-----YY--G--G------------.---TP-S-KN--NPVPE--L--A-..--.------S--EM-SK---S-*-...............-	101
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.BR.10.10BR_RJ015	--L---N-D--N------A-P-NK----R-----YW--T--G------------.-H-HR---L--DPVP---S--A-..--.------S--E--SK-----CA...............-	102
F1.CY.08.CY222	--------D--N------T-P--R----------YY--A--G------------.-H-TP-S--N--DPV----L--A-..--.------S--E--SK-----C-...............-	102
F1.ES.02.ES_X845_4	--V---N-D--N------T-P--T----V-----YR--T--G------------.---PP-SNKN--NPVPE--L--A-..--.------S--E--SK---G-*-...............-	101
F1.RO.96.BCI_R07	------N-D--N------T-P-NR----------YH--T--G-D----------.-H-TP-S-PI--NPVP---L--A-..-N.------S--E--NK--A--*-...............-	101
F1.RU.08.D88_845	--L---N-D--N------A-P--K----------R--TM-G------------.--GTX-S--I--NPVPE--S----..--.---Q--S-----SKA----*-...............-	101
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.10.DEMF210CM001	--V---KID--N------E-P-NK-----------L--T--G------------.---TP-S-EI--NPV----L--A-..--.---------S-----SKAK---*-...............-	101
F2.CM.10.DEMF210CM007	--I---N-D--Q------E-P-N------------R--TS-G------------.---SP-S-KV--DPVP---L--A-..--.----E-S-----SK-K-G-*-...............-	101
G.BE.96.DRCBL	-D----K-------------P-N-----SVAAL.----LN-G----------KH.--GTPHS-KD--TPVP---F-TT-..-N.----Q-S--E--SK-------...............-	101
G.CM.10.DEMG10CM008	-D----K-------------P-NK----A--W------LN-G----------KH.--GSSHG-KD--NPVPE--SPIS-..-N.---Q--S--E-ASKA---Q--...............-	102
G.CN.08.GX_2084_08	-D----N----N----K-----NK----V--W---L--LN-G----------KH.--ATPSSNKD--NPVP---IPTT-..-N.--D---S-----SK-K-----...............-	102
G.CU.99.Cu74	-D----N----N-----------N-----R--W---S--LH-G----------KH.--GTP-G-KD--NPVP---LPTT-..-N.------S--E-ASK-------...............-	102
G.ES.09.X2634_2	-D----N----N------------K----V--W------LN-G----------KH.---TP--NTD--NPVPE--LPTI-..-N.----E-S-----SK---N---...............-	102
G.GH.03.03GH175G	-D----N----N------T-P-NK----V--W------LN-G----------K-.--GTP-G-KD--NPVP---LPIS-..-N.Q--*--P--E-ASK---G-CA...............-	101
G.KE.09.DEMG09KE001	-D----N-----------R---N-----V--W---L--LN-G----------KH.----SPG-KD--NPVPE--LPTT-..-N.-A---S--E--S--K-----...............-	102
G.NG.09.09NG_SC62	-D----KI---N-----------NK----V--Y------LN-G----------NH.---TPSG--N--NPVP---FPTT-..-N.------S--E--SK-------...............-	102
G.ZA.01.TV546	-D----N----N--------P-NK----V--W------LN-G----------K-.--GTP-D-KD--NPVP--SLPTT-..-I.----E-S-----SK--AN-C-...............-	102
H.BE.93.VI991	-D----NQ---N------R---N--------Y---L--LK-G---Y--------.--GTPKSL-D--TLIP---L-RTS..--.----EK-----ASK------W...............-	103
H.CF.90.056	-D----K----N------Q---N--------Y---M--LK-G----------S-.-H-TPASL-D--N-I----L-RTH..--.-------Q--E-ASK----*-...............-	101
H.GB.00.00GBAC4001	-D----K----N-----------NG----N--Y---L--LK-G------------.--GTPAGL-DN-DPIP---LPRT-..--.---------------ASKA---LH...............-	102
J.CM.04.04CMU11421	------NI---N------R----Q---------------LK-G---N--------.--A-PPG-KX--DXVXE--LPLTS..RK.Q-----Q--E--SKA-PXR--...............-	102
J.SE.93.SE9280_7887	------NR---N--------------------Y------LQ-G------------.--S-PPG-K---DLIP---L--TQ..RK.----E-S--E--SKA-P-R--...............-	102
J.SE.94.SE9173_7022	------NR---N--------------------Y------LQ-G------------.--S-PPG-KN--DLIPE--LF-TQ..RK.----E-S--E--SKA-P-R--...............-	102
K.CD.97.97ZR_EQTB11	------NI---NQ--------NQ-------Y---L--LQ-G---C---E----.-TTTPYA-KN-KDPIP---LP-A-..--.------S--E--SK-K---*-...............-	101
K.CM.96.96CM_MP535	-D----NI---NQ--------NQ----R--Y---I--LK-G---N------P.--TTPY--EN--DP-R---L----..-E.Q-D---S-----SK-K--Q--...............-	102
U.CA.01.TV749	-D-I--NI---NQ---------N--------Y---L--LK-G---H--------.--A-PLS-KD--NPVP--S-P-T-..RV.S--S--SQE---SQ--P-R--...............-	102
U.CA.99.TV721	-D----NI---NQ---------N---------------LK-G-----------P.--A-PLS-KD--NPVP--SIP-AQ..R-.S--S--SQE---SQA-P-R--...............-	102
U.CD.83.83CD003_Z3	-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.CY.05.CY090	------N----N----------NK-F--------VA--TR-G------------.----S-SN-N--NPIP---L-HA-..-N.------S-EE--SK-----C-...............-	102
U.CY.08.CY223	------NI---NQ---------NQ-------W---L--LK-G------------.--A-SRS-KN--DHIPE-SLP-T-..RV.S--SE-S-E---SK--P-R--...............-	102
U.ES.10.DEURF10DZ001	--L---S-----------T-P-N-----F--W---L--LN-G-----------R.-PGPP-GR-D--D-VP---L-IF-..-N.S--S--Q--E-ASK---GQ--...............-	102
U.GR.99.99GR303	-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
01_AE.AF.07.569M	--L---N----N------T-N-S--------W---L--LK-G----------KH.--GTP-S-KD--NPIP---LPII-..-N.---Q--S--E-ASKA----C-...............-	102
01_AE.CF.90.90CF11697	------S----N------T---SK-------W---L--LK-G----------KH.--GPPPG-KD--NPIP---LPTT-..-N.------S--E-AKKAK----A...............-	102
01_AE.CN.10.YNFL03	------N-----------T-P-SP-------W---L--LK-G---H-----KH.--GTP-S-KD--NPIP---LPII-..-NH--D---S--E-ASKA---*-...............-	102
01_AE.HK.04.HK001	------N----N----K-T----K----I-SW---L--LK-G----------KH.--GTPRS-EG--NPVP--SL-TS-..-N.Q-D-E-S---AGK-----*-...............-	101
01_AE.IR.10.10IR.THR48F	------N----N------P-V-SK----I--W---L--LK-G----------KH.---TP-SNKN--DPIPE-SLPTS..-N.--D---S--E-ASK---G-CA...............-	102
01_AE.JP.x.JRC77AE	--L---N----N----K-T---SS----R--W---L--LK-G----------KH.--GTP-SRKD--NPVPE--LPII-..-N.--D---S--E-ASKA---C-...............-	102
01_AE.TH.09.AA090a_WG11	------N-----------T-D-SK-------W---L--LK-G----------KH.--GT--SNKD--NPIP---FPIS-..-N.S----S--E-ASKA---YA...............-	102
01_AE.TH.90.CM240	--L---N----N------T---SK-------W---L--LK-G----------KH.--GTP-S-KD--NPIP---LPII-..RN.--D---S--E-ASKA--QC-...............-	102
01_AE.US.05.306163_FL	------N-----------T-P-S--------W---L--LK-G----------KH.--GTP-S-KD--NPVP---L-II-..-N.---Q--S--E-ASK---G-C-...............-	102
01_AE.VN.98.98VNND15	-K----N----N------T-N-NT-------W---L--LK-G----------KH.--GTP-S-KD--NPIRE--LPII-..-N.--D---S--EGASKA---*-...............-	101

```
                             Tat start              C-rich region          nuclear localization    exon 1 end exon 2 start                    Tat end
B.FR.83.HXB2                 MEPVDPRLEPWKHPGSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRQ.RRRAHQNSQTHQASLSKQPTSQPR..GD.PTGPKE*KKKVERETETDPFD...............*  100
02_AG.CM.10.DE00210CM013     ------S----N------A---S-----W--W---L--LN-G-----------H.--GTP--RRD--NPVP---LPTT-..-N.----E-P--E--SK-----Y-...............-  102
02_AG.CY.09.CY256            --L---K-----------T---NT-------W---L--LK-G--------Q-GR.---TP-SR-D--NPVP---LPTTS..-N.---Q--S--E--SK--A--CA...............-  102
02_AG.ES.06.P1423            K.----S----N----------S--------W---Y--LN-G-----------P.---TP-SR-D--NPVP---LPTT-..-N.------S--E-ASK---Q*-...............-  100
02_AG.GW.05.CC_0048          ------S----N------T---SK----I--W---L--LK-G------------Q-GTPPS--D--NPVP---FTTI-..-I.-A----S--E-ASK----RC-...............-  103
02_AG.KR.12.12MHR9           --L---S----N------T---SK-------W---L--LN-G-----------R.G-GTP-SH-D--NPVP---LPTT-..-N.------S--E-ASKA--NQC-...............-  102
02_AG.LR.x.POC44951          -D----SI---N------T---SK----A--W---L--LN-G----------KR.--GSPRSH-DR-NPVP---LPTA-..-N.------S--E-ASK---NLC-Y..............-  103
02_AG.NG.09.09NG_SC61        --L---S----N------T---N--------W---L--LN-G-----------R.--GTPKSR-D--NPVP---LPAT-..-N.------S--E--SK-----C-...............-  102
02_AG.NG.x.IBNG              --L---S----N------T---SK---M---#---L--LN-G-----------R.--GTP-SR-D--NPVP---LPTT-..-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---LPTT-..-N.-A----S--E-AGK-----C-...............-  102
02_AG.SN.98.98SE_MP1211      --L---K----N------T---SK----L--W---L--LN-G-----------R.--GTP-SR-D--NPVP---LPTT-..-E.Q----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.94.94CY032_3       ------D----N------T-D-NK-F-----W------LK-G----------KH.--GSL-G-KG--NLIP---L--QPN.--.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.............-  103
06_cpx.AU.96.BFP90           ------KI---NQ---R------K-------Y--P---LN-G-------------.--Q-PPG-KN--DPV----L--TQ..RE.Q---EKS--E--SKA-P-R--...............-  102
07_BC.CN.98.98CN009          ------K--S----------------------------MK-G---F---------.--S-P-S-ED--NLI----LPRTQ.---.---SE-S-----SK-K-----...............-  102
08_BC.CN.97.97CNGX_6F        ------N----N------E---N-----R-SY--L---QK-G-------------.--S-P-S-ED--NLI---SLPRTQ.---.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.....WTDS.....-  107
11_cpx.CM.95.95CM_1816       ------NI---NQ---------NQ-------Y---H--LK-G-------------.--P-SHS-KN--DHIPE--L--AT.RK.--V-E-S--E--SK--P-R--...............-  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.96.96CM_1849       -D----KI---NQ-------P--K------------------G-------------.--S-SH---N--DLIP---LPI--..-N.Q----S--E--SK-------W.....SADS.....-  107
14_BG.ES.05.X1870            -D----N----N-----------K----R--W------LN-G----------KH.--GPP-SXKD--TPVP---LPTT-..-N.----E-S-----SK-K-----...............-  102
15_01B.TH.99.99TH_MU2079     -D---LN----N----R-T---N-----I--W---I--LK-G----------KH.--GTP-S-KD--NPAP---L-IS-..-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-Q..RI.---TE-S-TE--SKA----C-...............-  102
17_BF.AR.99.ARMA038          --I---N-D--N------T-P--R-F--E-----YW--T--G---N---------.---PP-GNPP--V--P---AP-A-..-N.------S--E--SKAK--Q*-W...............-  102
18_cpx.CU.99.CU76            ------NI---NQ---------N-----R--Y---L--LK-G-------------.--TTPYD-KH--DLIPE--L-RTS..W-.------Q-E--ASKA-A-QC-...............-  102
19_cpx.CU.99.CU7             -D----N----N----------NT-----------L--TK-G-------------.--GPP-SH-D--NPVP---LPDT..EG.---T--S--E--SK-K-----...............-  101
20_BG.CU.99.Cu103            -D----N----N----------NK----V--W------LN-G----------KH.--GPPPG-KD--NPVP---IPTT-..-N.------S-EE--SK-T---CA...............-  102
21_A2D.KE.99.KER2003         -D----N----N------R---NK-H--R--Y--IK--L--G---R----------.---PP-GD-A--VPVP---S----.--.------------S-A------C...............-  103
22_01A1.CM.01.01CM_0001BBY   --L---S----N------T---N--F--V--W------LK-G----------KH.--GTP-G-KD--NPIP---LP-TK..R-.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---LPTT-..-N.------S--E--SK--A----...............-  102
24_BG.ES.08.X2456_2          -D----N----N----------NK-F--V--W------LN-G----------KH.--GSP-S-KD--NPVP---IPTT-..-N.----EKS-EE-ASK-KA----...............-  102
25_cpx.CM.02.1918LE          -D----NI---N------G-P-N-----R-----------K-G----------S-.--GTPPS-KD--NPVP---IPTT-..-N.Q---E-S----ASK-----YA...............-  102
26_AU.CD.02.02CD_MBTB047     -----RNID--N--------P-NK----Q--Y---L-----G-----------P.--GTSRSN-D--NP-P---S--AQ..RE.SP--E-S-----SK--Q-R--...............-  102
27_cpx.FR.04.04CD_FR_KZS     -D----N----NQ---------NQ----R--Y---H--LK-G--F-----------.--T-PYGNKN--DPIQE-SIP-TQ..R-.S--SE-S-----SKARA-R--...............-  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.01.BREPM16704       ------------------R---NK------------------G-------------.---T--D---D-V------A--F-.--------S--T----A-----...............-  102
31_BC.BR.04.04BR142          -D----K----N----------N-----R-SY------L--G-------------.--S-PPS-ED--NLI----VP-S-.---.Q--SEKS-----SK-----C-...............-  102
32_06A1.EE.01.EE0369         -----HKI---NQ---R---P--P-------Y--YM--LN-G---Y----------.--T-PPG-KN--DPE----V--TQ..RE.----EK---E--SKATPNR--...............-  102
33_01B.ID.07.JKT189_C        ------N----N-------AT-SK----I--W---L--LK-G----------KH.---PP-S-ED--NPX-E--L-IS-..-N.--D-E-S---ASK----RC-...............-  102
34_01B.TH.99.OUR2478P        ------N----N------T---SK----N--W---L--LK-G----------KH.--GTP-S-KD--NPIPE--LPII-..-N.Q-----S--E-ASKA----*-...............-  101
35_AD.AF.07.169H             -D----N-D--N------P-P-S-----I-SY------LN-G-------------.---TP-RD-D--NPIP---LPRAK..-I.SA-QE-S-----SK--P-R--...............-  102
36_cpx.CM.00.00CMNYU830      --L---NI---N------R---N--F-----W------LN-G-----------R.--G-P-SR-D--DPV----LPAI-..-KN------S--E-ASK---N-C-...............-  103
37_cpx.CM.00.00CMNYU926      -D----N---------------N---------------TK-G-------------.--T-A-SR-DN-H-VP---LPTT-..-N.------S--E-ASKA----CA...............-  102
38_BF1.UY.03.UY03_3389       --I---K----N------T-P-NK----R-----YW--T--G---H----------.---TP-SN-I--DPVP---V--A-..-N.------S--E--SK-K---CA...............-  102
39_BF.BR.04.04BRRJ179        -D---------------------K-------------T--G-------------.---TP---I--DPVPE--I--AH..RN.------S--E--SKAK---*-...............-  101
40_BF.BR.05.05BRRJ055        -D---------------------K----R-------------G-------------.---SP-GR-S--G--#--SAP-Q-..--.----E-S-----------E-...............-  101
42_BF.LU.06.luBF_18_06       -----HK----E----------------R----------R-G-------------.--NPPPD-E-----V----AT-H-.--.------Q-----------DL...............-  102
43_02G.SA.03.J11223          --L---S----N------T---SK-------W---L--LN-G-----------R.--GTP-SR-D--NPV----LPITS..EN.------S--E--SK-------...............-  102
44_BF.CL.00.CH80             --LI--N-D--N------T-P--R----R-----YW--V--G-------------.---PP-SD-I--DPVP---L--T-.--.------S--E--SKA----CA...............-  102
45_cpx.FR.04.04FR_AUK        -D----N----N------T---NK-H-----Y--YS--LN-G----------KH.--GTPYS-KN--NPIP---L--T-.--.------S--E--SK-K-----...............-  102
46_BF.BR.07.07BR_FPS625      --L---N-D--N------T-P--K----R------W--V--G-------------.--GPP-S--I--DLVP---I--AQ..RN.----E-S--E--SKA--G-W-...............-  102
47_BF.ES.08.P1942            ------------------R----K----R-------------G-------------.--XPP-SG---D-------T-L-.--.----E-S-----------V-R.....RADS.....-  107
48_01B.MY.07.07MYKT021       ------N---------K-T---SK-------W--RL--LK-G----------KH.--GTSHRRED--NPIPE--L-IS-..-N.--D---S----ASKA--G-C-...............-  102
49_cpx.GM.03.N26677          ------KI---N-----------Q-------Y---I--LK-G---C----------.--T-PPG-KN--DLVQ---L--TQ..RK.----E-S--E--SKAGP-R--...............-  102
50_A1D.GB.10.12792           -D----N----N------T-P--K-------Y--PA--LN-G----------K-.---GPPKGD-A--VPIP---S----.--.-K----------SKA-A-Q--W.............-  103
51_01B.SG.11.11SG_HM021      ------N----N------G-P--A----R-S----A---K--------------.---T--D-E---------A----.-E.------S-E--------HSSRQ.......RLDS...-  107
52_01B.MY.03.03MYKL018_1     ------N----N------T---SS----I--W---L--LK-G---H-----KH.--GTP-S-KD--NPVPE--FPII-..-N.------S--E-ASKAA---C-...............-  102
53_01B.MY.11.11FIR164        ------S----N----X-T---SK-------W---L--LK-G-----------H.--GTP-S-KD--NPIP---LPIX-..-NNS-D---S--E-ASKA--X-C-...............-  103
54_01B.MY.09.09MYSB023       ------N-----------T---SK----R--W---L--LK-G----------KH.--GPP-S-KD--NPIPE--LPTT-..-N.------S--E-ASKA--G-*-...............-  101
55_01B.CN.10.HNCS102056      ------N----N----K-T---SK-------W---L--LK-G----------KH.--GNP-SR-D--NPVPE--LPII-..-N.--D---S--E-ASKA----C-...............-  102
56_cpx.FR.10.URF5_patient_A  -D-I--N----N----------NK----A--W---L--LN-G-X--------SH.---TPPG-KD--NPVP---LPTIN..-N.---QE-S----ASK-----*-...............-  101
57_BC.CN.09.09YNLX19sg       ------N----N----------N-----R-SY--L---QK-G-------------.--STPPS-ED--NLI----LPRTQ.---.Q--SE-S-----SK-KP-R--...............-  102
58_01B.MY.09.09MYPR37        ------N----N------T-D-SK----I--W---L--LK-G----------KH.--GTP-S-KD--NPVPE--LPTT-..-N.Q-N---S--E-ASQ-----CA...............-  102
59_01B.CN.09.09LNA423        ------N----N----K-T---SK-------W---L--LK-G---H-----KH.--GTP-S-KD--NPVPE-SSPII-..-N.------S--E-ASKA----YA...............-  102
60_BC.IT.11.BAV499           ---LETN----NQ-E-----P-N-----R-SY------L--G---Y----------.--S-PSSNKD--NPV----L--A-.---.Q-DLEKS----KSK-K-----...............-  102
61_BC.CN.10.JL100010         ------N----N------E---N--F--R-SY--L---Q--G-------------.----P-S-ED--SLI----LPRTQ.---.T--SE-S-----SK--P-----...............-  102
62_BC.CN.10.YNFL13           ------N----N------E---N-----R-S---L---QK-G-------------.----P-S-ED--NPIP---LP-IH.--.---SE-S-----SK-K----A...............-  102
63_02A1.RU.10.10RU6637       -D----N----N----------S--------W---L--LN-G----------GR.--GTS-SR-D--NPVP---LPTT-..-N.----E-S--E-ASK----QC-...............-  102
64_BC.CN.09.YNFL31           ------N----N------E---N--F----SY--L---HR-G--V-----------.----PPSGED--NPI---SLPKAQ.--.---SE-S-----SK-RA-Q--...............-  102
65_cpx.CN.10.YNFL01          ------N-----------Q---N--F--------X---TK--------------.---TP-S-EDY-NP-----L-II-..-NN--DQ--S--E-ASKA----S-...............-  103
67_01B.CN.11.2011.ANHUI.HF115 ------N-----------T-D-SP------VW---L--LK-G---H-----KH.---SPRS-KD--NPIP---L-IS-..-N.----E-S----ASKA----C-...............-  102
68_01B.CN.11.2011.ANHUI.WH73 ------N-----------T---SK------AW---L--LQ-G----------KH.---SPRS-KGY-NP-PX--L-II-..-N.----EKS---ASK----QG-...............-  102
70_BF1.BR.10.10BR_PE004      --L---K-----------R---K-------Y--------G--------W--.----SHS--N--V--P--------.--.------PE----------QDN...............-  102
71_BF1.BR.10.10BR_PE008      ----------Q-------R---T-------SY---S--L--G-------------.--G-P-SG-----------A--L-.--.------SEE---G----HLGNQ..........WMDS-  107
72_BF1.BR.10.10BR_MG002      ------K---------------N--F-----Y---A----G-------------.--KPP-D--------P----T-L-.--.----E-S--------A--QKH...............-  102
O.BE.87.ANT70                -D----EVP--H------QIP-N-----R--Y--Y---VR-G----------GR...P-AASHPD-KDPVP--SPTITK..RK.QERQE-QEEE--KKAGPGGYPRRKGSCHCCTRTSEQ-  115
O.CM.98.98CMA104             -D----EMP--H------Q-P-N-----S--Y--Y--LVR-G-----------R...P-TASHPDNKDPVPE-SP-ITN..RK.*ECQE-QE-E--K--GPG*YHC............SK.  100
O.CM.98.98CMABB141           -D----EMP--HQ-----PNP-N--------Y--Y---TS-G-----------R.-SA-AASRPDNKDLVP--SP-ITX..RK.QERQE-QE-E--KKAGP-**HR............RK.  102
O.CM.98.98CMABB212           -D----EMP--H----K-Q-P-NS----R--Y--L---TK-G---H------R..PA-AASYPDNKDPVPE-SP-NTG..RK.QKRQE-QE----T--DPGGRHCLQDSCNSCTRISGQ-  116
O.CM.98.98CMU5337            -D----EIP--H------PNP-N-----N--Y--Y---TR-G---------Y-R...--TTRHXD-KHLVP--SP-ITN..RK.QER-E-QE-E--KKAGPS*YPR............LK.  101
O.CM.99.99CMU4122            -D----EMP--H------Q-P-NK-------Y--Y---TS-G-----------R...P-AAHHPDNKDLVPE-SPTITN..RK.QERQE-QE----K--GP-*YPC............RQ.  101
O.FR.92.VAU                  -D-I--EMP--H------Q-P-NA----Q--Y--YL--TK-G---H------R..PA-AASYPDNKDPVPE-SP-NNK..RK.*ERQE-QE-E--TK-GAGGLHC............RK.  102
O.GA.11.11Gab6352            -D----EMP--H------Q-P-NT-------Y--Y---TS-G-----------R.-PAVGASRPDNK-FVPE---AITD..RK.QKRQE-QE-E--KKAGP-*QPR............HK.  103
O.SN.99.99SE_MP1299          -D----EMP--H------Q-P-NK----A--Y--Y---AS-G-----------R..PA-AARHPDN-DIVPE-L-YITN..RK.QKRQE-QE-E--N-ACPR*YPG............RK.  102
O.US.10.LTNP                 -D----EVP--H------Q-P-NT-------Y--Y---TK-G-----------R...P-TASHPD-KDPVP--SL-ITK..WK.QERQE-QE-E--KKAGPGGDHCSKGSWHCCTRTSEQ-  115

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.02.SJGddd               -----------N----------N-----R--Y--MC--TK-G---L------S-.---TPKS-KN--DLIPE--L--QQ..-X.Q--Q--Q--AL-SK--A--C-...............-  102
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.06.U14296               -----------------R----N--F--R--Y--MY--TK-G----------S-.---TLRS-KS--DLIPE--L--QQ.---.Q--QEKQE-TL-SK--A--C-...............-  102
N.CM.06.U14842               ----------------------NK----R---X-I---TK-G----------SR.---TP-S-KN--DPIPE--L--QH.---.Q--Q-KQE-A--SK-----C-...............-  102
N.FR.11.N1_FR_2011           ----Y------N----------N-----R--Y--VC--MK-G----X----S-.----PKS-KN--DIIPE--L--QH..-I.Q--Q--Q--AL-SK--A--C-...............-  102

P.CM.06.U14788               -D----D-P--QQ-----SSP-N-----A--Y--Y---TK-G-----------R.PA--DN-N-N--DPV--*-FACTG..AQ.QE*HQ-QT----KQATA-RS-C............WK.  102
P.FR.09.RBF168               -DQ---D-P--QQ-----SSP-N-----A--Y--Y---TK-G-----------R.PA--DN-N-N--DPV--*-LA*TG..IQ.QE*HQ-QT----KQATA-RS-CWTDSCLSSGTT...-  111

CPZ.CD.06.BF1167             ------D----L--S-T-S-P-NS----V--Y---L--TK----------R-G...KPSASNKNN-NPVRE-SL-KQ-..--.QVSQ--Q--E-ANQA--GGG-I............Q-  101
CPZ.CM.05.SIVcpzMT145        ---IE-D----N----K-R-E-NK-F-----Y--IL--TK-G----------G..-KPASSN-DNKDP-R---L---I..R-.Q--SEKQ--E---TS-A-QS-W...............-  102
CPZ.GA.88.GAB1               -D-I--D-----------R-V-N-----A--Y--IY--TK-G---------TTR--T-PAG-KNN-D-IP---L--S-..-N.KE-SEKST-E-ASK--A-Q*-...............-  102
CPZ.TZ.06.SIVcpzTAN13        -N-I--QVA--M---AA-E-P-------------PL--TK----------R-G...-KSAGDNK---DPVRQ-SLPKRS..RI.QSSQE-SQ-E--T-AGSGGRPRPEDSSASSGRTSGTS  115
CPZ.US.85.US_Marilyn         -D----NI----Q---------N-----------V---TK-G---------GRR--TTAKS-ADN-DP-R---F--QL..-N.Q--Q--Q--AL-KQ-TS-QR-...............-  103

GOR.CM.07.SIVgor2139_287     -D-T--E-P--QQ-----P-P-NT-------Y--Y---TK-G-----------R.PA-TADKD-NN--PV---SLAGT-..SQ.QE*YQK-E----GQ-T--RS-C............WQ.  103
GOR.CM.07.SIVgorCP2135con    -D-T--E-P--QQ-----P-P-NT-------Y--Y---TK-G-----------R.PA-TADKD-NN--PV---SLPGT-..SQ.QE*YQK-E----GQ-A--RS-C............WQ.  103
```

```
                            Rev start          exon 1 end exon 2 start     NLS                     Leu-rich effector domain                              Rev end
B.FR.83.HXB2                MAGRSGDS.DEELIRTVRLIKLLYQSNPP.PNP..EG.TRQARRNRRRRWRERQRQIHSISERILGTYLGRSAEPVPLQLPPLERLTLDCNEDCGTS.GTQ..............GVGSPQILVESPTV.LESGTKE*   116
A1.AU.03.PS1044_Day0        --------.--D-LKAI-I--X------Y.-K-..R-.-----K-------A-----D-------S-C---P----------I-------H------.---QSQ.......GTET---R--VS----GI.------N-   124
A1.CH.03.HIV_CH_BID_V3538   ------ST.-----QA--I--I------Y.-K-..R-.-----K-------A-----D-------S-C--------A---------R---S--S---.---QSQ......GAET---R--VSG--SVI.-G---ES-   124
A1.CY.08.CY236              --------.--D-L-A--I--I------Y.-K-..R-.S---K-------A--K--D-----V-SAC----E---------I---H---C------.---QSQ......RVET---R--VSG-P-GI.------N-   124
A1.ES.05.X1608_8            --------.----L-A--I--I------Y.-KR..R-.S---Q-------A-----D-L-----S-C---P-----------G-H---H------.---QSQ......GVET---R--VX----G-.-G---ET-   124
A1.KE.11.DEMA11KE001        --------.-A--LTA--I--I-------.-K-..R-.-----K-------A-----D-------S-C----T-------------H---S------.---QSQ......GVET---R--VP----A-.-G----N-   124
A1.RU.11.11RU6950           -------T.-A--LTA--I--I------Y.-T-..R-.-----K-Q-----A-----D-L-----SSC------S-SF----I---H---S------.---QSQ......GTET---G---SG-PSGI.-G------   124
A1.RW.11.DEMA111RW002       ------SG.----L-AI-I--I------Y.-K-..R-.----Q--------A-----D---Q---S-H--G-E---HF----------N-S------.---QSQ......GAAT---RS-VSG--D-I.-D---EN-   124
A1.SN.01.DDI579             ------N-.----L-AI-I--I------Y.-K-..Q-.-----K-------A-----D-L-----S-----E-------------H---S------.---..................TGS*-S--CSAI.-GE---KQ   112
A1.UG.11.DEMA110UG001       --------.-V--LTA--I--I------Y.-K-..K-.S---K-------A--K--D-------S-C-------------------H---S--G---.---QSQ......GTET---G--VS---SII.-G----K-   124
A1.ZA.04.04ZASK162B1        -------D.V-D-LQA--I--I--K---Y.-K-..G-.-----K-------A-----D-------S-C---P--------------H---S---..............GTET--EGHS------A-.--P-----   117
A2.CM.01.01CM_1445MV        -------P.--D-LKA--T--I------Y.-T-..R-.S---K-Q-----A---H-S-------S-C-----------------HI--S---D--.---QSQ......GTET---GH--S---SVI.-G---E--   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-------.--LQSQ......GTET---RS-ES---SVI.-G---E--   124

B.BR.10.10BR_MG029          --------.----LKA-----T------L.-S-..K-.N-----------R-----QD--GWV-CNL---------H---I---------------------.---..................-----------X-AI.----A---   117
B.CA.07.502_1191_03         ----R---.--D-LKA-----R--------.-ST..---.--------------------H-RLA-DW--R---D-PT---D-P----------SS------.---..................-E-------------.-GT--E--   117
B.CH.08.M2_0803101_NFLG8    -------R.--D--Q---------------.-S-..---.-----------------Y--Q--G-L-S-----P-------------------------.---PE..................-------F----A-.--------   117
B.CN.10.DEMB10CN002         --------.--D-LK------T---------.---..---.------------------R----Q--H------V--------EI--------K---N-.---..................---------G--D--.-D--AA-E   117
B.ES.10.DEMB10ES002         --------.----LK------F--------.-S-..---.--------------Q--K--QTL-AW--SNH---------------------SK------.---..................---G--V-----A-.--------   117
B.FR.11.DEMB11FR001         --------.E---LK------R--------.-S-..---.----------------K--H-RML-GW--DN----PE------------------S-----A.---..................---G--V-----A-.----A---   117
B.GB.05.MM45d213_GN1        --------.--D-LKA-----I---------.---..---.-------------------RG-GQ---D-----XX-------------------------.---..................----H-V-----A-.-D-----E   117
B.HK.06.HK003               --------.----LQK--I--F--------.-SA..---.------------KQ-----Q-L-GW-ISNH---P-------------------DQE----.---..................---N----L-P---.----AE-E   117
B.HT.05.05HT_129389         ----R---.----LK------Q-----L.-S-..---.--K-----------------K-G----A-----E------------S---S-G--D-.---..................-------V-----A-.--------   117
B.JP.12.DEMB12JP001         -------N.--D-LKA-----F------L.-SS..---.-------------------K-------A-----PE------------------D---N-.---..................---N--------AI.-------E   117
B.KR.07.HP_18_07JHS10_3909  -------N.----L--I-I--F------L.-K-..---.-------------KK--W-R----W--N-H---P-----H-------------S-E----.---..................------E------A-.-------K   117
B.PE.07.502_0525_wg5        --------.----L------R---------.---..---.-------------QK--H-RT--DWL-T-----P-------------------K-S---.---..................---N-------------.-----E--   117
B.RU.11.11RU21n             --------.--D-LK-I-------------.---..---.------------------H-R------IC-----PE-----P------------S-------.---..................--------------A-.-D------   117
B.TH.08.MERLBDTRC10         --------.--D-LK-------------L.-S-..---.--R-----------------RD------D-----PP---------I-------S-------.---..................---G--V---P-A-.-G------   117
B.US.11.ES38                --------.----LKA--I--F--------.-S-..A-.-----------------K-R----W--SNS---P-----F----I------SS----I-.---..................---N--V-----A-.--------   117

C.AR.01.ARG4006             --------.--A-LLA--I--T------Y.-K-..---.--R----------A-------------I-C---PT---------I---NIN-S-SG---.---..................R--N-*-FRKPCA-.-G-RA-KE   116
C.BR.07.DEMC07BR003         --------.--A-LQA-KI--I------Y.-K-..---.------------A-------------T-C---P-----------I---SIN-S-SG---.---QSQ......GTTE---N-*V-GKPCA-.-G-RA-KE   123
C.BW.00.00BW5031_1          -------N.--A-LQA--I--I------Y.-K-..---.-----K-------A--------G----S-C---------------------HI--S-SS--P.---QPQ......GDTE---R-*VSGK-CA-.-G----KE   123
C.CN.10.YNFL19              --------.--A-LKA--I--I------Y.-E-..R-.-----K-------A--K--R-L-----SSC---PT----F--------HIGDS-SG-A-.---QSQ......GTTD--E--*-SGR-CA-.-G--A-KE   123
C.CY.09.CY260               --------.-AA-LE--KI--I------Y.-K-..---.------------R-ATKKRLN*F-H-L-SSGCVGPP-QL-FF-----E-L-IDS--SRS-C--HPPG......GTAE-G-HHVFCE*CCV-.----RA--E   123
C.ES.08.X2363_2             -------N.--A-LLA--A--I------Y.-K-..---.------------A--K---Q------S-C---------------I---YI--S--S-I-.---..................----S*VPGK-CA-.-G----KE   116
C.ET.02.02ET_288            --------.-AA-LQA--A--I------Y.-K-..---.----Q--------A-------L-----SDC---HP--------------N---S-SG---.---QPQ......GTTE---N-*VSGQYCA-.-G-RA-KE   123
C.IN.09.T125_2139           --------.--A-LKA-KI--I------Y.-E-..R-.-----K-------A-------L---V-SAC---P-----------I---HISDS-SG---.--PQSQ......GTTE-----*-SGKPCA-.-G--A--G   123
C.KE.00.KER2010             -------N.-D--L-V--I--I------Y.-K-..---.-----K-------A--R-----------SAC---P-----F----I---HI--S-GG---.---QSQ......GTTE-----*VSGKLCA-.-G--A-KE   123
C.MW.09.703010256_CH256.w96 --------.--A-LQA--I--I------Y.-K-..---.-----K-------A--K-----G----SS----P-----------I---HIGDS-SG---.---QLQ......GTTE-----*VSGKPCA-.-G-RA-KE   123
C.TZ.08.707010457_CH457.w8  ------E-.--A-LQA--I--I------Y.-K-..---.-----K-------A-------------S-C---------------I---HI--S-GG---.---QPQ......GTTE---N-*VSGK-CA-.-G-R--KE   123
C.YE.02.02YE511             -------N.-AA-LQ--KI--I------Y.-KL..---.------------A--K-----G----SCC-E-P-----------I---HI--S-SS--T.---QSQ......GNTE----S*VSRK-CA-.-GARA-KE   123
C.ZA.08.705010534_CH534.w12 -------N.--A-LQA--I--I------Y.-K-..---.-----K-------A-------------V-C---P-----F----I---HI--S-SS---.---QHQ......GTAE---N-*VSGK-CAI.-D-RA-KE   123
C.ZA.10.DEMC10ZA001         --------.--V-LQA--I--I------Q.-K-..---.------------A-----NQA-----ASC--------------------HI--S-GG---.-E-QSQ......GTTE-----*-SGK-CA-.-G--A--E   123
C.ZM.11.DEMC11ZM006         --------.--A-LQA--I--I------H.-KL..---.-----K-------A-----TL------S-C---P-----F---------HI--S-SS---.---QSQ......GTTE-------SGK-Y-I.-G----KE   124

D.CM.10.DEMD10CM009         ----RE--.--N-LK-IQ---F---I---.-ST..---.------------A-------G----S---------------I---N-N-S---EP-.---..................----T--S----A-.-----E--   117
D.CY.06.CY163               ----RE--.--D-LKAI----V--------.-SI..QA.P----TH------A-------G----S-----P-------------N-N-S------.---..................----T--S----A-.--T--E--   117
D.KE.11.DEMD11KE003         --------.----LKA--F--I--------.-S-..A-.------------A-------LG--F-S-----PE----------------------------.---..................----H-------S-.-D------   117
D.KR.04.04KBH8              --------.--D-LKAI----F------L.-S-..---.------------A----N--G---IS--------------I----IN----------.---..................----S-VSG--SAI.-DT-----   117
D.SN.90.SE365               --------.----LKAI----F---N---.-ST..---.------------A----N--G---S------------------N-N-S-------.---..................RM----VS----A-.-DT--EK-   117
D.TZ.01.A280                --------.----L-AI----I--------.-SL..---.------------A----D--G---S------------------N-N-HT------.---..................----H-VS--P-AI.-D---E--   117
D.UG.10.DEMD10UG004         --------.-DD-LKA-----I--------.-S-..---.------------A----N--AG--ID-----PE--D-----------NIN-I--L---.---..................----Q--P--P-A-.-D--S---   117
D.UG.11.DEMD11UG003         --------.----LKA-----F--------.-SS..---.------------A----RT-G----S-----P-----F----------IN-S-N----.-A-..................---GH--S----P-.-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.AO.06.AO_06_ANG32        -------N.----L-A--I--I------Y.-K-..---.------------A----VRE------ASC---P-------------HIN-S---EQG.AEE..............R------S---RAI.--------   117
F1.AR.02.ARE933             --------.-T--LTAL-Y--I------Y.-K-..---.------------A----RA--D---SSC---PE-------------HIN-S----PG.TEE...............-L-G---AG--HA-.----I-N-   117
F1.BR.10.10BR_RJ015         --------.-TDIV-I-NY--I--------.-K-..---.------------A----RAL-----RFG---PE-------------HIN-S----QG.TKE..............EL-----SG-HH--.--Q-----   117
F1.CY.08.CY222              --------.-T--L-A--T--I------Y.-K-..---.------------A----RA----L-LSC---PE--------S----HIN-SK--EQG.TEE..................---N---S---H--.-D------   117
F1.ES.02.ES_X845_4          --------.--D-L-AI-T--I------Y.-K-..---.------------A---VRE------SSC----P-------------HIN-S--GEQG.TEE..............R------S---RA-.--------   117
F1.RO.96.BCI_R07            --------.-T--LKA-QY--I------Y.-K-..---.------------T----RE------SSC---P-------------HIN-S---EQG.AER..................-------S---HA-.--------   117
F1.RU.08.D88_845            --------.-A--LKA-KY--I--------.---..---.------------A----RE--Q---DDC---P-------------HIN-S-G--QG.TEK..................----S--P---S-I.-G------   117
F2.CM.02.02CM_0016BBY       --------.--A-LTA--T--I------Y.-K-..---.------------A----RE------S-C-------------------HI--S--S-QG.TER..................-------S---RS-.-G---EN-   117
F2.CM.10.DEMF210CM001       --------.----LKA--Y--I------Y.-K-..---.------------A----RE------IACV--P-D-----------HIN-S---DQG.AAT..................---N---P--PHAA.-G-----E   117
F2.CM.10.DEMF210CM007       --------.--D-LKA--Y--I------Y.-K-..---.------------A---VRE------RAC---PE-------------HIN-S---EQG.TGK..................--E----S--PRAA.--------   117

G.BE.96.DRCBL               ------ST.----LTA--I---------S.-P-..---.------------A-------------S-C---PE------------H---S--G---.---QSQ......GTEI-------F---SV-.-G------   124
G.CM.10.DEMG10CM008         ------ST.--A-LTA--I--I------H.-SA..---.-----K-------A----N-------SSC---P-----V-------HI--S--S-NF.---QSP......GTEP.......G---SV-.-G------   117
G.CN.08.GX_2084_08          ------ST.--Q-LQAI-I--I------Y.-P-..---.---T-K-------A----R-------S-CV--LE-----P---I---H---S--SR-P.E--QSP......GTET---G---S----V-.-G------   124
G.CU.99.Cu74                ------ST.----L-A--I--I------Y.-P-..---.-----K-------A--K--------V-S-C----X--------------S---SK-S---.---QSQ......GTET---R---S----G-.-G------   124
G.ES.09.X2634_2             ------NT.----L-IIQT--I------Y.-PS..---.-----K-------A-----------L-S-CV---T-------------HI-----S--L.---QSQ......GTET---G---P---SV-.-G------   124
G.GH.03.03GH175G            ------SN.----L-A--I--I------Y.-SA..---.-----K-------A---VRAL---F-SAC---PT-------------C---S--S------.---QSQ......GTET---R---S----V-.-G------   124
G.KE.09.DEMG09KE001         ------ST.-G--LQA--I--I------Y.-P-..---.-----K-------A-------------SAC---P------P-------HIA-S--S--.....................GTETEL-G---S----G-.-GA-A---   117
G.NG.09.09NG_SC62           ------TT.----LQA-KI--I------S.-P-..---.-----K-------A------------IS-C---CT-------------H---S--SR-P.---QSP......GTET---G---S----V-.-G-R----   124
G.ZA.01.TV546               ------S-.----L----I-XI------Y.-P-..---.S---Q--------A-----RA----V-SSCV--LE----F---------IN-S--I------.---QSQ......GTET---G---SE--SGI.-GT-----   124

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.CF.90.056                 ------A-.-T--LQVCKI--I------C.-E-..T-.------------A-----RE------TSC---PP---T-----------N-S------.-EK...............-E-----SL--S-I.-GT-----   117
H.GB.00.00GBAC4001          --------.--G-LQVCKI--I------Y.-ER..AE.------------A--------------H-C-AGPV-----------------S------.-EK..................-------S--P-A-.-GT-A---   117

J.CM.04.04CMU11421          --------.--Q-LLA-XX--I--R---Y.-S-..R-.S-----------A--N-XD-------C-C---------------I---HI--S----D-.---..................---D---SG-PCM-.-GA-----   117
J.SE.93.SE9280_7887         --------.-DQ-LLA-----I------Y.-K-..N-.S-----------A--N--D------PSSC---P----------I---R---S----N-.---..................---D---SG-PCM-.-GA-----   117
J.SE.94.SE9173_7022         -------N.-DQ-LLA--I--I------Y.SK-..N-.S-----------A--N--D-------SSC---P----------I---R---S----N-.---..................---D---SG-PCM-.-GA-----   117

K.CD.97.97ZR_EQTB11         ----R---.EQQ-LTP--I--I------Y.-K-..---.------------A-----RE--Q-V-SSC----T-------------S-N-D--S-QG.TEG..............EL-----P--PD--.-G--D---   117
K.CM.96.96CM_MP535          ----R--P.--Q-LT---T--I--E---Y.--L..---.S--T--------A--K--S-----L-SAC-------------I-K-N-N-D--P-KG.TEG...............-L-----S--PC--.--------   117

U.CA.01.TV749               -------N.--R-LSA--I--I--------.-K-..G-.S---K-------A--N--D-------SAC-E-----I----------S---R--------.---QPQ......GTEAR--R---SG--SV-.-G---ED-   124
U.CA.99.TV721               -------R.--Q-LLA--T--I------Y.-K-..R-.-----K-------A--N--D---R---S-C----E---------------S---S--------.---QPQ......GTEAR--R---S---SGI.-G---ED-   124
U.CD.83.83CD003_Z3          ------E-.----L-A--I--I-------.---..---.-----K-------R--Q--R-------S-C---P----H---------N---SK..................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.CY.05.CY090               --------.-A--L-VI-I--I------Y.-T-..---.------------A-----RA------SSC----T-----P-----N-S-N-SK--E........Q......GTTGEL-NS--SG--RV-.-G------   117
U.CY.08.CY223               --------.--Q-LAA--I--II-----Y.-K-..G-.S---K-------A--N--D---R---SDC----E----F----I---HIN-S--S-N-.---..................---T--PG--GA-.--PR-E--   117
U.ES.10.DEURF10DZ001        -------A.-QD-L-A--I--I------Y.-SS..---.-----K-------A----VN-------SAC---P----------I-K-S---S-S------.---QSQ......GTET-----*-SL-YSVI.-G------   123
U.GR.99.99GR303             --------.--Q-LT-I-I--I------Y.-K-..---.------------A-------------L-C----E--------SI---H---S----N-GD-G..............#-AD---SGGTFLL#MG---A-E   117
U.NL.95.U_NL_95_H10986_D1   ----R--Q.--Q-LT---I--I------F.-KL..---.---------K--A-------------S-C-------------------S--SREG.TEG..............EL-----P--PCA-.--------   117

01_AE.AF.07.569M            ------ST.----L-A--I--I------F.-SS..---.------------A-----RA------S-C----T-------------H------------.---QSQ......GTET---RS--SG-PSVI.-GP---N-   124
01_AE.CF.90.90CF11697       ------NT.--D-LQA--I--I------Y.-P-..---.-----K-------R-------L-----VACV---T-------------HI--S---------.---QSQ......GTET---G---SG--SVI.-G----N-   124
01_AE.CN.10.YNFL03          ------ST.----L-A--I--I------Y.-SS..--T---T-K-------A-----RE-----R--------------------H---S---------.---QSQ......GTET---R---SG--SVI.-GP-AEN-   125
01_AE.HK.04.HK001           ------ST.----LGA--A--I------S.-PA..---.---T-K-------A-----RK------SSVV---T-------S---H--------------.---QSQ......GTET---R---SG--SVI.-GA---N-   124
01_AE.IR.10.10IR.THR48F     ------ST.-D--L-AI-I--I------Y.-PA..---.---T-K-------A---VRAL-D---SDC----T---------I---S---S--------.---QSQ......GTET-L-R---SG-PSGI.--P--EN-   124
01_AE.JP.x.JRC77AE          ------ST.----L-A--I--I------Y.-SS..---.---T---------A-----RA------SSCV---T------------------S--------.---QPQ......GTET---R---PGK--DL.-GP---N-   124
01_AE.TH.09.AA090a_WG11     ------NT.------AI-I--I------F.-SA..---.------------A------TL-----S-C---P-------------H--SS--S---.---..................EA---R---SG--SV-.-GP--EN-   117
01_AE.TH.90.CM240           ------ST.----L-A--I--I------F.-SS..---.---T---------A-----SA------S-C----T---------V---N---S--G---.---QSQ......GTET---R---SG--SVI.-GP---N-   124
01_AE.US.05.306163_FL       ------ST.----L-A--I--I------Y.-SS..---.-----K-------A----VRA------SACV---T---------I---H---S--------.---QSQ......GTET---R---SG-PSSI.-GP--ED-   124
01_AE.VN.98.98VNNDI5        ------ST.----L-A--I-RI--E---Y.-SS..---.---T-K----KV-A-----RE------S-C---PT-------------H---S--------.---QSQ......GTET---R---SG--SVI.--P---N-   124
```

```
                           Rev start          exon 1 end exon 2 start      NLS                             Leu-rich effector domain                                        Rev end
B.FR.83.HXB2               MAGRSGDS.DEELIRTVRLIKLLYQSNPP.PNP..EG.TRQARRNRRRRWRERQRQIHSISERILGTYLGRSAEPVPLQLPPLERLTLDCNEDCGTS.GTQ.............GVGSPQILVESPTV.LESGTKE*   116
02_AG.CM.10.DE00210CM013   -------T.----L---EI--I------Y.-P-..G-.-----K-------A--K--RT------SAF---P---HF----I---C-----------.---QSQ......GTET-------S---SVI.-G------   124
02_AG.CY.09.CY256          ---N---A.----L-A--I--I------Y.-P-..A-.-----K-------A--K--RAL-----N-C----E----F----I---C--SS---------.---QSQ......GTETR------SG--SII.-G--A---   124
02_AG.ES.06.P1423          -------P.-A--L-V--I--I------Y.-P-..--.-----K-------A-----SE------SSC---PT-----L---I---S---C-GS-D-.---QSQ......GTGT-------SG--SII.-G------   124
02_AG.GW.05.CC_0048        -------NS----L-A--T--I------L.-PS..--.S----K-------A-----GA------RAC-----------------S-N----S---.---QSQ......GTKT-------S---SVI.-G------   125
02_AG.KR.12.12MHR9         -------A.----L-VI-I--I------Y.-P-..--.-----K-------A-----SA------S-----P----------I---S---SK------.--RQSQ......GPETR--G--TP---SSI.-G------   124
02_AG.LR.x.POC44951        ------NA.--G-LGVI-IVRI------Y.-P-..--.-----K-------A-----CA-I--V-S-C----E----------I---C---S------.---QSQ......GTET-------S---SII.-G------   124
02_AG.NG.09.09NG_SC61      -------A.----L-A--I--I------C.-P-..--.-------------A--K--RA------S-C---------------I---S---SK------.---LSQ......GAGT-------S---SII.--------   124
02_AG.NG.x.IBNG            -------A.----L-A--I--I------Y.-P-..--.---T-K-------A-----RA------S-C---------------I---N---S------.---LSQ......GTET-------S---YII.-G------   124
02_AG.SE.94.SE7812         ------NA.----L-A--T--I------Y.-P-..--.-----K-------A-----RA------S-C---------------I---R---S---------.................--------S---SI-.-G------   117
02_AG.SN.98.98SE_MP1211    -------A.--G-L-A--I-RI------Y.-P-..--.S------------A----VRA------S-C--P-----P--I---C---S--S------.---QSQ......GTET----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.94.94CY032_3     ------NI.--D-FKAA-A--I------YPN--..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
07_BC.CN.98.98CN009        -----E--.--A-LKA--T--I------Y.-E-..R-.-----K-------A--------------SAC---------------I---HISGS-SG---.---QSQ......GTTE-----*-SGK-CA-.-G--A-KE   123
08_BC.CN.97.97CNGX_6F      --------.--A-L-A--I--I------Y.-E-..R-.-----K-------A--N-L---------SAC---PT-------------HITDS-RG---.---QPQ......GTTE-----*-SGK-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.CM.95.95CM_1816     -------N.-DQ-LTA--I--II-----Y.-K-..Q-.S--S---------A--N--D---Q---SSC---P--------------N-----S-N-.---...............---DX--PG--Y--.--PR-E--   117
12_BF.AR.99.ARMA159        --------.----LKAA----------Y.-K-..--.-------------A--S--RA----ID-----PE----F-----------G---------.................................---------   117
13_cpx.CM.96.96CM_1849     --------.--V-LT---I--I---N--Y.-SQ..--.-----K-------A---------GQ---S-C-----------------H-----------.---QSQ......GTET---G---P----GI.-G----N-   124
14_BG.ES.05.X1870          ------ST.--D-L-AX-I---------Y.-P-..--.-------------A---------------SACV---E-------------H-G-S--S--V.---QSQ......GTET---GS--S---SVI.-G------   124
15_01B.TH.99.99TH_MU2079   ------ST.----L-A--I--I-H----Y.-SA..--.S--T---------A--KR-RE------SPC-------------------S-----------.---QSQ......GTET---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---------.---QPQ......GTET---R---S--PSV-.-G---E--   124
17_BF.AR.99.ARMA038        --------.--D-LKAI----F-------.-K-..--.-------------A-----NK-G----A-----P-----L----------SK------.---..............-E----V---PS-I.--------   117
18_cpx.CU.99.CU76          --------.--Q-LT--SI--I------Y.-E-..A-.-------------A-----SA------N-C--GP-----------------S-----F.---QSQ......GTET---GS-VP---SG-.-GT-D---   124
19_cpx.CU.99.CU7           --------.--D-L-VI-I--I------Y.-T-..QR.A---K-------A---------------S-C----E----F--------H-N-S---------..............R--G---S---SV-.-G--A-N-   116
20_BG.CU.99.Cu103          ------ST.--D-L-VA-I--I------Y.-P-..--.-------------A-----RAL-----T-C---PP-----------------S--S------.---QSQ......GTET---R--VS----G-.-G------   124
21_A2D.KE.99.KER2003       ----R---.-DD-LKAI----F-------.-S-..--.-------------A-------A---VSYH---E---------I---N-N-R----A-.---.............-G-R---S----AI.-G---E--   117
22_01A1.CM.01.01CM_0001BBY ------ST.----L-A--I--I------Y.-K-..R-.-----K-------A--K--D-L-----SAC-------------------H---S-----------.---QSQ......GTET---R---S----DI.-GA---K-   124
23_BG.CU.03.CB118          ------ST.----L----I--F------Y.-P-..--.-------------A---------Q---ADC---PP-----------------S--I----.---QSQ......GTETR--G---S----A-.-G------   124
24_BG.ES.08.X2456_2        ------ST.--D-L-A--I--I------Y.-P-..--.-------------A---------Q---NSC---P------------------S-...............GTET---GS-VS---SV-.-G------   113
25_cpx.CM.02.1918LE        ------VN.----LQAA-I--I------S.-P-..--.-------------A------TL-----ASC---LE-------------H---S--S------.---LSQ......GTET---R--VS----A-.-GT-----   124
26_AU.CD.02.02CD_MBTB047   -------P.--A-LGAI-I--I-------L.-K-..R-.S-----------A--NK-D-------S-C--GPP---------------S---S--F-N-.---QPQ......GTET---R---S----GL.-GA--E--   124
27_cpx.FR.04.04CD_FR_KZS   ----NE--.--Q-LTAI-I--I--K---Y.-K-..R-.-----K-------A--E--D--K-V-S-C---P------P-------C---S---------.---QPQ......GAET---R---S---SA-.-GT-A---   124
28_BF.BR.99.BREPMI2609     --------.--D--K-I-I--F-------.-S-..--.----------------------L-RW--SNH---PE----F---I---------------...............---------V-----A-.-G------   117
29_BF.BR.01.BREPM16704     --------.-----K------F-------.-SS..--.-----------------------RW--SNH---PT---------I---------S-------...............-------------AI.--------   117
31_BC.BR.04.04BR142        --------.--A-LQA--I--I------Y.--..--.-------------A-----RA------S-C---P-----FL---I---NIN-S-SG-----.---HPQ......GTAE---N-*V-KKPCA-.-G-RA-KE   123
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---PEKPGVL.-GT-----   117
33_01B.ID.07.JKT189_C      ------ST.-DD-L-A--I--I-X----Y.-SA..--.---T-K-------A-----GA------SAC---P--S-----------S---S---------.---QSQ......GTET---R---SG--SVI.-D----N-   124
34_01B.TH.99.OUR2478P      ------ST.----L-A--T--I------Y.-SS..--.-----K-------A--K--RE--D---ISC----T-------------------SK------.---..............-------V-----A-.-D------   117
35_AD.AF.07.169H           --------.----LKGI-I--I------Y.-E-..R-.S----K-------A--N--D---Q---S-C---P--------------Q---Q---------.---QSQ......GVET---R--VF----V-.-G----N-   124
36_cpx.CM.00.00CMNYU830    -------A.----L-V--T--I------Y.-PS..--R-----K-Q-----A-----RA------S-C---FE-------------H---S---------.---QPQ......GTEA---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-.---QPQ......GTET-L-G---S--P-G-.-G----N-   124
38_BF1.UY.03.UY03_3389     ------G-.----LKAI-Y--I------F.-K-..--.-------------A-----RAL-----SSC---PE-------------HIN-S----QG.TE-..............---N-PV---P-AI.-D------   117
39_BF.BR.04.04BRRJ179      --------.----LK---Y--I------Y.-K-..T-.-------------A-----RE----F-SAC---RE-------------HIS-S---EQG.AEE...............------TSG--LA-.--------   117
40_BF.BR.05.05BRRJ055      --------.--D-L-A--A--V--#----.-SS..--.-----------------RK---W-ISNH---PT----F-------------S------.---...............------V-----E-.--------   116
42_BF.LU.06.luBF_18_06     --------.--T-LQ------R--------.-ST..--.-----------------QTF-GW--N-----PE-----------------S-------.---...............-----------A-.--T-----   117
43_02G.SA.03.J11223        -------A.----L-X--I--I------Y.-S-..AR.-------------A---------------SAC---R--------------H---S--SR-P.---QSP......GTET---G---P----V-.--------   124
44_BF.CL.00.CH80           --------.-VD-LKVI-Y--I------Y.-K-..--.-------------A-----RAL-----SSC---LE-------------HIN-S----QG.TEG...............-------SG---A-.-G------   117
45_cpx.FR.04.04FR_AUK      ------NT.----LTA--I--I------Y.-K-..--.-------------A---------------S-C---P--------------H---S----D-.---...............--AD--TSG--CAI.-G------   117
46_BF.BR.07.07BR_FPS625    --------.--D-LKA--Y--I------Y.-K-..R-.----Q--------A----VRG------RSC---PE-------------HIN-S----QG.TEE...............------TSG--RV-.-G------   117
47_BF.ES.08.P1942          --------.--X-LKA-----I-------.-SS..--.-----------------R--G----N---D-PT-----------------S-------.---...............---T------P---.--------   117
48_01B.MY.07.07MYKT021     ------ST.----LTA--I--I------S.-SA..--.---T-K-------A----VRA------S-C---P----------------S-----P.---QSH......GTET---R---SG--SDI.-GA---N-   124
49_cpx.GM.03.N26677        -------N.--Q-LLA--I--I--K---Y.-K-..N-.S------------A--D--D---Q---IDC---P---------S----C--SS---N-.---...............---D--VSG--CPI.-GE-----   117
50_A1D.GB.10.12792         ------S-.--D-LKAI----F-------.---..--.---K---------A-----N--G----D-----PE--G-----------N-N-S---N-.---...............----N-V-----AI.-D------   117
51_01B.SG.11.11SG_HM021    --------.-----K---------------.---..--.S-----------------H-RVV-DW--S-----T-----------------S-G----.---...............------V-----P-.--------   117
52_01B.MY.03.03MYKL018_1   ------ST.----L-A--I--I------F.-SS..--.-----K-------A-----RA------C-CV--PT----P-----------SK------.---..............-G---P------A-.-D------   117
53_01B.MY.11.11FIR164      -------T.----L-A--I--I------Y.-SX..--T---T-K-------A----XRA------SXCV---T---------I---N---S-G-----.---QSQ......GTET---R---SG-PSVI.-G---XN-   125
54_01B.MY.09.09MYSB023     ------NT.--D-L-A--I--I------Y.-P-..--.-------------A----VRE------RD---G-T---------I---H---S---------.---QSQ......GTET-L-G---SG--SG-.-G--X-N-   124
55_01B.CN.10.HNCS102056    ------ST.---IL-A--I--I------F.-SS..--.---T-K-------A-----RA------AACV---T-----P--------HI--SK--E--.....T......GTET---R---SG--SVI.-G---EN-   120
56_cpx.FR.10.URF5_patient_A ------VT.----LQA--X--I------Y.-PS..T-.-----K-------A-----RE------S-C----T---H-P---I---C-----GG---.---LSQ......GTETR-----TS----IL.-G-----E   124
57_BC.CN.09.09YNLX19sg     --------.--A-L-A--I--I------C.-E-..R-.-----K-------A--N--D---A---S-C---P-----F------G-HISDS-SG-A-.--HQPQ......GTTE-----*-SGKYCA-.----E--KE   123
58_01B.MY.09.09MYPR37      ------ST.----L-A--I--I------Y.-P-..--.---T---------A-----RAL-----S-C---P--------------H-N-S---------.--HQSQ......GTET---R---SG--SVI.-GQ---N-   124
59_01B.CN.09.09LNA423      ------ST.----LKA--I--I------L.-SS..--.-------------A------TL---V-S-CV--PT-------------H-GDS-----P.---QSQ......GTET---R---SG--SGI.--P---N-   124
60_BC.IT.11.BAV499         --------.--A-LQAI-I--I------Y.-R-..--.---T*---K--*-A---------------S-C---P----------I---NINHS-SS---.R--QPQ......RTAEK--N-*VFRKPCA-.--------   121
61_BC.CN.10.JL100010       --------.----L-A--I--V------Y.-E-..R-.----QK-------A--N--R--------SSC---------------I-G-HISDS-SG---.---QSQ......GTTE----S*-SGKPCA-.-G--A-KE   123
62_BC.CN.10.YNFL13         --------.----L-A--I--I------Y.-KS..T-.-------------A-------L-----SAC---P----------I---HISGS-SGE--.---QSQ......GTTE-----*-SGKHCA-.-G----KE   123
63_02A1.RU.10.10RU6637     -------V.----L-V--T--I------Y.-P-..--.-------------A-----SA------SSF--------------I---S---S---AA-.---QSQ......GTAT-------S---SSL.-G------   124
64_BC.CN.09.YNFL31         --------.----L-V--I--I------Y.-K-..R-.-----K-------A--E--N-------R-C---PE---S---------HISDS-SG--P.---QSQ......GTTD--E--*-SGK-CA-.-G--A--E   123
65_cpx.CN.10.YNFL01        --------.----L-A--T--I------F.-SS..--T---T---------A------Q--D---RAC---PE----------I---YISD--SG---.---QSQ......GAET---R---XG-PSV-.-GP---N-   125
67_01B.CN.11.2011.ANHUI.HF115------NT.--D-LAA--I--I------X.-SA..--.-------------A-----RA------XPF----E-------------H-----------.---QSQ......GTET---R---SG--SVI.-G---EN-   124
68_01B.CN.11.2011.ANHUI.WH73-------NT.--D-LAA--A--I------F.-SS..--.-------------A-----KE------A-X----T---------D--S-----------.---QSQ......GTET---G---SGK-SVI.-G---EN-   124
70_BF1.BR.10.10BR_PE004    --------.----LTA--I--F------L.-S-..--.-----------------RT--DW--SNH---PT----F---I--------S------.---...............----------------.--P-----   117
71_BF1.BR.10.10BR_PE008    --------.----L-A-------------.-SS..--.--R----------KH-WA--GW--SNH-------E--H--Q--------S------.---...............-----------S-I.--------   117
72_BF1.BR.10.10BR_MG002    --------.E-N-LK-------------.-SS..--.-----------------RN--GW--N---D-PE-----------------------.---...............---D---R--HT--.--------   117
O.BE.87.ANT70              -----E-D...Q-LQAIQI--I------Q.-S-..R-.S-N--K-------R--A-VDTLAA-V-A-VVHGPQNNNIVD-----Q-SIRDP-GDQL-.EAWTVD......PRAEDN*CL-N-CSCN-I.-ATRIA--   121
O.CM.98.98CMA104           -----D-D...Q-LQAI-I--I------H.-S-..I-.S-S----------R--A-VDT-AA---A-VVHGPQNNNLVD-----Q-SIRDP-SDQP-.-NWTVD......PRAEDN*WL-T-*SCN-I.-ATRIA--   120
O.CM.98.98CMABB141         -----D-D.VQP-LQA--I--I------H.-S-..I-.S-N--K-------R--A--D--AA-V-A-VVHGPEDNNLVE-----Q-SIRDP-GDQP-.E-WTVD......SRAEDN*CL-N-CSCN-I.-ATRIA--   123
O.CM.98.98CMABB212         -----EED..QQ-LQAIQI--I------H.-TQ..A-.S-S--K-------R--T-VDA-AS---A-VVHG-QDNNLVD-----Q-SIRDP-ANQLP.E-WTVD......PGTKDN*-L-T-WSCN-I.-ATRIA--   122
O.CM.98.98CMU5337          -----T-E...--L-VI-I-NI------H.-S-..T-.S-N----------R--A-VNTLAS---A-VVHGPQDNNLVE-----Q-SIRDPDGDQP-.--WTED......PRAEDN*CL-N-YSCNSI.-ATRIA--   121
O.CM.99.99CMU4122          -----D-V...P-LH-I-I--I------Q.-S-..I-.S-S----------R--A--DTLAA-F-ASVVHGPQDNNLVE-----Q-SIRDPDGDQP-.--GTVD......PRTEDN*CL-N-CSCS-I.-ATRIA--   121
O.FR.92.VAU                -----EED..QQ-LQAIQI--I------H.-TT..S-.S-N----------R--A-VD--AA-V-ASVVHGPQDNNLVD-----Q-S-RDP-VDQ-P.--WTVD......HRAEDN*-L-T-*SYN-I.-ATRIAN-   121
O.GA.11.11Gab6352          -----DAD.-QQ-VQA--I---------Q.-S-..I-.S-N--K-------R--V--D-LAT---A-VVHGPQDNSCVE--S--H-SIGDS-GDQP-.E-WTVD......PRREDN*CL-N-YSCS-I.-VTRIAK-   123
O.SN.99.99SE_MP1299        -----DGD..QP-L-AIQI--I-----SH.TS-..T-.S-S----------T-HAHVDTLAA---A-VVHGPQDNNLVE-----Q-SIRDPDGDQP-.--WTVD......SGTEDN*CL-T-HSCN-I.-ATRVA--   122
O.US.10.LTNP               -----D-D...R-LQAIQI--I------Y.-S-..S-.S-N----------R--A-VET-AA-V---VVHGPQNNNIVE-----Q-SIRDS-GDQP-.E-WTVD......LGTETN*CL-N-YSCN-I.-ATRIA--   121

N.CM.02.DJ00131            ------VN.----L-A--I--I------Y.--NNNQ-.-------------T-----RE--Q---D-C--GLQ---D-P----D-----PA--S-AP.--EPHQ......GTATTE*TQNT--GNNSI.-GKRV-G-   125
N.CM.02.SJGddd             ------VN.----L-A--I-RI------Y.--N..K-.X------------A-----RA------SSC--GPP---D-P----D-----PK--S--P.E-ESQQ......GTATTE*THNT--GNTSI.-GRRX-R-   123
N.CM.04.04CM_1015_04       ------VN.--K-LTA--I--I------Y.--N..Q-.-------------A-----CA------S-C--GLQ--GDPP----------NT--S-DP.--EPQQ......GTATIE*TQKT--GNTSI.-GKRV-R-   123
N.CM.06.U14296             ------VN.----LGA--A--I------Y.--N..K-.-------------A-----RA------L----GPQ--GD-P----D-----PE--S----.--ESQQ......GTATTE*TQNT--GNNSI.-GKRV-R-   123
N.CM.06.U14842             ------VD.--G-L-A--I-RI------Y.--N..T-.-------------A-----RA------S-C--GPQ---D-P----D-----ST--S----.--ESQQ......GTATTE*TQDT--GNNSI.-GXRV-R-   123
N.FR.11.N1_FR_2011         X-----VN.----L-A--I-RI------Y.--S..T-.S------------A-----RA------S-C--GPQ---D-P----D-----TK--S--P.--ESQQ......GTATTE*IQN---GNTSI.-GK-V-R-   123

P.CM.06.U14788             -----DED.LR---T-I-I--I------S.-AR..GP.S-NST--------S-----DQ-AG--IASH---PE-LGGAD--DISQ-HIEDQGPPDN-.VDTPGP......TTEQ..*L-AR-LCCVWL.-D-RI-TX   121
P.FR.09.RBF168             -----DED.LR---T-I-I--I------W.-ER..GS.S-NST--K-----S-----DQ-AG---ASR---PE--GGAD--D-S--HIGDQGPPNNP.VDTPGP......TADQRA*L-AR-LCCVWL.-D-RI-TK   123

CPZ.CD.06.BF1167           ----EEE...N-LQAIKI--I--E---Y.--S..--.--S----------I----VEG--S---A-----LPTAP-VD--D-SK---APLDIDQVD.PSVEPIAVGDCETIAAQAHTP-DTISTPTGN.-GNSN*TA   127
CPZ.CM.05.SIVcpzMT145      -----EG..--N-L-AI-I-RI--E---Y.-SQ..S-.--R-----K----------NQ-G--V-ASC----S--PQ-P---I---CISDL-TT-DT.--IQPQ......GSTTNRY--TPTERD-A-.-DNRNQ--   123
CPZ.GA.88.GAB1             -----EPQD-AR-LQA-KI--I------Y.-S-..--.--K----------A--K--SE--G-V-A-----PPK-GD-E--E-DK-S-Q-V-TTQDV.--SNTS......QPQTAT-ETVPAGGNYSI.-GK-A-N-   125
CPZ.TZ.06.SIVcpzTAN13      ----EE-...AN-LE-I----I--D---Y.-SG..A-.S-A----------Q--D-VDALAQ---QHR--GPP--P--D--D-SQ-R-AAGDHTPSP.E-GDH*LHTDHNNST*KLSA*VETI*RLACQ.CLCDC-DF   125
CPZ.US.85.US_Marilyn       ------GDA--Q-L-A--I--I--A---F.--N..S-.------------KN-----NE--Q---S-C----S--PD-P---IGG--INSE--STNP.--EHQQ......GTATVE*TQST-LGGNRL.-GK-A-S-   124

GOR.CM.07.SIVgor2139_287   -----DED.LQ--LTKI-I-R-------W.-E-..GP.S-NST-------KG-----DQ-AG---TSH---YQN-GCVD--DISH--IGDQG-NTQP.DRVPDQ......TTEKSE*FDKR-LCFYCL.-GTR-*AK   122
GOR.CM.07.SIVgorCP2135con  -----DED.LQ--LTKI-I-R-------C.-E-..GP.S-NST-------KG-----DQ-AG---AAH----QN-DCVD--DISH--IGDQGNNTQP.DRAPDQ......TTEKSE*FDKR-LCFYCL.-GTR-*AK   122
```

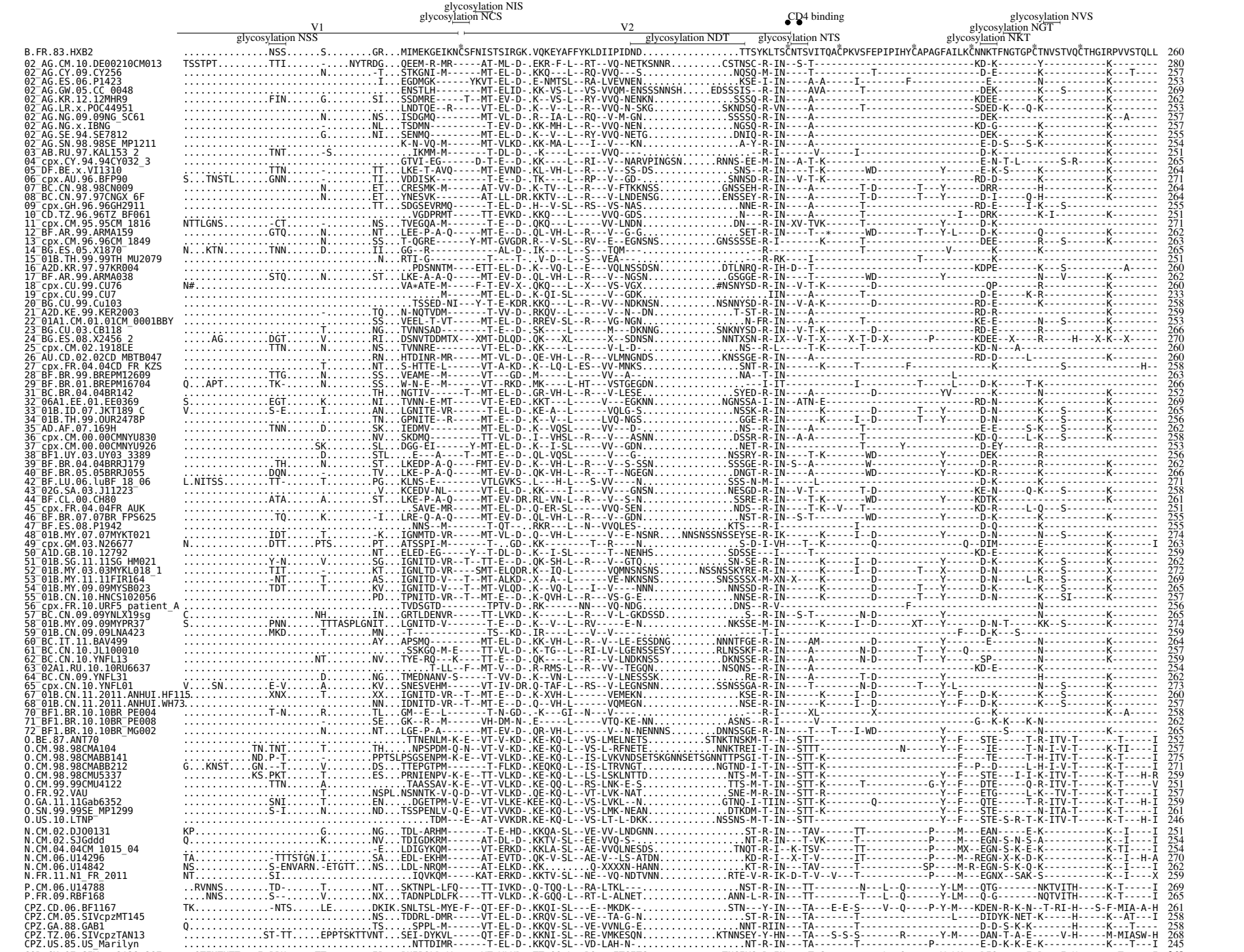
```
                                                                                           phos phos
                                   transmembrane domain                              cytoplasmic domain
                              Vpu start                                        α-helix                α-helix                    Vpu end
B.FR.83.HXB2                  TQPIP.....IVAIVALVVAIIIAIVVWSI.........VIIEYR.......KILRQRKIDRLIDRLIERAEDSGNESEGEISAL.VEM....GV.EMGHH...APWD.VDDL...*   82
A1.AU.03.PS1044_Day0          MT-LE.....-WS--G----L-------T-.........-G---K.......-L-K--------E-IR------------DTDE-.AL.....LI.---NY...DLGN.DIN-...-   82
A1.CH.03.HIV_CH_BID_V3538     MS-LE.....-Y--AG-I--L-------T-.........-G---K.......RL-K--R-----E-IR----------D-DAEE-.ST.....L-.-V-DR...-LLV.-NN-...-   82
A1.CY.08.CY236                MT-LE.....-S---G---VS-------T-.........-G--IK.......-V---------K-IR------------DTDE-.AL.....L-.---NY...DLG-.DNN-...-   82
A1.ES.05.X1608_8              MS-LE.....-S---G---TL-------TL.........-A----.......-V-K----E---E-IR------------DTDE-.AA.....LI.---DY...DLGN.DNN-...-   82
A1.KE.11.DEMA11KE001          MNSLT.PWN.-C---G-I--L-L-----T-.........-A----.......---K--------E-IK--E-------D-DTDE-.AA.....LC.---NY...GLG-..NN-...-   84
A1.RU.11.11RU6950             MT-LE.....-Y--A----VF-------T-.........-G----.......RL-K--------E-IR------------DEEE-.ST.....LM.---NL...-LL-.DNNV...-   82
A1.RW.11.DEMA111RW002         M-LLV.....VW--IG-I--L-L--I--T-.........-GL--K.......-L-K-------LE-IR----------D-DTEE-.WA.....T-.---NY...DLGN.DNN-...-   82
A1.SN.01.DDI579               MTSLE.....-W--IG-I--L-L-----T-.........-G---K.......-V-K-KR-----E-IR----------D-DTEE-.ST.....L-.-V-DD...I-LY.DNN-...-   82
A1.UG.11.DEMA110UG001         M-LLE.....-W--AG---L-L--I--T-.........-G---K.......RLQ--K------E-IR----------D-DTDD-.AK.....L-.---NY...DFG-.-NN-...-   82
A1.ZA.04.04ZASK162B1          MT-LE.....-C--AG---L-------T-.........-G----.......-L---------K-IT----------D-DTEE-.ST.....L-.D--NY...DLG-.DNNW...-   82
A2.CM.01.01CM_1445MV          ML-LA.....-LS--G-I--F-L-----T-.........-F-Q-K.......ELKK-----W--K-IS----------D-DTEE-.ST.....L-.---NL...DF--.ANN-...-   82
A2.CY.94.94CY017_41           ML-LV.....-L---G-I--L-L-----T-.........-F---K.......--KK-----W--K-IS----------D-DTEE-.SA.....L-.-R--L...DFG-.-NNV...-   82

B.BR.10.10BR_MG029            M-TLI.....-AS------VA----A---V.........-L----.......-----------------IR------------DQEE-.S-.....XL.----L.......-N--...-   82
B.CA.07.502_I191_03           M--LA.....-L--A-----A-L-----T-.........-G----.......--------N--L--IA------------DQEE-.SALVERGHL.-----.......-N--...-   87
B.CH.08.M2_0803101_NFLG8      MT-LY.....-WS-------A-------T-.........AL----.......----------IL--IR------------DQEE-.SA.....L-.----D.............-   82
B.CN.10.DEMB10CN002           MFSLQ.....----------A-L-----T-.........-Y-Q--.......---K----------I---T---------DQEE-.SA.....L-.D---D......N.-N--...-   82
B.ES.10.DEMB10ES002           MTTLV.....-CS-A----VA----I--T-.........-L----.......------R---I---IR------------DQEE-.SKL....V-.D---L.......I----...-   83
B.FR.11.DEMB11FR001           M-SLT.....-L---------L-----T-.........-F----.......---------S----IR-----------EDQ#...#A.....LM.----L.......I----...-   78
B.GB.05.MM45d213_GN1          M--LI.....VI-------VG----------.........-L----.......-----------------IR------------DHEE-.AA.....L-.-R--L.......IN---...-   82
B.HK.06.HK003                 M-SLE.....-L-V------A-L--I------.........-F----.......R-R-------------IR------------DQED-.SA.....LL.---N-.......I----...-   82
B.HT.05.05HT_129389           M-SLX.....XL--------X-X--I--X-.........-F----.......X-K--KR-------IRD-----------DQEE-.SA.....L-.D---L.......-N--...-   82
B.JP.12.DEMB12JP001           M--LQ.....-L--------A-L-----T-.........-G----.......---K--------N-IS------------DQEE-.SA.....LM.-----........-N-M...-   82
B.KR.07.HP_18_07JHS10_3909    M-SLT.....-FT-------A---------.........-F---K.......---K-----------IR------------DQEE-.SA.....L-.-------------------   82
B.PE.07.502_0525_wg5          M-SLV.....-L--------A-L-----T-.........-FL---.......---------------IA------------DHEE-.SA.....LM.---YN......N.---M...-   82
B.RU.11.11RU21n               M--LV.....-AT---F--VG--------L.........--V---.......-------------A-I----------D-DQ#...#A.....LM.----L.........---M...-   78
B.TH.08.MERLBDTRC10           M-ALY.....-L------I-A-------T-.........-------.......-----K--------IR------------DQEE-.SAL....V-.----D.............-   83
B.US.11.ES38                  MPSLY.....--S-------A-L--I------.........-F----.......---K-Q-------IR----------D-DQEE-.SA.....L-.---------------------   82

C.AR.01.ARG4006               #....LXXDYRLG-G--I--L----I--T-.........AY-----.......-LV--KR--W--E-IR------------DTEE-.ET.....M-.D---L...RLLH.-X--...-   82
C.BR.07.DEMC07BR003           MLTL-DRVDYRLG-G--I--L----I--T-.........AY-----.......-LV---R-----E-IR----------D-DTEE-.ET.....L-.D--NL...RLL-.-HA-...-   87
C.BW.00.00BW5031_1            MFALF.EVDYRLT-G-FI--LFL-----T-.........AYL----.......-LV---R--Q-VK-IR----------D-DTEE-.ST.....M-.D---L...RLL-.-N--...-   86
C.CN.10.YNFL19                MVD....LDYK-TVG-FI--L-L--I--T-.........AY-----.......-LTK-----W--K-IG------------DTEE-.ST.....M-.D--RL...RLL-.-N--...-   83
C.CY.09.CY260                 MTGLIEKVDYKI-V--FII-LV------I-.........-YL----.......-LV---------E-IR----------D-DHEE-.SK.....M-.D--NL...RLL-.AH--...-   87
C.ES.08.X2363_2               MLDLLAKVDYRI-V--FLI-L-L-----I-.........AY-----.......-LVK-TR--Q--K-IR------------DTEE-.AT.....L-.D---L...RLL-.-N--XCVG   90
C.ET.02.02ET_288              MFSFTERVDAR-G-G--I--L-L--I--T-.........AY-----.......-L----------E-IR------------DTEE-.ST.....M-.D--TL...RLL-.-H--...-   87
C.IN.09.T125_2139             MVE....LNYKL-VG--I--L-L-----I-.........-Y-----.......-LVK-----W-VK-IR----------D-DTEE-.ST.....M-.D---L...RLL-.-N--...-   83
C.KE.00.KER2010               MLDLLARVDYRLGVG--I--L----I--T-.........-Y-----.......-L-------W--E-IR------------DTEE-.AT.....M-.D---L...RLL-.-N--...-   87
C.MW.09.703010256_CH256.w96   MLDLVARVDYRIGVG--I--L-------IL.........AY---K.......RW-K-K---W--K-IR------------DTEE-.ST.....M-.D-E-L...GLL-.-NG....-   86
C.TZ.08.707010457_CH457.w8    MLDLLARVDYRIGVA--LI-L-------T-.........-Y---K.......-L---KR--W--K-IR----------D-D-EEV.ET.....L-.D--NL...RLL-.-IN*...-   86
C.YE.02.02YE511               MLDLLAKVDYRIG-A-F---L----I--T-.........AY-----.......-LVI-----W-VK-IR----------D-DTEE-.ST.....M-.D---L...RLL-.IN--...-   87
C.ZA.08.705010534_CH534.w12   MLN....LDYR-G-G--I--LN------I-.........-Y-----.......-LV---R--W-VK-IR----------D-DTEE-.ST.....M-.D---L...RLL-.-I--...-   83
C.ZA.10.DEMC10ZA001           MLSLLEKVDYRI-VA-FTI-L-L-----T-.........AY-----.......-VV--T------K-IR----------D-DQED-.ST.....M-.D---L...RLL-........-   85
C.ZM.11.DEMC11ZM006           MISLIERVDYRLGVG--II-L-------T-.........AYL----.......-L---K--W--E-VR------------D-EE-.ST.....M-.D---L...RLL-.GNE-...-   87

D.CM.10.DEMD10CM009           M--LV.....-IS------TL-------T-.........-Y-----.......--Q--K---W---IR--E---------DKEE-.ST.....L-.D-----.......IVA---...-   83
D.CY.06.CY163                 M-SLE.....-L--A-----L-L-----T-.........-Y-----.......--K------W---IG------------DKEE-.SQ.....LM.-----.......NVA--I...-   83
D.KE.11.DEMD11KE003           M-SLQ.....-A--------L-L-----T-.........-F--C-.......RLR--#---W---IR------------DEEE-.SA.....F-.G--Y-.........IN-I...-   81
D.KR.04.04KBH8                M--VV......I--X----TL----I--T-.........-F--CC.......RLRK--------N-IR------------DEEE-.ST.....LM.-----.......N------...-   81
D.SN.90.SE365                 M-SLE.....-L--A-I---L-L-----T-.........-F--V-.......--R------W---IK------------DEKE-.SA.....L-.D----.......N------...-   82
D.TZ.01.A280                  MT-LE.....-T--A---I-S-L-----T-.........-Y-----.......R-KK-K---W---IR----------D-DTEE-.ST.....L-.-----.......N.---I...-   82
D.UG.10.DEMD10UG004           M--LV.....-L--------LLL-----T-.........-F--C-.......-LR--KR--W---IR------------DEEE-.SK.....FM.-----.......N.I--M...-   82
D.UG.11.DEMD11UG003           MNLLV......IG---------L-----A-.........-F--W-.......RLK------KI-N-IA------------DREE-.SA.....L-.G---L.........IG-I...-   81
D.YE.02.02YE516               M-TLE.....-LS-----I-A----I--T-.........-Y-----.......--R-----Q---IR------------DEEE-.ST.....LM.----A.......NVA---...-   83
D.ZA.90.R1                    M--LV.....-L--------L-L-----T-.........-F-----.......R-RK----Y---IR------------DGEE-.SK.....L-.---YD.......N------...-   82

F1.AO.06.AO_06_ANG32          MSDLL.....AI--------L-------T-.........AYL----.......-VVK-KR-N--YE-IR----------D-DAEE-.A-.....LG.-V-PF...I-GG.INN-...-   82
F1.AR.02.ARE933               MSYLL.....AIG-T--I--L-------T-.........AY---K.......-LV---R-N--YK-IR------------DAEE-.AA.....LG.---PF...I-R-.I-N-...-   82
F1.BR.10.10BR_RJ015           MTDLL.....VIS-A--I--L--V----T-.........-YL----.......-LV-----N--YQ--R------------DAEE-.AA.....LG.---PF...L-G-.INN-...-   82
F1.CY.08.CY222                MTNLL.....AI--A--IL-L-V-----T-.........AY---K.......-VI-----N--YKKIR------------DAEE-.AA.....LG.-V-PF...I-G-.INN-...-   82
F1.ES.02.ES_X845_4            MSNLL.....AIS-------L--V-I--T-.........-FL----.......-L------N--YE--R------------DAEE-.AA.....LG.-V-PF...I-GN.INN-...-   82
F1.RO.96.BCI_R07              MSHVL.....AI-----I--L----I--T-.........AY-----.......-LV-----N--YE-IR------------DAEE-.AA.....LG.---PF...I-G-.INN-...-   82
F1.RU.08.D88_845              MTE-I.....-IG----I--L-V-----T-.........-F-----.......--------N--YE-IR----------D-DAEE-.AA.....LG.---PF...V-G-.-NN-...-   82
F2.CM.02.02CM_0016BBY         MSYLI.....-LV---FI--L-A--I--T-.........-Y---K.......-Q---KR-N--YE-IR------------DAEE-.AA.....LG.-V-LF...I-GN.INN-...-   82
F2.CM.10.DEMF210CM001         MPYLI......IL---FIAVL----I--T-.........-Y-----.......---K---NK-YK-IR----------N--DAEE-.AA.....LG.-V-PF...I-G-.INN-...-   81
F2.CM.10.DEMF210CM007         MSLSI.....--V-A-FI-VL----I--T-.........-----K.......-L---KR-NK-YE-IR------------DAEE-.AA.....LG.---PF...I-GN.INN-...-   82

G.BE.96.DRCBL                 M--LE.....-S---G-I--S-A-----T-.........-F-----.......--RK-KR-EK-L--IR------------DTEE-.AT.....LM.-L-DF...D--V.G-N-...-   82
G.CM.10.DEMG10CM008           M-ALV.....-A---G-I--FLA--T--T-.........-F-Q--.......E-RK-N---K-L--IR--I---------DTEE-.AA.....L-.DV-DF...DL-I.GNN-...-   82
G.CN.08.GX_2084_08            M--LE.....-A---G-I--F-A-----T-.........-Y-----.......--R--KR--KIL--IR------------DTEE-.AT.....L-.D-VEF...D--V.G-NM...-   82
G.CU.99.Cu74                  MKSLE.....-S---G-II-G-A-----T-.........-F-Q--.......E-RK-K----ILE-IR------------DTEE-.SA.....L-.---DF...DH-V.G-N-...-   82
G.ES.09.X2634_2               M--LE.....-A---G-I--GLG--I--T-.........-F-L-K.......E-RK-QR-Q-IL--IG------------DTEE-.AK.....LM.---DF...D--V.G-N-...-   82
G.GH.03.03GH175G              M-SLE.....-S---G-I--S-A--I-------.........-F-----.......R-RK-K--EK----IR------------DTDE-.AA.....L-.---DF...D--V.G-N-...-   82
G.KE.09.DEMG09KE001           M--LE.....-AS--G-I--FLA-TI--T-.........-F-Q--.......E-RK-------Y--IR------------DEEE-.AT.....L-.GR-AF...D--V.G-N-...-   82
G.NG.09.09NG_SC62             M-SLE.....-A---G-I--F-A--I--T-.........-F-----.......--RK-KR--S-LN-IR------------DTEE-.AT.....L-.D-VDF...D--V.GNN-...-   82
G.ZA.01.TV546                 M--LE.....-S---G-I--F-A-----T-.........-LL----.......--RKE---EK-L--IR------------DTEE-.AT.....LM.---XF...D--V.G-T-...-   82

H.BE.93.VI991                 MNILG......IG-G-----F-------T-.........AY-----.......-L#K--------E-IR----------D-DTEE-.SK.....L-.----L...NLGY.-A--...-   80
H.CF.90.056                   MYILG......LG-G----TF---VI--T-.........-Y---K.......-LV--K------E-IG----------D-DTEE-.SK.....LM.----L...NLGY.-A--...-   81
H.GB.00.00GBAC4001            MYILG......IG-G--I--G-LC-I--T-.........-YL----.......-LVQ-KR----LE-IR----------D-DTEE-.ST.....L-.-R--L...NLGY.-A-V...-   81

J.CM.04.04CMU11421            MNSLQ.....-AS-------FFL-VXA-XV.........AY-----.......-LV---R--K-X--IR--E-----D-D-DTEE-.AK.....L-.---P-...DL-N.-------   82
J.SE.93.SE9280_7887           MI-LQ.....-A----FI---FL--GM-T-.........-Y---K.......-L-------K---IR----------D-DTEE-.AD.....L-.-R-P-...DL-N.-N--...-   82
J.SE.94.SE9173_7022           MVSLQ.....-------I--FFL--C--T-.........-Y---K.......-L-------K--N-IR----------D-DTDE-.A-.....L-.---P-...DL-N.-N--...-   82

K.CD.97.97ZR_EQTB11           MV-LT......-G-I---A-L-L--I--T-.........AYL----.......-VV--KR-NW-F--IR------------DTEE-.AA.....LG.-T--L...ILG-.INN-...-   81
K.CM.96.96CM_MP535            MVSLA......IS-------L-L--I--T-.........-Y-----.......-LVK-KR-NW----IR------------DAEE-.AD.....IG.-L--L...ILGN.I-N-...-   81

U.CA.01.TV749                 MYTLQ.....--------I--TL---C--T-.........-YV----.......-L---K--------IT----------D-DTEE-.SK.....L-.---P-...DL-N.-N-....-   82
U.CA.99.TV721                 M-TLQ.....--S----I--AL---C--T-.........-YV----.......RL---K---K----IR----------D-DEEE-.SA.....L-.---P-...DL-N.-N--QD.-   84
U.CD.83.83CD003_Z3            MSILQ.....------II--L-L-----T-.........-Y---K.......RL-------W----IR------------DTEE-.ST.....L-.--EPD...NFRN.DN-M...-   82
U.CD.90.90CD121E12            MWIL-.......-G---II--LGL-----T-.........-YL--K.......-L---QR--W----IR------------DTEE-.SL.....L-.--EPD...NFRN.DNNM...-   81
U.CY.05.CY090                 M-TLL.....-AS-I-----SL-V-A--T-.........-L---K.......-LR--ER-N--YE-IR--T---------DTEE-.AA.....LG.----L...V-G-.IHN-...-   82
U.CY.08.CY223                 MNTPT.....-LS---I--F--VVI--T-.........-F---K.......-LRK--------Q-IR----------D-DTEE-.ST.....L-.-V-DN..VHLLV.N-GM...-   83
U.ES.10.DEURF10DZ001          M--LE.....-A---G----F-A-----------.........-Y-----.......#-RK-K---K-L--IR------------DTDE-.AK.....L-.---DF...D--V.G-N-...-   81
U.GR.99.99GR303               M-SLE.....-AS--G-I--F-A-----D-.........-L-V--.......-*EN--R-----V-IR---------I--DTEE-.SK.....LM.---PD...DF-N.-N--...-   81
U.NL.95.U_NL_95_H10986_D1     MNILV......-G-A-----F-LVVI--T-.........-YC--K.......-LK--ER--KI-N-IR------------DAEE-.AA.....LG.---YL...VLR-.I----...-   81

01_AE.AF.07.569M              MT-LE.....-S---G-I--L-S-----T-.........-A--AK.......-L---------VK-IR------------DTEEW.AK.....L-.---DF...D--V.G-N-...-   82
01_AE.CF.90.90CF11697         MSALQ.....-*---G-I--L-L-----T-.........-F---K.......------------E-IR--------G---DTNE-.AT.....L-.-V-DF...D--V.G-N-...-   81
01_AE.CN.10.YNFL03            MT-LE.....-S---G-I--L-L-----T-.........-A--F-.......R-----------K-IR------------DTDE-.AK.....L-.---DF...D--V.G----...-   82
01_AE.HK.04.HK001             MT-LQ.....-S---G-I--L-L-----T-.........-GL-V-.......------------KKIR--E---------DTDE-.AK.....L-.---DF...D--V.G-N-...-   82
01_AE.IR.10.10IR.THR48F       MT-LE.....-S--AG---L-L-----T-.........-A-----.......---N--------K-IS--X---------DTEE-.AK.....L-.---DF...D--V.G-N-...-   82
01_AE.JP.x.JRC77AE            MT-LE.....-S--AG-I--L-L-----T-.........-A---K.......-----K-----LN-IR------------DTDE-.AK.....L-.---DF...D--V.G-N-...-   82
01_AE.TH.09.AA090a_WG11       MT-LE.....-S---G-I--L-L-----T-.........-G--IK.......RL----------K-IS------------DTDD-.AK.....LL.---DI...D--V.G-N-...-   82
01_AE.TH.90.CM240             MT-LE.....-S---G-I--L-L----*T-.........-A--VK.......-----------VK-IR------------DTDE-.AK.....L-.---DF...D--V.G-N-...-   81
01_AE.US.05.306163_FL         MS-LE.....-S---G-I--A-S-----T-.........-A--VK.......--F--------VK-IR------------DTDE-.AK.....L-.---DF...D--V.G-N-...-   82
01_AE.VN.98.98VNND15          MT-LE.....-C---G-I--L-L-----T-.........-G--F-.......---K-------VK-IR------------DTDE-.AK.....L-.---DF...D--V.G-N-...-   82
```

```
                                                                                   phos phos
                                           transmembrane domain                  cytoplasmic domain
                              Vpu start                                         α-helix            α-helix                   Vpu end
B.FR.83.HXB2                  TQPIP.....IVAIVALVVAIIIAIVVWSI.........VIIEYR.......KILRQRKIDRLIDRLIERAEDSGNESEGEISAL.VEM....GV.EMGHH...APWD.VDDL...*   82
02_AG.CM.10.DE00210CM013      M--LE.....-S---G----F-A-----T-.........-Y---.......--KK-K-----L--IR----------D-DEEE-.ST.....LM.-I-YD...NLL-.NNN-...-   82
02_AG.CY.09.CY256             M--LE.....-A---G----F-A-----T-.........-Y---.......--RK-K-----LE-IR--------D-D-DTEE-.ST.....LM.-I-YD...NIL-.NNN-...-   82
02_AG.ES.06.P1423             MT-LE.....-A---G----F-A-----T-.........-Y---.......--RK-K---K-L--IR-----------DTEE-.ST.....LI.-V-DY...N-L-.NNI-...-   82
02_AG.GW.05.CC_0048           M-SLA.....-A---G----F-A-----T-.........-Y---.......--RKE---SN-L--IR----------D-DTEE-.AL.....LM.---DN...NIL-.NNN-...-   82
02_AG.KR.12.12MHR9            M--LE.....-S---G-I--F-A-----T-.........-F---.......--RKE---SK-L--IR-----------DTEE-.ST.....LM.---DY...NIL-.N-N-...-   82
02_AG.LR.x.POC44951           M--LE.....-A---G----F-A-V---T-.........-V---.......--RK-K--H--L--IR----------D-DTEE-.ST.....LM.-I-YD....IL-.NNN-...-   81
02_AG.NG.09.09NG_SC61         M-VLG.....-AS--G-I--S-A-----T-.........-F---.......--RKE---S--L--IR----------D-DTEE-.AT.....LM.-R-AD...NVL-.N-N-...-   82
02_AG.NG.x.IBNG               M--LT.....-T---G----F-A---------.........-Y---.......--RK-K-----L--IR----------D-DTEE-.ST.....LM.---YE...YIL-.N-N-...-   82
02_AG.SE.94.SE7812            MNSLD.....-----G----F-A-----T-.........-Y---.......--RK-K-----L--IR----------D-DTEE-.ST.....LM.---YD...NIL-.N-N-...-   82
02_AG.SN.98.98SE_MP1211       M--LA....VA---G----F-L-----T-.........-F---.......--RK-G---K-L--IR----------D-DTEE-.ST.....LL.---YD..N-AL-.I-N-...-   83
03_AB.RU.97.KAL153_2          M-SLA.....-A-------VG------G--.........-F---.......---------------IR-----------DQE....A.....LM.----L...V---.A----...-   79
04_cpx.CY.94.94CY032_3        MLFWE.....-W---G----L--V----TL.........-F---K.......-LR---R--S-YN-IR----------D-DAEE-.ST.....L-.G--NF...D--V.G-N-...-   82
05_DF.BE.x.VI1310             MSDLL.....TI-V---I--L-------I-.........-Y---K.......-LV-----N--YK-IR-----------DAEE-.AA.....LG.-V-PF...I-G-.INN-...-   82
06_cpx.AU.96.BFP90            M-ALE.....-A---G----FLA-----T-.........-F-Q--.......E-RK-K--EK-L--IR-----------DTDE-.AT.....LM.---DF...D--V.G-N-...-   82
07_BC.CN.98.98CN009           M-TLT.....-L--------A-------T-.........-L---.......------KR-------IR----------D-DQEE-.SA.....FM.------...------------...-   82
08_BC.CN.97.97CNGX_6F         MLDLE......L-VG--I--L-------T-.........-Y---.......RLVK-----W--K-IR-----------DTEE-.ST.....M-.D--RL...RLL-.-N--...-   81
09_cpx.GH.96.96GH2911         M-SLQ.....-AS--GII--F-A--IA-T-.........-F--V-.......--RK-KR-------IRD-----------DTEE-.AK.....L-.---P-...DS-N.-----...-   82
10_CD.TZ.96.96TZ_BF061        MLSLLARVYYKLRVG--I--L-------IR.........-Y-Q--.......-LVEPT---W--K-IR-----R-HD-D-DTEE-.ST.....M-.D--NL...RLL-.-N--...-   87
11_cpx.CM.95.95CM_1816        MISLQ.....-AS-I-----F--VV---T-.........-GL---.......-L------------IR-----------DTEE-.AQ.....L-.---P-...DL-N.-----...-   82
12_BF.AR.99.ARMA159           M-SLV.....TL--------A-L---A---.........-L---.......---------------IR-----------DAEE-.AA.....FG.-L-LL...I-GN.I-N-...-   82
13_cpx.CM.96.96CM_1849        MNAL-.....-IS----I--F-L-V---T-.........AY---.......-LVH-K--EK-L--IR-----------DTDD-.AK.....L-.---NF...D--V.G-N-...-   82
14_BG.ES.05.X1870             M-SLE.....-A---G-I--G-A--X--T-.........-F-Q-K.......E-RK-K--QGIL--IR-----------DTEE-.AV.....LM.---GF...D--V.G-N....-   81
15_01B.TH.99.99TH_MU2079      MS-LE.....-S---G-I----L-----T-.........-A--I-.......N--K-------VK-IR--E---------DTDE-.AK.....L-.---DL...D--V.G-N-...-   82
16_A2D.KR.97.97KR004          MNSLQ.....-LS--G-I--F-L-V---T-.........-F---K.......--KK-N---W--K-IS-----------DTEE-.ST.....LM.---NL...DFG-.AN-...-   81
17_BF.AR.99.ARMA038           M-ALE.....-L-------T------------.........-L---.......-L---------E-IR-----------DAEEI.SA.....LM.-----.....-----.IN--...-   82
18_cpx.CU.99.CU76             MXXFE.....--G-I--I--F---VC--T-.........-Y---.......---K-K---G-FE-IRD------Y----DTEE-.SK.....LM.---N-...N-GY.-A-V...-   82
19_cpx.CU.99.CU7              M-ALQ.....-A---G-I--F-A-----T-.........-Y---.......--RKE---EK-L--IR----------D-DTEE-.ST.....LM.-L-DF...D--V.G-NV...-   82
20_BG.CU.99.Cu103             M-SLE.....VA---G-I--G-A--I--A-.........-F-L-K.......E-RK-K----ILE-IR-----------DTDE-.AV.....L-.D--DF...DH-V.G-N-...-   82
21_A2D.KE.99.KER2003          M-ALQ.....-I--A-----T-------T-.........-F---.......R-K---R--C----IR-----------DREE-.SK.....L-.----Y....---N.I----...-   82
22_01A1.CM.01.01CM_0001BBY    M--LE.....-CT--G-I--L-L-----T-.........-A--IK.......-VR--K-----LE-IR-----------DTDD-.AK.....L-.---NF...D--V.GNN-...-   82
23_BG.CU.03.CB118             MHSLE.....-A-V-G-I--G-A--I--A-.........-F-L-K.......Q-RK-K----ILE-IR-----------DTEE-.AA.....L-.D--DF...DH-V.G-N-...-   82
24_BG.ES.08.X2456_2           M-SLE.....-A---G-I--C-A--I--T-.........-W-Q-K.......E-RK------ILE-IR-----------DTNE-.AT.....L-.D--DF...DH-V.G-N-...-   82
25_cpx.CM.02.1918LE           M-SLE.....-F---G-I--F-V-VI--T-.........-YL-I-.......--RK-K---Q-LE-IRT----------DTEE-.ST.....LM.---DF...DA-V.G-N-...-   82
26_AU.CD.02.02CD_MBTB047      MT-LQ.....-Y---G-I--L-L---------.........-LV--K.......---K-------L--IRD----------D-DTEE-.SA.....L-.---D-...N--G.YNV-...-   82
27_cpx.FR.04.04CD_FR_KZS      MISLK.....TIS-A--I--FL-VV---T-.........-Y---.......-L---KR----LE-IR----------D-DTEE-.AR.....L-.---P-...DLGN.DN--...-   82
28_BF.BR.99.BREPM12609        M-ALI.....-A-------VA-------T-.........-F---.......---------------IR-----------DAEE-.EA.....LG.D--PL...VLGN.INN-...-   82
29_BF.BR.01.BREPM16704        M-SLQ.....-I-------VA-------T-.........-F---.......-------R---I-N-IR-----------DHED-.LA.........LV--D.........-N--...-   80
31_BC.BR.04.04BR142           MFTL-ERIDYR-G-G--LI-L----I--T-.........-Y---.......-LV---R--W-VK-IR-----------DTEE-.ET.....M-.D--NL...RLL-.-P--...-   87
32_06A1.EE.01.EE0369          M-ALE.....-S---G-II-SLA-----T-.........-L-Q--.......E-RK----EK-L--IR-----------DTDE-.-T.....LM.---NF...D--V.G-N-...-   82
33_01B.ID.07.JKT189_C         MS-LE.....-S--AG-II-L-L-----T-.........-G--CQ.......---Q-------VK-IR-----------DTEE-.LK.....F-.---NL...D--V.G-N-...-   82
34_01B.TH.99.OUR2478P         MS-LE.....-S---G-I--L-L-----I-.........-A---.......-----E------VK-IQ-----------DTDE-.AK.....L-.---NF...D--I.G-N-...-   82
35_AD.AF.07.169H              MTSLY.....-W--IG----L-L-----T-.........-G--I-.......RV---------VK-IS--T-------D-DTDE-.AK.....L-.---NY...DLG-.DNN-...-   82
36_cpx.CM.00.00CMNYU830       M-SLQ.....-A--AG----F-A--GG-T-.........-L---K.......--I--K-----LE-IK---K--------NTEE-.ST.....L-.---DF...DH-V.G-N-...-   82
37_cpx.CM.00.00CMNYU926       M--LE.....-YS--G----F-A-----T-.........-L---K.......--RK---------N-IR----------D-DTEE-.ST.....LM.-L-DF...D--V.G-N-...-   82
38_BF1.UY.03.UY03_3389        MSNLL.....VLG----I--L-------T-.........-Y---.......--V---R-N--YE-IR-----------DAEE-.AA.....LG.-VQLV...I-GN.INN-...-   82
39_BF.BR.04.04BRRJ179         MSDLL.....AIG-T--I--L-------T-.........AYL---.......-FV---R-N--YE-IR-----------DAEE-.AA.....LG.-V-PF...ILG-.INN-...-   82
40_BF.BR.05.05BRRJ055         M-SLQ.....-A---G---VA-------T-.........-LL---.......-L-K--------E-IA-----------DQEE-.ST.....LM.-----.....-----------...-   82
42_BF.LU.06.luBF_18_06        M-SSV.....-L--------AL---------.........-L---.......---K--------K-IS----------D-DHEE-.SA.....L-.----------.....I--Q...-   82
43_02G.SA.03.J11223           M-XLE.....-A---G-I--FLA-----T-.........-F---.......--R--KX----LX-IR-----------DTEE-.AI.....L-.DNVXX...D--V.G-N-...-   82
44_BF.CL.00.CH80              MTDLL.....AIS-A--I--A--V----I-.........-Y---.......-LV-----N--YE-IR-----------DAEE-.AA.....LG.-V-PF...I-GN.I-N-...-   82
45_cpx.FR.04.04FR_AUK         MNYLI......-G-A--I--ALL--I--T-.........-Y---.......-VKK-ER-NW-YN-IR-----------DTEE-.AA.....LG.----L...ILG-.IAN-...-   81
46_BF.BR.07.07BR_FPS625       MSDLL.....AISLA--I--L-------T-.........AYL--K.......-LV-----N--YK-IR-----------DAEE-.AA.....LG.---PF...I-GN.INN-...-   82
47_BF.ES.08.P1942             M-SLE.....-L--------A-L---------.........-L---.......----E--VYK----IR-----------DQEE-.SA.....M-.-R--L......-----.IN--...-   82
48_01B.MY.07.07MYKT021        MT-LQ.....VS---G----L-L-----T-.........-G--F-.......-------------K-IR--T---------DTGE-.AK.....L-.---DF...D--V.GGN-...-   82
49_cpx.GM.03.N26677           M-SLQ.....-A-----I--FFLV-I--T-.........-Y-A--.......QLV-----E-----IR----------D-DMDD-.#K.....L-.-R-SL...DFRN.AN--...-   81
50_A1D.GB.10.12792            M--LV.....-LS-A---IVS-------T-.........-F---K.......RLR-----------IR-----------DREE-.SA.....ML.---D.........---M...-   82
51_01B.SG.11.11SG_HM021       MSSLD.....----------A-L--I-------.........-F---.......-L-K-K-------IA-----------DQEE-.AA.....L-.-R--L...V-G-.I----...-   82
52_01B.MY.03.03MYKL018_1      M--LE.....-S--AG----L-L-----T-.........-A--FK.......-------R----VK-IR-----------DTDE-.AT.....L-.---DF...D--V.G-N-...-   82
53_01B.MY.11.11FIR164         MT-LQ.....-S---G----L-L-----T-.........-VL---.......-----------VN-I------------DTDE-.AK.....L-.---NF...D--V.G-X-...-   82
54_01B.MY.09.09MYSB023        MT-LQ.....-S--AG-I--L-L-----T-.........-L---.......---K----------IR-----------DTDD-.AK.....L-.D--DF...DD-V.GNN-...-   82
55_01B.CN.10.HNCS102056       MT-LE.....-S--AG-I--L-L-----T-.........-A---K.......--QKE---EK-LE-IR-----------DTDE-.AQ.....L-.D--DF...DH-V.G-NV...-   82
56_cpx.FR.10.URF5_patient_A   M--LE.....-A---G-I--F-A--I--T-.........-Y---.......--RK-QR----L-KIR-----------DKEE-.AT.....LM.D-VDF...DS-V.G-N-...-   82
57_BC.CN.09.09YNLX19sg        MLN....INYRLGVG-----L-------T-.........-Y-K--.......-LV------C--KKIR-----------D-EE-.ST.....M-.D-E-L...RLL-.-NE-...-   83
58_01B.MY.09.09MYPR37         MS-LE.....-S---G-I--L-L-----T-.........-A--LL.......--RK-----K-VE-IR-----------DTDE-.AK.....L-.---NF...D--V.GGN-...-   82
59_01B.CN.09.09LNA423         MLSLE....T-GV-I-F-IVG-------T-.........-YL---.......-------K-------IA----------D-DQEE-.SA.....LM.----R......---N.-------...-   83
60_BC.IT.11.BAV499            MLG...SIDYRLGVG--I-RL----I--I-.........-Y---.......-LV---R--W-VK-IK----N------EDAEE-.DN.........MV-YL...RLL-.-N-V...-   82
61_BC.CN.10.JL100010          MLD....LESAL-VG-----L--V----T-.........-Y---.......-LVK--------K-IR-----------DTEE-.ST.....M-.D--NL...RLL-.-N-W...-   83
62_BC.CN.10.YNFL13            M*SLT.....-LS-------A-------T-.........-F---.......--R-----N-----IR--S-------D-DQED-.SA.....LM.---------....---N.IN---...-   81
63_02A1.RU.10.10RU6637        M-TLE.....-A---G----F-A-----T-.........-Y---.......--RK-K---K-L--IR----------D-DSEE-.ST.....LM.-V-NY....-LL-.DNNV...-   82
64_BC.CN.09.YNFL31            MLD....LNYTITVG-FI--L-------T-.........-Y---.......RLKK-----W--K-IR----------D-DTEE-.ST.....M-.DL-NL...RLL-.ANN-...-   83
65_cpx.CN.10.YNFL01           M--LV.....-S---G-I--L-L-----T-.........-L---.......--VK------I-K-IQ-----------DTDE-.AK.....L-.---DF...D--V.G-N-...-   82
67_01B.CN.11.2011.ANHUI.HF115 MT-LE.....-S---G-I--L-L--X--T-.........-G--F-.......-------------K-IR-----------DTDX-.AK.....L-.---DF...DH-V.G-N-...-   82
68_01B.CN.11.2011.ANHUI.WH73  MT-LE.....-S---G-I--L-L-----T-.........-G--F-.......-------------K-IR--E---------DTDE-.AK.....L-.---DF...DH-V.G-N-...-   82
70_BF1.BR.10.10BR_PE004       M-SLL.....-YS------VA--------.........-L---.......-------R-N-----IR-----------DAEE-.AA.....LG.D-ELL...I-G-.IGN-...-   82
71_BF1.BR.10.10BR_PE008       MP-LY.....-L-----I--G---------.........-----------.......----------V--IR----------D-DAEE-AAA.....IG.---PL...VLE-.IIN-...-   83
72_BF1.BR.10.10BR_MG002       MT-LY.....-S--------G-L-----T-.........-L---.......--RKEKR-E---N-IA----------D-DQEE-.SA.....L-.-R--L......-----.I--Q...-   82
O.BE.87.ANT70                 MHHRD.....LL--IIISALLF-NVIL-GF.........ILRK-LE...QKEQDRKE-E-LERLR-IR-IRD--DY--N--EEQ...-.....VM.DLVL....SHGF.DNPMFEP-   86
O.CM.98.98CMA104              MHYRD.....LL-LIIVSALLL-NVLL--F.........ILRK-LE...QKKQDR-E-E-LERLRKIR-IKD--DY--NE-EEQ...-.....VM.DLI-....SHGF.DNPMFEL-   86
O.CM.98.98CMABB141            MLHRD.....LLLLIIISALLLTN-IL-MF.........-LRK-LE...IKKQER-E-E-LERLRXIR-IRD--DY--NE-EEQ...-.....VRDHLV-....TFGF.ANPMFEI-   87
O.CM.98.98CMABB212            MHQRD.....LLILI-VSILCL-C-L--TF.........NLRK-LE...HRKQDK-E-E-LERLR-VR-IRD--DY--X--EEQ...-.....VM.DLI-....SHGF.ANP-FEL-   86
O.CM.98.98CMU5337             MHHRD.....LLXLIIISALLX-NVIL-MF.........ILRQ-LK...XXKQDR-E-E-LERLR-IR-IKD--DY--NE-EEQ...-.....VR.DLV-....SHGF.DNPMFEL-   86
O.CM.99.99CMU4122             M-HKD.....LLILIITSALLL-NVIL-LF.........-LKQCLE...QKKQTK-E-E-I-RLR-IR-IED--DY--N--EEQ...T.....VR.DLI-....SHGF.DNPMFEL-   86
O.FR.92.VAU                   MHHQG.....LLVLIVIIALLL-NVLL-MF.........NLRA-LE...QKKQDRTA-E-LER-R-IR-VED--DY--N--EEQ...-.....VR.DLVL....DFGF.DNPMFEL-   86
O.GA.11.11Gab6352             M-HKD.....LIIVIISSALLFLNVLI-TF.........-LKQCLD...HKKQDR-E-E-LERVR--R-IKD--DY--N--EEQ...-.....VM.DLV-....SHGF.YNPMFEL-   86
O.SN.99.99SE_MP1299           MHHRD.....LL-LITTSALLLTNV-L-TF.........ILRQ-LK...QKKQDK-E-E-LERLR-IRQIED--DY--D-TEEQ...-.....VR.DLV-....SYGF.DNPMFEL-   86
O.US.10.LTNP                  M-HKD.....L..LIIVSALLL-NVI--MY.........-LRK-LE...QKRQDR-E-E-LER-RKIR-IRD--DY--N--EEQ...-.....VM.DLV-....SHGF.DNPMFDL-   84

N.CM.02.DJO0131               ML........EWGFA--G------VII-VL.........LYK--K.......-LKL-E--EQIRQ-IRD-T----S-ND-DADW-.AI.....LL.SHDKF...DN-V.........-   75
N.CM.02.SJGddd                ML........FLGFL--GL-L---VII-VL.........LYR--K.......E-KL-ER-KHIRQ-IRD---------D-DTEL-.-T.....RL.SLSKL.DQGD-V.........-   77
N.CM.04.04CM_1015_04          ML........SLGFS--G------VII-VL.........LYK--K.......E-KL-E--EQIKQ-VRD---------D-DAEW-.AI.....LL.SPDKL...DN-V.........-   75
N.CM.06.U14296                MLQ.......SLGFI--GL-----V-I-VL.........IYR--K.......--KL-GR-KTI.............---GNN-NEEW-.AI.....LL.SPDKL...DN-V.........-   66
N.CM.06.U14842                ML........FLGFI--G------V-I-VL.........LYK--K.......--KL-EE-KQ-RQ-IRS-----D---DE-AEWV.AT.....LL.FPTKL...DN-V.........-   75
N.FR.11.N1_FR_2011            ML........LLGFI-VG--F---V-I-VL.........LYR--K.......--KV-E--RHIRQ-IRD---------D-DEEL-.-T.....LL.PPDKL.DQGN-V.........-   77

P.CM.06.U14788                MH-RD.....EAVLIIAG-LLLCC--X-GK.........-LLLVL.......-ERERD-FVQRLA-WR-GQ--E-Y--NE-EEE...Q.....LR.-L-N....LLGF.DHV-...-   79
P.FR.09.RBF168                MH-RD.....EAVLIIAG-LL-CI----GK.........-LLFVL.......-ERERD-FVQRLV-WR--Q--G-Y--NE-EEE...Q.....LR.-L-N....LLGF.GNI-...-   79

CPZ.CD.06.BF1167              VAIFT.....V---FR-G...#G---I-I-.........ALLVI-KAYLYIKQVQEEDR-EQA-IDRYQ-RGSTDSGI-EDENG-.IN.....WM.DPDNY...FLL.........-   80
CPZ.CM.05.SIVcpzMT145         M.QLE......--L-ILFIALMLV--FA-IA.........AYK--K.......-LQQV-R-E--Q--IRS---------D-DEIL-VE-.....LM.QVHQ-QVNPD-M.DRI-FW.-   87
CPZ.GA.88.GAB1                MT........LLVGLV-IL...VGLIA-N-CIWGYIIKWGYRR-K.......RHRLETE-E--NLI-R----------N--EEERLEQ.....LI.HNYN-...NNHF.ANPMFDL-   89
CPZ.TZ.06.SIVcpzTAN13         VGNLEQNV.VG-LVIII-LVGGG-LIA-I-.........IKERN-.......ERQH--L-E--VR--S.....IDSGI-EDEEF..QWL.....NF.DPDN-.NPRD-I.........-   79
CPZ.US.85.US_Marilyn          MLNWF.....EIGLI--GIEG-LVVII-GL.........-ARLW-Q...IKI-ENT-QE-QN-LE-IRI-E-------D--EEETLAK.....LLSSLELD...N-RIV.........-   85

GOR.CM.07.SIVgor2139_287      MH-RD.....-IV-IIGITLLAVTVII-LK.........IFAL-L.......RDR-E-RFFDRLE--LSNK--E-Y--NE-E..A.A-.....LM.---N....ELGF.DFN-H..-   80
GOR.CM.07.SIVgorCP2135con     MH-RD.....-IV-IIGITLLAATVI--LK.........ALAL-L.......RDR-E-RFFNRLE--LSTK--E-Y--NE-E..A.A-.....LM.-R-N....ELGF.DFN-H..-   80
```

```
                            Env start                    signal peptide end   gp120 start                                                  glycosylation NVT                                  CD4 binding   glycosylation NDT / V1
B.FR.83.HXB2                MRVKEKYQH.L.WRW....GW.RWGTM...LLGM.LMIC.SA...T.EKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVL.VNVTENFNMWKNDMVEQMHEDI.ISLWDQSLKPCVKLT.PLCVSLKCT.DL...KNDTNT....... 140
A1.AU.03.PS1044_Day0        --A-GIQMN.-.HC.....LL.K----...I---.IL--.--...A.-QR------------D-E----------------------------------IN-.X----E------N-----Q---.---------------.----T-N-S.HE...VIFNS-LNNSTHS 146
A1.CH.03.HIV_CH_BID_V3538   ---MGMQRN.Y.PH.....W-.K--IV...-FWL.-I--.N-...-.---------------D-V--------------------------------H-IP-.A----E------N------T--.--------------A.----T-E-N.-A...NRTNIN....... 139
A1.CY.08.CY236              ---RGIQMN.S.HC.....LL.-----...I---.II--.--...V.--T------------D-E------------T--A--I---------------IN-.E----E-----LN------T--.---------------.----T-D-R.HN...VTIH-N...VTNN 143
A1.ES.05.X1608_8            ----GIQMN.S.Q-.....LL.-----...I---.II--.--...A.-N-------------D-E---------------K------------------IR-.D----E------N------T--.------------E--.----T-N--.EA...Y---DV....... 139
A1.KE.11.DEMA11KE001        ---MGTLMN.W.R-.....FV.---I-...V--..II--.--...A.QN-----H-------D-D-I----------E--K----------------R--Y-.E----D------S------T--.---------------.----T-N-S.NA...NVSVSN...VSVT 142
A1.RU.11.11RU6950           -KARGMKRN.Y.QH.....L-.----W...-FW-.IIM-.K-...A.-H------------RD-E-------------K----------------D---IM-.E------D----N-----QT--.------------Q--.----T-H-A.GS...NSTSSN...NSTS 143
A1.RW.11.DEMA111RW002       -K-MGIQRN.Y.GQ.....L-.-----...I---.II--.--...-.-N-------------N-E-------------K---------------------Y-.E----E------N------T--.------------Q--.----T-N--.NN...ITVN......... 137
A1.SN.01.DDI579             ---MGIQRN.Y.LP.....L-.---MI...S-C-.II--.K-...-.-N-------------D-E-------------Q--------------------LD-.E----E------N------T--.------------Q--.----T-N--.-I...PKNFTN....... 139
A1.UG.11.DEMA110UG001       ---MGIQMI.W.QN.....L-.-----...I---.-I--.-V...A.GN------------RD-E----------------------------A-----IE-.K----K------N-----A---.---------------.----T----.--...DLNGTQ...SLNG 143
A1.ZA.04.04ZASK162B1        ---M-TQRN.Y.P-.....W-.I--I-...I---.II-G.R...S.DN-------------D-D------------SK--------------------MN-.E----K------G------T--.--------T------.----T-N--.NS...NFTR-S...NLTD 142
A2.CM.01.01CM_1445MV        -K-MGTQRN.C.QH.....L-.---IL...IF--.-I--.K-.....TN------------RD-E-------------R---------------------D-.K-----------N---------.-N------M------.----T-N-S.EV...-DN-......... 136
A2.CY.94.94CY017_41         ---MGTQRN.Y.QH.....L-.-G-IL...I---.-IM-.K-......D-------------D-D-I--------------------------------IN-.E-----------N-----Q---.---------------.----I-N-S.NA...NTS-HS...NS.. 140

B.BR.10.10BR_MG029          -----IRKN.Y.QX.....Y-.---IL...----.----.--...-.DR---------------I------------------A----------------R-.E-------V---N--D------.---------------.----T-N--.---...YANNTA...NGQN 143
B.CA.07.502_I191_03         ----GIRRN.Y.QHSWRGGI-.---I-...----.----.---.AE.DQG--------------K----------G----A--------------S-----M.E----E--------A-------.SN-------------.----T-N--.NA...NVSET-...Y... 146
B.CH.08.M2_0803101_NFLG8    -K--GTKKN.Y.QH.....L-.---M-...----.----.S-...A.DQ-----------------------------------------------------.G-----------N---------.---------------.----T-N--.-D...LKNN-N...SSSN 144
B.CN.10.DEMB10CN002         --A-VIRKS.Y.QP.....W-.I----...----.----.--...K.-----------------I-P----------A--A-------------------D-.E----D------N---------.-----E---------.----T-N--.-W...-PNNS-...TNNN 143
B.ES.10.DEMB10ES002         ----GIRKN.Y.QNL....W-.I---L...---I.----.-T...A.-N-----H---------S------------------------------S----L-.E-----------N--Q---D--.-----E---------.----T-N--.KV...SCTNS-...SSNN 144
B.FR.11.DEMB11FR001         -K--KTR.....KH.....L-.---IL...---I.----.--...A.-Q---------------N-------------------------------------.G----D------N---------.---------L-----.----T-N--.G-...N-A--I...... 136
B.GB.05.MM45d213_GN1        -----IMKS.W.QH.....W-.KG-IL...---I.----.-N...A.-N-------------------------------A------------------M--.E----E-D----N-----Q---.T--------------.----T-N--.-A...DVT-TN...ST.. 141
B.HK.06.HK003               -K--GIRKT.C.Q-.....F-.K----...---I.----.--T..K.-----------------N---------------A---------------------.G----E-DI---N---------.-N-------------.----T-N-S.NN...ITI-TS...TNR. 143
B.HT.05.05HT_129389         ----GIRKN.Y.QH.....W-.T---L...----.----.--...A..-M---------------------------E---------------------IE-.K-----------N---------.-----E---------.----T-N--.KP...V-C-KE...ENGT 142
B.JP.12.DEMB12JP001         -K--GTRKN.C.QH.....L-.K--I-...----.----.--.VEK.DN-------------------------------A-------------------R-.E-------I-E-N--------V.----E----------.----T-N--.-A...G-T-EA...GNT. 144
B.KR.07.HP_18_07JHS10_3909  ----GIRKN.Y.QH.....W-.---I-...----.W---.--...-.DEM------------------------------------------I-------E-.------------N---------.---------------.----T-N-S.N-...R-E--I...KNA. 141
B.PE.07.502_0525_wg5        --ARGIMRN.Y.QH.....L-.-----...----.----.N-...A.DQ--------------TA--------------P-------------------IL-.E----D------N---------.---------------.----T-N--.-K...LK.......... 135
B.RU.11.11RU21n             ---M-IR.....KH.....L-.----W...----.----.--...-.QN----------------------------SK-A------Y-------S---IP-.K--------G--N--D------.------------R--.----I-N--.NV...AS.......... 132
B.TH.08.MERLBDTRC10         -KA-GIRKN.Y.QHL....W-.K----...----.----.--...-.KQ-------------D-S---------------A------------------IA-.G----K----E-N--D------.------------R--.----T-N--.-V...WRN-TN...... 140
B.US.11.ES38                ---MGIRKN.Y.QH.....W-.-----...----.----.--...-.-----------------N----------G-KK-A--I-----------S---IK-.E---------------------.------------S--.----T----.NA...TTS--V..T.... 140

C.AR.01.ARG4006             --A-GIQRN.W.KQ.....W-.I--IL...GFC-.X---.NV...R.GN---------------K----------G-----------------------L--.K---------E----D---Q--.---------------.----T-N--.-A...NRTISX...TISN 143
C.BR.07.DEMC07BR003         ---MGTQRN.W.KH.....W-.I--IL...GFW-.F-L-.NV...M.GN-------------D-K--------------------------------H-MI-.E--------------D--QQ--.---------------.----T-N--.NA...TSTPLCVTPNCTN 146
C.BW.00.00BW5031_1          ---MGTQRN.C.QQ.....W-.I--IL...GFW-.---Y.NV...G.GN---------------K------------EK----------------S---I--.E----T---------D------.---------------.----T-N-S.AA...N-........... 135
C.CN.10.YNFL19              ---R-ILRN.Y.QQ.....W-.I--VL...GFW-.----.-V...G.GNR------------D-K----------G-EK--------------------M--.E--------------N-----V.---------------.----T-N--.NV...T-VNSN...STNN 143
C.CY.09.CY260               ---MGIMRN.C.Q-.....W-.I--IL...GFW-.----.NG...G.D----------------KA------E-R--ER--------------------L--.E--------------D------.-----E---------.----T-N-S.-Y...-GTH-N...TTSN 143
C.ES.08.X2363_2             ---RGILRN.W.QH.....W-.I--IL...GFW-.---X.NV...W.GN---------------K-------------X-A--------------D---L--.D-A--Y--------XD#-L---#---------------.----T-H--.NYLIPPT..IN...XTIT 143
C.ET.02.02ET_288            ---RGIQRN.C.QQ.....W-.I--LL...GFW-.F---.NV...V.GN-------------D-SP---------------------Y-----------L-M.E------------------Q--.-----E---------.----T-H-SSNV...TS.NSN...VTSN 143
C.IN.09.T125_2139           ---M-ILRN.Y.QQ.....W-.I--IL...GFW-.----.NV...G.GN---------------K-----------HE-----------------K---I--.G--------------D-----V.---------------.----T-E-G.NV...SVSNVS...VSNT 143
C.KE.00.KER2010             ---RGIQRN.W.QQ.....W-.I--IL...GSW-.--T-.-V...E.GN---------------KAP----------EK----------------S---I--.E--------------D------.---------------.----T-N-S.NV...TSNATS...NSTN 143
C.MW.09.703010256_CH256.w96 ---RGIQRN.Y.QQ.....W-.I-SIL...AFW-.--V..-V...R.GN---------------K-------------K--------------------IK-.E--------------D------.---------------.----T-N-S.HN...ITVNG-...MGNG 142
C.TZ.08.707010457_CH457.w8  -K-M-ILRK.W.KH.....W-.I--IL...GFW-.-L-S.NV...M.GN------------R--K------------ER--------------------MR-.E--------------D------.---------------.----T-R-E.NA...TYNDTN....... 139
C.YE.02.02YE511             ---TGILRN.Y.QQ.....W-.I--IL...GFWI.----.NG...G.GN---------------K------------ES--------------------I--.E----S---------D--Q---.---------------.----T-N-N.NN...ITNNV-...ND.. 141
C.ZA.08.705010534_CH534.w12 ---MGIQKN.C.QQ.....W-.I--IL...GFW-.-L--.NV.....QSS--------------K-------------E----------------S---MF-.E----D---------D------.---------------.----T-E--.NA...TFN-......... 136
C.ZA.10.DEMC10ZA001         -K-MGIKRT.C.QQ.....W-.I--IL...GFW...--Y.NV...L.GN---------------K------------ER----------------D---I--.G---------E----D------.---------------.----T---G.NV...SRNVTN...TNSS 141
C.ZM.11.DEMC11ZM006         ---RGILRN.Y.PQ.....W-.I--IL...GFW-.V-S-.-V...R.GN---------------K------------EK--------------------IW-.E--------------D--N---.-----E---------.----T-S--.NF...TTNSTI...ITIN 143

D.CM.10.DEMD10CM009         --ARGIKRN.Y.QH.....L-.T--I-...---LL--TY.--...A.DN---------------K-----------TKA-A--I-----------S---II-.E----P----------------.---------------.----T-N--.ET...HLNSTQ...SVN. 143
D.CY.06.CY163               ---RGIKRN.Y.HN.....L-.K----...----L--TY.--...A.D----------------KA---------VTKA-A--I---------------I--GE------D--------------.-----E---------.----T-D-I.-V...ETNNK-....... 141
D.KE.11.DEMD11KE003         --A-GMKRN.Y.QH.....L-.G--I-...---I.--TY.-V...A.-Q--------------------------S-KE-A--I----------------L-.E-----------N---------.---------------.----T-N--.EW...-DNSTN...SNST 143
D.KR.04.04KBH8              ---RGMKRN.Y.QH.....L-.K----...----.W---.--...A.GN-------------------------NSLKA-A--I-----------S------.A-----LTC---N--D------.----E----------.----T-N--.-I...R-K-IN...N... 140
D.SN.90.SE365               --AR-MKRN.Y.QH.....L-.I----...----.--T-.-V...A.----------------------------S-E--K--I---------------IE-.A-----------N---------.---------------.----T-N-R.-I...SS-ATS...NT.. 141
D.TZ.01.A280                ---M-TQRN.Y.QH.....L-.---I-...----.W-TY.-V...A.-Q----I-------R--N-----------SF--A--I---------------MD-.---S--------N---------.----EE---------.----T-H-S.-A...NTTNSG...NG.. 141
D.UG.10.DEMD10UG004         ---R-MKKN.Y.QN.....S-.---I-...---I.--T-.-V.....GQ---------------K------------KP-A--I---------------IY-.E-----------N---------.---------------.----T-N--.KW...TSN-TS...NST. 142
D.UG.11.DEMD11UG003         ---RGIEKN.Y.RH.....L-.G---L...---I.-VTY.-V...A.-Q------------R-------------S-E---------------------IH-.E-----------N---------.---------------.----T-N--.EW...EPNATN...ANNA 143
D.YE.02.02YE516             ---RGMKRN.Y.QH.....L-.K---L...----L--TY.--...A.D----------------KA---------SVK--A--I--S------------IE-.K---------------------.---------------.----T-S--.-I...--V-VD...NN.. 142
D.ZA.90.R1                  --ARGMERN.C.QN.....L-.K--M-...----.----.--...A.-N--------------------------G-KK-A--I----------------G-.E---------R----N------.---------------.----T-N--.-A...IFNSTS...FS.. 141

F1.AO.06.AO_06_ANG32        ---MGMQRN.W.QN.....LG.K--LL...F--V.-I--.N-...A.-N--------------------------G-ER--------------------IL-.K---------Q-N------T--.---------------.----T-H-K.-A...NTSSTE....... 139
F1.AR.02.ARE933             ---RGMQRN.W.QH.....LG.K--PL...F--I.-I--.N-...-.-N-L----------R-------------S-ER----------------D-R----.E------D----N------T--.---------------.---XT-X--.-A...IR........... 135
F1.BR.10.10BR_RJ015         -K-RGMQRN.W.QH.....LG.K--LF...F--I.-I--.N-...E.-----------------N------------KK-A------------------Q--.G----K-D--N--------T--.-----E---------.----T-N--.NA...AR........... 135
F1.CY.08.CY222              ---RVTQRN.W.QH.....LG.---LL...F--I.-I--.--...A.DN--------------------------G-EK--------------------IH-.E-I--D------N---------.---------------.----T-N--.AK...NSTE......... 137
F1.ES.02.ES_X845_4          ---RGMQRN.W.QH.....LG.K--LL...F--I.-I--.N-...A.DN-----------------------N--G-ER----------------D---IW-.E------D----N------T--.V--------------.----T-N-S.-A...RDANSQ....... 139
F1.RO.96.BCI_R07            -.....QKN.W.QH.....WG.K--LL...F--I.-I--.--...-.D---------------------------G-EK--------------------ID-.------------N---------.-----E---------.----T-N--.NA...NGTLSA...NQSS 138
F1.RU.08.D88_845            ---MGMQRN.W.QH.....WG.K--LL...F---.-I--.N-...-.NN------------R----------------E-A-------------------F-.E----D-D--N-S---------.---------------.----T-N--.NA...TIPNGN...NN.. 141
F2.CM.02.02CM_0016BBY       ---RGMQRN.W.QH.....LG.K--FL...F--I.-I--.N-...A.DN----------------------------EK-A--------------D---IF-.D-----------N--D------.--------------I.----T-H-S.-V...NITA-N...T... 140
F2.CM.10.DEMF210CM001       --MR-MQRN.W.QH.....LG.---LL...F--I.-I--.N-...-.-N------------R--N-V--------G-ER--------------------M--.A----S----E-N--N------.--------Q------.----T-N--.NV...TTTN--...TYAK 143
F2.CM.10.DEMF210CM007       ---RGMQRN.W.QH.....LG.K--LL...F--I.-I--.N-...A.DN------------------V----E----EK----A---Y-----------L--.E--------------D--N---.T--------------.----T-S--.-A...NDTSAV...NS.. 141

G.BE.96.DRCBL               ----GIQRN.W.QH.....L-.N--IL...I--L.VI--.-....A.--------------ED-NAP---------HS--S--I-----------S---INM.R-----------N---------.-----E---------.----T-N--.EI...N-NSTR...NI.. 140
G.CM.10.DEMG10CM008         ----GIQRN.W.QH.....L-.T---L...IF-L.VI--.--...S.-N------------E-TD------------S--R--------------D---LF-.E-------A---N---------.T----E---------.---IT-H-N.NG...NLNYTD...ANGE 143
G.CN.08.GX_2084_08          ---RGIQRN.W.QH.....L-.I-WNL...I--L.VI--.-T...S.-N------------E--D-P-------R--S--K------------------IKM.D----P------N---------.-----E---------.----T-N--.AV...TCANA-...HGT. 142
G.CU.99.Cu74                ---RGIQRN.C.QH.....L-.K---L...II-L.AI--.--...S.XN----I-------ED-DAP----------S--S------------------IP-.E----K------N---------.-----E---------.----T-N-V.NV...SCTNS-...G... 140
G.ES.09.X2634_2             -K--GTQKS.W.QN.....L-.K---L...I--L.VI--.--...S.NN------------ED-D---------R--S--S------------------IP-.---S-E------N---------.-----E---------.----T-N--.NV...TTSSTS...TN.. 141
G.GH.03.03GH175G            ----GIQTN.W.QH.....L-.K---L...I--L.VI--.--...S.D-------------ED-D----------S-SS-K------------------I-M.K----Y------N---------.-----E---------.----T-N--.-I...VYSNC-...RKHP 143
G.KE.09.DEMG09KE001         --A--MKRN.W.Q-.....L-.EG-LL...I--L.VI--.--...S.IN--A---------ED-D------------S--K--I---------------LP-.T-----------N---------.-----E---------.---IT-N--.NV...TSTGIE...NNTE 143
G.NG.09.09NG_SC62           -K--GIQKN.W.PP.....L-.T-W-L...I--L.VI--.N-...A.DN------------ED-E-P----------S--K------------------IG-.E----Y------N---------.-----E---------.----T-N-V.T-...NCTDA-...YIGN 142
G.ZA.01.TV546               ---RGIQRN.W.QH.....L-.K---L...I--L.VIL-.N-...S.NN------------ED-DAP--------S-S--R--I---------------II-.K-----------P---------.-----E---------.----T----.-Y...NGTVTN...KT.. 141

H.BE.93.VI991               -....-TQRN.Y.PS.....L-.----L...I---.-L--.-V...V.GN---------------K---------------R------------------M--.E----T----V--------T--.---------------.----T-D-S.SV...NATNV-...KS.. 138
H.CF.90.056                 -....-TQRN.Y.PS.....L-.----L...I---.-L--.--...A.QN---------------K------------E--K------------------M-M.E----S----E-N------T--.---------------.----T-N--.NV...R-N-SN....... 136
H.GB.00.00GBAC4001          -....-TQRN.Y.PH.....LL.SG-IL...I---.-LM-.-T...G.-N---------------K------------S--K--------------S---M--.E---------E-N--D------.---------------.----T-E-G.EV...G-K-SN...SN.. 138

J.CM.04.04CMU11421          -T-M-TQRS.W.QN.....L-.K--L-...IF--.----.K-...A.D-M--------------K----------S-S--S------------------MII.A--------------D------.--V--E-------M-.----T-N-S.HN...ISTNS-...NSRT 143
J.SE.93.SE9280_7887         -....-TQKN.W.QT.....L-.-G-L-...IF--.----.K-...K.-D-------------D-K------------S--K--------------S---MN-.P--------------D--Q---.--V--E-------I-.----T-N-S.NI...TSNS--...TS.. 138
J.SE.94.SE9173_7022         -....-TQTS.W.LS.....L-.---L-...IF--.----.--...R.-N------------RD-K------------S--K------------------MS-.P--------------D--Q---.--V--E-------I-.----T-N-S.-V...NSNNS-...DSN. 139

K.CD.97.97ZR_EQTB11         --AR-IQRN.W.QH.....LG.KR-IL...F--I.-I--.--...A.NN----------------------------E------------------------.E-----------N------T--.-----E---------.----T-T--.NV...T-NRTN...ANKN 143
K.CM.96.96CM_MP535          ---RGMQRN.W.QT.....LG.N--IL...F--I.-I--.-N...A.D-----------------P-----------EK---------------------EM.E-----------N------T--.-----E------E--.----T-N--.-Y...-GTNS-...NN.. 141

U.CA.01.TV749               ---MGTQRN.Y.QS.....L-.---L-...IF--.----.NTK..G.-N--A---------R--KA--------QS-S--K------------------IX-.K-------V-------------.-----E---------.----T-N--.-I...TSSNA-...IST. 143
U.CA.99.TV721               ---M-TKRN.Y.Q-.....L-.---L-...IF--.----.KI.EGK.DN---------------K---------QS-S--K--------------D---ID-.K-------A------D------.-----E---X-----.----T-T--.SV...A-NIT-...TNVT 145
U.CD.83.83CD003_Z3          -----IQRN.Y.QH.....L-.K-SLI...I---.I---.K-...I.--S------------D-E------------EK-S--------------S---L--.G-----------K---------.---------------.F---T-N-I.-V...--S--N...N... 140
U.CD.90.90CD121E12          --A--IQRN.Y.HY.....L-.K-SLI...I---.II--.KG...E.-QS------------D-D----------G-EK-A--------------S---L--.------------N---------.--------------S.S---T-N--.-V...GTNNTA....... 139
U.CY.05.CY090               --A--TQKN.W.QH.....LG.K--IL...F--I.FI--.-TAEKA.-N---------------K------------SP-A------------------I--.E----Y-------------K-V.---------------.----T-D-V.NI...TANITR....... 142
U.CY.08.CY223               ---MGTQRN.C.LL.....SL.------.CIFWL.M-V-.N-.....N--------------D-K-------N-Q--N--K------------------ML-.I--------------D------.-----E---------.----T-N--.NA...RANKV-...VNAT 143
U.ES.10.DEURF10DZ001        --ARGIQTN.W.QN.....L-.K---L...I--L.VI--.--...-.NN------------RD-E-------------Q------------------R-IP-.K-------A---N---------.---------------.----T-N-K.PF...NSN-SG....... 139
U.GR.99.99GR303             --L-GIQRN.C.QN.....L-.--AL-...IF--.----.K-...-.DN-------------D-E-----------HSA-K---L----------S---MI-.G-----------N------T--.---------------.----T-N-K.NP...SCVALN....... 139
U.NL.95.U_NL_95_H10986_D1   ---RGMQRN.W.RH.....LG.---IL...F--I.----.--...A.D-----I-------QD-N----------S-SN-R------------------LP-.------------N----V----.-----E---------.----T-N--.-V...T-R--S....... 139

01_AE.AF.07.569M            ----GTQRN.G.PN.....L-.K---L...I--L.VI--.--...S.DN------------RD-D-----------HE-----------------S---IY-.E-----------N-----Q--V.---------------.----T-N-S.NA...NLTNIN...NTIT 143
01_AE.CF.90.90CF11697       -G--GTQMN.W.PH.....L-.K---L...I--L.VI--.--...S.DT------------RD-D---------N-QE----------V----------IH-.E------D----N-----Q--V.---------------.----T-H--.KA...-LNDTY...NGTA 143
01_AE.CN.10.YNFL03          -----TQMN.W.PS.....W-.K---L...I--L.V---.N-...S.NN------------RD-N-----------HE--------IY-----------IH-.------------N-----Q---.-----E---------.----T-T-I.NA...NLTKL-..ANASD 144
01_AE.HK.04.HK001           -----TQMN.W.PN.....L-.----L...I--L.VI--.--...S.DN-------------D-D-----------HE---------------------IK-.GD--------X-K-A---Q--V.---------------.----T-N--.QA...NWKS-N...TTQN 143
01_AE.IR.10.10IR.THR48F     -K---TQRN.W.LN.....W-.K---L...I--L.VT--.--...S.DN------------RD-E-----------QE--A------------------IY-.E-----------P-A---Q--V.---------------.----T-N--.KA...NLTKLA...NASN 143
01_AE.JP.x.JRC77AE          ----GIQMN.W.PN.....L-.----L...I--L.GI--.--...S.D-------------RD-D-----------HE---------------------IH-.E-----------N-----Q--V.---------------.----T-N--.EA...NWTNVS...STSN 143
01_AE.TH.09.AA090a_WG11     ---RGTQMT.W.PN.....F-.---IL...I--L.VI--.--...S.DN-------------D-D-----------HME-A--I---------------IH-.E-----------N-----Q--V.---------------.----T-N--.NA...DLRYAD...NSTN 143
01_AE.TH.90.CM240           -----TQMN.W.PN.....L-.K---L...I--L.VI--.--...S.DN------------RD-D-----------HE---------------------IH-.E-----------N-----Q--V.---------------.----T-N--.NA...NLTNGS...SK.. 141
01_AE.US.05.306163_FL       ----GTQMN.W.PN.....L-.K---L...I--L.AI--.--...S.NN------------RD-E-----------HE---------Y-----------MH-.E-----------N-----Q--V.---------------.----T-N--.-F...-VP.......... 136
01_AE.VN.98.98VNND15        -----TQMN.W.PN.....L-.K---L...I--L.VI--.--...S.D-------------RD-E-----------HVP--------------------IP-.E-----------N-----Q--V.---------------.----T-N--.SA...TLPNSN...ANL. 142
```

```
                             Env start            signal peptide end   gp120 start                                              glycosylation NVT                               CD4 binding   glycosylation NDT  V1
B.FR.83.HXB2                 MRVKEKYQH.L.WRW....GW.RWGTM...LLGM.LMIC.SA...T.EKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVL.VNVTENFNMWKNDMVEQMHEDI.ISLWDQSLKPCVKLT.PLCVSLKCT.DL...KNDTNT....... 140
02_AG.CM.10.DE00210CM013 -K-MGMKKN.Y.PL.....L-.-L-MI...IFWI.II--.N-.....-N-------------D-K----------------------------------ID-.K---------------------.------------Q--.----T-N-S.-I...IFRSSS...INNS 142
02_AG.CY.09.CY256 -T-TGTQKS.Y.P-.....L-.-L-MI...IFWI.II--.N-.....KQ------------RD-E---------Q-----A--------------T---IY-.E-----------N------K--.--------Q------.----T-N-S.-A...NVNSTY...I... 139
02_AG.ES.06.P1423 ---RGIQRN.Y.PP.....LL.---II...I-WI.IIF-.N-.....QN------------RD-E------------H--A------------------MN-.K-----------N---------.-N-------------.----T-R-H.--...N-TNI-....... 138
02_AG.GW.05.CC_0048 ---MGIQKN.W.PF.....L-.K--II...IFWI.II--.E-...N.R--------------D-K-------------A--------------------LP-.E----D------N---------.---------------.----T-E-S.NV...NRTNVN...GTIQ 143
02_AG.KR.12.12MHR9 ---RGIQKN.Y.PL.....L-.----I...IFWI.MI--.N-.....-N------------RD-E----------------------------------IQ-.E----D------N------V--.---------------.----T-D-H.NF...NSSNSS...T... 139
02_AG.LR.x.POC44951 ---MGIQKN.Y.PL.....L-.-L--....IFWI.II--.N-.....NNM-----------RD-E-------N--T----A-------------------L-.E-----------N--D------.---------------.----T-D-H.NI...NSS.......... 134
02_AG.NG.09.09NG_SC61 -...-TQRN.W.LL.....L-.-G-LI...MFWI.MI--.N-.....QP----I-------R--E------------VK-S------------------IY-.K-----------N--D------.-----E---------.----T-E-N.NF...NSSN-S...TSA. 138
02_AG.NG.x.IBNG ---MGIQKN.Y.PL.....L-.----N...IFWI.MI--.N-.....-Q-------------T-E----------------------------------IH-.E----K------N---------.---------------.----T-D-H.NF...N-SYSN...S... 139
02_AG.SE.94.SE7812 ---MGIQKN.Y.PL.....L-.---MI...IFWI.MT--.--.....GN------------RD-E----------------------------------IH-.G----D------S---------.------------Q--.----T-H-Q.-N...LTSS......... 136
02_AG.SN.98.98SE_MP1211 ---MGILKS.C.PP.....F-.---MI..M--WI.-I--.N-.....-N-----------RRD-E------------E------------------S--MH-.E-----------N------V--.---------------.----T-E-H.NY...NYTR-N...S... 140
03_AB.RU.97.KAL153_2 -----IR.....KH.....L-.----L...F---.----.--.....-N----------------------------SK--------Y-------S---IP-.K--------G--N---------.---------------.----T-N--.--...-KEVTS....... 136
04_cpx.CY.94.94CY032_3 ---MGMQRN.Y.PH.....L-.E---L...I--L.VI--.--...S.NN------------RD-E-------E----EK----I----------------A-.I---------------------.----NEG----A---.S---TFT-I.NA...TTTNS-...N... 140
05_DF.BE.x.VI1310 ---RGMQRN.W.PH.....LG.K--LL...F--I.-I--.--...-.D-F-----------R--K----------G--K-----------------------.G------------------T--.---------------.----T-N--.-F...-ANSTA...NS.. 141
06_cpx.AU.96.BFP90 ----GIQTS.W.QH.....L-.K---L...I--L.VI--.--...S.KNM---------A-ED-D-I----------SPDK------------------IS-.K-----------N--D------.-----E---------.----T-T--.NA...TLGNK-...LGNN 143
07_BC.CN.98.98CN009 ---TGIKKN.Y.RH.....L-.-----...----.----.--...V.GN--------------------------------------------------MF-.E-----------E--N-----V.---------------.----T-E-K.NV...SSNS-G...TC.. 141
08_BC.CN.97.97CNGX_6F ---RGTRRN.Y.QQ.....W-.I--VL...GFW-.----.NV...E.GN---------------K------------E-----------------S---I-M.E---------N----N-----V.T--------------.----T-E-R.NV...SSNG-G...TY.. 141
09_cpx.GH.96.96GH2911 -----IQRN.W.QN.....L-.---L-...I---.----.N-...A.-N---------------E---------Q--SK-K------------------IG-.E-----------N------K--.---------------.----T-D-H.RV...N-SKPN....... 139
10_CD.TZ.96.96TZ_BF061 -T-TGIQRN.C.RQ.....W-.I--IL...GFW-.----.N-...R.----------------T-------------KA-A--I-------------R-I--.E-----------G--D------.------G----I---.----T-N-N.GI...NATND-....... 139
11_cpx.CM.95.95CM_1816 -----TQRS.W.HN.....L-.---L-...IF--.----.N-...-.DN------------RD-D-P-----N----S--K------------------IH-.E---------E-N---------.-----E-----I---.----T-N--.NV...-KNCTN...VEKN 143
12_BF.AR.99.ARMA159 ---RGMQRN.W.QH.....LG.N--FL...F--I.-I--.N-...-.-N--------------------------S-ER-----------------------.E------D----N------T--.---------------.----T-N-S.NA...NATA-N...ATAN 143
13_cpx.CM.96.96CM_1849 --ARGIQMT.W.QN.....L-.K---L...I--L.VI--.--...S.NN------------RD-E------------S--K------------------IN-.G---------E-N------T--.--I---------S--.----T-N-N.NI...TS-SSI....... 139
14_BG.ES.05.X1870 -KA-GTQRN.W.QS.....L-.---AL...I--L.VI-S.G-...S.ND---------------------------------------------------N-.----------E-N--D--Q---.---------------.----T----.-F...NGTNTN...NTDV 143
15_01B.TH.99.99TH_MU2079 -----TQMN.W.PN.....L-.K---L...I--L.VI--.--...A.-N-------------D-S---------------A--G-----------S---EI-.G---------E-------Q---.---------------.------N--.--...R----I....... 139
16_A2D.KR.97.97KR004 ---RGIQRN.Y.QH.....L-.K--IL...I---.---S.K-...-.-D------------RD-E---------------A------------------IN-.E-----------G---------.---------------.----T-N-S.RV...--TISS...TQS. 142
17_BF.AR.99.ARMA038 ---RGMQRK.S.QH.....L-.K-----...---I.----.NT...A.GT-----------------I--------G-E---------------------MP-.E------D--N-N------T-V.---------------.----T-N-N.-V...NVTATN....... 139
18_cpx.CU.99.CU76 ---RGTQRN.Y.PS.....L-.K---I...I---.-LM-.--...-.DH-------------D-V----------S-S--S---------------D--IT-.G-----------N---------.-----E---------.----T-N--.NV...XXSG-N...TXAS 143
19_cpx.CU.99.CU7 ---MGTRRS.Y.PP.....L-.N---L...I--L.VIM-.--...F.DN------------RD-E------------------------------S---IN-.E----E------N------A--.---------------.----T-D-H................... 131
20_BG.CU.99.Cu103 ----GIQTN.W.QC.....L-.T---L...II-L.GI--.--...S.NN------------ED-N-P----------S--S------------------IP-.E-----------N-----N---.-----E---------.----T-N--.N-...SVNSTV....... 139
21_A2D.KE.99.KER2003 ---R-IERN.C.QN.....L-.K--I-...---I.--T-.--...A.GN---------------N------------S--K--I---------------ME-.G-----------N---------.---------------.----T-N-N.-Q...VRNSEG....... 139
22_01A1.CM.01.01CM_0001BBY ---RGTQMT.W.QN.....L-.----L...I--L.VI--.--.....QN------------R--E-------------P----------------S---IE-.A-----------K-----Q---.---------------.----T---E.-A...RFNITN....... 138
23_BG.CU.03.CB118 ----GIQKN.W.Q-.....L-.T---L...II-L.GI--.--...S.DN------------ED-Y-P----------S--S------------------IP-.X-----------H---------.-----E---------.----T-N--.KV...DCE-RN...CTT. 142
24_BG.ES.08.X2456_2 --ARGIQMN.W.QL.....L-.T---L...II-L.GI--.--...S.DN----I-------ED-DAP--------S-S--S------------------IS-.Q-----------N---------.-----E---------.----T-S--.NV...NSTR-N...HVNS 143
25_cpx.CM.02.1918LE -K-RGIQKN.Y.PL.....L-.---IL...M--L.VI--.--...S.D--------------D-K------------K--A--I---------------IDI.E---------------------.-----E------Q--.----T-N-S.NI...NINSTG...N... 140
26_AU.CD.02.02CD_MBTB047 ---MGTRKN.C.QH.....LL.K---I...I--V.I-F-.N-...A.D-------------RD-E----------------------------------MN-.T----Y------N------A--.---------------.----T-N--.-I...TTNRTS....... 139
27_cpx.FR.04.04CD_FR_KZS -K-MGIQRN.W.QG.....L-.---L-...I---.I---.-V...Q.GN---------S--R--D-----------HE--A------------------IL-.D---------R-N-A-------.-----E------R--.-F--T-N--.-V...NTTRIH...DS.. 141
28_BF.BR.99.BREPM12609 ---RGMQRN.W.KH.....LG.T--LL...F--I.-I--.N-...-.NQS------------------------N-VK--A--I-----------I------.K-----------N---------.-----E-I-------.----T-N--.NN...VTV--M...NATN 143
29_BF.BR.01.BREPM16704 --A--TMKT.Y.-H.......S--M-...----.----.--.....-Q---------------S------------V--A------------------L--.E---------------------.-----E---------.----T---I.N-...NATNLN...DAST 141
31_BC.BR.04.04BR142 ----GTQRN.W.KQ.....W-.I--IL...GFW-.FL-Y.NV...R.GN---P--------RD-K------------E---------------------I--.Q---------E----D--Q---.---------------.----T-H-S.-A...NR........... 135
32_06A1.EE.01.EE0369 ---RGTQMN.W.SH.....L-.K---L...I--L.VI--.N-...S.KD----------A-ED-D-I---------HS--R------------------IH-.K--I--------H---------.K----E---------.----T-H-K.NA...TFKNA-...INGT 143
33_01B.ID.07.JKT189_C -----TQRN.C.SN.....L-.K---L...I--L.VI--.--...S.NN------------RD-E-----------HER--------------------MYM.E-----------K-AD--Q--V.--------Q------.----T-N-A.YA...NLTKSN...NTVN 143
34_01B.TH.99.OUR2478P -----TQMN.W.LN.....L-.K---L...I--L.VI--.--...S.DN------------RD-D-----------HE---------Y-----------IH-.E-----------N-----Q--V.---------------.----T-N--.NA...ALTNV-...I... 140
35_AD.AF.07.169H ---MGTQTN.W.QN.....L-.-----...I---.II--.--...V.-N---------------E------------E-------------------R-IH-.Q----E------K---------.---------------.----T-T--.NA...NVN--N...DSN. 142
36_cpx.CM.00.00CMNYU830 --ARGTQRN.C.LH.....L-.K---L...II-L.VI--.--...S.DL------------RD-N----------------------------------ID-.------------N-----Q---.---------------.----T-E-Q.NF...NGSN-H....... 139
37_cpx.CM.00.00CMNYU926 ---MGIQRN.Y.PP.....L-.N---L...I--L.VI--.--...N.NL-------------D-D---------T--VP--------LP#-----#--KINV.G---KD------N-----QI--.----GPK----E---#----TF-**.P*G#.TPK--G....... 135
38_BF1.UY.03.UY03_3389 ---RGMQRN.W.QH.....LE.KCSLL...F--I.-I--.N-.....------------------------------E----------------------F-.K----E-D----N------T--.---------------.----T-N-S.-A...SAPKT-....... 138
39_BF.BR.04.04BRRJ179 ---RGMQRN.W.QH.....LG.---LL...F--I.-I--.--...A.DN----------------------------ER----------------S-H-EF-.E-----------N------D--.---------------.----T-N--.-A...VRA--G....... 139
40_BF.BR.05.05BRRJ055 ----GIRKN.C.QP.....L-.---I-...----.----.N-...-.NNW----------------------------------------------------.G-----------N---------.---------------.----T-H-A.AF...NGTQ-S...TQGG 143
42_BF.LU.06.luBF_18_06 ---MGIMKN.Y.QH.....L-.K--I-...---I.---S.--...A.-E--------------------------G------------------------R-.------------N---------.---------------.----T---A.-A...TPHNAN...SNTT 143
43_02G.SA.03.J11223 -----IQRN.W.QY.....L-.TMWXL...I--L.XI--.--...S.NN------------ED-X-P-----N----S--K------------------MT-.E----P------N------D--.-----E------M--.----T-T--.NA...TCNKTN....... 139
44_BF.CL.00.CH80 ---RGMQRN.W.QH.....LG.K--LL...F--I.-I--.N-.....----------------------------------------------------ID-.P-----------N------T-V.V--------------.----T-E--.NA...TATNHS...D... 140
45_cpx.FR.04.04FR_AUK ---R-TQRN.W.QH.....LG.---IL...F--I.-L--.--...A.DN-------------D-Q---------------K------------------ID-.-----R------N------T--.------------Q--.----T-N-N.-I...T-RNG........ 138
46_BF.BR.07.07BR_FPS625 ---RGMQRN.W.QH.....WG.K--LL...F--I.-T--.-T.....--------------R----------------K-A------------------I-M.G------D----P------T--.---------------.----T-E-N.AA...-TND......... 136
47_BF.ES.08.P1942 --AR-IKKN.Y.QQ.....W-.KG-IL...---I.----.N-E.KS.---------------------------------A------------------ML-.K-----------N--D------.---------------.----T-N-N.NT...VTTNAS...M... 142
48_01B.MY.07.07MYKT021 -----TQGN.W.LN.....W-.K---L...I--L.VV--.--...S.DN------------RD-E-A---------QVE--------------------IH-.E-----------S-----Q--V.---------------.----T-N-S.EK...ITISGN...GN.. 141
49_cpx.GM.03.N26677 ---M-TWMI.C.#N.....W-.-G-LL...I---.----.-G...A.GN---------------K------------SK-A--------------H-R--K-.E-------V-------------.--V--E---------.----T-N--.NV...-TNNS-...ANNT 142
50_A1D.GB.10.12792 ---R-IERN.Y.QQ.....C-.---M-...----.--T-.-V...A.-Q--------------------------S-KA-A--I---------------ID-.E---------------------.---------------.----T-Q-S.NI...N-A--N....... 139
51_01B.SG.11.11SG_HM021 ----GIRKN.W.QH.....L-.-G-IL...F--I.----.N....A.---------------------------R------------------------LP-.E-----------Y-A---Q--V.---------------.----T-N--.TA...-FE--S...TT.. 140
52_01B.MY.03.03MYKL018_1 -----TQMN.W.PH.....L-.K---L...I--L.VI--.--...S.-N-------------D-D-----------HE--A------------------IH-.E-----------K-A---Q--V.---------------.----T-N--.NA...TVADHN...KNK. 142
53_01B.MY.11.11FIR164 -----TQMN.W.PN.....L-.K---L...I--L.VI--.--...S.D-------------RD-D-----------HE-----I---------------IH-.E------X----K-A---Q--V.-D-------------.----T-T--.NA...NLTNIN...TTM. 142
54_01B.MY.09.09MYSB023 ----GTQMT.W.PS.....L-.T--IL...MI-L.VI--.--...S.NN------------RD-E-----------HE--M--I---------------IL-.E-----------K-AD--Q--V.---------------.----T-H-A.SA...NVTGKN...IT.. 141
55_01B.CN.10.HNCS102056 -K-RVTQMS.W.LN.....L-.I---L...IT-L.VIM-.-T...S.DN------------RD-E-----------HV--M------------------IY-.E------K.---N--D-----V.----E----------.----T----.NA...-FNNT-...GG.. 140
56_cpx.FR.10.URF5_patient_A ---RGIKRN.W.PH.....L-.T-W-L...I--L.VI--.-T...S.TN-------------D-E---------R--------------------S----K-.E-----------N---------.---------------.----T-N--.NA...NVTSIS...NS.. 141
57_BC.CN.09.09YNLX19sg ---RGTLRN.C.QQ.....W-.I-SIL...GFW-.--S-.NV...G.GN------------R--K------------E-----I-------------R-IW-.KK-------------D--Q--V.V-----------RM-.H---T-E-K.NV...TYCK-D...TTNI 143
58_01B.MY.09.09MYPR37 -----TQMN.W.PN.....L-.K---L...I--L.VA--.--...L.DN------------RD-D-------------E-----------------S-R--M.E-----------N---------.-----E---------.----T-N--.KA...NFVKDN...TTTN 143
59_01B.CN.09.09LNA423 --AMGIRKN.C.QH.....S-.K---V...----.----.-T...A.-Q------------R--N------------E--K---------------------.A---------E-------Q---.------------E--.----T-N--.NY...N-TK-V....... 139
60_BC.IT.11.BAV499 ----RMQRN.W.IT.......W--IL...GFW-.--MY.N-...-.--------------R--K-------Y--------------------------LL-.E---------E----D--Q--V.---------------.----T-N-S.NV...NRTE-N...AT.. 139
61_BC.CN.10.JL100010 ---RGILKN.Y.PQ.....W-.I--IL...GFW-.---G.NG...G.KG------------RD-K------------E---------------------II-.E--------------N--Q--V.---------------.----T-N--.-V...------....... 135
62_BC.CN.10.YNFL13 ---MGTRKT.C.QH.....L-.-----...---I.----.--...A.-N---------------S------------------------------S---LF-.E--------------K-----V.----E----------.----T---S.-V...S-........... 135
63_02A1.RU.10.10RU6637 ---MGIVKN.Y.QH.....L-.---I-...-FW-.IIM-.K-...A.-N-------------D-K-------E---H---M--I---------------TH-.A-----------N---------.---------------.----T-D-Y.EC...RSNVSN...S... 140
64_BC.CN.09.YNFL31 ---MGILRN.C.QQ.....W-.I--IL...GFW-.-I--.NV...G.GN-------------D-K-----------HK---------------------M--.E---------R----D-----V.--I------------.----T-D-Q.-Y...NGTAGN...YTST 143
65_cpx.CN.10.YNFL01 ---RGTQMN.W.PN.....L-.----L...I--L.VI--.-T...S.DN-------------D-S-------E------------------------R-IN-.E--------------N--Q--V.------G--------.----T---G.NV...TVRKYN...ESVN 143
67_01B.CN.11.2011.ANHUI.HF115 ----GTQMX.W.LN.....L-.K---L...II-L.VIT-.--...S.DN------------RD-D-----------HE---------------------IP-.------------S-A---Q--V.--------X------.----T-N-S.NA...XLNXT-...N... 140
68_01B.CN.11.2011.ANHUI.WH73 ----GTXMN.W.LN.....L-.K--IL...II-L.VIT-.R-...LDNN----I-------RD-E---------R-HE---------------------IP-.------------S-A---Q--V.---------------.----T-H-S.NA...-LNVTA...N... 141
70_BF1.BR.10.10BR_PE004 ---RGMQRN.W.QH.....WG.T-SFL...F--I.-V--.NT...A.DN------------X--S------------XX-----------------------.G-----------N---------.---------------.----T-H-N.NX...TKSGTN...S... 140
71_BF1.BR.10.10BR_PE008 ---MGMQRN.WQQQ.....LG.K--LW...F-RT.-L--.--.....-N-----------------------------A-A--------------S----K-.K------------I----D---.---------------.----T-N--.-W...G-T-SN...N... 141
72_BF1.BR.10.10BR_MG002 -KAMGIKRN.Y.QH.....W-.KG-IL...---I.---S.--...-.D----------------------------N----------------------ME-.K------------------A--.---------------.----T-N-S.NH...G-A-QS....... 139
O.BE.87.ANT70 -KAM--RNK....K.....L-.TLYL....AMAL.ITP-.LS...L.RQ-YA---A-----ED--PV-------NLTS--K--I--SQ-------T-Y-YP-.H---DD--I---Y-----Q---.------------QM-.F---QME--.NI...AG........... 132
O.CM.98.98CMA104 -KAMK-KNK....-.....L-.ILYI....VM-L.ITP-.LS...S.SQFYA---A-----EN-NPV-------NLTSP-Q--I--AQ---------H-YP-.H---DR-DI-E-Y-----Q---.TD--E--------M-.FM--QME-N.-I...NT........... 132
O.CM.98.98CMABB141 -K-MK-KNK....K.....LGIILCI....--AL.-IP-.LR...S.KD-YA---A-----E--NPV-------NLTS--Q--I--SQ-------T-H-YP-.H---DK--I---Y--D--Q---.----E-------QM-.FM-IQMN--.-F...-T........... 133
O.CM.98.98CMABB212 ---M--KNK....K.....X-.ILSI....AMAL.-IP-.LS...YS-Q-YA---S-----E--KPV-------NLTS--K--I--SQ----------N-YT.T---DS--I---Y--D------.------------QM-.F---QMN--.-P...-I-STG...NNST 140
O.CM.98.98CMU5337 -K-MK-XNR....K.....LE.ILYI....VM-L.IIP-.LS...S.N-MYA---A-----ED--P-------VNLTS--Q--I--SQ-------T-H-YP-.H---D---I---Y-----Q---.----E-------QM-.F--IQMN--.-E...TEG-......... 134
O.CM.99.99CMU4122 --AM--RNR....Q.....LG.ILYI....VMAL.IIP-.LS...Y.DQ-YA---S-----E--KP-------VNLTS--K--I--SQ-------T-H-YP-.K---DK--I-E-Y-----Q---.----E-------QM-.F---QMN--.NV...ND........... 132
O.FR.92.VAU -KAMG-RNR....K.....LG.I-CL....I-AL.IIP-.LS...C.NQ-YA---S-----ED-KP--------NLTS--Q--I---Q-------S-N-YE-.K---DK--I-E-Y--D--Q---.-D----------QM-.F---QMN--.-I...--SIT-....... 136
O.GA.11.11Gab6352 -KAM--RNR....K.....L-.I-CI....VM-F.ITP-.LS...SNQQ-YA---A-----ED-SPV-------NLTSP-K--E--AQS------T---YP-.K---DK-DI---Y---------.-D--E------TQ-S.F---RMN--.-P...N-N-SK....... 137
O.SN.99.99SE_MP1299 -K-M-QRNR....K.....LG.ILCI....VMAL.ITP-.LS...Y.NQHYA---A-----E---PV-------NLTS--Q--I--SQ-------S-Y-YP-.TK--D---I---Y-----Q---.----E-------QM-.F---QMN--.NY...VQGNY-NN..... 138
O.US.10.LTNP -KAM-RKNR....-.....S-.ILCI....VMVL.IIP-.LT...S.KQ-YA-I-A-----ED-APV-------NLTSQ-Q-----AQ-------S-N-YT-.H---DK-DI---Y-----Q---.----E-------QM-.FM--QMN--.-H...NT........... 132
N.CM.02.DJ00131 -.....QIG.W.PF.....FC.LMISL...TI........G...S.K-Y-A---------RDVE-V---------HS--A--I---Q------------P-.D----P----E-K-A--VQ---.----E----------.----TMN-S.NS...NGNRT-...DEKE 131
N.CM.02.SJGddd ---MGIQSY.W.SL.....VC.LLVSL...IKV.......IG...S.-QH-A---------RD-E-I---------HS--A--I---Q------------L-.P-----------N-A---Q---.----E-------Q--.----TMF-N.NS...NGEEPV...DDLT 138
N.CM.04.04CM_1015_04 ---MGMQSG.W.PF.....FC.LLISS...II........G...S.-QH-----------RDTE-V---------HS--A--I---Q------------P-.I----H-D----N-A---Q---.----E-------X--.----TMH-X.NS...NGXNX-...VDN. 135
N.CM.06.U14296 -GIMGMKSG.W.PF.....FC.LLISL...II-Y.EQ...IR...S.-PH-----------RD-E-I---------HS--A--I---Q------------E-.P----Y----E-K-AD--Q---.V---E----------.----TMN--.-S...NGNVTK...NEKK 141
N.CM.06.U14842 ---MRRQSG.W.LL.....CC.FLLSL...TI........G...F.-PK-----------RD-E-V---------HS--A--I---Q------------P-.I----H------KVAD--Q---.----E----------.----TML--.NS...NGNVTE...GSGT 136
N.FR.11.N1_FR_2011 ---MGMKSY.W.SL.....FY.LLXSL...IKV.......IG...S.-QY------------D-E-I-----X---HS--A--I---Q------------L-.P----D----E-N-A---Q---.----E----------.----TMN--.-S...DG-GTN...STIN 138
P.CM.06.U14788 -K-MK-KKN....S.....LG.NLEI....C-AL.I-YF.N-ISY-SGTRY----------RD-KI-----A--SLTSX-Q--I---Q-------R-I--RI.D----S--I-D-Y--T--Q---.---------------.V---TMX-S.-C...STVDC-...NNSS 144
P.FR.09.RBF168 --AMK-KKN....S.....LG.NL-I....C-AL.VIYF.N-ISCASGIHY----------RN-EV-----A--SLTSK-Q--I---Q-------T-I-IP..I----S--I---Y--T--Q---.---------------.I---TMN-S.EC...HREASC...N... 140
CPZ.CD.06.BF1167 -.....KMA.-.LI.....--.ILTII...FCCK.-G........I.GTYYT-------------P------S--VAS-QP--I----N---L--Q-Y-IPI.NIT--.---ET-Y---E-KK-L.---FQ----------.-F--KMN-S.EF...TRNST-...GNNA 133
CPZ.CM.05.SIVcpzMT145 -K-MGMKYY.WWRN.....SC.-SISIKLI-I-W.IAS-.FG...E.-NW-----------R--K----------S-S--A--I---Q-------T----L-.P----E----E-Y--D--Q---.----E----------.----T-T-N.NP...T-TSC-....... 143
CPZ.GA.88.GAB1 -K-M--KKRDW.NS.....LS.IITII...TIIL.-TP-.LT....SE------------HD-DPV---------HS--A--I---Q-------S----F-.P--I-S------N--D------.---------------.----T-Q-S.KA...NFSQAK...NLTN 143
CPZ.TZ.06.SIVcpzTAN13 -.....KNS...SGL....IL.TLI-T...I--........IG...F.STYYA--F-------D--PP-------DVAS-QA--I----N---L--H-Y--P-.P--SIE-D-ST-Y---E-KT-L.---FQ--F-------.-F-IRMT-A.NVTE.ATT-TP...PTT. 134
CPZ.US.85.US_Marilyn -K-M--KKRLW.LS.....YCLLSSLI...IP-L.............SS--A---------RDVE------------KQ-A--I---Q------------H-.P---K-D---E-N-A---Q---.-----------I---.----TMT-L.NP...DSNSSA..V.... 135
GOR.CM.07.SIVgor2139_287 --AMR-K...Q.LN.....L-.K-VM....S--L.ILT-.IE.....ATHYA---------R--NV-----A--NYVSK-Q--I---Q-------R-T-FP-.E----T--I---Y--D--QD--.V--------------.VM--T-N--.Q-...NTTNT-....... 135
GOR.CM.07.SIVgorCP2135con --AMR-K...Q.LN.....L-.KGVM....S--L.ILT-.IE.....ATHYA---------Q--NV-----A--NYVSQ-Q--I---Q-------S-I-FP-.EX---T-DI-N-N--D--QD--.---------------.VM--T-N--.KT...TET.T-...QTNX 138
```

```
                                                        glycosylation NIS
                                                    glycosylation NCS                                                    CD4 binding                                   glycosylation NVS
                                  V1                                                V2                                                                     glycosylation NGT
                       glycosylation NSS                                                        glycosylation NDT    glycosylation NTS                   glycosylation NKT
B.FR.83.HXB2                ..............NSS......S........GR...MIMEKGEIKNCSFNISTSIRGK.VQKEYAFFYKLDIIPIDND................TTSYKLTSCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLL 260
A1.AU.03.PS1044_Day0        N.............KTL......N........NN...T-EM-E-VR---Y-VT-VL-D-.K--I-SL--R--VV--G-N................SDSE-I-IN----T-----------------T---Y------D-D-------K-----------K---T---- 268
A1.CH.03.HIV_CH_BID_V3538   ...........................ITE-EFKV------MT-IL-D-.K--MH-L--RT-VVQ-GSN.............SSERYK-R-IN----A---------------------------------D-E-------M-----------K--A------ 256
A1.CY.08.CY236              N.............-IT.....-........-I...NDTIAT-V------TT-ELKD-.R-NV-SL--R----VQ--SN.............SNNSH-R-IN----T--------T-K---------------------D-K-S-KR--R------W----------- 266
A1.ES.05.X1608_8            ...............N......R........TI...AQGMEK--Q-----MT-ELTDR.Q--VHSL--S--LAQMG-SSYKGQIQT.SNNNLSVTYHT-R-IH----T-R-----------------------------DTE-------K-----------K--------- 273
A1.KE.11.DEMA11KE001        NV............TVN......N......TEVN...LRE-G----------T-EL-D-.KK-V-SL--RS-VVSMNSN...............ESSE-R-IN----T----------Q-----------------R-E--D---S-K-----------K--------- 267
A1.RU.11.11RU6950           YN............---......-........NS...SDSWVEVMR-----VT-EL-D-.KK-V-SL--R---VSTGSN...............DSEQ-R-IN----AV-------T--------------------KDTN-T-K---K-----------K--------- 266
A1.RW.11.DEMA111RW002       ........................G........NT...KGT-IEGMT-----AT-EL-D-.N--M-SL------VQ-NE-KNNSS........KNNSSSLNR-IN----A-------------------------------RDPK-------K---S-------K--------- 263
A1.SN.01.DDI579             ........................T........EN...VTKAYE-M-------T-EL-D-.K--I-SL--R--VV--E-GSNS...........TENS-E-R-IN----A--------------------------------KE-K-------R-----------K--------- 262
A1.UG.11.DEMA110UG001       ........................T........Q-...LNGTEVGMT-----TT-EL-D-.IR----L--R---VQ-NGS...............DNST-R-IN----A-K------------------------------KDAE-------N--------------------- 261
A1.ZA.04.04ZASK162B1        DT............DNN......P........NN...T-NMEN--------VT-E--D-.R--A--L-----VVQ---N................SSE-I-IN----A--------T-D--------------------D-Q---S-I-K-----------K-------S 264
A2.CM.01.01CM_1445MV        .....................................TK-M----YLMP-ELKD-.T-EV-SL------VQ-NSTDNVTQ........QNND-QP-R-IR-------------T-------------------KD-K-----S-R---S----------------- 255
A2.CY.94.94CY017_41         ........................-........ST...QSPINE------Y-TT-IL-D-.T--V-SL--R--VVQL-ESENKN.........TSGSN-L-R-IN----T--------T-------------------KDPR-----S-K---S-------K--A------ 265
B.BR.10.10BR_MG029          AT...........N.VT-.......NS......PIG...ERIAR--V-------VT-A-KD-.-K----Y--N--VVQ-TEN..............ANNN--R-A-------------I--D--------------------------D-K---N-----------K--------- 270
B.CA.07.502_1191_03         ...............A--.......N........NS...NNNGWE-M-------T-AM---.-R--Q-L-HS---V--KDEDN............NNTNAT-M-R-------------I-----------------------S-R------------------------------ 271
B.CH.08.M2_0803101_NFLG8    T.NSTSN........AN-.......T........SE...AP-----------DVD-G-SK-.I-----L-----VV--ED-NKTTNG....SYNNTKYSN-R-I-----I-------T-------------------D-K------------------------------ 282
B.CN.10.DEMB10CN002         ...............-TN.......N........SV...ANL-I--V---------.KTGN-.-E-GQ-L-----VEQ-N-NAN.............TSYAN-V--H------K---------------------------KD-K-------K-----------K--------- 266
B.ES.10.DEMB10ES002         T..ASNNT....LT.CND.......T.......I-G...QGWA-E--------ATKG-GD-.---T--L-NT--VMQ--G-................N---T-IN------R-T--------------------------D-K------------------------------ 273
B.FR.11.DEMB11FR001         ...............T-........-..................NGMT------T--L-D-.W--V--L-----LVK--D-NN....................--R-I-------------------------------------D-K-------K-----------K--------- 249
B.GB.05.MM45d213_GN1        ...............IN-.......-........SG...EL--V----------KGVT-D-.-----N-Y-----VVE--DSNS.............SYR--R-I------M-H----------------T-----L----D-K-----L-K-------------------- 264
B.HK.06.HK003               ...............TTC.......N........SS...RVEKEVGM-------VT----D-.-K--H-L--N---VQMN-.................-T-R-I-------------I-----------T-----L--------M-K-S-------------------------- 261
B.HT.05.05HT_129389         C...PEYS.....K.TKP.......NM......TRE...LNETIE--R----D-APITKD-.---N--L--T--LV--KKEN...............N--R---------T-------------------------------D-W-----------S-------------------- 270
B.JP.12.DEMB12JP001         ...............T-T.......N........SS...K-E-----------VT-T-DN-.AR--S-L--R--VV---D-NNS...........DSSTSA--R-IN--------T------------------L-------S-K-------------------------H-- 271
B.KR.07.HP_18_07JHS10_3909  ...............TN-.......-........SG...KV-----M------VT-N-GN-.------L-----VVQ---....................---T-IN--S-I-------I-----------T------Q--D-K---S---S-----------Q-------- 259
B.PE.07.502_0525_wg5        ......................................NETD-----------T-HLKD-MQKE---L-----VV----SDN..................---R-IN---------------------------T---Y----KD-N-----T-H------------------------- 250
B.RU.11.11RU21n             ...............TNN.......-S.................I-GS-M-------T--MKKG.-K----L--N--VVQ---NDT...............S--R-I------V------L-----------------I--------------D----------K--------- 250
B.TH.08.MERLBDTRC10         ...............TT-.......G........-E...EV--S----------T--TKD-..KNT--L-----VT-M--EVSRNNIT......NDTNYSR-R-LN--------------------------T------IQ--D-K-------K-------------------- 270
B.US.11.ES38                ........................A........SN...WTTIGE-M-------T--SE--.---K--S--N---RQ-ED-NTS..........TNNWN--R-I-------------I-----------T-----L----ETN---S---R--T----------------------- 263
C.AR.01.ARG4006             ........................D........TI...LDDM----------MT-EL-D-.QK-VH-L--R--XV-LN-E..............NDSSD-R-IN----A--------T-D-----------Y-----------------K-----------K--------- 262
C.BR.07.DEMC07BR003         DIR............TTANS....T........KN...NSSISK-MMS----MT-EV-D-.KE-VN-L-------V-LNISSGNNNS....SDDNNSSGKY-R-IN----AV---------D-----------Y------------------N-----------K--------- 284
C.BW.00.00BW5031_1          .................................T...VAEM---------H---EM-D-.R---F-L-NI---V-LN-ENNN.............TKNFSD-R-I-----T----------WD-----------Y------------------N-----------K--------- 256
C.CN.10.YNFL19              S..............-G.......T........NI...SNSTDSGM------AT-EL-D-.RK----L------V--EGENSSE...........GKNFSE-R-IN----A---------N-D------T---Y------------------S-----------K--------- 270
C.CY.09.CY260               S...............TTA.....P........TT...TVSNTV---------T-EL-D-.EK----L--R---VELGSN.................SSE-R-IN----A---------T-D-----------Y------------------N-----------K--------- 264
C.ES.08.X2363_2             P....TN.........MPN.....I........NE...TGNM-E-M------AT-E--D-.KN-V--L--R---V-LS--NS............SSNSSD-R-IN----T-----------------------Y--------K-------Q-----------K--------- 271
C.ET.02.02ET_288            P....NV.........T--.....-........TD...ASKVI-QM-------T-EV-D-.NK-A--L--R---V-LT--..............NSSDE-R-IN----T----------D---L-------Y------D---------K-----------K--------- 268
C.IN.09.T125_2139           TQ..............GDA.....N........VT...HNEGAN-------K-A-EL-D-.E--V--L--R---VQLNDSSNSSS.....SNNSSKYSK-R-IN-D--TV--------T-D-----------Y--------E---------------------K--------- 276
C.KE.00.KER2010             G...............TGN.....D........TY...ENDMTR-M------TT-E--D-.Q--V-SL------VSLE-G................NNSST-R-IN----A---------I-------------Y------------------Q-----------K--------- 266
C.MW.09.703010256_CH256.w96 T...............RGN.....D........ST...IGKMGD-MT-----AT-E-KD-.KKL---L-----VVTLEEN.................SSE-R-IN-----V-----------H-----------Y------------------NKI---------K--------- 263
C.TZ.08.707010457_CH457.w8  .................................Y...KNITD--VM------T-E--D-.RR--S-L--RV--V-L--N..................SAE-R-IN-K--SL-------I--D-----------Y------------------NS----------K--------- 254
C.YE.02.02YE511             ........................V........TF...NDTMR----------T-E--D-.RKQ---L--R---V-L-GK...............EENPSE-R-IN----AV---------D-----------Y-------R---------N-----------K--------- 261
C.ZA.08.705010534_CH534.w12 ......................................SI-K-MRQ----VT-VV-D-.KR--N-L-------V-LNGNSSG..........NGNVSE-R-IN----T---------T-D------T---Y-----DN--------------------------K--------- 254
C.ZA.10.DEMC10ZA001         ........................-........AN...INDDMEVM------AT-V-----.T-N---L---T-VV-LNEN................ASE-I-IN--S-T-------------D-----------Y------------------H-----------K--------- 258
C.ZM.11.DEMC11ZM006         ........................N........AN...VTNSTD---R----MT-EL-D-.SKQ-H-L--RD--T-LGGN................DS---R-IN----A-K------------------------------N-------N-----------K--------- 261
D.CM.10.DEMD10CM009         .......................................TTDTDMR------T-MLQD-.K--K--L---M-LVE-GKN.................S-K-R-IN----A--------------------------------D-Q-------K---S---------------- 255
D.CY.06.CY163               ..................................TV...DNSSI-DL------VT-IL-N-.KKQVH-L--R--VVK--SQNNR..........ANQSNGT-R-IN----A--------------------------------DPR-E-K-T-K----------------------- 264
D.KE.11.DEMD11KE003         ...............--T.....-........IK...DESKEIGM-------T-E--DR.KK-F--L-----VVQ--GSNS.............TDN---R-IN----A--------I-----------------------D-M-------K---S---------------- 267
D.KR.04.04KBH8              .......................................NST--M------VT-VV-D-.KKQ---L--R---V-M---...................N---R-IN----TV-------I-----------------------------------------K--------- 252
D.SN.90.SE365               ...............I-N.....V.......TG...PM-G---MQ------T-E--N-.R-NV-SL-----LV--G-SNKN...........STNY---R-IH----T---------N-------------------KD-K-------K-----------K--------- 267
D.TZ.01.A280                ...............TNT.....T........DP...RLI----M-------T-E--D-.RKQVQ-L-----VV---KKNN................N--T-MH----A-K-----------------------------KD-K-------KK-------------------- 262
D.UG.10.DEMD10UG004         .......................................E-GLGV------VT-EV-D-.KKQVQ-L--SH-VVQ-R-N..................T-R-I-----A--------IT--------------------D-K---Q-R------------------------ 252
D.UG.11.DEMD11UG003         ........................T........IA...HNVTEKIMT----KVT-E-ND-.KK-VH-L-----VVK--DN................D-SA-R-IN----AL--------T-------------------D-N-T-----N---S-------K--------- 261
D.YE.02.02YE516             .......................................STRE-M------VT-VV-D-.QKQ---L------VAM--S...................T-R-IN----T-----------------------G-----E-NYT-K---N-----------K------F- 252
D.ZA.90.R1                  ...............S-G.....-........TT...TEKPTL-VQ-----VA-VV-D-.RKEAS-L--RP--V---Q-NK.............NGNK-T-R-IN----S-------------------------------RD-K---------------------K--------- 266
F1.AO.06.AO_06_ANG32        ........................N........ST...LPE-A-A-------MT-E--D-.RL-VN-L--R---V--NSSD..............SNSNE-R-IN----T-K-------WD-------------------------------N-----------K--------- 259
F1.AR.02.ARE933             ...............-TT......N........NT...LAD-P-A-Q-----VT-EV-D-.KL-VH-I--R---V--SGSN..............DNSST-R-IN----T---------WD-------T---Y-----XD-R-------R-----------T------S- 258
F1.BR.10.10BR_RJ015         ...............--T......N........DT...LKEDP-ALQ-----MT-EV-D-.QL-VQ-L------VQ-N-NSSNN.........SSSNFSD-R-IN--S-I-K-------WD-----------Y-----KNN-----------I-------K--------- 263
F1.CY.08.CY222              ...............Q--......Q........SS...LEE-S-A-Q-----MT-EV-D-.KL-VH-L--R---V--NSSD..............NESSE-R-IN----T---------*D-----------Y---R--Q-D-------H-----------K--------- 259
F1.ES.02.ES_X845_4          ........................N........ST...LKE---A-------MT-EL-D-.KL-VN-L------V-LTSK...............GSSE-R-IN----T-K-------WD---------------R-------------N-----------K--------- 257
F1.RO.96.BCI_R07            PI..SPA.........PTP.....N........DT...LKEDT-A-Q------T-EVGD-.KLRAQ-L------V-LS-TNGP..........LNNTSGE-R-IH----T-K-------WD-----------Y-----KD-E-------K-----------K--------- 270
F1.RU.08.D88_845            ...............STQ......N........DS...L-EDQ-A-Q--T--M--EVKD-.TKQMRTL--S---EL-NKQSKX...........NRTXKE-I-R-----TM------I--D----X--T---Y---X-------K---X---G--X---K-X------ 267
F2.CM.02.02CM_0016BBY       ...............-QP......N........NI...TLE-Q----------T-E-KD-.RKN-F-T---H--V--EKN........................R-------TV---------D-----Y-----Y-----D-R---K-L-----------------K--------- 260
F2.CM.10.DEMF210CM001       N..............HTT......P........NI...TME-RE-V------VT-E--D-.QK----L--R--VV--GDNNN.............ADSGE-R-IN----TV-----------------------------D-K-----L-K---------------T---- 268
F2.CM.10.DEMF210CM007       ........................T........EA...PPI-APPM------AT-EVKD-.KK----L--R--VV-LNAN................ANS-E-R-IN----T------------------------------E-N-T-K-L-R---I-------K--------- 260
G.BE.96.DRCBL               .......................................T-EYRMT-----MT-EL-D-.KKA---L--RT-VV--NEMNNEN.........NGTNS-W-R--N--V-T-K------T-------------------VD-K-----T-N-----------K--------- 260
G.CM.10.DEMG10CM008         ........................K........IN...ANASGM-M------MT-E--D-.KETK--L--R--VV--NGN..................NST-R-IN--V-TVK-------------------------D-N-------N-----------K--------- 260
G.CN.08.GX_2084_08          ........................N........VT...CNNATSDNF--T--TT-E--DR.QK----L-----VV----NTN.............NNS-SN-R-IN--V-T-K-----------------T-----L---RD-N-----Q-N---S-------K--------- 264
G.CU.99.Cu74                ........................R........NC...SVEKGE---------T-E-KD-.KKQ---L--R---V--NDNRDNSTV.......SNSSPEN-R-IN--V-A-K--------D-----------------RD-E-------N-----------K--------- 267
G.ES.09.X2634_2             ...............-NN......V........NI...TSEG-N-L-------T-E-QD-.KK----L--R--VL--ESNAN.............SSSNYSN-R-IN--V-T-K--------D------------------R-N-----L-N-----------K--------- 267
G.GH.03.03GH175G            N...GTV.........DYN.....-........TV...DNSSCEI--------T-EL-D-.SK----L--RP-VV---GNDN..............STYSD-R-IN--V-Y-K-------N-D-----------------GD-K-----L-N-----------K--------- 271
G.KE.09.DEMG09KE001         ........................-........TN...STV-AK-MT----T-T-E--D-.QKN---L--RD-VTQ-RDEGN.............STNS-L-R-IN--V-T-K-------T---------------------D-N-T---S-Q---I-------K--------- 265
G.NG.09.09NG_SC62           ...............YTK......T........NY...TCE--DK-V--T---T-E--DR.RK--H-L--R--V-----N.................NSSE-R-IN--V-T-K-----------------------------RD-N-T-K-E-K---S----------------- 263
G.ZA.01.TV546               ......................................TVXNAR--T-----VT-E--D-.RRT---L--R--VV--NE-QN...............A-SD-R-IN--V-T-K-------N-D----------------I--RD-N-T---T-R------------------------ 258
H.BE.93.VI991               ...............-N-......T........DI...N-G-IQ-QR-----VT-A--D-.N--VH-L--RA--VQ--EGER............NKSDNH-R-IN------K-----------------------------G-K------------------------------ 263
H.CF.90.056                 ........................-........TS...SMEAG--LT-----VT-VL-D-.Q--VH-L--R--VV----N..................S-Q-R-IN--------------------------------------------------L------------------------- 253
H.GB.00.00GBAC4001          ........................-........SA...SLEGEQKLT-----VT-A--DR.KK-VQ-L--RI-VV--NDEDYNGT........NGTNR-L-R-IN-------------T------------------------------------------------K--------- 264
J.CM.04.04CMU11421          T...PAS.........PT-......D........AG...RVNISEIM-----KMT-E--D-.TRR---L--RQ--V--SGR................PND--I-IN------K------N------------------------DRN--------------------K--------- 268
J.SE.93.SE9280_7887         ........................-........NS...SVSSPDIMT------T-E--N-.RKQ---L--RQ-VV---SN.................NKN-I-IN------K--------Q-----------------------D-N-----S-K-----------K--------- 254
J.SE.94.SE9173_7022         ........................-........SA...SNNSPEIM------VT-E--N-.RKQ---L--RQ-VV--NS-.................NK--I-IN------K--------Q------------------------------K-----------K--------- 256
K.CD.97.97ZR_EQTB11         ...............DTN......I........NA...TVTSTD---------T-ELKD-.KKRVS-L------VQ-KQSEI............NQSESEDR-IN----TV-------------------------------N-C-----------------K--------- 268
K.CM.96.96CM_MP535          ........................A........TS...TVVSPA---------T-E-KD-.KK--S-L--R--VL-LNGEGN.............NSS-E-R-IN----T---T----T---T--------------------KD-R-------K-----------K--------- 262
U.CA.01.TV749               ........................N........AT...LNATT------T--MT-E--N-.NKQ---L-----LVQL-SS..................--D-V-IN--V-T-K------T--------------------XD-E-------R-----------K--I----- 260
U.CA.99.TV721               N...STNT........TTG.....N........NT...TAE-GQ-M-------T-E--D-.KKQ---L---A-LV-L--N...................-D-V-IN--VTT-K------T--------------------KD-N-------S-----------K--I----- 269
U.CD.83.83CD003_Z3          .......................................T-EAT-T----KVP-ELKD-.TETVHTL-----VV-LNVTNN............SSISST-R-IN----T---------------------------------D-K-------K------------------------ 257
U.CD.90.90CD121E12          ..........................................AN-LT----AVT-ELKD-.KKNVH-L-----VV-LNET..............NSSYGT-R-IN----T---------------------------------D-E-------R------------------------ 252
U.CY.05.CY090               ......................DK........NI...T-ED-D-MT-----TT-IL-DE.K--VH-L--R--VVQ-SDK..................SN-L-R-IN----A-------------------F-----Y-----RDVN-------K-----------K--------- 261
U.CY.08.CY223               TGNNETANANQTAN.ATN......-N......TTT.TELD-AA--L-------T-EV-D-.QKQ---L---P--EQ--G-...................NS--R-IN--V-S-K------T-------------------KD-NY-----H-----------K--------- 281
U.ES.10.DEURF10DZ001        .......................................NATDD-Q-----AT-EL-D-.KK-IH-L---I--VK-NDS................ESNS--M-IN----T--------T-D------S----------D-N-T-K-L-R-----------K--------- 253
U.GR.99.99GR303             ...............C-N......V........NA...TDEMR--M----Y-MT-EM-D-.K--V-SL--R--VVQMGS-NS.............KSNYSE-R-IN-S--A---------T---------T-------R--EEK-------K---S-------K--------- 264
U.NL.95.U_NL_95_H10986_D1   .......................GN........NA...TVE----L-------T-VLQD-.KR----L--RF-VV-LNS-NN..............SSNYSE-R-IN--I-T-K------D---------------------KD-R-------K-----------K--------- 262
01_AE.AF.07.569M            D..............KIG......N........LT...IGNITDD-------MT-EL-D-.KK-A--L------VS-EKN.................-SE-R-IN--S---K-----I--D-------T---Y------D-K-------K---S-------K--------- 264
01_AE.CF.90.90CF11697       K..................................L...NDTIGD-VR-----VT-EL-D-.K-EVH-L--VP--VR-GEKNKN..........SSGNSSE-I-IH------K-------T-D-------T---Y------E-N-------K---S-------K--------- 266
01_AE.CN.10.YNFL03          Y...............SIT.....K........-K...VENITN-VR----MT-EL-D-.K--V--L---P--VE--SN..................NSE-R-IN------K---------D-------T---Y------D-N-------K---S-------K--------- 265
01_AE.HK.04.HK001           ...............IN-......X........VT...IGNMTD-VR-----MT-ELTD-.Q--V--L------VE-N-.....................ST-R-IN------K---------D-------T---YV-----D-K-S-----N---S-------K--------- 261
01_AE.IR.10.10IR.THR48F     I..............TNA......-........DI...LGNVTE-VR--T--MT-VLQD-.K-RVH-L-----LVQMNGS..................NNT-R-IN------K-----I--D-------T---Y--X---D-R-----------S-------K--------- 264
01_AE.JP.x.JRC77AE          T..............-DT......L........DT...IGNLTD-VR-----MT-DL-D-.K-QVH-L--R---V---........................ST-R-IN------K---------D-------T---Y-L----D-K-------R---S-------K--------- 262
01_AE.TH.09.AA090a_WG11     ...............GNN......N........TT...VLSLR--V------MT-EL-D-.Q-NF--L--R---VQ-SDT................DNST-R-IN------K-----I--D-------T---Y------D-K-------K---S-------K--------- 264
01_AE.TH.90.CM240           ...............TNV......-........NI...IGNITD-VR--T--MT-ELTD-.K--VH-L------VQ-EDK...............KTSSE-R-IN------K-----I--D-------T---Y------D-N-------K---S-------K--------- 263
01_AE.US.05.306163_FL       ........................................NLTD-V------VT-E-AD-.K--VH-L------VQ-NKS....................SN-R-IN------K-----I--D-------T---Y------D-K-----T-N---S-------K--------- 247
01_AE.VN.98.98VNNDI5        ...............TNI......N........NT...IGNITD-VT--T--MT-ELKD-.Q--VH-L------V--H-S...............NNSSSE-R-IN------K-----I--D-------T---Y-----KD-N-------K---S-------K--------- 265
```


glycosylation NIS
glycosylation NCS
CD4 binding
glycosylation NVS
glycosylation NGT
V1
V2
glycosylation NSS
glycosylation NDT
glycosylation NTS
glycosylation NKT
B.FR.83.HXB2
02_AG.CM.10.DE00210CM013
02_AG.CY.09.CY256
02_AG.ES.06.P1423
02_AG.GW.05.CC_0048
02_AG.KR.12.12MHR9
02_AG.LR.x.POC44951
02_AG.NG.09.09NG_SC61
02_AG.NG.x.IBNG
02_AG.SE.94.SE7812
02_AG.SN.98.98SE_MP1211
03_AB.RU.97.KAL153_2
04_cpx.CY.94.94CY032_3
05_DF.BE.x.VI1310
06_cpx.AU.96.BFP90
07_BC.CN.98.98CN009
08_BC.CN.97.97CNGX_6F
09_cpx.GH.96.96GH2911
10_CD.TZ.96.96TZ_BF061
11_cpx.CM.95.95CM_1816
12_BF.AR.99.ARMA159
13_cpx.CM.96.96CM_1849
14_BG.ES.05.X1870
15_01B.TH.99.99TH_MU2079
16_A2D.KR.97.97KR004
17_BF.AR.99.ARMA038
18_cpx.CU.99.CU76
19_cpx.CU.99.CU7
20_BG.CU.99.Cu103
21_A2D.KE.99.KER2003
22_01A1.CM.01.01CM_0001BBY
23_BG.CU.03.CB118
24_BG.ES.08.X2456_2
25_cpx.CM.02.1918LE
26_AU.CD.02.02CD_MBTB047
27_cpx.FR.04.04CD_FR_KZS
28_BF.BR.99.BREPM12609
29_BF.BR.01.BREPM16704
31_BC.BR.04.04BR142
32_06A1.EE.01.EE0369
33_01B.ID.07.JKT189_C
34_01B.TH.99.OUR2478P
35_AD.AF.07.169H
36_cpx.CM.00.00CMNYU830
37_cpx.CM.00.00CMNYU926
38_BF1.UY.03.UY03_3389
39_BF.BR.04.04BRRJ179
40_BF.BR.05.05BRRJ055
42_BF.LU.06.LuBF_18_06
43_02G.SA.03.J11223
44_BF.CL.00.CH80
45_cpx.FR.04.04FR_AUK
46_BF.BR.07.07BR_FPS625
47_BF.ES.08.P1942
48_01B.MY.07.07MYKT021
49_cpx.GM.03.N26677
50_A1D.GB.10.12792
51_01B.SG.11.11SG_HM021
52_01B.MY.03.03MYKL018_1
53_01B.MY.11.11FIR164
54_01B.MY.09.09MYSB023
55_01B.CN.10.HNCS102056
56_cpx.FR.10.URF5_patient_A
57_BC.CN.09.09YNLX19sg
58_01B.MY.09.09MYPR37
59_01B.CN.09.09LNA423
60_BC.IT.11.BAV499
61_BC.CN.10.JL100010
62_BC.CN.10.YNFL13
63_02A1.RU.10.10RU6637
64_BC.CN.09.YNFL31
65_cpx.CN.10.YNFL01
67_01B.CN.11.2011.ANHUI.HF115
68_01B.CN.11.2011.ANHUI.WH73
70_BF1.BR.10.10BR_PE004
71_BF1.BR.10.10BR_PE008
72_BF1.BR.10.10BR_MG002
O.BE.87.ANT70
O.CM.98.98CMA104
O.CM.98.98CMABB141
O.CM.98.98CMABB212
O.CM.98.98CMU5337
O.CM.99.99CMU4122
O.FR.92.VAU
O.GA.11.11Gab6352
O.SN.99.99SE_MP1299
O.US.10.LTNP
N.CM.02.DJO0131
N.CM.02.SJGddd
N.CM.04.04CM_1015_04
N.CM.06.U14296
N.CM.06.U14842
N.FR.11.N1_FR_2011
P.CM.06.U14788
P.FR.09.RBF168
CPZ.CD.06.BF1167
CPZ.CM.05.SIVcpzMT145
CPZ.GA.88.GAB1
CPZ.TZ.06.SIVcpzTAN13
CPZ.US.85.US_Marilyn
GOR.CM.07.SIVgor2139_287
GOR.CM.07.SIVgorCP2135con

HIV-1 Proteins

```
                                                
```

```
                                                     CD4 binding    glycosylation NNT
                                                              glycosylation NCT          V3 tip
                                        glycosylation NFT                                   V3
                      glycosylation NGS               glycosylation NTS                          glycosylation NNT                 glycosylation NKT          CD4 binding        V4  glycosylation NST
B.FR.83.HXB2                  LNGSLAE.EEVVIRSVNFTDNAKTIIVQLNTSV..EINCTR..PNNN.TRKRIRIQ.RGPGRAFVTIGK.....IGNMRQAHCNISRAKWNNTLKQIASKLR..EQFG.........N.N.K.......T......IIFKQ..SS..GGDPEIVTHSFNCG.GEFFYCNS 387
02_AG.CM.10.DE00210CM013      --------.--IM---E-I-N-----L---KEP-..K-------.------.---GVH...L----V-YAT-....DIT-DI---------TE--K--QGV-TQ--..GH-P...........-AT.......A.......---AN..H-..---L--T-------.------ST 406
02_AG.CY.09.CY256             ----T--.GNIT---E-M-N---I-----ADP-..--------..-S--.--R-V-MG.I---QT-YAT....DI---I------V-KVQ--R--Y-V-TQ--..KY-N...........-.D..............R---NR..---..---V--T------A.------DT 384
02_AG.ES.06.P1423             ------K.DQIM---E-IS---------DSP-..N-T-I-..-----.--ESV-...I---Q--YAR-....DI--DI-----EVNKTE-VK--HKVSAQ-E..KY-K...........-.-..............-N-TK..---..X--I--T-------.------XT 378
02_AG.GW.05.CC_0048           --------.K-IE---E-I-N---I-----KEP-..N-T-S-..-----.---SV-...I---QT-YAT-....DI--DI------V-KRT--E--QNV-TQ--..KY-N...........T.........-......-V-AN..---..---L--T-------.-------T 393
02_AG.KR.12.12MHR9            --------.GK-A---E-I-N-V-N------KP-..Q-----..-----.---SV-...I---QT-YAT-....EI---I-----TV-ETE-T---RKV-IQ--..-H-K...........-.N..........---AN..PL..---L--T-------.-------T 387
02_AG.LR.x.POC44951           --------.KKI----E-I-N---------DKP-..R-----..-----.---SV-...I---QT-YAT-....DI--DI---Y--V--TE--T--QKV-TR--..-SL-...........-NH..........---TN..H-..---L--I-------.-------T 379
02_AG.NG.09.09NG_SC61         --------.KD-----E-ISN---------KSP-..K-----..-----.---GVH...I---QT-YAT-....DI---I------V-GGD--TA-Q-VVEQ-G..KY-N...........K.S..........-V--N..H-..---L--T-------.-------T 382
02_AG.NG.x.IBNG               --------.G------E-I-N---------ANP-..K-----..-----.---GVH...I---Q--YAT-....DI--DI------V-KTE--K--H-VVTQ-K..TY-K...........-.T..........---AN..PL..---V--T-------.-------T 382
02_AG.SE.94.SE7812            --------.--I----E---N---I-----HE--..K-----..-G---.---SV-...I---QT-YAT-....DI--DI------V-WQQ--K--HDV-T-----..-Y-N...........-.T..........---DE..P-..---L--T-------.-------T 380
02_AG.SN.98.98SE_MP1211       --------.GDI----E-ISN----------KP-..W-----..-S---.---SV-...I---QT-Y-T-....AV--DI------V--E---S--QKVVT--G..KH-N...........S.S.......K.....---TN..---..---L--T-------.-------T 379
03_AB.RU.97.KAL153_2          --------.------------T-------KEP-..-----------.---G-H...I------YAT-....DIT-DI---------IT---------VI---..K---...........-.-.......-......-V-N-..---..---------M-------.-------T 376
04_cpx.CY.94.94CY032_3        ------T.-------K-I---T-N-----AKA-..K-----..-G---.---SVH...I---LTWYAT-....EI--DI---------GND--D---V-SEE-K..RL-P...........-.-.......-......-K-AP..PV..---L--T------K.-------T 390
05_DF.BE.x.VI1310             ------K.-GII---Q-IS---N---H--E--..H-----..-----.---S-H...L---Q--YAT-....DI--DI-K----V--EQ--K--I-V-KE-Q..SH-P...........-.-.......-......-K-NS..---..---L--TM-----R.------DT 389
06_cpx.AU.96.BFP90            --------.GNIT-KTE-I---T-N-------QP-..--R---..-G---.---S-S...F---Q--IAT-....DI--DI------V---N-TDI-GEVKV--E..-V-N...........-.T.......H......-T--S..-A..---L--T-------.-------T 396
07_BC.CN.98.98CN009           --------.--II---E-L-N-V-----H--Q--..--V---..-G---.---S--...I---QT-YAT-....EI--DI---------ED---K--QGVSE--A..-H-Q...........-.-.......-......-K-AS..---..---L--T------R.-------T 389
08_BC.CN.97.97CNGX_6F         --------.R-II---E-L-N-V-----H--Q--..--V---..-----.---S--...I---QT-YAT-....DI--DI---------KDA-YE--QRVSK--A..-H-P...........-.-.......-......-K-AS..---..---L--T------R.-------T 389
09_cpx.GH.96.96GH2911         --------.-DII---E-I-N-V-V-----ADP-..T-D-V-..TG---.--TSV-...I---QT-YAT-....DI--DI-KV--Q-NGT---EI-GKV-E--K..GH-N...........-.-.......-......---RP..---..---V--T-------.------DT 380
10_CD.TZ.96.96TZ_BF061        --------.--II---E-L-----N------E--..T---V-..-----.---RS--...I---QT-YAT-....DI---I---YY---GTE--K--Q-V-K--G..DL-N...........R.P.......-......----P..---..-------T--T----.-------T 375
11_cpx.CM.95.95CM_1816        --------.G--R---K-I----------FDSP-..H-----..-----.---S--...I---Q--YAT-....AI--DI-----V-NATI-----RVVEQ--QI-I-N...........K.T.......-......---NS..---..---L-TT-------.-------T 398
12_BF.AR.99.ARMA159           --------.-DII---Q-IS--T-----H--E--..Q-----..-----.---S-Q...L------YAT-....DI--DI-K----V-G-Q-SK-VE-VKA---..S--...........-.-.......-......-K-NS..---..-------TM-----R.-------T 386
13_cpx.CM.96.96CM_1849        --------.G-IK--AK--S--VGN-----A-P-..R-----..-----.---S--...I---Q--YAT-....DI--DI---------Q------Q-V-AQ--..KYLK.........D.-.T.......-......---AN..P-..---L-VT-------.-------T 389
14_BG.ES.05.X1870             --------.--I----K-----T-------KDP-..-----------.------T..M----VLY-T-....XI--DX-K------KE------G--VE------K..........-.-.......-......-V-NR..---..---------M-------.-------T 390
15_01B.TH.99.99TH_MU2079      --------.----I---S---N---I------E--..-----------.---S-H...I----SWY-T-....EI--DI------L-ST--G---N--TE-----...........-.-.......-......-V-N-..P-..------IV-------.-------T 376
16_A2D.KR.97.97KR004          -------GNKTI---A-I---T-N----FTKP-..Q------..-DVG.Q-RSV-...I------Y-RQTYTRQAK-DI---Q----SRQ--D--QKV-EQ--..KY-S...........-.-.......-......---TN..---..---I--T---V---.------DT 390
17_BF.AR.99.ARMA038           --------.GDII---K-IS--T-----H--E--..Q-----..-----.---S-Q...L------YAT-....EI--DI-K----V-KTQ--K--E-VRA--E..SH-P...........-.-.......-......-K-NS..---..---L--TM-----R.-------T 387
18_cpx.CU.99.CU76             ------T.-K-M---E-I---T-N-----T-P-..N-T---..-S---.--RS-H...I---Q--YAT-....DI--DI---------EKD--D--HEV-TQ-G..GHLN...........K.S.......-......-Y--P..---..---I--T--T----.-------T 385
19_cpx.CU.99.CU7              --------.GK-T---E-I-----------FTXP-..K-----..-----.--TSV-...I---QV-YAT-.......-DI------V-KKE-KEA-QEV-KQ-T..VL-N...........-.-.......-......-S-AN..---..---L--T-------.-------T 355
20_BG.CU.99.Cu103             --------.K-IM---E-L-----V------K-I..-----------.---S-G...F---Q--YAT-....DIT-DI---F-VVNKT---EM-QKVKA--A..-SYN...........K.N.....ITN......-K-SP..PT..---L--T------R.-------T 385
21_A2D.KE.99.KER2003          --------.--II---E-L-----------E--..V-----..-G---.--QG-H...I---Q-LY-T-....KI--DI---------G-E--K--QRV-I----..RL-N...........-.-.......-......-T-NH..---..---------------.-------T 383
22_01A1.CM.01.01CM_0001BBY    --------.K-IM---E---N---I----XVKP-..T---S-..-G---.--NS--...I---QV-YAR-....DI--DI-K-Y-EVNKTE--E--HGVVKQ--..KHWN...........-.-.......-......---TN..---..---L--T------R.-------T 377
23_BG.CU.03.CB118             --------.K-IM---E-L-N---V------S-I..T-----..-----.---S-T...F---Q--YAT-....DI---I-E-F-VVN-TA--E-MQKVKA--A..-SYN...........T.S.....TEN......-T-GP..P-..---L--T----I-R.-------T 393
24_BG.ES.08.X2456_2           --------.KK-M---E--AN---V------E-I..V-----..-G---.--RSXG...I------Y-T-....DI--DI-----IVNGT--GEMINKVKAE-S..KKYN...........T.T.....TAN......-T-SP..H-..---L--T------R.-------T 397
25_cpx.CM.02.1918LE           --------.--I----E-I-N---N-L-H--E--..M-----..-----.---S-H...I---Q--YAT-....DI--DI---------SQ--K--QGVVAQ--..KH-K...........-.TI.......-......-S-NS..---..---L-VT--I----.-------T 386
26_AU.CD.02.02CD_MBTB047      --------.GG-KV--E---N---N----FDKA-..V---S-..-----.---S-H...L---QS-YAT-....DI--DI---A-----KD-DKA-SKV-KQ-S..DH-P...........-.-.......S......-A-N-..---..---L--TM-----R.-------T 385
27_cpx.FR.04.04CD_FR_KZS      --------.--II-K-E-L---T-----H--E--..K-----..-F---.--SVS...L---SVWYRP-....EV--DI---Y---N-KE--G--YKVVAQ-G..RHLN...........-.MS.......-......-N-QP.PP-..---I--T--T----.------DT 385
28_BF.BR.99.BREPM12609        --------.--II---E-L-----I-------ET-..-------------.RN---H...L-R--IMYVTE....RIQ-GI---------E--E---R-VVT--E..---TKLGEQFKDFK.-.R.......-......-A-N-..---..---------M-T----.-------T 398
29_BF.BR.01.BREPM16704        --------.-DI--K-E-------------ETI..-------------.---S-H...I------YAA-....EV--KI-----SL-ES--KR--E---T--G..---K..........I.-.E.......-......-A-D-.PPP..---------M-------.-------T 393
31_BC.BR.04.04BR142           --------.--I----K-LS--V-------Q-P-..N-T-M-..-----.---ES--...I---QT-YAT-....DI--DI---------KTA--R--QEVGK--A..-H-P...........-.-.......-......-N-TK..H-..---L--T------R.-------T 377
32_06A1.EE.01.EE0369          ------G.-SII---E-I---T--------Q--..D-I---..-Y---.--RS-H...I---Q--FAT-....DI--DI---------G-----EM-HKVIA--K..-S-P...........G.-.......N......-T-NS..-A..---L--T-------.-------T 394
33_01B.ID.07.JKT189_C         --------.--II---E-L---------H--K--..V-----..-S---.--TS-P...I---QV-YRR-....DI--DI-K-Y-EVNETN-YKV---VTE--K..-H-N...........-.N.......-......-V-EP..---..---L--TM-H---R.-------T 389
34_01B.TH.99.OUR2478P         --------.--II---E-I-N-------H--K--..-----..-S---.--TSVH...I---Q--YK--....DI--DI---Y-E-NGT---E----V-G---..KH-N...........-.-.......-......-S--P..P-..---L--TM-H---R.-------T 380
35_AD.AF.07.169H              --------.G-I----E-IS-----------ET-..T-----..-----.---S--...I---Q--YAT-....DI--DI------V--SA-KK--E-VKTQ--..KHWN...........-.-.......-......-L-ND..---..---V--T------R.-------T 385
36_cpx.CM.00.00CMNYU830       ------K.N-TI---E-I-N------I--VKP-..N-T-V-..-----.---GTP...I---Q--YARN....DV--DI-K---EVNATE--D--Q-VVQ---..KFWN...........-.T.......-......---TE..P-..---L-VI-------.-------T 382
37_cpx.CM.00.00CMNYU926       --------.GK-M---E-I-N--------F-ET-..-------..F---.---SV-...I---QT-FAT-....DI--DI-K-Q-AVNKTE--KA-QKVVAQ--..NH--...........-.T.......-......-M-AN..---..---V--T------A.-------T 378
38_BF1.UY.03.UY03_3389        --------.--II---Q-IS--------HF-E--..Q-----..-----.---S-H...L---Q--Y-T-....DI--DI-K-Y--V-GEH-----ERVKE---..SY-P...........-.E.......-......-K-NS..---..---L--TM-----R.-------T 381
39_BF.BR.04.04BRRJ179         --------.-KII---Q-IS--------H--E--..Q-----..-----.---S-H...L---H--H-T-....KI--DI-K--Y---GTQ--K--R-VKEE-K..SH-P...........-.G.......-......-S--P..A-..---L-VTM-----R.-------T 387
40_BF.BR.05.05BRRJ055         --------.-NI----Q-IS----V---H--E--..Q-----..-----.--TS-N...I------YAT-....DIV-DI-K----V--TQ-DK----VGAE-K..RL-H...........-.-.......-......-K-NS..---..---L--TM-----R.-------T 391
42_BF.LU.06.luBF_18_06        --------.-DII---Q-IS--------HF-E--..Q---S-..-----.---S-P...I------YAT-....DI--DI-K-Y--V--EQ--K--ERVRE--G..SL-N...........-.-.......-......-K-NS..---..-----VTM-N---R.-------T 395
43_02G.SA.03.J11223           --------.--I----K--S---QVL-----X-I..-----------.---S-G...F---QT-YAT-....DI--DI-E----V--TR--E---NVKAQ-H..KI-N...........-.-.......-......-L-NS..-I..---L--T------R.-------T 382
44_BF.CL.00.CH80              --------.-KII---Q-I---T-----H--E--..Q-----..-----.---S-H...I---Q--YAT-....DI--DI-K----V-KTE--K--ERVKAE-K..SH-P...........-.A.......-......-K-ES..---..---L-VTM-----R.-------T 386
45_cpx.FR.04.04FR_AUK         I-------.-KIM---E-L-N-V--------VRP-..Q---S-..-T---.--TSV-...I---QV-YAT-....DIT--I-K----VD--N-TK---EV-AA-----..---D...........-.-.......-......-A-NS..-A..---L--T-------.-------T 376
46_BF.BR.07.07BR_FPS625       --------.-DII---Q-I-----I---H--E--..Q-----..-----.---S-Q...I------YAT-....DI--DI-K----V-G-Q--K--ERVKA--N..SY-I...........-.E.......-......-K-NS..---..---L--TM-----R.-------T 380
47_BF.ES.08.P1942             --------.-DII---Q-IS---S-------E-I..S---I-..-G---.---S-H...M---KV-YAT-....DI---I---------K-------R---R--G..---N...........-.-.......-......-V-N-..---..---------M-T----.-------T 380
48_01B.MY.07.07MYKT021        --------.--II---E-L-N---N---H--K--..-----------.-ST---.---MT...M---HV-YST-....EIT-EVKK-Y-EVNGTT-SKV-T-V-E--K..-H-D...........-.-.......N......-T-Q-..PA..---I--TM-H---R.-------T 398
49_cpx.GM.03.N26677           ----I--.KGII---E-IS--T-N-----DKP-..T-K-I-..-A---.---G-H...L---QVLYAT-....AVV-QI---------ESD-----SKV-I----..-Y-N...........E.TT.......-......-K-ES..AT..P-I--T--T----.-------T 389
50_A1D.GB.10.12792            --------.KDII---E-I-N---------ASP-..N-S-V-..-----.---SV-...I---QT-YANN....DI---I------V-GS---K--Q-V-R--G..NLLN...........Q.T.......-......----P..---..-------TM-------.-------T 384
51_01B.SG.11.11SG_HM021       --------.--II---E-L-N-------H--K--..-----------.F-Y.--TST-...I---QV-FQT-....DI--DI-K-Y-E-NGT---E--R-VTR---..-H-N...........-.-.......-......---RP..PL..---L---M-H---R.-------T 386
52_01B.MY.03.03MYKL018_1      --------.--II---E-I---T-----H--X--..-----------.-S---.--TS-G...M----V-FRT-....EIT-DI-K-F-E-NGT---KV---VKX--K..-HXN...........-.-.......-......XD-QP..P-..---L--T--H---R.-------T 397
53_01B.MY.11.11FIR164         --------.--II---E-L-----------K--..-----------.-G---.--TSLT...I---QV-YRT-....DI--DI-R-Y-EMNGTE-KEV-E-VTN--K..-H-N...........-.H.......-......-S-QP..P-..---L--TM-H---R.-------T 394
54_01B.MY.09.09MYSB023        --------.-DII---E-L-N--RN---H--E--..--K---..-S---.---S-H...I----V-YRS-....DI--DI---Y-E-NGT---K--N-VSK--K..-H-N...........-.-.......-......---HP..P-..---L--TM-H---R.-------T 389
55_01B.CN.10.HNCS102056       --------.--II---E-L-N-------H--N--..T-----..-A---.--TG-T...I---QV-YRT-....DI--DI-K-F-E-N-TE-KEV---VT---K..GH-N...........T.E.......K......--LQP..P-..---V--TM-H---R.-------T 382
56_cpx.FR.10.URF5_patient_A   --------.K-IM---E-I-N---I------E--..S-----..-S---.---S-H...L---QT-YAT-....DI--DI------V--NQ-MT--QRV-RQ--..KH-K...........-DT.......E......-T-AS..---..---V--T------A.-------T 382
57_BC.CN.09.09YNLX19sg        -----S-.DGII---QDLKN-------H--K--..--E-I-..-S--.R--SV-...I---Q-IYAT-....DI--DI---Y---REKN--K---MVGK--A..-H-P...........-.-.......I......-K-NI..---..---L--T------R.------DT 390
58_01B.MY.09.09MYPR37         --------.G-II---E-L-N-------H--K--..----A-..-SD-.--TSV-...I---QV-YRT-....DIT-DI---Y-E-NGTV--D--Q-VSE--K..-H-N...........-.-.......-......-E-QP..H-..---L--TM-H--RR.-------T 398
59_01B.CN.09.09LNA423         --------.---IV--R---N---N------E--..V-----..-----.---S-H...M----T-FAH-....EI--DI-K-Y--L-EEQ------R--I--Q..---E...........-.-.......-......-R-NR..-A..---S-V-M---T--.-------- 384
60_BC.IT.11.BAV499            --------.G-II---K-L---------H--EA-..X-T-E-..-----.--ES--...I---QT-YAT-....DI--DI-----------A--K--QAVGK--A..-H-P...........-.R.......-......---RN..H-..---L-VT--#---R.-------T 388
61_BC.CN.10.JL100010          --------.G-II---E-LA--V---------Q--..--V---..-----.--RS--...I---QT-YAT-....GIT--I---Y----KEG-KEA-DRVER--T..-H-P...........-.-.......-......-E-AP..---..---L--T------R.-------T 382
62_BC.CN.10.YNFL13            --------.K-II-S-E-L---------H--K--..V-E---..-----.---S--...I---QT-YAT-....DI--DI---------AND--EA-RNVSK--A..-L-P...........-.-.......-......-N-AP..PV.E---L-VI------R.-------T 385
63_02A1.RU.10.10RU6637        --------.G--T---E-L-N---V------E--..A-----..-----.--VGVH...I---Q--YAT-....AIV--I---Y--V-G-G--KA-QKVVTQ--..KH-K...........-.T.......-......---DK..H-..---L--TS-T----.-------T 379
64_BC.CN.09.YNFL31            --------.--II---K-L-----I---H--E--..--V---..-S---.---S--...I---QT-YATE....KI--DI---------KRN-TE--QKVGK--A..KL.Y...........-.-.......-......-K-EP..---..---L--T------R.-------T 386
65_cpx.CN.10.YNFL01           --------.-DII---E-L-N-V-----H--E--..--V---..-H---.---S--...I---QT-YAT-....DI---I---Y----K----E--QRVGK--A..GY-S...........G.-.......P......---AP..---..---L--T------R.-------T 398
67_01B.CN.11.2011.ANHUI.HF115 --------.--II---E-LXN---I---H--K--..-----------.-FX-.--TSV-...I---QV-YRTK....XIX-DI-K-Y-XXNGTE--E----VKA--G..-R-N...........-.-.......-......---QP..P-..---X--TM-H---K.-------T 384
68_01B.CN.11.2011.ANHUI.WH73  --------.--II---E-L-N-------H--K--..-----------.-S---.--TSV-...I---QV-YGT-....DI--XI-K-Y-E-NGTE--E----VTA--K..-Y-N...........-.T.......-......---QP..P-..---L--T--H---R.-------T 381
70_BF1.BR.10.10BR_PE004       I-------.--I----S-L---T-I-LI---A--..-----------.---G-H...M----V-Y-T-....XI--DI-X------XXX--X--Q--XT--G..-X-K...........-.-.......-......-X-XX..---..-------TM-------.-------T 383
71_BF1.BR.10.10BR_PE008       --------.-------E-------N-----KK--..K-----..-S---.--RS-H...I---K--WAT-....GI--DI------L--TE--------VI----..A--N...........-.-.......-......---N-..----R----M-------.-------- 387
72_BF1.BR.10.10BR_MG002       --------.--II---Q-IS--------H--E--..Q-----..-----.--QG-H...I---Q--YAM-....DI--DI-K----V-GMQ--K--EGVKG----..SL-P...........-.-.......-......-N-NS..---..---L--TM-----R.-------T 390
O.BE.87.ANT70                 ---T-SK.GKIRMMAKDILEGG-N---T--STL..NMT-E-..-QID..IQEM-...I--.M-WYSM-IGG.TAGNSS-A-Y-KYNATD-GKI---T-ERYL..-LVN...........-.T....GSIN......MT-NH..---..----L-VTHLH---H.-------T 381
O.CM.98.98CMA104              ---T-SK.GKIR-MGKDI-AGSRN-L-T--STI..N-T-E-..-F-L.-VEQMK...I--.M-WYS-DIGR.GNTSDP-R-F-KFEET--KEA-X-T-ERYL..-LVN...........-.TI...DNVS.....V--NR..-N..D-----AFLH---H.-------T 388
O.CM.98.98CMABB141            ---T-SK.GKIRLMAK-IS--SQN---T---TI..SMT-Q-..-G-T.-VQE-K...I--.MSWYSMST...EGNPES-R-Y-EYNTTN-VEA---T-ERYL..-LVN...........-.T....GNIS.....M---N..-T..---L-VTQLH---H.-------T 403
O.CM.98.98CMABB212            ---T-SR.GGIR-MGT-I--S--N---T---P-..NMT----..EER..KIQE-H...V--.MRWRSMASVQ.ENRKSS-A-N-TYNSTE-R---EKT-KRYL..-LVS...........K.TGNGAGDV-.......---NK..T-..---E--SNLH---H.-------T 404
O.CM.98.98CMU5337             ---T-S-.GKIR-MGK-IS-S-EN-L-T--YP-..N-T-Q-..--.-V.-VQE-K...I--.M-WYSMAI...EKNSTS-I-S-KYNGTE-EKA---T-ERYL..-LMN...........-.NE..TNI-......MT-NR..-L..D--A-TTHLH---H.-------T 387
O.CM.99.99CMU4122             ---T-S-.GKIR-MGK-I--SG-N---T-DSAI..N-T-E-..-W-Q.SVQELK...I--.MSWYSM-ISS.PLSNAS-L-Y-RYNATE-KKA----ERYL..-LVN...........-.TNNTXDNVS......M--SK..-T..---S--AHLH-S-H.-------T 385
O.FR.92.VAU                   ---A-SK.GNIT-MGKDIS-SGEN-LIT---DM..T-A-E-..-GSQ.-IRK-M...A--.M-WYSMALS..NTK-DT-A-Y--Y-ATD--KA--N-TERYL..-LVE...........Y.NQ...TDV-......-K-GN..H-..---A-VTNFF---H.-------T 387
O.GA.11.11Gab6352             ---TISK.GKIR-MGT-ISAS-QN-L-T---TI..KMT-E-..-GD-V-VGQVG...P--.M-VISMNLAAEGRGN-S-R-Y-KYNTSN-EEA-RETTERYL..-LVN...........-.S....QLV-......VQ-NR..-T..D--A--T-LH---H.-------T 391
O.SN.99.99SE_MP1299           ---T-S-.GNIR-MGK-IS--M-N---T--STI..NMT-V-..QG-Q.SVQE-Q...I--.M-WYSMSLAQEGKPN-S-I-Y-KYNISD-EKA---T-ERYL..DLRN...........-.T....NTVN......-T-ER..-I..---S-VTHLH---H.-------T 392
O.US.10.LTNP                  ---TISK.GKIRMMAK-I---T-N-L-T--YTI..-MI-E-..-WRE..IAPMS...V--.M-LYSYRLAQ.SGNNSS-E---RYNPTD-ERA---T-E-YL..-LVN...........-.T....RNI-......V---N..--NN---L-TTRLH---H.------DT 377
N.CM.02.DJO0131               -----DK..DI---N....NSGGNLL--W-EI-..TM-----..-G--.-GGQVQ...I--AMT-YN-E....KIV-DI------V-NE.-RSMWNKTKE-IK..SLL-...........-.-.......N......-T--AQEKN..-----VTHLM----.----I---T 371
N.CM.02.SJGddd                -----NT.DDII--N....-SSGNLL--W-KT-..S------..-G--.--GQVQ...I--AMT-YN-E....RIV-DT-----IV-KKD-TKMWNRTKE-IG............S.......N......KT-QARNKT..E----VTHLM---R.-------T 372
N.CM.04.04CM_1015_04          X----DT.DDI-X-X....X.XDNML--W-XT-..S------..-G--.XGGQVQ...I--AMT-YN-E....KI--DI------V-EE.-KSMWDRTKE-IK..GLL-...........-.-.......N......TT--SRVNI..-----VRHFM----.----L--T 374
N.CM.06.U14296                ------D..DI-V-.....N-SGNML--WXET-..S------..-G--.-GGQVQ...I--AMT-YN-E....KI--DI-----IV--KD-ESMWNKTKA-IK..-TL-...........E.......N......-T-RARKKN..E----VXNLM---H.-------T 390
N.CM.06.U14842                -----DT.DDI---NSSDSNTSGDLL--W-KT-..S------..-G--.-GGQVQ...I---MT-YN-E....RI--DI-----SV--KD-GAMWNSTKE-I-..NLL-...........S.-.......N......-T-RARGKN..D----VTHLML---.-------T 389
N.FR.11.N1_FR_2011            -----GT..-I-XXN....-SRGNLW--W-ET-..SM-----..-G--.-GGQVQ...I--AMT-YN-E....KIT-DI------V-KEE-KEMWNKTKEEI-..KIGN...........W.......N......-T--VRENN..-----VTHLML---.-------T 380
P.CM.06.U14788                ---T--K.G-PLVITQ-VS-TG-V---K--E--..SLT-I-..-G--.--GQVQ...L-V.MTWYNMK....HYV-DI-A----V--EN-KR--EWVSEAIW..KAYPQH......S.-HTS..KHNH-......FV--N..-T..-----VSSLH-S-H.-------T 402
P.FR.09.RBF168                ---T---.--PLVITQ-VS-TRYV---K--KN-..S-T-V-..-G--.--GQVQ...I--.MTWYNMK....FYT-DI-K-Y----MQN-TE---SVSEAIW..KAYPQP......P.-.....QNH-......FV-RN..-T..-----VSFLHX--H.-------T 394
CPZ.CD.06.BF1167              ----VE-.G-TKAYY-EAKH--.PLLIKF-KATVFN-T-V-..-G--.--GQVQ...I---MT-YN-R....NV--DT-K-F-YVNKTL-E-ATEAAKVAIN..-TLK...........-.TT..FKNT-YTFRTDGDL-W..I-..---Q-TQ--W---Q.------DL 398
CPZ.CM.05.SIVcpzMT145         ----T-D.NQT-A-I.DPSE-L..A-I--KDP-..K-T-R-..-G--.--GQ-Q...I--AMT-YN-E....NVV-DT-K-Y-E-NGTQ-AKA-NETKEV---..NILR...........-.-.......N......-S-MV..P-..-----VTN-H----.-------T 379
CPZ.GA.88.GAB1                I-------.GNITV-VE-KSK-TDVW----VEA-..SL--H-..-G--.--GEVQ...I---MT-YN-E....NVV-DT-S-Y-K-NGTT--R-VEEVKKA-A..TSSN...........R.T.....AAN......-TLNR..A-..-----VTH-M----.-------T 385
CPZ.TZ.06.SIVcpzTAN13         ---TYEK.-KTKVYHT-V-N-P.PLL-KF-ETI..H-T-E-..TG---.--GQVQ...I---MT-YNRE....NIF-DT-K-F-YVNATE-RR--EMAKTA---..-ATK...........K.........P......TLNIT..RP..K---VENFW---Q.-------L 391
CPZ.US.85.US_Marilyn          I-----T.KN-TV--K--A-...I-L--FSEG-..NMT-I-..-G--.-VGNVQ...L---MT-YN-P....KIV-DV-E------KLT-EKQR-YTLEIIK..KEAN...........L.T.......K......VELIP..NA..-----V-NMML---.-------T 367
GOR.CM.07.SIVgor2139_287      ---T---.GNAT-I-K-SS-SG-D---K-ARP-..--T-E-..TG--.--GQ-Q...I--.MTIYNSE....NIV--T-K-F-KYNETN-QGA--DTVQA-N..-S.....................L......-H-RN..---..-----VTFLH---H.-------T 389
GOR.CM.07.SIVgorCP2135con     ---T--K.-HAT-I-K-S--SG-D---K-AR--..--T-E-..TS--.--GQ-Q...V--.MTIYNSE....NIV--T-K-F-KYNKTN-Q-A--DTVRA-K..AN.....................Q......-R-QN..---..-----VTFLH---H.------DT 383
```

```
                                      glycosylation NST
                              glycosylation NST
                                                         V4                                                          V5
                         glycosylation NST                            glycosylation NNT         glycosylation NIT  glycosylation NES
                                                                                      CD4 binding                  CD4 binding       CD4 binding
B.FR.83.HXB2               TQLFNSTWFN.......STWS..TE.GS............NNT..EG......SD.....TITLPCRIKQIINMWQKVGKAMYAPPISGQIRCSSNITGLLLTRDGGN.SNN..........ESEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRRVV....  506
A1.AU.03.PS1044_Day0       SE---R--NA.......TDQL.......N..S.....TE..DS-..AL......NE.....--I-------V-----TP-Q-------R-A---E------I--------.D-T.........STN-T----------------------R-----I---T-------....  513
A1.CH.03.HIV_CH_BID_V3538  -E------ND.......TSNT.................S..QG-..NN......NS......-------------FV----R--Q-------K-K-Q-I-----------K.GKS.........NIT-----E--N-K---------------------------....  497
A1.CY.08.CY236             SK---T--NV.......TDS......................IEE..S-......NT.....I-V---K----------RA-Q-I-P---R-I---E-------------D.TE-.........TT-------------------------E-R-------....  505
A1.ES.05.X1608_8           -G--DG--NG.......TKS........N..I.........TWE..SI......N-......------K---VVR---R--Q-V-----T-A-S-N-T----I------D.NSS........GR---T--------------------------R-------....  515
A1.KE.11.DEMA11KE001       SG------SL.......NNTWS.....LN..S...TQS..-D-..GI......N-......-----------------R--Q-------Q-V-Q-V---------------.V-S........SNGN----------------------------R-------....  518
A1.RU.11.11RU6950          -E------DG.......NGT......................ATN..NT......E-.....NV--------------R--Q-------R-S-K-V--------I----G.G--.........SNN-T---I--N--------------Q---I-I---T-------....  506
A1.RW.11.DEMA111RW002      -G------SR.......N-TDN.NT.WFRN.N.....AS..SS-..GL......NG.....----Q-K---------RT-R-------P-T-M-N-S--------------.-G-.........NSN-T-------K--------------------R-------....  514
A1.SN.01.DDI579            SG------ND.......TAIN.......D..T......I..-VQ..NS......TY.....N-------R--VH---R--Q-------Q-V-N-T-------------D.NSS.........TN-T------N-K----------------------------....  504
A1.UG.11.DEMA110UG001      SG------N-.........--QES...........................NS......TE.....-----------V----RA-Q-I-----K-I-N-R------I-----H.N-..........TN-T-----------------------V--------T-R-------....  498
A1.ZA.04.04ZASK162B1       SG---R--ND.......TA-N.......E..T.....TW..--I..NS......NG.....S---Q------L----R--Q-------P-E---E------I-----L.EL-........NTNN-T--------------------I---------R-------....  509
A2.CM.01.01CM_1445MV       -G------TK.......EGN-...T.W.K..K...NDSV..SER..-S......N-......N---Q-K----------R--R-------A-V-K-I-----MI------K.R-D.........SS--T-------------------------I-----R-----....  505
A2.CY.94.94CY017_41        -G---G--W-.......NGTW.......N..G.....PY..TPN..NT......NG.....S-I-------------R--R-------A-I-K-T-----II--------.NG..........TN-T----------------------L--------R-------....  508
B.BR.10.10BR_MG029         -X---------.........---N..XT.X.N..G.........XIE..GS......DA.....N---Q-K----------Q----------R-K----------------G.NGS.........STN-T--------K----N--------------I-----------....  518
B.CA.07.502_I191_03        -L------GS.......NETR.......................-E-..ID......NS.....---IQ-K---L-----G--R--------R-N-S----------------.NGT.........NGT-T---E--N-K----N--------------I-----R-----....  512
B.CH.08.M2_0803101_NFLG8   S-------NA.......TGTQ.......N..T......TGTQ-E..TK......IG.....N-----K------R--E----------K-L-N-V-----I-------H.-K-TN....STDNDT-V--------------------------------------....  535
B.CN.10.DEMB10CN002        SK------PS.......DGTD.....-.N...........VTG..GS......N-......----Q---R----R--E----------R-V-N-T--------V----I.N-S..........TN-T----------------------RL--------E---I---....  508
B.ES.10.DEMB10ES002        -----N--NE.......TA-........N..E......T..--V..NN......-E.....-----K---------R--G----------R-S-----D-------------.T-E..........TT-----------------------Q---I-I-----------....  515
B.FR.11.DEMB11FR001        -P--D---NG.......GEP.....V-N...........-TG..GN......DT.....I---Q-------------E----------R--------------I-----DL.N-T..........QN-T--------------------------------------....  490
B.GB.05.MM45d213_GN1       --------NS.......TEG-..I-.--N...........STE..KN......IT.....L----------V-L--G----I-----R-L-S---------I-------.GX-I........XNNT-T--------------------------------------....  512
B.HK.06.HK003              -K--D-S-NG.......TN-G.......N...........DTD..AE......N-......---------R-------E----------K---K-----------------T.NTT.........NGT-----------------------R-----I-----R-E----....  503
B.HT.05.05HT_129389        -K---N--NG.......TDF........I...........YPW..NK......TE.....N----------------G-------------P---E--------------D.Q-GT.......DATN-T----------------------R---------E-------....  512
B.JP.12.DEMB12JP001        -K------N-........-N-P.......N...........DTM..PN......NN.....-----------V-----------N---A-L------------V------.TT-..........TTT-T--------------------------------------....  513
B.KR.07.HP_18_07JHS10_3909 -K------RF.......NSTW.......N...........GTE..-I......N-......P-----------R--E----------R---K-----------------T.E--K.....TGSNRT----------------------R-----I-----R-----....  506
B.PE.07.502_0525_wg5       -K----SGD-.......--D.....................SDQ..DK......N-......-----------------E----------Q---S----------------.-S-.........TNGT-------------------------I-----------....  491
B.RU.11.11RU21n            --------DK.......NITE.....FN............KTNELNN......TE.....I----------------A-------------K--------------------.............NDT-T----------------------R-----I-----------....  492
B.TH.08.MERLBDTRC10        SA---H-FN-.......TEGE.......N...........ITE.SIN......TT.....I-K-------------E--------------S-------I-------G.N-T..........GN-T------N-K-------------------I-----------....  512
B.US.11.ES38               -K------EK.......TNDT.....W.S...........G-E..TL......NG..TITP----------------I----------P-P-N--------I------G.DRS........NATN-T-------------------------L-----Q-----....  511
C.AR.01.ARG4006            -------YXT.......NGTE.......I.........T..GW-..ND......-X.....P--IS-----------E--R-------R-M-T-R-----I-------T.NGT.........NNT-------------------------E-N---I---X--------....  505
C.BR.07.DEMC07BR003        SS----ELDS.......NGTF.......K.........I..-G-..-NGTGTE.NS.....N---Q-----------E--R-------A-E-T-R----------------D...........TSD-------------------------EVK-------E-------....  529
C.BW.00.00BW5031_1         S------YRA...............................-S-..NSTS.....NS.....N----------------G--R-------A-N-I-K-----V---Y---E.GV..........K-N-T---A--N-K--------------E-K---I-----------....  493
C.CN.10.YNFL19             S----G-YNG.........TY.......M.........S..-G-..-S......NA.....---I-----------E--R---N--A-N-T-K--------V----E.GT-.........ATN-T----------------------E-K---L---E-------....  511
C.CY.09.CY260              SK---Y-YPY.......NGTH......................SH.......NG.....-TITQIV-I-LVIR--E--R-IS---TA-ENISG-Y--RV-WPTG--K.N*G.........QNDT-T----------------------E-K---I---E-------....  502
C.ES.08.X2363_2            -K---G-YS-.........GTY.......S.........PG.Y-S..SSHS.....NN.....S-IIQ--------------Q---------A-N-T-N--------------S.--..........DTK-----T--N-K--------------E-K---I-----------....  515
C.ET.02.02ET_288           SK---T-GLS.......D..........................................NA.....-L-I-----------E--R-------K-I-T-R--------------M.N-T..........NNT-T----------------------E-K-------G---A--....  501
C.IN.09.T125_2139          SG---R-YTT..DL..FNKTY.......N.........D..-D-..-DNS.....GP.....V-II-----------E--R-------A-N-T-N-S--------------.D-K........TSTN-T---L--N------N---------E-K---I---T-------....  525
C.KE.00.KER2010            SG---E-LTR......LNDTY.......S.........SN.........................GT.....N-----------------R-------A-N-T-T-K------Q-----.-T..........NKT-T------N-K--------------E-K---I---E-------....  505
C.MW.09.703010256_CH256.w96 SKI-D-..............................D-A..NKT.....-G.....N-----K---------G--R-------A-N-T-T-----------------.VT..........NDT--------N--------------E-K---I-----------....  496
C.TZ.08.707010457_CH457.w8 SK---N............................-I-..SDN.....-S.....N--I--K----------------------A-N-T-T-----I---------.NSD.........SDP-T---T------------------EV----I---G--------....  488
C.YE.02.02YE511            SK---R-YMF.......NSTY.......M.........P..-S-..NSTS.....NS.....-----------------G--R-------A-N-T-T---------------G.KD..........N---T----------------------E-K---I---A-------....  505
C.ZA.08.705010534_CH534.w12 SR---G-YMS...................M.........S..----..KNDT.....NA.....-----------------G--R-------A-D-T-K-T----I------E.ESK.......TNNT-----A----------------E-K-------E-------....  496
C.ZA.10.DEMC10ZA001        SK-WDN.................................Y..TE-..T-......NG.....--I-----------------R-------A-N-T-----------------.TTDN......TTNNNT-V------NIK-------------E-K---I-----------....  496
C.ZM.11.DEMC11ZM006        SG--ENRTYT..............................L..-D-..GSTSNNT.NS.....---IQ-----------R--Q-------A-N-T-K--------------A.KK-E.......SDPP-T----------------------EVK---L-----Q------....  506
D.CM.10.DEMD10CM009        SY---G--I-.......G-RN.......D...........TR-..N-......TG.....DTII-------------G----------K---K---TV---I--V-EV-.NSS.........PEAT----------------------R-----I-----------....  494
D.CY.06.CY163              SK---T--SS.......E--E.......N..S.....TG..----..-L......NG.....---I--T-----I-C-G--V---T--PA-I-Q-K----ITG--------.NS-.........CN-T----------------------R-----I---R-------....  507
D.KE.11.DEMD11KE003        SK--TTNITS..................N...........................TS......EG.....N----------------G--G------RE-L-T--------------QEG.---.........IN-T------N-K-------------R-----L---R-------....  500
D.KR.04.04KBH8             SG------MV.......NSTV.......N............QTG..LN......NS.....----Q-----------E--R-------E-V-K--------I-----A-.N-..........TN-T------N-K-------------QV--I-L-----------....  492
D.SN.90.SE365              SG------KT..................................KS......G-.....N-----------L--R----------E-V-S----------L----I.H-T........SN-T-T----------------------RL--------T--------....  500
D.TZ.01.A280               SL---T--IK.......G-Q........N...........-TE..TN......NS.....---I--------------G----------A-L---T---------------.V--.........SRE----------------------R---I-----R-------....  502
D.UG.10.DEMD10UG004        SG------QE.......NSTL.....LN..K......T..-Y-..-E......NK.....-----------------RI---------E-L-K---------E---SV.E--.........SQ-TV--V--E--N-------------R-K-I-L---R-------....  497
D.UG.11.DEMD11UG003        SE----I-L-.......N-GN....WTGN.T.....TN..STE..KN......NG.....-----------------G----------E-E-K-I----------------.N-A..........TN-T----------------------T-----I-------G---....  507
D.YE.02.02YE516            ST--TR-INT.......TNDN...T.V.N..S.........TGN..NT......N-......---I-------------G--R-------A-I-K----------------S.N-T.........NST-T----------------------R-------------------....  496
D.ZA.90.R1                 SG------NR.......TGNN.......S...........SS-..GS......NE.....--------R-------G----------E-L----------I------D.NSS.........SHN-T----------------------R---I-L---R-------....  508
F1.AO.06.AO_06_ANG32       SG---NSD.......................................TS......NR.....-----------V----G--Q----A-VA-K-T-N---------------.............NSHT-T------N-K-------------EL----------R-----AR..  490
F1.AR.02.ARE933            SG---D-........................................DD......NG.....-----------V----E--R-NNH--HA-N-T-G-------------T.G-CTD...GGTGNCT-T------N-K-------------E-----------R-Q-------....  495
F1.BR.10.10BR_RJ015        ----STK........................................S-......N-......-----------V----E--R---N-VA-K-T-N-T-------------.DKST.......XNNT-----W--N-----------------I-------P--....  496
F1.CY.08.CY222             SE---NSN........................................SN......NG.....------K----------G--Q-------A-K-T-N-S-----------Q.--KSK......DNG-------R-NIK-------------E-----I-------Q--....  494
F1.ES.02.ES_X845_4         SR--TDSRT........................................-F......N-......------K---V----G--Q----A--A-N-T-N-------------.N-T.........S-N-T------N-K-------------E-----I-----------GR..  492
F1.RO.96.BCI_R07           SG----STSP.......EGV.......................F--..SH......KG.....N---L-----------G--Q----A--A-S-T-N--------------Q............NN-T-T------N-K-------------E-----I---R-Q-Q--....  508
F1.RU.08.D88_845           SK---H-V--.......D-NN.......I.........V..-GS..SL......NG.....-----------V----G--R----A--A-K-V-N-S------------X.NDT.........RTN-T---A--N-----------------R-----------H--....  512
F2.CM.02.02CM_0016BBY      -A---T-LL-.......T-M.............................DN......NG.....--II-------V-V--R--R-------A-K-Q-N---------------S.KA-...........NTD-L--I--E--------------Q-Q---I--SR---Q--....  495
F2.CM.10.DEMF210CM001      SA---TEALK.................................----..GQ......NG.....-T-IY-----------R----------A-K-Q-N-S----I--I--NR.T--..........K--TL-------K-------------------------R-------....  505
F2.CM.10.DEMF210CM007      D----ISQSE.......KA..............................GN......NG.....S-F-----R--V--------R-------A-V---N-S-----------E.D-S..........STDTL-------K-------------------------R---Q--....  497
G.BE.96.DRCBL              SG---NSILK.......-NI-...........................-N......N-......-----N-K----VR---R--Q-------A-N-T-R------I------D.N-S...........T-----------KN-----------T---KS--I---R-R-----....  497
G.CM.10.DEMG10CM008        SR--.......................-.........PL..-S-..NS......TG.....N----------VR---R--Q-------A-K-T-I-----------------.N-............SI-T---A----K---------------K---I---R-R-----....  497
G.CN.08.GX_2084_08         SG---NSKL-.......--T.............................TN......NT.....-----------VR---R--Q-------A-N-T-R------------A-.NT...........GN-T-------------------I---K---I-----R-----....  501
G.CU.99.Cu74               SK----SXMX.......NS...............................T-......NG.....PFK---K----VR---R--Q-------A-N-T-V--------------T.N-T.........NRT-V---A-----------R------K-------X-R-----....  508
G.ES.09.X2634_2            SE----SYL-.......-.................................TS......N-......-----------VRL--R-EQ-------E-N-T-R---------V------.NDT..........K-N-T---A------------------K---I---R-R-----....  502
G.GH.03.03GH175G           SR---NETG-......................................ST......NT.....N-----K----VR---R--Q-------E-E---D-T----I--------.---........TNNT-T--------------------T---KS--I---R-R-----....  505
G.KE.09.DEMG09KE001        SR---M-ELD..................K...................S-......NE.....------K----VR---R--Q-------E-K-T-N--------------E.N-S..........TN-T---A-----------------IR-K---------R-----....  501
G.NG.09.09NG_SC62          -G---TSG--.......D-...............................IT......N-......--M--------VR-------Q-------E-S-T-R---------M-----.N-T.........NKT-T---S-----------------I---K--EI--G--R-----....  498
G.ZA.01.TV546              SL---RNETS.......ENLT.......V.........-GN..KT......GE.....N-----N----VR-------Q-------E-N-T-X-------------K.N-............S------------------------K---------R-----....  498
H.BE.93.VI991              -K------T-.......-SYT.......N..DT..YNSN..STE..DI......TG.....N---Q-K----V----R--Q-------R-N-T-I-------I--F-RN-............TNNVT----------------------------E-R-----....  509
H.CF.90.056                SG----S-EM.......H-NY.......T.........S..-D-..K-......NE.....N-----------V----R--R-------Q-N-M-V-------I--I-E--.ASA..........-NYT-------------------------I----TR-----....  495
H.GB.00.00GBAC4001         SR------NS.......NDTG.......N..G.............TD......NS.....------K---------R--Q-------Q-N-T-V-------I--L-RYV.D-G..........TNVTL----------------------Q-----I-----R-----....  502
J.CM.04.04CMU11421         SS----S-SS.......KNDT.......N..S..............TA......NY......TII--K-R--VR------Q-I-----E-N-T---------I-------D.K-ST......SNNST-----T----K-------------E-----------R--------....  510
J.SE.93.SE9280_7887        SK----S-DK.......NSIE.......A...........T-D..TS......XA.....---I--K----VR---RT-Q-I-----A-N-T-T----------------.RG-G......SENGT-T---T--N-K-------------E----------------....  498
J.SE.94.SE9173_7022        ST----S-DE.......NNIK.......D..T.........-S-..ND......NT.....---I--K----VR---RT-Q-I-----A-N-T-K----------------.R-G.......SENGT-T---T--N-K-------------EL---------------....  500
K.CD.97.97ZR_EQTB11        -DTVDD-EE.........................................-E......DT.....---I---------------Q-I----TA-N-T-R-----MI--------.D--.........TRTE-T----------------------Q-----I---R-R-----....  501
K.CM.96.96CM_MP535         -K---E-GE...........................................NG.....--------------------I-----A-S-N-------MI--------.NT...........HN-T----------------------Q-----I---R-R-----....  492
U.CA.01.TV749              SK----S-WF.......N-TN.......S..I...GINT..-S-..-I......N-......--I--------VR---R--Q-------A-N-T-V-------------KY.RP-.........STN-T---T--E---I--N---N----E-------------R-----....  511
U.CA.99.TV721              SG---KS-GS.......WEPN.......T..T...EV.N..-V-..KE......NG.....----Q-K----VR---R--Q-V-----A-N-T-T----------------.N-S..........DN-T---T------I------N--I-E-------------R-----....  514
U.CD.83.83CD003_Z3         SE--TGI-NG.......TWDK.......N......CTS..TES..NC......TG.....N----------VVRT--G--Q---------E-T--------------------.G-A..........TQN-T--------K-------------------------R-------....  501
U.CD.90.90CD121E12         SG------NG.......TWAE.......N..N....CTNSES-S..SC......NE.....---I-------VRT--G--Q-------K-T-M-N-S-------------D.NE-..........DTK-T-------------------------------R-------....  499
U.CY.05.CY090              SK-W-D-VGS.......NGIA.......G.............SN......EN.....-----RPFRL-LS----L-L-L-----A-KST-T--S--PI---A-DW............TRT-T----------------------R-----L---R-R-----....  495
U.CY.08.CY223              SK---G----.......G--V.NGT.WYN..N....TLD..S-D..TE......NG.....------K-R--VR---R--Q-------Q-V-H-Q--------------DK............N-T-T---T---------------------------R-R-----....  530
U.ES.10.DEURF10DZ001       ----D-S-GI.......NDTV.......S..N......S..-DA..DS......D-.....N---Q-------K---R--Q-I-----R-N-S-H----------------.NS...........TN-T--------------------I-----------H-------....  496
U.GR.99.99GR303            SG------RP.......NDMQ.....W-..............-D-..#AI......*R.....I------#---------R--Q-------Q-E-K-N-T--------------.-S-..........D--T-------------X-------------------G-------....  504
U.NL.95.U_NL_95_H10986_D1  -R---T-HLY.......NSTG.......N..D....TQD..--S..TN......-K.....N-----Y----VR---R-EQ-------A-R-Q-I--------------K.N-E.........TNGN-T---A--N-K-------------------I-----R-----....  511
01_AE.AF.07.569M           -R---N-CIK.......NATV.......R.............GC......N-......--I-------------EA-Q----------I-N-V-----I---------.-S..........TN-T---E--NIK-------------Q-----I---R-------....  501
01_AE.CF.90.90CF11697      -K------TT.......NEIM.......E..E.........FKG..TN......-S.....-----------VV----E-----------EAVN-V-----II---------.............ATN-T------N-K-------------Q-----I---R-R-----....  506
01_AE.CN.10.YNFL03         -K---K-NIE.......NETV.......K.............SD......N-......--I-------VR---G--Q----------N-M-V-----I------D.N-............ST--------N-K-------------Q-----I-----------....  502
01_AE.HK.04.HK001          ----D-S-NT.......T-TN.......R.............-N......RS.....NFI-------------ET----------R-S-Q-V-----I--------.N-G.........SN-T-T------I---------------Q---------R-------....  501
01_AE.IR.10.10IR.THR48F    -K------G-...RTENNGTI.....................--......NG.....--I---K---------ET-Q-------V--I-N-L-----I---S-T-G.K--.....GTGENGT-T---A--N-K-------------E-----I-----------....  510
01_AE.JP.x.JRC77AE         SK------G........NGTI.......E.............GH......NG.....------K----------GA-Q-------R-K-N-V-----I---------.-SA..........NN-T------NIK-------------Q-----L---R-------....  499
01_AE.TH.09.AA090a_WG11    -----R-IG........NVTG.......E..G......S..-DE..NN......NK.....--I-------------GA-Q-------R-K-K-V-----I---------.I--..........ATN-T------Q-K-----R--------Q----------R-R-----....  507
01_AE.TH.90.CM240          -K---N-CLG.......NETM.......A.............GC......N-......------K----------GA-Q----------R-N-V-----I-------V.N-T..........DN-T------NIK-------------Q-----I---R-------....  500
01_AE.US.05.306163_FL      -S---N-YIG.......-NG........T...............KG......--I---K---------G--Q----------R-N-V-----I-------L.N-............TN-T------NIK----X-------Q-----L-----------....  479
01_AE.VN.98.98VNND15       -K---D-CIK.......NNS........R.............GC......N-......--I-------------R--Q----------R-N-V-----I-------T.NSRTN...ETNETNE---------IK-------------Q-----I---R-R-----....  509
```

```
                                 glycosylation NST
                        glycosylation NST
                                                   V4                                                                    V5
              glycosylation NST                          glycosylation NNT              CD4 binding        glycosylation NIT  CD4 binding   glycosylation NES    CD4 binding
B.FR.83.HXB2                 TQLFNSTWFN.......STWS..TE.GS............NNT..EG......SD.....TITLPCRIKQIINMWQKVGKAMYAPPISGQIRCSSNITGLLLTRDGGN.SNN..........ESEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRRVV....  506
02_AG.CM.10.DE00210CM013     SA------LF.......NETY..ST.WHS............SWN..SS......AE.....N---Q-----FV----R--Q-------R-M---E--------------.K-T..........SAN-T--------K------------I----------P-------....  527
02_AG.CY.09.CY256            SR------DT.......--GY.......D.............-SS..RD......N-.........----K-K-R-VVR---R-AQ-------Q-E---N-T-------------.-SG..........AN-T-------------------------------H-------....  500
02_AG.ES.06.P1423            SR------D-.......GNLTFNGSQSTLK.N.........TS-..NL......NE.........-----------FVR---R--R-------P-V---E-----I---------.K-E........TEGN-T---------------------Q---I-----H-------....  505
02_AG.GW.05.CC_0048          SA------KH.......D-KF.......D..N.........TSV..GP......Y-.....N---Q-----------R--Q-------P-V-Q-E-------------K.I--..........TN-----------------------E---------Y-------....  510
02_AG.KR.12.12MHR9           SK------D-.......NSTWDNNSTWEN..N.........S--..NS......-G.....---FQ-K----V----R--Q----L--Q-V---E--------------.K--..........SNN-T-------------V-----------------H-R-----....  512
02_AG.LR.x.POC44951          SI------D-.......--VN.......T.............-I-..-S......N-.........----Q--------------Q-------K-I---D--------------.N-S..........TT-T-----------------------D-I-----H-------....  495
02_AG.NG.09.09NG_SC61        -P--K-V-NK.......NGST.....WNS..N.....TT..-S-..-S......NN.........----Q-K------L--R--Q-I-----P-V---E--------------.IDD..........DAN-T---------------------------R-R-----....  504
02_AG.NG.x.IBNG              SK------D-.......-NST.......A.............-H-..GS......N-.........----Q------V--------Q-------Q-I---D--------------.NS...........TN-T---------------------------R-------....  497
02_AG.SE.94.SE7812           SN---R--NH.......NGTW.......NAPGPFNDTEDKTI-G..TE......DK.........----Q------VR-------Q-------P-E---E--------------.D--..........NT-T--------------------------SH-------....  508
02_AG.SN.98.98SE_MP1211      AE------AS.......N-NG...I.WAS..N...INAS..--K..DA......N-.........----K-K-------------Q-I-----E-V---D-----I-------D.NT-..........GD-T--------------------------SH-------....  503
03_AB.RU.97.KAL153_2         -K------NG.......TEEL.......N.............-TE..........G-.....IV--------------E----R-----A------------------.QS-..........VT----T-------------------------R-------....  490
04_cpx.CY.94.94CY032_3       -P-----HMQ.......NGTN.......I..T.........STD..ST......NS.........----Q--L--FVR---E--Q----S--A-S-N---D---II------T...........NNT-------------------------I----N--R-----....  505
05_DF.BE.x.VI1310            SK---A-V--.......D-VF.......N..A.....TM..F-N..DS......DK.....N-I---K-----VR---G--Q----A--A-N-A-N-T---I--A------.G-DS.......SNDT---------------------E-Q---I---R---Q---....  511
06_cpx.AU.96.BFP90           SN---TSDLF.......N-S......................RGN..DT......NT.........------K-----VR---R--Q-------A-N-T-V------II------N-.E-..........V--T----------------------K---I---W-R-----....  509
07_BC.CN.98.98CN009          SG---G-YM.........................P..-D-..KSNS....-S.....I--I--------------E--R-------E-N-T-K---------V----T.ET...........NDT-T---------------------E-K-------A-------....  502
08_BC.CN.97.97CNGX_6F        SG---G-YNG.........TY.......M.........L..-G-..RDNS....-S.....I--I--------------E--R-------E-N-T-K---------V----R.TEP..........NNT-----------N---N--------E-K-------A-------....  507
09_cpx.GH.96.96GH2911        SG---N--NG.......TWE.......................--N..NS......IE.....NY-IH------VR---R--Q-------P-E---R-------I------S...........DDN-T---T--N-K---------------------R-R-----....  492
10_CD.TZ.96.96TZ_BF061       SE------TS.......NNT......................STG..VN......GS.........--------R--------G----I-----A-L-N---------V------A.N--..........SQN-T------------------------I-L---T-------....  490
11_cpx.CM.95.95CM_1816       SK------NS.......NGTV.......G.........N..STI..GP......N-.....N---S-K----VR---R--Q-------Q-K-M-N-T------------K.NSS..........KT-T---I----K--------------E-K-------R-------....  515
12_BF.AR.99.ARMA159          SG---D-...........................................IS......N-.....--I---K----V-L--E--R----A--E-N-T-N--------------L.NET..........NQT-T-------N-K-------------E---------R---Q---....  492
13_cpx.CM.96.96CM_1849       SA------E-.......D-V-...T.Q.G.........S..-D-..GS......N-.........-----------FV----R--Q-------Q-I---E---------S----V.G--..........RTN-T-------N-K---------------------SR-R-----....  509
14_BG.ES.05.X1870            --------SS.......AWNG.......T.............TG-..NN......TG.....V---------V--L--E-----------------------V----R.--D........SSKP--------N-K-------------------------R-----....  508
15_01B.TH.99.99TH_MU2079     S--------Q.......NGN-.......T.........W..-D-..A-......NG.........---------R--V----E--R-------K--------------------K.NE-T.......TNTT-----A---------------I--------------------....  496
16_A2D.KR.97.97KR004         SG------PA.......NASR......................E-E..-K......DR.....NV---------V----R--R-------N-T-K-T-----MI-----NS.GG-..........ATN-T---------------------L--------R-R-----....  506
17_BF.AR.99.ARMA038          SK---DSRD-..................................D-..-S......N-.....--I--------V-R--E--Q----A--A-N-V-N--------------Q.G-G..........SGP-----E--N-K-------------E-------------------....  498
18_cpx.CU.99.CU76            -A------NK.......TXGT.....WIKL.......DN..TTN..XQ......NG.........----Q------VR---RT-Q------SR-N-T---------I--S-LIX.HXT.............-NVT---------------------------R-------....  506
19_cpx.CU.99.CU7             -D----..............................................#EE.........----Q-----VV----#--Q-------P-V-Q-K-----------------...............N-T---A--G--------------------------------....  452
20_BG.CU.99.Cu103            SG----SN...................................NN......-G.....N----------VR---R--Q-------A-S-T-I-------I-------.NST..........NRT-----A---------------------K-------R-------....  492
21_A2D.KE.99.KER2003         ---------NS.......TES-...T.W.N...........STE..SN......NS.........------------------E-----------E-H-N--------------T.N-T..........SNM---------N-K-------------R---------R-------....  502
22_01A1.CM.01.01CM_0001BBY   SG------HD.......N-T......................GLD..NT......NG.........--I---------------RA-Q-------P-T-T-X-------------D.K-S...........AN-T------------------------I---R-------....  491
23_BG.CU.03.CB118            SG----SN...................................NN......-S.....N----------VR---R--Q-------A-S-T-V------II------P.NSX............N-T--LA-------------------K---------R-------....  498
24_BG.ES.08.X2456_2          SN----SN-...........................................NS.....N----------VR------Q-------A-S-I-V------IIS-----Y.E-S..........SGT-T---A-------------R------K---------R-R-----....  503
25_cpx.CM.02.1918LE          SS----S-T-.......--A......................-G-..D-......N-.....N---Q---R-------R--Q-------Q-V---------I-----A-.TA-..........NT--------------N--------R---------R-------....  500
26_AU.CD.02.02CD_MBTB047     SG---R--KD.......GN-T.....................T..QEN..GT......NE.....N----------VR---E--Q------VA-M-Q-I-------I------K.DK-...........DT---------N-K--------------Q-----L-----R-------....  501
27_cpx.FR.04.04CD_FR_KZS     KE------NK.......--N...GT.W.T..K...NSTT..GE-..GN......TA.....N---Q-----V----RT-Q-I-----R-I-Q-E-----I-----RH-............HTN-T-------N-K-----N--------E-----I---R-------....  506
28_BF.BR.99.BREPM12609       -L------KW.......NNTW.......N..N..........TGK..GA......DK.....N---L--------R--E--------------E--------------------S.N-E........TNMT-T---A---------------------------M-R-----....  517
29_BF.BR.01.BREPM16704       SR----V-RL.....DNN-SW.......EQDNNTRN.....GTN..SN......N-.........----------------LA----TSP---R-------T--L--S-N---S.N-T........TNRN-T----------------------R---------------------....  520
31_BC.BR.04.04BR142          SG-----YYS.........TY.......N..S.....TV..S--..TA......DV.....N--I--------------E--R-------K-A-T-I-------------DK..............NDT-T---------------------E-K---I---------....  492
32_06A1.EE.01.EE0369         SK----SSRS.................................................NE.........------K-R--VR-----AR--------A-N-T-T-------------SK.ES-.............-T--S-----N-K---------------K---I--S--R-----....  500
33_01B.ID.07.JKT189_C        -R---N-CL.......NETM.......G...........................NC......NG.....--I---K----------RA-Q---------T-N-V-----I-S-----V.NSS..........RNN-T------N-K--------------E-----I---------....  502
34_01B.TH.99.OUR2478P        -K------PE.......NGTM.......E...........................GS......NG.....--I---------------G--Q---------T-N-I-----I----------.-TT...........DT-T---E--NIK-------------Q-----I---R-------....  493
35_AD.AF.07.169H             SG---NSL--.......-SGE...............................SD......-T.....N---------------RT-Q-------R-E---V-------------A.NE-..............---T------------------------------------....  497
36_cpx.CM.00.00CMNYU830      SG------GY.......NNSN.......V..T.........RLN..NT......N-.....--I---K-----K---R--Q-------A-R---E--------------K.N-S...........TN-T---------------------Q---------R-------....  499
37_cpx.CM.00.00CMNYU926      SG------NT.......T-ST.......Q.........I..P-R..TT......DG.....N-----------V----R--Q-------P-V---E-----M------I.NKT...........DN-T-------------------------I-I------R-------....  495
38_BF1.UY.03.UY03_3389       SK---D-AS-...........................................NG.........------------V----E--R---TA--A-D-T-R--------------Q.K--GS...CTEATCTQ----Q--N-K-------------E-----I---R-------....  494
39_BF.BR.04.04BRRJ179        SN---D-..................................I-......NG.....--I---K-R--V----E--R----A--A-N-A-N-T------------S.GS-G.......TNQN-T------N-K-------------E-----I---R-Q-Q--....  495
40_BF.BR.05.05BRRJ055        SG--DNNTGT................................DT......N-.........------K----V----E--R----A--A-N-T-N-S------------Q.N-E........TNNT-T------N-K--------------------R-R-----....  501
42_BF.LU.06.luBF_18_06       -I---E-G--.......D-......................RS......N-.........------------------E--R---TA--A-N-A-------------K.TTD..........NQTDT---------------------E-----I--N--R-----....  506
43_02G.SA.03.J11223          SR--G...................L..............F-X..S-......--.....NL-I-------VR---R--Q-------E-I-T-R--------------.I-S...........TNDT--------K---------------K---X-----R-----....  489
44_BF.CL.00.CH80             SG---DSIPI...............................SNA.........--.....------------V----E--R----A--A-R-S-N--------------K..........TDXN-T------N-K-------------E-----I-------Q---....  494
45_cpx.FR.04.04FR_AUK        S-------SH.......NDT........G.............STK..-Y......NG.....--Q--------------RA-R--F----Q-A---E-----------------.A-EA.......RN-T----------------------Q---------R-R-----....  494
46_BF.BR.07.07BR_FPS625      SG---D-GS-.........................................NE.....--I-------M-----E--R----A-VA-N-T-N-------------P............GNKTQ---------K-------------E-----I-------Q---....  485
47_BF.ES.08.P1942            -S------E-.......D-NI.......T..E.........ESN..SS......D-.........----Q-K------L--E-------------Y-N---------I-V------.NRT...........S---T---E--N-K-------------------I-------------....  498
48_01B.MY.07.07MYKT021       -K---D-CIE.......NNT......................RC......NH.....A-I---------------G--Q-------N-K-N-V------I-------R.NSS..........DGN-T------NIK-------------Q-----I---R-------....  510
49_cpx.GM.03.N26677          ------S-NK.......NANG.......T.............TGN..TT......NS.........----Q------VR---RI-Q-I-----A-N-T-T-----------------.NT-..........NT-----T-------------------ELK---I---R-------....  505
50_A1D.GB.10.12792           SG----I-NS.......---E.......-N..S...TEL..-S-..-L......N-.....N---Q--------------G-----------E-K-I-R--------------A.N-T..........NQN-TL------------N---------RL----I---R-R-----....  506
51_01B.SG.11.11SG_HM021      -R----N-TG.......DENV.....................TGH..G-......N-.....I-I-------------RA-Q-----------K-N-V-----I-------D.N-S...........TNDT---A--NIK----N-------IQ-----I--------------....  501
52_01B.MY.03.03MYKL018_1     -L---NNCTE.......NETG.......T.............GC......NE.........--I--------------G--Q---------I-N-I-----I---S--DK.-K-..........D---------N-K-------------Q-----I--S----------....  510
53_01B.MY.11.11FIR164        -K--K...TE.......NETR.....................................-I--------------G--R---------S-N-V-----I-------R.N--...........GT-----E--NIK-------------Q-----I---R-------....  498
54_01B.MY.09.09MYSB023       -K--D-I-YG.......NETI.......N..E.........TIK..GN......-S.....--I---K--------------Q---------N-S-V-----I-------I.NET..........RNN-T---A--NIK-------------Q-----I---R-------....  507
55_01B.CN.10.HNCS102056      -K------IMENDTWTDNKTM.......EE.G.....VN..GTM..GN......NS.....IFII-------V----GT-Q-------R-P-N-V-----I------------.--G...........SN-T-------IK-------------QL---------R-------....  509
56_cpx.FR.10.URF5_patient_A  SG------ND.......TSTW.......N.............DT......DR.....NV---------F-----R-ER-------R-D-K-A-----------------.N-S...........TN-T---A------------------------I---R-------....  495
57_BC.CN.09.09YNLX19sg       SG---G-YGT..........Y.......D.........P..-G-..--NS....-SEGNSKF-I--------------E--R-------E-N-T-K---------V----T.Q-E.........TKNE-----E------------------E-K---------E--------....  513
58_01B.MY.09.09MYPR37        ST---R-...................................TS......RN.....VTII--K----------E--Q---------S-N-V-----I-------F.N-ET........GN-T-T------NIKN-------------Q---------R-------....  506
59_01B.CN.09.09LNA423        -P----I-WP.......NGTG.......Q..S.....N..R-G..TI......NG.....--I-Q-------------------T--SK-P-------------------.ND-..............-------------K------------------L-----------....  502
60_BC.IT.11.BAV499           SG--D--YYW.................................DKE..GNITIQI#-G.....N---------------E--R-------K-K-T--------------T.N..........STN-T---------------------ELK---I-------------....  503
61_BC.CN.10.JL100010         S----INDTR..................................-D-..-KN..........IT----K---------G--R-------E-N-T-R---------V----S.N...........TTT-T---------------------E-K---I---P-------....  491
62_BC.CN.10.YNFL13           SG---R-YI-.........GT.......V.........H..-----..-N......IT.....I--I--------------E--R-------A-N-T-T---------V----K.T-E.........NTT-----V--N---------------E-K---I---T-------....  501
63_02A1.RU.10.10RU6637       SA----S-NS.......TGN........S.............TGV..NI......T-.....P-R---K----------R--Q-I-----P-V---N-----I-------T.GT-S.......TN-T-T------------------------------G-------....  497
64_BC.CN.09.YNFL31           SD---G-FNG.........TY.......I........PF..-D-..........IT.....N--IR-K----------E--R-------E-E-T-K---------L-----.T--........TNNT---------------------I-E-K---L---A-------....  502
65_cpx.CN.10.YNFL01          SG---G-YNG.........TY.......T.........P..--D..TD......NS.....--I-Q--------------E--R-------VA-N-T-K-----I-------V.N-..........NTN-T-----------------------------------------....  513
67_01B.CN.11.2011.ANHUI.HF115N---E-M...................................-S......NS.....--I-------V----GA-Q-------E-N-N-V-----I-------D.NXX..........XN-T------NIK-------------E-----I---R-------....  490
68_01B.CN.11.2011.ANHUI.WH73-K-----E.........NETR.......E.........................-L......NG.....-VI---K----V----GA-Q-------M-N-N-V-----I-------K.NA...........TN-T------N-K-----N------I-E-----I---R-------....  491
70_BF1.BR.10.10BR_PE004      -K------I-.......G-RN.......X.............TXE..SD......XX.....X-X--------------R----------Q---K-------II------D.NTT..........T-T-----A------------------------I-------------....  500
71_BF1.BR.10.10BR_PE008      -P------NT.......TERH.......N..W.....N..STR..RS......N-.....I---R------------GIK---------Q-------------I-----SS.EG-..............-T-T------N-K-------------R---------------------....  505
72_BF1.BR.10.10BR_MG002      SN---NS.................................I-......NG.........------------V----E--R----A--A-N-T-N--------------S.G--..........TNDT--------N-K-------------E-K---I-------Q---....  497
O.BE.87.ANT70                AKM--Y-FSC.......NGT........TC.S......V..S-V..SQ......GN.....NG----KLR-VVRS-IRGQSGL-----K-NLT-M-----MI-QM-NTW.NSS..........NNNVT---I----K-I--T--FN----RVK-FS----RIA-P-ISTRT  504
O.CM.98.98CMA104             SGM--Y-FTC.......NKT........NC.T......N..-R-..DN......-T.....ND-I--KLR-VVRS-IQG-SGL-----K-TLT-M-----MI-EMERSW.NGSATKDINATWNINAT---I----K-I--TQ-FN----RVK-FS-----IA-P-IGMGT  520
O.CM.98.98CMABB141           S-M--Y-FTC.......DGT........TC.K......N..LGF..NS......TN.....NTQI---L--VVRS-I-GESGL-----Q--LT-R-----MI-QM-QPW.NGSQ.......SNSNST---I--N-K-I--T--FR---IR-K-FS----RIT-P-IGFGH  528
O.CM.98.98CMABB212           S-M--Y-CNC.......TSSE......CNC.T......R..PGY..N-T......NG.....NE-I---LR-VVRS-I-GESGL-----P-NLT-Q-----MI-QL-APW.NSTC..........TKT---V----K-I--T---N----Q-K-FS-----VS-PIISINT  529
O.CM.98.98CMU5337            SGM--Y-FSC.......NGT........TC.N......NT.......QS......NN.....NTLI--KLR-VVRS-MRG-SGL-------YLT-M-----MI-QM-EPW.KGS...........NAS---T--Y-Q-I--T--F------VK-FS-----IA-P-ISTGT  506
O.CM.99.99CMU4122            --M--Y-FSC.......NGT........SC.T......N..-QS..NSX.....NK..NGTLTEI--K---VVRS-IRGESGL-----K-PLK-E-----MI-QK-EPW.NKTN.........NTHAT--------K-I--T--------RVK-FS-----IA-P-IGTGT  513
O.FR.92.VAU                  N----H-FSCKKNM.TYNKI.......NC.T......N..ISN..NS......NG......-QAI---LR-VVRD-MRG-SGL-----P-NLV-R-----MI-QL-TPW.NKTH........PN-TTL-------K-I--TQ-F-----RVK-FS-----IA-PTIGTST  517
O.GA.11.11Gab6352            SKM-CY-FNC.......-RSN.......NC.TLIKNNCT..I-S..TS......NS......-SEI---LR-VVRS-MGGQSGL-----K-NLT-I-----MI-QL-KPW.NIS..........NRNITL--T----K-I--T--FQ-----VK-FS-----IA-P-IGTGT  521
O.SN.99.99SE_MP1299          SKM--Y-FSC.......IGT........NC.T......S..-QN..SSN......-N.....DTRIY-----VVRS-IQG-SGL----RK-NLT---L---MI-QL-MPW.NSTN.........NSNAT---T----K-I--T--F------VK-FS-----IA-P-IGTGT  517
O.US.10.LTNP                 GKM--Y-LEC.......NGT........IC.ENI....S..SES..NSSSES..KN.....KTRI---L--MVRS-ATG-SGI-----Q-NLT-T-----MI-QT--VW.DYPH.........PNNVT---T--N-E-I--T---N-----VK-FS----RIA-P-IGTGT  507

N.CM.02.DJ00131              SR---ESMNT................N...........................GT......NG.........---------R--V-L-TR---GI-----R-NLT-N-T----I-EHS--S.N.................GTVY-T--N-VNL--Q------T-S---I----G-----T-.....  476
N.CM.02.SJGddd               SK--DKNLL-..................................KT......NE.....--V--------V-L-TR---GI----VR-NLT-N------I-EYS-NS..............GN-MVY-S--H-VNL--Q--HN----S---I----S----HT-.....  477
N.CM.04.04CM_1015_04         SR--DENGTV.......................................NG.....--I--------V-L-TR---GI-----R-N-T-N-S----I-EVS-NS.................TVY-S--N-VNL--Q--------S---I----G-----T-.....  474
N.CM.06.U14296               SK--DEXQNK.......................................T......NX.........------------V-L-TR---GI-----R-NLT-H------I-EYSEDX.KDG........TGTNTTVY-S--N-V-L--Q--H-----S---I----G-----T-.....  499
N.CM.06.U14842               SK---ENET.......................................HS......NQ.........--------VR--V-L-T-I--GI-----R-N-T-N------I-ESS--G.N-I........S..NVSVY-T--Y-VNL--Q--H--E--SV--I----G-----T-.....  496
N.FR.11.N1_FR_2011           SNI-DEGKNE........................................TT......-E.........---------R--V-L-TR---GI-----R-NLT--------M-EYS-TG...............DNTTVY-S--N-VNL--H--------S---I----G----HT-.....  485

P.CM.06.U14788               SS---FSYTW......ND-TW.......RA........A..G-H..NT......-E.....N---S--L--VV-S-MR--SGLF-T--E-ELK-Q------M-E-ELPY.NDS..........SKNTTLS-I----TNI-------P----QVKA-A-----IS-PTI.MAH  525
P.FR.09.RBF168               ST---YSYTC.......N-S..........GA......C..TGKHTGE......NM.........-------L--VV-S-M---SGLF----E--LQ-H------I-D-ASPY.NA-..........SSNTTLS-T--E--NI-------P----QVKA-A----RVS-PTI.MAH  517

CPZ.CD.06.BF1167             KNWT-.......................................-GI..-I......DG.....MLIA---LR--V-H-GI-SRGI-L--RE--VK-V-S---FIM-AETND.TDG..............T-TPTFSAKVE-Y-KV-MSR----E-Q--S-----G--P--GDPE  506
CPZ.CM.05.SIVcpzMT145        SEII-I-KI-..................................KT......EN.....MTII----R--V-S-MR---GIF----R-N-T-T-----M--EIHKNR.EDQGEDQD....QNNTYVCLT--N-K-I----------I-E-Q---------SR-YA-EKQH  497
CPZ.GA.88.GAB1               S-I-TDN.....................................IT......NG.....I-I----R--VSS-MR--RGI-----R-N-T-N----------S-TPV.T--..........SGNLT---T--N-K-I-----------R----S------R-HT-ARQK  495
CPZ.TZ.06.SIVcpzTAN13        -KWI-E--L-E.................................TD......QT.....NLVA---LR--V-H-GI-A--I-L--RR-TVK-V-----F-M-NEETG.ATFS..................-KVE-Y--V--A---I-E-Q--S----TG--PEIKANH  496
CPZ.US.85.US_Marilyn         IP---M-YN-..................................TD......NT.........----K---R--V-Q-MR---GIF----K-VLS-N-----MI-DISISAVN-D..........SRNITVM-T----TAL-KN--H-----S---I----G----HT-.....  477

GOR.CM.07.SIVgor2139_287     -KM-SYNCTT.......EGC......................-T-..ETS......NT..TET-KYI--HL--VVRS-MR--SGLF----R-TL--I-----II-Q--APL..KT..........D-NSTL--I----TNI-------P----RVKA-T---------P-I..GH  509
GOR.CM.07.SIVgorCP2135con    -KM--YNCTK.......ESC......................DCI..T-.......NC......-NYI--HL--VVMS-MR--SGLF----R-TL--K-----II-Q--VPL.NKT.....NENSTDYNTL--I----TNI-------P-----VKA-S-------R-P-I..SH  504
```

```
                                          fusion peptide                                                                               glycosylation NAS              glycosylation NHT
                             gp120 end   gp41 start                                                               immunodominant region          glycosylation NKS           glycosylation NYT
B.FR.83.HXB2                 ..QREKR..A.VG.IGALFLGFLG.AAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWS.NK.............SLEQI.WNHTTWMEWDREINNYTSLIHSLIEESQNQQEK 655
A1.AU.03.PS1044_Day0         ..-----..-.--.---V-I----.----------I-----------------S-----------M-K------------V--L-----------------------N----NT--.--.............NKSE-.-DKM--LQ--K--S---QI-YN------T---I 662
A1.CH.03.HIV_CH_BID_V3538    ..E----..-.--.L--V------.----------I---A-------------S-------------K------------V-------Q------------------N----S---.--.............-Q-D-.-GNM--LQ-EK--S----I-YE-L--------- 646
A1.CY.08.CY236               ..-----..-.--.---V------.----------I-----------------S----------QM-R------------V-------G------------------N----ST--.--.............--DK-.--NM---Q--K--S---Q--YT----------I 654
A1.ES.05.X1608_8             ..E----..-.--M---V------.---------AL-----------------S-----------M-R-----V------V--L----R-------------IV---N----S---.--.............THX--.-ENM--LQ-----S---EI-YN-L-X------X 665
A1.KE.11.DEMA11KE001         ..G----..-IG-.L--V------.M-----------------N---------S-------------K------------V--M----R-------------V--N-N----S---.--.............-YNE-.-DEM--QQ-EK--S---DI-YN-L--------- 668
A1.RU.11.11RU6950            ..E----..-.I-.L--A------.----------I-----------------S-------------K------------V-------R------------------S----S---.--.............-RQE-.-DNM---Q--K-VF---N--YE-L--------- 655
A1.RW.11.DEMA111RW002        ..E----..-.--.M----I----.----------V-----------------S---K---------R---------------------------------------S----S---.--.............-QDE-.-QNM--IQ--K------E--YN---Q------- 663
A1.SN.01.DDI579              ..E----..-.--.L--V------.----------------------------S-----------M-K------------V--I-----------L-----------N----SN--.R-.............-Q-E-.-GNM--LQ--K--S---DI-Y----------- 653
A1.UG.11.DEMA110UG001        ..-----..-VF-.L--V-I----.----------I-----------------S-------------K------------V--L-S--R------------------N----S---.--.............-QDE-.-QNM--LQ-----S---D--YT------I---- 648
A1.ZA.04.04ZASK162B1         ..E----..-.--.L--V------.----------I-----------------S-------------K------------V---------------------I----T----S---.--.............-YDD-.-DNM---Q-EK--G---E--Y--L----V---- 658
A2.CM.01.01CM_1445MV         ..E----..-.--.L--A------.----------I-----------------S---K------Q--K------------V--I----Q------L-----------T----S---.--.............NQS--.-DNM--L---K--S---ET-YK-L-D------- 654
A2.CY.94.94CY017_41          ..E----..-.--.L--V------.----------L-----------------S---Q---------K------------V------------------------A-T----T---.--.............-QDE-.-DNM--LQ--K--S---NI-YR-L--------- 657
B.BR.10.10BR_MG029           ..-----..-.--.----------.-------------------------------------------------------L-------M---R--------R-----N--------.--.............TYDE-.-QNM--RD-E---D---G--YN-L-K------- 667
B.CA.07.502_1191_03          ..---R-..-.A-.L----I---S.----------L-----------------------------M--------------V-------R------L------I----S-T--ST--.--.............--DE-.-KNM---Q-EK--D---DI-YT-L-K------- 661
B.CH.08.M2_0803101_NFLG8     ..-----..-.-T.L--M------.----------L-----------------------------M--------------V-------R-----------------------TT--.--.............T----.-GNM---Q-EK--D---NV-YT---DA------ 684
B.CN.10.DEMB10CN002          ..-----..-.--.L--M------.----------I--M-------T------S-----------M--------------V-------Q--R---------------T--------.--.............--ND-.--NM---Q-EK--D---G--Y-------I---- 657
B.ES.10.DEMB10ES002          ..-----..-.A..F---------.----------L-----------------S-----------M--------------V-------R---------------------------.--.............TYN--.-DNM---Q-E---D----I-YD---K------- 663
B.FR.11.DEMB11FR001          ..-----..-.--.----------.----------------------------------------M--------------V-------------------------------T---.--.............T-NE-.-ENM---Q---------N--YT---K------R 639
B.GB.05.MM45d213_GN1         ..-----..-.I-.L--V------.----------V---------------------K----------------------V--I-S---------------------T----T---.--.............--NE-.-DNM---Q--K-------T-YT---KA------ 661
B.HK.06.HK003                ..-----..-.--.----------.----------L-----------------------------M--------------V-------R------------------N--------.-R.............T-TE-.-DNL-----EK--D---G--Y----------- 652
B.HT.05.05HT_129389          ..-----..-.--.L--M------.----------IA--A-----------------------------------------V-------Q------L------R----T----T---.--.............--DY-.-QNM---Q-E---S---G--Y-------I--D- 661
B.JP.12.DEMB12JP001          ..-----..-.--.---V------.----------LM----------------------------M--------------V-------R------------------T----T---K-R.............N-TD-.-DNM--VQ-E---D---KE-YT-L--A------ 663
B.KR.07.HP_18_07JHS10_3909   ..-----..-.--T---M------.------------------L------------------------------------V--------------------------N----T---.--.............-QKE-.-DNM---Q----------F-YT-L--------- 656
B.PE.07.502_0525_wg5         ..-----..-.--TL--V------.----------I---A---------------M-K----------------------V-------Q----------------A----------.--.............--N--.-ENM---Q-EK--D---E--YT-L-K------- 641
B.RU.11.11RU21n              ..-----..-.--.---V------.----------I----------A---------------------------------------------------------------------.--.............--DD-.--N------E---D---EV-Y----QT------ 641
B.TH.08.MERLBDTRC10          ..-----..-.--V---M------.----------L--------------------------------------------V-------------------------------D---.--.............--TE-.--NM-----E-------GI-Y---------R-- 662
B.US.11.ES38                 ..-----..-.--.F---------.----------V--------------------------------------------V--------------L----------------N---K--.............T-GE-.-DNM--IQ-EK--D---DI-Y--L--A-T--DI 661
C.AR.01.ARG4006              ..E----..-I--.L--V------.----------I-----------------S-----------M------------T-V--L----R------------------N----S---.--.............--DN-.-DNM---Q---------DT-YR-L--------Q 655
C.BR.07.DEMC07BR003          ..E----..-.--.---V---I--.----------------------------S-----------M--------------V--I---------------------P------T---.--.............-QQE-.-DNM---Q---------DT-YR-L--------- 678
C.BW.00.00BW5031_1           ..E----..-.A..L--V------.----------I-----------------S-----------M--------------V--I----R-------------I---------S---.--.............NE-D-.-GNM-*-Q-----S---DT-YR-L-D--S---Q 640
C.CN.10.YNFL19               ..E----..-.--.---V------.----------I-----------------S------------------------T-V--I-----------L----------------S---.--.............THDE-.-DNM---Q-----S---NT-YQ-L----S---S 660
C.CY.09.CY260                ..E----..-.--.---V------.----------L-----------------S---K----------------------V--I-----------------------T----I---.--.............-KD--.-DNM---Q--K--S---DT-YR-L--------- 651
C.ES.08.X2363_2              ..---REKR-V--.L--V------.----------I--------------------------------------------V--I-----------M------F----T----S---.--.............-ETE-.-ENM---Q-----S---YT-YN-L-K--I---T 667
C.ET.02.02ET_288             ..G----..-.A..L---------.----------I-----------------S-----------M-----------RT-V--I----Q----------------P-T----S---.--.............-Q-E-.-ENM---Q-----S----T-YR-L-V------- 649
C.IN.09.T125_2139            ..E----..-.--MM--M------.----------I-----------------S--------------------------V--I----------------------------S---.--.............-YDE-.-ENM---Q--K------DT-YR-L-D--I---R 675
C.KE.00.KER2010              ..E----..-.--.---VL-----.----------I-----------------S-----------M--------------V--I-----------------------N----S---.--.............-QTD-.--NL---Q--K------DT-YR-L----S---R 654
C.MW.09.703010256_CH256.w96  ..E----..-.--.---V------.----------IA----------------------------M--------------V--I----E-----------------------N---.--.............-QDD-.-GNM---Q-----S---DT-YR-L-D--I---- 645
C.TZ.08.707010457_CH457.w8   ..E-Q--..-.--.M--M------.----------I-----------------S-----------I------------T----I-----------------------S----S---.--.............-ESE-.--NM---Q--K--S---NT-YR-L--------- 637
C.YE.02.02YE511              ..E----..-.A-.M--------S.----------I-----------------S------------------------T-V--I--------------------------------.--.............-QGD-.-DNM---Q-----S---GI-YK-L--------E 654
C.ZA.08.705010534_CH534.w12  ..E----..-.--.----------.T---------I---------------------K-------M--------------V--L----R------M-------------Y--S---.--.............TQ-E-.-ENM---Q-----S---NT-YK-L--------- 645
C.ZA.10.DEMC10ZA001          ..E----..-.--.---M------.----------I-----------------S-----------M--------------V--L----Q---------------------------.--.............--DD-.-KNM---Q---------YT-YR-L----T---Q 645
C.ZM.11.DEMC11ZM006          ..G----..-.--.---V------.----------I-----------------S--------------------------V--I--------------N-------N----S---.--.............-Y-D-.-ENM---Q-EK------DT-YR-L-V--T---Q 655
D.CM.10.DEMD10CM009          ..E----..-.I-.L--M------.----------I---------------------------------------L--------------------------H----N----S---.-R.............--ND-.-GNM---Q-E---DK--GV-YN---D--I---- 643
D.CY.06.CY163                ..E----..-.I-.L--M------.----------IA--A----V---------------------------------------------------------H----N-A--ST--R--.............--NE-.-DNM---Q-E---E---GV-YT-L----L---R 657
D.KE.11.DEMD11KE003          ..E----..-.I-.L--M------.----------L-------------------------------------------------S---------V------H----N----S---.--.............-QNE-.-GNM--I--E---D---E--Y----S--I---A 649
D.KR.04.04KBH8               ..E----..-.I-.L--M------.----------V--------------------------------------------V--L-----------------------T----S---.-R.............-VDT-.-DNM---Q-E---D---GI-YD-L--------- 641
D.SN.90.SE365                ..E----..-.I-.L-VM------.-----------A---------------------------------------------------R------------RH----N----S---.--.............--TE-.-QNM-----E---D---G--Y-------T---- 649
D.TZ.01.A280                 ..E----..-.I-.L--M------.----------L--------------------------------------------V----S----------------H---------S---.--.............--DD-.--NM-----EK--D---GV-Y-------V---- 651
D.UG.10.DEMD10UG004          ..EKQ--..-.I-.L--M------.-----------A-------V-----------------------------------V----S--R-------------RH---------ST--.--.............T--D-.-GNM-----EK--D---E--YN---V--T---- 646
D.UG.11.DEMD11UG003          ..E----..-.I-.L--M------.--------V-L-----V-----------------------M-------V-----------S----------------H----N----S---.--.............-V-Y-.-ENM---Q-EK--E---GV-Y-------I---- 656
D.YE.02.02YE516              ..E----..-.I-.L--M------.----------V-------------------------------------------------Q----------------H----S----S---.-R.............--DA-.-QNM-----EK--D---G--YA-L----I--D- 645
D.ZA.90.R1                   ..A----..-.I-.LASW------.----------V--------------------------------------------V-------R-------------H----T----DPE-.--.............TYDE-.-GNM-----E---D---GY-YT------V---- 657
F1.AO.06.AO_06_ANG32         ..G----..-.--.-----F----.----------I-----------------------------M--------------L----------R-------------P-N----S---.--.............-QAE-.-ENM-----EK--S--SNT-YE---K------I 639
F1.AR.02.ARE933              ..K----..-.--.M----I---A.---G------I-----------------------------M--------------L-------R------L-----------T----S---.--.............TQKE-.-ENM-----EK--S--SNE-YR---A------ 644
F1.BR.10.10BR_RJ015          ..-----..-.I-GL--V------.----------L-----------------S--------------------------V-----------I--L-----------S----S---.--.............-YTD-.-ENM---Q-EK--S--SEE-YR---Q------ 646
F1.CY.08.CY222               ..K----..-.-E.---V--T---.----------L---------------------K----------------------V--------------------------N----S---.-R.............-Q-D-.-KNM---D-EK--D---DT-YR-L-Q------Q 643
F1.ES.02.ES_X845_4           ..E----..-.--.----------.----------I--------------------------------------------V--------------------------N----S---.--.............--DY-.-ENM-----EK--S--SNT-YR---Q------I 641
F1.RO.96.BCI_R07             ..K----..-.A-.L-L-L-----.----------I---------------------K----------------------V--------------------------N----S---.--.............-QD--.-DNM------K--S--SNI-YN-L-Q--I---- 657
F1.RU.08.D88_845             ..E-Q--..-.L-.M---------.----------I-----------------------------M--------------V-------Q--R-----------F---N----S---.--.............TYDE-.--NM---Q-EK--S-HSET-YR-L----S--DR 661
F2.CM.02.02CM_0016BBY        ..K--R-..-.--.---V------.----------I---------------------K-------------------------------------------------N----S---.--.............-Q-E-.-GNM---Q-EK--D---DT-YR---A------ 644
F2.CM.10.DEMF210CM001        ..S----..-.--.L--M------.----------I-----------------S---K----------------------L--------------------------T----S---.S-.............PHDD-.-QNM---Q-EK--A---GT-YQ---A-D---- 654
F2.CM.10.DEMF210CM007        ..N----..-.--.---V------.T---------I-----V-----------S-----------M--------------L-----F--------L-----------S----S---.--.............-Y-E-.-DNM---Q-EK--S---ET-Y----NA------ 646
G.BE.96.DRCBL                ..E----..-.--.V--I------.T---------I-----V-----------S-----------------------R--V--L-----------------------N----T---.--.............-YNE-.-ENM--I--E---D---YH-Y----Q--I---- 646
G.CM.10.DEMG10CM008          ..G----..-.--.L--V------.----------I-----V--------------------------------------V------------------------P-N----T---.--.............GYDE-.-ENM--------VS---QQ-Y----K------- 646
G.CN.08.GX_2084_08           ..-----..-.--TM---------.T---------L-----V-----------S--------------------------V--------------------------N----T---.--.............-YSD-.-DKM--IQ-E--VS---QH-Y------------ 651
G.CU.99.Cu74                 ..E----..-.--.MR-V------.----------I-----V-----------S--------------------------V--------------L-----------N--------.--.............-VDE-.-GNM--I--E-------AQ-Y------------ 657
G.ES.09.X2634_2              ..G----..-.A..L--I------.----------I-----V-----------S-----------M--------------V--L----R------------------N----T---.--.............-YDD-.-DNM--VQ-E-------QQ-YT-L--------- 650
G.GH.03.03GH175G             ..E----..-.--.L--V------.----------I-----V-----------S-----------M--------------V-------R------------------N----T---.--.............TYNE-.-DNM--IQ-EK------KQ-Y------------ 654
G.KE.09.DEMG09KE001          ..G----..-V--.V--V-----S.----------I-----V-----------S--------------------------V--I-GF------------T-----P-N----T---.--.............TY---.-ENM--TQ-----S---QE-YR---N--T---I 651
G.NG.09.09NG_SC62            ..-----..-A--.M--------S.----------L--------------------------------------------V--M-----------------------N----V---.-R.............-YD--.-DNM--IQ----VR---QQ-YT--K-A------ 648
G.ZA.01.TV546                ..G----..-.--.L--V------.----------I-----V-----------S--------------------------V--------------------------N--------.--.............-XSE-.-DNM--I--E-------QX-Y----X--X---- 647
H.BE.93.VI991                ..E----..-.--.M--F------.----------I-----------------S------Q----M--------------V--------------------------N----S---.--.............--DE-.-DNM------KQ-----DE-YR-L-V------- 658
H.CF.90.056                  ..E----..-.--.M--S------.----------I-----------------S------Q-R--M--------------V-------R------------------N----S---.--.............-QSE-.-DNM------KQ-S---EE-YR-L-V--T---- 644
H.GB.00.00GBAC4001           ..E----..-.--.M--F------.----------I-----------------S------Q----M-------V------V--------------------------N----S---.--.............-YXX-.-XNM-----E-Q-D---XE-YX-L-I------- 651
J.CM.04.04CMU11421           ..E----..-.L-.---V------.T---------I-----V-----------S---K---------K------------V--------------------------T----TT--.-R.............-HDD-.-GNM--IQ-E-------GI-YT----A------ 659
J.SE.93.SE9280_7887          ..E----..-.--.---V------.T---------I-----V-----------S---K---------K------------V--------------------------N--------.--.............-Y-D-.-ENM--IQ-E-------GI-Y-----A----N 647
J.SE.94.SE9173_7022          ..E----..-.--.---V------.T---------I-----V-----------S---K--X------K------------V--------------------------N--------.--.............-Y-D-.-ENM--IQ-E-------GI-Y-----A----T 649
K.CD.97.97ZR_EQTB11          ..-----..-.--.----------.----------I----------------------------QM-----------R--V-------R------------------N----S---.--.............-QSE-.-ENM---Q-EK--S-H--T-YR------I---- 650
K.CM.96.96CM_MP535           ..-----..-.--.L--V-F----.----------I-----------------S-----------------------R-----------------------------N----S---.--.............-W-E-.--NM-----EK--G--SDT-YK------T---- 641
U.CA.01.TV749                ..-----..-.--.----------.----------I-----------------S-------------K------------V--------------------------N----T---.--.............--DE-.--NM---D----V----AI-YG---Q------- 660
U.CA.99.TV721                ..-----..-.--.----------.T---------I-----------------S-------------K------------V--------------------------N----T---.--.............-YNE-.--NM-------------GI-YN---Q------- 663
U.CD.83.83CD003_Z3           ..E----..-.--.M---------.-----------A--A----------------------------------------V-------ES-----L-----------T----S---.--.............--DN-.-DNL---------S---QV-YG-L-D--K---- 650
U.CD.90.90CD121E12           ..E----..-.--.M---------.----------IA--A----------------------------------------V-------EG-----L-----------T----S---.-R.............--ND-.-DNM------K--S---QI-YG---D--K---- 648
U.CY.05.CY090                ..E----..-.--.L---------.----------L-----------------------------M-----------R--V--L----Q------------------N----S---.--.............TYND-.-ENL---Q--K--S---GI-YG------I---- 644
U.CY.08.CY223                ..E----..-.--.M--I------.----------I----------A------S---K------Q--R------------V--L-----------------------N----S---.--.............-Y-E-.-DNM--LQ-E---D---VQ-YT----A-Y---- 679
U.ES.10.DEURF10DZ001         ..G----..-.I-.M--F------.----------I---A-------------S---K------N--R------------V--L-A--R------------------N----S---.--.............-Y-D-.-ENM--QQ--K------GI-YE---K------- 645
U.GR.99.99GR303              ..E----..-.--.MA-II-----A----------V--------------R--S---Q---------R------------V-----------I-E--A---------T----S---.--.............-Q-LTYGKNM--L---K--S---QQ-YE------I---- 655
U.NL.95.U_NL_95_H10986_D1    ..-----..-.--.L----F----.----------I-----------------------------------------------M-----------------------N----S---.--.............-Q-E-.-DNM--I--E---SK--GT-YR------------ 660
01_AE.AF.07.569M             ..E-Q--..-.--.---MIF----.----------I-----------------S-----------M--------------V----------KF--L------I---------ST--.--.............-Y-D-.--NL--T--E---S---NQ-YA-LT---D--DR 650
01_AE.CF.90.90CF11697        ..E----..-.--.---MIF----.----------L-----------------S--------------------------V----------KF--L------I----N----T---.--.............-YAE-.-DNM--I---K------NQ-YE-LTK------- 655
01_AE.CN.10.YNFL03           ..E----..-.--.---MIF----.----------I-----------------S-----------M----------------------Q--KF--L------I---------ST--.--.............TY---.--NM--I--E---G---NQ-YDILIK--D---R 651
01_AE.HK.04.HK001            ..E----..-V--.---MIF----.----------L-----------------S-----------M--------------V----------RF--L------I---------NT--.--.............-Y-E-.--NM--TQ-E---S---DQ-YAILT------D- 651
01_AE.IR.10.10IR.THR48F      ..E-Q--..-.--.---MIF----.----------I-----------------S--------------------------V----------KF--L------I--P------ST--.--.............T----.--NM--TQ-E---S---EQ-YEILT-------- 659
01_AE.JP.x.JRC77AE           ..E-Q--..-.--.---MIF----.----------I-----------------S--------------------------V----------KF--L------I---------ST--.-R.............-Y-E-.-DNM--IQ-E---S---NQ-YG-LT-A-D--D- 648
01_AE.TH.09.AA090a_WG11      ..E----..-.--.---MIF----.----------I-----------------S-----------M--------------V-----------F--L------I----N----ST--.--.............-Y-K-.--NM--I--E---S---NQ-YDILT------D- 656
01_AE.TH.90.CM240            ..E----..-.--.---MIF----.----------I-----------------S--------------------------V----------KF--L------I---------ST--.-R.............-F-E-.--NM--I--E---S---NQ-YEILT------DR 649
01_AE.US.05.306163_FL        ..E-Q--..-.--.--TMIF----.----------I-----------------S----------RM--------------L----------RF--L------I---------ST--.--.............-Y-E-.-DNM--R--E---S---DQ-YG-LT-------R 628
01_AE.VN.98.98VNND15         ..E----..-V--.---MIF----.----------I-----------------S--------------------------V----------TF--L------I----N----S---.--.............-F-E-.-DNM--I--E---S---NQ-YAILT------D- 659
```

```
                                           fusion peptide                                                          glycosylation NAS          glycosylation NHT
                              gp120 end  gp41 start                                              immunodominant region     glycosylation NKS          glycosylation NYT
B.FR.83.HXB2                  ..QREKR..A.VG.IGALFLGFLG.AAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWS.NK.............SLEQI.WNHTTWMEWDREINNYTSLIHSLIEESQNQQEK 655
02_AG.CM.10.DE00210CM013      ..G----.-.--.L--V-F----.----------I-----------------S--K--------R-----------V--L-----------L----R-----T---S---.--.............-YKY-.-DNM--LQ-E---D---GI-YD---------- 676
02_AG.CY.09.CY256             ..E----.-.--.---VL----.----------I-----T-----------S--K-----Q--R-----------V--L---R-----M-----------S---S---.--.............NYTD-.-DNM--LQ--K--S---DT-YR---------- 649
02_AG.ES.06.P1423             ..E----.-.--.L--V------.X---------V-----------------S------------K-----------V--L---R-----------------N---S---.-R.............TY-D-.-GNM--LQ-E---D---NI-YG---------- 654
02_AG.GW.05.CC_0048           ..-----.-.--.L--------.----------X-----------------S----------M-K-----------V--L---R-----------------------S---.-R.............TYDD-.-DNM--LQ--K--S---Q--YG---D--I---- 659
02_AG.KR.12.12MHR9            ..G----.-.--.L--VL----.----------L-----------------S-------------K-----------V--L-G-R-----------------T---S---.--.............TY-N-.-GNM--LQ--K--S---NI-YD---------- 661
02_AG.LR.x.POC44951           ..EK---.-.--.L--V-I---.----------I-----------------S--K--------R-----------V--L---R-----------------N---T---.--.............TYHD-.-DNM--L---K--S---NK-YE---------- 644
02_AG.NG.09.09NG_SC61         ..G----.-.--.L--V------.----------V-----------------S-------Q--K-----------V--L---R-----------R-----N---Y---.--.............-YND-.-DNM--LQ--K--S---AI-YN---------- 653
02_AG.NG.x.IBNG               ..E----.-.--.L--V------.--------R-I-----------------S--K--------K-----------V--L---R-----------------T---S---.--.............TFND-.-DNM--IQ-EK--S---DI-YN-------R--- 646
02_AG.SE.94.SE7812            ..E----.-L--.L--F-F----.----------I-----------------S--K--------R-----------V--L-A-------------------T---S---.--.............TYDH-.-GNM--LQ--K--S---HI-YD---------- 658
02_AG.SN.98.98SE_MP1211       ..E----.-.--.L--V------.----------I----------------------K--------K-----------V--L-------------------T---S---N--.............TFKD-.-DNM--LQ--K--S---EK-YT----A------ 653
03_AB.RU.97.KAL153_2          ..-----.-.--.---V------.----------I-----------------------------------------V-----------------------------T---.--.............P-DE-.*-NM-----E-------G--YN---------- 638
04_cpx.CY.94.94CY032_3        ..----*.-.--.---M------.------------------------------S------------R-----------V--L-S---------------N---S---.--.............-YND-.-DNM--LQ--K------QI-YG-L--------- 653
05_DF.BE.x.VI1310             ..K----.-.M-.---M----A.----------IA----------------Q------------------------V-------------L--S--H--S-S---S---.--.............-EGE-.-DNM-----EK--S--S-T-YR---Q--I---- 660
06_cpx.AU.96.BFP90            ..G----.-.--.L--M-----.T---------I-----V-----------S--------------------------------------P-N-L--T---.--.............TYDE-.-GNM--I----------QQ-Y----L--T---- 658
07_BC.CN.98.98CN009           ..E----.-.--.---V------.----------I-----------------S-----------------T-V--I-----------------------------S---.--.............-QGE-.-DNM---Q--K--S---NT-YR-L--------R 651
08_BC.CN.97.97CNGX_6F         ..E----.-.--.L--V------.----------I-----------------S---------M-----------T-V--I-----------------------------S---.--.............-QQE-.-DNM---Q--K--S---NT-YR-L-D------R 656
09_cpx.GH.96.96GH2911         ..E----.-.--.M---------.----------I-----------------S-------------K-----------V-----------------------N---S---.--.............-QHE-.-ENM--L------S---Q--YR---------- 641
10_CD.TZ.96.96TZ_BF061        ..E----.-.I-.L--V------.----------L-------V-----------------------------V---S----------------RH---N---S---.--.............--GE-.-DNM-----E---D---G--YN------I---- 639
11_cpx.CM.95.95CM_1816        ..E----.-.--.---VL----.----------I-----------------S--K--------K-----------V--I-----------------------N---V---.--.............-YNE-.-DNM--I--E-------QT-YN-L--------- 664
12_BF.AR.99.ARMA159           ..T----.-.--.----------.----------V---------------------Q------------------V-------------L-------A-N---S---.--.............-Q-E-.--NM---Q-EK--S--SHA-YR---D------- 641
13_cpx.CM.96.96CM_1849        ..A----.-V--.L--V------.-------V--L---------------------M--------K--------------I-----------------------N---S---.--.............-QSE-.-DNM--LQ--K-VS---EM-YR-L--------- 659
14_BG.ES.05.X1870             ..-----.-.--.----LF----.---------AV----------------------------M-----------V--M---Q--R-----------------N---------.--.............---D-.--NM-----E---D---G--YT---Q------R 657
15_01B.TH.99.99TH_MU2079      ..-----.-.--T---M------.T---------I-----------------R-----------------------V-----------------------------------.--.............--NE-.-DNM-----E---S---KE-YT---Q------- 646
16_A2D.KR.97.97KR004          ..E----.-.A..V-LF-----.----------V-----------------S--K--------M-R-----------V-------Q-------------------F---S---.-R.............TQDD-.--NM--LQ-E---D---GT-YR-L--------- 654
17_BF.AR.99.ARMA038           ..-----.-.--.MA-V-I----.----------I-----------------S--------------M-----------V-----------M-----------------T---.--.............-Q-E-.--NM-----EK--S--SEE-YR--G-------- 647
18_cpx.CU.99.CU76             ..G----.-.--.L--V------.----------I-----------------S--------------M-K-----------V-------------------P---S---.--.............TYND-.--NM--LQ--K--S---Q--YT---------- 655
19_cpx.CU.99.CU7              ..E----.-.M-.L--V------.----------I-----------------S--K-----Q--R-----------V--L-----------------------T---S---.-T.............-YXD-.--NM--L---K-----PQI-YN---------- 601
20_BG.CU.99.Cu103             ..-----.-.--.L--V------.----------I-----V-----------S-------------------------------V-----------L---------N---T---.--.............-YNE-.-DNM--I--E-------QQ-YT---------- 641
21_A2D.KE.99.KER2003          ..D----.-.I-.L--M------.----------V---------------------M-------------------V-----------M-----H---D---S---.--.............--DD-.-QNM---Q-E---D---GK-Y--L----I--D- 651
22_01A1.CM.01.01CM_0001BBY    ..E----.-.L-.F--V------.----------I-----------------S----------M-K-----------V-----------------------T---S---.--.............-Q-E-.--NM--LQ--K--S---DI-Y----------- 640
23_BG.CU.03.CB118             ..E-G--.-.--.L--V------.----------I-----V-----------S-----------------------V--L-----------L---------N---T---.--.............T--E-.-DNM--IQ-E-------HQ-YN------Q---- 647
24_BG.ES.08.X2456_2           ..G----.-.--.L--V------.T---------I-----V-----------S-----------------------V-----------L---------N---S---.--.............-FNA-.-DNM--I--E-------QH-Y----------- 652
25_cpx.CM.02.1918LE           ..E----.-.I-.L--A------.T---------I-----V-----------G-----------------------V-----------------------N---T---.--.............-FNE-.-DNM--I--E---D---QQ-YN---L------- 649
26_AU.CD.02.02CD_MBTB047      ..E----.-.--.M---------.----------I-----------------S-----------M-K-----------V---------------I---N---T---.--.............-WNE-.-DNM--LQ--K------NT-YN-L-VA------ 650
27_cpx.FR.04.04CD_FR_KZS      ..E----.-.--.----IF----.----------I---H-------------S-----------M-----------V--------F-------I-----------S---.--.............-Y-A-.--KM--I------S---KQ-YE-LT---D---- 655
28_BF.BR.99.BREPM12609        ..-----.-.--.AL--------.--------------------M-------------------------------------------------------N---S---.--.............-----.-DNM---Q-EK--D---DM-YT---KA------ 667
29_BF.BR.01.BREPM16704        ..-----.-.--.---M------.----------VA----------------------------------------V---S--R-------------------------------.--.............-QSE-.-DNM-----E---G-H-RF-Y--L----H---- 669
31_BC.BR.04.04BR142           ..E----.-.--.L---F-----.----------I-----------------S----------M-----------T-V--I----E-----------------N---S---.--.............-QAD-.-DNM---Q--K--S---GI-YRML-D------- 641
32_06A1.EE.01.EE0369          ..G----.-.--.L--V------.T---------I-----V-----------S-----------------------V--L---V---------T-----P-N---D---.--.............TF-E-.--NM--I--E--------H-YT---------- 649
33_01B.ID.07.JKT189_C         ..E----.-.--.---MIF----.----------I-----------------S----------M-----------V----------KF--L-----T---------ST--.-R.............-Y-E-.--NM--T--E---S---NQ-YEILT------D- 651
34_01B.TH.99.OUR2478P         ..E----.-.--.---MIF----.----------I-----------------S--------------------------------F--L-----I---------ST--.-R.............-Y-E-.--NM--IK-E---S-H-GQ-YQILT------D- 642
35_AD.AF.07.169H              ..E----.-.--.L---I-----.----------V-----------------S--K--------M-R-----------V--L---R-----------------T---S---.--.............TQNE-.--NM--LQ--K--SS--NT-YT-L--------- 646
36_cpx.CM.00.00CMNYU830       ..E----.-.--.L--V------.----------I-----------------S-------------K-----------V--L---R-----------------N---S---.--.............TYDS-.-ENM--LQ-----S---QI-YE---------R 648
37_cpx.CM.00.00CMNYU926       ..E----.-.--.L--V------.----------I-----------------S-------------K-----------V--L-------------------H-T---S---.--.............--ND-.-ENM--LQ--K--E---QI-YN---------- 644
38_BF1.UY.03.UY03_3389        ..-----.-.--.MA-V-I----.----------I----------------------------M-----------V-----------------------T---T---.--.............-HDE-.-DNM--RQ-EK--S--SEG-YR----A------ 643
39_BF.BR.04.04BRRJ179         ..K----.-.A..L-VV----A.----------I---------------------K--D----------------L----F-------L-----------N---S---.--.............-Q-E-.-GNM-----EK--S--SDQ-YR---Q------- 643
40_BF.BR.05.05BRRJ055         ..-----.-.-T.L--M------.---------AT-----------------S-----------M-----------V-----------L-----------N---I---.--.............--DD-.-KNM---Q-EK--D---GI-YG---------- 650
42_BF.LU.06.luBF_18_06        ..-----.-.--...................----V-------------------------M-----------V-------R-----L-----------N---S---.--.............TH-E-.-GKM-----EK-----SQE-YR---D------- 639
43_02G.SA.03.J11223           ..-----.-.--.M--V------.----------I-----V-----------S-----------------------V-----------------------N---T---.--.............-YDE-.-GNM--IQ-E--VG---QH-YD---------- 638
44_BF.CL.00.CH80              ..T--R-.-.-T.---M------.----------I-----------------R--M--------M-----------V-----------------------N---S---.--.............-Q-E-.-GNM--I--E---S--SEE-YR---------- 643
45_cpx.FR.04.04FR_AUK         ..E----.-V--.L--V-I----.---------AV-----------------S-------------K-----------V-----------------------T---S---.--.............--DK-.-GNM--L------S---QI-Y----------- 644
46_BF.BR.07.07BR_FPS625       ..T----.-.--.L---------.----------I---------A-----------------------------------------L-----------N---S---.--.............TQD--.-QNM-----EK-VSS-SNE-YR--T---S---- 634
47_BF.ES.08.P1942             ..--Q--.-.--.F---------.----------VA-----L-----------------M-----------V---------IM-------Y---------T---.--.............-YD--.-KNM---Q-EK--D----E-Y---AL--D---- 647
48_01B.MY.07.07MYKT021        ..E----.-.--.----IF----.----------I-----------------S----------M-----------V---------TF--L-----I---N---ST-G.--.............---E-.-DNM--TQ-E---S---NQ-Y--LT------DR 659
49_cpx.GM.03.N26677           ..E----.-.--.M--V------.----------L-----------------S--Q--------M-K-----------V------Q-------------F---S---S---.--.............-Y-E-.-ENM--R--EK--D---G--YT--G-A---D- 654
50_A1D.GB.10.12792            ..A----.-.I-.L--M------.----------L-------------------------V--------V---S--------------RH---N---T---.--.............--QE-.-DNM-----E---D---GV-Y----K--T---- 655
51_01B.SG.11.11SG_HM021       ..E----.-.A-.----IF----.------------------------------S----------M-----------V-------------------------N-------.--.............T-DY-.-ENM---Q-EK--D---G--Y--L-K------- 650
52_01B.MY.03.03MYKL018_1      ..E----.-.--.---MIF----.----------L-----------------S----------M-----------V----------KF--L-----I---------ST--.--.............-HDE-.-DKM--IQ-E---S---DK-YE-LT------DR 659
53_01B.MY.11.11FIR164         ..E-N--.-.--.---MIF----.----------I-----------------S----------M-----------V----------KF-SL-----I---------ST--.--.............-Y-E-.-GNM--I--E---S---NH-YEILT------DR 647
54_01B.MY.09.09MYSB023        ..E----.-.--.---MIF----.----------I-----------------S-----------------------V----------KF--L-----I---------ST--.--.............-Y-E-.-DKM--I--E---S---DQ-YDILTQ-----D- 656
55_01B.CN.10.HNCS102056       ..E----.-.--.---MIF----.----------I-----------------S-----------------------L----------KF--L-----I---N---ST--.--.............-F-E-.-QNM--I--E---S---NT-YD-LM------D- 658
56_cpx.FR.10.URF5_patient_A   ..G----.-.--.M--V------.----------I---------------------K--------M-R-----------V--L-----------------P-------S---.--.............NYTD-.-DNM--LQ--K--S---DI-Y----------- 644
57_BC.CN.09.09YNLX19sg        ..E----.-.--.L--V------.----------VA--------------------------M-----------V--I---E-------------------------S--N.-R.............-QDE-.-KNM---Q---------NT-YQ-L----H---R 662
58_01B.MY.09.09MYPR37         ..-----.-.-RA---M------.----------I-----------------R-----------------------V-----------------------T---------.--.............--DA-.-DNM--I--E---S---KE-YT---------- 656
59_01B.CN.09.09LNA423         ..-----.-.--.L---------.----------V--------------------------M-----------V--L---R-----------I-----------I---.--.............---E-.-DNM--TQ-EK--G---YK-Y--L----I---R 651
60_BC.IT.11.BAV499            ..E----.-.A..L--M------.----------I-----------------S----------M-----------T-V--L---R-----------------------S---.--.............---A-.-ENM---Q-----X---NT-YN-L-K------- 651
61_BC.CN.10.JL100010          ..E----.-.--.---V------.----------I-----------------S-----------------------T-V--I---R-----------------------T---.--.............-Q-E-.-QNM---Q-----S--SNT-YR-L-D------R 640
62_BC.CN.10.YNFL13            ..E----.-.--.---V------.------------------------------S--K--------M-----------T-V--I-------------------P-------I---.--.............-R-E-.-ENM---Q-----S---NT-YR-L--------- 650
63_02A1.RU.10.10RU6637        ..E-G--.-V--.L--V------.T-------AL--------------------------M-K-----------V-----------------------N---S---.--.............-KNE-.-DNM--LQ-----S---DI-YG-L--------- 647
64_BC.CN.09.YNFL31            ..E----.-.--.---V------.----------I-----------------------------------------V--I-----------------------------S---.--.............N-TD-.-DNM---Q-----S---ET-YR-L-D------R 651
65_cpx.CN.10.YNFL01           ..-----.-.A..L--M------.----------I-----------------R--M--D------------------V-------------L-----------T---V---.--.............TQK--.-QNM--I--E---D---GE-YR---------- 661
67_01B.CN.11.2011.ANHUI.HF115.D----.-.--.---MIF----.----------I-----------------S----------M-----------V----------KF--L-----V---N---ST--.--.............-F-E-.--NM--T--E---S---NK-YE-LT------D- 639
68_01B.CN.11.2011.ANHUI.WH73..D----.-.--.---MIF----.----------I-----------------S-----------------------V----------KF--L-----I---N---ST--.--.............-Y-E-.-DNM--T--E---S---NK-YE-LT------DQ 640
70_BF1.BR.10.10BR_PE004       ..-----.-.--.L-VM------.----------I-----------------------------------------------------L-----------N---------.--.............--DN-.-DNM---Q-EK--G---D--YD---K------- 649
71_BF1.BR.10.10BR_PE008       ..-----.-V--.----------.----------------L-----------------M-----------V-------R-----------------------T---.--.............-QG--.-ENM-----EK--D---D--YT-L--------- 655
72_BF1.BR.10.10BR_MG002       ..T----.-.L-.M---------.----------V----------------------M-----------V--L-----------L-----------N---S----L.............-HD--.-GNM--KK-EE--S--S-E-YM------F---- 646
O.BE.87.ANT70                 ..H----.-.--.L-M----V-S.---------AT--A--THT--K------D------Q---Q--R-SX---R--R--L--L-TL-QN----SL---K---V-Y-S-K--RT-I.GN..............-S-.-DTL--Q----Q-S-IS-T-YEE-QKA-V---Q 651
O.CM.98.98CMA104              GTH----.-.--.L-M----V-S.---------ATA---THT-IK------D------Q--HQ--K-S----R--R--L--L-TFIQN----NL---K-N---Y-S-K--NT-G.GDR...........D-GN-.-DNM--Q---QQ-S-IS-T-YEE-QKA-V---H 672
O.CM.98.98CMABB141            ..HK---.-.--.L-M----V-S.---------ATA---THT-IK------D------Q------R-S----R--R--L--L-TLIQN----NS---K---V-Y-S-K--KT-T.GN.............DS-.-GNL--QQ--QQ-D-VS-I-FDE-LKA-V---E 675
O.CM.98.98CMABB212            .NH-R--.-.--.L-M--F-V-S.---------AT--A-RTHSV-R------D------Q------R-S----R--R--LQ-L-TLIQI--R-DL---K-RI--Y-S-E--KT-G.DYN............DS-.--QL--QQ--QQ---VS-I-YDE-QAA-D---- 678
O.CM.98.98CMU5337             ..H----.-.--.L-M----V-S.---------ATA---THS-MK------D------Q------R-S----R--R--L--L-TFIQN----NL---K-R---Y-S-A--NT-T.KYN.........ETNIDK-.-GNL--Q---KQVA-ISDI-YEE-QKA-V---Q 658
O.CM.99.99CMU4122             ..H----.-.--.L-M----V-S.--------TAT--A--THT-MK------D------Q---Q--R-S----R--R--LQ-L-TLIQN----NL---K-R---Y-S-K--NT-A.-GT...........N-NT-.-DNL--QQ--QQ-D-ISDV-YGE-QKA-D---Q 663
O.FR.92.VAU                   ..H----.-.A-.LAM----I-S.---------ATA---RTQH-IK------D------Q------R-S----R--R--L--L-TFIQN---HNL---KNR---Y-S-K--KT-G.GDN.............-S-.-DEL--QQ--QQ---VS-F-YEK-Q-A-E---- 665
O.GA.11.11Gab6352             ..H----.-.--.L-M----V-S.---------ATA---THT-IK------D------Q------R-S----R--R--L--L-TLIQN----NL---K-R---Y-S-Q--KT-V.--T.........NTD--N-.-GNL--Q----Q-D-ISDI-YEE-QRA-V---Q 673
O.SN.99.99SE_MP1299           ..-----.-.--.L-M----V-S.---------ATA-A--TQS-MK------D------Q---Q--R-S----R--R--L--L-TLIQN----NL---K-----Y-S-K--TT-T.-CT.......N.TNK-DD-.-DKL--QQ--QQ-S-VS-I-YEE-RNA-V---Q 670
O.US.10.LTNP                  ..H----.-.A-.L-M----V-S.---------ATA-----HT-MK------D------Q------R-S----R--R--LQ-L-SLIQN----NL---K--QV-Y-S-K--D--T.GN..............-S-.--NL--QQ--QQ-S-VSAI-FGE-HKA-I---Q 654

N.CM.02.DJO0131               ..S----.-AF-.L---------.----------I------T-------------V-------------SI------R-KV--I----R---I-SL------T--Y-T----ET--.-N...........T-YDT-.-GNL--QQ---KVR--SGV-FE---KA-E--NT 627
N.CM.02.SJGddd                ..S----.-AF-.L-------R.----------I------T----L----------L---H-----S---------KV--I--F-R---I--L------X--Y-T----ET--.SN...........T-YDT-.--NL--Q---KKVR--SGV-FN---QA-E--NT 628
N.CM.04.04CM_1015_04          ..N----.-AF-.L---------.----------I------N-------------------Q----SI---------KV--I----R---I-SL------T--Y-T----ET--.-H...........T-YDS-.-GNL--QQ--EKVR--SGV-FD---QA-E--NT 625
N.CM.06.U14296                ..S----.-AF-.L---------.----------I------TV----------------------S-----------V--I---------I--L------P--Y-T----ET--.SN...........T-YDT-.-DNL--QQ--QKVR--SGV-FG---QA-E--NX 650
N.CM.06.U14842                ..S----.-AF-.L---------.----------I------T----X--------------------SI------R-KV--I---------I-SL------T--Y-T----ET--.SN...........T-YDT-.--NL--QQ--XKVR--SGV-FD---QA-E--NT 647
N.FR.11.N1_FR_2011            ..S----.-AFX.L---------.----------I------K------------------------S---------KV--I----R---I--L------T--Y-T----ET--.-N...........T-YDT-.--NL--Q---KKVRD-SGV-FQ---QA-E--NT 636

P.CM.06.U14788                DGH-K--.-.A-.L-M------S.---------GI------K---H--------M--------E--R-S----R--R--L--I-T--R------L-----QIX-Y-N----ST-T.--N.........ETE-DG-.-GNL--Q---KLVD---DT-YLE-QKA-E--KE 679
P.FR.09.RBF168                DAH-K--.G.A-.L-M-----MS.---------AV-----------H--------M--------E--R-S----R--R--L--I-T--R------L-----QTV-Y-N----R--T.--S.........ETE-DG-.-TNL--Q---KXVD---DT-YLE-QRA-D--KA 671

CPZ.CD.06.BF1167              AKS-S--.E.I-.M-MI-M---S.----------IA------G----------A---Q------------S----------L-----K--R------L---AN----HSSLE--S--L.--TLCGNKTNNYENNY-C-.-SQY--QQ--E--R-ISDV-YLSL-TA-V---R 669
CPZ.CM.05.SIVcpzMT145         ..H----.-.L-.L---------.----------VV---------T------------------------S-----------V--------R------L---T--T--P---R--KT-G.-I...........SDYQV-.--NY--QQ----V----G--YT-L--ANT---- 647
CPZ.GA.88.GAB1                ..D-Q--.-AF-.L---------.---------AV----------------------K------------SI--V-----L-------Q---I--L------AV-Y-T----N--PGSN.............-TDD-.-GNL--QQ--KLVS---GK-FG-L--A-S---- 646
CPZ.TZ.06.SIVcpzTAN13         ..T-S--.D.--.--L------S.----------IA-VA---GV--------H---Q-----------------V------L--M-K-IR-----SV---AN--V-HST-V--DT-A.ARKPC..SNTSGPTDYDC-.-QNL--QQ-SELVD-S-DT-YT-L-VA-I---- 655
CPZ.US.85.US_Marilyn          ..K----.-AF-.L---------.----------VV-------------------------------S-----------V-----------I--L------T--Y-T----DT--.-N...........L-YDA-.-GNL--Q----KVR--SGT-F----QA-E--NT 628

GOR.CM.07.SIVgor2139_287      ..N----.-.A-.L-M-----MS.---------AV--------V-H--------M----A---E--R-S----R--R--L--I-T--R------L---T-R---Y-N----KT-T.G-N.........DTE-DD-.-GNM--QQ--KLVD---DT-FLE-QRA-E---A 661
GOR.CM.07.SIVgorCP2135con     ..N----.-.A-.L-M-----MS.---------AV--------V-H------X-M----A---E--R-S----R--R--L--I-T--R------L---T-R---Y-N----KT-T.G-N.........ETE-DD-.-GNM--QQ--KLVD---DT-FLE-QKA-E---V 656
```

```
                                                                            transmembrane domain  gp41 cytoplasmic tail start              glycosylation NGS        *
B.FR.83.HXB2                 NEQELLELDKWASLWNWFNITNWLWYIKLFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTHLPTP..RG.PDRPEGIEEEGGERDRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELL.GRR..............GWEALKYWWNLLQY.WSQE 806
A1.AU.03.PS1044_Day0         ------A-----N-----D-SQ------I--------I-----------IS--------------T-N-..E-.L---GRT------QG--------S-F---A-------------------S----T----.-HSSLK.......GLRL---G---L----V-.---- 820
A1.CH.03.HIV_CH_BID_V3538    ---D--A------------S-------I--------I------T----I---------------T-N-..E-.L---GR------QG-G-------S-F---A-------L---------FA--AA-T----.-HNSLK.......GLRL---G---LG---L-.-GR- 804
A1.CY.08.CY236               ---D--A-----N-----D-SH------I--------I-----------IR-------------T-N-..G-.L---GR--G---Q--S------S-F---A-----N----------FI--AA-T----.-HSSLK.......GLRL---G---LE---V-.--R- 812
A1.ES.05.X1608_8             --KD--A-----N-----D-SK------I--------I------T-F--I--------------LI-SA..GD.----R----G---Q--G------S-F---A-----------------FI--A--T----.-HNSLK.......GLRL---G---L----V-.-GR- 823
A1.KE.11.DEMA11KE001         ---D--AF----N--S--K---------I--------I----------VI-K-------------N-..G-.L---GR-------QG-G------S-F---A-----------------FT--AA-T----.-HSSLK.......GLRL---G---L----L-.-GR- 826
A1.RU.11.11RU6950            ---D--A----T------D-S-------I--------I----I-T----I--A-------L--LT-H-..E-.----GR-S-----QG--------S-F---A-----------------FI--AA-T----.--SSLK.......GLRLV--G---LG---G-.-G-- 813
A1.RW.11.DEMA111RW002        ------A--------T--D-S-------I--------I-----------I--------------I-T-NQ..GE.----R----G---QG-------AS-F-----E---N--I-------------AA-T----.--SSLK.......GLRL---G---LG--T--.-I-- 821
A1.SN.01.DDI579              ---D--A--------S--D-SK------I--------I-----------I--------------LT-N-..--.------R------Q---------S-F-----------------------FT--AA-T----.-.................QG---L--V-L-.-GR- 800
A1.UG.11.DEMA110UG001        ---D--A-----N-----D-S-------I--------I----IC---A-I--------------I-T-N-..--.----H-R------Q--N--T-SAS-F---A-------F-----L---FI--AA-T----.-HSSLK.......GLRL---G---L----L-.-GR- 806
A1.ZA.04.04ZASK162B1         ------A-----------D-SR------I--------I----I-----ML--------------I-N-..E-.L---GR------QG--------S-F---A-----N-----------FI---A-T---.................GLRL--KGIQ-L----L-.---- 809
A2.CM.01.01CM_1445MV         ---DV-A-----N-----S---------I--------I--K--IVIF-VI------------L-IPT-R-..E-.L----R-K-G---Q---T-V---S-F---A-----------------CIS-AA-T-T--.-HSSLK.......GLRL---GI--L----L-.-GR- 812
A2.CY.94.94CY017_41          ---D--A-----D--S----SH-----RI--------I-------IITV----------V---IPT-S-..E-.----R-T--G---QG-----------FF--A-----------------CI--AA-T----.-HCSLK.......GLRL---G--NL----L-.-GR- 815

B.BR.10.10BR_MG029           -----------------D--H------I---X----I------T-------------------RF-A-..--.T---G--------G----RI--D-FF-IF-V--------I-------------------------.---..................---------I---.---- 818
B.CA.07.502_I191_03          ---------------S--D---------I--------I------T-I---S------------R--AQ..--.-------T--G-------T-GW-AT-F-E---I---N-WI-L-----------LA------.---...............-R-V----------.-G-- 812
B.CH.08.M2_0803101_NFLG8     ------------N--S--D-S-------I--------IV---I--------R------------V-..--.-------T----------T-----D-F-------------------------------T-G--.---...................L---------.---- 835
B.CN.10.DEMB10CN002          --L---K--Q--------D--R------I--------I--K--------------------R----..--.--------------------GP---RFFI------WN-------Q---------------I-.---..................G--I---.-T-- 808
B.ES.10.DEMB10ES002          ---D--------------E--S------I-------I-----I-T-------------------A-..--.L-----------D---N--RP--H-F--I--V-----------------------A-T---.---....................---------.---- 814
B.FR.11.DEMB11FR001          -----------S------D--S------I-------------CT--------------------AQ..--.---------G----KS--T-GC--D-F-TI--V---N--R-----------S-AA------.---...........................---- 790
B.GB.05.MM45d213_GN1         ------------------D---------R-------------------L-----------L-------..--.---------G----------GG----F-T---V--QN--------L-------T--------.---...............---I----------.-I-- 812
B.HK.06.HK003                ---D--A-----------D---------I-----A---------T-I-----------------V-..--.------G-------N--S---SP--G-L--I--V-----------------------I-SA----.-H-...............I---C---------.---- 803
B.HT.05.05HT_129389          ------Q-----------D-S-------I-------------------T----------------RF-A-..--.-E----T-------------T-----F-P--------------H---F---AA-V--I-.---....................-----------.---- 812
B.JP.12.DEMB12JP001          ------------------D--S------I----I----------------------------------RF-A-..--.---------G----------S---S-F-P-----------------------------I-.---...............I-----------.---- 814
B.KR.07.HP_18_07JHS10_3909   -----------S--------SQ-----RI--------I----I-S--------------------F-S-..--.---------G-----N--G-GP----F-T---V--------I-RH---------A-S----.-L-...............V----------.---- 807
B.PE.07.502_0525_wg5         ------------------D--Q------V----I---I-----CV-IN--------------------..--.-G----------D-N--T-A---T-F-P-------------------------A---L---.---...............I-----------.---- 792
B.RU.11.11RU21n              ------A--------S--D---------I--------I----I-----V--------------R---Q..--.------------------T-V-----LF--------------L-----------AA------.-L-.......................-I-- 792
B.TH.08.MERLBDTRC10          --L---------------D--K------I--------------------------------L--RF-A-..--.-GG------------------GT----F-T------QN-------R----------V------.---.......................-G-- 813
B.US.11.ES38                 -------------------S--------I--I----II------T------------------R--AQ..--.------T-----------E--VP----F--IL-V---------------------LA-T--F-.-H-...............I-----------.---- 812

C.AR.01.ARG4006              --KD--A----QN-------S-------I----I---I----I------------------LI-N-..--.-GGLG--------Q-K-------S-F-S-A-------------RQ----I--AA-AA-R-.--S...............I---LGS-V--.-GL- 806
C.BR.07.DEMC07BR003          ---D--A----QN-----G-S------RI--------I----I------------------LI-S-..--.----GE-------Q--G------S-F-P-A-----------------F---AA-A----.--SSLR.......GLQR---I---LGS-V--.-GL- 836
C.BW.00.00BW5031_1           --KD--A--S-KN--S--D-S-------I--------I----I------------------LT-SQ..-E.L--LGR-------Q---------A--F---A-E----------RH---FI---V-A----.--SSLR.......GIQK--D----LGS-V--.-GL- 798
C.CN.10.YNFL19               --KD--A--S-KN--T--D--------RI--------I----I------------------LI-NQ..G-.L--LGR-------Q-KN--V-----F---A-----N-------L---FI-VTA-A----.--NSLR.......GLQT--R----LGG-V--.-GL- 818
C.CY.09.CY260                ---D--A--S-N----------K-----I--------I----I------------L--LI-N-..--.----G--V--E--QQ-RD--S--I-F-VVVAW----S---FC----SYF-LIA-T--V-V-HILLG.......GLRR-GVIMYFVS-VVL-.---L- 810
C.ES.08.X2363_2              --KD--------NN--T--D--------I--------I------I---------------L--LT-N-..--.----LG--------Q--S---------F---A-----------------FI--AA-T--Y-.-H-...............---LGS-V--.-GL- 818
C.ET.02.02ET_288             ---D--A----KD--S--S--Q------I--------I----I----L-------------LT-N-..-E.----LR--------Q-----------F--IA---I-----------------I--AA-A----.--SSLK.......GLQR---I---LGSIV--.-GL- 807
C.IN.09.T125_2139            --KD--A--N-KN-------S-------I--------I----I------------------LI-S-..G-.----L-R-------Q-----------F---A-----N-------------FI-VTA-A----.-HHSLR.......GLRK-------LGS-V--.-GL- 833
C.KE.00.KER2010              --KD--A--S-NN-----D--R---------I--VS-I----I-V-F--------------LT-N-..--.----LGR-------Q-K-G------F---A-----------------CI---A-V----.--SSLR.......GLQK-------LG-CV--.-G-- 812
C.MW.09.703010256_CH256.w96  --KD--A--S-QN---------K------I--------I----I------------------LI-NQ..-E.----LGR-------P--N---------F---T-----------------FI-VTV-V----.--SSFR.......GLQR-------LGS-V--.-GL- 803
C.TZ.08.707010457_CH457.w8   ---D--A--S-KN--S--S----------I--------I----I--I---------------LT-N-..--.----LGR---G---Q-N---T---S-F---A-------F--------FT---A-V----.--SSLR.......GLQK---T---LGS-V--.-GL- 795
C.YE.02.02YE511              --KD--AS-S-NN-----------------I--------I-V--I-T----I-----------L---LT-NS..--.----LG--------Q--N---Q-A--F--VV-N-----------------FT---A-A----.--SSLK.......GIQR-------LGS-V--.-GL- 812
C.ZA.08.705010534_CH534.w12  ---D--T--S-NN-----S--K------I--------I----I------------------LT-NQ..-E.L--L-R--G---Q-------------F-S-A-----------------FI---V-A----.--SSIK.......GLQR---I---LGS-V--.-IL- 803
C.ZA.10.DEMC10ZA001          --KD--A----QN--S------------I--------I----I------------------LN-N-..--.----L---------Q--G------S-F-P-A--------------H---FI---V-V--R-.-SSSLR.......GLQR-------LGS-V--.-GL- 803
C.ZM.11.DEMC11ZM006          --RD--A----N---T--D-S-------I--------I----I------------------LT-SS..-E.---LGR-------Q-KG------S-F---A--------------H---FI--AA-AA---.--SSLR.......GLQR-------LGS-I--.-GL- 813

D.CM.10.DEMD10CM009          --K---------------S--Q------I---V----I-------------K---------L--A-..--.-----------G---Q--G---------FS--------N-----------I--AA-T-N--.---..................L-------.--R- 794
D.CY.06.CY163                --KD-----------R---D-SK------I--------IR-K-------PIH---R---P--H-L--AS..-P.-A---P---G---Q--G---------FS--------N-----------I--AA-T-G--.---..................L-------.-R-- 808
D.KE.11.DEMD11KE003          ------K--Q--------S--K------I--I-----I----------V--K------------L--A-..--.--------------QG----------FS--------N------RH-------A-------.---...............I--L-------.-I-- 800
D.KR.04.04KBH8               --KD--------------S--R------I--I-----I----I----L---------------PF-A-..--.-------T--G---Q--GK-----S-LS--------N-------------A-------.---...............V--LG------.-I-- 792
D.SN.90.SE365                ------A-----------D---------I---V----I----I----L----------------L--A-..--.--------------QG----------FS--------N---------I--AA------.---...............V---L-S-----.-I-- 800
D.TZ.01.A280                 --K---------------S--K--------------I--K---T-F-L-----------L--L--AS..--.--------------QG-G---------FS--------N------RH----I--A---G--.-H-...............I--L-------.-I-- 802
D.UG.10.DEMD10UG004          ------Q--Q----------S------RI--------I----------L--------------L--A-..--.--------------QG-------LG-LST-------N-I-----------I--AL--L---.---...............SR-L-------.-I-- 797
D.UG.11.DEMD11UG003          -----WQ-----------T---------I--------I--------------------------L--A-..--.---------------G----V----FFP-------N------------------AA------.---...............I--L-------.-I-- 807
D.YE.02.02YE516              --K---------------S-SQ------I------S-I----------M-----------L--L--A-..--.--------------PGKGT-------FS--F-E---N---Y---Q----I---A-T-D--.---..................L-------.--R- 796
D.ZA.90.R1                   --K-------------------------I--------I----------V---------------L--AQ..--.---------------G--K--------FYS-------N-------R-----------AA------.---..................L----------.---- 808

F1.AO.06.AO_06_ANG32         ------A--------S--D---------I--------I-------------------K------L--LI-S-..--.---------GD--Q---K-V---S-F-P-V-----N------RH---FI--AA-T-NRG.L-S...............L---V--.--R- 790
F1.AR.02.ARE933              ------A-----------S-----------I----FI-------------------K------L--LI-S-..-E.---------G---QGK---V---T-F---A-----N------RH---FI--AA-T-DRG.L--...............LLG--TL-.---- 795
F1.BR.10.10BR_RJ015          --L---------------D--H------I--------I------T-----------K--------L--S-..--.---------G---Q-----V-----FF--A-----N------RL---FI-T-A---GKG.L--.............S------LG-IIL-.--K- 797
F1.CY.08.CY222               --K---A-----------D-S-------I----I---I-------------------K------L--LI-S-..--.---------G---Q-----V-----FFS-V-----N------RH---FI--AA-T-NRG.L--...............I---L---I--.-I-- 794
F1.ES.02.ES_X845_4           --K---A-----------D-S-------I--------I-------------------K---------LI-S-..--.---------G---Q-----V-----F---V-----N------RH---FI--AA-T-NRG.L--...............L---A--.--R- 792
F1.RO.96.BCI_R07             ------A--------T--D--K-----RI--------I-------------------K------L--LI-S-..-E.---------G---Q--G--V-----F---V-----N------RH---FI--AA-T-NRG.L-G...............L---T--.---- 808
F1.RU.08.D88_845             --K-----------S--D-S------RI--I------I--------F---S---K--------L--L---Q..--.---------------QG----V---S-F-TIV-----N----N-RLW--FI--AA-V-DRG.LK-...............L---A--.-G-- 812
F2.CM.02.02CM_0016BBY        ---D--A----D---S--T--------RI---V----I-----------I-------------L--LI-S-..--.-----G---------Q-K---V---S-F---A-----------------RH---FI--AA-T-DRG.LKG...............V---L---A--.-GR- 795
F2.CM.10.DEMF210CM001        ---D--S-SQ-D-------S--Q------I--------I------------R---------V-L--LI-S-..--.-----------------QGK---V---S-F-S-V-----I------RH---FI--AA-T-TRG.LQL...............T---L---TLH.-G-- 805
F2.CM.10.DEMF210CM007        ---D--A----D------A--H------I--------I---------A-IKK---------L--LI-S-..--.-----G---------Q-K---V---S-FF--V-----N------RH---FI--AA-T-NRG.LE-...............K----L---A-H.---- 797

G.BE.96.DRCBL                ---D--A--Q-----S--S-S------RI-V------I--------------------------L-HHQ..-E.----A----G---Q----------S-F---A-----------------FI--AA-T----.--NSLK.......GLRL-------L----L-.-AR- 804
G.CM.10.DEMG10CM008          ------A-----------D-S------R--VI-----I--------------------------LTHQQ..-E.-----R---G---QG---------S-F---A-----N---WC------FT--AA-T--I-.--SSLQ.......GLSL.......V----L-.-G-- 797
G.CN.08.GX_2084_08           ---D--A-----------D--K------I--------I---I-----------------------THHQ..-E.-----R--------Q-K---V---S-F---A-----------L------FI--AA-T----.K-SSLQ.......GLRLV--G---L----L-.-G-- 809
G.CU.99.Cu74                 ---D--A--------T--D-S-------I--------I------T---LI---K---------PTHHQ..-E.-----R---G---Q--N------S-F---A--------------------V--AA-T----.-HSSLK.......GLRL---G---L----V-.-G-- 815
G.ES.09.X2634_2              ---D--A--T--------Q-SQ------I--------I-----------I---K---------LTHHQ..-E.----GR--------Q---K-----S-F---V--------------H---FI--A--T----.--NSLK.......GLRL---G---L----L-.-GR- 808
G.GH.03.03GH175G             ------A--------S--D--K------I--------I------T-------------------LTHQQ..-E.---L-RT--G---Q-------V---S-F---A-----------------FA--AA-T----.--SSLK.......GLRL---G---L----L-.-V-- 812
G.KE.09.DEMG09KE001          ------A--------S-LS--------RI--------I----I---------------------LTHHQ..-E.-G---R---G------K-------S-F---A-----N----L------FI-LAA-T---...............GLRLS--G---L----V-.-G-- 802
G.NG.09.09NG_SC62            ---D--A--S-----D--A-SK------I--------I-FK-----F-----------------LPHHQ..-E.-----R---G---Q---------S-L---A---V--------------FIS-AA-TA---.--SSLQ.......GLRL--AG---L----L-.-G-- 806
G.ZA.01.TV546                --HD--A-NN-----S--D--K-----RI-V------I--X----------------V---IPTHHQ..-D.----R--------Q--GK-V---S-F---A--------------------S--V-TL---.-HNSLK.......GLRL---G---LKS--E-.-G-- 805

H.BE.93.VI991                ---D--A-----N-----S--------RI-------II-------------------L--LI-NQ..--.----RE-------Q-----------F-P-V-E---N------R------S--A-T----.---..................LLG---L-.-G-- 809
H.CF.90.056                  ---D--A--------T--D-SH------I--------I----I----------------------LV-N-..--.-------T--G---Q-----V-----F-PVV-------S----RL--------V-T----.---.............-R-----L------.-G-- 795
H.GB.00.00GBAC4001           ---D--S----S------D-SX------I---VI---I------T----I-S------------LT-NA..-R.-------------QG----TP--S-FFI-V-Q--WN------R----------A-T----.---..................L------.-G-- 802

J.CM.04.04CMU11421           ------A--Q-GN-----D-SK-----RI--------I----I---FA----------------LT-HL..AE.A-----T--G---QG-TX--------FF--A-E---N-----------FI---A-T--T-.---...............I---LG--AW-.-G-- 810
J.SE.93.SE9280_7887          --KD--A----TN-------S-------I----I---I----I----A-----------------LI-N-..TE.A---G----G---QG-T---------F---A-----N----------FV--AA-T-GT-.-L-...............I---LV--VW-.-G-- 798
J.SE.94.SE9173_7022          --KD--A----TN-------S-------I----I---I----I----A-----------------LI-N-..TE.A---G----G---QG-T---------F---A----------------FV--AA-T-GT-.-L-...............I---LV--VW-.-G-- 800

K.CD.97.97ZR_EQTB11          ---D--A-----------D-S-------I--------I------T---V---------------LT-S-..--.---------G---Q-K---V---S-F---A-----N------RH----V--A---LDRG.LKG.............S------L---IL-.-G-- 801
K.CM.96.96CM_MP535           ---D--A-----------D--K------I----I---I---A-----V---------------LI--S.--.A-----------Q-KN--V---S-F---A-----N------RQ--N-I------L-RG.L-G...............L---V--.---- 792

U.CA.01.TV749                ---D--A--Q----S--D---------RI--------I----I-T-I-M-----------L-I-P-N-..E-.L--L-R-AG----P--T---------F---A-N-----F-----------V--AG-T----.-HSSLK.......GLRL---G---LG---L-.-G-- 818
U.CA.99.TV721                ---D--A----T------D--K------I---V---XI----I---I--I--------------I-T-S-..E-.L--L-R-AG----PG-T-----AD-F---A-------F-----------V--AA-T----.--SSLK.......GLRL---G---L----A-.-G-- 821
U.CD.83.83CD003_Z3           S-KD--------------D---------I--------I------T-F--I---------------L-----..--.----GRT------E-NN--V-----F---A-E-------I----------I---VK.................GLRR-------LG--VL-.-G-- 797
U.CD.90.90CD121E12           ------------------D-S-------I--I-----I------T----I---------------LI-----..--.----G-T------K-NT--V-----F---A-----------------FIS--AK.................GLRR-------LG--V--.---- 795
U.CY.05.CY090                --KD----SN--------D-S-------I--------I------S-----K--------------LI-R-..-E.---------GG----Q-----V-----F---A-----N---L--H--T-V---A-T-N.........K.......GLQGS--I---LG--AL-.-G-- 795
U.CY.08.CY223                --KD-----N--------T-SK-----RI--------I-----------I---------------IPT-N-..E-.L--LGR--G----Q--T------G-F--IA-----N-------Q---FI---V---T-.---..................LGS-V--.---- 830
U.ES.10.DEURF10DZ001         ---D--A----S---S--D-S-------I---V----I--------FNLI---------------PIHLQ..-E.L--L-RT--G---Q---------S-F---A-----------------Q--N-V--AA-A----.-HSSLK.......GLRL-------L--I-S-.-G-- 803
U.GR.99.99GR303              ---D--A--TX-------K-S------RI--------I---------V-K-------------IPI-S-..--.L-----------Q-K-------S--FS-A-----N-----------FI---A-T--I-GTQ#...............-QX---LGEPFC-#-V-- 806
U.NL.95.U_NL_95_H10986_D1    ---D--A--T-----S----S-------I--------I------T---I----------------LF-SS..--.-------T--R---Q-----------F---A-----N------RH---------V--L-RG.L-G.............S------L-S-V--.---- 811

01_AE.AF.07.569M             --MD--K--Q--------D---------I--------I----I--------------------PSHHQ..-E.---------G---QG----V---S-F---A-----------------FI--A--T----.-HSSLK.......GLRR---G---LG---L-.-G-- 808
01_AE.CF.90.90CF11697        ---D--A-----N-----S----------I---V----I-------------------------LTHHQ..-E.-----R---G---EGK--------S-F-S-A-----------------FI--AA-T----.-HSSLK.......GLRR---G---LG---S-.-G-- 813
01_AE.CN.10.YNFL03           --KD--------------D--K------I--------I----I---------------------PTHHQ..-EP----R---G---QG----V-----FF-----------------L---FIS-AA-T----.-HSSLK.......GLRR---G---LG---L-.-G-- 810
01_AE.HK.04.HK001            --KD-----Q--------S--K------T--------I----I--------------------IP-HQQ..-E.----GR------GQ-----V---S-F---L-----N----------FI-----T----.-HSSLK.......GLRL---G---LG---L-.-G-- 809
01_AE.IR.10.10IR.THR48F      --KD--------------D--K------I--------I----I------------------L-IPTHQQ..-E.-----R---G---Q-----V---T-F--IA----------------K--A--AA-T----.-HSSLK.......GLRR---G---LG---A-.---- 817
01_AE.JP.x.JRC77AE           --KD--------------S--H------I-------II---I-----------------------PTHHQ..-E.----------G---QG----V---S-F---V---------------------AA-T----.-HSSLK.......GLRR---G---LG---I-.-G-- 806
01_AE.TH.09.AA090a_WG11      --KD-------T------S---------I--------I----I------------------L--LSHQQ..-E.L--------G---QG---------S-F---A-----------------FI--AA-TLG--.-...............RQ---G---LG---L-.-G-- 807
01_AE.TH.90.CM240            --KD--------------D---------I--------I----I--------------------PSHHQ..KE.---------G---QG----V---S-F---A-------------------T--AA-T----.-HSSLK.......GLRR---G---LG---L-.-G-- 807
01_AE.US.05.306163_FL        --KD--------------D---------I--------I----I-------X-------------PIHHQ..-E.-----R---G---Q-----V---S-F---A-----------------FI---A-T----.-HSSLK.......GLRR---G---LG---V-.-G-- 786
01_AE.VN.98.98VNND15         --KD-----------------------I--I---S-I----IC--HT-------D---S-L--PTHHQ..-E.-----R---GRC-QG----V---S-F---A--------------L---FI--AA-T----.-HSSLK.......GLRR---G---LG---L-.---- 817
```

```
                                                                     transmembrane domain   gp41 cytoplasmic tail start                         glycosylation NGS
B.FR.83.HXB2                 NEQELLELDKWASLWNWFNITNWLWYIKLFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTHLPTP..RG.PDRPEGIEEEGGERDRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELL.GRR..............GWEALKYWWNLLQY.WSQE  806
02_AG.CM.10.DE00210CM013     ---D--A----D---S------------I---------I--------FA-I---------------LTHHQ..GE.----GRT--G---Q--N--V---S-F---F-----N--I---------FV-----T----.-HNSLK.......GLRL--G----L----S-.-G--  834
02_AG.CY.09.CY256            ---D--A-----N-----D--R------I---------I------T-FALI--A-----------LTHHQ..-E.-----R---G---Q--S--V---S-F-T-V-------------------FA--AA-T----.--SSLK.......GLRL-------LG---S-.-G--  807
02_AG.ES.06.P1423            ---D--A---------------------I---------I---------N-I---------------LTHHQ..-E.-----R---G---Q-----V---S-F---A-E---N----C-------V--AA-G--I-.-HSSRR......GLEL------FLG---S-.-GR-  812
02_AG.GW.05.CC_0048          ---D--A--------S--D--K------IV-I-----I------T--T-I---------------LYHHQ..-D.-G---R---G---Q-----V---S-FF--A-------------------F---A--T----.-HSSLK.......GLRL-------L----L-.-G--  817
02_AG.KR.12.12MHR9           ---D--A--R--------D---------I---------I---------T-I---------------LTHHQ..-E.-----R---G---QG--K-V---S-F------------------------V---A-T----.-HGSLK......GLRL------HL-S--A-.-G--  819
02_AG.LR.x.POC44951          ---D--A-----------D--K------I---------I---------T-I---------------LTHRQ..-E.-----R---G---Q---K-V--LS-F---A-------------------FA--AA-T----.-HSSLK.......GLRL-------L----S-.-G--  802
02_AG.NG.09.09NG_SC61        ---D--A-----------D--K-----RI---------I---------ALI---------------LARHQ..-E.----------G---Q--N--V---S-F---A-------------------F---AA-T----.-HSCLK.......GLGL-------L----S-.----  811
02_AG.NG.x.IBNG              ---D--A-----------D-------RI---------I---------T-I---------------LTHHQ..-E.-----R---G---Q-K---V---S-F---A--------------------I--AA-T----.-HNCLK.......GLRL--G----L---IS-.-V--  804
02_AG.SE.94.SE7812           ---D--A--------D--S-SS-----RI--I-----I---------A-I---------------LTHHQ..-E.-G---R---G--GQ-----V---S-F---A-------------------FVS--A-T----.-H-..............--------L----S-.-G--  809
02_AG.SN.98.98SE_MP1211      ---D--A-----------D-------LI-L-V----I---------A-I---------------LTHHQ..--.A---------G---Q-----V---S-F---A----------L------FV--AA-T----.-HSSLK.......GLRL------HL----S-.-G--  811
03_AB.RU.97.KAL153_2         ----I-A-----------D-SK------I--------------I--------G----------R--AQ..--.------------D-----T-----RF------------F-I--H-------AA------.--..................-------------.-I--  789
04_cpx.CY.94.94CY032_3       ---D--A-----------S--K------I---------I---I----------------L--LI--TQ.--.L--G-T------Q--S--------F-P------N--------RH--N-----A-T----.-I-.................-------L--F-L-.-G--  805
05_DF.BE.x.VI1310            ------S--Q--------D--K------I---------I------T---------K----------PF-A-..--.------ET--G---Q--G--------FS--------N----------H----T---V-------.--...................-------L-S-P--.--R-  811
06_cpx.AU.96.BFP90           ---D--A----N--S--D----------I---------I--------I-------------L--LI-N-..T-.A--GE---G---QG-T--------F---A-------------------FG--AA-T--I-.--...............---I---LG--IC-.-G--  809
07_BC.CN.98.98CN009          --KD--A--S-KN--S--D---------I--I-----I---I-------------------LT-N-..G-.---LGR-------Q-K---------F---A----N---------------FI-VAA-V----.--NSLR.......GLQR-------LGS-V--.-G--  809
08_BC.CN.97.97CNGX_6F        --KD--A--S-KN--S--D---------I---------I---I------------------ILT-N-..G-.-G-LGR-------Q-KTS-------F---A----N---------------FI-LTA-G----.--NSLR.......GLQR-------LGS-V--.-GL-  814
09_cpx.GH.96.96GH2911        ---D--A--S-----S--D--------R----------I---I----V-------------------LT-N-..-E.----------G---Q-KV-----S-FF--A----N---------------FI--AA-T--I-.-H-..............V-QI---LG--A--.-G--  792
10_CD.TZ.96.96TZ_BF061       ------Q--------S--S---------I---------I----C---L-----------LH-L--A-..--.-----------------G-G-----D-F-T-------N----------------I--A--------.--M.................-I--L-------.-I--  790
11_cpx.CM.95.95CM_1816       ---D--A--------T--D---------I---------I-------------C------V-L--LI-SH..KE.A-S-G----G---Q--T------S-F---A-----N---Y---X----I------LGT-.---...............---T---LG--T--.----  815
12_BF.AR.99.ARMA159          ------A-----------D-SK-----RI---------I---------------K------L---I-S-..-E.----------G---QGKV--V-----L-T------------------------V-------.--.......................----------R-  792
13_cpx.CM.96.96CM_1849       ---D--A--Q--------S-SK------I---------I------------KS-------------LTHPK..-E.-----R---G---Q---------S-F---A-------------------FI--A--T----.-HSSLK.......GLRL---G---L----V-.-GR-  817
14_BG.ES.05.X1870            ---------------S--D---------I-----I---I----------I----K----------PTHHQ..-E.----G---------Q-K-------S-F---A------------X---FI-VAA-T----.--NSLK.......GLRL---G---L----L-.-GR-  815
15_01B.TH.99.99TH_MU2079     ------A-----------D---------I---------I-------------------------L--PIHQQ..-E.A----E---G---QG-NG-V---S-F---A-------------------K--VS-A--T----.--SSLK.......GLRL---G---LG---L-.-G--  804
16_A2D.KR.97.97KR004         --KD--A-----------D-S-------I---------I-----M-II-V----------V---IPP---..ED.---HGR--DG---QG----V---S-F---V-E------------------CIS-AA-T----.-HSSLK.......GLRL---G---L----L-.-GR-  812
17_BF.AR.99.ARMA038          ------A-ET--------D-S-------I-------------V--T----------K---V-L--RP-S-..-E.----------G---QGK----------F-P-------------C--------W--A--------.--..................-R---R-----------  798
18_cpx.CU.99.CU76            ---D--A--------T--D--------RI---------I------------K-------------LI-NQ..L-.------T-G----QG-G--V---S-F-T-A-E----------------------W-AA-T----.--SSLK.......GLRL-------L----V-.-G--  813
19_cpx.CU.99.CU7             ------A----GK-----D---------I---------I---------------------------LTRH..--.---H-R--G---Q-K--------F---A-------------------FIS-AA-T----.-HS..............---G---L----L-.-G--  751
20_BG.CU.99.Cu103            ---D--A--------------S-------------------I------T---I--------------------THHQ..-E.-------GG---Q-N---V---S-F-P-A------------------------AA-T----.--SSLK.......GLRL---G---L----G-.-G--  799
21_A2D.KE.99.KER2003         --K------Q----------SQ-----RI---V---I----------L--------------L--A-..--.------------------G------L--LS-I------N---------------I--AG------.--..................-E-G---L-------.-GR-  802
22_01A1.CM.01.01CM_0001BBY   ---D--A---SN--T--D-S--------I---------I----------I------------------LT-NQ..E-.L---GR---G---Q--NK-----S-F---A-------------------FI--AA-T----.-HSSLK.......GLRL---G---L----I-.-G--  798
23_BG.CU.03.CB118            --RD--A-----------D-S-------I--I-----I-----M-----I-----------------LTHHQ..-E.----------G---Q--G-----AS-F-PIV-----------------------AA-TL---.--SSLK.......GLRL---G---L----L-.-G--  805
24_BG.ES.08.X2456_2          ---D--A--------T--D-S------RI---------I------T-I-LI-------------------THHQ..-E.----GE--GG---Q-KG------S-F-T-A-----------------------AA-.....................GLRL---G---L----L-.-GR-  799
25_cpx.CM.02.1918LE          ---D--A--------S--S-S------RI---------I------T-I---------------------PHHQ..-E.----G--------Q-----------F-P-V-E---N---------------FI--AA-T----.-HSCLK.......GLRL---G---L-------.-GR-  807
26_AU.CD.02.02CD_MBTB047     ---D--A--E--------S-SK------I---------I----AL--IT----------------LF-S-..E-.VA--G-------Q--T---------F---A-E---N---C---H---F---AA-TL-I-.-HSSLK......GLRR---G---L----A-.-G--  808
27_cpx.FR.04.04CD_FR_KZS     --KD--A-----------D-S-------I---------I------T-F--I-------------I-T-N-..E-.L--LGR-------QGKS-----S-F---A-----N----L----K-CA--AA-T----.-HSSLK.......GLRL---G---L-------.-G--  813
28_BF.BR.99.BREPM12609       ---D--Q--T--------D---------I------------C----------------------R--VQ..-E.---G---------N------AD-F--I--E--------------------------------.--......................-------.-G--  818
29_BF.BR.01.BREPM16704       ------Q----N--S--D--K-----R-----------------------------------R--V-..--.------D----G-------D-F--I------------------------------A-------.--......................-----------L-  820
31_BC.BR.04.04BR142          ---D--A---QN--S--D--K------I---------I---I-------------------RT-I-..--.----LGE-------Q------V---S-F---A----------C-------T--AA-A----.--SILR.......GLQR---I---LKS-V--.-GL-  799
32_06A1.EE.01.EE0369         ------A--------S--D-S-------I---------I--------------------V-L--RI-N-..T-.A--GE---G---QGNTK---AT-F---V-------------------FG---A-T--T-.-H-..............---I---LK--VC-.-G--  800
33_01B.ID.07.JKT189_C        --KD-------T---S----S-------I---------I---I---------------------PIHQQ..-E.----GR-------Q-----V---S-F---A--------------------V--AA-T----.-HSSLK.......GLRL---G---LG---S-.-I--  809
34_01B.TH.99.OUR2478P        --RD-------T------D-S-------I---------I---I---------------------PTHHQ..-E.-----R---G---QG-N--V---T-F-S-A-------------------------A-------.-H-..............-----R---------.-I--  793
35_AD.AF.07.169H             ---D--T-----N-----S-S------RI---------I---------A-I---------------LT-SQ..GD.LG--GR-------Q--T-------S-F---A-------------------FS---R-T----.-HSSLK.......GLRL---G---L----L-.-G--  804
36_cpx.CM.00.00CMNYU830      ---D--A--Q--------D--K------I---------I---------NVI----------P---LTRHQ..-EE----R-K-G---Q---K-V-----F---A-------------------FI--AA-T----.-HNSLK.......GLRL---G---L----V-.-G--  807
37_cpx.CM.00.00CMNYU926      ------A----TN--S--S--------RI---------I---------------------------LTHHQ..-E.-----R---G---QG----V---S-F---A-------------------FI--AA-T--I-.-HSSLK.......GLRL---G---L----V-.-G--  802
38_BF1.UY.03.UY03_3389       --K---------------D-SK------I---------I----I--------------------RF-S-..-E.----------G---Q-K---V---S-F---A-----N--------RH---FT--AA---GRG.LN-..............---I-Q---------.-I--  794
39_BF.BR.04.04BRRJ179        ------A--Q------------------I---------I---------A------K----------I-S-..PE.----------G---QGK---V---S-F---V---V-N--------RH---FI-VAA---NRG.L--................------LLG---L-.----  794
40_BF.BR.05.05BRRJ055        ---------------S--S---------I----------------T---------------------IRP-AA..--.----G--------------G----L--I--------------------------AA------.--......................-----------K-.----  801
42_BF.LU.06.luBF_18_06       ---D--A-----------D---------I----------------S---------------------RH-AQ..--.------T-------------RP--D-F-T---V-----F--------H-------AA------.--.....................V-----------.-R--  790
43_02G.SA.03.J11223          ------A-----X-----D-S------RI---------I------------X--------------LTHHQ..-E.----------G---Q-----------F---A-E-V-N---------------FI--AA-T----.--SSLQ.......GLRL---G---L----L-.--R-  796
44_BF.CL.00.CH80             ---D--A-----------S--------RI---------I------T----------K------L--PI-N-..--.----------G---QG----V-----F---V-----N--------RL---FI-TAA---DRG.LKG...............------LG-----.-G--  794
45_cpx.FR.04.04FR_AUK        ---D--------TN-----D-SH------I---------I--------V-------------------LI-N-..--.----------G---Q-K--------S-F---V-----N---------------FI--AA-T--T-.--..................--QI--HLG--A--.-G--  795
46_BF.BR.07.07BR_FPS625      ---D--AF----------D-S------RI---------I---I----------K------L---I-S-..EE.----R----G---QG----VG----FF--V-----N--------RH---FI--AA---DRG.LK-...............------LLG--AW-.-G--  785
47_BF.ES.08.P1942            ------------------D-S-------I------------------I-----------------------H-A-..--.-----G---------------G-SAS-F-T---I--------------H-------AA------.--.....................L---------------  798
48_01B.MY.07.07MYKT021       --KD-----------------------RI---------I-------...A------------------P-HQQ..-E.-----R-------QG----V---S-F---A-----------------------AA-T----.-HSSLT.......GLRR---G---LG---I-.-A--  814
49_cpx.GM.03.N26677          ---D--S----TN--S--D-S------RI---------I---I----A-I---------V-L--LI-N-..TE.A---G----G---QG-T---------F-SIV----------C--RH---FV--AA-T--T-.-H-..............---I---LG--VL-.-GK-  805
50_A1D.GB.10.12792           --K------------------K------I---------I---I----L-------------------L----..--.-E---------QG-G--------F-T---E------V---------------I--AA---I-.-H-...............----I--L-------.-I--  806
51_01B.SG.11.11SG_HM021      -------------N-------S------I---------I------T--N-I---------------R-------..--.A------G--------TFESSAT-F-----V-----F------R---------A-V-------.--.....................-----------------  801
52_01B.MY.03.03MYKL018_1     ------------------D---------I---------I---I-----------------------LSHHQ..-E.-----R---G---QGS---V---S-FF--A-------------L---------AA------.--..................-G---R---------.-I--  810
53_01B.MY.11.11FIR164        --K--------TN-----D-S------RI---------I---I-----------------------LTHQQ..-EQL----R---G---QG----V---S-F---A-X--------------------I--AA-A----.-HSSLK.......GLRR---G---LGX--S-.-G--  806
54_01B.MY.09.09MYSB023       --KD-------T---S----S------RI--V----I-----------------------------LTHHQ..-E.----------G---QG----V-----F-EI--E-------------------FI--AA-T----.-HSSLK.......GLRL---G---LG---V-.-G--  814
55_01B.CN.10.HNCS102056      ------------------D--K------I-------II----IL-------K--------------PSHHQ..-E.-----R---G---QG----V---S-F-P-A------------L------FI--AA-T-K--.-....Q.......GLRR---G---LG---L-.-G--  812
56_cpx.FR.10.URF5_patient_A  ---D--A--------T--D-S------RI--I-----I---------------------------LTHHQ..-E.----GR-------Q-----V---S-F---A---------I-L------FV--A--G----.-HSCLK.......GLRL------HL-S--S-.-G--  802
57_BC.CN.09.09YNLX19sg       --KD--A--S-Q--------S-------I---------I---I---A-----------------LA-N-..G-.---LGR-------Q-KT------Q-F---A-----N---------------FI-VTA-V----.--ISLR.......GLQK-------LGSIV--.--LK  820
58_01B.MY.09.09MYPR37        --L---------------D--K------I---------I---------------------------LTHHQ..-E.--K-------G---P-----V-----F---A-------------------FI--AA-T----.--TSLK.......GLRR---G---LG---L-.-GK-  814
59_01B.CN.09.09LNA423        ------------------D--K------I--------------I---------------------ILSHHQ..-E.----------G---QG-------S-F---A-------------------FISVTA-T----.-HSSLK.......GLRR---G---LG---A-.----  809
60_BC.IT.11.BAV499           ---D--A---QN--X--D--S------I---------I---I--------IF-------L--LI-G-..--.---LRE--K--K-Q-K------S-F---A--------------------I--TA-A----.--SSLK.......GLQRR--I---LGS-V--.----  809
61_BC.CN.10.JL100010         --KD--A--S-KN--S--D--R------I--------------I-----V------------------PT-N-..G-.H--LRR-------Q--T--V-----F---A-----N---------------FISVTA-V-V--.--NSLR.......GLQK-------LGS-V--.-GL-  798
62_BC.CN.10.YNFL13           --KD--A--S-KN--S--S-S-------I---------I---I-----------------------LT-N-..--.---LG--------Q-K---------F---A-------------------FI-VAA-AGK--.-HSSLK.......GLQR-------LGSIV--.-GL-  808
63_02A1.RU.10.10RU6637       ------A----TN--S--D-S------RI---------I------T----I----K----------LTHHQ..-E.----G----G---Q-----V-----F---F-----------------Q----A--AA-T-Q--.-HSSLK.......GLRL-------L----A-.-G--  805
64_BC.CN.09.YNFL31           --KD--A--S-KN-----D---------I---------I---I-------------------L-IPTQS-..G-.---LGR-------Q-KS--------FF--A-----N---------------FI-VTA-V----.--SSLR.......GLQT--R----LGS-V--.-GL-  809
65_cpx.CN.10.YNFL01          --LD-M---R--------D--K------I---------I---I-----------------------PFHHQ..-EQ----------G---QG---------T-FF--A-----N---------------FI-VT--V----.--SSLK.......GLRQ---G---LG---L-.-G--  820
67_01B.CN.11.2011.ANHUI.HF115--KD--------------D--K------I---------I---I-----------------------PXHQQ..-E.----G--------QG----V---S-F---F-------------------FI--A--T----.-HSSLK.......GLRR---G---LG---L-.-G--  797
68_01B.CN.11.2011.ANHUI.WH73 --KD--------------D---------I---------I---I-----------K---------PFHHQ..-E.----GE-------Q-----R---S-F-P----------------R--I--V--A--T----.--SSLK.......GLRL---G---LG---L-.-G--  798
70_BF1.BR.10.10BR_PE004      ------A-----------D--K------I---------I-----------I---------------F-A-..--.------TG------------GQ--T-F--I--V-----------------------A-------.--.......................-----------------  800
71_BF1.BR.10.10BR_PE008      ---D--A-----------G---------I---------------V----I--------------R--AQ..--.----G----GG---G---NT-GQS-D-F--I--V-----S---I--S------T-A-------.--..................------S--------.----  806
72_BF1.BR.10.10BR_MG002      --K-----------S--D-S-------RI---------I------S---------------------H-A-..--.-----G------------S---ET-AG-F-T---I----------FRH---------T------.--..................---I---G--I----.----  797
O.BE.87.ANT70                --KK-----E---I---LD--K------IA-I---A---V-VIMI--N--KNI----Q---L-IPNHHQ..EE.AGT-GRTGGG---EG-P-W-PSPQ-F-P-LYT---TII-WT--L-SN-ASGIQKVISY-.RLGLWI.......LGQKIINVCRICAAVT--.-L--  809
O.CM.98.98CMA104             --KK----SE---I----D--K------IA-I---A---V---MI--NL-RNI----Q-----IPIHHQ..SE.VGV-G-TG-G---G--P-L-P-QQ-F-Q-LYT---TII-WT--L-SN-ASGTQ-VISH-.-IGLWI.......LGQKIIDGC-LCKAVI--.-L--  830
O.CM.98.98CMABB141           --KK-----E---I---LD--K------IA-I---A-I-V---MV--NL-RNI----Q---L-IPIHHX..SE.AGT-GRTG-G---EG-P-L-ASPQ-F-P-LYT---TII-W---L-SN-ASGIQKVISH-.RLGLWT.......LGQKIIS-CRICAAVT--.-L--  833
O.CM.98.98CMABB212           -VKQ-----E--------D--K------IA-I---A-I-I-VIMI--N--KNI----Q---L-IPIQHR..QE.AEA-GRTG----DG--P-WTP-PP-F-QQLYT---TII-WI--L-SN-ASGIQ-LISY-.RLGLWI.......LGQKII--CRLCGAVI--.-L--  836
O.CM.98.98CMU5337            --RK-----E---I---LD--K------IA-I---A-I-V-VIMI--NL-RNI----Q---L-IPIHHQ..PE.AGT-G-TG-G---EG-PKL-PSPQ-F-X-LYT---TII-W---L-SN-ASGIQTVISH-.-LGLRI.......LGQKIID-CRICTAVT--.-L--  816
O.CM.99.99CMU4122            --KK----NE-----D-LD--K------IA-I---A-I-I---MI--NL-RNI----Q-----IPNHHQ..SE.AGA-G-TG-----G--P-L-P-PP-F-H-LYT---TII-W---L-SN-ASGIQTVISH-.-LGLWI.......LGQKIIN-CRICAAVV--.-L--  821
O.FR.92.VAU                  --K------E---I---LD--K------IA-I---A-I-V-V-ML--NL-KNI----Q---L-IPIQQQ..AE.VGT-G-TG-G--DE--R-WTP-PQ-F-H-LYT---TII-WI--L-SN-ASEIQKLIRH-.-LGLWI.......IGQRTI--CRLFKAII--.-L--  823
O.GA.11.11Gab6352            --KK-----E---I----D--K------IA-I---A---V--IMVI-NL-RNI----Q-------SHHR..SE.AET-GRTG-G---EG-S-LTASPQ-F-P-LYT---TIVVW---L-SI-ASGIQKAISH-.RLGLWI.......LGEKIIS-CRICTAVV--.-L--  831
O.SN.99.99SE_MP1299          --KK-----E---I---LD--K------IA-I---A-I-V---MI--NL-KNI----Q----K--IHHQ..PE.AEA-G-TG-G----GMPTL-PWPQ-F-P-LYT---TII-W---L-SN-ASGIQTVISH-.-LGLWT.......LGQKIIS-CRLCIAVI--.-L--  828
O.US.10.LTNP                 --KS-----E-V-I---LD--K-----TA-I---A-I-V-VIMI--NL-RNI----Q---L-IPIHHQ..VE.AGT-G-TG-G---EG-P-WRP-QQ-F---LYT---TII-W---L-SN-ASGIQKVISH-.RLGLWI.......LGQKLIN-CRICTAVT--.-L--  812

N.CM.02.DJ00131              --KS-----Q--------S---------IA--V-A-II-I--ISVIIT-IA----------L--LI--TTT--.-----ET--GV-GQG----V---S-F-T-V-E-F-N-LI-L----T-S---LQ-TL---.--SLIR.......GLQLLN-LRTRL-GIIA-.-GK-  787
N.CM.02.SJGddd               --RS-----Q----T--DL--------IA----A-II-I--ISVIIT-IA----------L--LI--T.---.X-----T--GV--Q--G--V---S-F---V-E---N-SI-L----T-S---LR-TL---.RQSLSR.......GLQLLN-LITRL-GI-A-.-GG-  786
N.CM.04.04CM_1015_04         --KS-----Q-T------D-S-------IA--V-A-II-I--ISVIIT-IA----------L--LI--T.--.------T--GV--Q--G--V-----F---V-E-C-N-VTLL----------TQ-TL-I-.-QSLNR.......GLQLLN-LR-RL-GI-A-.-GK-  783
N.CM.06.U14296               --KS-----Q--------D-S-------IX----A-II-I-XISVIIT-IA----------L--LI--T.--.----GETG-DV--Q--G--V---S-FFS---E---N-VI-L-R--T-S---LX-TL---.-QSLSR.......GLQLLN-LRIHL-GIIA-.-GK-  808
N.CM.06.U14842               --KS----XH--------D--K------IA----A-II-I--IS-IIT-IA----------L--LI--T.--.-----ETG-G---Q-----V-----F---V-E---N-LT-L----T-S---LR-TL---.-QSLSK.......GLQLLN-LRIHL-GIIA-.-GK-  805
N.FR.11.N1_FR_2011           --KS-----Q-T--------S-------IA----A-II-I--IS-IVT-IA----------L--LI--A..-D.-----T--GV--Q--G--V-----FS--V-E---N-LI-L----T-S---LR-TL---.-QSLSK.......GLQLLN-FRI-L-GI-A-.-GK-  794

P.CM.06.U14788               --RK--------Q--S-MD--K------I-------II--K-LM-IGNV-RK-------V-L--LRLHG..GP.AGIAP-TD-----AGNG-----LE-L-P-------N-VVQI-QI-VSCI-RIKDLLTI-.WIHLGQ.......LLNS..NC-RDCFAACG-.-T--  835
P.FR.09.RBF168               --KK-----Q--Q----LD--Q------I-------II----LL-IINV-R-I------V-L--LGLNG..DP.AGIAP-TN-----AGNG-----LD-F-P-V---KN-VVQI-QI-VGCI-GIKDLLTI-.WIHLGQ.......LLTR-LNC-RDCFAACG-.-T--  829

CPZ.CD.06.BF1167             -HKQ-------S------D---------IA-I--AS-------MFIINL-GKL-----L--P-IPIQ-A..--.-GQ---T--G---SG---WRGYPV-F-Q---E-C-Q-L-WI-QTSVN--WHLWTLIK-I.HQWSLLLWETVRQLLRRLIHL-TQFLH--ET.-GIL  834
CPZ.CM.05.SIVcpzMT145        --K------S--N--S--D---------M-LIV---II----C--IG-LI----K-------L--LI-AN..Q-.--GLGET-KG---NV-G-------S-F-P-V-E---N-LS-L--Q---CASLIW-LL-I-.-QYSLR.......GVQQIGTL-HQLRGT----.-TT-  805
CPZ.GA.88.GAB1               --RD-----Q--------D--K------I-L-A---II---IMT-F-V-R-----------L--LI-VQ..-E.QG-LGE-D-G---Q--S--V---E-C-P-------N-GIW--QS-TS-ACN-W-QLKT-.-HLILH.......SLRLLR-R-CLLGGII--.-GK-  804
CPZ.TZ.06.SIVcpzTAN13        -KK--------ST--D--D--Q-------A-I-IA--I----LMFIVNVLRQ-----M--FS-IPTQAE..QD.-EQ-G--A-G--D-G-I-WTPSPR-FFSIV-E---N-L-WI-QTCLNFVWLLGT-LQA-.KQGTISFIQTIIIVHRNLVLKWRLFRDW-AN.AYVI  820
CPZ.US.85.US_Marilyn         --KS-----Q-S------D---------I-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-.-QNINK.......GLQLLN-LRARC-GVIA-.-AR-  786

GOR.CM.07.SIVgor2139_287     --KA--------D--S-LD--Q------I--I-IA-------LM-IINMCH------------ILGRNQ..VP.AGIVPE-G----RA-N------LA-F-P-----I-T-VVWI-RT-VI--SGIKE-TLS-.IEYLTR.......LLRRVNNLTRDCFAFIA-.-G--  819
GOR.CM.07.SIVgorCP2135con    --KA--------D--S-LD--Q------I----IA-------LM-IINMCH------------ILARNQ..VP.AGIVPE-G----RA-S------LA-F-P-------T-IVWI-RT-VI--SGIKE-TLS-.IEHLTR.......LLRRVNNLTRDCFAFIA-.-G--  814
```

```
                                                                          Env end
                                                                         gp41 end
                                    glycosylation NAT            cytoplasmic tail end
                                                              *
B.FR.83.HXB2                 LKNSAVSLLNATAIAVAEGTDRVIEVVQGACRAIRHIPRRIRQGL.ERILL*   856
A1.AU.03.PS1044_Day0         --I---N-YDTI-----GW------IG--I----LN---------.--A---   871
A1.CH.03.HIV_CH_BID_V3538    -RV--IT--D----T--GW------IG-RF----L----------.--A-I-   855
A1.CY.08.CY236               --I--IN-FDTI-----GW------I---IG---IN---------.--A---   863
A1.ES.05.X1608_8             --L---N-YDTI-----GW------IG-RL----LN---------.--A---   874
A1.KE.11.DEMA11KE001         --I--T---DTV-----GW------IR-RIG---L---------F.--A---   877
A1.RU.11.11RU6950            -----IN-VDTI--V--GW------IG-------CN--------A.--A-Q-   864
A1.RW.11.DEMA111RW002        --I--I--FDTI--T--GW------I--RI----LN--T-----F.--A---   872
A1.SN.01.DDI579              --S--I--------V---W------IG-R-G--LL-----V----.--A---   851
A1.UG.11.DEMA110UG001        -----IN--DTI-----GW---I--IG-SIG---L---------A.--A---   857
A1.ZA.04.04ZASK162B1         -R---I---DTI--T--GW-G----IG-RIV---L----------.--A-Q-   860
A2.CM.01.01CM_1445MV         ------N-VD-I-VT----------LG-RI----LN--T-----F.--A---   863
A2.CY.94.94CY017_41          -----I--FDTI-V----W------IG-R-F---LN---------.--A---   866

B.BR.10.10BR_MG029           ---------------------------RI#---LN--T-MG---#--A---   868
B.CA.07.502_1191_03          --------F-----------------L-R-G--VL-V-------F.--S---   863
B.CH.08.M2_0803101_NFLG8     -----TN---TI----------I--LA-RGF--FL----------.--A-Q-   886
B.CN.10.DEMB10CN002          -QK-------------------I--A--RIG---I----------.----V-   859
B.ES.10.DEMB10ES002          ------N---T---V-------I---A-R-W--FL---T-----A.--V---   865
B.FR.11.DEMB11FR001          -----------------------I--RI----I---V------.--A---   841
B.GB.05.MM45d213_GN1         --K--I----------------A-DII-RXF---L----------.--S-V-   863
B.HK.06.HK003                --K-------FI----------I--A--RGV---L----------.--A-V-   854
B.HT.05.05HT_129389          ---------D---V--G-----I--I-RRVV---L-V-T------.--A---   863
B.JP.12.DEMB12JP001          --K--I----------------I---I-RL----L---V------.--A---   865
B.KR.07.HP_18_07JHS10_3909   --K------H---R-I------A--IL-R-Y---L---------F.--A---   858
B.PE.07.502_0525_wg5         -----------I----------I---GLNIF--FI---A------.--L---   843
B.RU.11.11RU21n              --------F-------G-----------RTF--VL---------F.--A---   843
B.TH.08.MERLBDTRC10          ----------V-----------------R-Y---I----------.--T-V-   864
B.US.11.ES38                 --S--I--F--I------------GL-R-W--FI----------.------   863

C.AR.01.ARG4006              --K--I---DTI----------I---I-RIWEI-CN--------F.-AA-Q-   857
C.BR.07.DEMC07BR003          --KR-I--FDTV----------I--II--IW---CN--------F.-AA-Q-   887
C.BW.00.00BW5031_1           --K--I---DTI----------I-----RLY---LN--------F.-AA-Q-   849
C.CN.10.YNFL19               --K--I---DT-----------I-AIX-R-W---....T--G-DP.-VA-Q-   865
C.CY.09.CY260                --K-VS--DDT-TT-A-GAA--IV-LLRRISGI-CS-R----RC-.AAA-Q-   861
C.ES.08.X2363_2              --K--I---DTI----------I--FI-RI----SNL-------F.-AA---   869
C.ET.02.02ET_288             --K--I----T---V-------I--L--RFW---CNA-------F.-AA---   858
C.IN.09.T125_2139            --KG-I--FDTI----------I--FL-RL----YN---------.-AA-Q-   884
C.KE.00.KER2010              --R--I---DTI----------I--LI-RI----LN--T------.-AA---   863
C.MW.09.703010256_CH256.w96  --K--I---DT-----------I--FL-RI----Y---------F.-AA-Q-   854
C.TZ.08.707010457_CH457.w8   --K--I---DTI---T------I--L--RI----CNV-T------.-AA-Q-   846
C.YE.02.02YE511              --K-----VDTI----------I--I--RS---LCN--------F.-AA-Q-   863
C.ZA.08.705010534_CH534.w12  --K--I---DT------T----I---I-RG----CN--------F.-TA---   854
C.ZA.10.DEMC10ZA001          --K--I---DTI----------IF-A-LRF-----N--T-----F.-AA---   854
C.ZM.11.DEMC11ZM006          --K--I---DTI--RT------I--L-LNI----LS--------F.-AA-Q-   864

D.CM.10.DEMD10CM009          -----I--FD--------------A-LR-F---L----------.--A---   845
D.CY.06.CY163                -----IN--D--------------T---RT---VLR--------F.--A---   859
D.KE.11.DEMD11KE003          ---------DTI--VI------L--IGLRIG---IN---------.--A---   851
D.KR.04.04KBH8               -----IR-FD-----------------RR-F--VL---------F.--A-I-   843
D.SN.90.SE365                -----I---DTI----------I-D---R-----L---T------.-GA---   851
D.TZ.01.A280                 -----I----V-----------I--II-R-F--VL-----V----.--A---   853
D.UG.10.DEMD10UG004          V----I--FGT---V-------I---L-RIG---LN--T-----A.--AF--   848
D.UG.11.DEMD11UG003          -----I--F-T---V----------L--R-G--VLN--------F.--A---   858
D.YE.02.02YE516              --I--I--F-------------I--I--R-G---LN---------.--A---   847
D.ZA.90.R1                   -----I---DTI---T-----------R----VLNV-V------.--L---   859

F1.AO.06.AO_06_ANG32         -----IN--DT---V-----------L-R-G--VL----------.------   841
F1.AR.02.ARE933              -RI--I--F-T---V----------A--R-V--VCN---------.--AF--   846
F1.BR.10.10BR_RJ015          --------------V-G-----I--AL-RFG---LN--------F.--A-I-   848
F1.CY.08.CY222               -----I---DTI--V-----------AL-R-G--VLNV--K-----.------   845
F1.ES.02.ES_X845_4           -----I---DT---V-------I--AL-RTG---L-----V----.--A---   843
F1.RO.96.BCI_R07             -----I----TI--V-----------AL--VLT--LN---------.--A---   859
F1.RU.08.D88_845             -----I----T---V---W----L-AL-R-G--VL----------.--L---   863
F2.CM.02.02CM_0016BBY        --I--I----T---V----------IL-R-G---L---------F.--A---   846
F2.CM.10.DEMF210CM001        --K--I--FDT---V-------I---L-R-G---L---------A.--A---   856
F2.CM.10.DEMF210CM007        -----I---DT---V-------I-D-L-RIG--VL---------F.--A---   848

G.BE.96.DRCBL                -----IN--DTI-----NW-------A-R-G--VLN---------.--A---   855
G.CM.10.DEMG10CM008          -----IN--DT------DW-------A-RVG--FLN---------.------   848
G.CN.08.GX_2084_08           -----IN--DTI-----NW-------AHR-----CN---------.--A-V-   860
G.CU.99.Cu74                 -----IN-VDTV-V---NW------I--RT---FLN--------F.--A-Q-   866
G.ES.09.X2634_2              --S--IN--DTV-----NW------I--RVG---LN---------.--A---   859
G.GH.03.03GH175G             -----I---DTV-----NW-------I-R-G---LN--T-----F.--A---   863
G.KE.09.DEMG09KE001          -----IN--DTV-----N-------I--R----VL---------F.-SA---   853
G.NG.09.09NG_SC62            -----I--VDTL--V-GNW-------A-R-G--FL-----V----.--A---   857
G.ZA.01.TV546                -----IN--DTX-V---NW------IA-RTG---CN---------.--A---   856

H.BE.93.VI991                -----I----T----------I--L--R-W---L---------F.--A---   860
H.CF.90.056                  -----ID---T---------GI-VI--R-W---L---------F.--S---   846
H.GB.00.00GBAC4001           ------N---T---V-------I--L--R-G---I---------F.--A---   853

J.CM.04.04CMU11421           -----I---DT---VX------I--IXRR-F--VL---------F.--A---   861
J.SE.93.SE9280_7887          -----I----T-----------I--IA-R-F---L----------.--A---   849
J.SE.94.SE9173_7022          -----I----T-----------I--IA-R-F---L----------.--A---   851

K.CD.97.97ZR_EQTB11          I----IN---T-----------I--I-YR-F--LL---------F.--L---   852
K.CM.96.96CM_MP535           -----I----T------G----I--IG-R-F--LL----------.--A---   843

U.CA.01.TV749                --T--IN---TV-------------G-RI-----N--------F.--A---   869
U.CA.99.TV721                --T--IN---TV--V-------I---G-RI----CN--------F.--A---   872
U.CD.83.83CD003_Z3           -----I--------V-------I---G-RI----LN--------F.--A---   848
U.CD.90.90CD121E12           -----I--------V-------I--L--RVG---L---------F.--A---   846
U.CY.05.CY090                -----I---DT---V-------I--IG-RIF--FIN---------.--A---   846
U.CY.08.CY223                -----I----TI----------I--IG---I---LN---------.--A---   881
U.ES.10.DEURF10DZ001         -----I---DTI---T-NWI--A--GI-RIG--LLN---------.--A---   854
U.GR.99.99GR303              -QK--T---DI---V--GW------TG-RIG---L---------F.--A---   857
U.NL.95.U_NL_95_H10986_D1    -----IN---T-----------I--IG-R-F--VL---------F.--A---   862

01_AE.AF.07.569M             -RI--I---D-I-----GW------I----W--FI----------.------   859
01_AE.CF.90.90CF11697        --T--IT-FD-I-V---GW---------R-W--LI----------.--A---   864
01_AE.CN.10.YNFL03           --T--I--FD-------G----------R-W--LL---------F.--A---   861
01_AE.HK.04.HK001            --I--I----T-------W-------A-R-W---L----------.--T-V-   860
01_AE.IR.10.10IR.THR48F      --I--S---D-------GW-------A-R-W--LL----------.--A---   868
01_AE.JP.x.JRC77AE           --T--I---D-A------W-------L-R-----L----------.--T---   857
01_AE.TH.09.AA090a_WG11      --I--I---DTI-----GW-------A-R-W---IQ---------.--A-V-   858
01_AE.TH.90.CM240            --I--I---D-----A-GW-------A---W---L----------.--T---   858
01_AE.US.05.306163_FL        --T--I---D-----T--W-------A-RTG---LN---------.--A-I-   837
01_AE.VN.98.98VNND15         --I--I--IDTA-----GW-------A-R-W--FL----------.--A-Q-   868
```

```
                                                                    Env end
                                                                    gp41 end
                                  glycosylation NAT     cytoplasmic tail end
                                                     *
B.FR.83.HXB2                 LKNSAVSLLNATAIAVAEGTDRVIEVVQGACRAIRHIPRRIRQGL.ERILL*   856
02_AG.CM.10.DE00210CM013     -----TN--DTI-----NW------I--R-----YN---------.--A---   885
02_AG.CY.09.CY256            -----IN--DTI-V--GNW--G----G-R-----CN--T------.--A-Q-   858
02_AG.ES.06.P1423            -----IN--DTI-----NW------IA-R-G--FL---------F.--A---   863
02_AG.GW.05.CC_0048          -----IN---TI--V--NW------IG-S-----CN--T------.--L---   868
02_AG.KR.12.12MHR9           -----L--FDTI-----NW------IG-R-G---LN--------F.--A---   870
02_AG.LR.x.POC44951          -----IN--DTI-----NW------IG-R-G----N--T------.-KT---   853
02_AG.NG.09.09NG_SC61        -----------I-----NW---I--IG-R-G-....................-   844
02_AG.NG.x.IBNG              -----IN---TI--V--NW---A--IG-RVG----N--------F.--A---   855
02_AG.SE.94.SE7812           -----I---DTI--V--NW------L--R-G---LN--------F.--A---   860
02_AG.SN.98.98SE_MP1211      -----IN--DT--V---NW------I--RTG---CN---------.--A-Q-   862
03_AB.RU.97.KAL153_2         --S--IN-IGTI-----GW------IG-RF---M-N--------A.-KA-Q-   840
04_cpx.CY.94.94CY032_3       -----IN-F-T-----------I--A--R-----CN---------.--A---   856
05_DF.BE.x.VI1310            -----I----T--VV----------AL-R-G---LN---------.--A---   862
06_cpx.AU.96.BFP90           -Q---I--FD-A-----NW----A----RIF--FLNV-------F.--A---   860
07_BC.CN.98.98CN009          --K-TI--VDTI----------I--L---L----YN--------F.-AA-Q-   860
08_BC.CN.97.97CNGX_6F        --K-TI--VD-I----------I-NR---I----HNV-------F.-AA-Q-   865
09_cpx.GH.96.96GH2911        -----I----T---V-----------L-R-G---L---------F.--A---   843
10_CD.TZ.96.96TZ_BF061       -----I---DTI--E--GW------I--R-V---LN--T------.--A---   841
11_cpx.CM.95.95CM_1816       -----I----T---V-------I---IHRXL--VL----------.--A---   866
12_BF.AR.99.ARMA159          -----I----T---V--G----IL-AL-RIG--VLNA---V----.--A---   843
13_cpx.CM.96.96CM_1849       --I--IT--DS-------W---I--TG-RIG-G-L---------F.--A-Q-   868
14_BG.ES.05.X1870            -----I---DTI-----NW---A-X----VG---LNV-V------.--F---   866
15_01B.TH.99.99TH_MU2079     --I--I---DT------GW-------A-R-W---LR--------F.--A---   855
16_A2D.KR.97.97KR004         -----I--F---------W------I--R-----IN---------.--A---   863
17_BF.AR.99.ARMA038          -----I----T---V----------AL-R-G--VLN--------F.--A---   849
18_cpx.CU.99.CU76            I-S--IN--DTV---T-NW------I--RTG---LN--------A.--A-Q-   864
19_cpx.CU.99.CU7             --T--I----T---V--GW---A--II-R-W--VL----------.--A---   802
20_BG.CU.99.Cu103            -----IN--DTV-----NW---I--I--RGF--FL---------F.--A---   850
21_A2D.KE.99.KER2003         -R---N--Y---------------I-RR-----LN---------F.--A---   853
22_01A1.CM.01.01CM_0001BBY   -----IT--D----T--GW------IT-R-W--FL---T------.--A---   849
23_BG.CU.03.CB118            -----IN--DTV-----N----I--I--RGF--FL---------F.--A---   856
24_BG.ES.08.X2456_2          -----IN-YDTV-----N----I--I--RGF--FL---------F.--A---   850
25_cpx.CM.02.1918LE          -----IN--DTI---T-DW------IA-R-G---LNV-V------.--A---   858
26_AU.CD.02.02CD_MBTB047     -----IN---TI-V----C------IA-RGG---L----------.--A-Q-   859
27_cpx.FR.04.04CD_FR_KZS     -----IN--DTV---T-NW-------A-R-G---IN--------A.--A-Q-   864
28_BF.BR.99.BREPM12609       -------------------------IL-RI-----N--T------.--A---   869
29_BF.BR.01.BREPM16704       ----------------------------RIW---L---T------.--A---   871
31_BC.BR.04.04BR142          --K--I---DTI----------I---I--IG---YN--------F.-AA-Q-   850
32_06A1.EE.01.EE0369         -----IN-IDTI-----DW-------G-RVF--FLN---------.-IA---   851
33_01B.ID.07.JKT189_C        --I--I-----------GW-------I-IVW---L---------F.--S---   860
34_01B.TH.99.OUR2478P        ----------------------I-----R-Y---L----------.--A-I-   844
35_AD.AF.07.169H             --T--TN-IDTI-----GW-----DIG-RFG---L----------.-GA---   855
36_cpx.CM.00.00CMNYU830      --T--IT--D---VT--GW-------A-RTG---LNV-T------.--A---   858
37_cpx.CM.00.00CMNYU926      --I--IN--DT---V--GW------IA-R-----L---------F#--A-Q-   853
38_BF1.UY.03.UY03_3389       -----I--F-T---V-------I--AL-R-G--VLN---------.--A-I-   845
39_BF.BR.04.04BRRJ179        -----I----T---V-------IL-AL-TIG-G-L----------.--S---   845
40_BF.BR.05.05BRRJ055        ----------TL----------I--A--RIG-G-L----------.--S---   852
42_BF.LU.06.luBF_18_06       -----I----------------I---L-RIG--VL---T------.--A---   841
43_02G.SA.03.J11223          -----IN-VDTI-----GW-------A-R-----IN--------F.--A---   847
44_BF.CL.00.CH80             -----I----T---V--GW-------L-R-G--VL---------F.--A---   845
45_cpx.FR.04.04FR_AUK        ---------D-------GW------T--SL---FLN---------.--A---   846
46_BF.BR.07.07BR_FPS625      -----I----T---V-------I--AL-R-G--VLN---------.--A---   836
47_BF.ES.08.P1942            -----I----T---V-------I--AL-S-G--VL----------.--A---   849
48_01B.MY.07.07MYKT021       --I--I--------T-------I---ARR-W---I----------.--A---   865
49_cpx.GM.03.N26677          -----I--FD-I----------I---ARR-F--LLN----T----.--A---   856
50_A1D.GB.10.12792           -----I--FDTI--V--SW---A---G-RIG---L---------A.--A---   857
51_01B.SG.11.11SG_HM021      --------F----V-----------I--RVW--FL----------.--S---   852
52_01B.MY.03.03MYKL018_1     ---------DT-----------I-----R-Y---L---T------.--A---   861
53_01B.MY.11.11FIR164        --T--I---D-A-----GW---------R-W---L----------.--A-Q-   857
54_01B.MY.09.09MYSB023       --I--I---DT--V---GW---------R-W---L----------.--A-Q-   865
55_01B.CN.10.HNCS102056      --I--IA---TI--V--GW------AA---G---L----------.--S---   863
56_cpx.FR.10.URF5_patient_A  --K--I---DTI---I-NW------LG-A-G--VL-V-------F.--A---   853
57_BC.CN.09.09YNLX19sg       --K-T---YDSI----------I--FI-RI--S-LN---K----F.KAA-Q-   871
58_01B.MY.09.09MYPR37        --I--I---D-------GW---------SIW---L----------.--A---   865
59_01B.CN.09.09LNA423        --T--I---DTL-----GW------LA-R-W--LL-V--------.--A-Q-   860
60_BC.IT.11.BAV499           -----I----T-------R--KI--IL-RF-K--L---AK---D-.-------   860
61_BC.CN.10.JL100010         --K-TI--YDTI----------I--L--RLY----NV-T-----F.--A---   849
62_BC.CN.10.YNFL13           --K--IN--DTI----------I--TI-R-Y---CN--------F.-AA-Q-   859
63_02A1.RU.10.10RU6637       -----IN-FDTI-----NW-------G-RVG--LL----------.--A---   856
64_BC.CN.09.YNFL31           --K--I---DTI-------------I-R-----L---------F.-AA-Q-   860
65_cpx.CN.10.YNFL01          --I--I---D---VT--GW-------A---W---L-V--------.--A---   871
67_01B.CN.11.2011.ANHUI.HF115-I--I--FD----XI-GW------AA-R-W--FL---------F.--A-V-   848
68_01B.CN.11.2011.ANHUI.WH73--I--I---D---V-T-GW------AA-R-W---L---------F.--A---   849
70_BF1.BR.10.10BR_PE004      ----------T---V--------V-AL-R-G--VLNV-------F.--A---   851
71_BF1.BR.10.10BR_PE008      ----------T---V---------GAL-SIG--FL----------.--A---   857
72_BF1.BR.10.10BR_MG002      -R--------TI--V--------L-AL-TLG---I----------.--A-Q-   848
O.BE.87.ANT70                -Q---T---DTL-V---NW--GI-AGI-RIGTG--N---------.--S---   860
O.CM.98.98CMA104             -Q---T---DTI-V--GKW--DI-LGI-RIG-G-LN---------.--S---   881
O.CM.98.98CMABB141           -Q--VI--VDTI-V---NW--NI-LGI-RLG-G-WN--T-----F.--S---   884
O.CM.98.98CMABB212           -Q---I--FDTI-V-A-NW--NI-LGI-RIG-GXLN--------A.--X-X-   887
O.CM.98.98CMU5337            -Q---T---DTI-V-T-NW--SI-LGI-RIG-G-YN---------.--S---   867
O.CM.99.99CMU4122            -Q---TN--DTL-V---NR--SIFLGI-RIG-G-LN---------.-QA-V-   872
O.FR.92.VAU                  -QT--TN--DTV-V---NW--SI-LGI-SIG-G-LN---------.--L---   874
O.GA.11.11Gab6352            -Q---TN-VDTL-V---NW--GI-LGI-RIG-G-LN--T------.--S---   882
O.SN.99.99SE_MP1299          -Q---T---DTI-V---NW-VTI-LGI-RIG-G-LN---------.--S---   879
O.US.10.LTNP                 -Q---T---DTF-V---NW--GI-SGI-RIG-G--N---------.--S---   863

N.CM.02.DJ00131              --D--I----TI--V-------L--LA-RIG-G-L----------.--A---   838
N.CM.02.SJGddd               -RD--I----T---V-------L--LA-RIG-G-L---------A.--A-V-   837
N.CM.04.04CM_1015_04         --D--I----T------G----L--LA-RIG-G-L----------.--T---   834
N.CM.06.U14296               --D--I----T---V-------L--LA--IG-G-L----------.--A---   859
N.CM.06.U14842               -XD--I----T---V-------L--LA-RIG-G-L----------.----V-   856
N.FR.11.N1_FR_2011           --D--I----T---V-------L--LA-RVG-G-L----------.--A-I-   845

P.CM.06.U14788               --Q--T----TV--S--NW--Q--A-G-QIG-GFLN----L----.--S---   886
P.FR.09.RBF168               --Q--T---DTV--S--GW--Q--I-G-QIG-GFLN--------I.--S--X   880

CPZ.CD.06.BF1167             IRQHCIVCIDTL-EFTGWW--G---ALRV-VDI-----T-------.-IA-N-   885
CPZ.CM.05.SIVcpzMT145        I----I---DT-----G-----IL-AITRLG-G-L----------.--A---   856
CPZ.GA.88.GAB1               --I--I---D------------I--AF-VTL-I--N---------.--A---   855
CPZ.TZ.06.SIVcpzTAN13        V-T-FIRN-DRL-NFT-WW--IL--GAVNIF-G--N--T------.-IA-N-   871
CPZ.US.85.US_Marilyn         --V--T---DT-----------I--LTRRLFLG-I----------.--S--X   837

GOR.CM.07.SIVgor2139_287     --Q--I---DCV-VWT-NW--Q--AIA-RIG-G-LN---------.--S---   870
GOR.CM.07.SIVgorCP2135con    --Q--I---DCV-VWT-DW--Q--AIA-RIG-G-LN---------.--S---   865
```

```
                            myristoylation                                                                                              phosphorylation
                            Nef start                                                                    acidic cluster  poly-P helix                                HXB2 premature Nef end
B.FR.83.HXB2                MGGKWSKSSVIGWPTVRERMRR.......AE.................PAADRVGAASRDLEKHGAITSSNTAATNAACAWLEA...QEE..E..EVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQ.RRQDILDLWIYHTQGYFPD*QNYTPGPGVRYPL 136
A1.AU.03.PS1044_Day0        -------R-SK--SEI--KI-Q.......TP.................T--EG---V-K--D----V----V..NHPS-T----.----.G.------R---------F-G-F---F--------D---Y-X.K--E-----V-N-------W--------I---- 135
A1.CH.03.HIV_CH_BID_V3538   ---------IV---EI------.......TP.................A--EG---V----R-------------S---------E.G.------R---------F-G--------------D---Y-R.K--E-----V-N-------W--------T-F-- 138
A1.ES.05.X1608_8            ---------IV---EI-D-I--.......XP.PA..............---SG---V-Q--D-R--V----I..NHPS----N-.----X.-..D-----R-----------G-F----------------Y-R.K--E-----V-N-------W------------- 137
A1.KE.11.DEMA11KE001        ---------IV---E-------.......-P................S--PG---V-Q--A----V----I..NHPS------.----D.-.-------R-------------I-----------D---Y-R.K--EL----V---------WH------------ 135
A1.RU.11.11RU6950           ---------IV---Q----I--.......-P.AP............A---SG--PV-Q--DR---V--------------------.-------R-----------G-F------------D---Y-K.---------V---------W--------I---- 140
A1.RW.11.DEMA111RW002       --S------IV---E------Q.......-R.RAT...........PP---AG---V--------------V..NHPS------.----D.-.-------R-------------F---F------------Y-K.---E-----V-----F---W------------- 140
A1.SN.01.DDI579             --------RQE--HK--D---Q.......TP.PT...............---EG---V-Q--S----V--T-Q-..S-S-T-----.-.-.-------R-----------G-F------------D---Y--.Q---------V---------W--------T---- 137
A1.UG.11.DEMA110UG001       ---------IV---E----I-N.......TP..............AP---KG---V-Q--D---------L..NHPSNT------.-.-.-------R----T------------------------Y-R.K--E-----V-----F---W--------I-F-- 137
A1.ZA.04.04ZASK162B1        ---------IV---QI---I-Q.......TP.TA............A--EG---V-----R---------E-N-ESL--VR-.----E.-.-------R-----------G-F------------D---W-R.K--E-----V-N-------W------------- 140
A2.CM.01.01CM_1445MV        --S-L--.......---A-------.......TP.................---EG---V-Q--GTR--V-IR------SD-------.----S.D.-------R-----------G-F------------D---Y--.K--E-----V---------W----------F-- 134
A2.CY.94.94CY017_41         -------R-IP---AI------.......TP.PTAQRTE.......AVS---PG---V-Q--AT---V---------PD---V-.----E.S.-------R---------F-G-F---F--------D---Y--.K------M-V---------W--------I---- 148
B.BR.10.10BR_MG029          -------R-GK-----------.......--.................---EG---V-----TR---------E--D---------.-.-.-------R------------------------------------Y--.---------V---------W--------I---- 137
B.CA.07.502_I191_03         -------GGMG--NA--D----.......T..................---EG---------R-----------N-PD------.----N.-.-------R--------------------------D-------TW--.K--------V-N-------W--------T---- 137
B.CH.08.M2_0803101_NFLG8    --S-L--C-PV--S-I------.......T-.PRNE..........QNEQ--VG---------R-----------D--------.-.-.-------R-----------G-L-------------------------.K-----------------W------------- 144
B.CN.10.DEMB10CN002         -------PRLG--SAI------.......T-.P.............VE---EG---------ER--L-T-----N------------.----D.-..D-----R-----------G-L-------------------------.K--E-----V---------WH-------T-F-- 140
B.ES.10.DEMB10ES002         ----------V---AI-D----.......--.PA............AE---AG---------DR---L-------N--DY-----.----DD.G.-------K-----------G-L------------------Y--.Q---------V-N-------W----Q-------- 142
B.FR.11.DEMB11FR001         ---IF--R-GG---A-------.......TG.TEPKAERLGKERVRRTE---EG---------R-----T----DN--DS-----.----D.-.-------R-----------G-L------R-----------Y--.K---------V---------W--------I---- 154
B.GB.05.MM45d213_GN1        ----------V---X-------.......-R...............AE----G---V-Q--GR-----------N------V--.--T-G.-.-------R-------------F-------------------W--.Q---------V-N---F---W--------T-F-- 140
B.HK.06.HK003               -------RGRD--DAI---I--.......--.PR............AE---AG---V----AQR-------Q..N--DRV-----.----EE-.-.-------R-----------D-F-----------------VY--.K-K-------V-----F---W-C-------I---- 141
B.JP.12.DEMB12JP001         ---------IV---A----I--.......--.PR............AE---EG---V-------------I..N--D-V------.-.-.-------R---------F-E-F------------------Y--.K-K---N--V-----I---W-E-------T-F-- 139
B.KR.07.HP_18_07JHS10_3909  -------R--P--DAI------.......T-..................---EG---V-----QR----T----SN---------.-.-.-------R-----------G-----------------VY--.K---------V---------W--------T-F-- 137
B.PE.07.502_0525_wg5        -------C-PV----I-D----.......--..................---EG---------D----L-A-----N---------.-.-.-------K---------F-G-----------------VW--#K---------V---------W--------T-F-- 137
B.RU.11.11RU21n             ---------L----RI----Q-.......T-..................---EG---V---------------------------.-.-.-------R-----------G------------------Y--.K------------------W--------T---- 137
B.TH.08.MERLBDTRC10         ---------LV---AI----K-.......--..................----G---V-----R-----NN--P-N--D-------.----E.-.-------K-----------G-L------R---------VY--.K--------------F---W--------T-F-- 138
B.US.11.ES38                -------I-AD--SR-------.......-G.T.............AE---EG---------R----------VTN------VA-.----DD-.-.-------R---------F---L------N---------VW--.K---------V-N-------W--------P---- 142
C.AR.01.ARG4006             --N----...L---A-------.......-P.PGVGDXR......RQDA---EG--P--Q--DRY--L-----P-N--D-------.--Q-E.-.-------R---------F-G-F---F------------Y--.Q-K-------V-N-------W------------- 146
C.BR.07.DEMC07BR003         --N----C-PV--SAI------.......T-..................---EG--P--Q--DRY--L-----PGN---------.--Q-E.-.-------RX------------F---F------------Y-K.K----I---V-N-------W--------I---- 138
C.BW.00.00BW5031_1          ---------LV---E--D-I--.......TD..................---EG-----Q--D----L-N-------KD------.----E.G.-------R-----------G-F--GW--------D---Y-K.K--E-----V---------W------------- 138
C.CN.10.YNFL19              --A------IV---AI------.......T-..................----G---------R--------I..N-D--T--Q-.----E.-.-------R----V------G-F---F------------Y-K.K--E-----V-N---F---WS------------ 136
C.ES.08.X2363_2             --------RRV--TE-------.......TK.PAAEP.........AVE---VG--P--Q--D-Y--L-----PSN--D----A-.----E.DE.-------R-----------G-F--GF--------D----XK.K---------V---------W--------I---- 147
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------------- 138
C.IN.09.T125_2139           --S----C-IV---A--D----.......T-..................---EG-----Q--D-Y--------GT---DV----T.----A.G.-------R-----------G-F---F------------Y-K.K--E-----V---------WH------------ 138
C.KE.00.KER2010             ---------IV---A-------.......--..................---EG-----Q--D-Y--L-----GN--D----Q-.----E.-.-------R---------R-G-I---F--------D---Y-K.K---------V-N---F---W------------- 138
C.MW.09.703010256_CH256.w96 ---------IV---N----I--.......TD..................---EG---------R---L-T---VSN----------.----E.K.-------R-----------G-F---F------------Y-K.K---------V-N---F---W------------- 138
C.TZ.08.707010457_CH457.w8  -------C-MV---E----I--.......TA..................---EG---------DR---------T-----------.----A.-.-------R-----------G-----F------------Y-K.K---------V---------W------------- 138
C.YE.02.02YE511             ---------IV---V--D-I--.......TN.PAAAGVG...A..ASQDS--AG-----Q--A----L-----PTN----------.---QD.-.-------R-----------G-----W------------Y-K.K--E-----M-N-------W--------T---- 150
C.ZA.08.705010534_CH534.w12 ---------IV---E----I--.......TR.PAAEGE.........KT--EG--T--Q--D----L-T----RN-D---------.----D.G..D-----R---------F---F---F------------Y-R.K--E-----VHN---F---W------------- 145
C.ZA.10.DEMC10ZA001         --N------IV---A-------.......TR.PTAKRER.......RAK---VG-----Q--D----------Q..---------Q-.----D.D.-------R-----------G-F--GF--------D---W--.K---------V---------W------------- 146
C.ZM.11.DEMC11ZM006         --N----......---AI---I--.......T-..................---EG-----Q--D-F--L-T----KN--DV--Q-.----EGDG.-------R-----------G-F--GF--------D---Y-K.K--K-----V-----F---W------------- 136
D.CM.10.DEMD10CM009         ---------IV---AI------.......TQ.PT.............E----G---V-----------------Q---D-------.----D.G.-------R--------------------------D--VY--.K--E-----V---------W--------T---- 141
D.KE.11.DEMD11KE003         ---------IV---AI---I--.......T-..................---AG---V-----R---------SQ---D-------.-.--.D.-------R-------------------------------VW-R.K--E-----V-----F---W--------I---- 137
D.KR.04.04KBH8              ---------.....--AI-------.......T-..................----G---------R-----R------D-------.----E.-.-------R----T-------L-------D-------VW--.K---------V-N---F---WH-------I---- 133
D.SN.90.SE365               ---------IVE--AI---I--.......T-..................----G---V---------V-----KD---D-------.----D.G..K-----R-------------------------------VW-K.K--E-----V-----F---WH-------T---- 138
D.TZ.01.A280                ---------IV---AI---I-K.......-D..................----G---V-----R---------H---D-------.----A.D.-------R--A-----------L-------D-------VW--.---------V-----F---WH-------T---- 138
D.UG.10.DEMD10UG004         ---------IV---AI---I--.......TD..................---EG-----K--------------Q--P-------.-.-.-..G-----R-----------G-L-------------D-I-W--.K---------V---------W--------T---- 137
D.UG.11.DEMD11UG003         ---------IV---AI---I--.......TA..................---EG---------------------Q--S-------.----E.G.-----IR-----------G-L-----------------VW-P.K--E-----V---------W--------I---- 138
D.YE.02.02YE516             ---------IV----I--KI-K.......TD..................----G---V--------------..N-P-------.---D.-.-.-------R-------------------------------V---.---------V---------W--------I-H-- 135
D.ZA.90.R1                  ---------IV---AI---I--.......TD..................----G---V--------------S---D-------.----S.-.-------K---------F-E------------------W-K.K--E-----V-N---I---W--------I---- 138
F1.ES.02.ES_X845_4          ---------IV---A--D----.......TP.................A--EG---V-Q--DRR------------SDL-----.-.----.A.-------R-------------L-----------D-----R.K---------V-----F---W--------I---- 137
F1.RO.96.BCI_R07            ---------LA---A----I--.......TP.PAAE...........RG--EG---V----DRR--------G---PDL--V--.----D.G.-------R----P------G-L------R---------Y--.K-RE-----V-N-------W--------T---- 143
F1.RU.08.D88_845            --A------IV---AI-D----.......-P.PAAPP........AAEGX--EG---V----DQR--------R---PDL------.-.-.-.-------R-------------F-X-------------Y-K.K--E-----V---------W----------L-- 146
F2.CM.02.02CM_0016BBY       ---------IV---K------Q.......TP.................V--EG--KV-Q--D----------G---NDI-----.----D.-.-------R---------F-G--------------D---Y-R.K--E-----V---------W-------X..... 132
F2.CM.10.DEMF210CM001       ---------IV----I---L--.......TP.AAAAE..........GG--EG---V----D-Y--------R---DL----E.-.-.-.-------R---------F-G-F--GF-------------Y-K.K--E-----V---------W--------I---- 143
F2.CM.10.DEMF210CM007       ---------IV----I---IKQ.......TP.................V--EG---V-Q--D-Y--------KE--PDL------..-D-.-.-------R---------F-G-F---F-------------Y-K.K--E-----V---------W--------T-F-- 137
G.BE.96.DRCBL               --N----RK-A---E----L-Q.......HP.................A--EG---V-Q--DR------R---G--PD------.-.---..DS.-------R-------------F---F--------D--VY-K.K--E-----V-----F---W--------T-V-- 138
G.CM.10.DEMG10CM008         ---------IV---E----I--.......-P.PG............QTPQ--AG---V-Q--DR-----------N-PD------.-.---..DS.D-----R-----------G-F---F--------D---Y-K.K--E-----V-N---F---W--------T---- 143
G.CN.08.GX_2084_08          #-------CIV---EI---I-Q.......TP..................---EG-----Q--DR-------------PD------.-.---..-S.-------R---------F-G-F---F--------D---W-K.K---------V-N-------W--------T-F-- 137
G.CU.99.Cu74                ---------IV---A----I-Q.......TP..................---AG---V-Q--AR-----------N-PD----Q-.-.---..DS.-------R-------------F---F--------D---Y-K.K---------V-N---F---W--------T---- 138
G.ES.09.X2634_2             -------R-IV---EI---L--.......TP.S.............A--EG---V-Q--AR-----------N-PDR------.-.---..DS.-------R-----------G-F---F--------D---Y-K.Q---------V-N-------W--------T-F-- 139
G.GH.03.03GH175G            I------R-IKE--K-K-QI-Q.......-P.................A--EG-----Q--D----L-------N-PD------.-.---..DS.-------K-------------F---F--------D---W-K.K--E-----V-N---F---W--------T-L-- 138
G.KE.09.DEMG09KE001         --N-------V---Q----I--.......TP.PAAAGEG......AASP---AG-----Q--A-----------TN-PD-----...QD..-P.-------R-------------F--GF--------D---Y--.K---------V-N---F---W--------I---- 149
G.NG.09.09NG_SC62           -EN------......--A----I--.......TP.................A--EG-----Q--TR---------SDR-PDV------...Q-..DS.-------R-------------F--GF--------D--TY-K.Q---------V-N-------W--------I---- 134
H.BE.93.VI991               -------GCIS---A----I-Q.......T-..................---EG---V-Q--DRR--V-IN-I-SN--DS-----.----E.-.-------R-----------G-F----------------Y-K.K--E-----V-N-------WH-------E---- 138
H.CF.90.056                 --------RMG--S-I------.......--..................-V-EG---V----DRR--V-IN---S--RDA-----.---DG.-.-------R-----------G-F------------D---Y-K.Q---------V-N-------W--------E-F-- 138
H.GB.00.00GBAC4001          ---R-----IV---RI---I--.......--..................---EG--PV----DXX--V-IN---DR-PDS-----.----E.-.-------R-----------G-L---------------VY-K.K--E-----V-N-------WH------------ 138
J.CM.04.04CMU11421          ---------IV---Q-------.......-A..............APV----G---V-Q--AT----------T---D------...T-.-.-.------------------G-F--GF--------D---Y-K.K--E-----V-----F---W--------X---- 140
J.SE.93.SE9280_7887         --N-----......--Q-------.......-P................A----G---V-Q--A---------------D------...T-.-.-.------K--I-------G-----F--------D---Y-K.K--E-----VHN-------W--------I---- 134
J.SE.94.SE9173_7022         --N-----......--Q--D----.......-A.P..............A----G---V-Q--A--------------DD------...T-.-.-.-------R-----------G-F---F--------D---Y-K.K--E-----VHN-------W--------TX--- 135
K.CD.97.97ZR_EQTB11         ---------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---- 137
K.CM.96.96CM_MP535          ---------IV---AI------.......-R.PAADRV........GTQ----G---V-Q--AR---V-----SHN-PD-----.-.-.-.-.-------R-------------F--GF--------D---Y-K.K--E-----V-----F---W--------I---- 146
U.CA.01.TV749               -------R-IV---AX------.......-X.PXP...........A--XG---V-K---------T-.--SN-PDL--V-.----E.G.-------R-------------------------X---Y--.X---------V-N---F---W--------I---- 140
U.CA.99.TV721               --A-----......---AI------.......TP.PTP...........T---EG---V-KS-Q----V-T-.A--N-PDL--V-.----G.-.-------R----T--------F--GF--------D---Y--.Q---------V-N---FX--W--------I---- 137
U.CD.83.83CD003_Z3          --N-----....Q--AI------.......-R.PAA............E----G---------Y---------E--E---Q---..AQ-D.G.-------R-------------F--GF--------D---Y-K.K--E-----V-N-------W--------I---- 138
U.CD.90.90CD121E12          --N-----....Q--AI-----Q.......-R.PAA............E----G---V----DR-------------S---Q---.----E.G..D-----K-----------------F-------------Y-K.K--E-----V-N-------W--------I---- 138
U.ES.10.DEURF10DZ001        ---------IV---EI---I--.......TP.TAAE...........GA--EGA--V-Q--AR---V------T--PDL------.-.-.-.-------R-------------F---F--------D---Y-K.K---------V-N---F---WH-------I---- 142
U.GR.99.99GR303             ---------IV---EI---I--.......TP.PKDR..........QPL---EG---V-Q--DR---V-----R-N--NL-----.-.---.-.-.-EA--R--L--------E-L---------------W-----.Q---------F-N-------W--------I---- 144
U.NL.95.U_NL_95_H10986_D1   ---------IV---A----I--.......TP.AAEGR..........RA----G---V----DR--------------D------...T-.-.-.------R-------------F---F--------D------.K---------V---------W--------I---- 144
01_AE.AF.07.569M            --S------IV---Q---KIKK.......TP.PAAEGVGAVSQDLDKH.G--EG---V----D----V----M..N--DNV--R-.----EGDDEG-----R---------F-G-F---F--------D---Y-K.K--E-----V-----F---W--------T-F-- 155
01_AE.CF.90.90CF11697       -------NRIV---Q----I--.......TP.................A--EG-----Q--D---------M..N--D--------.----.-.-.-------R-----------G-F---F--------D---Y-K.K--E-----V---------W--------I---- 135
01_AE.CN.10.YNFL03          --N-----.....---R----IQQ.......TP..................---EG---V-Q--D----V----M..N--DST--R-.----D.-.-------K-----------G-F---F--R-----D---Y-K.K--E-----V-----F---W------------- 132
01_AE.HK.04.HK001           ---------IV---Q----I-Q.......TP.................--TEG---V-Q--D----V--N-M..N-DDSV--R-.----DE-..G-----R-----------D-F---F--------D---Y-K.K--E-----V-N---F---W--------I---- 137
01_AE.IR.10.10IR.THR48F     --S------IV---Q----IEQ.......TP..................---EG---V-Q--D----V----M..N--D-V------.----X.-.-.-------R-----------G-F---F--------D---Y-K.K---------VHN---F---W--------I---- 135
01_AE.JP.x.JRC77AE          ---------IV---QI---I--.......TP..................---EG---V-Q--D----V----M..N--ESV--R-.----DE-E.G-----R---------F-E-F---F-------------Y-K.K--E-----V-N---F---W--------T---- 138
01_AE.TH.09.AA090a_WG11     --S-----......--Q----I-K.......TP.PAA...........EEV.--KG---V-Q--A----V----M..N-PDSV--R-...H--DE-.-.-------R-----------E-F---F--------D---Y-K.K--E-----V-N---F---W------------- 138
01_AE.TH.90.CM240           --S------IV---Q---KIKQ.......TP.................--TEG---V-Q--D---------I..D--D-V--R-.----D.-.-.-------M-----------G-F---F--------D---Y-K.K--E-----V-N---F---W--------I-F-- 135
01_AE.US.05.306163_FL       ---------......--Q----I--.......TP.................--TEG---V-Q--D---------#..N--D-T--R-.-----.D.-------K---------F-G-F---F--------D---Y-K.K--E-----V-N---F---W--------T-F-- 130
01_AE.VN.98.98VNND15        ---------IV---QI---I--.......TP.................--TEG---V-Q--D----V----M..N--DSV--R-...H--ED-..G-----R---------F-G-F---F--------D---Y-K.K---------V-N---F---WH-------I---- 137
```

```
                              myristoylation                                                                       phosphorylation
                              Nef start                                                    acidic cluster  poly-P helix                           HXB2 premature Nef end
B.FR.83.HXB2                  MGGKWSKSSVIGWPTVRERMRR.......AE.................PAADRVGAASRDLEKHGAITSSNTAATNAACAWLEA...QEE..E..EVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQ.RRQDILDLWIYHTQGYFPD*QNYTPGPGVRYPL  136
02_AG.CM.10.DE00210CM013      ---------IV---KI---I-H.......TP.ARGERAGG.......EEAP-EG-----Q--GRR---------S--PDV-------.--D.-.-------R-------------H---F------------Y-K.K--E-----V-----F---W--------T-F--  147
02_AG.ES.06.P1423             --A------IV---QI-------.......TP.....................---E---X--Q-KD--E-S--N--T-------GA--.......-.-------V-------------I------------D--VW--.K--E-----V-N-------W--------T-F--  137
02_AG.GW.05.CC_0048           ---------IV---KI-D-IKQ.......TP..................---EG-----Q--D----L------E--PD------.-.-...D-----R-----------S-F---F------------Y-R.K--E-----V-----F---W--------I----  137
02_AG.KR.12.12MHR9            ---------IV---R----L--.......TP.PAAEGVGA.....ASTP---EG-----Q---RR--------------------------.-.-.-EEG----R-----------G-F-F----R---------Y-K.K--E-----V-----F---W--------I----  151
02_AG.LR.x.POC44951           ---------LV---K----I-Q.......TP.VS.............ER--KG---V-Q--DRR--------GS---D-----------.-.-.-------R-------------F-------------D-------.K---------V-----F---W--------T-F--  140
02_AG.NG.09.09NG_SC61         ---------IVE--K----I-Q.......TP.................-T-TG-----Q--D-----------E--PD-------.--D.-.-.-------K-------------F---F------------Y-K.K--E-----V-N---F---W--------T-F--  137
02_AG.NG.x.IBNG               ---------IV---K-MK---Q.......TP.................T--TG-----Q--DR-----------Q--PD-------.--D.-.-.N-----R-----------G---------------D---Y-K.K--E-----V-N---F---W--------T-F--  137
02_AG.SE.94.SE7812            ---------IV---QI-D-I-Q.......TP....................---RG-----Q--DR-------------------------.-.-.-------R--------------------------------Y-K.K--E-----V-----F---W--------T-F--  137
02_AG.SN.98.98SE_MP1211       ---------LV---Q----I--.....TQPTP.................S--IG---V-Q--DR------R---H---D-----------.-.-.-------R------------------------Q-------Y-K.K--E-----V-----F---W--------I----  140
03_AB.RU.97.KAL153_2          ---------IV---Q----I--.......-P.................A---RG--PV-Q--D-Y--V-------N--D-----------.-.-.-------R-----------G-F-------------D---Y-K.K--E-----V----------W......--I-F--  133
04_cpx.CY.94.94CY032_3        ---------IV---EI-------.......-R.AEPERMR...R..AQAE---AG---V-Q--D------IN------PDKT-------.----E.-.-------R----------F-G-L-------------D---Y-K.K--E-----V-----F---WD--------E-F--  150
05_DF.BE.x.VI1310             ----------V---AI--K----.......TP.................---EG---VFQ--DNR--V-IR--V---PDM--V------.----E.-.-------K--------------------------------Y-K.K---------V-N----S--W--------I----  138
06_cpx.AU.96.BFP90            --N----.....--SQ-------.......TP.PT.............ER--EG---V-Q--D-----------T-------V---.-T-.-.-------R----V--------F---F----------D---Y-K.K--E-----V-N---F---W--------T-F--  136
07_BC.CN.98.98CN009           ---------IV---DI-----T.......EQ.....................----G---V---------------P---DD---------.---K.-.-------R-----------G-L---F-------------Y-R.K--E-----V---------WH---------F--  137
08_BC.CN.97.97CNGX_6F         ---------IV---AI---I--.......T-.....................----G---V---------------D---D-----T.----.-.-.-------R---------F-G-F---F-------------Y-K.K--E-----V---------WH---------F--  137
09_cpx.GH.96.96GH2911         --------XIV---EI---I--.......TS.................PSP-EG-----Q--GR---L------S--HD-------.--D.-.-.D-----R-----------G-----F---------D---Y-K.K--E-----V---------W----------F--  138
10_CD.TZ.96.96TZ_BF061        ---------IV---Q-------.......TS.....................---EG-----Q--A-F--L-----PGN--D-------.---AE.-.-------R---------F-G-F--GF---------D---Y--.K-K------V-N---F---W--------T----  138
11_cpx.CM.95.95CM_1816        --------RIV---EI---I--.......TS.PT..............A--PG---V-K-------V-----GE------------.----E.-.-------R-----------G-F--GF---------D---Y-K.K--E-----V---------W--------I----  140
12_BF.AR.99.ARMA159           ---------IV---EI---L--.......-P.....................---EG---V-Q---NR--------R-N-P-L-------.-.-.-.-------R---------F-G-F-----------------Y-K.K--E---------------W----------L--  137
13_cpx.CM.96.96CM_1849        -------R-TV---A----I.........................---TG---V-Q--DR---V----L-..-KDYV--------.--G.-.-------K-----------G-F--GF---------D-----K.K---------V---------W--------------  130
14_BG.ES.05.X1870             -------R-IV---EI---L--.......SP.....................---EG---V-Q--AR---V-------N-PD-------.-Q-..DS.-------R-----------G-F---F---------D---Y-K.Q---------V-N---F---W--------I----  138
15_01B.TH.99.99TH_MU2079      --S-----.....--Q---KIKQ.......TP.....................---EG---V-Q--DN---V----M..N--DSV--R-..E-D.-.-.-------R---------F-E-F---F---------D---Y-K.K--E-----V-N---F---W--------I----  131
16_A2D.KR.97.97KR004          -------R-LP---AI-------.......TP.PAAERT........PPAA--EG---V-Q--ATR--V-IN--------DS-------.----E.-.-------R-----------G-F------R-----D-------.K---------V-----F---W-G-------I----  147
17_BF.AR.99.ARMA038           --------.....---A-------.......--.P..............E---EG---V-----RR--V-----S-N-PDL-------.-.-.-.-------R-----------G-L-----------------Y-K.---E-----V-N-------W--------I----  135
18_cpx.CU.99.CU76             ---------LV---EI---I--.......TQ.....................ELPRTPP-E-QEER---SG-N-S-----PTHGR--E..E--.-.-------R-----------G-F---F---------D---FF-.K-PE-----VFN---FL--W----------F--  137
19_cpx.CU.99.CU7              #--------I----A-----Q-.........--.....................----G---V--------------R----D-----------A.D.-.-------R-----------G-L---------------IVW--.K---------V---------WH--------I----  137
20_BG.CU.99.Cu103             --N-----.....--A----L--.......TP.....................---AG---V-Q--A-------------N-PD-------.-..DS.-------R---------F---F---F---------D---Y-K.Q---------V-N-------W--------T-F--  134
21_A2D.KE.99.KER2003          ---------IV----I---I--.......TD.....................----G--P-----R-----------T--PD-------.---D.-..G----R---I-----G-F---------------D---Y--.K--E-----V-----F---W--------T-F--  138
22_01A1.CM.01.01CM_0001BBY    ---------IV---Q----I-Q.......TP.................A--EG---V-Q--D----V----M..NH-S-V-------.--D..N.-.-------R-----------G-F---F---------D---Y-K.K--E-----V-N-------W-------X.....  130
23_BG.CU.03.CB118             --N-----.....-AA-------.......-P.PAXN............P---AG---V-Q--AR------------------DR--V--.-.-.-..DS.-------R-------------F---F---------D---W-K.Q---------V-N---F---W--------I----  139
24_BG.ES.08.X2456_2           --N-----.-SE--R----I--.......TP.LTTTRGGASQNSARQES---VG-----Q--AR-----------N-PD-------.-Q-..DS.-------R-------------F---F---------D---Y-K.Q---------V-N-------W--------T-F--  154
25_cpx.CM.02.1918LE           -------NR.V---EI---I--.......TP.PAE...........AAA--EG-----Q--A-----------S--PD-------.---S.-.-------R-----------G-F---F-----------VW--.K--E-----------------W-------X.....  137
26_AU.CD.02.02CD_MBTB047      ----------V---E--------.......TP.....................---TG---V----D----V-----SF--PD-------.----E.G.-------R-------------F--GF---------D---Y--.Q-K------V-N---F---W--------I----  138
27_cpx.FR.04.04CD_FR_KZS      ------Q--PV---A---K-----.......-P.....................----G---V-Q--A------NN---N-PD---V---.---D.S.-------R---------------------------D---W-K.K--E-----V-N-------W--------E-F--  138
28_BF.BR.99.BREPMI2609        -------RGTG--EA--------.......TR.PA............AE---G---V-----R-----------HN--D-----------.-.-.-.-------R-----------G-L------R------------P.K--E-----V---------W--------T-F--  141
29_BF.BR.01.BREPM16704        -------C-MS--AA--------.......T-.PGP............AE---EG----------Y------------------------.-.-.-.-------R-------------I-I-------------------.K--E-----V---------W-------------  142
31_BC.BR.04.04BR142           --N----CIPV---AI---L--.......-H.PAAERIG..SERVRGAE---EG--P--Q--DR---L-----RDN--D-------.-Q-E.-.-------R-------------F---F-------------Y-K.Q---------V-N-------W--------T-F--  153
32_06A1.EE.01.EE0369          ---------IV---EI---L--.......TP.....................---EG---V-Q--D--------H-PT-------------.--D.-.-.-------R-----------G-F--GF---------D---Y-K.K--E-----V---------W--------I----  137
33_01B.ID.07.JKT189_C         --S-C-E-CIV---Q----I-Q.......TP.....................---EG---V-Q--D----V--T-M..N--DSV--------.-.-.-.-------R---------F-G-F---F--R-----D---Y-K.K---------V-N---F---W-------------  135
34_01B.TH.99.OUR2478P         ---------IV---Q----I--.......TP.....................---EG---V-Q--D----V----M..N--D-V--R-..---NE-.-.-------R-------------F---F-------------Y-K.---E-----V-N---F---W--------T-F--  137
35_AD.AF.07.169H              ---------KKE--E-K-K--Q.......TP.................A--EG---V-Q-------------I..NHPS-T---TQ..--D.-.-.-------R---------F-G-L--------Q---D---Y-R.K--E-----V-N-------W--------I----  136
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-..DP.D----R-------------F---F---------D-------.K-----------N-------W--------......  132
38_BF1.UY.03.UY03_3389        -------RRIV--ADI-------.......--.....................---EG-----Q--DRR--------KDN-PDV-------.-.-.-.-------R-----------G-----------------Y-K.---E-----V---------W--------I----  137
39_BF.BR.04.04BRRJ179         ---------K----AI-D----.......TP.PTAER.........GAAK--EG----------SR--L------KDN-PDV------.----E.-.-------R-------------L---------------VW--.K---------V---------W----------V--  146
40_BF.BR.05.05BRRJ055         --------R-P---EI-------.......--.PAADG...........ER--G-----Q--A----L-----VSN--D---------.---N.-.-------R----------HR--I------RKE---------K.---------V-----F---W--------I----  143
42_BF.LU.06.luBF_18_06        ---------LV---A------K.......T-.....................---VG--P----------L-T----TK---V---------.-.-.-.-------R-----------G-L-----------------Y--.------------------------W--------I----  137
43_02G.SA.03.J11223           -------R-MG--SEI---I--.......TP.PTAAE.........GAGT--EG-----Q--AR---------PQ---D-----------.--D.-.-------R-R-------------F--GF-----------VY-K.K--E-----V-----F---WH-------I----  145
44_BF.CL.00.CH80              ---------IV---DI-------.......TP.....................---EG--KV-Q---RR--------S-N-PDV-------.--D.-.-------R---------F-----------------------.K---------V---------W--------I----  137
45_cpx.FR.04.04FR_AUK         -------C-IV---EI---I--.......TP.....................---TG---V-Q--D----V---------S-----------.---N.-.-------R-----------G-I---F-----------VY--.K---------V---------W--------I----  138
46_BF.BR.07.07BR_FPS625       --NI--R-RIV---A----L--.......TP.....................---EG-----Q---RR----N---RDN-PDV-------.--D.-.-------R-------------L---------------TY-K.K-HE-----V---------W--------I----  137
47_BF.ES.08.P1942             --N------IV----I---I--.......TP.................I-EG---V----G------------N-PDL-------.--G.-.-------R-----------G-F---F-------------Y-R.K--E-----V---------W-------------  137
48_01B.MY.07.07MYKT021        ---------KV---Q----I-Q.......TP.....................---EG---V-Q--D----V--T-M..N--DSV--R-.----D.-.-------R-----------G-F---F--------G---Y-K.K---------V-N---F---W--------T-F--  136
49_cpx.GM.03.N26677           --SS----RIV---EI----------.......-R.ATP..........TAPA---EG---V-Q--D----------P-----------E..-T-.-.-------R-------------F--G---X.........................................   99
50_A1D.GB.10.12792            -------R..V---ETEK---M.......GP.................AP---PG---V-Q--A---------V..NHPS---------.--D.-.-------R-------------F--GF---------D---Y-K.K--E-----V---------W--------I----  135
51_01B.SG.11.11SG_HM021       --A------PV---KI----K-.......T-.PAA.............EA--EG--P---------L------DN--D-------..H-D.-.-------R-----------F-G-F-----------------W--.K---------V---------W--------E-F--  141
52_01B.MY.03.03MYKL018_1      -------C-..---R-----E-.......-G.....................----G---------R---------------D-----------.-.-.-.-------R-----------G-F---F--R-----D---Y-K.K--E-----V-----F---WH-------I----  135
53_01B.MY.11.11FIR164         ---------IV---Q----I--.......TP.....................---EG---V-Q--D----V-VT-M..N--DST--R-.----E.D..G----R---------F-D-F---F---------D---Y-R.K--E-----V-N---F---WS-------T----  136
54_01B.MY.09.09MYSB023        -------GRLG--TQI-------.......TP.AAAEE..........VGA--EG---V-K--D----V--T-M..N--DKV-------.----D.-.-------R-------------F--GF---------D---W-K.K---------V-N-------W--------T-F--  143
55_01B.CN.10.HNCS102056       --S------IV---Q----IKQ.......TP.................S--EG---V-Q--D----------I..N--DSV--R-.----E.-..G----R---------F-G-F---F---------D---Y-T.K--E-----V-N---F---WH------------  136
56_cpx.FR.10.URF5_patient_A   -------.RIV---Q----I-Q.......TT.PX..............---VG-------D----L-T----QN-PD-------.--D.-.-.-------R-----------G...................................................   84
57_BC.CN.09.09YNLX19sg        ---------IV---AI-------.......TQ.....................---EG-----Q--D-Y--L-----DTN-TD----R-.----E.-..G----R---------F-G-F---F-------------Y-K.K--E-----V---------W-------------  138
58_01B.MY.09.09MYPR37         ---------LV---QI--KI-Q.......TP.....................---EG---V-Q--D----V--T-L..N--D-S--RT.----E.G.-------R-----------G-----F-------------Y-K.Q---------V-N---F---W--------I----  136
59_01B.CN.09.09LNA423         ---------KA---Q----IEQ.......TP.................V-TEG---V-Q--D----V---I..N-TDSV--K-.-----DE-..G----R---------F-G-F---F---------D---Y-K.K--E-----V-----F---WH---X.........  129
60_BC.IT.11.BAV499            ---------IVR--AIK--L--.......T-.....................---EG---V------K------L..N-PDRT--K-...-K-D.K..--S--R-----------G-L-----------------Y--.Q-R------V-N-------W-------RI----  136
61_BC.CN.10.JL100010          -------C--V---S-------.......--.....................----G---V-----Y-------------D-----------..K.-.-------R-----------G-L-----------------Y-K.K--E-----V---------W--------I-F--  137
62_BC.CN.10.YNFL13            --S----R-IV---DI-------.......--.....................---VG-----Q--D-Y--L-----NT---D----S-.----E.G.-------R---------F-G-F---F-------------Y-K.K--E-----V-----F---WD---------F--  138
63_02A1.RU.10.10RU6637        -------R-IV---Q----I--.......-P.................AS--TG--PV-Q--D----V-------N--D-----------.-.-.-.-------R-----------G-L-----------------Y-K.K--E-----V---------W--------T-F--  138
64_BC.CN.09.YNFL31            ---------IV---AI---I-Q.......TD.....................----G---V-------------------DD-------.----E.G..G----R---------F-G-F---F-------------Y-K.---E-----V---------W--------I-F--  138
65_cpx.CN.10.YNFL01           --N----R-IV---EI---L--.......TR.TD...........RTE----G---V----AR-------------D-----------.-.-.-.-------R---------F-G-F---F-------------Y-K.K--E-----V---------W----------F--  142
67_01B.CN.11.ANHUI_HF115      -------X-IV---Q--D-IKQ.......TP.................A-TEG---V-Q--D---------I..N--DNX--R-.----EG-E.-------R-------------F---F---------D--LY-K.K--E-----V-N---F---W--------I-F--  138
68_01B.CN.11.ANHUI_WH73       ---------IV---Q----IKQ.......TP.................A-TEG---T-Q--D---------I..N--DNV--R-.----EG-E.G----R---------F-G-F---F---------D--VY-K.K--E-----V-N---F---W--------I-F--  138
70_BF1.BR.10.10BR_PE004       ---------IV---EI---L--.......NP.....................---VG--------QR--L-IN---N-D---F----..H--E.-.-------R-----------G-L------------..........................................   97
71_BF1.BR.10.10BR_PE008       ----------V----I---I--.......TP.....................----G---V-----R--------..-N--D-V---------.-.-.-.-------R-----------G-F-----------------W--.K---------V---------W--------T-W--  135
72_BF1.BR.10.10BR_MG002       ---------IV---QI-------.......TP.TAAERRERIRQ..TPTAA--EG---V---------------..FN--D-V-------..-K-DE-.-------R-----------G-L------R---------Y--.K---------V---------W--------I----  152
O.BE.87.ANT70                 --NALR-GKFE--AA-------TR.....TF.PES............E-C-PG--QI--E-AAR-G-P--H-PQN---L-F--S..HQ-.-.-.-------A-----------G-F---F-------------Y-H.K-AE-----V-N---F---W--------T-F--  143
O.CM.98.98CMA104              --NVLG--KFA--S---------.......-PXPEP............E-C-PG--XL--E-AAR-G-P--Y-PQN---L-F--S..HQD..-D.------R----V----F-G-F---F---------D---Y-D.A-AE-----VHN---FX--W--------T-F--  143
O.CM.98.98CMABB141            --NALR--KFE---A------K.......TS.PEP............KDC-PG--QI-KE-AAR-G-P--Y-PQN---L-F--S..HQDE.-.-------R----V----F---F---F---------D---Y-H.K-AE-----VHN---F---W--------T-F--  142
O.CM.98.98CMABB212            --NA----KFA--SE--N---Q.......SS.PDP............Q-C-PG---V--E-ADR-G-P--Y-PQN---L-F-T-..H-D.-..D-----R---------F-G-F---F-------------Y-H.K-AE-----VH----F---W-C-------I----  141
O.CM.98.98CMU5337             --NALR-HRDQR--A------N.......-P.S.............A-C-PG--QV-QE-AAR-G-STAH-PQN---L-F--S..HQD.-..D-----R---------F-G-F---F-------------Y--.A-AE-----V-N---F---W--------T-F--  139
O.CM.99.99CMU4122             --NALR--RFP--SAI-----K.......TS.PEP............E-C-XG--PV--E-AAR-G-P--Y-PQN---L-F--S..HSD.-.-.-------R----V----F---F---V----N-------Y-H.G-AE-------N---F---W--------T-F--  141
O.FR.92.VAU                   --NVLG--KWA--SE--R---Q.......TS.PDP........QSSDTQ.P-PG-R-V--A-ADRRG-P--Y-PQN-Q-L-F-GS..HQG..K..GI---R-----------G-F---F---------D---Y-H.K-AE-----VH----F---W--------T-F--  145
O.GA.11.11Gab6352             --NVLG-DKFG---A-------TR.....TS.PES............E-P-AG---V--AIAEK-G-PN-Y-PQN-E-L-F-DS..HQ-.-.-.-------R---------F-G-H---V-------------Y-H.A-AE-----V-----F---W-H-------T-F--  143
O.SN.99.99SE_MP1299           --NVLG-DIFK--SA------G.......TS.PDP............E-C-PG--QI--E-AAR-G-P--Y-PQN---L-F--S..HQD.-.-.-------R-----------G-F---F---------D---Y-H.K-AE-----V-----F---W-G-------P-F--  141

N.CM.02.DJ00131               --KI-----IV---EI--------....QRP.H.EPA............VEP-VG-----Q--ANR--L-T---RTN-PTV--V--.-.-.-.-G.------R-----------------F-------------W-R.K---V----V-----F---W--------T-F--  144
N.CM.02.SJGddd                --KA-----IV---A---------..ARQ..TQ.EXA..........AGAE--VG---V-Q--ANR-GV-T---KDN-ETV-------.-.-.-.-------C-------------F---F-------------W-R.K--E-----V-----F---W.............  133
N.CM.04.04CM_1015_04          --KS------V---EI--------...Q..RQ.TQA............AA--VG-----Q--ANR----T---RDN-ETV--V--.----E.-.-------R---------F---F---F-----------VW-R.K--E-------N-------W-------------  143
N.FR.11.N1_FR_2011            --KI-----IV---EI--------...Q..TP.GAA............VEP-AG-----Q--ANR----IN--KDN-QSV-------..H-D.-.-------R---------F-MGF---F---D-------VW-R.K---------V-----I---W--------T----  142

P.CM.06.U14788                --NA-K---LV---A---KIKQ.......TT.PTTPDPTTP.......VTP-PG--EI-KE-AQGKG-P-KFSSKN---L-F-D-...H--.-.-.-------R------C------F---F---------D--VY-P.E-AE-----------F---W--------E----  147
P.FR.09.RBF168                --NA-K---LV---A-----NQ.......TA...P............ETP-SG--EI-KE-AQGKG-P-KYNSKN---L-F-----..H-D.-.-.-------R----V--------F---F-------------Y-P.G-AE-----V-----F---W-H-------E-F--  139

CPZ.CD.06.BF1167              --AVLG....TRF-GIDNMLN-.......IR.................NTPAADAPG-Q--HNR-GL-TNTIGTEHDVLQHSGD...HT-.-...A-----R----M----E-L-----W-----------YW-P.K-AA---T-MFN---I---W-------------  133
CPZ.CM.05.SIVcpzMT145         --A------LV---EI---IK-.......-P.................T---G--EV-Q-IASR--V-IR--PQ--ETL----E...MQD.-.-.-------K---------F-G-F----------------W-R.K--E-----V---------W-------------  137
CPZ.GA.88.GAB1                --T------LV---E--R-I-E.......-P.................T--EG--EV-K---R------R--PE--QTL----E...MDN.-.-.-------R----T--------F-----------------VY-R.---E-----V-----F---W----T---T-F--  137
CPZ.TZ.06.SIVcpzTAN13         --NI.....FGR-AGAT-VL--.......LH.................A-PAHD-Q--KE-HER-GL-NNTIGTEKDVAHYSQD...HT-.-.-.-------R-A--M----E-L-T---W---------D---Y-T.--AA---T-M-N---V---W-------------  132
CPZ.US.85.US_Marilyn          --N------IV---E--N-L-Q.......TQ.T...............TA--EG--PV-Q--AE-----TR--PQN-QTL---DEMTNH-...S.-------R-----------Q-F--GF-----------VY-R.---E-----V-----I---W-------------  141

GOR.CM.07.SIVgor2139_287      --NT-K---F----A------Q.......TN.PEP............K.P-PG--EV--A-AQR-G-PAKY-PQN-Q-L-F----..H-D.-.-.-------R---------F-G-F---F--N--------VF--.A-S------V---------W--------I----  140
GOR.CM.07.SIVgorCP2135con     --NT-K---F----A------Q.......TS.PEP............K.P-PG--EV--V-AQR-G-PAKY-PRN-Q-L-F----..H-D.-.-.-------X---------F-G-F---F-------------VF--.A-SE-----V---------W--------I----  140
```

normal Nef end

```
B.FR.83.HXB2          TFGWCYKLVPVEPDKI.EE.ANKGENTS.LLHPVSLHGMDD..PEREVLEWRFDSRLAFHHVARE.LHPEYF.K............NC*  205
A1.AU.03.PS1044_Day0  -----F-----D--EV.-K.-TEA--N-.-----ICQ------..DD--T-M-K---K--LK-----.----FY.............D-.  204
A1.CH.03.HIV_CH_BID_V3538 -----F-----DQTEV.---.-TV---N-.-----ICQ------..E-----M-K---H--R-F-K-.-----FY.............D--  208
A1.ES.05.X1608_8      -----F-----D--EV.-K.--E--DNC.----M-Q--V---..E---T-M-K------LK-R---.I----Y.............D--  207
A1.KE.11.DEMA11KE001  C----F-----D-EEV.-K.--E---N-.-----MCQ------..D-K-T-V---------K-T---.M----Y.E...........---  205
A1.RU.11.11RU6950     ------------D-ADV.---.-TE---N-.-----ICQ------..E-K---K-Q------LT-R---.M----Y.............D--  210
A1.RW.11.DEMA111RW002 -----F-----N-EEV.-R.-TE---N-.-----ICQ------..E-G-T-M--------LR-----.E----FY.............D--  210
A1.SN.01.DDI579       -----F-----DQASV.---.-TE---.-.------CQ------..K-K---M--------LK-T----.M---FY.............D--  204
A1.UG.11.DEMA110UG001 -----F-----D-EEV.-K.-TE---N-.-----MCQ------..D---T-M-K------Q-T---.M---FY.............D--  207
A1.ZA.04.04ZASK162B1  ------------QAEV.-Q.DTQ---N-.-----ICQ--I---..Q-K---M-K------LK-L---.M---FY.............D--  210
A2.CM.01.01CM_1445MV  -----F-----N-SEV.---.-TE---N-.-----ICQ--I---..------K-----H--LR-R----.-Y-DFY............D--  204
A2.CY.94.94CY017_41   -----F-------SEV.---.-TQ---N-.-----ICQ--V---..------R-E--RS--RR-R---.------Y.............D--  218

B.BR.10.10BR_MG029    -----------D---V.---.--E---NC.----MG----E-..------M-K------------.M---Y.-TQELLTASNQ#.---  217
B.CA.07.502_1191_03   -----F-------EQV.---.-TV--TNN.----G-----E-..--K---V-----K-------K-.-K--F-.RDR............-  207
B.CH.08.M2_0803101_NFLG8 C----F-----D---V.---.--T---IN.----M-Q---E-..--K---V-K--------M---.-----FY.............D--  214
B.CN.10.DEMB10CN002   C----F----L--EE-.-K.-TE--D-C.------Q---E-..A-----V-K---------M---.-------.R............D--  210
B.ES.10.DEMB10ES002   C----F-----DREQV.---.--E---N-.-----M-----E-..--K---Q-K-------------.------.............D--  212
B.FR.11.DEMB11FR001   -----F-----DKE--.---.--Q---NC.----M-Q-----..------V-K---L--VRYM---.K---FY.............D--  224
B.GB.05.MM45d213_GN1  -----F-------E-V.---.--A----R.----I-Q------..------Q-K---H---X-----#-----Y.............N-  209
B.HK.06.HK003         -----F----L--EEV.-K.--E---NC.----IN------..--G---V-----H--LR-M---.------Y.............---  211
B.JP.12.DEMB12JP001   -----F-----D--QV.-K.D-E---IN.----M-Q------..------*M-K---H--LQ-----.K----Y.............D--  208
B.KR.07.HP_18_07JHS10_3909 -----F-------E-V.---.-TV---NC.----MN------..--G---V-K-------------.K----Y.............D--  207
B.PE.07.502_0525_wg5  C----F-----DQEQV.-K.D-ER-DKC.----M-Q------..--K---I-K---S-----K---.M---FY.-D...........-  206
B.RU.11.11RU21n       -----F-------EE-.-K.-TT---NC.----ANM-----..------V-K------Y------.K---FY.............D--  207
B.TH.08.MERLBDTRC10   -----F------EEV.-K.--E---NC.----M-Q---E-..------K-K---S--Y--R----.----FY.............D--  208
B.US.11.ES38          -----F-----D---V.---.--T---NC.----M-Q------..T-K---V-K---H---N-----.-----Y.............---  212

C.AR.01.ARG4006       -----F-----D-REV.---.--T---N-.-----M-----E-..SDG---Q-K---L--RR-M----.------Y.............D--  216
C.BR.07.DEMC07BR003   -----F-----D-REV.---.--T---N-.-----M-----E-..T-G---Q-K--RL--RR-M---.------Y.............D--  208
C.BW.00.00BW5031_1    ------------D-KEV.---.-TE---NC.----ICQ---E-..ED----R-K---E--RR-I---.R---FY.............D--  208
C.CN.10.YNFL19        -----F-----D-KEV.---.--E---NC.------CQ---E-..EH----K-Q--IQ--HR-M---.----FY.............D--  206
C.ES.08.X2363_2       ----PF-----D-KEV.---.D----AC.------A-------..AD--I-K-Q---S-VRI-K---.------Y.............D--  217
C.ET.02.02ET_288      -----F-----D-SEV.---.--E---NC.----I--------..EDK---M-K--IS--HR-L-----.------Y.............D*-  207
C.IN.09.T125_2139     -----F--I--D-KEV.---.--E---NC.------CQ---E-..EH----M-K---Q--RR-L---.Q---FY.............D--  208
C.KE.00.KER2010       -----F-----D-REV.---.---E--NC.----M-Q--IE-..E-----R-Q---L--HR-L---.Q---Y.............D--  208
C.MW.09.703010256_CH256.w96 -----F-----D-REV.---.--E---NC.----M-Q------..DH----K-K---Q--HK-M---.----FY.............D--  208
C.TZ.08.707010457_CH457.w8 -L---------D-KEV.---.ETE---N-.-----M-----E-..EH----K-K---M--RR-M---.K----Y.............D--  208
C.YE.02.02YE511       -----F-----D-REV.---.--E---N-.-----I-----E-..GDG---M-K------HR-----.----FY.............D--  220
C.ZA.08.705010534_CH534.w12 -----F-----D-REV.---.--E---NC.----M-Q---E-..ED----M-K--IS--HR-M---.I----Y.............D--  215
C.ZA.10.DEMC10ZA001   ------------D-REV.---.-----DNY.----I-----E-..E-G---K-Q---S--RR-L---.------Y.............D--  216
C.ZM.11.DEMC11ZM006   ----PF-----D-QEA.---.H--RDDSC.-M--M-Q--IE--.EH--I-M-K--TQ--RR-L---.----WY.............D--  206

D.CM.10.DEMD10CM009   -----FE----D-RVV.-G.ETE----C.------Q---E-..T-----V-K-N-----E-K--Q.K---F-.------------D--  211
D.KE.11.DEMD11KE003   -----FE---L--EEV.-K.-TE---NC.------Q------..Q-----I---N-----E-K--Q.------Y.------------  207
D.KR.04.04KBH8        -----FE-----SEEV.---.--E---NC.----M-Q--IE-..---Q--K---N-K---E-K--M.----FY.------------D--  203
D.SN.90.SE365         C----FQ----D-EEV.---.-TE--DNC.----LCQ--IE--.E--Q--V---N-K---E-K--K.-----FY.------------D--  208
D.TZ.01.A280          -----FE----D-KEV.---.TTE---SC.----I-Q---E-..T-K---V---N-----E-K--S.------Y.------------D--  208
D.UG.10.DEMD10UG004   -----FE----D-KEV.---.DTE---NC.----MCQ---E-..Q-----K---N-----E-K-KI.Q---Y.------------D--  207
D.UG.11.DEMD11UG003   -----F----IDAKEA.---.DTE---NC.----ACQ---E-..T-----M---N-----E-K--V.---DFY.------------D--  208
D.YE.02.02YE516       -----FE----D-KMV.---.E-E--DKC.----MHQ---G-..--K---V---N-K---E-K--K.T---F-.------------D--  205
D.ZA.90.R1            -----F-----D-QEV.-Q.-TE--TNC.----MNQ---E-..------K-K------E-K---.----Y.------------D--  208

F1.ES.02.ES_X845_4    -L----------D-EEV.-K.--E---NC.----M-Q------..ADG---K-K---T--LR-I----.RF---Y.-D...........-  206
F1.RO.96.BCI_R07      --E--F-----D-EEV.-K.--E---NC.----M-Q---E-..ED----K-Q------K-E----.R---FY.QD...........-  212
F1.RU.08.D88_845      -----F-----D-EEV.-R.--E----C.----M-Q------..EDG---R-K------LR-L----.K---FY.-D...........-  215
F2.CM.02.02CM_0016BBY ------------------------------------------------------------------------------------  132
F2.CM.10.DEMF210CM001 -----F-----S-EEV.-K.--E---NC.----M-Q---E-..EDG---M-K---G--RR-I----.K---FY.-D...........-  212
F2.CM.10.DEMF210CM007 -----F-----D-EEV.-K.--E--TNC.----I-Q--IE--.ED----M-K---T--MR-I-K-.K---FY.-D...........-  206

G.BE.96.DRCBL         -----F----M--SEV.---.------N-.----ICQ---E-..EDG---V-----S--RR-L----.-----Y.............D--  208
G.CM.10.DEMG10CM008   ------------MD-AE-.---.-T----NC.----ICQ---E-..ED----V-----S--RR-I----.K---FY.------------D--  213
G.CN.08.GX_2084_08    -----F----MD-AEV.---.--E--N-.----ICQ--Q---..ED----V-----S--RR-R----.----FY.------------D--  207
G.CU.99.Cu74          -----F----LD-AE-.---.-T----N-.----ICQ---E-..DD----I-----S--RR-L----.----DFY------------D--  208
G.ES.09.X2634_2       -----F--E-MD-AEV.---.-TR---N-.----IC---AE--.GDK---R-K---S--RR-L----.-----FY.------------D--  209
G.GH.03.03GH175G      -----F----ID-AEV.---.-----TN-.----ICQ--IE--.EDK---V-K---S--RR-L----.------Y.------------  208
G.KE.09.DEMG09KE001   -----F-----D-AE-.---.-TE--RNT.----ICQ---E-..SD----K-K---S--RR-I----.-Y---Y.------------  219
G.NG.09.09NG_SC62     ----PF----MN-AE-.---.------N-.----I-Q---E-..DDK---V-K---S--RR-I----.---D-Y.----------*-  203

H.BE.93.VI991         -----F-----D-QDV.-K.--E---N-.----MCQ--IE-..------M-K------LR-R-K-.-----FY.------------D--  208
H.CF.90.056           -----F-----N-QEV.-Q.--E---N-.----M-----E-..DG----M-K------LT-L--V.K---.Y.------------D--  207
H.GB.00.00GBAC4001    C----F-----D-QEV.-K.--E---N-.----MCQ---E-..E-K---V-----H---R-R---.Q---FY.------------D--  208

J.CM.04.04CMU11421    -----------D-REV.---.--E---NC.----X-Q--IX--.X-X---M-K--XS--RR-X---.----FY.------------D--  210
J.SE.93.SE9280_7887   -----------D-SEV.---.--E---NC.----ICQ--IE--.E-----Q-K---S--RR-I---.----FY.------------D--  204
J.SE.94.SE9173_7022   -----F-----D-SEV.---.--E---NC.----ACQ--IE--.E-----K-K---S--RR-I---.----FY.------------D--  205

K.CD.97.97ZR_EQTB11   -----------D-REV.---.-TE---NC.-----NQ---E-..EH----K-K---S--RK-----.M----Y.------------D--  207
K.CM.96.96CM_MP535    -----------D-AEV.---.TTE--DNC.----INQ---E-..EH--I-M-K---S--RR-----.---D-Y.------------D--  216

U.CA.01.TV749         ------------D-XEV.---.-TE---N-.----MCQ--AE--.S-----R-K---K--LK-I-K-.---DFY------------D--  210
U.CA.99.TV721         -----F-----D--EV.---.-TE---N-.----ICQ--V---..--K---K-Q---Q--RR-I---.-Y-DFY------------D--  207
U.CD.83.83CD003_Z3    ----PF-----D-KEV.---.--E---NC.----ICQ------..E-K---M-K---S--RR-L----.----FY.------------D--  208
U.CD.90.90CD121E12    -L---F-----D-VEV.---.--E---N-.----ICQ---E-..E-K---M-K---H--RS-K----.------.------------D--  208
U.ES.10.DEURF10DZ001  -----------D-REV.---.ET---NC.----LCQ--AE--.E-----R-K---S--RR-L----.K---FY.------------D--  212
U.GR.99.99GR303       -----------D-REV#--.---.--E---NC#----I-Q---E--.T-----I-K------LR-R--G.E----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

01_AE.AF.07.569M      C----F-----D-STV.---.D-----NC.----M-Q--IE--.E-----I-K---A--RK-I----.-----Y.------------D--  225
01_AE.CF.90.90CF11697 C----------D-KEV.---.D--D-SNC.----M-Q--V---.E-----M-K---S--RR-I---.-R---Y.------------D--  205
01_AE.CN.10.YNFL03    C----F-----D-RAV.---.ET-E-.SC.----INQ--V---.ED----M-K---A--RK-I---.Q---FY.---------------  201
01_AE.HK.04.HK001     C----F-----D-REV.---.D-----NC.----A-Q--I---.E-----M-K---A--RK-L----.-----FY.------------D--  207
01_AE.IR.10.10IR.THR48F C----F-----DRQEV.---.------NC.----M-Q--QE--.E-----R-K---X--RK-----.------Y.------------D--  205
01_AE.JP.x.JRC77AE    C----F-----D-TEV.---.D----DNC.----M-Q--I---.ED----M-K---A--RK-I----.------Y.------------D--  208
01_AE.TH.09.AA090a_WG11 C----F-----D-REV.---.DT----NC.----M-Q--V---.E-----M-K---A--RK-I----.-R---Y.------------D--  208
01_AE.TH.90.CM240     C----F-----DQREV.---.D-----NC.----M-Q--IE--.E-----M-K---A--RK------.Q----Y.------------D--  205
01_AE.US.05.306163_FL C----F-----D-REV.---.------NC.----L-Q--IE--.E-----M-K---A--RK----D.-----Y.------------DX.  199
01_AE.VN.98.98VNND15  C----F-----D-REV.---.DT----NC.------Q--LE--.G-----M-K---A--RK-I----.---D-Y.------------D--  207
```

```
                                                                                                       normal Nef end
B.FR.83.HXB2                   TFGWCYKLVPVEPDKI.EE.ANKGENTS.LLHPVSLHGMDD..PEREVLEWRFDSRLAFHHVARE.LHPEYF.K............NC*   205
02_AG.CM.10.DE00210CM013       -----F----LD-VE-.-R.--E---N-.----ICQ-----..ED----V-----T---K-K---.----FY.-D.............-   216
02_AG.ES.06.P1423              ----------M--SEV.-K.--E---NC.----ICQ-----..ED----V-----L---K-T---.-----Y.-............D--   207
02_AG.GW.05.CC_0048            -----F----MD-TEV.--.L--E--N-.------Q---E..ED----V-----X---X..........................   192
02_AG.KR.12.12MHR9             -----F----LD-AEV.--.---E--N-.-----CQ---E-..ED----V---G-----T-K-K-.----FY.-............D--   221
02_AG.LR.x.POC44951            -----F----MD-AEV.--.------N-.----ICQ--IE-..TD-Q--V-----S---K-R-Q-.----FY.-............D--   210
02_AG.NG.09.09NG_SC61          -----F----MD-AD-.--.------N-.----ICQ---E-..E-----V-----S--RR-I---.-----Y.-............D--   207
02_AG.NG.x.IBNG                -----F----MD-AEV.--.--E---N-.----ICQ---E-..DD----I---------R-T---.-----Y.-............D--   207
02_AG.SE.94.SE7812             -----F----MD-AD-.-K.DTE---N-.----ICQ---E-..ED----V---------T-K---.M---FY.-............D--   207
02_AG.SN.98.98SE_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.94.94CY032_3         C----F-----D-QEV.--.-TE----C.----I-Q---E-..E-----K-K------YK-----.----FY.-............D--   220
05_DF.BE.x.VI1310              -L---F-----N-EEV.-K.--E--DNC.----M-----E-..DD----Q-K---S--LR-I---.R---FY.QD.............-   207
06_cpx.AU.96.BFP90             -----------D-EEV.---.LT---NC.----ICQ--AE-..E-----K-K---S--RR-I---.K---FY.-............D--   206
07_BC.CN.98.98CN009            -----F-----D-REV.--.--E---NC.-----CQ---E-..EH----K-K---Q--HR-R---.----FY.-............D--   207
08_BC.CN.97.97CNGX_6F          -----F-----D-REV.--.--E--DNC.-----CQ---E-..EH----K-K---Q--HR-R---.----FY.-............D--   207
09_cpx.GH.96.96GH2911          -----F-----D-KEV.--.-TG---N-.----MCQ--V.-..SGG---M-----T--LK-I---.K---FY.-............D--   207
10_CD.TZ.96.96TZ_BF061         -----F-----D-REV.--.------NC.----M-----E-..KHG---V-K---T--HK-I---.-----Y.-............D--   208
11_cpx.CM.95.95CM_1816         C----F-----D-REV.--.--A---NC.----L-Q--T--..E-----M-K---S--RR-I---.---DFY.-D.............-   209
12_BF.AR.99.ARMA159            -----F-----D-VEV.-K.--E---NC.----M-Q---E-..ED----I-K------LR-L---.K---WY.-D.............-   206
13_cpx.CM.96.96CM_1849         C-E--F----ID-KEV.--.--E---NC.----M-Q--IE-..A-----M-K---S--RK-I---.----FY.-............D--   200
14_BG.ES.05.X1870              -----F--E--D-AEV.--.-TG---N-.----ICQ--IE-..ADK---R-K---G--RR-I---.-R-DFY.-............D--   208
15_01B.TH.99.99TH_MU2079       C----F-----D-SDV.---.N-----NC.----M-Q---E-..E-----K-Q---A--RK-I---.-----Y.-............D--   201
16_A2D.KR.97.97KR004           -----F-----S-AEV.--.-TE--TN-.----ICQ--IE-..------K-V---H--LV-K---.----FY.-............D--   217
17_BF.AR.99.ARMA038            -L----------EE-.-K.--E---NC.----M-Q-E---..ED----K-K---S---R-----.R---FY.QN.............-   204
18_cpx.CU.99.CU76              -----F----DSNEV.--.--E---N-.----ICQ-----..E-----K-K---A--RR-I---.Q---FY.-............D--   207
19_cpx.CU.99.CU7               -----FE----D-KEA.-K.DTE---NC.----ACQ-----..--K---V---N-----E-K--I.K---FY.-............D--   207
20_BG.CU.99.Cu103              -----F----MD-AEV.--.--N---N-.----ICQ---E-..N-----I-----S--RR-L---.----FY.-............D--   204
21_A2D.KE.99.KER2003           -----F-----D-SEV.---.DT----N-.----MCQQ-I--..------R-K--PH--LK-R-Q-.K---FY.-............D--   208
22_01A1.CM.01.01CM_0001BBY     ........................................................................................   130
23_BG.CU.03.CB118              -----F-----D-ADV.--.-SN---N-.----ICQ---E-..DDK---M-----S--RR-I---.R--DFY.-............D--   209
24_BG.ES.08.X2456_2            -----F-----D-TEV.--.--S--TN-.----ICQ---E-..SDK---I-K---S--RR-I---.---DFY.-............D--   224
25_cpx.CM.02.1918LE            ........................................................................................   137
26_AU.CD.02.02CD_MBTB047       -----F-----D-AEV.--.------N-.---AM-----E-..S-G---M-K---Q--RK-T---.-----Y.-............D--   208
27_cpx.FR.04.04CD_FR_KZS       -L---F-----D-KEV.--.--E---N-.----M-----E-..E-----M-K------LR-----.K---FY.-............D--   208
28_BF.BR.99.BREPMI2609         -----F-----D-EQV.-A.--E---N-.----M-Q-----..--K---M-K---------M---.-----Y.-D.............-   210
29_BF.BR.01.BREPM16704         -----------D--QV.---.E----N-.----M-Q-----..------M-K---------M---.----F-.-............D--   212
31_BC.BR.04.04BR142            -----F-----D-REV.--.--A---N-.----M-----E-..THG---M-K---L--RR-M---.----FY.-............D--   223
32_06A1.EE.01.EE0369           -----------D-REV.--.-T----NC.----IC---AE-..E-G---M-----S--RR-I---.K---FY.-............D--   207
33_01B.ID.07.JKT189_C          C----F-----D-TEV.---.E-----NC.----M-Q---E-..ED----K-Q---S--RR-I---.Q----Y.-............D--   205
34_01B.TH.99.OUR2478P          C----F-----D-GEV.K-......--SC.----M-Q--IE-..E-----M-K---A--RK-I---.----FY.-............D--   203
35_AD.AF.07.169H               -----------D-AEV.-K.D-E---N-.----MCQ-----..E---T-M-K-------T-K---.----FY.-............D--   206
36_cpx.CM.00.00CMNYU830        ........................................................................................   131
37_cpx.CM.00.00CMNYU926        ........................................................................................   132
38_BF1.UY.03.UY03_3389         -M---F-----D--EV.-K.--E---NC.----M-Q--A--..EDK---V--------LR-I---.R---FY.QD.............-   206
39_BF.BR.04.04BRRJ179          -----F-----S-EEV.-R.--E--DNC.----M-Q-----..ED----V-K---K-------------.-----Y.-............D--   216
40_BF.BR.05.05BRRJ055          -----F-----D-EEV.-K.--E---NC.----M-----E-..E-K---Q----AT---------.----FH.-............---   213
42_BF.LU.06.luBF_18_06         C----F----L-Q-QV.-K.-TE-----.----I-----E-..T-----V-K---K-------H-.----FY.-............D--   207
43_02G.SA.03.J11223            -----F----MD-ADV.--.---E--N-.----ICQ---E-..ED----V-----S--RR-----.-----Y.-............D--   215
44_BF.CL.00.CH80               -L---F-----D-EEV.-K.--E----G.----M-Q-----..ADK---M-K---S--LR-I---.RF--FYQDD.............-   207
45_cpx.FR.04.04FR_AUK          -----F-----D-KEV.--.-TE----C.----I-Q--IE-..T-----I-K---H--QR-L---.K---FY.-............D--   208
46_BF.BR.07.07BR_FPS625        -----F-----N-EEV.-R.--E---NC.----I-Q---E-..E-----Q-K---S-----R---.---DFY.-NCCI......EL---   213
47_BF.ES.08.P1942              -----F-----D-EEV.-K.--E---NC.----M-Q---E-..ED----M-K--RH--SK---------.-----Y.-............D--   207
48_01B.MY.07.07MYKT021         C----F-----D-QEV.---.D-----SC.----M-Q---E-..E-----M-K---S--RR-I---.-R---Y.-............D--   206
49_cpx.GM.03.N26677            ........................................................................................    99
50_A1D.GB.10.12792             ------------QEEV.-K.-TE-G-NC.----ICQ-----..E-K---Q-K---S--RK-I---.----FY.-............D--   205
51_01B.SG.11.11SG_HM021        -----FX----D--QV.-K.--EK----.----M-Q--N--..--K---M-K------LR-M-K-.K--D-Y.-............D--   211
52_01B.MY.03.03MYKL018_1       C----F-----D-KEV.---.D-----SC.----I-Q--QE-..E-G---M-K---A--RT-T---.----FY.-............D--   205
53_01B.MY.11.11FIR164          C----F-----D-AEV.--.---E--NC.----M-Q---E-..D-----I-K---S--RR-I---.-----Y.-............D--   206
54_01B.MY.09.09MYSB023         C----F-----D-REV.--.------NC.----M-Q--AE-..E-K---M-K---A--RT-----.M----.-............D--   213
55_01B.CN.10.HNCS102056        C----F-----D-REV.---.TT-E--NC.----M-Q-----..E-K---M-K---A--RK-R---.----FY.-............D--   206
56_cpx.FR.10.URF5_patient_A    ........................................................................................    84
57_BC.CN.09.09YNLX19sg         -----F-----D-RE-.--.--E---NC.-----CQ---E-..EH----R-K---Q--HR-R---.----FY.R............D--   208
58_01B.MY.09.09MYPR37          C----F-----D-REV.--.------NC.----M-Q--IE-..N-----M-K---A--RK-I---.K----Y.-............D--   206
59_01B.CN.09.09LNA423          ........................................................................................   129
60_BC.IT.11.BAV499             C----F------TKD......IEE-SSC.----MQQ-RVE-..------M-K---S-----M---.K----.-............D--   202
61_BC.CN.10.JL100010           -----F-------KEV.--.-TE---NC.-----CQ---E-..EH----M-K---H---S-----.K---WY.-............D--   207
62_BC.CN.10.YNFL13             -----F-----D-REV.--.--E---NC.----M-----E-..EHG---R-K---Q--HK-R---.----FY.-............D--   208
63_02A1.RU.10.10RU6637         -----F----MD-AE-.---.-TE---N-.----ICQ--LE-..ED----V-----K---D-----.R---FY.-............D--   208
64_BC.CN.09.YNFL31             -----F-----D-REV.--.--E---NC.-----CQ---E-..EH----M-K---Q--HR-R---.----FY.-............D--   208
65_cpx.CN.10.YNFL01            -----F-----D-REV.---.S-E---NC.----M-Q-----..EHG---V-K---M--HR-M---.Q---FY.-............D--   212
67_01B.CN.11.ANHUI_HF115       C----F-----D-REV.---.D-----KC.----M-Q---E-..E-----M-K---A--RK-I---.---X-Y.-............D--   208
68_01B.CN.11.ANHUI_WH73        C----F-----D-RDV.---.D--R-DSC.----I-Q--I--..E-----M-K---A--RK-I---.---D-Y.-............D--   208
70_BF1.BR.10.10BR_PE004        ........................................................................................    97
71_BF1.BR.10.10BR_PE008        -----F-----D-E-V.--.---K--S-.----M-----E-..A-K---V-K---S-G-----------.-----Y.-............---   205
72_BF1.BR.10.10BR_MG002        -----F----L--EEV.-K.-SE#--NC.----AN------P.G-G---V-----------M---.-----Y.-N.............-   221
O.BE.87.ANT70                  ----LF-----SEEEA.-RLG-TC-RAN.----ACA--FE-..THK-I-M-K--RS-GNT---MI.T---L-Q-D.............-   214
O.CM.98.98CMA104               ----MF-----SEAEA.--LG--H-RAK.----ACA--FG-..AHG-I-K-Q--RS-GLT---KI.T---L-P-D.............-   214
O.CM.98.98CMABB141             -Y--L------TREEA.-RLG-EC-RA-.----ACN--FE-..-HG-I-K-Q--RS-GNT---KI.T---X-X-D.............-   213
O.CM.98.98CMABB212             ----L------TAEEA.-RLG-TR-DA-.----ACN--AE-..AHG-I-K-Q--RT-GLT---LQ.K---L-.PK.............-   211
O.CM.98.98CMU5337              ----MF-----SEAEA.DXLG--H-RAK.----ACN--FE-..-HG-I-K-Q--RS-GGT---MI.T--HL-N-D.............-   210
O.CM.99.99CMU4122              ----LF-----SEAEA.--LG--CVTAK.----ACN--SE-..-HG-I-K-Q--RS-GNI---KV.T---L-L-D.............-   212
O.FR.92.VAU                    ----LF-----SEAEA.-ALG--C-RA-.----ACN--YE-..QHK-I-K-Q--RS--NT---LI.T---L-S-D.............-   216
O.GA.11.11Gab6352              ----LF-----TE-EA.KRLG-ECVRAH.----ACA--FE-..-HG-I-K-K---S-GLT--SKV.I---L-T-D.............-   214
O.SN.99.99SE_MP1299            ----LF-----SEAEA.--LG--C-RA-.----ACN--FE-..NHGQI-K-Q--RS-GST---MV.TN--L-N-D.............-   212
N.CM.02.DJO0131                -----F----LSAEAV.--.--E-D-NA.----LCQ--V--..GHK---Q-Q---S--RR-I---.---DFY.-............---   214
N.CM.02.SJGddd                 ........................................................................................   133
N.CM.04.04CM_1015_04           -----F----LSAEAV.--.--E-D-NA.----ICQ--V--..DHK---V-----S--RR-I---.---DFY.-............---   213
N.FR.11.N1_FR_2011             -----F----LSAEAV.--.-TE---N-.----ICQ--V--..DHK---V-----S--RR---K-.----F-.-............D--   212
P.CM.06.U14788                 ----LF-----SEVEA.--MGD-Q-KAK.----ACTY-FS-..-HK-I-V-K---S-GRE---LQ.K---L-I-D.............-   218
P.FR.09.RBF168                 ----LF-----SEVEV.--MGDEQ-KAK.----ACTY-DS-..HH--I-M-K--RS-GER---LQ.K---L-T-D.............-   210
CPZ.CD.06.BF1167               CR--LF----AD-PE.......ED-RNL.-M--ACSY-R--..-NG-L-V-K---E--RR-I---.R-----.............-   196
CPZ.CM.05.SIVcpzMT145          -----F----LP-EEV.-K.--E---NC.----MCQ---E-..E-----V-T---K--LR-L---.K----Y.R............D.-   206
CPZ.GA.88.GAB1                 C----F----LTEEQV.-Q.--E-D-NC.----ICQ---E-..EDK---V--------LR-I---.Q----Y.-............D.-   206
CPZ.TZ.06.SIVcpzTAN13          CR--LF-----T-PE.......DD-RNI.----ANS--RE-PDGDG-T-M-Q--TS--RR-I---.R-----.-.............-   197
CPZ.US.85.US_Marilyn           -Y---F----LTEEEV.-Q.-----TNI.----MCQ---E-..EHG---I-Q--TE--RR-R-K-.-------.R............-.-   210
GOR.CM.07.SIVgor2139_287       ----LF-----SEAEA.--MGTSN-KAM.----ACS--R--..-YG-I-V----RA-GTT---LQ.K---L-KRD.............-   211
GOR.CM.07.SIVgorCP2135con      ----LF-----SEAEA.--MGTAS-RAI.----ACS--R--..-HG-I-V----RA-GTT---LQ.K---L-KRE.............-   211
```