V

HIV-1/SIVcpz Proteins

Contents

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

V-1 Introduction

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

V-2 Annotated features

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

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

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

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

V-3 Sequences

Sequences included in the HIV-1 protein alignments.

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

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

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

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

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

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

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

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

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

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

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

```
                                                                          p24 end p2 start    p2 end p7 start       Zn motif            Zn motif          p7 end p1 start    p1 end p6 start    Vpr binding
B.FR.83.HXB2                  NWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNS.....ATIMMQRGNFR.NQRKIVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCT....ERQANFLGKIWPSY....KGRPGNFLQSR...PEPT...........  456
01_AE.CF.90.90CF11697         -----------------S------TG------------------------------QHA......N--------K.G-KR.I----------L---------R------Q-------------....------------LN....-------P---...L---...........  454
01_AE.CN.05.FJ051             -----------------ST-----IG----------------S------------AQHT......N--------K.G-KR.I----------L--------------------------....-------------N....-------P---...----...........  454
01_AE.CN.06.FJ054             -----------------S------TG-S--------------S-----------S-QHT......N--------K.G-KR.I----------L---------E----------------...N-------------N....-------P---...----...........  455
01_AE.HK.x.HK001              -----------------S------AG----------------S----------N-AQQT......-V------XK.G-KR.I------R---L--------------------------....------------PN....-------P---...---S...........  454
01_AE.TH.01.01TH_R2184        -----------------S------A-----------------S------------AQHA......N-L------K.GRKT.I------R---L--------------------------....---------L---N....-------P-N-...----...........  452
01_AE.TH.02.OUR769I           -----------------S------TG----------------S-----------HAQQA......N--------K.G-KR.I----------L--------------------------....-------------N....-------P---...----...........  452
01_AE.TH.90.CM240             -----------------S------TG----------------S-----------HAQHA......---------K.G-KR.I------R---L---------Q----------------....-------------N....-------P---...----...........  454
01_AE.US.00.00US_MSC1164      -----------------S------TG-S--------------S-----------H-QHA......N--------K.G-KR.I------R---L--------------------------..ERQG-----------H....-------P---...----...........  454

02_AG.CM.02.02CM_4082STN      --------------------RG---G-----------------------------AQQA......N---------.G--T.I----------L----K----R----------------....-------------N....-------P---...----...........  449
02_AG.EC.x.ECU41              --------------------R----G-----------------------------AQQA......N---------.GP-T.I----------L-K--K----R----------------....-------------S....-------P---...----...........  417
02_AG.GH.03.GHNJ196           --------------------R----G-S--------------S------------AQQ-......NV--------.G--T.I----------L----K----R----------------....-------------N....-------P---...---S...........  457
02_AG.NG.01.PL0710            G----------------S--R----G--------------------------------QQT......NV--------.S--T.I----------L----K---------------------....-------------S....-------P---...----...........  447
02_AG.NG.x.IBNG               -----------------S--R---TG--------------------------------QQA......NV--------.G--T.I----------L----K----R----------------....-------------S....-------P---...----...........  451
02_AG.SE.94.SE7812            -----------------S--R----G----------------S-------------QQP......N---------.G--P.I----------L----K---------------------....-------------S....-------P---...----...........  451
02_AG.SN.98.MP1211            -----------S----R--------G-------S---------------------AQQ-......N---------.G--T.I----------L----K----R----------------....-------------S....-------P---...----...........  450

03_AB.RU.97.KAL153_2          --------------------R---SG--------V----------------------Q-A......N----KS---.GPKR.I----------L--------------------------....--------R----S....-------P---...---S...........  454

04_cpx.CY.94.CY032            -----------------S------TG----------------S------------AS-AAA....A----KSK-K.G--RTI----------L--------------------------....--------RM---S....---------N-...----...........  457

05_DF.BE.x.VI1310             ------------------------QQ----------------S------------A---AA....-A-------K.GP---I----------I-K-----G--------R---------....---------V---H....-----------...---S...........  458

06_cpx.AU.96.BFP90            ----D---------------R----G-----------------------------ASVG......A----KS--K.GPKRSI----------L--------------------------....-------------N....---------N-...----...........  454

08_BC.CN.97.97CNGX_6F         ----D---------------R---SG-S--------------S------------TN-T.......-L---S--K.GSKR------------I-K------------------------....---------F---H....-----------...----...........  449
09_cpx.GH.96.96GH2911         -*------------------R----G---------------------X------Q-T......N-L------K.G-KR.I---------XX--------------------------....-------------N....-------X---...---S...........  453

10_CD.TZ.96.96TZ_BF061        ------------------------------------------S------------A-GG.....N---------K.GPK-SI------------K------RR------R----L---P...ER------------N....-----------...----...........  457

11_cpx.GR.x.GR17              -----------------S--R---VG----------------S------------AQHT......N--------K.G-KR.I----------L------------*-------------....-------------S....-----------...----...........  451

12_BF.AR.99.ARMA159           ----D-------------------SG-----------------------------A-ST......AVL--KSS-N.G--R---------T--I-K--K--------------------P....-------------N....---------N-...----...........  451

13_cpx.CM.96.1849             ----D---------------R----G-----------------------------ASSAAA....-A---KS--K.GPKR-I------R---L-------------------------M....-------------N....-----------...-G--....APPESRP  465

14_BG.DE.01.9196_01           S---D---------------R----G----------------S-X----------ASGA......---L-KS--K.GP-RQI----------L--------------------------....-------------N....---------N-...----...........  454

15_01B.TH.99.99TH_MU2079      --------A--------S------SG-S--------------S-----------H-QQT......NV-------K.G-KR.I----------L----------------E---------....--------RL---#....f------P---...---S...........  452

16_A2D.KR.97.97KR004          -----------------S--R----G-S--------------S--------------Q-T..NSNIM--K----.G-K#.I----------L--------------------------....-------------H....S------P---...T---...........  454
17_BF.PE.02.PE02_PCR0155      G-------------------R---SG---------------------------------GT......-V---KS--K.G--R--------R---I-K--------------R---------....---------V---N....R--------N-...T---...........  449

18_cpx.CU.99.CU76             -----------------N--R----G----------------S-----------H-Q-T......N--------K.G-KR.I----------I--------------------------....-------------N....---------K-...----...........  454

19_cpx.CU.99.CU7              -------------------------G----------------S------------A---....H------K---K.GL--V-----------I--------------------------....-------------H....-----------...----...........  457

20_BG.CU.99.Cu103             G---D---------------R----G------------------------------SGA....G-S-L--KS--K.GPKRNI----------L--------------------------....---------L---H....-----------...----...........  457

21_A2D.KE.91.KNH1254          ----D------------S-------G----------------S------------A-S-....A-AV-------K.GP---I----------I-K--------------R---------....---------V---H....-----------...----...........  456
22_01A1.CM.01.01CM_0001BBY    -----------------S--R---TG--------------------------------QHT......G--------K.G-KR.I----------LT-------R-----------------L.TL-------------S....-------P-G-...----...........  455

23_BG.CU.03.CB118             G---D---------------R----G-S---------------------------IPGA....G-S-----S--K.GPKRNI-------D--L-----------------D--------...EG--------L---H....-------P-N-...L---...........  458

24_BG.CU.03.CB378             G---D---I-----------R----G------------------------------SGT....G-S----KS---.GP-RNI----------L--------------------------....---------L---H....-------P-N-...L---...........  457
24_BG.CU.03.CB471             G---D---------------R----G------------------------------SGT....G-S----KS--K.GTKRNI----------L----------------R---------....---------F---H....-------P-N-...L---...........  457

25_cpx.CM.01.101BA            G-------------------R----G-S----------------------------SGT....ATA--------K.GPKRNI----------L--------------------------....-------------S....---------N-...----...........  456

27_cpx.CD.97.97CDKTB49        ----D-------------------SG----------------S------------ASGAA....-A----KS--K.GP--TI----------L---------R----------------....-------------N....-----------...---S...........  457

28_BF.BR.99.BREPM12609        G---D--------------------G----------------S------------A--T......-V-------K.G---------------I----------------R---------....-------------H....---------N-...----...........  455

29_BF.BR.02.BREPM119          ----D--------------------G----------------S-----------HA--T......A-L--KS--K.G--RV-----------I-K--------------R---------....-------------H....---------N-...----...........  443

31_BC.BR.02.110PA             ----D---------------R----G-S----------------------------N-T......N-----S--K.GPKRPI----------I-----------------------E--....-------------H....----X------...----...........  453
32_06A1.EE.01.EE0369          --------------------R----G-----------------------------AQ-A......N----KS---.GPKR.I----------L----------------Q---------....-------------S....---------N-...----...........  451

33_01B.MY.05.05MYKL007_1      -----------------S------TG----------------S-----------HAQHT......G--------K.G-KR.I----------L--------------------------....---------L---H....-----------...----...........  454

34_01B.TH.99.OUR2478P         -----------------S------TG----------------S------------TQQA.....NMM-------K.G-KR.I----------L--------------------------....-K-I-----F---N....-------P---...----...........  453

35_AD.AF.05.05AF095           G----------------S--R---TG----------------S------------AQH......TN----K---K.GHKR.I-------D--L----------------R---------....-------------H....-------P---...----...........  451

36_cpx.CM.00.00CMNYU830       -----------------S------TG--------------------------------QH......-G--I--S---.G--T.I----------L--------------------------...N-------------S....-------P---...----...........  454

37_cpx.CM.00.00CMNYU926       -----------------S------TG-----------------------------T-AQQ......TS-----S--K.GPKRNI----------L--------------------------....--------R----S....-------P---...----...........  451

38_BF1.UY.05.UY05_4752        ----D----------------G--IG----------------------------------P......IV---KN--K.G--RSI----------L----------------R---------....-------------N....---------N-...----...........  451
39_BF.BR.03.03BRRJ103         ------------------------------------------S------------A--Q.....TNV--------.----T-----------V-----------------------E--....-------------H....---------T-...----...........  456

40_BF.BR.04.04BRRJ115         ----D--------------------G-----------------------------I--TA...NTAV---KS--K.G--R------------I----------------R---------....-------------H....---------N-...L---...........  460

42_BF.LU.03.luBF_05_03        G---D--------------------G-----------------------------A-G......TAVL---S--K.GP--------------I-K--------------R---------....-------------.....---------N-...---R..........P  455

43_02G.SA.03.J11223           ----D------------N--R---SG----------------S------------ASGA....A-A----K---K.GP-RNI----------L--------------------------....-------------H....-----------...----...........  454

U.CD.83.83CD003               ----D---I---------------TG----------------S------------TN.......TA--------K.GP-R------------I-K--------------R---------....-------------N....---------N-...----...........  454
U.CD.90.90CD121E12            ----D---T---------------TG-----------------------------M-.......T-V-------K.G---------------I-K--------------R---------....-------------N....---------N-...T---...........  452
U.GR.99.GR303                 ------------------------SG--------------------------------QS......PN----K----.G-KR.I----------L---------X---------------P....--------R----S....-------P---...----...........  452
U.NL.95.U_NL_95_H10986_D1     ----D----------------------S--------------------I-------A-K......TVV-------K.GTKR-I----------I--------------------------....--------RV---N....-----------...----...........  452

CPZ.CD.90.ANT                 A-------I--------H------TG-S----L---------A----------ASAN-AQ...GTAVFL----GNRGGKRPL---------------K---R----R--Q----L-N-PA.TNTGKV-----PT-TW...WGC-----V-KE.EVV---...........  484
CPZ.CM.01.SIVcpzCAM13         T---D------------Q-------G-S--------------A----------TM-QSQART...D-FF-K-PGAT.P-RKI--Y-------L----K---R-------Q--------S....G--V-----G---R....S------V-N-...T---...........  463
CPZ.CM.05.SIVcpzEK505         -----------------N--R----G-S--------------A----------T-AQTA....TSVF-------K.GI--TI----------L----K-----------Q--------R..SG---------V--LS....-------P-TT.TRK---...........  463
CPZ.CM.05.SIVcpzLB7           ----D------------S--R----G-S--------------A------------ANAA....TV---------K.GPKRVI----------I----K---R-------Q---R-Q--N....MGKV-----L---S....Q------P---...----...........  456
CPZ.CM.05.SIVcpzMB66          ----D------------S--R----G----GDEP-F------S------------AQH-........NDAKRQ-K.GPKR------------I----K---R-------Q-----RN--...N---------L---S....---------K-...----...........  452
CPZ.CM.05.SIVcpzMT145         ----D---I--------N--R----G----------------S-------------NKTE.......-F---GQNRPPN-KI----------L--------R-------Q--------N..QNNGRI-----N--PW....G-----YV-G-...K---...........  448
CPZ.GA.88.GAB1                ----D------------Q-------G----------------S----------.SMVQNQ..GR-DVFF-K-QGA.GPKRKI----------L----K---R----R--Q---------....G--V-----G---R....S------V-N-...T---...........  466
CPZ.TZ.00.TAN1                T----------------H------QG------L---------S---KI-----ASA-AGG.....VN-L-G-KRPPLKKGQLQ------V----------------R--Q---------TRNNSTGV-----RT-LW....GC-----V-N...................  482
CPZ.US.85.CPZUS               T-------------------R----G------V---------A----------C-MK-P.....SSVFL-K--AG.KPGRKI----------L--------R-------Q---------A..GN--------H-SPSWSGGSK------EN-...K---...........  463
CPZ.CM.x.SIVcpzDP943          T---D--X---------N--R----G----------------A-------------Q--.....TSVFL----NG.RPTRKI----------L--------R-------Q----I----A..GD--------G-SPW.SGGSK-------G-...K---...........  458
CPZ.CM.x.SIVcpzMB897          -----------------Q--R----G----------------S------------ANT.......-V------VK.I-KR------------I--------R-------Q----Q-N-P...DMQR---------PR....---------K-.PE-PEP...........  457

N.CM.02.DJO0131               -----------------Q------------------------A-------------QTA.....TSVFA-GR--K.GI-RTIR---------L----K---RG------Q--------K..NEG--------G-SPF....-------P-TTTTRK---...........  466
N.CM.04.04CM_1015_04          ----------------XQ-------S-S-----------------------A-AQAA.....TSVFV-----K.GT-RTI----------L----K---RX------Q--------K..NEG--------G-SPF....-------P-TA.TRK---...........  462
N.CM.04.04CM_1131_03          -----------------Q-------GSS----------------------TA-AQAA.....TSVFV-----K.GT-RTI----------L----K---RG------Q--------K..NEG--------G-SPF....-------P-TT.TRK---...........  462
N.CM.95.YBF30                 -----------------QL------G----------------A-------------QQP....TTSVFA-----K.GI--PI----------L----K---RG------Q--------K..NEG--------S-SPF....-------P-TT.TRK---...........  463
N.CM.97.YBF106                -----------------Q-------G----------------A----------A-AQTA.....TSVFV-----K.GI--TI----------L----K---RR------Q--------K..NEGX-------G-SPF....-------P-TT.TRR---...........  462

O.BE.87.ANT70                 -----------------Q---S---G-------V--------T----------ATAQQDLKGGYTAVF----QNP.IRKGTI----------I----------------Q--------R...NGK-------Y--PG....GT----YV-RP...AH-S...........  461
O.CM.91.MVP5180               -----------S-----Q-------E-------V--------T---KI-----ASAQQDLKGGYTAVF----QNP.-RKGPI----------I-K-------R------Q--------K...NG--------Y--PG....GT----YV-KQ...VS-S...........  461
O.CM.96.96CMABB637            ---------------X-Q---S---G-------V--------T----I--X--AS-NQDLKGGYTAVF----QNP.TRQRPI----------L----------------Q--------R...NGK------RY--PG....GT----YV-KQ...VP-S...........  460
O.CM.98.98CMA104              -----------------Q-------G-------V--------T----------ASAQQDLKGGYTAVF----QNP.IR-GP-----------I-------------N--QD-------R...NGK-------Y--PG....GT-----M-KQ...VA-S...........  461
O.CM.99.99CMU4122             -----------------Q---S---G-------V--------T----------ATAQQDLKGGYTAVF----QNP.GRKGPI----------L----------------Q--------R...NGK-------Y--PG....GT----YA-RQ...VS-S...........  461
O.FR.92.VAU                   -----------------Q-------G-S-----V--------T----------ASAQQDLKGGYTSVF----QKP.IRKGPI----------VTK--------------Q--------KN...R--T---K-Y--PR....GT----YV-EQ...VS-S...........  460
O.SN.99.SEMP1299              -----------------Q---S---G-------I--------T----------.AAAQDLKGGYTAVF----QNP.SRKGPI----------L----------------Q--------K...NG--------Y--PG....GT----YA-RQ...VS-S...........  460
O.US.99.99USTWLA              -----------S-----Q-------G-------V--X-----T---KI-----ASAQQDLKGGXTAVF----QNL.GRKGP-----------V----------------Q--------R...NG--------Y--PG....GE-----M-KQMPP-A-S...........  461
```

HIV-1/SIVcpz Proteins — Gag

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

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

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

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

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

protease end p66, p51 RT start — M41L — D67N — K70R — D110 catalytic site

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

```
                                                                                                        K219Q
                                                                   polymerase motif              T215Y
B.FR.83.HXB2              QGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAELELAENREILKEPVHGVYYD 475
A1.GE.99.99GEMZ011        -----------C----------SK--E-----------------------------A--MS--F------------------------------M-LD----------------------------K--------A----DIVT----------------R--------- 475
A1.KE.00.KER2008          -----------C----------SK--E-I--------------------A------A---S--F------------------------------E----E--------------------------K--------A----DIVT-----------------D-------- 475
A1.KE.00.KNH1144          ----------------------SK--E-I--------------------A------A---S--F------------------------------E--N-E---------------------A----KH--R----A----D-VT------------K------------- 475
A1.KE.00.KSM4024          ----------------------AK----I--------------------A------A---N--F-----------------------R---------Q----------------------------K--------A----D-VT-----------------D-------- 475
A1.KE.00.MSA4069          ----------------------SK--E-I--------------------A-V----A---S--------------------------------------E---------------------A----K--------A----DIV------------------D-------- 475
A1.KE.00.NKU3005          ----------------------LK--EMI--------------DL----A-V----N---S--F------------------------------E----E---------------------A-------------A----DTVT-----------------D-------- 475
A1.RU.00.RU00051          -------S---C---R------SK--E------------------------V----A---S--FAX--X-------------------------M--D-------------------------------------A----DIVT-------------------------- 475
A1.RW.93.93RW_024         ----------------------SR----I--------------------A------A---S--F-----------------------------------E---------------------A----K--------A-----IVT-------------------------- 475
A1.SE.95.SE8891           --------------I-------V---E-I--------------------A-V----A---S--F------------------------------Q--N-EN--------------------V----K--------A----DIVT-----------------D-------- 470
A1.SE.95.UGSE8131         ----------------------SK--E-I----------------E----------A---S--FF---Q-------------------------K----E---------------------A----K--------A----DIVT-----------------D-------- 469
A1.TZ.01.A173             ----------------------SK--E-I---------------------------N---K--F---------------------------------D-E---------------------A----K--------A----DIVTM------------------------- 475
A1.UA.01.01UADN139        -------S--------------LK--E-----------------------------A---S--F------------------------------M--D-------------------------------------A----DIVT--------------L----------- 475
A1.UG.92.92UG037          -----------A----------SK--------------------------------E---K--F------------------------------E----E---------------------A----K--------A----DIVT-----------------D----A--- 476
A1.UG.99.99UGA07072       ----------------------SK--E-I----------------E-----V----A---S--F------------------------------Q--N-E--------------------------K--------A-T--DIV------------------D-------- 475
A2.CD.97.97CDKTB48        ------------------D---AR--EM-----------------D---A------A------F------------------------------K--------------------------A------------------DIV---R------E-------D-------- 468
A2.CY.94.94CY017_41       ----------------------SK-TELI----------------S---V------A---K--FY-----------------------------K--------------------------A----K--------A----DIVT--K------E-------T-------- 475
A.SN.01.DDI579            -----------Y----------S---E-I------------N--------------A---K--F------------------------------Q--------------------------A-------------A----DIVT-----------------D----A--- 474
A.SN.01.DDJ369            -----------Y----------S---E-I---------------------------A---K---------------------------------Q--D----------------------------K--------A----D-VV-----------------D-------- 474
A.SN.96.DDJ360            -----------Y----------L---E-I---------------------------A---K--F------------------------------Q--------------------------A----K---R----A----DIVT-----------------D----A--- 474
A.CD.02.02CD_KTB035       -----------C----------S---------------------------------N---K--F------------------------------Q--D----------------------------K--------A----DIVT----------------------E--- 475
A.CD.97.97CD_KCC2         ------------T---------SK--E----------------------A------A---K--F------------------------------Q-------------------------------K---R-I--A----DIVT-------------------------- 475
A.CD.97.97CD_KTB13        ----------------------AK---M---------------------A------D---K--F------------------------------Q-------------------------------K--------A------VE-------------------------- 475

B.AR.04.04AR151516        ---------------------------------------------------V----K---K--F---------------------------------------------------------G----K--------A-S----V---K----------------------- 480
B.AU.87.MBC925            -----------C--------------E------------------------------------F---------------------------------------------------------A----K-----I--------IV---K----------------------- 477
B.BO.99.BOL0122           -----------C------------------------------------------------K-------------------------------------H----------------------------K------------------------------------------- 476
B.BR.03.BREPM2012         ------------------D-----------------------------------------K--F------------------------------M--------------------------A----K---------------V--------------------------- 479
B.CA.97.CANB3FULL         -------S---------------------------------------------------WK--FY--------------------------------------------------------A----K------------------------------------------- 478
B.CN.05.05CNHB_hp3        -----------C-------------------------------------A-V----E------F---------------------------------------------------------A----KE----------------------------------T------- 475
B.CO.01.PCM001            -----------C---------------------------------------V----E---K--FA-----------------------------I-------------------------------K---------------V--SA----------------------- 475
B.GB.83.CAM1              ----------------------------------------------------------------------------------------------M--------------------------A----K---------------V--------------------------- 475
B.GB.86.GB8               --------------------------E------------------------------------FS----------------------------------------------------------------------A-------T-------------------------- 481
B.GE.03.03GEMZ010         --------------------------------------------------------K---S--F-----------------------------V---------------------------A--------------------VS----------------R--------- 475
B.IT.05.SG1               ---------C-----R--------------------------------SI-------X-----F-----X---X--------------------K---X---------------X-----------K--------------I--------------------S---A--- 471
B.JP.05.DR6538            -----------C--------------E---------------------------------G--F---------------------------------------------------------A----K--------A-----I-V----------------------A--- 475
B.KR.05.05CSR3            --------------------------N----------------------A---D------K--F--------------------------------------------------D----------AK--------A----------K----------------------- 475
B.NL.00.671_00T36         --------------------------------------------M-----R---------K--FD-----------------------------------X-X----------X---------------------------IV---A----------------------- 485
B.RU.04.04RU128005        -------S-------------------------------------------------------F----------------------XX--X---K--D-------------------------------------A----------QG---------------------- 475
B.TH.00.00TH_C3198        ----------------------------I-------------------------------K--F------------------------------I--V-----------------------G--R-KE--------------V--------------------------- 475
B.UA.01.01UAKV167         ---------------R-----------------------------------------------F------------------------------K--------------------------A----K---------------V--------------------------- 482
B.US.04.ES10_53           ---------------------------L-----------------E-------K---------F------------------------------E--------------------------A----K---R-----------V--------------------------- 475
B.US.99.PRB959_03         ----------------------E---E-------------------H---------E---K--F------------------------------E----E--------------------------K----------S----------------------R--------- 475

C.AR.01.ARG4006           -------S----X---------A---E----------------------A---K--E---K---------------------------------Q-------------------------------K-------------DIV---A----------------------- 471
C.BR.04.04BR013           ----------------------AK--E----------------------A------E---K--FI---Q-------------------------Q---------------------------------------------DIV--------------------------- 479
C.BW.00.00BW07621         ---------------R------T---E----------------------V------R---K--F------------------------------Q----ED----------------------------------A----DIV--------------------------- 471
C.CN.98.YNRL9840          ---------------R------A---E----------------------A------E------F------------------------------Q----------------------------------------A----DIV--------------------------- 471
C.ET.02.02ET_288          ----------------------AK--E-----------------------------E---K--F------------------------------Q----E--------------------------K--------A----DIV--------------------------- 477
C.GE.03.03GEMZ033         ----------------------TR---M---------------------A------E---K--F------------------------------Q---EE-----------------------------------A----DIV--------------------------- 471
C.IL.99.99ET7             ---------------R------A---E----------------------A------G------F------------------------------Q----------------------------------------A----DIV--------------------------- 471
C.IN.99.01IN565_10        -----------A----------A---E----------------------A------A---K--F------------------------------Q-------------------------------K---R---------DIV--------------------------- 471
C.KE.00.KER2010           ----------------------AK--E----------------------A------E------F------------------------------Q--D--------------------------R--H------------DIV--------------A------------ 471
C.MM.99.mIDU101_3         ---------------R------A---E-----------------------------E---K--F------------------------------Q-------------R--------------------------A----DIV-----------------------A--- 471
C.MW.93.93MW_965          ---------------R------A--------------------------A------E-------------------------------------Q----------------------------------------A----DIV--------------------------- 472
C.SN.90.90SE_364          -------S--------------A---E------------------M---A-V-K--E---K--F------------------------------Q--D-E---------I-------------------------A----D-V--------------------------- 467
C.SO.89.89SM_145          -----------C---R------AK--E----------------------A------E---K--F------------------------------Q----------------------------------------A----DIV--------------------------- 482
C.TZ.02.CO178             ----------------------A---E----------------------A------E------F------------------------------Q---------------------------------------------DIV-----------------R--------- 471
C.UY.01.TRA3011           ----------------------A---EV---------------------A-V----E---K---------------------------------Q----------------------------------------A----DIV-----------------------A--- 471
C.YE.02.02YE511           ----------------------A---E------------------K---A-V----E---K--F------------------------------Q--D-------------------------------------A----DIV--------------------------- 469
C.ZA.04.04ZASK164B1       ----------------------A-----------------------L--E---Q--D---K--F---G--------------------------Q--T------------------------------------------DIV--------------------------- 471
C.ZA.05.05ZAPSK240B1      -----------C------D---AK--E------------------E---A------T---K--F------------------------------Q----------------------------------------A----DIV--------------------------- 471
C.ZM.02.02ZM108           -----------C----------TK--E----------------------A------E---K--F-----------------------------------E---------------------A-------------A-S--DIV-----------------------A--- 466

D.CD.83.ELI               --------------------------EM-------------------------K--E------F-R--------------------------S-K----E--------N--ER--------------------------------------------------------- 474
D.CM.01.01CM_0009BBY      -----------C--------------ELI--S-----------------S------E---K------------D--------------------K----E---------------------S----K--------A----------K-------------R--------- 476
D.KE.01.NKU3006           --------------------------EV---------------------A------E---K--F------------------------------K----E-------------------------------C---A----------K----------------------- 474
D.KR.04.04KBH8            --------------------------E----------------------A---K--E---K---------------------------------T----E---------------------A----K--------A----------K----------------------- 474
D.TD.99.MN011             -----------C------D-------EL---------------------A---K--E---K---------------------------------T----E--------------------------K---R----A-T----V---K-------------R--------- 476
D.TZ.01.A280              --------------------------E-----------------------------D---K--F------------------------------K----E---------------------S----K--------A----------K-------------R-T---A--- 476
D.UG.99.99UGK09259        --------------------------EV----------------------T-----K---K---------------------------------K----E---------------------A---------CI--A------V---A----------------------- 474
D.YE.01.01YE386           --------------------------EV---------------------I------E---K--F------------S-----------------K----E-------------------------------C---A------V--------------------------- 474
D.YE.02.02YE516           -----------C--------------EL-------------------------K--E-------------------------------------T--Q-EN-------------------------K---R----A----D-----R-------------R--------- 476
D.ZA.90.R1                -----------C-------------------------------------A------E------F------------------------------K----E------------------------------------------V--------------------------- 474

F1.AR.02.ARE933           -----------C----------TK---M----------------------------K---T--F------------------------------Q-------------------K-----------K--------A----DIV---A----------------------- 475
F1.BE.93.VI850            -----------C----------MK--------------------------------E------F-----------------H------------Q--N-----------------------------P-------A----DIV---A-------K-----R--------- 475
F1.BR.01.01BR125          -----------C------D---AK--G----------------------A------E------F--------R---------------------Q--N--------------------------R-KN-------A----DIV---A----------------------- 475
F1.ES.x.P1146             -----------C------D---AK--E-----------------L----A------E---K--F------------------------------Q--N-------------------------------------A----DIV---T-------------------A--- 473
F1.FI.93.FIN9363          -----------C----------TR--------------------------------E---K--F------------------------------Q--D------D-------------X----------------A----DMV---A--N-------------------- 474
F2.CM.02.02CM_0016BBY     ----------------------AR--E--V-------------------A------E------F----------------------------A-Q--N-S------------------------R-KH-------A----D-V---A----------------------- 475
F2.CM.95.MP257            --------------I-------E---E----------------------A------E------F----------------------------A-Q--D-S------------------------R-KH------------D-V---A----------------------- 475
F2.CM.97.CM53657          ----------------------VK--E-----------------------------E------F----------------------------A-Q--D-S------------------------R-KH-------A----D-V---V-----M----------------- 475

G.BE.96.DRCBL             ----------------------T---E----------------------A------E------F------------------------------Q--N-EN-------------------------K-----I--A----DIVSM-A---M------------------- 471
G.CM.01.01CM_4049HAN      ----------------------AK--E----------------------A------K---E--F------------------------------Q--D-E--------------------------KH--R----A----DIV---A---M-----K------------- 474
G.CU.99.Cu74              --------------------------E-E--------------------A----------S--F------------------------------Q--D-E--------------------------K--------A----DIV--------------------------- 471
G.ES.99.X138              ----------------------IK--E----------------------A------K---S--F------------------------------Q--D-E--------------------------K--------A----DIV---A----------------------E 473
G.GH.03.03GH175G          --------------I-------TK-------------------------A-V----E---K--F------------------------------Q--D-E--------------------------KH--R----A-----IVS--A---M------------------- 475
G.KE.93.HH8793_12_1       -------P--------------IK--EM---------------------A--K---E------F------------------------------Q----E-----------------------------------A----DIV---A----------#------------ 472
G.NG.01.01NGPL0669        ----------------------T---EM---------------------A------N------F------------------------------Q--D-EN--------------------A----K-------------DIV-----------------R--------- 474
G.PT.x.PT2695             ----------------------IK--E----------------------A------K---N--F------------------------------Q--D-E-------------------------IK--------A----DIV---A----------------------E 474
G.SE.93.SE6165            ---------------R------AN--EM---------------------A------E---K--F------------------------------Q--D-E--------------------------TH-------A----DIVS--A---M---------R--------- 472

H.BE.93.VI991             --------------------------EVI--------------------E------A------F----Q------------------------VK-------------------------------K---X----A-----IV---K-------------------A--- 473
H.BE.93.VI997             --------------------------E-I--------------------A------A------F-----------------------------VK----------------------------------------A----D-V---K-------------R--------- 473
H.CF.90.056               ------------------A---E---EM---------------------A------A---K--F----------------------------TVK---------------------------N---K--------A----DI----K-------------R--I------ 474

J.CD.97.J_97DC_KTB147     -----------C----------AK--EV-----------------E---K------E------F-----------------------------VQ----E--------------------------K-------------DIVA--A----------------------- 472
J.SE.93.SE7887            -----------C------K---ER--E------------------E---R--K---E---K--F------------------------------Q----ED-------------------------K------K-A----DIV---R---------K------------- 473

K.CD.97.EQTB11C           -----------C----------RK---M-L-------------------A------E------F------------------------------Q--D----------------------F-----K--------V----DIV---A----------------------- 475
K.CM.96.MP535             -----------H----------IK--EM-------------------P--------E---K--F------------------------------Q--D----------------------------K--------V----DIV---A----------------------- 475
```

```
                                                
```

```
                                                                                                                    p51 RT end p15 RNase H start
B.FR.83.HXB2              PSKDLIAEIQKQGQGQWTYQIYQEPF.KNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGK.TPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEK.EPIVGAETFYVDGAANRETKLGKAGYVTNRGRQKVVTLTDTTNQKTELQAIYLALQ 642
A1.GE.99.99GEMZ011        -----V--------------------.---------KKGS----------AV---VA---------.----R---------A--M----------------------------.--------------------I-------D-------P--E--------H--H---- 642
A1.KE.00.KER2008          -----T--------D-----------.----------K-S------R--A-V---VVM--------.-----------------MD---------------------------.D--A-----------H------------D-------S--E--------H------- 642
A1.KE.00.KNH1144          ---E----V-----D-----------.----------Q-S---------A-V----AM-G------.----R------------RD---------------------------.D---------------------------D-------S--E--------H------- 642
A1.KE.00.KSM4024          --------------D-----------.----------R-S---------ADV---VVM--------.-----------------MD---------------------------.D--A----------------I-----I-D-------S--E--------H------- 642
A1.KE.00.MSA4069          --------------D-----------.----------K-S-----------V---VVM--------.-----------------MD---------------------------.D-----------------N---------D-------S--E--------H------- 642
A1.KE.00.NKU3005          --------------D-----------.----------K-S------R--A-V--RVVM--------.-----------------MD---------------------------.D--A------------------------D------ISX-E--------H--H---- 642
A1.RU.00.RU00051          -----V--------D-----------.---------KKGS----------AV---VA---------.----R----R----A--M----------------------------.----------------------------DK------S--E--------H--H---- 642
A1.RW.93.93RW_024         --------------D-----------.----------K-S------R--A-V---VVM--------.-----------------MD---------------------------.D--E----Y-------------------D------IS-AE--------H--H---- 642
A1.SE.95.SE8891           --------------D-----------.----------K-S------R--V-V---VVM--------.-----------------MD---------------------------.D---------------------------D---------AE--------H------- 637
A1.SE.95.UGSE8131         --------------D-----------.----------KKS------R--A-V---VVM--------.----------D--D---MD---------------------------.D--A-V----------------------D-------S--E--------H--H---- 636
A1.TZ.01.A173             ------V-------D-----------.----------K-S---------A-V---VV---------.-----------------MD---------------------------.D--I------------------R-----D-------S--E--------H--Q---- 642
A1.UA.01.01UADN139        -----V--------D-----------.---------KKGST---------AV---VA---------.----R---------A--M----------------------------.--------------------I-------D---K---P--E--------H--H---- 642
A1.UG.92.92UG037          --------------D-----------.----------K-S---------A-V---VVM--------.-----------------MD---------------------------.D--A------------------------D-------S--E--------Y--H---- 643
A1.UG.99.99UGA07072       --------------D-----------.----------KKS---------A-V---VVL--------.-----------------MD--------D------------------.D--S-V----------------------D-------S-DE-----A--H--L---- 642
A2.CD.97.97CDKTB48        -----------------S--------.---------KRKST---------------A---------.I---R--------------W-------------------------T.---A------------------------D-----I-P--E--------H------- 635
A2.CY.94.94CY017_41       --------------D-----------.---------KR-ST----I-----------M--------.-----------------A----------------------------.---A------------------------D-----I-S--E--------H------- 642
A.SN.01.DDI579            --------------N-----------.----------K-SS-----R--V-V---VV---------.----R------------M------Y---------------------.D--I------------------------D-------S-ME--------H--H---- 641
A.SN.01.DDJ369            -----V--------D-----------.----------K-S------R--V-V----V---------.-------------D----------C---------------------.D---------------------------D-------S--E--------H--H---- 641
A.SN.96.DDJ360            --------------D-----------.----------KKSS-----R--V-V---VV---------.-------------D---M------C---------------------.D---------------------------D-------S--E--------H--H---- 641
A.CD.02.02CD_KTB035       -----V-------HD-----------.---------KK-S------R----V---VAV--------.----R---------A--MD---------------------------.D---------------------------D-------S--E--------H------- 642
A.CD.97.97CD_KCC2         --------------------------.---------KK-S------R--S-V---VA---------.----R---------A--MD---------------------------.D---------------------------D-------S--E--------H------- 642
A.CD.97.97CD_KTB13        ---E----------D-----------.---------KK-S------R--------VAL--------.----R---------A--MD---------------------------.D--A-V----------------------D-------S--E--------H------- 642
B.AR.04.04AR151516        ------------D-------------.---------X----------------N--A--G------.I---R--------DA--M----------------------------.---A------------------R-----D-------S----------X---H---- 647
B.AU.87.MBC925            --------------------------.----------T----------------------------.I---------------------------------------------.---I-------------------------K------S------------------- 644
B.BO.99.BOL0122           --------L----E-------F-D--.----------AK------I----------A---------.----------------------------------------------.D--E-V------------------------------S------------------- 643
B.BR.03.BREPM2012         ---E----------------------.---------------------------------------.I---R---------A--I----------------------------.----------------------------D-------P--------------H---- 646
B.CA.97.CANB3FULL         --------------R-----------.--------------------------KQ-----------.--------------A-------------------------------.----------------------------D-------SIP------------N---- 645
B.CN.05.05CNHB_hp3        -------------L----------QY.-----------------------------A---------.--------------A-------------------------------.---E-------------------------K---R----N---------H--H---- 642
B.CO.01.PCM001            --------L--------S--------.-------------T---------------A---------.--------------A-------------------------------.-----------------D--S-------D-------S--------------H---- 642
B.GB.83.CAM1              --------L-----------------.---------K---T---------------A---------.-------------DA--ID---------------------------.-------------------R--------D-------P------------------- 642
B.GB.86.GB8               -----V-------L------------.---------K-------------------A---------.--------------A--M----------------------------.---I------------------------D-------P--------------H---- 648
B.GE.03.03GEMZ010         --------V---E-------------.-----------------------------A---------.I---R------------A---------------------------T.----------------------------D-------S--------------H---- 642
B.IT.05.SG1               ----X--------------X-F--X-.---X--------------I----------SX--------.----R---------A-------------------------------.---E---------------R--I-----T---R---S-P---I--------H---- 638
B.JP.05.DR6538            -----V--V-----------------.-----------K-----------------A---------.--------------A--M----------------------------.---------------S------------D-------H-A------------H---- 642
B.KR.05.05CSR3            -----V-------L------------.----------V------------------AQ-------R.I-------------I--M----------------------------.---------------S----A-----I-D--------------------------- 642
B.NL.00.671_00T36         -T------------------------.---------KV------------------A---------.--------------A-------------------------------.----------------------------D------IPIK------------H---- 652
B.RU.04.04RU128005        -X------------------------.---------K---T---------------A---------.----R---------A-------------------------------.---E--------------------------------S--------------H---- 642
B.TH.00.00TH_C3198        -------------L-----------Y.-----------------------------A---------.--------------A----------V--------------------.---E-------------------------K---------------------H---- 642
B.UA.01.01UAKV167         --------V---E-------------.---------K-------------------AI--------.----R------------M---------------------------T.----------------------------D-------S--------------H---- 649
B.US.04.ES10_53           --------------------------.-------------------R---------A---------.I-------------A-------------------------------.-----------------------------K---------------------H---- 642
B.US.99.PRB959_03         --------L---EY------------.------------------I----------A---------.----R-----------------------------------------.---L------------------------D-------S-P------------R---- 642
C.AR.01.ARG4006           ------------E-------------.---------K--T------R---------AQ---I----.----R---------S---D---------------------------.---T----------------M-------D-----I-A--------------Q---- 638
C.BR.04.04BR013           -------------ED-----------.---------K--T------R---------AQ--------.----R---------A---D---------------------------.---P------------------------------I-S--E-----------K---- 646
C.BW.00.00BW07621         -------D-----ND-----------.---------K--T----------------SR--------.----R-------------D---------------------------.D-L--V-----------D--I-------DK--R--I---E-----------Q---- 638
C.CN.98.YNRL9840          S-------------D---------S-.---------K--T----------------AM--------.----R-------------D---------------------------.D--A-V--------------I-------D---K-I-S--E------------I--- 638
C.ET.02.02ET_288          -------------DD-----------.---------KR-T-------L--------AL--------.I---R-------------D---------------------------.---A----------------I-------D-----I-S--E-----------Q---- 644
C.GE.03.03GEMZ033         -------------HD-----------.---------K--A----------------AM--------.----R-------------D---------------------------.---A-V---------S------------D-----II-I-E-----A-----Q---- 638
C.IL.99.99ET7             -------------DD-----------.---------K--T---------VG-----AM-G------.I---R-------------D--------D------------------.-----V--------------I-------D-----I-S--E-----A-----QI--- 638
C.IN.99.01IN565_10        --------L----HD-----------.---------K--T---------A------AL--------.----R-------------D---------------------------.-----V-----------D--I-------D-----I-S--E---------------- 638
C.KE.00.KER2010           -------------HD-----------.---------K--T----------------A--------R.----R-------------D---------------------------.D--A----------------V-------D-------SISE-----------Q---- 638
C.MM.99.mIDU101_3         --------------D-----------.---------K--T----------------AM--------.----R-------------D---------------------------.D--A-V--------------I-------D---K---S--E------------I--- 638
C.MW.93.93MW_965          ---E---------HD-----------.---------K-KS-----------T----AM--------.----R-------------D---------------------------.---A----------S-----I-------DK------S--K-----A-----Q---- 639
C.SN.90.90SE_364          ---E---------HD-----------.---------K--T----------------A---------.----R-------------D---------------------------.D--I----------------I-------D-------SI-E-----------Q---- 634
C.SO.89.89SM_145          -------------HD-----------.---------K--T----------------AQ--------.----R-----------------------------------------.---A----------------M-----I-D-----I-S--E-----------QI--- 649
C.TZ.02.CO178             -------------HD-----------.---------K--T----------------AM--------.----R---------A---D---------------------------.---A------------------------DK----IIS--E-----------Q---- 638
C.UY.01.TRA3011           ------------D-------------.--------TK--T------R---------AL---I----.----R---------A---D---------------------------.---A----------------M-------D-----I-S--E-----------Q---- 638
C.YE.02.02YE511           -------------YD-----------.---------KI-T----------------AM--------.I---R-------------D---------------------------.--M--V--------------M-------D-----I-S-NE-----A-----H---- 636
C.ZA.04.04ZASK164B1       -------------HD-----------.---------K--T----------------AL-------R.----R-------------D---------------------------.D-MI-V----------------------D------IS--E-----------L---- 638
C.ZA.05.05ZAPSK240B1      -------------E------------.---------K--T----------------AI--------.----R-----------------------------------------.---A-----------TKD--I-------D-----IIS--E-----------Q---- 638
C.ZM.02.02ZM108           -------------HD-----------.---------K--T----------------AL--------.V---R-------------D---------------------------.---S----------------I-------D-------SI-E-----------Q---- 633
D.CD.83.ELI               -------------H------------.----------------------A----R-S--------R.----R------------A----------------------------.---I------------------------D-------P--------------N---- 641
D.CM.01.01CM_0009BBY      --------V----ES--SW-------.---------K-KS----------------A----I----.I---R---------A--M----------------------------.---I------------------------DK------PF-------------N---- 643
D.KE.01.NKU3006           --------------D---------QY.---------KL------------------AQ-C------.----R-----------------------------------------.---I----------S-------------D--------I-------------N---- 641
D.KR.04.04KBH8            --------L-----------------.---------K---T---------------AQ-------R.----R---------A------------------------------R.---I---------------R--R-----D------IP--------------N---- 641
D.TD.99.MN011             --------V---------W-------.----------IK-----------------SL--------.----R---------A--M----------------------------.---M------------------------D-------PF-------------N---- 643
D.TZ.01.A280              ------------E-------------.---------K---T---------------AL-G------.I---R-----------------------------------------.---I---------------R--------DK------S--------------N---- 643
D.UG.99.99UGK09259        --------------D---------QY.---------KR------------------AQ-C------.----R------------A----------------------------.----------------------------D-------S--------------N---- 641
D.YE.01.01YE386           --------------D---------QY.---------K-------------------AQ-C------.----R-----------------------------------------.---A-----------S----I-------D-------S--------------N---- 641
D.YE.02.02YE516           --------V---------W-------.------------------------V----AQ--------.----R---------A--M----------------------------.---I------------------------D-------PF-------------N---- 643
D.ZA.90.R1                -------------D------------.----------T-----------A----NLS---------.----R-----------------------------------------.--------------------V-------D-------PF-------------N---- 641
F1.AR.02.ARE933           -------------------------L.------------S----------------AQ--------.--------L----D---ISI-------------------------T.---A----------S-----K-------D-----A-P--E-----A--K-VQ---- 642
F1.BE.93.VI850            -------------D---------N--.---------KV-S----------------AL--------RS-------L----D----D--------------------------T.---A--D-------S-----K-------DK-K----S--E-----A---------- 643
F1.BR.01.01BR125          --------------------------.---------K--S----------------AQ--------.----R--------D-------------------------------T.---A----------S---S-K-------DK----A-S--E-----A-----Q---- 642
F1.ES.x.P1146             ------V-----------------T-.---------K--ST---------------AQ-C------.--------------A---D--------------------------T.---M----------S-----Q-------D-------S--R-----A-----H---- 640
F1.FI.93.FIN9363          ------PKL-------------R---.---------K--S----------------AL--------.--------L----D-------------------------------T.---A----------S-----K-------D-------S--E-----A-----H---- 641
F2.CM.02.02CM_0016BBY     -------------HD----------H.R--------KR-S-----------V----A--G------.V---R----------------------------------------T.---I-----------S----I-------D-------S--E-----------H---- 642
F2.CM.95.MP257            -------------HD----------H.----------RKS-----------V---VA--G------.V---R---------I------------------------------T.---I----------------------I-D-------S--E-----------Q---- 642
F2.CM.97.CM53657          --------V----HD----------H.----------R-S-----------V----A--G------.V---R----------------------------------------T.---L----------------I-------D-------P--E-----------H---- 642
G.BE.96.DRCBL             --------V----L------V----Y.---------KGGS-----------V----A--G-I----.I-------K-----V------------------------------T.---P-V--Y-------------------DK---III---E-----A--H--Q---- 638
G.CM.01.01CM_4049HAN      ---E----V----LD----------Y.---------KKGS-----------V----A---------.I-------R-----V------------------------------T.---P----Y-------------------DK-K--IIA--E-----A--H--H---- 641
G.CU.99.Cu74              ---EI---V----ND----------Y.---------KRGS---------V-V----S---------.--------R-----V---D--------D--------------R--T.---L----Y-------------------DK-K---I---------------Q---- 638
G.ES.99.X138              ---E----V----LD----------Y.---------KRGS-----------V----A----I----.--------R-----V------------D--------------R--T.---A----Y-------------------DK-K--II---E-----A-----QI--- 640
G.GH.03.03GH175G          ---E----V----LD----------Y.---------QRGST----------V----A--G------.V-------R-----V------------------------------T.---P----Y-------------------DK-K--II---E-----A-----Q---- 642
G.KE.93.HH8793_12_1       ---E----V---E------------Y.---------KRGT-----------V----A---------.--------R-----V------------D--------------R--T.---P----Y-------------------DK-K--II---E-----------Q---- 639
G.NG.01.01NGPL0669        --------V----PD----------Y.---------KRGS------R---------A---------.V-------R-----V---D--------D--------------R--T.---P------------------------DK-K---I-IQE--------H--Q---- 641
G.PT.x.PT2695             ---E----V----P-----------Y.---------KRGS-----------V----A---------.--------R-----V------------D--------------R--T.---I----Y-------------------DK-K---I---E-----A-----Q---- 641
G.SE.93.SE6165            ---E----V----LD----------Y.---------KRGS-----------V----A---------.--------R-----I---D-----------------------R--T.---P----Y-------------------DK-K--II---E-----A-----Q---- 639
H.BE.93.VI991             ---E---------PD-----------.---------K--S-----------V----A---------.I---R--------------H---------------H---------T.---E----Y-----------M-------D--K--I-S--E---------------- 640
H.BE.93.VI997             -------------PD-----------.---------K--N---S------------A---------.I------------------H--------*------H---------A.---A----Y-----------I-------DK-K----A--E---------------- 639
H.CF.90.056               ---------R----------------.---------K--T-----I----------S---------.I---R------------------------------H---------T.---A----Y-I-----------------D--K----S--E---------------- 641
J.CD.97.J_97DC_KTB147     SA-E----V----LD-----------.---------KR-T------R----V----SC--------.I---R-------------D------------II-------------.---M----------------I-------DK--------N----------------- 639
J.SE.93.SE7887            SA-E----V----LD-----------.---------KR-S---------A-V----AL-A------.----R----R--------D---------------------------.---M----------S-----T-------DK------------------H------R 640
K.CD.97.EQTB11C           -------------H-----------Y.----------I-S-----------V---VAM--------.----R--------G-------------------------------T.--------------------Q-------DK-----ISI-E-----------H---- 642
K.CM.96.MP535             -------------ND----------H.------------S----------------A--G------.----R----------------------------------------T.---------------H----K-R-----D-------SI-E-----A-----C---- 642
```

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

p66 RT, p15 Rnase H end p31 Integrase start

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

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

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

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

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

p31 Integrase end
Pol end

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

HIV-1/SIVcpz Proteins

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

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

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

Vif end

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

```
                                              amphipathic α-helix                     Vpr end in HXB2
                          Vpr start           oligomerization             frameshift in HXB2
                                                                            H(S/N)RIG motifs         Vpr end
B.FR.83.HXB2              MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIW.LHGLGQHIYETYGDTWAGVEAIIRILQQLLFIH#FRIGCRHSRIGVTRQ....RRARNGASRS*  96
A1.GE.99.99GEMZ011        ------------------M-----D--H--------P-.---------N------E-------T------V-.-----Q-----IEQR....--G---S----  97
A1.KE.00.KER2008          --------------Y--------------------GP-.------Y--H------E---ML---------V-.-----T-----IIPR....----D-SG---  97
A1.KE.00.KNH1144          --------------Y---M-----D--H--------P-.---------D------E-------T------V-.-----Q-----II-G....--G---S----  97
A1.KE.00.KSM4024          --------------Y---M--------H--------P-.---------N------E--------------V-.-----Q-----II-G....--GI--SG---  97
A1.KE.00.MSA4069          ------------------M-----D--H--------P-.---------D------E-------L------T-.-----Q-----IIPGRRG.--V--------  100
A1.KE.00.NKU3005          --------------Y---M--------H--------P-.----------------E--------------V-.-----Q-----II-G....--V-D-SG---  97
A1.RU.00.RU00051          --------------Y---M--------H--------P-.---------N------E--------------V-.-----Q-----IN-R....--V---S----  97
A1.RW.93.93RW_024         --------------Y---M-----D--H--------P-.---------------QE-----V--------V-.-----Q-----IIPG....--G---S----  97
A1.SE.95.SE8891           --------------Y---A---V----H--------P-.---------N------E-------T------V-.-----Q-----II-R....--V-D-S----  97
A1.SE.95.UGSE8131         --------------Y---M-----D--H--------P-.------Y--Y------E-----V-T------V-.----......#II-G....----D-S----  90
A1.TZ.01.A173             --------------Y------------H--------P-.-------V--------E--------------V-.-----Q-----II-GG...--V---S----  98
A1.UA.01.01UADN139        ---V--------------M--------H--------P-.---------N------E--------------V-.-----Q-----IVQR....--G---S----  97
A1.UG.92.92UG037          --------------Y---M-D---D--H--------P-.---------H------E--------------V-.-----Q-----INIRG...--V-D-SG---  98
A1.UG.99.99UGA07072       --------------Y------------H--------P-.---------N------E-------T------V-.-----S-----I--R....----D-S----  97
A2.CD.97.97CDKTB48        ----------P-------V--------Q--------P-.---------N-----RE-----V--------V-.-----Q-----IIPR....--V-D----P-  97
A2.CY.94.94CY017_41       --------------Y---M--------Q--------H-.------Y--N------E---V---Y------V-.-----Q-----II-R....--V-D----P-  97
A.SN.01.DDI579            --------------Y---M--------Q--------P-.---------N------E-------M------V-.-----Q-----II-RG...--V-D----P-  98
A.SN.01.DDJ369            --------------Y---M-----D--H--------P-.------YV-N------E------Q-------V-.-----Q-----I--RG...--G-D------  98
A.SN.96.DDJ360            --------------Y---M------I-Q--------P-.------Y--N------E-------M------V-.-----Q-----II-RG...--V-D----P-  98
A.CD.02.02CD_KTB035       --------------Y---M--------H--------P-.------Y--N------E-----T---H------.Y---Q-----II-S....--R-D-TG---  97
A.CD.97.97CD_KCC2         --------------Y---------------------V-.---------N------E----------------.Y---Q-----I-PS....-GV--------  97
A.CD.97.97CD_KTB13        ------N-------Y---V--------H--------M-.---------N------E-----V-M--------.Y---Q-----IIPS....--I--------  97

B.AR.04.04AR151516        ------N-------Y------------S---X----X-.-QS---Y--------------------------.-----------I-L-....-----------  97
B.AU.87.MBC925            --------------Y---------------------P-.---------------------------------.-----Q-----I---....---------P-  97
B.BO.99.BOL0122           ---------------------------S--------T-.-Q----------------------------L--.-----------I-Y-....K----------  97
B.BR.03.BREPM2012         -----Q--------F------------S----------.--N------------------LL-S--------.-----Q-----II--....-----------  97
B.CA.97.CANB3FULL         --------------Y------------S--------V-.-----------------------------M---.-----Q-----I---R...-AR--------  98
B.CN.05.05CNHB_hp3        --------------Y----------I-R--------R-.--E---Y---------V--------------T-.-----Q-----I---....-----------  97
B.CO.01.PCM001            --------------Y------------RV------SS-.-------------N------VL--T--------.------L----I.Q-....-G---------  96
B.GB.83.CAM1              ---T--------Q-Y------------S--------P-.-LA-------------P----L--S--------.-----------IIQ-....--S--------  97
B.GB.86.GB8               --------------Y------------S-------S--.------Y--------------------------.-----------I---....-----------  97
B.GE.03.03GEMZ010         ---T----------S---A--------S--------P-.--N---Y---------T-------T--------.-----------I---....------S----  97
B.IT.05.SG1               --------------Y-R-------------------T-.------Y-----R--*------T--M-------------H-----M--H....-----------  97
B.JP.05.DR6538            ---------T----Y------------R--------P-.--S------------------------------.-----Q-----IIP-R...-AR--------  98
B.KR.05.05CSR3            --------------Y---S--------R--------P-.--S---Y---------T----L--M--------.-----Q-----II--....-----------  97
B.NL.00.671_00T36         --------------Y-A----------S--------G-.---------------------L-----------.-----Q-----I--R....G-V------P-  97
B.RU.04.04RU128005        ------N-------Y------I--------------V-.------Y---------X--------------V-.-----------I---....-----------  97
B.TH.00.00TH_C3198        ---------------------------S----------.---------D-----------LL-T--------.-----------I-Q-....-----------  97
B.UA.01.01UAKV167         ----------P---Y-A-A--------S--------P-.--S---Y--------------------------.-----Q-----II--....-----------  97
B.US.04.ES10_53           --------------Y--------------------TV-.---------------------------------.-----Q-----I---....-----------  97
B.US.99.PRB959_03         --R---N-------Y---------------------P-.--S---Y------------------------V-.-----Q-----II--....-----------  97

C.AR.01.ARG4006           --------------Y------------Q--------P-.--S---Y---------M--------------V-.-----------II--....-----------  97
C.BR.04.04BR013           --------------Y------------Q--------P-.--S---Y---------T--------------V-.-----N-----IM--....-----------  97
C.BW.00.00BW07621         --------------Y------------Q--------P-.--S------N------T----------------.-----Q-----IM--....--T--------  97
C.CN.98.YNRL9840          ---P-----------------------Q--------P-.------YV--------T---T---L--------.-----Q-----IW--....-----------  97
C.ET.02.02ET_288          --------------YS--------D--Q--------P-.--S---Y--N------R----L-TF---S----.-----H-----I---....-----------  97
C.GE.03.03GEMZ033         --------------Y------I-----Q--------P-.--SI--Y---------T---V---V--------.-----Q-----IL--....-----------  97
C.IL.99.99ET7             ---------------------------Q--------P-.--S---Y---------T----L--T------T-.-----Q-----IMP-....-------R---  97
C.IN.99.01IN565_10        --------------Y------------Q--------P-.------YV--------T-------L--------.-----Q------L--....-----------  97
C.KE.00.KER2010           ---P----------Y------I-----Q--------P-.--SI--Y---------T----------------.-----H-----IL--....-----------  97
C.MM.99.mIDU101_3         ---P-----------------------Q--------P-.------YV--------T---I------------.L----Q-----IV--....-----------  97
C.MW.93.93MW_965          --------------Y---A--------Q--------P-.--N-------------T---V------------.-----H-----IL--....--T---PN---  97
C.SN.90.90SE_364          --------------Y------------Q--------P-.--S-------------T----------------.-----Q-----IIQ-....-----------  97
C.SO.89.89SM_145          ---S----------Y------I-----Q--------P-.--N-------------T--------------T-.-----Q-----IL--....-----------  97
C.TZ.02.CO178             ---P----------Y---A--I--D--Q--------P-.--S------A-H----N----L--M------V-.-----Q-----ILQ-....-----------  97
C.UY.01.TRA3011           --------------Y------------Q--------P-.--S-------------M--------------V-.-----Q-----IL--....-----------  97
C.YE.02.02YE511           -------H------Y------------Q--------P-.--S---Y---------T-----L--------V-.-----Q-----TI-H....-------N---  97
C.ZA.04.04ZASK164B1       --------------Y------I-----Q--------P-.--S---Y--------------LL-----M----.-----H-----ILQH....-----------  97
C.ZA.05.05ZAPSK240B1      ---------------------------Q--------P-.--N---Y---------T--QV------------.Y---Q-----IW--....-------N---  97
C.ZM.02.02ZM108           --------------Y------------Q--------P-.------Y---------T-----------M----.-----S-----II--....-----------  97

D.CD.83.ELI               -----A--------Y---A--------S----------.--S-------------V----------------.-----Q-----II--....------S----  97
D.CM.01.01CM_0009BBY      --------------Y------I-----S--------V-.----------------------T-V--------.-----Q-----I---....------S----  97
D.KE.01.NKU3006           --------------YDA----------S--------M-.--S------D------T----L-----------.-----L-----II--....--TG-------  97
D.KR.04.04KBH8            -----G--------Y---A--------S--------P-.--D------A------V----LL-V------V-.-----Q-----II--....------S----  97
D.TD.99.MN011             -----A--------Y------------S----------.---------------------L-------M---.-----------IQ-T....A-----P----  97
D.TZ.01.A280              --------------Y------------S--------L-.--N---F--------------------------.-----Q-----I-L-....--T---S--P-  97
D.UG.99.99UGK09259        --------------Y------------S--------M-.----------------T----LL-F------V-.-----Q-----I---....--T--------  97
D.YE.01.01YE386           --------------Y------------S--------P-.--S---Y--D------T----L--T--------.-----Q-----IL--....--I--------  97
D.YE.02.02YE516           -----A--------Y------------S--------A-.------------------------T-----L--.-----Q-----ISP-....------S--P-  97
D.ZA.90.R1                --------------Y------------S-------IT-.--S------------------------------.-----Q-----II--....------S----  97

F1.AR.02.ARE933           --------------Y---A-D---------------P-.--S------N-----------LL----------.-----H-----I---....-----------  97
F1.BE.93.VI850            -----G--------Y---A-I---------------P-.---------N------E----------------.-----------IVP-....--V--------  97
F1.BR.01.01BR125          --------------Y---------------------P-.--S------N------E-------M--------.-----H-----IM--....--I--------  97
F1.ES.x.P1146             --R-----------Y---------------------P-.------Y--S------E----------H-----.-----Q-----I---....--V--------  97
F1.FI.93.FIN9363          -----G--------Y---A-----------------PF.---I--Y--N------E------K---------.-----H-----II--....--V--------  97
F2.CM.02.02CM_0016BBY     ---P----------Y------------H--------E-.---------N------E-------T--------.-----H-----II--....--I---NG---  97
F2.CM.95.MP257            ---P----------Y------------H--------E-.---------N------E----------------.-----H-----II--....--I---SG---  97
F2.CM.97.CM53657          ---P-----------------------H--------E-.---------N------E--Q----T--------.-----H-----II--....--L---NG---  97

G.BE.96.DRCBL             --R---------------A--------H--------P-.------L--N------E----------------.-----Q-----IIP-R...--V-D-SG---  98
G.CM.01.01CM_4049HAN      ---V----------Y---------------------P-.------Y--N------E--------------V-.-----Q-----I-P-....--V---PG---  97
G.CU.99.Cu74              ------------------------------------P-.------YV-N------E----------------.-----------ISLR....--V-D-PG---  97
G.ES.99.X138              --------------Y---A-----------------P-.------Y--N------E----------------.-----Q-----IVP-....--V-D-PG---  97
G.GH.03.03GH175G          ------N-------Y---A-----------------P-.------Y--N------E-------M--------.Y---Q-----IIP-....--V-D-P--P-  97
G.KE.93.HH8793_12_1       ---V----------Y---A-----G-----------P-.---------N------E----L-----------.-----Q-----I-LR....--V---P----  97
G.NG.01.01NGPL0669        ------N-------Y------------V--------P-.---------N------E----------------.-----Q-----I-PR....--V-D-SG---  97
G.PT.x.PT2695             --------------Y---A-----------------P-.------Y--N------E----------------.-----Q-----I-PR....--V-D-PG---  97
G.SE.93.SE6165            --------------Y---A-----------------L-.---------N------E----------------.-----Q-----I-PR....--V-D-PG---  97

H.BE.93.VI991             ---------------------------S-----Y--P-.--D------N------E-------T------A-.-----Q-----I---....--V---P----  97
H.BE.93.VI997             -----------------*---------S--------D-.--Q------N------G----L--T--------.-----H-----II--....--V---P----  96
H.CF.90.056               -------------------------I----------V-.--Q------N------V----L--T--------.-----Q-----I---....--V---P----  97

J.CD.97.J_97DC_KTB147     --------------Y------------R--------P-.--N------N------E----------------.-----H-----II--....--G---S----  97
J.SE.93.SE7887            -------H------YH--------------------P-.--S---Y--S------E----------------.-----H-----IIP-....--G--------  97

K.CD.97.EQTB11C           --------------Y----I-I-----R--------P-.--D-----NN------E----------------.-----H-----IL--....--GG-------  97
K.CM.96.MP535             --------------N------I-----R--------P-.--N------T------E-L--------------.-----H-----IIP-....--G---S----  97
```

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


HIV-1/SIVcpz Proteins

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

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

```
                          Rev start          exon 1 end exon 2 start   NLS                                  Leu-rich effector domain                                   Rev end
B.FR.83.HXB2              MAGRSGD..SDEELIRTVRLIKLLYQSNPPP..NPE.GTRQARRNRRRRWRERQRQIHSISERILGTYLGRSAEPVPLQLPPLERLTLDCN.......EDCGTSGTQ.......GVGSPQILV.ESPTVLESGTKE*   116
A1.GE.99.99GEMZ011        -------..T-A--LTA--I--F------C-..T-R.------K-Q-----A-----D-------S-C--------------I---H---S.......---------QSQGTET---G---SG.K-SGI-G------   124
A1.KE.00.KER2008          -------.------LKAI-I--I------Y-..K-K.------K-------A-----D-L-----S-C---P-------------H---S.......---------.......---R--VS-.---DI-G----N-   117
A1.KE.00.KNH1144          -------.------LTA--I--I------C-..K-R.-S----K-------A-----D-------S-C----E---------S--G-R---C.......---------QSQGVEA---G--VS-.---GI------N-   124
A1.KE.00.KSM4024          -------.----G-L-AI-I--I------F-..K-K.-S----K-------A-----D-------SDC----T-------------H---H.......---------QSQGVET---R--VS-.---AI-G----N-   124
A1.KE.00.MSA4069          -------.------FKA--I--I------Y-..K-K.------K-------A-----D-------S-C---P----------I---H---H.......K--------QSQGAET---G--VSG.---DI-G----N-   124
A1.KE.00.NKU3005          ------N.------LKA--T--I------H-..K-G.------K-------A-----D---Q---S-C-------------------H---C.......---------QSQGVET---RT-VSG.---GI------N-   124
A1.RU.00.RU00051          ------N..A--G-LTA--I---------Y-..K-R.-X----K-------X-----D---X---S-CX---X-----P-------H----.......---------QSQGTET---R---S-.---GI-X------   124
A1.RW.93.93RW_024         -------.------LKA--I--I--K---Y-..K-R.-S----K-------A--K--D---K---SAC---------FL-----G-H---S.......---------QSQGVET---R--VSW.---V--G---EK-   124
A1.SE.95.SE8891           -------.------LKA--I--I------Y-..K-R.------K-------A--K--D-L-----S-C---PQ-----L---I---H---S.......---------QSQGAET---G--VSE.--SII-G----T-   124
A1.SE.95.UGSE8131         -------.------LKA--T--I------Y-..Q-K.-S----K-------A-----D-L-Q---SAC----E-------------HI--H.......--------#.........fT-VSG.--SV--D----D-   112
A1.TZ.01.A173             ------N.----D-LKA--I--I------Y-..K-K.-S----K-------A-----D-------S-C----P-------------H---C.......---------QSQGVET---G--VSG.---G--GT---K-   124
A1.UA.01.01UADN139        -------..P-A--LTA--I--I------Y-..T-R.------K-------A-----D-------S-C---------SF----I---HIN-S.......---------QSQGTET---E---SG.--SGI-G----G-   124
A1.UG.92.92UG037          ------N..P----L-AI-I--I------Y-..E-K.------K-------A-----DTL---V-S-C---P-----------I---H---S.......---------QSQGVET---RT-VSG.---V--G----N-   124
A1.UG.99.99UGA07072       -------.------LKA--I--I------Y-..K-K.-S----K-------A-----D---K---S-C---------F----I---H---S.......---------QSQGTET---R--VSG.--SNI-G----N-   124
A2.CD.97.97CDKTB48        -------..P--D---AI-I--I------Y-..K-R.------K-------A-----D-------S-C---PT-------------H---S.......-----C---QSQGAET---R--TS-.--SGI-G--IED-   124
A2.CY.94.94CY017_41       T----D-..P--S-LQAI-T--I------Y-..K-R.-S---Q--------A-----D-----V-R-C---PT-------------HI--S.......--------LQSQGTET---RS-ES-.--SVI-G---E--   124
A.SN.01.DDI579            ------N.------L-AI-I--I------Y-..K-Q.------K-------A-----D-L-----S-----E-------------H---S.......---------.........GS*-S-.-CSAI-GE---KQ   112
A.SN.01.DDJ369            -------.------L-AI-I-RI--E---F-..K-R.-----Q--------A-----N-------S-C-E-PT----F------G-HI--S.......---------QSQGTET---G-S-SG.--SVI-G*---K-   123
A.SN.96.DDJ360            ------N.------L--I-I--I------Y-..K-R.------K-------A-----D-----L-S-C---PE-------------H---S.......---------.........GS*-S-.--SAI-GE---KQ   112
A.CD.02.02CD_KTB035       ------H..T----LTA--I--II-----Y-..K-S.-S----K-------A-----D-------S-C----E----F----I---H-N-S.......----A----QSQGTET---E---S-.---G--G-AAEK-   124
A.CD.97.97CD_KCC2         ------S.---A--LAAI-I--I------Y-..K-R.------K-------A--N--D-------R-C----E----FC-------H---S.......---------QSQGTET---G---S-.--SG------EN-   124
A.CD.97.97CD_KTB13        ------S.------LTA--I--I------Y-..K-R.------K-------A--N--D-L-----S-C---PE----F--------HI--S.......---------QSQGTET---R---P-.--SDI-V----N-   124
B.AR.04.04AR151516        -------.----D-LK------F---------..SS-.------X-------------QTL-HW--S-H---P-----------------D.......---------.......-X--S--P-.--X----------   117
B.AU.87.MBC925            -------.------LK------F------L-..S--.---------------------RE--G---S-F--GPT------------------.......K--------.......------V--.-------------   117
B.BO.99.BOL0122           -------.----D-LQ------F---------..---.------------------H-QAV-AW-IS-H----T------------------.......---------.......---D-----.------G------   117
B.BR.03.BREPM2012         -------.------L----I-F---------..---.---------------A-----RK--DW--S-H---P-----------------S.......---------.......----T--S-.--S---D------   117
B.CA.97.CANB3FULL         -------.------LK------F---------..SS-.-------------------RTL-----S-----T-------------------.......---------.......---------.-------------   117
B.CN.05.05CNHB_hp3        ----RE-..N----LK--K---F------L-..---.---K---------A-----RE-----D---G-T-----------------S.......---------.......R-------G.---A--D------   117
B.CO.01.PCM001            -------.----D--Q---F-R----L-----..SR-.-------------------G---AW--RSH-----------------N---S.......----A----.......------V--.---A------E--   117
B.GB.83.CAM1              -------.------LKA-----------L-..SSK.-----------------H-RA--NW--S-H---P-----------------S.......K--------.......---------.---A---------   117
B.GB.86.GB8               -------.----DFLK---F--F---------..--K.---------------A------Q-G-W--SAF---P----------I-------D.......---------.......------V--.---A--DP-----   117
B.GE.03.03GEMZ010         -------.------L-----------------..S--.---------------A-----RT--DW--R-H-D--K---------S--------S.......---R-----.......---------.----I-G------   117
B.IT.05.SG1               -------.------LK--Q---Y------FR..S-T.--------------K--W-Q-L-GW--SNH---PM---------------I-TG.......------S--.......---T--V--.-----------E--   117
B.JP.05.DR6538            -------.----D-LKA-----F------Y-..SL-GT------------------R-------C---E-P-------------S---S.......---------.......---------.---AI--------   118
B.KR.05.05CSR3            -------.---D--LK------F---------..S-Q.-------------------Q---AWV-RNH---------P------------S.......---------.......---D--V--.--S-I--------   117
B.NL.00.671_00T36         -------.------L-------Y---------..SS-.-----------------R-RA---W--SNH---P-----------------S.......---------.......---------.---AI-D----G-   117
B.RU.04.04RU128005        -------.----D-L-A----F---I-----..SQ-.-------------------RTL-GW--SD----PX------------------.......---------.......---N-----.-P-P---------   117
B.TH.00.00TH_C3198        -----E-.-------K-----R-------L-..S--.---------------A-----RE------D----G-T------------------.......---------.......--------L.------DP-A-K-   117
B.UA.01.01UAKV167         -------.----D-LK------Y---------..---.---------------A-----RT--DW--SNH---P-----------------S.......---------.......---G-----.---A---P-----   117
B.US.04.ES10_53           -------..NV---LT---I--F---------..SQ-.-------------------RTL-GW--SN--------------I-------------.......---------.......------V--.---A--D------   117
B.US.99.PRB959_03         -------.------LQKI---RF---------..SS-.---------------A--------GW-VS-L---P-----------------S.......---------.......-E-N-----.-PS------D---   117
C.AR.01.ARG4006           -------.----A-LLA--I--T------Y-..K--.---R----------A-------------I-C---PT---------I---NIN-S.......-SG-----.......R--N-*-FR.KPCA--G-RA-KE   116
C.BR.04.04BR013           -------.----A-LQA--I--I------Y-..---.----------------A-------------T-C----T----F----I---NIN-S.......-GS------QPQGTPE---D-*VFR.KPCAI-G-RA-KE   123
C.BW.00.00BW07621         -------..N-DT-LQA--I--I------Y-..K--.----------------A-------------S-V---P-----F----I---HIG-S.......-SG------QPQGTPE-M-N-*-SG.N-CA--G-RA-KE   123
C.CN.98.YNRL9840          -------..G--A-L-A--I--I------Y-..E-R.------K-------A--KR---------S-C---P----HF----I---HIGDS.......-SG-A----QSQGTTE--A--*-SG.K-CA--G--A-KE   123
C.ET.02.02ET_288          -------.---AA-LQA--A--I------Y-..K--.-----Q--------A--------L-----SDC---HP-------------N---S.......-SG------QPQGTTE---N-*VSG.QYCA--G-RA-KE   123
C.GE.03.03GEMZ033         -------.----A-LQA--I--I------Y-..E-K.-----Q--------A--------------S-C---------------I---HI--S.......QSS-A----QSQGTTE-----*-FG.K-GA--G----KE   123
C.IL.99.99ET7             -------.DQ-AA-LQA-KI--I------Y-..K--.------K-------A--K---T------S-C-------------------HI--S.......-SS------QPQGNTT-----*VSG.K-CA--G-RA-KE   124
C.IN.99.01IN565_10        -------.----A-LQA--I--I------Y-..E-Q.-----Q--------A-----C-------S-C---P-----F----I---HISDS.......-RS------PSQGTTE---N-*-SG.K-C---GF-A-KE   123
C.KE.00.KER2010           -------..N-D--L-V--I--I------Y-..K--.------K-------A---R---------SAC---P-----F----I---HI--S.......-GG------QSQGTTE-----*VSG.KLCA--G--A-KE   123
C.MM.99.mIDU101_3         -------.----A-L-A--I--I------Y-..E-R.-----QK-------A--K-------V-S-C---P------L---ID--HISDS.......-SG-A----QSQGTTG-----*-SG.R-CA--G--A-KE   123
C.MW.93.93MW_965          -------.----A-LQA--I--I------Y-..K--.----------------A-------G----A-C---P-----------I---HI--S.......-SS------QSQGTAE-----*VSG.K-CA--GARA-KE   123
C.SN.90.90SE_364          -------.----A-LQA--I--I------Y-..K--.------K-------A--------------DC---P-----------I---HI--S.......-SS--PD-E.........---N-*VSG.KPXA--GF-A-KE   116
C.SO.89.89SM_145          -------.----A-L-A--I--I------Y-..K--.------K-------A--------------S-C---PT----------I---HIGGS.......-SG------QSQGTTER----*VSG.K-WA--G--A-KE   123
C.TZ.02.CO178             -------.---DA--QV-KI--II-----Y-..E-K.----------------A--------------SAC-E-P--------------HI--S.......-SG------QSQGTTE---N-*VPG.RLRA--G---EKE   123
C.UY.01.TRA3011           -------.----A-L-A--I--F------Y-..K--.----------------A--------------T-C----P-S-------I---NIN-S.......-SG------QPQGTTE---N-*VFR.-PCAI-GPRA-KE   123
C.YE.02.02YE511           -------..N-AA-LQ--KI--I------Y-..KL-.----------------A--K-----G---SCC-E-P-----------I---HI--S.......-SS--T---QSQGNTE----S*VSR.K-CA--GARA-KE   123
C.ZA.04.04ZASK164B1       -------.S---A-LQA--I--I------F-..E-K.------K-------A---E-N-------S-C---P-----F----I---HIGDS.......-NL------.......---N-*VSG.KPY---G-RA-KE   117
C.ZA.05.05ZAPSK240B1      -------.----A-LQA--I--I------Y-.NPN-.----------------A------Q------V-----PT----------I---HI--S.......-SS------QSQGTTE---N-*-SG.K-R-I-G-RA-KE   124
C.ZM.02.02ZM108           ------V.----A-LQAT-I-RI------Y-..K--.------K-------A------Q------S-C----------FR---I---HI--D.......-SS------QSQGTTE---N-*VSG.K-CA--G-RA-KE   123
D.CD.83.ELI               -------.----D-LKA-----F---------..S--.---------------A-----RE-A----------P-------------N-N-S.......---R-----.......---H---S-.-----------E-Q   117
D.CM.01.01CM_0009BBY      ----RE-.----D-LKV-----F---N-----..S--.---------------A---HL-T-G---S-------------------N-N-S.......---------.......-------S-.------G---E--   117
D.KE.01.NKU3006           -------..N--D-L-AI----I---------..S--.---------------A--------G--IS-----E---F--------N-N--.......---------.......----H-V--.---AI-D------   117
D.KR.04.04KBH8            -------.----D-LKAI----F------L-..S--.---------------A-----N--G--IS---------------I----IN--.......---------.......----S-VSG.--SAI-DT-----   117
D.TD.99.MN011             ----REE..N--DFLK-I----W--K------..ST-.---------------A--------G----S---------S-----I---N-N-S.......---EP----.......----S--S-.----------E--   117
D.TZ.01.A280              -------.------L-AI----I---------..SL-.---------------A-----D--G----S-------------------N-N-H.......T--------.......----H-VS-.-P-AI-D---E--   117
D.UG.99.99UGK09259        -------..C-GA-LQA-K---I---------..SS-.---------------A----L-T-G----S-----LE-----------------S.......---------.......----H-VF-.---A--D------   117
D.YE.01.01YE386           -------.------LK-I----F---------..S--.----------------------G----S------E-------------N-N--.......----A----.......----H-VP-.-PSA--D------   117
D.YE.02.02YE516           ----RE-.----D-LKAI----F---N-----..S--.---------------A---HL---G----S-L--G----------I---N-N-S.......----P----.......-------S-.-PS----P--ED-   117
D.ZA.90.R1                -------.----N-LTAI-I--F---------..S--.---------------A-----N--G--V-F-------------------------S.......---------.......-------S-.---A---------   117
F1.AR.02.ARE933           -------.---T--LTAL-Y--I------Y-..K--.---------------A-----RA--D---SSC---PE-------------HIN-S.......---------.......PGTEE-L-G---AG.--HA-----I-N-   117
F1.BE.93.VI850            -------.---T--LKA-KC--I------Y-..K--.---------------A-----RAL-D---SSC----E-------------HIN-S.......---------.......QGPEE----S--SG.--HA---------   117
F1.BR.01.01BR125          -------.------LKAI-Y--I------Y-..K-R.-------------NA---H-RE------SSC---PE-------------HIN-S.......----E----.......QGTEE---N--TSG.-PRA-----D---   117
F1.ES.x.P1146             -------SDT-T--LK---C--I------Y-..K--.---------------A------K------RSC--GPV-------------HIN-S.......---------.......QGTEK-------S-.--HA--G---Q--   119
F1.FI.93.FIN9363          -------.---T--LKA-KY--I------Y-..S-D.---------------A-----RA------SSC---LE-------------HIN-S.......---------.......QGTEE-------SG.-HH---------   117
F2.CM.02.02CM_0016BBY     -------.----A-LTA--T--I------Y-..K--.---------------A-----RE------S-C-------------------HI--S.......---S-----.......QGTER-------S-.--RS--G---EN-   117
F2.CM.95.MP257            -------..R----LKA--Y--I------Y-..KL-.---K-----------A------Q------S-C---LQ---R----L--K-HIN-S.......---------.......QGTEK-------S-.--RA--G------   117
F2.CM.97.CM53657          -------.------LKA--Y--I------Y-..K--.---T-----------A--N--RQ------S-C---PT-------------HIN-S.......----E----.......QGTNK---N---SG.--RA--GP-N---   117
G.BE.96.DRCBL             ------S..T----LTA--I---------S-..P--.---------------A-------------S-C---PE-------------H---S.......--G------QSQGTEI-------F-.--SV--G------   124
G.CM.01.01CM_4049HAN      ------A..T-G--LQA-QI--I------Y-..P--.----T-K-------A--K---T------S-C---PT----F----I---C---S.......--S------QPQGTET---R--V-W.---AA-G------   124
G.CU.99.Cu74              ------S..T----L-A--I--I------Y-..P--.------K-------A--K--------V-S-C----X-------------S---S.......K-S------QSQGTET---R---S-.---G--G------   124
G.ES.99.X138              ------S..T--D-L-AI-I--I------Y-..P--.------K-------A-------------SACV--PE---S------------------.......--SX-X---EXEGTET---W---XG.-CSIE-G------   124
G.GH.03.03GH175G          ------S..N----L-A--I--I------Y-..SA-.------K-------A----VRAL---F-SAC---PT------------C---S.......--S------QSQGTET---R---S-.---V--G------   124
G.KE.93.HH8793_12_1       ------S..T--D-L--I-I--I------Y-..P--.------K-------A--------------SSC-------------------H--SS.......--G------QSQGTET---G---F-.--SV--G--A---   124
G.NG.01.01NGPL0669        -------..T--A-LQA--I--I------Y-..P--.----------K---A---------K---SDC---R-----F--------H---S.......--SRIT---QSPGTET---G--TS-.---A--G----N-   124
G.PT.x.PT2695             ------S..T----L-AA-I--I--E---Y-..P--.----T-K-------A-----------V-S-C----T-------------H---S.......--S--L---QPQGTEP-L-G---S-.--SV--G------   124
G.SE.93.SE6165            ------S..T----L-A-KA--I------Y-..P--.---------------A-----SA------TA----P-------------H---S.......--S------QPQGTET---RS*-F-.--SV--GQ-----   123
H.BE.93.VI991             -------..N--G-L-AC-I-R-------Y-..E-A.-----Q--------A--------G--V-A-C--GP-----------------S.......---------.......EK------TSG.---A--GT-A---   117
H.BE.93.VI997             ------A..G--Q-PQVCKI--II-----Y-..E-A.---------------A-----RA------TDC---PP------------------.......K--------.......EK-K-G---P-.--S---GT-----   117
H.CF.90.056               ------A.---T--LQVCKI--I------C-..E-T.---------------A-----RE------TSC---PP---T----------N-S.......---------.......EK-E-----SL.--S-I-GT-----   117
J.CD.97.J_97DC_KTB147     -------.----Q-LLA-----I---G--Y-..K-N.-S------------A--N--D-------SAC---P-----F----I---R---S.......---D-----.......TQ---D---SG.--GVL-GA-----   117
J.SE.93.SE7887            -------.---DQ-LLA-----I------Y-..K-N.-S------------A--N--D------PSSC---P-----------I---R---S.......---N-----.......TQ---D---SG.-PCM--GA-----   117
K.CD.97.EQTB11C           ----R--..-EQQ-LTP--I--I------Y-..K--.---------------A-----RE--Q-V-SSC----T------------S-N-D.......--S------.......QGTEGEL-----P-.-PD---G--D---   117
K.CM.96.MP535             ----R--..P--Q-LT---T--I--E---Y-..-L-.-S--T---------A--K--S-----L-SAC--------------I-K-N-N-D.......--P------.......KGTEG-L-----S-.-PC----------   117
```

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

```
                                                                        phos phos
                         transmembrane domain                       cytoplasmic domain
                         Vpu start                   α-helix                        α-helix        Vpu end
B.FR.83.HXB2             MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHYESPHPRISSEVHIPLGD..ARLVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPELADQLIHLYYFDCFSDSAIRKALLGHIVSPRCEYQAGHN.KVGSLQYLALAALITPK.....KIKPP  162
A1.GE.99.99GEMZ011       --------------------N----------K---E---------R--K-----------.----VK------A--K-----H-------Q------I--D---------H------E------IV------K-N-PP--S.-----------K--V--T.....RER--  162
A1.KE.00.KER2008         --------------------N--------I-K--K----------R--KV-----------..-K--VR------P--K-----H-------L------------------H--N---E------I--QV---S--------.-----------K--V---.....-T---  162
A1.KE.00.KNH1144         ---K----------------N----------K--KN-V----F--RN--V-----------..-K--VR------P--N-----H-------L-S-------D-----------------G------I--EL-------T--S.-----------K--V--Q.....-T---  162
A1.KE.00.KSM4024         --------------------N----------K--KN-----F--R--KV-----------..-K--VI---------K--------------RG-----I--D---------H--A--T-------I--QV-R-----PT---.-----------K--V--T.....-TR--  162
A1.KE.00.MSA4069         --------------------N----------K--SK-W-------R--KA-------V-E..-----R-----N---KA-Q--H--------LRK-------DQ-----------------E------I--QV-----D--T---.-----------K--V---.....-T---  162
A1.KE.00.NKU3005         --------------------N-------H--K--K-----------R---V-----------..--I--R----------K--Q--H-------L------LT-DM-------H--N----------I--QV---------T---.-----------K--V--P.....R-R--  162
A1.RU.00.RU00051         --------------------N----------------L-------R---V-----------..----VR---------K--Q--H-------Q--X---I--D------------------------E--R-----------.-----X----K--V--T.....RGR--  162
A1.RW.93.93RW_024        --------------------N----------K-TKN---------R---V-----------..---IVR---------K-----H-------L-----I--D---------H------E----R-I--QV---S---HT---.Q---------K--V--T.....-TR--  162
A1.SE.95.SE8891          --------V-----------N--------I-K--K--L----F--R--KV---------E..----VR-----Q---K-----H-------LR-----I--D---------H--K----------I--E---------------.-----------K--V---......--R--  162
A1.SE.95.UGSE8131        --------------------N--------I-K--A-----------R--KV-----------..---IVR------P-----Q--H-------LR-----I--D---------H--N----------I--RV---S----T---.-----------K--V---.....RT---  162
A1.TZ.01.A173            --------------------N----------R--KD----------R--KV-----------..----VR-----Q---K-----H-------L-----I--D---------H--------------I--QV-R---D-------.-----------K--V---.....-----  162
A1.UA.01.01UADN139       -----------------X--N----------K---K-V-K-----IN-KV-----------..----VK---------K-----H-------Q------I--D-----------------E------IV-Q------N-P----.---P------K--V--I.....RER--  162
A1.UG.92.92UG037         --------------------N--------I-RR-K----------R--KV-------I--..--I-VR-----Q---K-----H-------L-----I--D---------H--N----------I--QV-----D--T---.-----------K--V--S.....R----  162
A1.UG.99.99UGA07072      --------------------N----------K--K-------F--R--KV-----------..----V------Q---K--Q--H-------L-----I--D---------H--N--------R-I--QV---S---------.-----------R--V--T.....RTR--  162
A2.CD.97.97CDKTB48       --------------------N----------K---E---------R---V--------KE..---IVR------P--K-----H-------QG----I--D--H-----------E----R-I--E--R----------.-----------R--VAST.....RT---  162
A2.CY.94.94CY017_41      --------------------N--------I-R--K--V-K----RN-----------E..-IIVR------I--K-----H-------QN--H--I--D--H-----------E------II-E--------------.-----------K-VVAST.....RT---  162
A.SN.01.DDI579           --------------------N--------I-K--KD-------DCR---A-----------..-M--VR------P--Q-----H-------Q-----I--D---------H-----------R-I--Q--R-----T---.-----------R--VAS-.....-TR--  162
A.SN.01.DDJ369           -----------------K--N----------R--KK-V-K-----R--KV----Y----E..-I--VR------P-----Q--H-------Q-----I--D---------H--N-------------Q--R-S--------.-----------K--V--E.....RT---  162
A.SN.96.DDJ360           --------------------N--------I-K--KN----------R---V-----------..-I--VR------P--Q-----H-------Q-----I--D---------H--N----------I--Q--R-----PT---.-----------R--V---.....-----  162
A.CD.02.02CD_KTB035      --------V--------K--N--------I-K--K--V----F--R--KA-----------..---IVR-----Q----------H------QL-----I--D---------H--E----------I--N-----D-------.-----------K--VA--.....GT---  162
A.CD.97.97CD_KCC2        ------------------A-N-------H--K--I-----------R--K-----------..----VK---------G----H------KLR-----I--D---------H--T---E------I---R---K--------.-----------K--V--T.....RA---  162
A.CD.97.97CD_KTB13       --------------------N--------I-K--K-----------RP-K-----------..----VR-----QP-----Y--H------KMR-----I--D---------H--------------I-----------------.-----------K--VQ--.....RR---  162

B.AR.04.04AR151516       -----------------S-----------T----KR-I-------TN-------------..---I------------E-------------EK------DI-------------------------------R-S--------.--------------VK--.....RKR--  162
B.AU.87.MBC925           ---------------------------------K--C-------TN-----------E..-K---V------------------------------G-D---------------E----------R-------------.------------------.....-----  162
B.BO.99.BOL0122          -----------------K-----------I-K--K--C------NL-----G------P..-K---------------------------E--A---T-DI---------------A-------I--R----S-DFR----.-------------V---.....-T---  162
B.BR.03.BREPM2012        -----------------------------I----KE---------T-------------..------------------------------R-------D-------------N---E----N-I--R----K--------.R---------------R.....RV---  162
B.CA.97.CANB3FULL        -----------------K---------K-IT---KE-I-K----N-------------..-KV-V-----------------------------E-------D-----------------E----N----R----S--------.------------------.....Q----  162
B.CN.05.05CNHB_hp3       -A---------------------------I-R--K--C-K---D-T--K-----------..-K-----------------------------RD----AE-S-----------------E------I--RV--SS-N-PE--S.-------------V---.....RR---  162
B.CO.01.PCM001           -----------------------------I-K--K-----------T--K---------E..N---V-----------------------------R-------D-----------------E----N-I--NL-R-----------.--E-------T-------.....-----  162
B.GB.83.CAM1             -----------------------------I----KK-S-------T-------------E..----V-----------------------------T-G-N-----D-----------------E------IV-RL-------------.-----------T---A--.....-----  162
B.GB.86.GB8              -----------------------------I----KK-V-K----NT-------------E..------------------------------R--R-----G-------Q-------E----N----RT---S-K------.-----------T-------.....-----  162
B.GE.03.03GEMZ010        -----------------K--H----Y---I-K---D-R-------T--K-----------..-E--VI------S-----------------------------DS-----------------E----R-II-RL-------------.-----------T--VK--.....-K---  162
B.IT.05.SG1              --------------------H---I--K-M-K--KN-V------N-----G-----------.GK-------------Q--Q--H-------E-------D--R-----------E----N-I--RR---L---------.-----------------R.....RR---  162
B.JP.05.DR6538           -------------------------Y-----R-TER-V-------T--K-----------..DK-----------------------------R-------N----------E---EA---HTI--R-----------T.--------------K--.....-----  162
B.KR.05.05CSR3           -----------------------------I-R--K--V------NI---V-----------..-K--V-----------------------------G--N-----T-------R-------------------I-------S--------.-----------T-------.....-T---  162
B.NL.00.671_00T36        -------------------------Y---R----KK---------T-------------E..-K-----------------------------ER--T-----D-------V-------E----N-I--RL-------------.------------V------.....-----  162
B.RU.04.04RU128005       -----------------K--N-I------I-R--KK-V-------TN-K-----------..-K---K-------Q-----H-------RG-N-----D-----------------EXX--N-II-----------PT---.--------------IT--N.....-R---  162
B.TH.00.00TH_C3198       -----------------K-----------I-R--K--I-K---D-S--K-----------..-T-------------------------------N-N-----G-----------------E----N-IV-RV---S-D------.------------V-------.....-R---  162
B.UA.01.01UAKV167        --------------------N--------I-----K-C-K----TN--T-----------..-K--VI--------E---------------K---I--D-----------------G----S-I--RL---S----T--S.--------------VK--.....RQ---  162
B.US.04.ES10_53          -----------------S------Y------K--------------K---------E..------------------------------R-------------------------E----N-I--Q---------------.-------------------.....RR---  162
B.US.99.PRB959_03        ------------------A------------K--K-----------R--------------..-K---A-----------------------------R-------N-------M-------E----H-I---R---S--------.-----------T--V---.....-T---  162

C.AR.01.ARG4006          -------L----------A-N--------I-K-----Y-------R--KV-----------..-----I----Q---------H-------LXK-------G-------MH----T-----------R-------------.-----------T--LK--.....-R---  162
C.BR.04.04BR013          -------L--------K---N--------I-KR-K--Y-------R--K----------E..-K-------XQ---N----H-------LR-------D-------TH----A-----------R----D------.-----------T---K--......--R--  162
C.BW.00.00BW07621        -------L--------K---N----------KR-N----------R--KL-----------..-----K-----------------H------KLR-------G-------MH----A---------QV-I---D------.-----------T---K--.....-R---  162
C.CN.98.YNRL9840         -------L--R-----K---N--------I-RR-N--L-----D-R--KV---------E..-----K-----Q--------H-------LR--N--IE-G--------------A-------I-----I--D------N----------T---K--.....-----  163
C.ET.02.02ET_288         -------L-------K--A-N--------I-RR-N--L-K----RN--T---------E..---IVK-----Q--------H-------LRG-T-----C-------MH----AE------I--QR-----D------.-----------T--LK--.....-R---  162
C.GE.03.03GEMZ033        -------L----------K-N----------KR-S-------F--R--KT-----------..-----K-----Q----E---H-------LRK-------G-------MH----A-------I-----I---D------.-----------T---K--.....QR---  162
C.IL.99.99ET7            ------GL--------K-K-N----------KR-S----------R--KV-A---------..-----K-----------------H-------LR-------G-------MH----AK------I-----I---D------.-----------T---K--.....-R---  162
C.IN.99.01IN565_10       -------L--------K---N--------I-RR-K--------D-G--KV-----------..-----K-----Q--------H-------LRS-----E-G-------MH----A-------I-----I---D------.-----------T---K--.....-----  162
C.KE.00.KER2010          -------L------------N-------HI-RR-S-------F--R--KV---------E..-----K-----Q--------H------KL--N-----G-------MH----A-------I-----I-----------.-----------T---K-N.....-----  162
C.MM.99.mIDU101_3        -------L--------K---N--------I-RR-K--------D-R--KV---------E..---I-K-----Q----E---H-------LR------E-G--------------A-------I-----I---D------N----------T---K--.....-----  163
C.MW.93.93MW_965         -------L------------N----------RR-D----------R--KV-----------..-----K-----Q---------H-------LR-------G-------MH----A-------I---------D------.-----------T---K--.....-----  162
C.SN.90.90SE_364         ------GL--------K---N--------T-RR-K----------R--KV-----------..-----K-----------------H-------LR-------G-------MH----T-------I-----I-----------.-----------T---K--.....-----  162
C.SO.89.89SM_145         -------L------------N--------I-KR-S----------R--KV---------E..-----K-----Q--------H-------LR-------G-------MH----A-------I-----I---D------.-----------T---K--.....-----  162
C.TZ.02.CO178            -------L------------N----------RR-S----------R--KV---------G..-----------Q---------H-------LR-------G-------MH----A-------I-----I---D-P----.-----------T---K--.....-----  162
C.UY.01.TRA3011          -------L----I---K---N----------KR-S--Y-------R--K----------E..-K---------Q----E---H-------LR-----I--G-------MH----A-----E-I---R-----D------.Q---------T---R--.....-----  162
C.YE.02.02YE511          ------GL------------N--------I-RR-S----------R--KV---------E..-----K-----------------H-------LR-------G-------MH----T-------I-----I---D----T.-----------T---K--.....-----  162
C.ZA.04.04ZASK164B1      ------AL------------N--------T-KR-K--Y-------R--KV---------E..-----------Q----E---H-------LR-------G-------M-----A-------I-----I---D------.-----------TT--K--.....R----  162
C.ZA.05.05ZAPSK240B1     -------L--------K---N-------HI-KR-K----------RN-KV-----------..-----K-----Q---------H-------LR-------G-------MH----A----Q-I---------D------.-----------T---K--.....RT---  162
C.ZM.02.02ZM108          -------L-------K--A-N--------I-KR-N---K----RN-K-----------..-----------------------------H-------LR-------G-------MH----A-------I---------D------.---------T-T---K--.....-R---  162

D.CD.83.ELI              -----------------K-------------K--NR-----------K---------E..-----K---------E------------R-------G-------M-------E------I--D-------------.-----------T---A--.....Q----  162
D.CM.01.01CM_0009BBY     -----------------N-------Y--H--K--K----Q-------K-----------..-K--V-----------------------------------G-------M----A-------I---R----D------.-----------T---E--.....-----  162
D.KE.01.NKU3006          --------------------N----Y-----K--Y--L-----C-N-K---------E..----VK---------E------------R-------G-------I-----AE------------T---N------.-----------T---K--.....R----  162
D.KR.04.04KBH8           -----------------N-------Y---I-K-----V----D-N---------------..-K--V------N---K-----------L--R--I--G-------MH------------------I---R-----D------.-----------T-------.....R----  162
D.TD.99.MN011            --------V--------N-------Y-----KE-K--------D---K---------E..----VK------N-------------------R-------G-------TH-----------------I--QR-----D------.-----------T---K-R.....-R---  162
D.TZ.01.A280             -------------------------------R--Q--L-----D-H--K-----------..----V----------E------------R-------G-------M-----T-------I---------D------.-----------I---K--.....-----  162
D.UG.99.99UGK09259       ------------------A--------IHI-K--K--L-----DC---K---------E..----VR---------E------------R-------G--------------AE------I--Q--T---N------.-----------T---AQ-.....-----  162
D.YE.01.01YE386          --------------------N----------K--Q--L-----DCL--K-----------..----VK-----------------------------R-------G-------I-----AE------I---V-T---N------.-----------T---K--.....GT---  162
D.YE.02.02YE516          -----------------N-----------I-XE-K--------D---K---------E..-K--VI-----N----------------QR-------S-------T-----E------I--LR-----D------.-----------T---A-R.....RR---  162
D.ZA.90.R1               -----------------N-------Y--H--K--K--------D---K---------E..E--VK---------K---P--------R--I-----G-------I-----E------I---R-----------N----------K---AS-.....-T---  163

F1.AR.02.ARE933          -----------------N-------Y--H--K--KK-----F--R---V-------EE..-K---I--------------------XGN-R--I--G-------I-S-----E----G-I---RI----N------.RL--------T---A-Q.....-T---  162
F1.BE.93.VI850           ------L----------N-------Y-----K--K--S----FQ-R---V-------EE..VK----------P---E-------------QGK-R--I--G-------I-------E------I---RI----N------.-----------T---A-E.....-T---  162
F1.BR.01.01BR125         -----------------N-------Y--H--K--KR-L----F--R---V-------EE..-K-----------------------------QG--R--I--G-------I-------E------I---KI----N------.Q---------T---A--.....-----  162
F1.ES.x.P1146            --------V--------S-------Y--HI-K--KK-C----F--RN--V-------KE..-KI-V----------E-------------QGK---I--D-------I-------E------I--YR----N------.--K-------T---A--.....-T---  162
F1.FI.93.FIN9363         -------T---------N-------Y--H--K--KR-----F--R--K---------EE..-E----------N----------------QG--R--I--G-------I-------E------I---RI----D------.-----------T--VS--.....-A---  162
F2.CM.02.02CM_0016BBY    --------------------N----------RR-Q----------R---V-------EN.DSK---V---------K-----H-------Q--R-----G-------K-------E------I---R----N------.-----------T--V---.....-V---  163
F2.CM.95.MP257           ----------------K---N--------I-KR-A----------R---V-------EE.DSK---I-----------------------Q--R-----G--------H-------------------I--QR----N------.-----------T-------.....-----  163
F2.CM.97.CM53657         --------------------N----------RR-K----------R---V-------K-.DSK---VS-----------------------Q--R-----G-----------------E------I--QR----N------.-----------T-------.....-----  163

G.BE.96.DRCBL            --------------------N--------I-K--K----------R--KV-----------..-K--VR---------DKE---H-------QG-------D---H----H--N---E------I--QT-R-S---P----.-----------KV-VA-T.....RRR--  162
G.CM.01.01CM_4049HAN     -------V------------N------KHI-K-T-H----------R--K-----------..-M-IVR---------K-----H-------QR-----I--D---H-----------E------I--E--------------.-----------K--VP-T.....RKR--  162
G.CU.99.Cu74             --------L-----------N----------K--KD-W-------R--K-----------..---IVR---------K-----H------KQRK----I--D---------H--K--------Q-I--EV-R---------K.-----------K--V--T.....-----  162
G.ES.99.X138             ----------------K---N----------K--K-------F--R--KV-----------..-T--VR---------K-----H-------QR-----I--D---------H--------------I--R-----------T.-----------K----T.....-A---  162
G.GH.03.03GH175G         --------------------N--------I-K--K-----R---R--KV-----------..-T--VK---------QA----N-T-----QR--N--I--D---H----H------E------I--E--------------.-----------K--VK-A.....RK---  162
G.KE.93.HH8793_12_1      --------------------N-----L----K----------F--R--KV-----------..-K--VS-----N---K--Q--H-------QR-----I--D---------H------E----N-I--R--R-----------.-----------K--V--T.....-TR--  162
G.NG.01.01NGPL0669       ----------------K---N--------I-K-------G-----R---T-------I--..-KI-VI---------KA-Q--H-------QR-----I--D---------H----T-------I--EL-R-----------.-----------K--VK--.....-----  162
G.PT.x.PT2695            ------------------A-N----------K--K--S----F--M--KV--D---------..-K--VR---------K-----H-------QR-----I--D---------H------------------G-----------T.-----------K--A--T.....-A---  162
G.SE.93.SE6165           --------------------H----------K--------P--A-R---V-----------..-T--V----------K--Q--H-------QR--R---E-D---H----H--------------I--Q--------------.Q---------KV-V-S-.....RSR--  162

H.BE.93.VI991            -----------------K--N----------K--KK-V-------TN-KT-------V--..-------------------------H-------QE-----I--D---------H--------------I---R---I-D-----R.-----------T---S--.....RT---  162
H.BE.93.VI997            -----------------NK------Y---I-K--NR-V-----D-----------------..-----------------------------Q---------G-------TH------E----G-I--RV-------------.Q---------T--VA--.....-T---  162
H.CF.90.056              -----------------N-------Y--HI-R----------F--T-----------E..-----------N---E-----------L-------E-G-------MH------E------I--RV-R---N-P---K.Q--T------T--VA--.....-----  162

J.CD.97.J_97DC_KTB147    -----------------N-------Y--HN-K---K-L-----D-N--KV---------C..-E--V-----------------------L---R-----G-------MHH-----EA-----I-----------------.-----------T---K--.....-R---  162
J.SE.93.SE7887           -----------------N-------Y--N--K---K-L-----D-N--K---------E..-I--------Q-----------------QR--R-----G-------MC---------------------I--Q-----D------.-----------T---R--.....RR---  162

K.CD.97.EQTB11C          -----------------N-------Y-----K--NR-R-----D-N--K-----------..-E--V-----------E------------L-K-R-----G-------I-------E----------R-----------T.Q---------T---A--.....-T---  162
K.CM.96.MP535            -----------------N-------Y--HI-K--NR-Y-------R--K-----------..-E--V------L-------------------L---R---E-D-------I-------E--V---I---R------C-----.-----------T--VA-R.....RP---  162
```

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

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

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


HIV-1/SIVcpz Proteins

```
                                                
```

```
                                         Env start                    signal peptide  gp120 start                                   glycosylation NVT                          glycosylation NDT
                                                                                                                                                                                        V1
                                                                                *                          *                    *                                           *      *    *
B.FR.83.HXB2                MRVKE...KYQ.HLWRWGW....RW.....GTMLLGMLMI.CSATEK..LWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLVNVTENFNM.WKNDMVEQMHEDIISLWDQSLKPCVKLT.PLCVSLKCTDLKND.............. 137
01_AE.CF.90.90CF11697       -G--GTQMNWP.---K-.............--LI--LVI-.---SDT..------------RD-D---------N-QE----------V----------IH-E------D-.---N-----Q--V---------------.----T-H--KA-LN.............. 136
01_AE.CN.05.FJ051           -----TQMNWP.N--K-.............--LI--LVI-.-N-SDN..-------------D-D-----------HK---------------------MH-E--------.---N-----Q--V---------------.----T-N--NAAVN.............. 136
01_AE.CN.06.FJ054           -----TQMNWP.N--K-.............--LI--LVI-.---SNN..-------------D-D---------R-HE-----I-V-------------IL-E--------.---N-----Q--V---------------.----T-N--TANLP.............. 136
01_AE.HK.x.HK001            -----TQMNWP.N----.............--LI--LVI-.---SDN..-------------D-D-----------HE---------------------IK-GD-------.-X-K-A---Q--V---------------.----T-N--QANWK.............. 136
01_AE.TH.01.01TH_R2184      -----TQMTWP.N--K-.............--LI--LVI-.---SDN..------------RD-D-----------HE--A------------------IH-E--------.---N-----Q--V---------------.----T-N---A-WT.............. 136
01_AE.TH.02.OUR769I         ----GTQMSWP.---TG.............-ILII-LVI-.---S-N..------------RD-D-----------HE---------------------IH-E--------.---N-----Q--V---------------.----T-N--NANLT.............. 136
01_AE.TH.90.CM240           -----TQMNWP.N--K-.............--LI--LVI-.---SDN..------------RD-D-----------HE---------------------IH-E--------.---N-----Q--V---------------.----T-N--NANLT.............. 136
01_AE.US.00.00US_MSC1164    -----TQMNWP.N--K-.............--LI--LVI-.---SD-..-------------D-E---------R-HE---------------------IY-E--------.---K-A------V---------------.----T-N--TANLT.............. 136
02_AG.CM.02.02CM_4082STN    ---MGIQKN-P.L----.............-MIIFCIMI-.-N-KD...-------------D-E------------------------------S---MP-E----D---.---N------V--------G-----R--.----T-H--NFNRS.............. 135
02_AG.EC.x.ECU41            ---MGIQKN-P.LF---.............-VIIFWTM--.-N-AN...------------RD-E------------Q-----------------S---IN-G--------.---N---------------------E--.----T-N-QNVT-I.............. 135
02_AG.GH.03.GHNJ196         ---MGTQKN-P.C----.............-IIIFWIML-.-KGED...------------RD-D----------S-------------------S---IY-E--------.---N---------------------E--.----T-E-HSVT-S.............. 135
02_AG.NG.01.PL0710          -G-M-IQKSCP.P--K-.............-MIIFWTMI-.-N-D-...-------------D-E---------------A------------------IH-E--------.---K------------------------.----T-N---RL-N.............. 135
02_AG.NG.x.IBNG             ---MGIQKN-P.L----.............--NIFWIMI-.-N-EQ...-------------T-E----------------------------------IH-E----K---.---N------------------------.----T-D-HNFN-S.............. 135
02_AG.SE.94.SE7812          ---MGIQKN-P.L----.............-MIIFWIMT-.---GN...------------RD-E----------------------------------IH-G----D---.---S---------------------Q--.----T-H-Q-NLTS.............. 135
02_AG.SN.98.MP1211          ---MGILKSCP.PF---.............MI---WI-I-.-N-EN...-----------RRD-E------------E-----------------S---MH-E--------.---N------V-----------------.----T-E-HNYNYT.............. 136
03_AB.RU.97.KAL153_2        -----IRK.H..----..............--LF------.-----N..----------------------------SK--------Y-------S---IP-K--------.G--N------------------------.----T-N-----KE.............. 133
04_cpx.CY.94.CY032          ---MGMQRN-P.---E-.............--LI--LVI-.---SNN..------------RD-E-------E---EK-----I----------------A-I--------.-----------------NEG----A---.S---TFT-INATTT.............. 136
05_DF.BE.x.VI1310           ---RGMQRNWP.--GK-.............-LLF--I-I-.----D-..F-----------R--K----------G--K-----------------------G--------.----------T-----------------.----T-N---F-AN.............. 136
06_cpx.AU.96.BFP90          ----GIQTSW-.---K-.............--LI--LVI-.---SKN..M---------A-ED-D-I----------SPDK------------------IS-K--------.---N--D-----------E---------.----T-T--NATLG.............. 136
08_BC.CN.97.97CNGX_6F       ---RGTRRN--.QW-I-.............-VLGFW----.-NVEGN..--------------K-------------E-----------------S---I-ME--------.-N----N-----VT--------------.----T-E-RNVSSN.............. 136
09_cpx.GH.96.96GH2911       -----IQRNW-.N----.............-L-I------.-N-A-N..--------------E----------Q--SK-K------------------IG-E--------.---N------K-----------------.----T-D-HRVN-S.............. 136
10_CD.TZ.96.96TZ_BF061      -T-TGIQRNCR.QW-I-.............-ILGFW----.-N-R--..--------------T-------------KA-A--I-------------R-I--E--------.---G--D------------G----I---.----T-N-NGINAT.............. 136
11_cpx.GR.x.GR17            -----TQRNWH.N----.............-L-IF-----.-N-EK...M-----------RD-D-I-------RT-S--K-------S--------R-LS-E--------.---N--------V-----E---------.----A-N---A................. 132
12_BF.AR.99.ARMA159         ---RGMQRNW-.--GN-.............-FLF--I-I-.-N---N..--------------------------S-ER-----------------------E------D-.---N------T-----------------.----T-N-SNANAT.............. 136
13_cpx.CM.96.1849           --ARGIQMTW-.N--K-.............--LI--LVI-.---SNN..------------RD-E----------S--K--------------------IN-G--------.-E-N------T----I---------S--.----T-N-NNITS-.............. 136
14_BG.DE.01.9196_01         -KA-GIQRNC-.R--K-.............-VT-I-----.-N-V-Q..--------------T----------R--------------------S-L----E--------.---N------------------------.----T-D--N-NET.............. 136
15_01B.TH.99.99TH_MU2079    -----TQMNWP.N--K-.............--LI--LVI-.---A-N..-------------D-S---------------A--G-----------S---EI-G--------.-E-------Q------------------.------N----R--.............. 136
16_A2D.KR.97.97KR004        ---RGIQRN--.---K-.............-ILI------.SK---D..------------RD-E---------------A------------------IN-E--------.---G------------------------.----T-N-SRV--T.............. 136
17_BF.PE.02.PE02_PCR0155    ---RGMQRN--.---K-.............-I----I---.-N--KD..------------R--X----------G-E------------------------E------D-.---N------T-----------------.----T-N--NANAT.............. 136
18_cpx.CU.99.CU76           ---RGTQRN-P.S--K-.............--II----LM.----DH..-------------D-V----------S-S--S--------------D---IT-G--------.---N--------------E---------.----T-N--NVXXS.............. 136
19_cpx.CU.99.CU7            ---MGTRRS-P.P--N-.............--LI--LVIM.---FDN..------------RD-E------------------------------S---IN-E----E---.---N------A-----------------.----T-D..................... 130
20_BG.CU.99.Cu103           ----GIQTNW-.C--T-.............--LII-LGI-.---SNN..------------ED-N-P----------S--S------------------IP-E--------.---N-----N--------E---------.----T-N--N-SVN.............. 136
21_A2D.KE.91.KNH1254        ---R-MKRNC-.N--K-.............--I--F----.----DN..------------R--NAS----------TK-A--I---------------IE-E--------.---N--------------E---------.----T-N-AGSTTS.............. 136
22_01A1.CM.01.01CM_0001BBY  ---RGTQMTW-.N----.............--LI--LVI-.---QN...------------R--E------------P-----------------S---IE-A--------.---K-----Q------------------.----T---E-ARFN.............. 135
23_BG.CU.03.CB118           ----GIQKNW-.R--T-.............--LII-LGI-.---SDN..------------ED-Y-P----------S--S------------------IP-X--------.---H--------------E---------.----T-N--KVDCE.............. 136
24_BG.CU.03.CB378           ----GIQMNW-.---T-.............--LII-LGI-.---SNN..-------------D-D-P--------S-S--S------------------IP-E--------.---N--------------E---------.----T-N--T-NSN.............. 136
24_BG.CU.03.CB471           ---RGIPMNW-.L--T-.............--LII-LGI-.---SSN..------------ED-Y-P----------VA-S----------------H-IP-Q--------.---N-----N--------E---------.----T-N--SVDCN.............. 136
25_cpx.CM.01.101BA          ---R-IQRN-P.L----.............-ILII-LGI-.---SDN..-------------D-E-X----------SP-A--------------S---IE-T--------.---N--------------E------Q--.----T-N-VT-NCS............NR 138
27_cpx.CD.97.97CDKTB49      ----GTQRNW-.G----.............-L-I---A--.---K-D..------------RD-D-----------HT---------------------II-E--------.---N--------------E---------.----T-N--NANAT.............. 136
28_BF.BR.99.BREPM12609      ---RGMQRNWK.--GT-.............-LLF--I-I-.-N--NQ..S------------------------N-VK--A--I-----------I------K--------.---N--------------E-I-------.----T-N--NNVTV.............. 136
29_BF.BR.02.BREPM119        --ARGTRKNC-.---E-.............-IL-------.---EQ...-----------------------N-------A--------------D---IP-K--------.----------Q----------Q------.----T-N-S--G-N.............. 135
31_BC.BR.02.110PA           ----GIQRNW-.QW-I-.............-ILGFW----.YNVRGN..------------R--K------------------------------S---LD-G--------.-E----D---Q-----*-----------.----T-N-S-AISN.............. 135
32_06A1.EE.01.EE0369        ---RGTQMNWS.---K-.............--LI--LVI-.-N-SKD..----------A-ED-D-I---------HS--R------------------IH-K--I-----.---H---------K----E---------.----T-H-KNATFK.............. 136
33_01B.MY.05.05MYKL007_1    -----TRMNWP.N--K-.............--LI--LVIM.---SNN..------------RD-D-----------HE---------------------IH-K-I------.---K-A---Q--V------------Q--.----I-H--NANLT.............. 136
34_01B.TH.99.OUR2478P       -----TQMNWL.N--K-.............--LI--LVI-.---SDN..------------RD-D-----------HE---------Y-----------IH-E--------.---N-----Q--V---------------.----T-N--NAALTNVTI.......... 140
35_AD.AF.05.05AF095         T--MGTQMNW-.N----.............---I---II-.---V-N..-------------D-E------------E--M------------------IP-G----D---.---K------T-----------------.----T----NATTNI............. 137
36_cpx.CM.00.00CMNYU830     --ARGTQRNCL.---K-.............--LII-LVI-.---SDL..------------RD-N----------------------------------ID----------.---N-----Q------------------.----T-E-QNFNGSNNH........... 139
37_cpx.CM.00.00CMNYU926     ---MGIQRN-P.P--N-.............--LI--LVI-.---NNL..-------------D-D---------T--VP--------LPGTHRPH.--KINVG---KD---WKN-MVEQMQIDI-SLWGPKL.---E---PS-RH#-NSSHKV-PQDQRE.......... 140
38_BF1.UY.05.UY05_4752      ---RGMQRNW-.P-GK-.............SLLF--I-I-.-N-ED...----I-------R---------------------------------D---I--G------D-.-R-P------T-----------------.----T-H--N-TTK.............. 135
39_BF.BR.03.03BRRJ103       ---RGMQRNW-.--GK-.............-LLF--I-I-.-N-A-N..------------RD--------------E----------------------E-K------D-.---N------T-----------------.----T-N-SNVNVT.............. 136
40_BF.BR.04.04BRRJ115       ----GIKKNC-.R--K-.............-M----L---.-N-NDQ..W-------------------------G-E---------------------I--G--------.----------------------------.----T---AAFNGT.............. 136
42_BF.LU.03.luBF_05_03      ---TGIMKN--.-X-K-.............-I----I---.S-----..---------------X------------------I---------------------------.---N------------------------.----T-N--NXXXXXXXXXX........ 141
43_02G.SA.03.J11223         -----IQRNW-.Y--TM.............WXLI--LXI-.---SNN..------------ED-X-P-----N----S--K--X---------------MT-E----P---.---N------D-------E------M--.----T-T--NATCN.............. 136
U.CD.83.83CD003             -----IQRN--.---K-.............SLII---I--.-K-I--..S------------D-E------------EK-S--------------S---L--G--------.---K------------------------.F---T-N-I-V--S.............. 136
U.CD.90.90CD121E12          --A--IQRN-H.Y--K-.............SLII---II-.-KGE-Q..S------------D-D----------G-EK-A--------------S---L-----------.---N-----------------------S.S---T-N---VGTN.............. 136
U.GR.99.GR303               --L-GIQRNC-.N----.............AL-IF-----.-K--DN..-------------D-E-----------HSA-K---L----------S---MI-G--------.---N------T-----------------.----T-N-KNPSCV.............. 136
U.NL.95.U_NL_95_H10986_D1   ---RGMQRNWR.--G--.............-ILF--I---.---AD-..----I-------QD-N----------S-SN-R------------------LP----------.---N----V---------E---------.----T-N---VT-R.............. 136
CPZ.CD.90.ANT               --KPIHII........................WGLA-L-QFIEKGTNEDY----F------RN--P-----TN-SMTS---------TS---I--D-IV-R.L-TSVW--A.Y--Y---S-T--MXQ-FQ--H-------.-M-IKMN--GYNGT.............. 127
CPZ.CM.01.SIVcpzCAM13       -K-M-ERKRSF.NGFYMT............MTICMA-L-P-LTSD...E------------R-VN-V---------HS--A--I------------------H----D---.-N-Q-----Q---S--------------.----TM--SNVTRQ.............. 136
CPZ.CM.05.SIVcpzEK505       -K-T-MQKNW...................LICC--IG-I...KIIGS.E------------RD-E-V---------HS--A--I---Q-------------LIP---R-D-.---N--D--Q-------E----------.----T-S-SSWRSV.............. 130
CPZ.CM.05.SIVcpzLB7         ----GTWTSCM..........PYYVVGILSLCS-IIS...STKN-N..--------------EA-----------QKA-A--I--SQ-------D-KAIE-S--------.-T-A--N--QQ-V-----------I---.----T-H-SIP-F-.............. 137
CPZ.CM.05.SIVcpzMB66        -KAM-TQRNCR.T-SLKEI...........ILCT-VLGI-.GIIKCEDNM-----------R--D-----------QNP-------SQ---S-N---E-IE-T----Y--A.-E-N--------VN------V------I.----T-N-SLF-CI...........KEN 143
CPZ.CM.05.SIVcpzMT145       -K-MGMKYYWW.RNSCRSI...........SIK-ILIGW-.A-CFGEENW-----------R--K----------S-S--A--I---Q-------T---L-P----E----.-E-Y--D--Q-------E----------.----T-T-NNPT-T.............. 140
CPZ.GA.88.GAB1              -K-M-.KK-RD.WNSLSII...........TIITIIL-TP.-LTS-...------------HD-DPV---------HS--A--I---Q-------S---F-P--I-S----.---N--D---------------------.----T-Q-SKANFS.............. 136
CPZ.TZ.00.TAN1              *KKMK.NLIG...ITLI..............LIITI-G-.GFS-....YYT--F--------QP---------DITSRDK--I----N---L----Y--T-A--SIR---.EE-Y--QE-K---L--FQ--F-------.-F-IKMT--MTNTT.............. 128
CPZ.US.85.CPZUS             -K-M-.KK-R...--LS.............YCL-SSLIIP.GLSS....--A---------RDVE------------KQ-A--I---Q-------------H-P---K-D-.-E-N-A---Q--------------I---.----TMT-LNPDSN.............. 131
CPZ.CM.x.SIVcpzDP943        ---M-RIRGS.....................LW-MG-I-P-LTGSD.Q--A-----------D-E------A--S-LNK-A--I--SQ------------HVP-I--T-D-.-E-N-----QT-----------------.----TMH-IPYNIS.............. 130
CPZ.CM.x.SIVcpzMB897        -K-MGTQRSWK.QFLSLR..........IVMCLW-LGV...IRSENN..--------------K------------QIA-A--I--SQ---------E-IT-E----E-DA.-N-N--D--Q--L---------------.----T-N-STNTTT.............. 137
N.CM.02.DJO0131             V-MMGMQIGW....................PFFC-MIS-T-.G-KK....Y-A--------RDVE-V---------HS--A--I---Q------------P-D----P---.-E-K-A--VQ-------E----------.----TMN-SNSNGN.............. 129
N.CM.04.04CM_1015_04        ---MGMQSGW....................PFFC--ISSI-.G-EQ....H----------RDTE-V---------HS--A--I---Q------------P-I----H-D-.---N-A---Q-------E-------X--.----TMH-XNSNGX.............. 129
N.CM.04.04CM_1131_03        K--MGMQSGWVTRRVLGMMSGW........PFFC--ISSI-.G-EQ....H----------R-TD-V---------HS--A--I---Q------------H-I----K-D-.---N-A---Q-------E----------.----TMH-N-SNGF.............. 141
N.CM.95.YBF30               -K-MGMQSGW.........MGMKSGWLLFYLLVSLIKV-.G-EQ....H----------R--E-----------HS--A--I---Q-------------L-P---K---.-E-K-AD--Q-------E----------.----TML-N-SYGE.............. 138
N.CM.97.YBF106              R--MGMQSGW....................PFFC--IS-T-.G-DP....H----------RD-E-V---------HS--A--I---Q-------------L-T---Y---.-E-K-A---Q-------E----------.----TML-NNSNGN.............. 129
O.BE.87.ANT70               -KAM-KRN-K...--TL.............YLAMALITP.-LSLRQ..-YA---A-----ED--PV-------NLTS--K--I--SQ-------T-Y-YP-H---DD--I.---Y-----Q---------------QM-.F---QME--NIAGT.............. 133
O.CM.91.MVP5180             -K-MKKNNRK...S-SL.............YIAMAL-IP.-LSYS-.Q-YA---S-----E--APV-------NLTS--Q--I--SQ---------H-FP-G---D--DI.---Y--D----------E------E-M-.F---QMN-V--QTN.............. 134
O.CM.96.96CMABB637          -K-M-MKN-K...-CIL.............CI-MAL-IP.-LNSDQ..-YA--FS-----ED--P--------NLTS--K--I--SQ-------S-H-YP-I---DK--I.-E-Y--D--Q-------E-------QM-.F---QMN---VNAT.............. 133
O.CM.98.98CMA104            -KAMKKKN-R...--IL.............YIVM-LITP.-LSSSQ.FYA---A-----EN-NPV-------NLTSP-Q--I--AQ---------H-YP-H---DR-DI.-E-Y-----Q---TD--E--------M-.FM--QME-N-INTT.............. 133
O.CM.99.99CMU4122           --AM-KRNRQ...-GIL.............YIVMALIIP.-LSYDQ..-YA---S-----E--KPX------VNLTS--K--I--SQ-------T-H-YP-K---DK--I.-E-Y-----Q-------E-------QM-.F---QMN--NVNDT.............. 133
O.FR.92.VAU                 -KAMGKRNRK...-GI-.............CLI-ALIIP.-LSCNQ..-YA---S-----ED-KP--------NLTS--Q--I---Q-------S-N-YE-K---DK--I.-E-Y--D--Q----D----------QM-.F---QMN---I--S.............. 133
O.SN.99.SEMP1299            -K-M-QRNRK...-GIL.............CIVMALITP.-LSYNQ..HYA---A-----E---PV-------NLTS--Q--I--SQ-------S-Y-YP-TK--D---I.---Y-----Q-------E-------QM-.F---QMN--NYVQG.............. 133
O.US.99.99USTWLA            -KAM-XKNRK...S-IL.............CXIMALIIP.-LNS-Q..-YA---S-----ED-SP-------VNLTS--Q--I--SQ-------S-N-YD-K---DY--I.---Y--D------------------QM-.F---QMN---V-DN.............. 133
```


HIV-1/SIVcpz Proteins

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

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

```
                                                          glycosylation NIS
                       glycosylation NDT          glycosylation NCS                                                                                               glycosylation NVS
                                   V1                                                    V2                                                                 glycosylation NGT
                               glycosylation NSS                                              glycosylation NDT           glycosylation NTS          glycosylation NKT               glycosylation NGS
B.FR.83.HXB2              ..........TNTNSSSGRMIMEKGEIKNCSFNISTSIRGKVQKEYAFFYKLDIIPIDND...........................TTSYKLTSCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAE 267
01_AE.CF.90.90CF11697     ..........DTY-GTAKLNDTIGD-VR-----VT-EL-D-K-EVH-L--VP--VR-GEKNKN.....................SSGNSSE-I-IH------K------T-D-------T---Y------E-N-------K---S-------K--------------- 273
01_AE.CN.05.FJ051         .........NSKSITV-NNTVGNITDDV---T--MT-EL-D-K--A--L--R--LV----.......................N-E-R-IN------K-----I--D-------T---YV-----D-N-------K---S---A------A------------ 267
01_AE.CN.06.FJ054         .........KNNQ-GEYIEIIGNMTE-VR-----MT-GL-D-Q--VH-L--R---A--NENNS.....................NENN-E-R-IN------K-P---I--D-------T---Y--F---D-N-------K---S-------K-------------S- 273
01_AE.HK.x.HK001          .......SNNT-QNIN-XVTIGNMTD-VR-----MT-ELTD-Q--V--L------VE-N-.......................ST-R-IN------K--------D-------T---YV-----D-K-S-----N---S-------K--------------- 268
01_AE.TH.01.01TH_R2184    .........KVS-PTDVPSIVGNVTD--------MT-EL-D-K-SVH-L--R---V-MKANS.....................SNS-E-R-IN------K-----I--D-------T---Y---R--E-N-T-----K---S-------K--------------- 271
01_AE.TH.02.OUR769I       ..........NVKSTNEAPSIGNITD-VR-----MT-EL-D-KK-VH-L--R---VQ-E-N......................SSE-R-IN------K-----I--D-------T---Y-L----D-K-----L-K---S-------K--------------- 267
01_AE.TH.90.CM240         .........NGSSKTNV-NIIGNITD-VR--T--MT-ELTD-K--VH-L------VQ-EDKK.....................TSSE-R-IN------K-----I--D-------T---Y------D-N-------K---S-------K--------------- 270
01_AE.US.00.00US_MSC1164  ..........NV-ITKD-NKIGNLTD-VR-----MT-EL-D-KH-VH-L--R--LV--E-N......................NSSE-R-IN------K-----I--D-------T---Y------D-N-------N---S-------K--------------- 268
02_AG.CM.02.02CM_4082STN  ...............NYNISDDMRE---------T-EL-D-KK----L--RQ-VVQ--SE........................VGQ-R-IN----A----------Q-----------------RD-S-----S-K-----------K-------------T-- 260
02_AG.EC.x.ECU41          ...................TEEMRE--------MT-EL-D-K--VS-L--R--VVQ-N-SQ......................DNSH-R-IN----A--------T--------------------------K---K-----------K--I-----------S- 258
02_AG.GH.03.GHNJ196       ............SE-KIGNIS-EMQ---------MT-EL-D-NR-MH-L--RQ--V-MNESLVSI..................NTTNS-DQ-R-IN----TV------------------------D-N-----L-R-----------K--------------- 273
02_AG.NG.01.PL0710        ...............N-TFNNSMV--------XMT-VL-D-K-NVH-L-----VVQ-NSNS......................NSSE-R-IN----A-------IT---------------------D-E-S-----K-----------K--------------- 262
02_AG.NG.x.IBNG           ............YSN--NLTSDMN----------T-EV-D-KK-MH-L--R--VVQ-NEN.......................NGSQ-R-IN----A----------------------------KD-G-------K-----------K--------------- 264
02_AG.SE.94.SE7812        ...............-GNISENMQ---------MT-EL-D-K--V--L--RY-VVQ-NETG......................DNIQ-R-IN----A-----------------------------DEK-------K-----------K--------------- 262
02_AG.SN.98.MP1211        ...............RNNSK-N-VQ-M------MT-VLKD-KK-MA-L---I--V---KN........................A-Y-R-IN----A-----------------------------E-D-S---S-K-----------K--------------- 261
03_AB.RU.97.KAL153_2      .............VT-TNTSSIKMM-M-------T-DL-D--K----L-----VVQ---.........................D--R-I------V------I-----------------------D-K-----------------K--------------- 258
04_cpx.CY.94.CY032        ...............N-TNGTVI-EG------D-T-E--D-KK----L--RI--V--NARVPING..................SNRNNS-EE-M-IN--A-T-K-------------------------E-N-T-L-------S-R-----K--------------T 272
05_DF.BE.x.VI1310         .........STA-STTN-TTLKE-T-AVQ-----MT-EVND-KL-VH-L--R---V--SS-D......................SSNS--R-IN----T-K-------WD----------Y------E-K-S-----K-----------K--------------K 271
06_cpx.AU.96.BFP90        .NKTLGNNSTNSTLGNN-TIVDDISK---------T-E--D-TK----L--RP--V--GD-S......................NNSD-R-IN--V-T-K--------------------------RD-D-------K-----------K--------------- 278
08_BC.CN.97.97CNGX_6F     ...........GNGTYNETYNESVK--------AT-LL-DRKKTV--L--R---V-LNDENSG....................ENSSEY-R-IN----A--------T-D-------T---Y------D-I-----Q-H-----------K--------------- 271
09_cpx.GH.96.96GH2911     ..............KPNTTSDGSEVRMQ------T-EL-D-H--V-SL--RS--VS-NAS.......................NNE-R-IN----A--------T------------------------RD-E-----I-K---S------------------------ 262
10_CD.TZ.96.96TZ_BF061    ..................NDTVGDPRMT-----TT-EVKD-KKQ---L-----VVQ-GDS........................N---R-IN----A--------T---------------I---DRK-------K-I---------K--------------- 258
11_cpx.GR.x.GR17          ....................RDNATD----T---T-ELED-KKN-R-L--R--VV--NDSS.....................SNIGQ-R-IN--V-AVK---S--T-----------------------RD-E-------K-----------K--------------- 255
12_BF.AR.99.ARMA159       .......ANNA-ANGTQNNTLEE-P-A-Q-----MT-E--D-QL-VH-L--R---V--G-G.......................SET-R-IN----T--*------WD-------T---Y-L----D-K-------Q-----------K--------------- 269
13_cpx.CM.96.1849         ..............S-INSST-QGRE------Y-MT-GVGDRR--V-SL--RV--E--EGNSNS...................GNSSSSE-R-I--------K------T-------------------DEE-------R---S-------K--------------- 270
14_BG.DE.01.9196_01       ............ESGN-TTNSSSW--M------VTSIT-D----Q--Y--N---VQ--SES......................NS-K-R-I-----I--------T-------------------D-K------------------------------------ 266
15_01B.TH.99.99TH_MU2079  ...............TNINRTI-G----------T----T--.-D--L--S--VEA---.........................---R-RK----I------------------T-------T-----------K------------------------------------ 258
16_A2D.KR.97.97KR004      ...............I-STQSPDSNNTM----ETT-EL-D-K--VQ-L--E---VQLNSSDSN....................DTLNRQ-R-IH-D--T------------------------KDPE-------K---S----------A------------ 267
17_BF.PE.02.PE02_PCR0155  .KSASAKNVST-ASXANDS#TLE-P-A-L-----MT-EV-D-QL-V-SL--X---V--S-N.......................GSE-R-IN----T--------WD----------------D-K-------KD----------K----LIVV.------- 274
18_cpx.CU.99.CU76         ...........G-NTXA-NSS-#ATE-M-----F-T-EV-X-QKQ---L--X---VS-VGXVf....................SNYSD-R-IN--V-T-K--------D------------------QP-------R-----------K--------------T 268
19_cpx.CU.99.CU7          ........................HM------MT-EL-D-K-QI-SL------V--GD...........................-IIN----A--------T--------------------------D-E-----K-R-----------K--------------- 240
20_BG.CU.99.Cu103         ..................STVTSSED-NI---Y-T-E-KDRKKQ---L--R--VV--NDKNSN....................NSNNYSD-R-IN--V-A-K--------D------------------RD-E-------K-----------K--------------- 265
21_A2D.KE.91.KNH1254      ......SVNSTNSSV-TNSTSEL-AAGMQ-----VT-VV-D-QKQ---L--RP--V----N......................DNS--R-IN----A-------I------------------------RD-Q-------K-----------K--------------- 272
22_01A1.CM.01.01CM_0001BBY...............ITNSSVEEL-T-VT-----MT-EL-D-RREV-SL--R---VG-NG........................NN-FR-IN----A--------T----------------------KE-E-------N---S-------K--------------- 260
23_BG.CU.03.CB118         ...........-RNCTTTNGTVNNSAD--------T-E--D-SK----L------M---DKNNG....................SNKNYSD-R-IN--V-T-K--------D------------------RD-E-------R-----------K--------------- 273
24_BG.CU.03.CB378         .........QTK-CTEE-NSTVINCK---------T-ELQD-KKQ---L---I--V--N-SSS....................NNNYSD-R-IN--V-A-K--------D-----------Y-----RDEE---K---R-----------K--------------D 273
24_BG.CU.03.CB471         ...........-RNCTEGINRTTNYED-------AT-E-KDRQKQ---L-----VVQ-NDKDNSS..................NNNNYSD-R-IN--V-T-K--------D------------------RDAE-------K-----------K--------------- 274
25_cpx.CM.01.101BA        INNSNCSHSINNS-EKNNNTIANGCQR--------T-ELKD-KK----L-----VV--EDGN......................NSND--R-IN---A-K------T------------------KD-K-------K-----------K--------------- 282
27_cpx.CD.97.97CDKTB49    ..............-NHTTINDNM---L------VT-A-KD-Q--VH-L-----VVQ-G--TN.....................TTGAE-R-IN------K------T----------------------------------S-------K-------------S- 267
28_BF.BR.99.BREPM12609    .......TNMNATNTTGNSSVEAME--M------VT---GD-M-----L-----VV--A--........................NA--T-IN---------------------------L------------------------------------------- 270
29_BF.BR.02.BREPM119      ............TNTN--DWGR-----M-------T-DL-D-K-----V-----VT--G--........................N---R-I------L------T---------------------D-K---S---------------------------------- 264
31_BC.BR.02.110PA         .ANKTNANQTNA-QT-NNSST-DM----------MT-EL-D-IR-VH-L--R---V-LNENN.....................SSSGD-R-IN----A----------D-----------Y-----K--N--R----N-----------K--I------------ 278
32_06A1.EE.01.EE0369      ......NATINGTSEGTKNITVNN-E-MT-----VT-E-ED-KKT---L-----V---EGKNN...................NGNSSA-I-IN--ATN-E-------------------------RD-N-------K-----------K--------------G 276
33_01B.MY.05.05MYKL007_1  .........NV-RQTNAPSILENVTE--S-----TT-EL-D-K--V--L------VS-GGNK.....................SNSSE-R-IN------K--------D-------T---Y------D-N-------K---S-------K--------------- 271
34_01B.TH.99.OUR2478P     ..................TNGPNITE--R-----MT-E--D-K--V--L-----LVQ-NGS.......................GGE-R-IN------K-----I--D-------T---Y------D-N-------K---S-------K--------------- 263
35_AD.AF.05.05AF095       ....................TENMN----------T-EL-D-R--V-SL-----VVQ-ED-........................NSS--R-IN----A--------T-------------------ENK-----S-K---S-------K--------------D 259
36_cpx.CM.00.00CMNYU830   ..................NVSKDMQ---------TT-VL-D-I--VHSL--R---V---ASN......................NDSSR-R-IN--A-A--------T------------------KD-Q-----L-K---S-------K--------------K 265
37_cpx.CM.00.00CMNYU926   ................#KSLDGG-EI------Y-MT-EL-D-K--I-SL-----VV--GDN.......................NET-R-IN-------------T------Y------------D-EY------R--------------------------- 264
38_BF1.UY.05.UY05_4752    .........SQNG-SN--STLKEDI-V-Q-----MT-EV-D-QL-V-SL--R---V--S-GNNGS...................RSNSSGE-R-IN--S-T---------WD-----------Y-------N--------------------K-M-----------S- 275
39_BF.BR.03.03BRRJ103     .............ND-QNSTLKE-A-ALQ-----MT-E--D-Q--VH-L--R---V--SDNSNG...................SNGSGK-R-IH----T---------WD-----------------D-K-------K-----------K--------------- 270
40_BF.BR.04.04BRRJ115     .........VHN--QGIDSTLKE-PEA-Q-----MT-EL-D-Q--VH-L------V-LNS-GN....................NSSEKY-R-IN----T---------WD-----------Y---------------------------K--------------- 273
42_BF.LU.03.luBF_05_03    ................XXXXXXXXXX--------VTLG-KS-M-----L--RS-VV----NS.......................NS-N-M-IX----XL------I-----------------------D-K-------K------------------------X-- 268
43_02G.SA.03.J11223       ................KTNVKCEDV-NL------VT-EL-D-KK----I-----VV---GNSN.....................NESGD-R-IN--V-T---------T-D-----------X----KE-N-----Q-K---S-------K--------------- 265
U.CD.83.83CD003           ..................TNNNT-EAT-T----KVP-ELKD-TETVHTL-----VV-LNVTNN....................SSISST-R-IN----T--------------------------D-K-------K-------------------------S- 264
U.CD.90.90CD121E12        .....................NTAAN-LT----AVT-ELKD-KKNVH-L-----VV-LNETN.....................SSYGT-R-IN----T--------------------------D-E-------R-------------------------S- 259
U.GR.99.GR303             ...........ALNC-NVNATDEMR--M----Y-MT-EM-D-K--V-SL--R--VVQMGS-NS...................KSNYSE-R-IN-S--A--------T---------T-------T-------R--EEK-------K---S-------K--------------K 271
U.NL.95.U_NL_95_H10986_D1 .............---GNNATVE----L-------T-VLQD-KR----L--RF-VV-LNS-NN....................SSNYSE-R-IN--I-T-K------D-------------------KD-R-------K-----------K--------------- 269
CPZ.CD.90.ANT             ...PTTPSTTTSTVTPKTTTP-VDGMKLQE-N--Q--GFKD-K--MK-I---G-LMKCQDNN.......................E-NC-Y-WH---TT---S-E-ST------------Y---R-EDED-T-V-M-K---V-H-----S-M-A-W-----TYQT 267
CPZ.CM.01.SIVcpzCAM13     .NSTSNSNNNP-LK-NT--KNETGVLQM---T--TT-EL-D-KKQV-SL--VD-LQSLGSG.......................TGDT-TMIN---TA----------------------DVN-S-K-K-R-----H-----K---T---II-----T 277
CPZ.CM.05.SIVcpzEK505     ................NNSVNQTNHVQMQ-----VT-EL-D-KKQV-SL--MG----L-TNNS....................SGNNSQ-R-IN---TAV------I-------Y----P----I---DQD-----E-N-----------K--I----I------T 260
CPZ.CM.05.SIVcpzLB7       ...................NSSINSSNETM------TT-EL-D-KKQT-SL--VD-LTQ-NKT......................-E----IN---TA-------I------------L-R-----Y--S---N-----H-----K--I----I------- 261
CPZ.CM.05.SIVcpzMB66      GNTTNCTVQIS-GND-TANNITVGTIDMY-----AT-EL-DRKKQV-SL--RQ-LE-LEGNKPP...................EGDKNAL-R-YN---TAM----S---L-------------L----D-N-T-I-Q-K-----H------------F----T-E- 291
CPZ.CM.05.SIVcpzMT145     ...............-CTNSTDDRL-DMR-----VT-EL-D-KRQV-SL--VE--TA-G-.........................NST-R-IN---TA-------T----------------L----DIDYK-NET-K-----H-----K--AT---I-----T-D 265
CPZ.GA.88.GAB1            ............QAKNLTNQTSSPPL-M-------VT-EL-D-KKQV-SL--VE-VVNLG-E........................NNT-RIIN---TA-------T--------------------D-D-S-K-K--------H-----K---T----I------- 265
CPZ.TZ.00.TAN1            ...NKTLNSAT-TLTPTVNLSSIPNY-VY-----QT-EF-D-KKQI-SL--RE--VKE-GN........................NN--Y-HN--------E-D-ST------R--------L---RDQN-T-K-Q-S---V-H-----Y-MIA-A-H------E- 266
CPZ.US.85.CPZUS           ...................SSAVNTTDIMR------T-EL-D-KKQV-SL--VD-LAH-N-.........................NT-R-IN---TA-------T-----------P---------E-D-K-K-E-K-----------K---T---II-----T 252
CPZ.CM.x.SIVcpzDP943      .............STGNDT-NSSTSQMV-K-F--MT-EL-D-QKQV-SL--VD--VX-NG.........................XX--R-IN---TA-------T-----------P----M---EPD-----K-R-I----------K---T-X-I-------- 257
CPZ.CM.x.SIVcpzMB897      .....................EASQVQY-----VT-EL-D-KKQV-SL--RE--TSL-SNKT.....................VKNGT-R-IN---TA-------T-------Y-------L----DQN-K-K-T-R-----H-----K------F--------- 260
N.CM.02.DJO0131           ........RTTDEKEKPGNGTDL-ARHM-------T-E-HD-KKQA-SL--VE-VV-LNDG........................NNST-R-IN---TAV------TT----------P----M---EAN-----E-K-----------K--I-----I------DK 263
N.CM.04.04CM_1015_04      .............-XTVDN-ELDIGYKQM------VT-ERKD-KKLA-SL--AE-VVQLNESD....................STNQT-R-I--K-TSV------TT----------P----MX--EGN-S-K-E-K-----------K-TI-----IX-----DT 261
N.CM.04.04CM_1131_03      ............-ETVGN-TLDIGYKQM------VT-EREDRKK-A-SL--TE-VVQLNDSD....................SDSNT-T-IN-K-T-VR-----TT----------P----MR--EGN-S-K-E-K-----------K--I-----IM-----D- 273
N.CM.95.YBF30             .........ERN---MTTREPDIGYKQM------AT-ELTD-KKQV-SL--VE-VV--NAY........................NKT-R-IN---TAV------T-----------P----M---EGN-S-N-S-------------K--I-----I------NT 270
N.CM.97.YBF106            ..........SAGNSTTNRTEDL-DRQM-------T-E--DRKKQV-SL--VE-VV--KDGT......................DNNT-R-IN---TAV------TT----------P----M---EGN-S-N-S-------------K--I-----I------DT 262
O.BE.87.ANT70             .....................TNENLM-K-E--VT-V-KD-KE-KQ-L--VS-LMELNETSST...................NKTNSKM-T--N--STT-----------------------Y--F---STE-----T-R-ITV-T--------T-----I---T-SK 259
O.CM.91.MVP5180           ................KT-LLNETIN-MR-----VT-VLTD-KEQKQ-L--VS-LSKVNDSNA....................VNG-T-M--N--STI-K------------------T-Y--F---DTD-----L-H-I-V-T-----K-T-----I---T-SR 263
O.CM.96.96CMABB637        ......TSNNTPR--TATSTPGI-EMKV-K-E--VT-ILKD-KE-KQ-L--VS-LTKVN-SETG...................NNTEPI-T-IN--STT-K-------N-----------Y--F---DTA-----L-K-I-V-T-----K-T-----I---TIS- 274
O.CM.98.98CMA104          ...............NTNTTTHNPSPDM-Q-N--VT-V-KD-KE-KQ-L--VS-L-RFNETEN...................NKTREI-T-IN--STTT---------------N-----Y--F----IE-----T-N-I-V-T-----K-TI----I---T-SK 264
O.CM.99.99CMU4122         ..................TNATAASSAV-K-E--VT-VLKD-KE-QQ-L--RS-LNK-E-ST.......................TS-M-T-IN--STT-K------T----------G-Y--F---DTE-----Q-R-ITV-T-----K-T-----V---T-S- 258
O.FR.92.VAU               ............I-TTN-PLNSNNTK-V-Q-D--VT-VLKD-QE-KQ-L--VT-LVK-NATS.......................NE-M-R-IN--STT-R------------------------Y--F---ETG-----L-K--TV-T-----K-T-----I---A-SK 264
O.SN.99.SEMP1299          .........NY--NS-INNDTSSPENLV-Q-E--VT-VVKD-KE-KQ-L--VS-LMK-NEAN.....................DTKDM-T-IN--STT-K-----------------------Y--F---STE--------N-ITA-T-----K-T-----I---T-S- 268
O.US.99.99USTWLA          .........TTNT-TAKTTNPEE-TNPV-K-D--VT-VVKD-QE-KQ-L--VS-LLK-G-G........................-NI-T-IN--S-T-K------T-----------------Y--F---ETG---------K-I-V-T-----K-T---H-IF--TISK 265
```

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

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

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

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

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

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

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

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

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

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

```
                                V4                                CD4                                                
```

```
                                V4                                CD4                                                
```

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

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


HIV-1/SIVcpz Proteins

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

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

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

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

HIV-1/SIVcpz Proteins

myristoylation / Nef start; acidic cluster; poly-P helix; phos; phos; HXB2 premature Nef end

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

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

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

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

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

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

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

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

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

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

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

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


HIV-1/SIVcpz Proteins