

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	344
	V-4.9 Nef	356

V-1 Introduction

The HIV-1/SIVcpz alignment contains the same sequences as the complete nucleotide alignment. Because of the abundance of HIV-1 sequences there was no need, and no more space, to include additional sequences in the compendium alignment.

V-2 Annotated Features

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

Feature	Protein	Location	Page
Gag start, p17 start	Gag	1	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	344
signal peptide	Env	30	344
gp120 start	Env	31	344
glycosylation NVT	Env	88-90	344
V1	Env	131-156	344
glycosylation NDT	Env	136-138	344
glycosylation NSS	Env	141-143	346
glycosylation NCS	Env	156-158	346
V2	Env	156-196	346
glycosylation NIS	Env	160-162	346
glycosylation NDT	Env	186-188	346
glycosylation NTS	Env	197-199	346
glycosylation NKT	Env	230-232	346
glycosylation NGT	Env	234-236	346
glycosylation NVS	Env	241-243	346
glycosylation NGS	Env	262-264	346
glycosylation NFT	Env	276-278	346
glycosylation NTS	Env	289-291	348
glycosylation NCT	Env	295-297	348
V3	Env	296-331	348
glycosylation NNT	Env	301-303	348
V3 tip	Env	312-315	348
glycosylation NNT	Env	339-341	348
glycosylation NKT	Env	356-358	348
CD4	Env	368	348
CD4	Env	370	348
V4	Env	385-418	348
glycosylation NST	Env	386-388	348
glycosylation NST	Env	392-394	348
glycosylation NST	Env	397-399	348
glycosylation NNT	Env	406-408	348
CD4	Env	427	350

Feature	Protein	Location	Page
CD4	Env	429	350
CD4	Env	438	350
glycosylation NIT	Env	448-450	350
V5	Env	460-471	350
glycosylation NES	Env	463-465	350
fusion peptide	Env	510-527	350
gp120 end	Env	511	350
gp41 start	Env	512	350
immunodominant region	Env	588-607	352
glycosylation NAS	Env	611-613	352
glycosylation NKS	Env	616-618	352
glycosylation NHT	Env	624-626	352
glycosylation NYT	Env	637-639	352
glycosylation NGS	Env	750-752	354
glycosylation NAT	Env	816-818	354
gp41 end	Env	857	354
Env end	Env	857	354
Nef start	Nef	1	356
myristoylation	Nef	2-7	356
acidic cluster	Nef	62-65	356
poly-P helix	Nef	69-78	356
phos	Nef	77	356
phos	Nef	81	356
HXB2 premature Nef end	Nef	124	356
normal Nef end	Nef	207	358

V-3 Sequences

Sequences included in the HIV-1 protein alignments.

Name	Accession	Proteins	Author	Reference
01_AE.CF.90.90CF11697	AF197340	All	Anderson, JP	<i>J Virol</i> 74 (22); 10752-65 (2000)
01_AE.CN.05.FJ051	DQ859178	All	Yan, Y	<i>Chin Med J(Engl)</i> 119 (19); 1622-8 (2006)
01_AE.CN.06.FJ054	DQ859180	All	Yan, Y	<i>Chin Med J(Engl)</i> 119 (19); 1622-8 (2006)
01_AE.HK.x.HK001	DQ234790	All	Tsui, SKW	Unpublished
01_AE.JP.93.93JP_NH1	AB052995	All	Sato, H	<i>J Virol</i> 75 (12); 5604-13 (2001)
01_AE.TH.01.01TH_R2184	AY945730	All	Watanaveeradej, V	<i>ARHR</i> 22 (8); 801-7 (2006)
01_AE.TH.02.OUR769I	AY358062	All	Tovanabutra, S	<i>ARHR</i> 20 (5); 465-75 (2004)
01_AE.TH.90.CM240	U54771	All	Laukkanen, T	<i>J Virol</i> 70 (9); 5935-43 (1996)
01_AE.US.00.00US_MSC1164	AY444804	All	Tovanabutra, S	<i>ARHR</i> 21 (5); 424-9 (2005)
02_AG.CM.02.02CM_4082STN	AY371141	All	Kijak, GH	<i>ARHR</i> 20 (5); 521-30 (2004)
02_AG.EC.x.ECU41	AY151001	All	Carr, JK	<i>ARHR</i> 19 (4); 329-32 (2003)
02_AG.FR.91.DJ264	AF063224	All	Laukkanen, T	<i>Virology</i> 247 (1); 22-31 (1998)
02_AG.GH.03.GHNJ196	AB231898	All	Tatsumi, M	Unpublished
02_AG.NG.01.PL0710	DQ168577	All	Carr, JK	Unpublished
02_AG.NG.x.IBNG	L39106	All	Howard, TM	<i>ARHR</i> 10 (12); 1755-7 (1994)
02_AG.SE.94.SE7812	AF107770	All	Laukkanen, T	Unpublished
02_AG.SN.98.MP1211	AJ251056	All	Toure-Kane, C	<i>ARHR</i> 16 (6); 603-9 (2000)
02_AG.UZ.02.02UZ710	AY829207	All	Carr, JK	<i>JAIDS</i> 39 (5); 570-5 (2005)

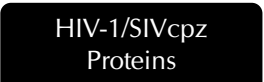
Name	Accession	Proteins	Author	Reference
03_AB.RU.97.KAL153_2	AF193276	All	Liitsola, K	<i>AIDS</i> 12 (14); 1907-19 (1998)
04_cpx.CY.94.CY032	AF049337	All	Gao, F	<i>J Virol</i> 72 (12); 10234-41 (1998)
05_DF.BE.x.VI1310	AF193253	All	Laukkanen, T	<i>Virology</i> 269 (1); 95-104 (2000)
06_cpx.AU.96.BFP90	AF064699	All	Oelrichs, RB	<i>ARHR</i> 14 (16); 1495-500 (1998)
06_cpx.RU.05.04RU001	DQ400856	All	Galkin, AN	Unpublished
07_BC.CN.97.CN54	AX149771	All	Shao, Y	Patent: WO 0136614-A 125 25-MAY-2001; Genearth GMBH Gesellschaft fuer angewandte Boitechnologie(DE) ; Shao, Yiming (CN)
08_BC.CN.97.97CNGX_6F	AY008715	All	Piyasirisilp, S	<i>J Virol</i> 74 (23); 11286-95 (2000)
09_cpx.GH.96.96GH2911	AY093605	All	McCutchan, FE	<i>ARHR</i> 20 (8); 819-26 (2004)
10_CD.TZ.96.96TZ_BF061	AF289548	All	Kouliniska, IN	<i>ARHR</i> 17 (5); 423-31 (2001)
11_cpx.GR.x.GR17	AF179368	All	Paraskevis, D	<i>ARHR</i> 16 (9); 845-55 (2000)
12_BF.AR.99.ARMA159	AF385936	All	Carr, JK	<i>AIDS</i> 15 (15); F41-7 (2001)
13_cpx.CM.96.1849	AF460972	All	Wilbe, K	<i>ARHR</i> 18 (12);849-856 (2002)
14_BG.DE.01.9196_01	AY882421	All	Harris, B	<i>ARHR</i> 21 (7);654-660 (2005)
14_BG.ES.99.X397	AF423756	All	Delgado, E	<i>JAIDS</i> 29 (5); 536-43 (2002)
15_01B.TH.99.99TH_MU2079	AF516184	All	Viputtigul, K	<i>ARHR</i> 18 (16); 1235-7 (2002)
16_A2D.KR.97.97KR004	AF286239	All	Gao, F	<i>ARHR</i> 17 (8); 675-88 (2001)
18_cpx.CU.99.CU76	AY586540	All	Thomson, M	<i>AIDS</i> 19 (11); 1155-63 (2005)
19_cpx.CU.99.CU7	AY894994	All	Casado, G	<i>JAIDS</i> 40 (5); 532-7 (2005)
20_BG.CU.03.CB471	AY900575	All	Perez, L	<i>ARHR</i> 22 (8); 724-33 (2006)
21_A2D.KE.91.KNH1254	AY945737	All	Visawapoka, U	<i>ARHR</i> 22 (7); 695-702 (2006)
23_BG.CU.03.CB118	AY900571	All	Perez, L	<i>ARHR</i> 22 (8); 724-33 (2006)
24_BG.CU.03.CB378	AY900574	All	Perez, L	<i>ARHR</i> 22 (8); 724-33 (2006)
25_cpx.CM.01.101BA	DQ826726	All	Carr, JK	Unpublished
27_cpx.CD.97.97CDKTB49	AJ404325	All	Vidal, N	<i>ARHR</i> 16 (18); 2059-64 (2000)
28_BF.BR.99.BREPM12609	DQ085873	All	De Sa Filho, DJ	<i>ARHR</i> 22 (1); 1-13 (2006)
29_BF.BR.02.BREPM119	AY771590	All	Sa-Filho, DJ	<i>ARHR</i> 21 (2); 145-51 (2005)
31_BC.BR.02.110PA	EF091932	All	Santos, AF	<i>AIDS</i> 20 (16); 2011-9 (2006)
33_01B.MY.05.05MYKL007_1	DQ366659	All	Tee, KK	<i>JAIDS</i> 43 (5);523-529 (2006)
34_01B.TH.99.OUR2478P	EF165541	All	Tovanabutra, S	<i>ARHR</i> 23 (6);829-833 (2007)
35_AD.AF.05.05AF095	EF158041	All	Sanders-Buell, EE	<i>ARHR</i> 23 (6); 834-9 (2007)
36_cpx.CM.00.00CMNYU830	EF087994	All	Powell, RL	<i>ARHR</i> 23 (8); 1008-19 (2007)
37_cpx.CM.00.00CMNYU926	EF116594	All	Powell, RL	<i>ARHR</i> 23 (7); 923-33 (2007)
42_BF.LU.03.luBF_05_03	EU170155	All	Struck, D	Unpublished
A1.GE.99.99GEMZ011	DQ207944	All	Carr, JK	<i>ARHR</i> 22 (5); 470-6 (2006)
A1.KE.00.KER2008	AF457052	All	Dowling, WE	<i>AIDS</i> 16 (13); 1809-20 (2002)
A1.KE.00.KNH1144	AF457066	All	Dowling, WE	<i>AIDS</i> 16 (13); 1809-20 (2002)
A1.KE.00.KSM4024	AF457077	All	Dowling, WE	<i>AIDS</i> 16 (13); 1809-20 (2002)
A1.KE.00.MSA4069	AF457080	All	Dowling, WE	<i>AIDS</i> 16 (13); 1809-20 (2002)
A1.KE.00.NKU3005	AF457089	All	Dowling, WE	<i>AIDS</i> 16 (13); 1809-20 (2002)
A1.RU.00.RU00051	EF545108	All	Thomson, MM	<i>ARHR</i> 23 (12); 1599-604 (2007)
A1.RU.03.03RU20_06_13	AY500393	All	Papuashvili, MN	<i>Infect Genet Evol</i> 5 (1); 45-53 (2005)
A1.RW.93.93RW_024	AY713406	All	Brown, BK	<i>J Virol</i> 79 (10); 6089-101 (2005)
A1.SE.95.SE8891	AF069673	All	Laukkanen, T	<i>AIDS</i> 13 (14); 1819-26 (1999)
A1.SE.95.UGSE8131	AF107771	All	Laukkanen, T	<i>AIDS</i> 13 (14); 1819-26 (1999)
A1.TZ.01.A173	AY253305	All	Herbinger, K-H	<i>ARHR</i> 20 (8); 895-901 (2004)
A1.UA.01.01UADN139	DQ823357	All	Saad, MD	<i>ARHR</i> 22 (8);709-714 (2006)
A1.UG.92.92UG037	AB253429	All	Sakamoto, Y	Unpublished
A1.UG.99.99UGA07072	AF484478	All	Harris, ME	<i>ARHR</i> 18 (17); 1281-90 (2002)

Name	Accession	Proteins	Author	Reference
A1.UZ.02.02UZ0659	AY829209	All	Carr, JK	<i>JAIDS</i> 39 (5); 570-5 (2005)
A2.CD.97.97CDKS10	AF286241	Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Gao, F	<i>ARHR</i> 17 (8); 675-88 (2001)
A2.CD.97.97CDKTB48	AF286238	All	Gao, F	<i>ARHR</i> 17 (8); 675-88 (2001)
A2.CY.94.94CY017_41	AF286237	All	Gao, F	<i>ARHR</i> 17 (8); 675-88 (2001)
B.AR.04.04AR151516	DQ383752	All	Carr, JK	<i>Retrovirology</i> 2006 Sep 7;3:59
B.AU.87.MBC925	AF042101	All	Deacon, NJ	<i>Science</i> 270 (5238); 988-91 (1995)
B.BO.99.BOL0122	AY037270	All	Carr, JK	<i>AIDS</i> 15 (15); F41-7 (2001)
B.BR.03.BREPM2012	EF637046	All	Sa-Filho, D	<i>ARHR</i> 23 (9):1087-1094 (2007)
B.CA.97.CANB3FULL	AY779553	All	Wang, B	<i>ARHR</i> 21 (8); 728-33 (2005)
B.CN.05.05CNHB_hp3	DQ990880	All	Tan, J	<i>Chin Med J(Engl)</i> 120 (9); 831-3 (2007)
B.CO.01.PCM001	AY561236	All		<i>Am J Trop Med Hyg</i> 74 (4); 674-7 (2006)
B.FR.83.HXB2	K03455	All	Wong-Staal, F	<i>Nature</i> 313 (6000); 277-84 (1985)
B.GB.83.CAM1	D10112	All	McIntosh, AAG	PhD dissertation, University of Cambridge (1989)
B.GB.86.GB8	AJ271445	All	Farrar, GH	<i>J Med Virol</i> 34 (2); 104-13 (1991)
B.GE.03.03GEMZ010	DQ207942	All	Carr, JK	<i>ARHR</i> 22 (5); 470-6 (2006)
B.IT.05.SG1	DQ672623	All	Calugi, G	<i>J Virol</i> 80 (23); 11892-6 (2006)
B.JP.05.DR6538	AB287363	All	Sakamoto, Y	Unpublished
B.KR.05.05CSR3	DQ837381	All	Cho, Y	Unpublished
B.NL.00.671_00T36	AY423387	All	Geels, MJ	<i>J Virol</i> 77 (23):12430-12440 (2003)
B.RU.04.04RU128005	AY682547	All		<i>ARHR</i> 22 (11); 1192-7 (2006)
B.TH.00.00TH_C3198	AY945710	All	Watanaveeradej, V	<i>ARHR</i> 22 (8); 801-7 (2006)
B.UA.01.01UAKV167	DQ823362	All	Saad, MD	<i>ARHR</i> 22 (8):709-714 (2006)
B.US.04.ES10_53	EF363127	All	Blankson, JN	<i>J Virol</i> 81 (5):2508-2518 (2007)
B.US.99.PRB959_03	AY331296	All	Bernardin, F	<i>J Virol</i> 79 (17); 11523-8 (2005)
C.AR.01.ARG4006	AY563170	All	Aguayo, N	<i>ARHR</i> 20 (9); 1022-5 (2004)
C.BR.04.04BR013	AY727522	All	Sanabani, SS	<i>ARHR</i> 22 (2); 171-6 (2006)
C.BW.00.00BW07621	AF443088	All	Novitsky, VA	<i>J Virol</i> 76 (11); 5435-51 (2002)
C.CN.98.YNRL9840	AY967806	All	Qiu, Z	<i>ARHR</i> 21 (12); 1051-6 (2005)
C.ET.02.02ET_288	AY713417	All	Brown, BK	<i>J Virol</i> 79 (10); 6089-101 (2005)
C.GE.03.03GEMZ033	DQ207941	All	Carr, JK	<i>ARHR</i> 22 (5); 470-6 (2006)
C.IL.99.99ET7	AY255824	All	Harris, ME	<i>ARHR</i> 19 (12); 1125-33 (2003)
C.IN.99.01IN565_10	AY049708	All	Khurana, S	Unpublished
C.KE.00.KER2010	AF457054	All	Dowling, WE	<i>AIDS</i> 16 (13); 1809-20 (2002)
C.MM.99.mIDU101_3	AB097871	All	Takebe, Y	<i>AIDS</i> 17 (14); 2077-2087 (2003)
C.MW.93.93MW_965	AY713413	All	Brown, BK	<i>J Virol</i> 79 (10); 6089-101 (2005)
C.SN.90.90SE_364	AY713416	All	Brown, BK	<i>J Virol</i> 79 (10); 6089-101 (2005)
C.SO.89.89SM_145	AY713415	All	Brown, BK	<i>J Virol</i> 79 (10); 6089-101 (2005)
C.TZ.02.CO178	AY734556	All	Arroyo, MA	<i>AIDS</i> 19 (14):1517-1524 (2005)
C.UY.01.TRA3011	AY563169	All	Aguayo, N	<i>ARHR</i> 20 (9); 1022-5 (2004)
C.YE.02.02YE511	AY795906	All	Saad, MD	<i>ARHR</i> 21 (7); 644-8 (2005)
C.ZA.04.04ZASK164B1	DQ056405	All	McKay, AR	<i>J Virol Methods</i> 136 (1-2); 118-25 (2006)
C.ZA.05.05ZAPSK240B1	DQ369991	All	McKay, AR	<i>J Virol Methods</i> 136 (1-2); 118-25 (2006)
C.ZM.02.02ZM108	AB254141	All	Tatsumi, M	Unpublished
CPZ.CD.90.ANT	U42720	All	Vanden Haesevelde, MM	<i>Virology</i> 221 (2); 346-50 (1996)

Name	Accession	Proteins	Author	Reference
CPZ.CM.01.SIVcpzCAM13	AY169968	All	Nerrienet, E	<i>J Virol</i> 79 (2); 1312-9 (2005)
CPZ.CM.05.SIVcpzEK505	DQ373065	All	Keele, BF	<i>Science</i> 2006 May 25;
CPZ.CM.05.SIVcpzLB7	DQ373064	All	Keele, BF	<i>Science</i> 2006 May 25;
CPZ.CM.05.SIVcpzMB66	DQ373063	All	Keele, BF	<i>Science</i> 2006 May 25;
CPZ.CM.05.SIVcpzMT145	DQ373066	All	Keele, BF	<i>Science</i> 2006 May 25;
CPZ.GA.88.GAB1	X52154	All	Huet, T	<i>Nature</i> 345 (6273); 356-9 (1990)
CPZ.TZ.01.TAN1	AF447763	All	Santiago, ML	<i>J Virol</i> 77 (3):2233-2242 (2003)
CPZ.US.85.CPZUS	AF103818	All	Gao, F	<i>Nature</i> 397 (6718); 436-41 (1999)
D.CD.83.ELI	K03454	All	Alizon, M	<i>Cell</i> 46 (1); 63-74 (1986)
D.CM.01.01CM_0009BBY	AY371155	All	Kijak, GH	<i>ARHR</i> 20 (5); 521-30 (2004)
D.KE.01.NKU3006	AF457090	All	Dowling, WE	<i>AIDS</i> 16 (13); 1809-20 (2002)
D.KR.04.04KBH8	DQ054367	All	Cho, Y	Unpublished
D.TD.99.MN011	AJ488926	All	Vidal, N	<i>JAIDS</i> 33 (2); 239-46 (2003)
D.TZ.01.A280	AY253311	All	Herbinger, K-H	<i>ARHR</i> 20 (8); 895-901 (2004)
D.UG.99.99UGK09259	AF484498	All	Harris, ME	<i>ARHR</i> 18 (17); 1281-90 (2002)
D.YE.01.01YE386	AY795903	All	Saad, MD	<i>ARHR</i> 21 (7); 644-8 (2005)
D.YE.02.02YE516	AY795907	All	Saad, MD	<i>ARHR</i> 21 (7); 644-8 (2005)
D.ZA.90.R1	EF633445	All	Jacobs, G	<i>ARHR</i> 23 (12); 1575-8 (2007)
F1.AR.02.ARE933	DQ189088	All	Aulicino, PC	<i>ARHR</i> 21 (2):158-164 (2005)
F1.BE.93.VI850	AF077336	All	Laukkanen, T	<i>Virology</i> 269 (1); 95-104 (2000)
F1.BR.01.01BR125	DQ358802	All	Sanabani, S	<i>Infect Genet Evol</i> 6 (5):368-377 (2006)
F1.ES.x.P1146	DQ979023	All	Sierra, M	Unpublished
F1.FI.93.FIN9363	AF075703	All	Laukkanen, T	<i>Virology</i> 269 (1); 95-104 (2000)
F2.CM.02.02CM_0016BBY	AY371158	All	Kijak, GH	<i>ARHR</i> 20 (5); 521-30 (2004)
F2.CM.95.MP257	AJ249237	All	Peeters, M	<i>ARHR</i> 16 (2); 139-51 (2000)
F2.CM.97.CM53657	AF377956	All	Carr, JK	<i>Virology</i> 286 (1); 168-81 (2001)
G.BE.96.DRCBL	AF084936	All	Debyser, Z	<i>ARHR</i> 14 (5); 453-9 (1998)
G.CM.01.01CM_4049HAN	AY371121	All	Kijak, GH	<i>ARHR</i> 20 (5); 521-30 (2004)
G.CU.x.Cu74	AY586547	All		<i>JAIDS</i> 45 (2); 151-60 (2007)
G.ES.00.X558	AF423760	All	Delgado, E	<i>JAIDS</i> 29 (5); 536-43 (2002)
G.ES.99.X138	AF450098	All	Delgado, E	<i>JAIDS</i> 29 (5); 536-43 (2002)
G.GH.03.03GH175G	AB287004	All	Takekawa, N	Unpublished
G.KE.93.HH8793_12_1	AF061641	All	Salminen, MO	<i>ARHR</i> 8 (9); 1733-42 (1992)
G.NG.01.01NGPL0669	DQ168576	All	Carr, JK	Unpublished
G.PT.x.PT2695	AY612637	All	Esteves, AM	Unpublished
G.SE.93.SE6165	AF061642	All	Laukkanen, T	<i>Virology</i> 247 (1); 22-31 (1998)
H.BE.93.VI991	AF190127	All	Janssens, W	<i>AIDS</i> 14 (11); 1533-43 (2000)
H.BE.93.VI997	AF190128	All	Janssens, W	<i>AIDS</i> 14 (11); 1533-43 (2000)
H.CF.90.056	AF005496	All	Murphy, E	<i>ARHR</i> 9 (10); 997-1006 (1993)
J.CD.97.J_97DC_KTB147	EF614151	All	Abecasis, AB	<i>J Virol</i> 81 (16); 8543-51 (2007)
J.SE.93.SE7887	AF082394	All	Laukkanen, T	<i>ARHR</i> 15 (3); 293-7 (1999)
J.SE.94.SE7022	AF082395	All	Laukkanen, T	<i>ARHR</i> 15 (3); 293-7 (1999)
K.CD.97.EQTB11C	AJ249235	All	Peeters, M	<i>ARHR</i> 16 (2); 139-51 (2000)
K.CM.96.MP535	AJ249239	All	Peeters, M	<i>ARHR</i> 16 (2); 139-51 (2000)
N.CM.02.DJO0131	AY532635	All	Bodelle, P	<i>ARHR</i> 20 (8):902-908 (2004)
N.CM.04.04CM_1015_04	DQ017382	All	Yamaguchi, J	<i>ARHR</i> 22 (1):83-92 (2006)
N.CM.04.04CM_1131_03	DQ017383	All	Yamaguchi, J	<i>ARHR</i> 22 (1):83-92 (2006)
N.CM.95.YBF30	AJ006022	All	Simon, F	<i>Nat Med</i> 4 (9); 1032-7 (1998)
N.CM.97.YBF106	AJ271370	All	Ayoub, A	<i>AIDS</i> 14 (16); 2623-5 (2000)
O.BE.87.ANT70	L20587	All	Vanden Haesevelde, M	<i>J Virol</i> 68 (3); 1586-96 (1994)
O.CM.91.MVP5180	L20571	All	Gurtler, LG	<i>J Virol</i> 68 (3); 1581-5 (1994)

Name	Accession	Proteins	Author	Reference
O.CM.96.96CMABB637	AY169810	All	Yamaguchi, J	<i>ARHR</i> 19 (11); 979-88 (2003)
O.CM.98.98CMA104	AY169802	All	Yamaguchi, J	<i>ARHR</i> 19 (11); 979-88 (2003)
O.CM.99.99CMU4122	AY169815	All	Yamaguchi, J	<i>ARHR</i> 19 (11); 979-88 (2003)
O.FR.92.VAU	AF407418	All	Vartanian, JP	<i>J Gen Virol</i> 2002 Apr;83(Pt 4):801-5
O.SN.99.SEMP1299	AJ302646	All	Toure-Kane, C	<i>ARHR</i> 17 (12); 1211-6 (2001)
O.US.99.99USTWLA	AY169814	All	Yamaguchi, J	<i>ARHR</i> 19 (11); 979-88 (2003)

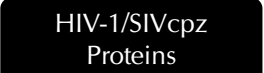
	membrane binding			Gag start, p17 start nuclear localization			phosphorylation site nuclear localization			p17 end p24 start		
B.FR.83.HXB2												
01_AE.CF.90.90CF11697	K	A	RM	L	A	Q	LIE	SN	K	I	K	F
01_AE.CN.05.F1051	K	A	RM	L	A	Q	LIE	SN	K	I	K	F
01_AE.CN.06.F1054	T	K	A	RM	L	A	Q	LIE	ST	K	K	F
01_AE.HK.x.HK001	I	T	E	K	A	RM	L	A	Q	LIE	ST	K
01_AE.JP93.93JP.NH1	K	A	RM	L	A	Q	LIE	SN	K	I	K	F
01_AE.TH.01.01TH.R2184	K	A	RM	L	A	Q	LIE	SN	K	I	K	F
01_AE.TH.02.OUR7691	K	A	R	M	LI	L	A	Q	ME	K	F	
01_AE.TH.90.CM240	K	A	R	R	L	L	SF	A	Q	LIE	ST	
01_AE.US.00.00US_MSC1164	T	K	K	R	L	L	A	Q	LIE	ST	K	
02_AG.CM.02.CM_4082STN	K	A	R	L	L	L	A	Q	LIE	ST	K	
02_AG.EC.x.EC141	K	S	A	R	L	L	A	Q	LIE	ST	K	
02_AG.FR.91.DJ264	K	A	R	L	L	L	A	Q	LIE	ST	K	
02_AG.GH.03.GHNJ196	K	A	R	L	L	L	A	Q	LIE	ST	K	
02_AG.NG.01.PL0710	K	S	R	I	L	L	D	A	Q	LIE	ST	
02_AG.NG.x.IBNG	K	A	R	L	L	L	A	Q	LIE	ST	K	
02_AG.SE.94.SE7812	K	A	R	L	L	L	A	Q	LIE	ST	K	
02_AG.SN.98.MP1211	X	K	A	R	L	L	A	Q	LIE	ST	K	
02_AG.UZ.02.02UZ710	K	A	R	L	L	L	A	Q	LIE	ST	K	
03_AB.RU.97.KAL153_2	K	A	E	RI	L	L	S	A	Q	LIE	ST	
04_cpx.CY.94.CY032	K	A	R	R	L	L	A	Q	LIE	ST	K	
05_DF.BE.x.V11310	K	A	R	R	L	L	A	Q	LIE	ST	K	
06_cpx.AU.96.BFP90	K	E	R	M	L	L	A	Q	LIE	ST	K	
06_cpx.RU.05.04RU001	K	E	R	Q	L	L	A	Q	LIE	ST	K	
07_BC.CN.97.CN54	I	R	K	K	H	M	L	L	A	Q	LIE	
08_BC.CN.97.97CNGX_6F	X	I	R	K	K	H	M	L	L	A	Q	
09_cpx.GH.96.96GH29P1	K	A	R	R	L	L	A	Q	LIE	ST	K	
10_CD.TZ.96.96TZ_BF061	K	E	Q	L	D	I	S	A	Q	LIE	ST	
11_cpx.GR.x.GR17	K	A	R	L	L	L	A	Q	LIE	ST	K	
12_BF.AR.99.ARMA159	K	A	R	R	L	L	A	Q	LIE	ST	K	
13_cpx.CM.96.1849	K	K	R	M	L	L	A	Q	LIE	ST	K	
14_BG.DE.01.9196_01	K	A	R	R	L	L	A	Q	LIE	ST	K	
14_BG.ES.99.X397	K	A	R	R	L	L	A	Q	LIE	ST	K	
15_01B.TH.99.99TH_MU2079	K	A	R	L	L	L	A	Q	LIE	ST	K	
16_A2D.KR.97.97KR004	K	A	R	L	L	L	A	Q	LIE	ST	K	
18_cpx.CU.99.CU76	K	K	R	L	L	L	A	Q	LIE	ST	K	
19_cpx.CU.99.CU7	L	K	A	S	Q	RM	L	L	A	Q		
20_BG.CU.03.CB471	K	A	R	R	L	L	A	Q	LIE	ST	K	
21_A2D.KE.91.KNH1254	K	A	R	L	L	L	A	Q	LIE	ST	K	
23_BG.CU.03.CB118	K	A	R	L	L	L	A	Q	LIE	ST	K	
24_BG.CU.03.CB378	K	S	R	L	L	L	A	Q	LIE	ST	K	
25_cpx.CM.01.101BA	K	A	R	L	L	L	A	Q	LIE	ST	K	
27_cpx.CD.97.97CDKTB49	E	Q	A	H	Q	RI	L	L	A	LD		
28_BF.BR.99.BREPM12609	R	L	R	L	L	L	A	Q	LIE	ST	K	
29_BF.BR.02.BREPM119	X	K	Q	V	L	L	A	Q	LIE	ST	K	
31_BC.BR.02.110PA	I	R	K	T	K	K	Q	MM	L	LD		
33_01B.MY.05.05MYKL007_1	K	A	R	R	L	L	A	Q	LIE	ST	K	
34_01B.TH.99.OUR2478P	K	A	R	S	Q	M	L	L	A	Q		
35_AD.AF.05.05AF095	I	K	A	R	L	L	S	T	Q	ME		
36_cpx.CM.00.00CMNYU830	K	A	R	L	L	L	A	Q	LIE	ST	K	
37_cpx.CM.00.00CMNYU926	T	K	A	R	L	L	A	Q	LIE	ST	K	
42_BFLU.03.luBF_05_03	X	K	Q	A	Y	R	M	L	C	MD		
N.CM.02.DJO0131	T	K	Q	A	Y	R	M	L	C	MD		
N.CM.04.04CM_1015_04	T	K	K	A	Y	R	M	L	C	MD		
N.CM.04.04CM_1131_03	T	K	K	A	Y	R	M	L	C	MD		
N.CM.95.YBF30	T	K	Q	S	Y	R	M	L	C	MD		
N.CM.97.YBF106	T	K	Q	A	Y	R	M	L	C	MD		
O.BE.87.ANT70	S	T	S	K	A	Q	K	S	R	L		
O.CM.91.MVP5180	T	S	K	A	R	S	A	R	L	Y		
O.CM.96.96CMABB637	T	S	K	A	Q	K	S	R	L	D		
O.CM.98.98CMA104	T	S	K	A	Q	K	S	R	L	C		
O.CM.99.99CMU4122	S	K	S	A	Q	K	S	R	L	C		
O.SN.99.SEMPI299	T	S	K	A	Q	K	S	R	L	D		
O.US.99.99USTWLA	I	T	K	T	Q	S	L	K	C	D		
O.FR.92.VAU	T	T	K	A	Q	K	S	LI	G	D		
CPZ.CD.90.ANT	G	R	E	K	T	S	MI	L	RS	Q		
CPZ.CM.01.SIVcpzCAM13	C	M	S	A	L	L	Q	E	A	K		
CPZ.CM.05.SIVcpzEK505	T	K	Q	Y	R	M	L	L	A	Q		
CPZ.CM.05.SIVcpzLB7	I	T	K	D	V	R	M	L	L	A		
CPZ.CM.05.SIVcpzMB66	T	K	Q	Y	R	M	L	L	A	Q		
CPZ.CM.05.SIVcpzMT145	I	T	K	A	V	R	M	L	L	A		
CPZ.GA.88.GAB1	R	D	K	T	S	K	R	LI	GS	Q		
CPZ.TZ.01.TAN1	T	R	A	R	MM	L	L	A	Q	LIE		
CPZ.US.85.CPZUS	T	R	A	R	MM	L	L	A	Q	LIE		



B.FR.83.HXB2
 A1.GE.99.99GEMZ011
 A1.KE.00.KER2008
 A1.KE.00.KNH1144
 A1.KE.00.KSM4024
 A1.KE.00.MSA4069
 A1.KE.00.NKU3005
 A1.RU.00.RU00051
 A1.RU.03.03RU20_06_13
 A1.RW.93.93RW_024
 A1.SE.95.SE8891
 A1.SE.95.UGSE8131
 A1.TZ.01.A173
 A1.UA.01.01UADN139
 A1.UG.92.92UG037
 A1.UG.99.99UGA07072
 A1.UZ.02.02UZ0659
 A2.CD.97.97CDKTB48
 A2.CY.94.94CY017_41
 B.AR.04.04AR151516
 B.AU.87.MBC925
 B.BO.99.BOL0122
 B.BR.03.BREPM2012
 B.CA.97.CANB3FULL
 B.CN.05.05CNHB_hp3
 B.CO.01.PCM001
 B.GB.83.CAM1
 B.GB.86.GB8
 B.GE.03.03GEMZ010
 B.IT.05.SG1
 B.JP.05.DR6538
 B.KR.05.05CSR3
 B.NL.00.671_00736
 B.RU.04.04RU128005
 B.TH.00.00TH_C3198
 B.UA.01.01UAKV167
 B.US.04.ES10_53
 B.US.99.PRB959_03
 C.AR.01.ARG4006
 C.BR.04.04BR013
 C.BW.00.00BW07621
 C.CN.98.YNRL9840
 C.ET.02.02ET_288
 C.GE.03.03GEMZ033
 C.IL.99.99ET7
 C.IN.99.01INS565_10
 C.KE.00.KER2010
 C.MM.99.mIDU101_3
 C.MW.93.93MW_965
 C.SN.90.90SE_364
 C.SO.89.89SM_145
 C.TZ.02.CO178
 C.UY.01.TRA3011
 C.YE.02.02YE511
 C.ZA.04.04ZASK164B1
 C.ZA.05.05ZAPSK240B1
 C.ZM.02.02ZM108
 D.CD.83.EL1
 D.CM.01.01CM_0009BBY
 D.KE.01.NKU3006
 D.KR.04.04KBH8
 D.TD.99.MN011
 D.TZ.01.A280
 D.UG.99.99UGK09259
 D.YE.01.01YE386
 D.YE.02.02YE516
 D.ZA.90.R1
 F1.AR.02.ARE933
 F1.BE.93.VI850
 F1.BR.01.01BR125
 F1.ES.x.P1146
 F1.FI.93.FIN9363
 F2.CM.02.02CM_0016BBY
 F2.CM.95.MP257
 F2.CM.97.CM53657
 G.BE.96.DRCBL
 G.CM.01.01CM_4049HAN
 G.CU.x.Cu74
 G.ES.00.X558
 G.ES.99.X138
 G.GH.03.03GH175G
 G.KE.93.HH8793_12_1
 G.NG.01.01NGPL0669
 G.PT.x.PT2695
 G.SE.93.SE6165
 H.BE.93.VI991
 H.BE.93.VI997
 H.CF.90.056
 J.CD.97.J_97DC_KTB147
 J.SE.93.SE7887
 J.SE.94.SE7022
 K.CD.97.EQTB11C
 K.CM.96.MP535

	CyPA binding										major homology region															
SPRTLNAWVKVVEEKAFSPVPIPMFSALSEGATPQDLNLTMLNTVGGHQAAMQLKETTINEEAAEWDRVHPVHAGTAPGQMPREPGSDIAGTTSTLQEQIGWMT . . . NNPP1.PVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEV	I	I	M	I	D	L	AQ	FP	I	D	P	T	S	D	S	V	K	F	V	T	D	306				
-----I--RG-----M--I-----D-----L--AQ--P--I--D-----P--T-----S-----D-----V--K-----F--V-----T--D-----	I	I	M	I	D	L	AQ	FP	I	D	P	T	S	D	S	V	K	F	V	T	D	311				
-----I-----M--I-----D-----T-----P-----I-----S-----V-----D-----V--K-----F--V-----T--D-----	I	I	M	I	D	T	P	I	S	V	D	V	K	F	V	T	D	313								
-----I-----M--I-----D-----L--AQ--P--I-----P--I-----A-----D-----L--I-----V--K-----F--V-----T--D-----	I	I	M	I	D	L	AQ	FP	I	P	I	A	D	L	I	V	K	F	V	T	D	310				
-----I--R-----D--H--M--I-----E--D-----L--AQ--P--I-----P--A-----G-----D-----L--I-----V--K-----F--A-----T--D-----	I	I	R	D	H	M	I	E	D	L	AQ	P	I	P	A	G	D	L	I	V	K	F	A	T	D	311
-X--X-I--R-----T-----M--I-----D-----L--AQ--FP--L-----S-----D-----V--K-----F--V-----T--D-----	X	X	I	R	T	M	I	D	L	AQ	FP	L	S	D	V	K	F	V	T	D	313					
-----I-----M--I-----D-----L--TQ--FP-----P-----S-----D-----V--K-----F--V-----T--D-----	I	I	M	I	D	L	TQ	FP	P	S	D	V	K	F	V	T	D	313								
-----I-----M--I-----D-----L--V-----P-----S-----D-----V--K-----F--V-----T--D-----	I	I	M	I	D	L	V	P	S	D	V	K	F	V	T	D	312									
-----I--G-----H--M--I-----E--D-----T--I--V-----P--A-----S-----V-----D-----V--K-----F--A-----T--D-----	I	I	G	H	M	I	E	D	T	I	V	P	A	S	V	D	V	K	F	A	T	D	311			
-----I--R-----H--M--I-----E--D-----L--Q--P--I-----P--LQ-----G-----N-----S-----D-----V--K-----F--A-----T--D-----	I	I	R	H	M	I	E	D	L	Q	P	I	P	LQ	G	N	S	D	V	K	F	A	T	D	311	
-----I-----M--I-----D-----L--AQ--FP-----P-----S-----D-----V--K-----F--V-----T--D-----	I	I	M	I	D	L	AQ	FP	P	S	D	V	K	F	V	T	D	311								
-----I--RG-----M--I-----D-----L--V-----P-----NI-----S-----D-----V--K-----E-----F--I-----T-----	I	I	R	G	M	I	D	L	V	P	NI	S	D	V	K	E	F	I	T	313						
-----I-----M--I-----D-----I--AQ--FP-----P-----S-----D-----V--K-----F--V-----T--D-----	I	I	M	I	D	I	AQ	FP	P	S	D	V	K	F	V	T	D	311								
-----I-----M--I-----D-----L--Q--P-----A--N-----S-----D-----V--K-----F--V-----T--D-----	I	I	M	I	D	L	Q	P	A	N	S	D	V	K	F	V	T	D	319							
-----I-----M--I-----D-----L--Q--P-----S-----D-----V--K-----F--V-----T--D-----	I	I	M	I	D	L	Q	P	S	D	V	K	F	V	T	D	309									
-----I-----L--I--V-----S-----D-----V-----X-----V-----F-----T-----D-----	I	L	I	V	S	D	V	X	V	F	T	D	312													
-----I-----L--AQ--I-----T-----S-----D-----V-----F-----T-----D-----	I	L	AQ	I	T	S	D	V	F	T	D	314														
-----L--V-----D-----V-----F-----T-----D-----	L	V	D	V	F	T	D	315																		
-----L--Q--V--L--D-----A--V-----G--M-----G-----F-----T-----D-----	L	Q	V	L	D	A	V	G	M	G	F	T	D	318												
-----L-----H-----F-----T-----D-----	L	H	F	T	D	314																				
-----L-----V-----F-----T-----D-----	L	V	F	T	D	311																				
-----L--V--I-----V--T-----S-----D-----V-----F-----T-----D-----	L	V	I	V	T	S	D	V	F	T	D	313														
-X--X-I-----T-----L-----V-----I-----A-----S-----D-----XS--K--X-----I-----D-----	X	X	I	T	L	V	I	A	S	D	XS	K	X	I	D	314										
-----L-----L-----M-----K-----F-----T-----D-----	L	L	M	K	F	T	D	313																		
-X--I-----T-----L-----I--X-----A-----S--A-----V--K-----X-----T-----	X	I	T	L	I	X	A	S	A	V	K	X	T	316												
-----L-----L-----A-----S--A-----V-----F-----T-----D-----	L	L	A	S	A	V	F	T	D	313																
-----I-----L--AQ--V--I-----T-----D-----	I	L	AQ	V	I	T	D	311																		
-----I-----L--AQ--I-----T-----D-----	I	L	AQ	I	T	D	316																			
-A-----I-----T-----D-----L--X--V-----N--A-----S-----D-----V--K-----F-----T-----D-----	A	I	T	D	L	X	V	N	A	S	D	V	K	F	T	D	309									
-----I-----T-----D-----L-----A-----S--Q-----V-----V--K-----F-----T-----D-----	I	T	D	L	A	S	Q	V	V	K	F	T	D	310												
-----I-----T-----D-----S--A-----G--SV-----D-----V--K-----F-----T-----D-----	I	T	D	S	A	G	SV	D	V	K	F	T	D	310												
-----I--G--N-----T-----D-----I-----Q-----A-----S-----D-----M-----V--K-----F-----T-----D-----	I	I	G	N	T	D	I	Q	A	S	D	M	V	K	F	T	D	309								
-----I-----T-----D-----L-----V-----P--LA-----S--V-----D-----V--K-----F-----T-----D-----	I	T	D	L	V	P	LA	S	V	D	V	K	F	T	D	308										
-----I-----T-----D-----L--P-----V-----VA-----S--V-----D-----V--K-----F--R-----T-----	I	T	D	L	P	V	VA	S	V	D	V	K	F	R	T	319										
-----I-----T-----D-----L-----V-----NP-----S--A-----G--V-----D-----V--K-----F-----T-----D-----	I	T	D	L	V	NP	S	A	G	V	D	V	K	F	T	D	308									
-----I-----T-----D-----L-----V-----N--VA-----G-----D-----V--K-----F-----T-----D-----	I	T	D	L	V	N	VA	G	D	V	K	F	T	D	310											
-----I--G--N-----T-----D-----L-----V-----N--VA-----G-----D-----V--K-----F-----T-----D-----	I	I	G	N	T	D	L	V	N	VA	G	D	V	K	F	T	D	309								
-----I--G--N-----T-----D-----L--Q--V-----N--VA-----G-----D-----V--K-----F-----T-----D-----	I	I	G	N	T	D	L	Q	V	N	VA	G	D	V	K	F	T	D	308							
-----I-----T-----D-----L-----V-----A-----S-----D-----V--K-----F-----T-----D-----	I	T	D	L	V	A	S	D	V	K	F	T	D	308												
-A-----I-----T-----D-----L-----V-----MA-----S--V-----D-----M-----V--K-----F-----T-----D-----	A	I	T	D	L	V	MA	S	V	D	M	V	K	F	T	D	305									
-----I--G--N-----T-----D-----L-----V-----S--A-----G--V-----D-----M-----V--K-----F-----T-----D-----	I	I	G	N	T	D	L	V	S	A	G	V	D	M	V	K	F	T	D	308						
-----I-----I-----T-----D-----L-----V-----S--A-----G--V-----D-----M-----V--K-----F-----T-----D-----	I	I	I	T	D	L	V	S	A	G	V	D	M	V	K	F	T	D	311							
-----I-----V-----D-----L--PQ--V--I-----T-----S-----D-----V--K-----F-----T-----D-----	I	V	D	L	PQ	V	I	T	S	D	V	K	F	T	D	302										
-----I-----L-----V-----MA--I-----G--V-----V-----F-----T-----D-----	I	L	V	MA	I	G	V	V	F	T	D	313														
-----I--N-----T-----A-----L-----I--N-----S-----D-----V--K-----F-----T-----D-----	I	I	N	T	A	L	I	N	S	D	V	K	F	T	D	307										
-----I--N-----T-----S--A-----S-----D-----V--K-----F-----T-----D-----	I	I	N	T	S	A	S	D	V	K	F	T	D	312												
-----I--N-----T-----D-----L-----V-----L-----D-----A-----S-----D-----V--K-----F-----T-----D-----	I	I	N	T	D	L	V	L	D	A	S	D	V	K	F	T	D	313								
-----I-----L-----A-----S-----D-----V--K-----F-----T-----D-----	I	L	A	S	D	V	K	F	T	D	308															
-----I--N-----T-----D-----L-----V-----A-----G-----V-----D-----V--K-----F-----T-----D-----	I	I	N	T	D	L	V	A	G	V	D	V	K	F	T	D	311									
-----I--N-----T-----D-----L--AQ--V-----I-----V-----E--R-----S-----D-----V--K-----F-----T-----D-----	I	I	N	T	D	L	AQ	V	I	V	E	R	S	D	V	K	F	T	D	310						
-----I-----L-----A-----P-----P-----O-----S--V-----D-----V--K-----F-----T-----D-----	I	L	A	P	P	O	S	V	D	V	K	F	T	D	309											
-----I-----L-----P-----N-----O-----S--V-----D-----V--K-----F-----T-----D-----	I	L	P	N	O	S	V	D	V	K	F	T	D	311												
-----I-----L-----P-----O-----S--V-----D-----V--K-----F-----T-----D-----	I	L	P	O	S	V	D	V	K	F	T	D	309													
-----I-----L--Q--P--I-----N--A-----S--V-----D-----V--K-----F-----T-----D-----	I	L	Q	P	I	N	A	S	V	D	V	K	F	T	D	311										
-----I-----L--Q--P--I-----N--A-----S--V-----D-----V--K-----F-----T-----D-----	I	L	Q	P	I	N	A	S	V	D	V	K	F	T	D	307										
-----I-----L--Q--P--I-----D--T-----A-----R-----S--V-----D-----V--K-----F-----T-----D-----	I	L	Q	P	I	D	T	A	R	S	V	D	V	K	F	T	D	306								
T-----L--PQ--P--I-----D--T-----A-----R-----S-----D-----V--K-----F-----T-----D-----	T	L	PQ	P	I	D	T	A	R	S	D	V	K	F	T	D	311									
-----I-----L--PQ--P--I-----D-----V--K-----F-----T-----D-----	I	L	PQ	P	I	D	V	K	F	T	D	313														
-----I-----M--XX--P--I-----N-----R-----S-----D-----V--K-----F-----T-----D-----	I	I	M	XX	P	I	N	R	S	D	V	K	F	T	D	312										
-----I-----L--Q0--P--I-----D--T-----N-----R-----S-----D-----V--K-----F-----T-----D-----	I	L	Q0	P	I	D	T	N	R	S	D	V	K	F	T	D	313									
-----I-----I-----T-----D-----L--Q0--P--I-----N-----R-----S-----D-----V--K-----F-----T-----D-----	I	I	T	D	L	Q0	P	I	N	R	S	D	V	K	F	T	D	311								
-----I-----I-----T-----D-----L--Q0--P--I-----T-----S-----D-----V--K-----F-----T-----D-----	I	I	T	D	L	Q0	P	I	T	S	D	V	K	F	T	D	312									
T-----L-----L-----M--Q0--FP-----I-----S-----D-----V--K-----F-----T-----D-----	T	L	L	M	Q0	FP	I	S	D	V	K	F	T	D	313											
-----I-----L-----VA-----G-----D-----V--K-----F--R-----V-----T-----D-----	I	L	VA	G	D	V	K	F	R	V	T	D	315													
-X-----A-----D-----L-----P-----A-----G--S-----D-----V--K-----F-----T-----D-----	X	A	D	L	P	A	G	S	D	V	K	F	T	D	313											
-----I-----A-----D-----L-----P-----A-----G--A-----D-----V--K-----F-----T-----D-----	I	A	D	L	P	A	G	A	D	V	K	F	T	D	313											
-----I--G--N-----D-----L-----V-----N--A--I--GMGG--Q-----V-----F-----T-----D-----	I	I	G	N	D	L	V	N	A	I	GMGG	Q	V	F	T	D	295									
-----I-----I-----D-----L-----V-----V-----N-----G-----V-----F--A-----T-----D-----	I	I	D	L	V	V	N	G	V	F	A	T	D	313												
-----I-----I-----D-----L-----V-----V-----N-----G-----V-----F--A-----T-----D-----	I	I	D	L	V	V	N	G	V	F	A	T	D	313												
-----I-----M--Q--P--I-----T-----S-----D-----V--K-----F--R-----V-----T-----D-----	I	M	Q	P	I	T	S	D	V	K	F	R	V	T	D	309										
-----I-----L-----P-----I-----A-----S--V-----V-----F-----T-----D-----	I	L	P	I	A	S	V	V	F	T	D	309														

	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	KNMVTEILLVQVANPDKCILKALGPAATLEEMMTACQGVGGPKHARVLAEAMSVQVNTS	... ATIMMQRGNFR	NORKIVKCFNCGKEGHTARNCRAPRKKGGWCKGKEGHQMKDCT	... ERQANFLGKIWPSYK	GRPGNFLQSRPEPTA P	458	
01_AE.CF.90.90CF11697	-----S-----TG-----	-----OHA-----N-----	-----K.G.KR.I-----L-----R-----Q-----	-----LN-----	-----P-----L-----	-----	456	
01_AE.CN.05.F1051	-----ST-----IG-----S-----	-----AQHT-----N-----	-----K.G.KR.I-----L-----L-----E-----	-----N-----	-----N-----	-----P-----	456	
01_AE.CN.06.F1054	-----R-----AG-----S-----	-----S.OHT-----N-----	-----K.G.KR.I-----L-----L-----E-----	-----N-----	-----N-----	-----P-----	457	
01_AE.HK.x.HK001	-----S-----TG-----S-----	-----N.AQQA-----N-----	-----K.G.KR.I-----R-----L-----	-----PN-----	-----P-----S-----	-----	456	
01_AE.JP.93.93JP_NH1	-----S-----TG-----S-----	-----HAOHA-----S-----	-----K.G.KR.I-----R-----L-----	-----N-----	-----P-----	-----	455	
01_AE.TH.01.01TH_R2184	-----S-----A-----S-----	-----AQHA-----N.L-----	-----K.GRKT.I-----R-----L-----	-----L-----N-----	-----P.N-----	-----	454	
01_AE.TH.02.OUR7691	-----S-----TG-----S-----	-----HAQQA-----N-----	-----K.G.KR.I-----L-----L-----	-----N-----	-----P-----	-----	454	
01_AE.TH.90.CM240	-----S-----TG-----S-----	-----HAQQA-----N-----	-----K.G.KR.I-----R-----L-----Q-----	-----N-----	-----P-----	-----	456	
01_AE.US.00.00US_MSC1164	-----S-----TG.S-----S-----	-----H.OHA-----N-----	-----K.G.KR.I-----R-----L-----	-----ERQG-----	-----H-----	-----P-----	456	
02_AG.CM.02.02CM_4082STN	-----R.G-----G-----S-----	-----AAQA-----N-----	-----G.T.I-----L-----K-----K-----R-----	-----N-----	-----P-----	-----	451	
02_AG.EC.x.EC141	-----R-----AG-----S-----	-----AQQA-----N-----	-----G.T.I-----L-----K-----K-----R-----	-----N-----	-----P-----	-----	421	
02_AG.FR.91.DJ264	-----S.R-----G-----S-----	-----QQ-----N-----	-----G.T.I-----L-----K-----K-----R-----	-----S-----	-----P-----	-----	453	
02_AG.GH.03.GHNJ196	-----G-----R-----G-----S-----	-----AQO-----N.V-----	-----G.T.I-----L-----K-----K-----R-----	-----N-----	-----P-----S-----	-----	459	
02_AG.NG.01.PL0710	-----G-----R-----G-----S-----	-----QOT-----N.V-----	-----S.T.I-----L-----K-----K-----R-----	-----P-----	-----P-----	-----	449	
02_AG.NG.x.IBNG	-----S.R-----G-----S-----	-----QQA-----N.V-----	-----G.T.I-----L-----K-----K-----R-----	-----S-----	-----P-----	-----	453	
02_AG.SE.94.SE7812	-----S.R-----G-----S-----	-----QQP-----N-----	-----G.P.I-----L-----K-----K-----R-----	-----S-----	-----P-----	-----	453	
02_AG.SN.98.MP1211	-----S.R-----G-----S-----	-----AQO-----N-----	-----G.T.I-----L-----K-----K-----R-----	-----S-----	-----P-----	-----	453	
02_AG.UZ.02.02UZ710	-----S.R-----G-----S-----	-----QQ-----N-----	-----G.T.I-----L-----K-----K-----R-----	-----S-----	-----P-----	-----	451	
03_AB.RU.97.KAL153_2	-----R.SG-----V-----	-----Q.A-----N.KS-----	-----GPKR.I-----L-----	-----R.S-----	-----P.S-----	-----	456	
04_cpx.CY.94.CY032	-----S-----TG-----S-----	-----AS.AAA.A-----KSK.K.G.RTI-----	-----L-----	-----RM.S-----	-----N-----	-----	459	
05_DF.BE.x.VII310	-----QO-----	-----A.AA.A-----A-----	-----K.GP.I-----I-----K-----G-----R-----	-----V.H-----	-----S-----	-----	460	
06_cpx.AU.96.BFP90	-----D-----R-----G-----	-----ASVG.A-----KS.K.GPKRST-----	-----L-----	-----N-----	-----N-----	-----	456	
06_cpx.RU.05.04RU001	-----G.D-----TG-----S-----	-----S.T.A-----T.S.K.GPKRST-----	-----I-----	-----N-----	-----N-----	-----	459	
07_BC.CN.97.CN54	-----D-----R-----G.SI-----	-----S.K-----K.TNSA-----L.S.K.GSKR-----	-----I-----	-----H.G-----	-----N-----	-----	453	
08_BC.CN.97.97CNGX_6F	-----D-----R-----SG.S-----	-----TN.T-----L.S.K.GSKR-----	-----I.K-----	-----F.H-----	-----X-----	-----S-----	452	
09_cpx.GH.96.96GH2911	-----*-----R-----G-----	-----X.Q.T-----N.L-----	-----K.G.KR.I-----XX-----	-----N-----	-----X-----	-----S-----	455	
10_CD.TZ.96.96TZ_BF061	-----S.R-----VG-----S-----	-----AQHT.N-----K.G.KR.I-----	-----L-----K-----R-----	-----*-----	-----S-----	-----	453	
11_cpx.GR.x.GR17	-----D-----R-----SG-----	-----A.ST.AVL-----KSS.N.G.R-----	-----T.I.K.K-----	-----P-----	-----N-----	-----	453	
12_BF.AR.99.ARMA159	-----D-----R-----G-----	-----ASSAAA.A-----KS.K.GPKR.I-----	-----R.L-----	-----M-----	-----N-----	-----G-----PPESRPEPTA	470	
13_cpx.CM.96.1849	-----S.D-----R-----G-----S.X-----	-----ASGA.L-----KS.K.GP.ROI-----	-----L-----	-----N-----	-----N-----	-----	456	
14_BG.DE.01.9196_01	-----G.D-----R-----G-----	-----ASGA.K-----K.G.RNI-----	-----L-----R-----	-----SK-----	-----N-----	-----N.L-----	456	
14_BG.ES.99.X397	-----A-----S-----SG.S-----S-----	-----H.QOT.NV-----K.G.KR.I-----	-----L-----E-----	-----RL.K-----	-----P.S-----	-----	455	
15_01B.TH.99.99TH_MU2079	-----S.R-----G.S-----S-----	-----Q.T.NSNIM-----K-----	-----H.X.I-----L-----	-----HS-----	-----P.T-----	-----	457	
16_A2D.KR.97.97KR004	-----N.R-----G-----S-----	-----H.O.T.N-----K.G.KR.I-----	-----I-----	-----N-----	-----K-----	-----	456	
18_cpx.CU.99.CU76	-----D-----R-----G-----	-----A.H-----K.K.GL.V-----	-----I-----	-----H-----	-----	-----	459	
19_cpx.CU.99.CU7	-----G.D-----R-----G-----	-----SGT.G.S-----KS.K.GTKRNI-----	-----L-----R-----	-----F.H-----	-----P.N.L-----	-----	459	
20_BG.CU.03.CB471	-----D-----S-----G-----S-----	-----A.S.A.V-----K.GP.I-----	-----I.K-----	-----V.H-----	-----	-----	458	
21_A2D.KE.91.KNH1254	-----G.D-----R-----G.S-----	-----IPGA.G.S-----K.GPKRNI-----	-----D.L-----D-----	-----EG.L-----	-----H-----	-----P.N.L-----	460	
23_BG.CU.03.CB118	-----G.D-----I-----R-----G-----	-----SGT.G.S-----KS-----	-----GP.RNI-----L-----	-----L.H-----	-----P.N.L-----	-----	459	
24_BG.CU.03.CB378	-----G.D-----R-----G.S-----	-----SGT.ATA-----K.GPKRNI-----	-----L-----	-----S-----	-----N-----	-----	458	
25_cpx.CM.01.101BA	-----D-----R-----SG-----S-----	-----ASGAA.A-----KS.K.GP.TI-----	-----L.R-----	-----N-----	-----S-----	-----	459	
27_cpx.CD.97.97CDKTB49	-----G.D-----V-----G-----S-----	-----A.T.V-----K.G-----	-----I.K-----R-----	-----H-----	-----N-----	-----	457	
28_BF.BR.99.BREPM12609	-----D-----R-----G-----S-----	-----HA.T.A.L-----KS.K.G.RV-----	-----I.K-----R-----	-----H-----	-----N-----	-----	446	
29_BF.BR.02.BREPM119	-----D-----R-----G.S-----S-----	-----N.T.N-----S.K.GPKRPI-----	-----I-----E-----	-----L.H-----	-----X-----	-----	455	
31_BC.BR.02.110PA	-----S-----TG-----S-----	-----HAQHT.G-----K.G.KR.I-----	-----L-----	-----L.H-----	-----	-----	456	
33_01B.MY.05.05MYKL007_1	-----S-----TG-----S-----	-----TQAQ.NMM-----K.G.KR.I-----	-----L-----	-----K.I-----F.N-----	-----P-----	-----	455	
34_01B.TH.99.OUR2478P	-----G-----R-----TG-----S-----	-----AQH.TN-----K.K.GHRR.I-----	-----D.L-----R-----	-----H-----	-----P-----	-----	453	
35_AD.AF.05.05AF095	-----S-----R-----TG-----S-----	-----QH.G.I-----S.G.T.I-----	-----L-----	-----N-----	-----S-----	-----	456	
36_cpx.CM.00.00CMNYU830	-----S-----R-----TG-----S-----	-----T.AQO.TS-----S.K.GPKRNI-----	-----L-----	-----R.S-----	-----P-----	-----	453	
37_cpx.CM.00.00CMNYU926	-----G.D-----R-----G-----	-----A.G.TAVL-----S.K.GP-----	-----I.K-----R-----	-----N-----	-----RPEPTA-----	-----	461	
42_BFLU.03.luBF_05_03	-----Q-----	-----OTA.TSVFA-----GR.K.GI.RTIR-----	-----L.K.RG-----Q-----	-----KN.EG-----	-----G.SPF-----	-----P.TTTRKEPTA-----	468	
N.CM.02.DJ00131	-----XO-----S.S-----	-----A.AQAA.TSVFV-----K.GT.RTI-----	-----L.K.RX-----Q-----	-----KN.EG-----	-----G.SPF-----	-----P.T.ATRKEPTA-----	464	
N.CM.04.04CM_1015_04	-----Q-----	-----TA.AQAA.TSVFV-----K.GT.RTI-----	-----L.K.RG-----Q-----	-----KN.EG-----	-----G.SPF-----	-----P.T.TTRKEPTA-----	464	
N.CM.04.04CM_1131_03	-----QL-----G-----	-----A.OQP.TTSVFA-----K.GI.PI-----	-----L.K.RG-----Q-----	-----KN.EG-----	-----S.SPF-----	-----P.T.TTRKEPTA-----	465	
N.CM.95.YBF30	-----Q-----	-----A.AOTA.TSVFV-----K.GI.TI-----	-----L.K.RR-----Q-----	-----KN.EGX-----	-----G.SPF-----	-----P.T.TTRKEPTA-----	464	
N.CM.97.YBF106	-----S-----G-----V-----T-----	-----ATAQD.LKGGYTAVF-----ONP.TRKGTI-----	-----I.K-----R-----Q-----	-----RN.GK-----Y.PGG.T-----	-----YV.RPAH-----	-----	460	
O.BE.87.ANT70	-----X-----Q-----S-----G-----V-----T-----	-----ASAQD.LKGGYTAVF-----ONP.RKGTI-----	-----I.K-----R-----Q-----	-----RN.GK-----Y.PGG.T-----	-----YV.KOVS-----	-----	460	
O.CM.91.MVP5180	-----X-----Q-----S-----G-----V-----T-----	-----I.X.AS.ND.LKGGYTAVF-----ONP.TRORPI-----	-----I-----R-----Q-----	-----RN.GK-----Y.PGG.T-----	-----YV.KOVP-----	-----	459	
O.CM.96.96CMAABB637	-----X-----Q-----S-----G-----V-----T-----	-----ASAQD.LKGGYTAVF-----ONP.TR.GP-----	-----I-----N.OO-----	-----RN.GK-----Y.PGG.T-----	-----M.KQVA-----	-----	460	
O.CM.98.98CMA104	-----Q-----S-----G-----V-----T-----	-----ATAQD.LKGGYTAVF-----ONP.GR.KGPI-----	-----L-----Q-----	-----RN.GK-----Y.PGG.T-----	-----YA.ROVS-----	-----	460	
O.CM.99.99CMU4122	-----S-----Q-----S-----G-----V-----T-----	-----AAAQD.LKGGYTAVF-----ONP.SRKGP-----	-----L-----Q-----	-----KN.G-----Y.PGG.T-----	-----YA.ROVS-----	-----	459	
O.SN.99.SEMP1299	-----S-----Q-----S-----G-----V-----X-----T-----	-----ASAQD.LKGGYTAVF-----ONL.GR.KGP-----	-----V-----Q-----	-----RN.G-----Y.PGG.E-----	-----M.KOMP-----	-----P-----	460	
O.US.99.99USTWLA	-----S-----Q-----S-----G-----V-----T-----	-----ASAQD.LKGGYTAVF-----OKP.TRKGTI-----	-----VTK-----	-----KN.R-----T.K.Y-----	-----PRG.T-----	-----YV.EQVS-----	459	
O.FR.92.VAU	-----A-----I-----H-----TG.S-----L-----A-----	-----ASAN.AO.GTAVFL-----GNRGGKRL-----	-----K.R.R-----R.Q-----L.N.PA.TNTGKV-----	-----PT.TWGC-----	-----V.KEEVPEPTA-----	-----	486	
CPZ.CD.90.ANT	-----T-----D-----Q-----G-----S-----A-----	-----TM.SOART.D.FF.K.PGAT.P.RKI-----Y-----	-----L.K.R-----R-----Q-----S.G.V-----	-----G.RS-----	-----V.N.T-----	-----	465	
CPZ.CM.01.SIVcpzCAM13	-----D-----N-----R-----G.S-----A-----	-----T.AOTA.TSVF-----K.GI.TI-----	-----L.K.R-----Q-----	-----R.SG-----	-----V.LS-----	-----P.TTTRKEPTA-----	465	
CPZ.CM.05.SIVcpzEK505	-----D-----S-----R-----G.S-----A-----	-----ANAA.TV-----K.GPKRVI-----	-----I.K.R-----Q-----R.Q.N-----	-----MGKV-----	-----L.SQ-----	-----	458	
CPZ.CM.05.SIVcpzLB7	-----D-----S-----R-----G.S-----GDEP.F-----S-----	-----AQH-----NDAKRO.K.GPKR-----	-----I.K.R-----Q-----RN-----	-----N-----	-----L.SQ-----	-----YV.K-----	454	
CPZ.CM.05.SIVcpzMB66	-----D-----I-----R-----G-----S-----	-----NKT.L.F-----GNRGR-----	-----L.R-----Q-----N.ONNRI-----	-----N-----	-----PHG-----	-----V.G.K-----	450	
CPZ.CM.05.SIVcpzMT145	-----D-----I-----R-----G-----S-----	-----SMVNOG.R.DVFF.K.OGA.GPKRKT-----	-----L.K.R-----R-----Q-----	-----G.V-----	-----G.RS-----	-----V.N.T-----	468	
CPZ.GA.88.GAB1	-----T-----D-----H-----Q-----G-----L-----S-----KI-----	-----ASA.AGG.VN.L.G.KRPP.LKGGLO-----V-----	-----R-----Q-----	-----TRNNSTGV-----	-----RT.LWG.C-----	-----V.N.T-----	484	
CPZ.TZ.01.TAN1	-----T-----D-----R-----G-----V-----A-----	-----C.MK.P.SSVFL.K.AG.KPKRKI-----	-----L-----R-----Q-----	-----AGN-----	-----H.SPSV.SGSKRPGNFL.NRKEPTA-----	-----	465	



	Vpr binding	Vpr binding p6 end, Gag end
B.FR.83.HXB2	PEE . SFRSGV ETTTPPQ . KQEPIDKEL	YPLTSLRSLFGNDPSSQ*
A1.GE.99.99GEMZ011	-A . H . GM . E I . PSL . KQEQKDREOH	P . SI . K L
A1.KE.00.KER2008	-A . I . GM . E I . S . PKQEQKEREOT	P . FV . K L
A1.KE.00.KNH1144	-A . I . GM . E I . AS . LKQEQKGEOT	P . SI . K L
A1.KE.00.KSM4024	-A . L . GM . E M . AS . LKQEQKGEOT	T . V . K L
A1.KE.00.MSA4069	-A . IYGM . E GI . AS . KQEQDKQEE	A . SI . K LL
A1.KE.00.NKU3005	-A . LLGIKE I . AS . PKQEQNDRGN	P . SV . K LL
A1.RU.00.RU00051	-A . N . GM . E IAPS . LKQEQKDREOP	P . SI . K L
A1.RU.03.03RU20_06_13	-A . X . GMXE I . PSL . KQEQKDREOR	P . SI . K L
A1.RW.93.93RW_024	-A . LLGM . E I . A . PKQEQDREOG	P . V . K L
A1.SE.95.SE889T	-A . I . GM . E I . S . PKQEQKRESE	AP . I . K L
A1.SE.95.UGSE8131	-A . M . M . E I . AS . PKQEQ . NNPP	S . V . K LL
A1.TZ.01.A173	-A . L . GM . E GI . ASLPKQEQKDREOV	P . V . K L
A1.UA.01.01UADN139	-A . N . GM . E I . PSL . KQEQKDREOQ	P . SI . K L
A1.UG.92.92UG037	-AAEI . GMRE I . VS . PKQEQNDROON	P . SV . K LL
A1.UG.99.99UGA07072	-A . L . GM . E M . S . PKQEQKDREOV	P . V . K L
A1.UZ.02.02UZ0659	-A . N . GM . E I . PSL . KQEQKDREOH	P . SI . K L
A2.CD.97.97CDKTB48 MEE I . SSL NREPST AI . K LL
A2.CY.94.94CY017_41	-A . NL . M . E I . SSL LETR . P YN . AI . K LL
B.AR.04.04AR151516 F P . S GT K * . T
B.AU.87.MBC925 F S E A . K
B.BO.99.BOL0122 F F QK . RG YP . D . K * . K
B.BR.03.BREP42012	-A . V . F
B.CA.97.CANB3FULL	-L F SP . S QT . Q
B.CN.05.05CNHB_hp3 F
B.CO.01.PCM001 F V . L T
B.GB.83.CAM1 F K . S
B.GB.86.GB8 F N . P
B.GE.03.03GEMZ010 F A . SP T . D
B.IT.05.SG1 F X M
B.JP.05.DR6538 F I . S M
B.KR.05.05CSR3 F A T . Q
B.NL.00.671_00T36 F K . S R . T
B.RU.04.04RU128005 F K . S VE
B.TH.00.00TH_C3198 F A . S D
B.UA.01.01UAKV167 F
B.US.04.ES10_53 I . GF AAA . S K
B.US.99.PRB959_03 F APA . S T . K
C.AR.01.ARG4006	-A FE SP K . R . P
C.BR.04.04BR013	-A FE SP K . R . P
C.BW.00.00BW07621	-A KR N . TP
C.CN.98.YNRL9840	-A FE AP . SK . R . P
C.ET.02.02ET_288	-A FE A P . LK . R . A
C.GE.03.03GEMZ033	-A YE AP . K . R . P
C.IL.99.99ET	-M FE AP . SK . R . P
C.IN.99.01IN565_10	-A FE TP . K . R . P
C.KE.00.KER2010	-A FE ALAP . LKER . P
C.MM.99.mIDU101_3	-A GF AP . K . R . P
C.MW.93.93MW_965	-A FE AP . KEDRE
C.SN.90.90SE_364	-A FE TP . OK . R . P
C.SO.89.89SM_145	-A FE D . X . AP LKEH
C.TZ.02.CO178	-A FE AP . KEG
C.UY.01.TRA3011	-A FE A . SP . R QK . R . P
C.YE.02.02YE511	-L FE AP . AK . EP
C.ZA.04.04ZASK164B1	-A FE AP . K . R . P
C.ZA.05.05ZAPSK240B1	-L FE TP . LK . G . S
C.ZM.02.02ZM108	-L FE AP . K . R . P
D.CD.83.EL1	-A GF . E I . S
D.CM.01.01CM_0009BBY	-I GF . E V . S
D.KE.01.NKU3006	-A GF . E I . S
D.KR.04.04KBH8	-A GF . E I . S
D.TD.99.MN011	-A GF . E I . S
D.TZ.01.A280	-A . I . GF . E I . K . S
D.UG.99.99UGK09259	-A GF . E I . S
D.YE.01.01YE386	-A GF . E V . S
D.YE.02.02YE516	-A GF . E I . S
D.ZA.90.R1	-A GF . E I . S
F1.AR.02.ARE933	-A GFRE I . SP
F1.BE.93.V1850	-A GFRE I . SP
F1.BR.01.01BR125	-A GF . E I . N . SP
F1.ES.x.P1146	-A GF . E EVTL . H
F1.FI.93.FIN9363	-A . LGFR . E I . SP
F2.CM.02.02CM_0016BBY	-A . G . GF . E I . SP
F2.CM.95.MP257	-A GF . E I . A . SP
F2.CM.97.CM53657	-A GF . E I . SP
G.BE.96.DRCBL	-A . N . GF . E I . A . SP
G.CM.01.01CM_4049HAN	-A GV . E I . A . SP
G.CU.x.Cu74	-A GF . E I . A . SP
G.ES.00.X558	-A SF . E I . A . SP
G.ES.99.X138	-A GF . R I . A . SP
G.GH.03.03GH175G	-A GF . E I . A . SP
G.KE.93.HH8793_12_1	-A GF . E I . A . SP
G.NG.01.01NGPL0669	-A GF . E I . A . S
G.PT.x.PT2695	-A GF . E I . A . SP
G.SE.93.SE6165	-A LGF . E I . A . SP
H.BE.93.V1991	-A GF . E I . SP . R
H.BE.93.V1997	-A GF . E M . SSSP
H.CF.90.056	-A GF . E M . SP
J.CD.97.J_97DC_KTB147	-A GF . E KI . S
J.SE.93.SE7887	-A LGL E . I . SP
J.SE.94.SE7022	-A LGF E . I . SP
K.CD.97.EQTB11C	-A GF . E KI . SL . R
K.CM.96.MP535	-A GF . E I . SP . R

500
492
497
499
496
497
497
500
499
498
498
493
497
497
500
497
497
497
497
497
498
497
504
502
502
503
511
504
502
499
501
508
499
502
501
501
513
501
502
506
504
504
492
502
493
492
498
491
491
501
491
492
486
505
491
488
489
494
491
491
501
497
499
500
503
498
499
496
501
501
495
494
497
493
496
493
499
497
497
500
499
498
497
498
501
500
501
479
498
498
496
496
496

	Vpr binding	Vpr binding p6 end, Gag end	
B.FR.83.HXB2	PEE . SFRSGV ETTTPPQ . KQEPIDKEL YPLTSLRSLFGNDPSSQ*		500
01_AE.CF.90.90CF11697	-A . -LGM-E -I . -SFP -OK - -H PS -V -K L - -		499
01_AE.CN.05.FJ051	-A . -NWEMRE -I . -SM -OK -DH PP -V -K L - -		499
01_AE.CN.06.FJ054	-A . -NWGM-D . GA -SSL -OK -RVH PP -V -K L - -		502
01_AE.HK.x.HK001	-A . -NWGM-E -I . -SLP -OK - -H PS -SI -K L - -		496
01_AE.JP93.93JP.NH1	-A . -NWGM-E -I . -SLP -OK - -H -YP -K L - -		496
01_AE.TH.01.01TH.R2184	-A . -NWGM-E -I . -SSL -OK - -H PP -V -K L - -		497
01_AE.TH.02.OUR7691	-A . -NWGM-E -S-L -OK - -O PP -V -K L - -		497
01_AE.TH.90.CM240	-A . -NWGM-E . EITGE . -I . -SLP -OK - -H PP -V -K L - -		504
01_AE.US.00.00US_MSC1164	-A . -NWGM-E -I . -SLL -OK - -H PP -SV L - -		499
02_AG.CM.02.ZCM.4082STN	-A . - -GM-E -I . -S -OQPRE KGLYP K *		496
02_AG.EC.x.EC041	-A . -V-I-E -I . -SSP . N -TR . -G YP -V -K *		462
02_AG.FR.91.DJ264	-A . -GM-E -I . -SSP . -R-OG YP -A -K *		495
02_AG.GH.03.GHNJ196	-A . -LGM-E -V . ASTP . -G -GI YP K *		501
02_AG.NG.01.PL0710	-A . -LGM-E -I . -SSP . Q -R -G YP K *L -		491
02_AG.NG.x.IBNG	-A . -GM-E -I . P-SP . Q -R -G YP K *L -		495
02_AG.SE.94.SE7812	-A . -LGI-E -I . -SS -G -G YP -A -K *		495
02_AG.SN.98.MP1211	-A . -GM-E -I . -SSP . -G -G YP K Y - -		496
02_AG.UZ.02.02UZ710	-A . -NVGM-E -I . -SSP . Q -R -G YP -A -K *		493
03_AB.RU.97.KAL153_2	-A . -N-GM-E -I . -SL -OK -R -O HP -SI -K D -L - -		499
04_cpx.CY.94.CY032	-A . -CLERKE - -SSL -R -K S -L - -		501
05_DF.BE.x.VI1310	-A . - -F-E -I . ASSP . -OK-EG YP -A -K L - -		503
06_cpx.AU.96.BFP90	-I . - -GF-E -I . A-SP . -SKE . -E KGL . -A -K S - * -		500
06_cpx.RU.05.04RU001	-A . - -FEE -AA-SS . -KE -A -K S - * -		501
07_BC.CN.97.CN54	- - - -F-E -S -K -K -		496
08_BC.CN.97.97CNGX_6F	-A . - -FE - -AP . -K-R-P -K S -L - -		491
09_cpx.GH.96.96GH29P1	-A . - -GMRE -X . -SSX . X-XKDGE P -X -K L - -		497
10_CD.TZ.96.96TZ_BF061	-A . - -GF-E -I . -S . - - -OK - -P H -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-EGQ YP -A -K * - * -		494
13_cpx.CM.96.1849	-A . - -GF-E -I . A-S -KETYP -AA-K S - * - * -		508
14_BG.DE.01.9196_01	-A . - -GF-E -I . A-SP . -KE-S -K S - * -		497
14_BG.ES.99.X397	-A . -LGF-E -I . A-SP . -KE-I -A -K S - * -		497
15_01B.TH.99.99TH_MU2079	-A . -NWGM-E - -SLL -OK-R-H TP -V -K LL - -		498
16_A2D.KR.97.97KR004	-A . -D-GM-E - -L -LKNR-Q HT-AI-K LL - -		500
18_cpx.CU.99.CU76	-A . -LGF-E -I . A-SP . -SKEEKY -FS -K S - - -		497
19_cpx.CU.99.CU7	- - - -GF-E -I . -S . - -DQK -A -K L - -		501
20_BG.CU.03.CB471	-A . - -GF-E -I . A-SP . -OKE-D -I -K S - * -		497
21_A2D.KE.91.KNH1254	-A . - -GF-E -I . -S . -R-KEELY -K L - -		499
23_BG.CU.03.CB118	-A . - -GF-E -I . A-SP . -G-KEQD -I -K S - * -		498
24_BG.CU.03.CB378	-A . -E -I . A-SP . -OKE-D -I -K S - * -		492
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 -K L - -		500
28_BF.BR.99.BREPM12609	-A . - -F-E - -SP . -OKEEG YP -A -K * - -		499
29_BF.BR.02.BREPM119	-A . - -F-E -I . -SP . -EKNEG YP -A -K * - K -		488
31_BC.BR.02.110PA	-A . - -FE - -SP . -OK-R-P -K L - -		494
33_01B.MY.05.05MYKL007_1	- - - -F-E - -V -V -K L - -		499
34_01B.TH.99.OUR2478P	- - - -F-E G . - -S -SA -K LL - -		498
35_AD.AF.05.05AF095	-A . - -GF-E -I . -S . -KQEQKDGELY -A -K -		495
36_cpx.CM.00.00CMNYU830	-A . -N-GI-E -M . SSS . -KQESRDKGLY P K *L -		498
37_cpx.CM.00.00CMNYU926	-A . - -EVKE -V . ASSLKQEGKELYPP -A -K S - * P -		493
42_BFLU.03.luBF_05_03	-A . - -GF-E -I . -SLKQEQKDEGPY P -		504
N.CM.02.DJO0131	-L . -NYG-OEER S-OGKE . -R -NTENS YP L - - -		513
N.CM.04.04CM_1015_04	-L . -YG-OEEGSTO . GKEMQE . -N -KTENS YP L - - -		512
N.CM.04.04CM_1131_03	-L . -YG-OEEERSTO . RKEMQE . -N -KTENT YP L - - -		512
N.CM.95.YBF30	-L . -YGF0EEKSTO . GKEMQE . -N -RTENS YP L - - -		513
N.CM.97.YBF106	-L . -YG-OEEKSTO . GKEMQE . -N -KTETS YP LX - - -		512
O.BE.87.ANT70	SAPP MEEEVKG . QENQEO . GG . PNEL . -FA -K T-Q*-L-Q		501
O.CM.91.MVP5180	SAPP MEEAVKE . QENQSO . GD . QEEL . -FA -K T-Q*-L-Q		501
O.CM.96.96CMABB637	SAPP MEEEVKG . QENONONOGGDODEL . -FA T-Q*-L-Q		501
O.CM.98.98CMA104	SAPP MEEEVKG . QENQEO . -D . SNEL . -FA -K T-Q*-L-Q		501
O.CM.99.99CMU4122	SAPP MEEERKG . PESQEO . GD . PNEL . -FA T-Q*-L-Q		501
O.SN.99.SEMP1299	SAPP M-EEKKG . QENQEO . -D . QNEL . -FA T-Q*-L-Q		500
O.US.99.99USTWLA	SAPP MEEVYK . XNQNQG . -D . SNEL . -FA -K XT-Q*-L-Q		501
O.FR.92.VAU	SAPP MEEEMKG . QE-DONQKG . EQNEL . -FA T-Q*-L-Q		501
CPZ.CD.90.ANT	-I . -IYQEEHK R . -OKGL . -G-EELPP SY -K K-Q*WK		525
CPZ.CM.01.SIVcpzCAM13	-A . -YGYOEK -K . NOEE . QE-GNRLYP -K S - -		507
CPZ.CM.05.SIVcpzEK505	-I . -YGYOEEKTTQ . G . EREE . -EKTESSLY P K S - L - -		512
CPZ.CM.05.SIVcpzLB7	-I . - -GFQPKTT OK . QEDR . EEDKTKENK YP L - - -		504
CPZ.CM.05.SIVcpzMB66	-I . - -GFREEEPA OK . EEVK . SE-TEN-LY PQ K L - - -		500
CPZ.CM.05.SIVcpzMT145	-I . - -GFOE -K . KVKG . EVRES-PP S P S - -		501
CPZ.GA.88.GAB1	-I . -YGYOEE -K . SOEK . EGESSLYP S P K S - -		509
CPZ.TZ.01.TAN1	EKG . KAQEO -AQT . VPVTPAPPLE MTMKGGF . -K -I -S-Q** -		526
CPZ.US.85.CPZUS	-I . -D-GYOE -VTQE . - -GKKEPF Q K S - - -E -		506

	Gag-Pol TF start				Gag-Pol TF end protease start					
B.FR.83.HXB2	FFREDLAF	LQ	GKAREFSS	EQTRANS	PTTR	ELQVWG	RDNNSPS	
A1.GE.99.99GEMZ011	N-Q	RE	K	I	T	WDG	PL	
A1.KE.00.KER2008	N-Q	RE	K	I	D	WDG	PL	
A1.KE.00.KNH1144	N-Q	RE	K	I	D	WDG	PL	
A1.KE.00.KSM4024	N-Q	RE	K	I	S	D	WDG	
A1.KE.00.MSA4069	N-Q	RE	K	I	S	D	WDG	
A1.KE.00.NKU3005	N-Q	RE	K	I	S	D	WDG	
A1.RU.00.RU00051	N-Q	RE	K	I	S	D	WDG	
A1.RU.03.03RU20_06_13	N-Q	RE	K	XXI	X	WDX	PL	
A1.RW.93.93RW_024	N-Q	RE	K	I	S	D	WDG	
A1.SE.95.SE8891	N-Q	RE	K	I	S	D	WDG	
A1.SE.95.UGSE8131	N-Q	RE	K	I	S	D	WDG	
A1.TZ.01.A173	N-Q	RE	K	I	S	D	WDG	
A1.UA.01.01UADN139	N-Q	RE	K	I	S	D	WDG	
A1.UG.92.92UG037	N-Q	RE	K	I	S	D	WDG	
A1.UG.99.99UGA07072	N-Q	RE	K	I	S	D	WDG	
A1.UZ.02.02UZ0659	N-Q	RE	K	I	S	D	WDG	
A2.CD.97.97CDKS10	N-Q	RE	K	I	S	D	WDG	
A2.CD.97.97CDKTB48	N-Q	RE	K	I	S	D	WDG	
A2.CY.94.94CY017_41	N-Q	RE	K	I	S	D	WDG	
B.AR.04.04AR151516	T	P	E	K	P	PRANS	T	
B.AU.87.MBC925	N	P	RE	K	I	G	PL	TE	
B.BO.99.BOI0122	N	P	RE	K	I	G	PL	TE	
B.BR.03.BREPM2012	N	P	E	P	A	QTRANS	G	
B.CA.97.CANB3FULL	N	P	E	P	A	QTRANS	G	
B.CN.05.05CNHB_hp3	N	P	E	P	A	QTRANS	G	
B.CO.01.PCM001	N	P	E	P	A	QTRANS	G	
B.GB.83.CAM1	N	P	E	P	A	QTRANS	G	
B.GB.86.GB8	N	P	E	P	A	QTRANS	G	
B.GE.03.03GEMZ010	N	P	E	P	A	QTRANS	G	
B.IT.05.SG1	N	P	E	P	A	QTRANS	G	
B.JP.05.DR6538	N	P	E	P	A	QTRANS	G	
B.KR.05.05CSR3	N	P	E	P	A	QTRANS	G	
B.NL.00.671.00T36	N	P	E	P	A	QTRANS	G	
B.RU.04.04RU128005	X	P	E	P	A	QTRANS	G	
B.TH.00.00TH_C3198	N	CP	E	P	A	QTRANS	G	
B.UA.01.01UAKV167	N	P	E	P	A	QTRANS	G	
B.US.04.ES10_53	N	P	E	P	A	QTRANS	G	
B.US.99.PRB959_03	N	P	E	P	A	QTRANS	G	
C.AR.01.ARG4006	N	P	E	P	A	QTRANS	G	
C.BR.04.04BR013	N	P	E	P	A	QTRANS	G	
C.BW.00.00BW07621	N	P	E	P	A	QTRANS	G	
C.CN.98.YNRL9840	N	P	E	P	A	QTRANS	G	
C.ET.02.02ET_288	T	P	E	P	A	QTRANS	G	
C.GE.03.03GEMZ033	T	P	E	P	A	QTRANS	G	
C.IL.99.99ET7	T	P	E	P	A	QTRANS	G	
C.IN.99.01INS565_10	T	P	E	P	A	QTRANS	G	
C.KE.00.KER2010	N	P	E	P	A	QTRANS	G	
C.MM.99.mIDU101_3	KT	P	E	P	A	QTRANS	G	
C.MW.93.93MW_965	N	P	E	P	A	QTRANS	G	
C.SN.90.90SE_364	N	CP	E	P	A	QTRANS	G	
C.SO.89.89SMT_145	N	P	E	P	A	QTRANS	G	
C.TZ.02.CO178	N	P	E	P	A	QTRANS	G	
C.UY.01.TRA3011	N	P	E	P	A	QTRANS	G	
C.YE.02.02YE511	N	P	E	P	A	QTRANS	G	
C.ZA.04.04ZASK164B1	N	P	E	P	A	QTRANS	G	
C.ZA.05.05ZAPSK240B1	N	P	E	P	A	QTRANS	G	
C.ZM.02.02ZM108	N	P	E	P	A	QTRANS	G	
D.CD.83.ELI	N	P	E	P	A	QTRANS	G	
D.CM.01.01CM_0009BBY	N	P	E	P	A	QTRANS	G	
D.KE.01.NKU3006	N	P	E	P	A	QTRANS	G	
D.KR.04.04KRBH8	N	P	E	P	A	QTRANS	G	
D.TD.99.MN011	N	P	E	P	A	QTRANS	G	
D.TZ.01.A280	N	P	E	P	A	QTRANS	G	
D.UG.99.99UGK09259	N	P	E	P	A	QTRANS	G	
D.YE.01.01YE386	N	P	E	P	A	QTRANS	G	
D.YE.02.02YE516	N	P	E	P	A	QTRANS	G	
D.ZA.90.R1	N	P	E	P	A	QTRANS	G	
F1.AR.02.ARE933	N	P	E	P	A	QTRANS	G	
F1.BE.93.VI850	N	P	E	P	A	QTRANS	G	
F1.BR.01.01BR125	N	P	E	P	A	QTRANS	G	
F1.ES.x.P1146	N	PE	E	P	A	QTRANS	G	
F1.FL.93.FIN9363	N	P	E	P	A	QTRANS	G	
F2.CM.02.02CM_0016BBY	N	P	RE	W	H	G	AS	RR	
F2.CM.95.MP257	NV	Q	E	K	I	AS	RR	RR	
F2.CM.97.CM53657	N	P	E	P	A	QTRANS	G	
G.BE.96.DRCBL	N	P	E	P	A	QTRANS	G	
G.CM.01.01CM_4049HAN	N	P	E	P	A	QTRANS	G	
G.CU.x.Cu74	N	P	E	P	A	QTRANS	G	
G.ES.00.X558	N	P	E	P	A	QTRANS	G	
G.ES.99.X138	N	P	E	P	A	QTRANS	G	
G.GH.03.03GH175G	N	P	E	P	A	QTRANS	G	
G.KE.93.HH8793_12_1	X	N	P	E	P	A	QTRANS	G
G.NG.01.01NGPL0669	N	P	E	P	A	QTRANS	G	
G.PX.P2693	N	P	E	P	A	QTRANS	G	
G.SE.93.SE6165	N	P	E	P	A	QTRANS	G	
H.BE.93.V1991	N	P	E	P	A	QTRANS	G	
H.BE.93.V1997	N	P	RE	K	P	A	AS	RR	
H.CF.90.056	N	P	RE	K	P	A	AS	RR	
J.CD.97.J_97DC_KTB147	S	Q	E	L	P	A	T	RR	
J.SE.93.SE7887	N	P	RE	K	P	A	T	RR	
J.SE.94.SE7022	N	P	RE	K	P	A	T	RR	
K.CD.97.EQTB11C	V	SO	E	K	I	AS	RR	RR	
K.CM.96.MP535	N	P	E	P	A	QTRANS	G	
MSPGRWPKMIGGIGGF	M	S	L	P	G	R	Q	Y	
QDILIEICGHK	Q	D	I	L	E	I	C	G	
KAIGTVLVGPT	K	A	I	G	T	V	L	V	
PNVI	P	N	V	I	I	

	Gag-Pol TF start	Gag-Pol TF end protease start	
B.FR.83.HXB2	FFREDLAFLO, GKAREFSSSEQTRANS	PTTRR. ELQVWG RDNNNSPS. EAGADROGT VSFNFPQVTLWQRPLVTIKIGQGLKEALLDLDGADDTVLEE. MSLPGRWPKMIGGIGGFKVRQYDQILIEICGHKAIGTVLVGPTPVNI	140
01_AE.CF.90.09CF11697	N-SQ-E	S-GDG-LLP-E	D-IN-KR
01_AE.CN.05.FJ051	N-Q-GK	S-S-K-GNE-PLN-T-E-P	D-IN-K-K
01_AE.CN.06.FJ054	N-Q-E-K	S-S-K-GDRG-LLT-E	D-IN-K-K
01_AE.HK.x.HK001	N-SQ-E	S-G-K-EG-SST-S	D-IN-K-K
01_AE.JP.93.93JP.NH1	N-Q-R-G	S-K-GDG-LLT-E	D-IN-K-K
01_AE.TH.01.01TH.R2184	A-Q-GK	S-K-GDG-LLT-E	D-IN-K-K
01_AE.TH.02.OUR7691	N-Q-G	AS-K-GDG-L-T-E	D-IN-K-K
01_AE.TH.90.CM240	N-Q-G	S-K-GDG-RDNGG-L-T-E	D-IN-K-K
01_AE.US.00.00US.MSC1164	N-Q-GK	S-GDG-LLT-E	D-IN-K-K
02_AG.CM.02.02CM.4082STN	KN-Q-E-K-T	S-WNG-TACTGEG.KGTLS-L-I	D-IN-K-K
02_AG.EC.x.EC041	N-Q-E-K-P-GT	S-S-DEG-G-LL-S-DEG	D-IN-K-K
02_AG.FR.91.DJ264	N-Q-E-K-K-GT	S-WDG-L-TEGP	D-IN-K-K
02_AG.GH.03.GHNJ196	N-Q-E-K-GT	S-GDG-G-LLNP	D-IN-K-K
02_AG.NG.01.PL0710	N-Q-K-G	S-KPGDG-LL-T-TEG-A-IS	D-IN-K-K
02_AG.NG.x.IBNG	N-Q-E-K-GT	S-WDG-T-L-T-TEG-A-IS	D-IN-K-K
02_AG.SE.94.SE7812	N-Q-E-K-GT	S-PWDR-LL-T-TGG-TS-LS	D-IN-K-K
02_AG.SN.98.MP1211	N-Q-E-KL-GT	S-WDG-LLP-TGG	D-IN-K-K
02_AG.UZ.02.02UZ710	N-Q-E-K-T	G-RDGG-LL-A-EG-A-IS-S	D-ID-K-K
03_AB.RU.97.KAL153_2	N-Q-RE-K	S-K-WDG-PLP-T-TEG-AS	D-IN-K-K
04_cpx.CY.94.CY032	NV-Q-RE-K-A	A-G-M-REER-G-LL-TEG	D-IN-K-K
05_DFBEx.VII1310	S-P-E-LPP-G-L	AS-G-SLL-EGR	D-IN-K-K
06_cpx.AU.96.BFP90	N-Q-E	H-RFRR-G-SPLP-T-VEGE-G-KGAI-LSL-I	D-IN-K-K
06_cpx.RU.05.04RU001	N-L-E-L-A	C-R-SPLL-EGK-A-I-LS-I	D-IN-K-K
07_BC.CN.97.CN54	N-LP-R	G-I-N-I	D-LN-K-K
08_BC.CN.97.97CNGX_6F	I-P-E-PP	S-R-G-PS-TE	D-IN-K-K
09_cpx.GH.96.96GH2911	N-Q-RE-L-X-G-IG	XS-LLX.XXXEG-GXA-P-X-I	D-IN-K-K
10_CD.TZ.96.96TZ_BF061	N-Q-R-LP	S-S-R-G-TL-T-E-A-ALS-I	D-IN-K-K
11_cpx.GR.x.GR17	N-Q-E-PT-A	S-R-R-G-PLP-T-OGE-I-Y	D-IN-K-K
12_BF.AR.99.ARMA159	N-Q-E-K-P-A	AS-W-RR-G-LL-E-R-VP-LS-I	D-IN-K-K
13_cpx.CM.96.1849	N-Q-RE	T-PTREQTRANS-S-R-RR-G-SPLP-TEGDIS	D-IN-K-K
14_BG.DE.01.9196_01	N-Q-E-P-A	R-RR-G-SPLP-R-EGK-V-IPL-L-I	D-IN-K-K
14_BG.EE.99.X397	N-Q-E-P-A	R-RR-G-SPLP-R-EGK-D-IPLSL-I	D-IN-K-K
15_01B.TH.99.99TH_MU2079	T-Q	S-GDG-LLT-E	D-IN-K-K
16_A2D.KR.97.97KR004	N-P-RE-N-T	S-G-WNG-G-PLA-EK	D-IN-K-K
18_cpx.CU.99.CU76	N-Q-E-P-KA	SG-R-RR-G-SPLP-RVEGGEI	D-IN-K-K
19_cpx.CU.99.CU7	N-P-E-KLP	R-GN-PL-T-SE	D-IN-K-K
20_BG.CU.03.CB471	I-P-E-A-D	C-R-RR-G-SPLP-R-EGE-L-L-I	D-IN-K-K
21_A2D.KE.91.KNH1254	G-P-E-G-L-A	S-R-PL-T-E-GTV	D-IN-K-K
23_BG.CU.03.CB118	T-P-L-A	C-R-RR-G-SPLP-R-EGA-L-L-I	D-IN-K-K
24_BG.CU.03.CB378	T-P-GKL-A	C-G-SPLP-R-EGE-L-L-I	D-IN-K-K
25_cpx.CM.01.101BA	Q-RE-KL-A-D	C-R-R-G-PLP-I	D-IN-K-K
27_cpx.CD.97.97CDKTB49	N-Q-E-P-A-I	S-R-GG-DPL-T-EG-S	D-IN-K-K
28_BFBR.99.BREPM12609	N-P-E-K-P-A	AS-RR-G-PL-E-G-VP-LS-I	D-IN-K-K
29_BFBR.02.BREPM119	N-P-E-K-P-A	S-R-G-PL-GKER-VP-LS-I	D-IN-K-K
31_BC.BR.02.110PA	N-P-E-X-PP	SG-R-G-PL-EG	D-IN-K-K
33_01B.MY.05.05MYKL007_1	A-P-G-LP-KP	G-S-I-C	D-IN-K-K
34_01B.TH.99.OUR2478P	I-Q-G-P	G-I-T	D-IN-K-K
35_AD.AF.05.05AF095	N-P-E-K	S-R-G-ALPETGAER-GVV	D-IN-K-K
36_cpx.CM.00.00CMNYU830	N-Q-E	S-WDR-VLF-EAGIEGQGTI	D-IN-K-K
37_cpx.CM.00.00CMNYU926	N-Q-RE-K-G	S-RGER-GS-SLL-EAG-QGTIPS	D-IN-K-K
42_BFLU.03.luBF_05_03	N-E-K-PP-A-KARANS	AS-W-RR-G-PL-EAGAER-GSV	D-IN-K-K
N.CM.02.DJ00131	KG-VS-RET-QLPPDNKKAHSAHS	AT-W-S-GE-EHTGEG-EH-ELS	D-IN-K-K
N.CM.04.04CM_1015_04	KG-VS-RET-LPPD-SNKKAHSAHS	AT-W-S-GGEHT-GKGDAG-P-E-ELS	D-IN-K-K
N.CM.04.04CM_1131_03	G-VS-RET-L-PD-SNKKAHSAHS	AT-W-S-GEEHT-EKGDAG-P-E-ELS	D-IN-K-K
N.CM.95.YBF30	E-VS-RET-KLPPD-SNKKAHSAHS	AT-W-S-GEEHT-GEGDAG-P-E-ELS	D-IN-K-K
N.CM.97.YBF106	KG-VS-RET-KLPPD-SNKKAHSAHS	AT-W-S-GEEHT-GKGDAG-P-E-DLS	D-IN-K-K
O.BE.87.ANT70	QT-SGG-HE-QLCA-TST	ISPT-DGGG-EG-TGESGTERG	D-IN-K-K
O.CM.91.MVP5180	V-SGG-HE-QLCA-TSV	ISPT-NGGG-EG-TRESSE-G-SGRA-PTCL	D-IN-K-K
O.CM.96.96CMAABB637	QV-SGR-HE-QLCT-ASA	ISPT-NGGG-EG-ARESESEPRGGPRAL-VCL	D-IN-K-K
O.CM.98.98CMA104	QI-SGG-HE-Q-YA-TSG	ISPS-DGGS-EG-TGEPGTERG	D-IN-K-K
O.CM.99.99CMAU4122	I-SGG-HE-QLCT-TSV	ISSS-DGGG-EG-TRESGTE-G-PERTL-VCL	D-IN-K-K
O.SN.99.SEMP1299	I-SGG-HE-QLCA-TSV	ISPT-D-GNNEG-TRESGTE-G-PERTL-VCL	D-IN-K-K
O.US.99.99USTWLA	V-SGG-RE-Q-YA-ANAPS	ISPT-NGGG-EG-AXXSESG-G-LERTL-VCL	D-IN-K-K
O.FR.92.VAU	K-I-SER-HE-QLCAGTSV	ISPT-NGRDEG-AR-GSESE-GAERAL-VCL	D-IN-K-K
CPZ.CD.90.ANT	KDPHVGVQT-LCA-GSSGANS	S-H-D-SGGAQ-E-SEG-QG-GTTS	D-IN-K-K
CPZ.CM.01.SIVcpzCAM13	G-P-RE-QLCT-D	S-W-PRE-GE-EPRG-TGRGQTIP	D-IN-K-K
CPZ.CM.05.SIVcpzEK505	G-S-RE-K-PPDNKKAHSAHS	SN-W-S-GEDHT-G-REGG-KGEDRELSV	D-IN-K-K
CPZ.CM.05.SIVcpzLBT	T-P-KL-H	AN-W-SAKNH-PE-TGGQ-GR-Q-EQA	D-IN-K-K
CPZ.CM.05.SIVcpzMB66	T-Q	S-N-R-Q-GGTC-PE-GGSE-R-DRE-AV	D-IN-K-K
CPZ.CM.05.SIVcpzMT145	KK-S-G-E-LCAR-EQT	S-YG-I-RIS-GE-DG-GGGE-AA	D-IN-K-K
CPZ.GA.88.GAB1	R-P-RE-QLCA-N-T-G	R-W-P-G-E-EPGE-R-REQTSI	D-IN-K-K
CPZ.TZ.01.TAN1	THPLVG.VQT-LCA-H	D-GSGA-G-STDT-SGANPTTG	D-IN-K-K
CPZ.US.85.CPZUS	T-VPIV.ERGIKETR-LPGKQEGAHSAHS	S-N-GFR-SR-SDTG-T-KGKGAL	D-IN-K-K

	protease end_p66, p51 RT start	M41L	K70R D67N	D110 catalytic site			
B.FR.83.HXB2	IGRNLITQIGCTLNFPISPIETVPVKLKPMDGPKVKQWLPTEEKIKALVEICT	EMKEGKISKIGPENPYNTVPVFAIKKDKTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKSVTVLVDVGDYFVPLD	EDFRKYTAFTIPSINNETPGIRYQYNVLPQGW		308		
A1.GE.99.99GEMZ011	M	T	TD N	D	S	T	308
A1.KE.00.KER2008	M	T	R	Q	I	L	308
A1.KE.00.KNH1144	M	T			I	H	308
A1.KE.00.KSM4024	M	T		Q	I	N	308
A1.KE.00.MSA4069	M	T			I	K	308
A1.KE.00.NKU3005	M	D	T	M	I	S	308
A1.RU.00.RU00051	M	LX	T	A	I	N	308
A1.RU.03.03RU20_06_13	XM	L	T	K	V	X	308
A1.RW.93.93RW_024	M		T		I	X	308
A1.SE.95.SE8891	M		RI	T	R	X	304
A1.SE.95.UGSE8131	M		T		I	N-DR	302
A1.TZ.01.A173	M		R	T	M	I	308
A1.UA.01.01UADN139	M			TD	K	V	308
A1.UG.99.99UGA07072	M	S	RI	T	A	D-R	309
A1.UZ.02.02UZ0659	M	L	T	RI	TK	K	308
A2.CD.97.97CDKS10	M	L	T	ID	K	I-V	308
A2.CD.97.97CDKTB48	M	V-L	T	K		D	301
A2.CY.94.94CY017_41	M	V-L	T	K		A	308
B.AR.04.04AR151516							313
B.AU.87.MBC925							310
B.BO.99.BOI0122							309
B.BR.03.BREP2012			R				312
B.CA.97.CANB3FULL							311
B.CN.05.05CNHB_hp3		L		M		L-D	308
B.CO.01.PCM001							308
B.GB.83.CAM1							308
B.GB.86.GB8							314
B.GE.03.03GEMZ010		L				S	308
B.IT.05.SG1	XE	*-XMV	D	X	*	X	308
B.JP.05.DR6538							304
B.KR.05.05CSR3							308
B.NL.00.671_00T36							308
B.RU.04.04RU128005			S				318
B.TH.00.00TH_C3198							308
B.UA.01.01UAKV167							308
B.US.04.ES10_53							315
B.US.99.PRB959_03	M	L				R	308
C.AR.01.ARG4006	M	L					308
C.BR.04.04BR013	M	L					304
C.BW.00.00BW07621	M	L					312
C.CN.98.YNRL9840	M	L					304
C.ET.02.02ET_288	M	L					304
C.GE.03.03GEMZ033	M	L					310
C.IL.99.99ET7	M	L					304
C.IN.99.01INS565_10	M	L					304
C.KE.00.KER2010	I	L					304
C.MM.99.mIDU101_3	M	L					304
C.MW.93.93MW_965	M	L					304
C.SN.90.90SE_364	M	L					305
C.SO.89.89SMT145	M	L					300
C.TZ.02.CO178	M	L					316
C.UY.01.TRA3011	M	L					304
C.YE.02.02YE511	M	L					304
C.ZA.04.04ZASK164B1	M	L					302
C.ZA.05.05ZAPSK240B1	M	L					304
C.ZM.02.02ZM108	M	L					304
D.CD.83.ELI							299
D.CM.01.01CM_0009BBY							307
D.KE.01.NKU3006							307
D.KR.04.04KBH8							309
D.TD.99.MN011							307
D.TZ.01.A280							309
D.UG.99.99UGK09259							307
D.YE.01.01YE386							307
D.YE.02.02YE516							307
D.ZA.90.R1							309
F1.AR.02.ARE933	F	I	M	H	K	L	307
F1.BE.93.V1850	M						308
F1.BR.01.01BR125	M						308
F1.ES.x.P1146	M						306
F1.FL.93.FIN9363	V	M					307
F2.CM.02.02CM_0016BBY	M						308
F2.CM.95.MP257	M						308
F2.CM.97.CM53657	M						308
G.BE.96.DRCBL	M						305
G.CM.01.01CM_4049HAN	M						307
G.CU.x.C674	M						307
G.ES.00.X558	M						306
G.ES.99.X138	M	K					306
G.GH.03.03GH175G	M						308
G.KE.93.HH8793_12_1	M						307
G.NG.01.01NGPL0669	M						307
G.PX.P12695	M						307
G.SE.93.SE6165	M						305
H.BE.93.V1991	I	M	L	T			306
H.BE.93.V1997	I						306
H.CF.90.056	I						307
J.CD.97.J_97DC_KTB147	M	L					306
J.SE.93.SE7887	M	L					306
J.SE.94.SE7022	M	L					306
K.CD.97.EQTB11C	M						308
K.CM.96.MP535	M						308

	polymerase motif			K219Q T215Y			
B.FR.83.HXB2							478
A1.GE.99.99GEMZ011	C	SK	E	A	MS	F	M-LD
A1.KE.00.KER2008	C	SK	E	A	S	F	K
A1.KE.00.KNH1144	C	SK	E	A	S	F	K
A1.KE.00.KSM4024	C	AK	E	A	N	F	R
A1.KE.00.MSA4069	C	SK	E	A	V	N	F
A1.KE.00.NKU3005	C	SK	E	A	V	N	S
A1.RU.00.RU00051	S	C	R	LK	EMI	DL	A-V
A1.RU.03.03RU20_06_13	S	C	R	LK	EMI	DL	A-V
A1.RW.93.93RW_024	S	C	R	LK	EMI	DL	A-V
A1.SE.95.SE889T	I	SK	E	A	V	N	S
A1.SE.95.UGSE8131	I	SK	E	A	V	N	S
A1.TZ.01.A173	I	SK	E	A	V	N	S
A1.UA.01.01UADN139	S	SK	E	A	V	N	S
A1.UG.99.99UGA07072	S	SK	E	A	V	N	S
A1.UZ.02.02UZ0659	S	SK	E	A	V	N	S
A2.CD.97.97CDKS10	S	SK	E	A	V	N	S
A2.CD.97.97CDKTB48	S	SK	E	A	V	N	S
A2.CY.94.94CY017_41	D	AR	EM	SK	TELI		
B.AR.04.04AR151516							
B.AU.87.MBC925	C	SK	E	A	V	N	S
B.BO.99.BO10122	C	SK	E	A	V	N	S
B.BR.03.BREP2012	C	SK	E	A	V	N	S
B.CA.97.CANB3FULL	S	C	R	LK	EMI	DL	A-V
B.CN.05.05CNHB_hp3	C	SK	E	A	V	N	S
B.CO.01.PCM001	C	SK	E	A	V	N	S
B.GB.83.CAM1	C	SK	E	A	V	N	S
B.GB.86.GB8	C	SK	E	A	V	N	S
B.GE.03.03GEMZ010	C	SK	E	A	V	N	S
B.IT.05.SG1	C	SK	E	A	V	N	S
B.JP.05.DR6538	C	SK	E	A	V	N	S
B.KR.05.05KRSR3	C	SK	E	A	V	N	S
B.NL.00.67L_00T36	C	SK	E	A	V	N	S
B.RU.04.04RU128005	S	C	R	LK	EMI	DL	A-V
B.TH.00.00TH_C3198	S	C	R	LK	EMI	DL	A-V
B.UA.01.01UAKV167	R	SK	E	A	V	N	S
B.US.04.ES10_53	E	L		E	H	K	
B.US.99.PRB959_03	S	X	A	E	E		
C.AR.01.ARG4006	S	X	A	E	E		
C.BR.04.04BR013	R	T	V	R	R	K	FT
C.BW.00.00BW07621	R	T	V	R	R	K	FT
C.CN.98.YNRL9840	R	T	V	R	R	K	FT
C.ET.02.02ET_288	AK	EM		AK	EM		
C.GE.03.03GEMZ033	TR	M		TR	M		
C.IL.99.99ET7	R	A		R	A		
C.IN.99.01INS565_10	A	AK	E	A	V	N	S
C.KE.00.KER2010	A	AK	E	A	V	N	S
C.MM.99.mIDU101_3	R	A		R	A		
C.MW.93.93MW_965	R	A		R	A		
C.SN.90.90SE_364	S	C	R	LK	EMI	DL	A-V
C.SO.89.89SM145	C	SK	E	A	V	N	S
C.TZ.02.CO178	C	SK	E	A	V	N	S
C.UY.01.TRA3011	A	EV		A	V	N	S
C.YE.02.02YE511	A	EV		A	V	N	S
C.ZA.04.04ZASK164B1	A	EV		A	V	N	S
C.ZA.05.05ZAPSK240B1	A	EV		A	V	N	S
C.ZM.02.02ZM108	C	D	AK	EM			
D.CD.83.ELI	EM			EM			
D.CM.01.01CM_0009BBY	C	ELI	S	S	K	E	F
D.KE.01.NKU3006	C	EV		A	V	N	S
D.KR.04.04KBH8	C	EV		A	V	N	S
D.TD.99.MN011	C	D	EL	A	V	N	S
D.TZ.01.A280	C	EV		A	V	N	S
D.UG.99.99UGK09259	C	EV		A	V	N	S
D.YE.01.01YE386	C	EV		A	V	N	S
D.YE.02.02YE516	C	EL		A	V	N	S
D.ZA.90.R1	C	EL		A	V	N	S
F1.AR.02.ARE933	C	TK	M	A	V	N	S
F1.BE.93.V1850	C	TK	M	A	V	N	S
F1.BR.01.01BR125	C	D	AK	G			
F1.EG.X.P1146	C	D	AK	G			
F1.FL.93.FIN9363	C	D	TR	V			
F2.CM.02.02CM_0016BBY	I	AR	E	V			
F2.CM.95.MP257	I	AR	E	V			
F2.CM.97.CM53657	I	AR	E	V			
G.BE.96.DRCBL	V	TK		V	TK		
G.CM.01.01CM_4049HAN	G	AK		G	AK		
G.CI.X.C174	G	AK		G	AK		
G.ES.00.X558	G	IK		G	IK		
G.ES.99.X138	G	IK		G	IK		
G.GH.03.03GH175G	I	TK		I	TK		
G.KE.93.HH8793_12_1	P	IK	EM	A	K	E	F
G.NG.01.01NGPL0669	P	IK	EM	A	K	E	F
G.PX.P2695	P	IK	EM	A	K	E	F
G.SE.93.SE6165	R	AN	EM	A	K	E	F
H.BE.93.V1991	H	EV	I	E	A	F	Q
H.BE.93.V1997	H	EV	I	E	A	F	Q
H.CF.90.056	A	E	EM	A	A	K	F
J.CD.97.J_97DC_KTB147	C	AK	EV	E	K	E	F
J.SE.93.SE7887	C	K	ER	E	R	K	E
J.SE.94.SE7022	C	K	ER	E	R	K	E
K.CD.97.EQTB11C	C	RK	M	L	E	E	F
K.CM.96.MP535	H	IK	EM	P	A	E	K

B.FR.83.HXB2
01_AE.CF.90.90CF11697
01_AE.CN.05.FJ051
01_AE.CN.06.FJ054
01_AE.HK.x.HK001
01_AE.JP.93.93JP.NH1
01_AE.TH.01.01TH.R2184
01_AE.TH.02.OUR7691
01_AE.TH.90.CM240
01_AE.US.00.00US.MSC1164
02_AG.CM.02.CM.4082STN
02_AG.EC.x.EC041
02_AG.FR.91.DJ264
02_AG.GH.03.GHNJ196
02_AG.NG.01.PL0710
02_AG.NG.x.IBNG
02_AG.SE.94.SE7812
02_AG.SN.98.MP1211
02_AG.UZ.02.02UZ710
03_AB.RU.97.KAL153_2
04_cpx.CY.94.CY032
05_DF.BE.x.VII1310
06_cpx.AU.96.BFP90
06_cpx.RU.05.04RU001
07_BC.CN.97.CN54
08_BC.CN.97.97CNGX_6F
09_cpx.GH.96.96GH2911
10_CD.TZ.96.96TZ_BF061
11_cpx.GR.x.GR17
12_BF.AR.99.ARMA159
13_cpx.CM.96.1849
14_BG.DE.01.9196_01
14_BG.ES.99.X397
15_01B.TH.99.99TH_MU2079
16_A2D.KR.97.97KR004
18_cpx.CU.99.CU76
19_cpx.CU.99.CU7
20_BG.CU.03.CB471
21_A2D.KE.91.KNH1254
23_BG.CU.03.CB118
24_BG.CU.03.CB378
25_cpx.CM.01.101BA
27_cpx.CD.97.97CDKTB49
28_BF.BR.99.BREPM12609
29_BF.BR.02.BREPM119
31_BC.BR.02.110PA
33_01B.MY.05.05MYKL007_1
34_01B.TH.99.OUR2478P
35_AD.AF.05.05AF095
36_cpx.CM.00.00CMNYU830
37_cpx.CM.00.00CMNYU926
42_BFLU.03.luBF_05_03
N.CM.02.DJO0131
N.CM.04.04CM_1015_04
N.CM.04.04CM_1131_03
N.CM.95.YBF30
N.CM.97.YBF106
O.BE.87.ANT70
O.CM.91.MVP5180
O.CM.96.96CMAABB637
O.CM.98.98CMA104
O.CM.99.99CMU4122
O.SN.99.SEMP1299
O.US.99.99USTWLA
O.FR.92.VAU
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzLBT
CPZ.CM.05.SIVcpzMB66
CPZ.CM.05.SIVcpzMT145
CPZ.GA.88.GAB1
CPZ.TZ.01.TAN1
CPZ.US.85.CPZUS

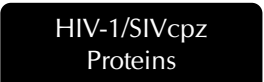
Table with columns for sequence identifiers and amino acid sequences. The sequences are aligned in columns, with gaps represented by dashes. The table is organized into two main sections: the top section contains sequences from B.FR.83.HXB2 to CPZ.US.85.CPZUS, and the bottom section contains sequences from O.BE.87.ANT70 to CPZ.US.85.CPZUS. The sequences are aligned to a common reference sequence, with gaps indicating insertions or deletions. The table is organized into two main sections: the top section contains sequences from B.FR.83.HXB2 to CPZ.US.85.CPZUS, and the bottom section contains sequences from O.BE.87.ANT70 to CPZ.US.85.CPZUS. The sequences are aligned to a common reference sequence, with gaps indicating insertions or deletions.

p51 RT end_p15 RNase H start

B.FR.83.HXB2	AYFLKLA.GRW.PVKTIHTDNGSNFTGATVRAACWAGIK.QEFGIPYNPQSQGVVSMNKLKIIQVQRDQAEHLKTAVMQVFIHFNKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYRDSRNLPLWKGPALKLLWKGEGAVVIQDNSDIKVVPRRK	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.....I.....P.....L.....D-I.....	979
01_AE.CN.06.FJ054-A-V.....S-A-K..NVR.....I.....L.....D-I.....	981
01_AE.HK.x.HK001-V.....S-A-K..NVR.....D.....I.....D-I.....	976
01_AE.JP93.93JP.NH1-V.....S-A-K..NVR.....I.....I.....D-I.....	977
01_AE.TH.01.01TH.R2184-V.....S-A-K..NVR.....I.....I.....R.....D-I.....	980
01_AE.TH.02.OUR7691-V.....S-A-K..NV.....I.....I.....D-I.....	979
01_AE.TH.90.CM240-V.....S-A-K..NVQ.....I.....I.....D-I.....	984
01_AE.US.00.00US_MSC1164-V.....S-A-K..NV.....E.....E.....I.....D-I.....	979
02_AG.CM.02.02CM.4082STN	---I---...-V.....IS-A-K..DVR.....I.....I.....S.....I.....D-I.....	982
02_AG.EC.x.EC141	---I---...-V.....S-A-K..NVT.....I.....S.....I.....I.....D-I.....	979
02_AG.FR.91.DJ264	---I---...-V.....S-A-K..NVT.....I.....S.....I.....I.....D-I.....	979
02_AG.GH.03.GHNJ196	---I---...-RV.....S-A-K..DV.....I.....S.....I.....I.....D-I.....	980
02_AG.NG.01.PL0710	---I---...-RV.....S-A-K..N-T.....I.....S.....I.....I.....D-I.....	979
02_AG.NG.x.IBNG	---I---...-V.....S-A-K..NVT.....I.....S.....I.....I.....D-I.....	979
02_AG.SE.94.SE7812	---I---...-V.....S-A-K..NVT.....I.....S.....I.....I.....D-I.....	979
02_AG.SN.98.MP1211	---I---...-GKMAS..-VV.....S-A-K..KVTX..L..X.....L.....R.....I.....S.....N-K.....D-I.....	980
02_AG.UZ.02.02UZ710	---I---...-V.....STA-K..NVT.....I.....S.....I.....I.....D-I.....	979
03_AB.RU.97.KAL153_2	---V---...-I.....STA-K.....Q-QT.....I.....I.....I.....D-I.....	979
04_cpx.CY.94.CY032	---I---...-M..A..P..S-A-K..D-N.....I.....S.....I.....E-I.....	978
05_DF.BE.x.VII310	---I---...-M.....P..SGA-K.....-Q.....G.....I..T.....I.....D-V.....	979
06_cpx.AU.96.BFP90	---I---...-V.....P..S-A-K..N-T.....E.....I.....S.....I.....D-I.....	981
06_cpx.RU.05.04RU001	---I---...-V.....S-A-K..NVT.....I.....S.....I.....I.....D-I.....	979
07_BC.CN.97.CN54	---I---...-V.....STA-K.....Q.....S.....A.....L.....L.....I.....D-S.....	978
08_BC.CN.97.97CNGX_6F	---I---...-V.....S-A-K..Q.....L.....L.....R.....I.....D-I.....	975
09_cpx.GH.96.96GH29PI	---I---...-V.....S-A-K..NVT.....I.....R.....I.....S.....R.....D-I.....	978
10_CD.TZ.96.96TZ_BF061	---I---...-VV.....S-A-K.....Q.....I.....I.....I.....G-I.....	978
11_cpx.GR.x.GR17	---I---...-V.....SNA-K.....Q.....G..E.....E.....G.....I.....L.....K.....D-I.....	978
12_BF.AR.99.ARMA159	---I---...-P.....IS-M-K.....S.....I.....S.....R.....I.....D-V.....	979
13_cpx.CM.96.1849	---I---...-V.....S-A-K..N-T.....E.....I.....L.....I.....D-I.....	986
14_BG.DE.01.9196_01	---I---...-I.....ISN-K.....X.....T.....I.....S.....N.....F.....D-I.....	978
14_BG.ES.99.X397	---I---...-I.....S-A-K..N-T.....I.....S.....N.....F.....D-I.....	978
15_01B.TH.99.99TH_MU2079	---I---...-V.....S-AMK..NVR.....WD.....I.....I..E.....P.....S.....I..L.....P.....PN..S.....D-I.....	977
16_A2D.KR.97.97KR004	---I---...-V.....P..IS-P-K..VQ.....I.....I.....S.....R.....D-I.....	979
18_cpx.CU.99.CU76	---L---...-V.....P..S-A-K..N-T.....I.....I.....SE..F.....D-I.....	977
19_cpx.CU.99.CU7	---I---...A...-V.....P..S-A-K..NVQ.....I.....S.....S.....D-I.....	978
20_BG.CU.03.CB471	---I---...-V.....P..S-A-K..S-T.....I.....SS.....D-I.....	975
21_A2D.KE.91.KNH1254	---WI---...-V.....P..ISTP-K.....D.....I.....R.....D-I.....	977
23_BG.CU.03.CB118	---I---...-V..X..P..IS-A-K..D-T.....I.....I.....D-I.....	975
24_BG.CU.03.CB378	---I---...-V.....P..IS-A-K..D-T.....I.....S.....F.....D-I.....	970
25_cpx.CM.01.101BA	---I---...-V.....P..IS-A-K.....E.....I.....I.....I.....D-I.....	978
27_cpx.CD.97.97CDKTB49	---I---...-V.....R..STA-K..NVT.....I.....S.....I.....D-I.....	977
28_BF.BR.99.BREPM12609	---I---...-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.....	979
31_BC.BR.02.110PA	---I---...-RV.....SN-K.....Q.....I.....I.....I.....D-I.....	975
33_01B.MY.05.05MYKL007_1	---I---...-V.....S-A-K..NVQ.....E.....L.....D-I.....	979
34_01B.TH.99.OUR2478P	---I---...-V.....S-A-K..NVR.....I.....I.....D-I.....	979
35_AD.AF.05.05AF095	---I---...-VV.....S-AFK.....E.....I.....D-I.....	978
36_cpx.CM.00.00CMNYU830	---I---...-V.....S-A-K..D-Q.....I.....S.....D-I.....	979
37_cpx.CM.00.00CMNYU926	---I---...-RV.....S-A-K..NV.....I.....S.....D-I.....	978
42_BFLU.03.luBF_05_03	---I---...-V.....X..XSTV-K.....E.....I.....D-V.....	983
N.CM.02.DJO0131	---I---...-V.....P..S-K..N-T.....I.....I.....E..T.....I.....TK..T..L.....D-I.....	984
N.CM.04.04CM_1015_04	---I---...-V.....P..S-K..S.....A.....I.....I.....TK..T..L..V.....D-I.....	986
N.CM.04.04CM_1131_03	---I---...-V.....P..S-K..S.....A.....I.....I.....TK..T..L..V.....D-I.....	986
N.CM.95.YBF30	---I---...-V.....S-K..N.....A.....I.....I.....TK..T..L..V.....D-I.....	986
N.CM.97.YBF106	---I---...-V.....P..IS-K.....A.....I.....I.....TK..T..L..V..X.....D-I.....	986
O.BE.87.ANT70	---I---...A...-V.....P..ST-MK..N-Q..H.....A.....S..Q.....R.....V.....I.....L..S0..T.....L..XHK.....D-I.....	975
O.CM.91.MVP5180	---I---...A...-V.....P..S-AMK..N-Q..H.....A.....S..Q.....R.....V.....I.....L..S0..T.....L..N.....D-I.....	975
O.CM.96.96CMABB637	---I---...A...-V.....P..S-MK..N-R..H.....A.....S..Q.....R.....V.....CT..L..ML..S0..T.....H..L..K..Q.....D-I.....	977
O.CM.98.98CMA104	---I---...A...-V.....P..S-MK..N-X..H.....A.....S..E.....R.....I.....L..L..S0..T.....L.....D-I.....	975
O.CM.99.99CMU4122	---I---...A...-V.....P..S-MK..TN..H.....A.....S..Q.....R.....V.....Y.....I.....L..S0..T.....F..K..Q.....D-I.....	975
O.SN.99.SEMP1299	---I---...A...-V.....P..S-AMK..N..H.....A.....S..Q.....R.....V.....Y.....I.....L..S0..T.....F..K..Q.....D-I.....	975
O.US.99.99USTWLA	---I---...A...-T...-RP..S-MK..N..H..X.....X..A.....S..Q.....R.....V.....Y.....I.....L..S0..T.....F..K..Q.....D-I.....	978
O.FR.92.VAU	---I---...A...-I..IL..P..S-AMK..N-Q..H.....A.....ES..MQ.....V.....V.....I.....L..S0..T.....L.....D-I.....	976
CPZ.CD.90.ANT	---I---...S.....A.....S-A-K..N-Q.....Q-Q.....I.....Q.....V..H.....TP-Q..L..L.....TQ..N..L..Q..H.....D-V.....	979
CPZ.CM.01.SIVcpzCAM13	---I---...S.....VL..T..SNA-K..N..K.....E.....I.....S.....I.....L..K.....D-I.....	978
CPZ.CM.05.SIVcpzEK505	---I---...S.....V.....T..S-K..N-Q.....A.....I.....I.....TK..T..L..V.....E-T.....	985
CPZ.CM.05.SIVcpzLB7	---I---...S.....I.....STA-K..Q-Q.....I.....I.....T.....S..V..K.....R..D-I.....	982
CPZ.CM.05.SIVcpzMB66	---I---...V.....V.....SS-K..Q-Q.....I.....I.....L.....TK..Q..SN-K.....D-I.....	982
CPZ.CM.05.SIVcpzMT145	---I---...V.....P.....SNA-K.....D.....E.....I.....S.....L..Q.....D-I.....	978
CPZ.GA.88.GAB1	---I---...V.....P..S-A-K..D.....L.....Q-Q.....E.....Q..R.....I..Y.....E..T.....LL..LT..N..Q.....L..V.....A..D-I.....	977
CPZ.TZ.01.TAN1	---I---...V.....K.....P..STA-K..Q-Q..H.....Q-Q.....E.....Q..R.....I..Y.....E..T.....LL..LT..N..Q.....L..V.....A..D-I.....	975
CPZ.US.85.CPZUS	---I---...V.....V.....SS..Q.....I.....I.....SEL..DL..L..V.....D-I.....	981

	p31 Integrase end Pol end ¹	
B.FR.83.HXB2	AKIIRDYGKQMGDDCVASRODED*	1003
A1.GE.99.99GEMZ011	-----G-----	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	-----G-----	1004
A1.RU.03.03RU20_06_13	-----G-----	1004
A1.RW.93.93RW_024	V-----G-----	1003
A1.SE.95.SE889T	-----G-----	1004
A1.SE.95.UGSE8131	-----G-----	1000
A1.TZ.01.A173	-----G-----	998
A1.UA.01.01UADN139	-----G-----	1004
A1.UG.92.92UG037	V---K-----G-----	1004
A1.UG.99.99UGA07072	-----G-----	1005
A1.UZ.02.02UZ0659	-----G-----E-----	1004
A2.CD.97.97CDKS10	-----G-----	403
A2.CD.97.97CDKTB48	-----G-----	998
A2.CY.94.94CY017_41	-----G-----	1004
B.AR.04.04AR151516	-----G-----	1009
B.AU.87.MBC925	-----G-----	1006
B.BO.99.BOL0122	-----G-----	1005
B.BR.03.BREP2012	V---K-----G-----	1008
B.CA.97.CANB3FULL	-----G-----	1007
B.CN.05.05CNHB_hp3	-----G-----	1004
B.CO.01.PCM001	-----G-----	1004
B.GB.83.CAM1	-----G-----	1004
B.GB.86.GB8	-----G-----	1010
B.GE.03.03GEMZ010	-----G-----	1004
B.IT.05.SG1	-----G-----	997
B.JP.05.DR6538	V-----G-----	1004
B.KR.05.05CSR3	---F-----G-----	1004
B.NL.00.671_00T36	-----G-----	1004
B.RU.04.04RU128005	-----G-----	1004
B.TH.00.00TH_C3198	-----G-----	1004
B.UA.01.01UAKV167	-----G-----	1011
B.US.04.ES10_53	-----G-----	1004
B.US.99.PRB959_03	-----G-----	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-----A-----G-----N-----	1000
C.GE.03.03GEMZ033	V-----A-----G-----	1006
C.IL.99.99ET7	---K-----A-----G-----	1000
C.IN.99.01INS565_10	---K-----A-----G-----	1000
C.KE.00.KER2010	---K-----A-----G-----	1000
C.MM.99.mIDU101_3	---K-----A-----G-----	1000
C.MW.93.93MW_965	V---K-----A-----G-----	1000
C.SN.90.90SE_364	V-----A-----G-----	1001
C.SO.89.89SM_145	V-----A-----G-----	996
C.TZ.02.CO178	V---K-----A-----G-----	1012
C.UY.01.TRA3011	V---K-----A-----M-----N-----N-----	1000
C.YE.02.02YE511	V---K-----A-----G-----	1000
C.ZA.04.04ZASK164B1	---K-----A-----G-----Q-----	998
C.ZA.05.05ZAPSK240B1	---K-----A-----G-----Q-----	1000
C.ZM.02.02ZM108	V-----A-----G-----Q-----	1000
D.CD.83.ELI	V-----G-----	995
D.CM.01.01CM_0009BBY	-----G-----	1003
D.KE.01.NKU3006	-----G-----Q-----	1005
D.KR.04.04KBH8	-----G-----	1003
D.TD.99.MN011	-----G-----	1003
D.TZ.01.A280	-----V-----G-----	1005
D.UG.99.99UGK09259	-----G-----	1005
D.YE.01.01YE386	-----G-----	1003
D.YE.02.02YE516	V-----G-----	1005
D.ZA.90.R1	V-----G-----	1003
FI.AR.02.ARE933	-----M-----G-----	1004
FI.BE.93.V1850	-----V-----G-----	1005
FI.BR.01.01BR125	-----G-----	1004
FI.ES.x.P1146	-----G-----N-----	1002
FI.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-----	1001
G.CM.01.01CM_4049HAN	---K-----G-----G-----	1003
G.CU.x.Cu74	-----A-----G-----	1000
G.ES.00.X558	-----G-----	1002
G.ES.99.X138	-----G-----	1002
G.GH.03.03GH175G	---K-----G-----	1004
G.KE.93.HH8793_12_1	-----G-----	1003
G.NG.01.01NGPLD669	-----G-----	1003
G.PL.x.PT695	-----G-----	1003
G.SE.93.SE6165	-----G-----	1001
H.BE.93.V1991	-----G-----	1002
H.BE.93.V1997	-----G-----	1001
H.CF.90.056	-----G-----	1003
J.CD.97.J_97DC_KTB147	-----G-----	1002
J.SE.93.SE7887	-----G-----	1002
J.SE.94.SE7022	-----G-----	1002
K.CD.97.EQTB11C	-----G-----	1004
K.CM.96.MP535	-----G-----	1004

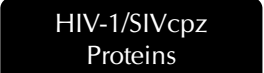
	p31 Integrase end Pol end ¹	
B.FR.83.HXB2	AKIIRDYGKQMAGDDCVASRODED*	
01_AE.CF.90.90CF11697	-----G-----N-	1003
01_AE.CN.05.F1051	-----G-----	1001
01_AE.CN.06.F1054	-----G-----	1004
01_AE.HK.x.HK001	-----G-----	1006
01_AE.JP93.93JP_NH1	-----G-----	1001
01_AE.TH.01.01TH_R2184	-----G-----	1002
01_AE.TH.02.OUR7691	-----G-----	1005
01_AE.TH.90.CM240	-----G-----	1004
01_AE.US.00.00US_MSC1164	-----G-----	1009
02_AG.CM.02.02CM_4082STN	-----G-----	1004
02_AG.EC.x.EC141	V-----G-----N-	1007
02_AG.FR.91.DJ264	V--K-----G-----	1004
02_AG.GH.03.GHNJ196	--K-----G-----	1004
02_AG.NG.01.PL0710	-----G-----	1005
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--	1004
02_AG.UZ.02.02UZ710	--K-----G---N-	1005
03_AB.RU.97.KAL153_2	-----G-----	1004
04_cpx.CY.94.CY032	-----N---G-----	1004
05_DF.BE.x.VII1310	-----G-----	1003
06_cpx.AU.96.BFP90	-----G-----	1004
06_cpx.RU.05.04RU001	-----G-----	1006
07_BC.CN.97.CN54	--K-----A---G-----	1004
08_BC.CN.97.97CNGX_6F	--K-----A---G-----	1003
09_cpx.GH.96.96GH29Π	--H-----G-----	1000
10_CD.TZ.96.96TZ_BF061	V--K-----A-----Q	1003
11_cpx.GR.x.GR17	-----GS-----	1003
12_BF.AR.99.ARMA159	-----G-----	1003
13_cpx.CM.96.1849	-----G-----	1004
14_BG.DE.01.9196_01	-----G-----	1011
14_BG.ES.99.X397	-----G-----	1003
15_01B.TH.99.99TH_MU2079	-----G-----	1002
16_A2D.KR.97.97KR004	-----G-----	1003
18_cpx.CU.99.CU76	-----G-----	1004
19_cpx.CU.99.CU7	-----G-----	1002
20_BG.CU.03.CB471	-----G-----	1003
21_A2D.KE.91.KNH1254	--K-----G----N--	1000
23_BG.CU.03.CB118	-----G-----	1000
24_BG.CU.03.CB378	-----G---N-	1002
25_cpx.CM.01.101BA	-----M-G-----	995
27_cpx.CD.97.97CDKTB49	--K-----G-----	1003
28_BF.BR.99.BREPM12609	-----G-----	1002
29_BF.BR.02.BREPM119	-----G-----	1004
31_BC.BR.02.110PA	V--K-----A--M-----	1004
33_01B.MY.05.05MYKL007_1	-----G-----	1000
34_01B.TH.99.OUR2478P	-----G-----	1004
35_AD.AF.05.05AF095	-----G-----	1004
36_cpx.CM.00.00CMNYU830	--N-----G-----	1003
37_cpx.CM.00.00CMNYU926	-----M-G-----	1004
42_BFLU.03.luBF_05_03	-----G-----	1003
N.CM.02.DJO0131	-----G---G-A-NO	1008
N.CM.04.04CM_1015_04	-----G---G---NO	1009
N.CM.04.04CM_1131_03	-----G---G---NO	1011
N.CM.95.YBF30	-----G---G---NO	1011
N.CM.97.YBF106	-----G---GX-NO	1011
O.BE.87.ANT70	----E-----T-SM--G-T-SF	1000
O.CM.91.MVP5180	----L-----T-SM-N--T-SEE	1000
O.CM.96.96CMA104	--L-----A-SM-G-TXSEE	1002
O.CM.98.98CMA104	--H-----A-SM-G-T-SEE	1000
O.CM.99.99CMU4122	--N-----T-SM-G-X-SEE	1000
O.SN.99.SEMP1299	--H-----T-SM-G-T-SEE	1000
O.US.99.99USTWLA	----T-----M-G-T-SF	1003
O.FR.92.VAU	----A-----S--G-T-GE	1001
CPZ.CD.90.ANT	----KE-R-KIEDR-DL-G-N--	1004
CPZ.CM.01.SIVcpzCAM13	-----G-----	1003
CPZ.CM.05.SIVcpzEK505	-----G---G---SQ	1010
CPZ.CM.05.SIVcpzLB7	-----E-AN-L-D-----N-	1007
CPZ.CM.05.SIVcpzMB66	-----A-S-----NQ	1007
CPZ.CM.05.SIVcpzMT145	-----A-S-----S-	1003
CPZ.GA.88.GAB1	-----T-----M-G-T-SF	1002
CPZ.TZ.01.TAN1	----KE-----AGGMDD--N-T-	1000
CPZ.US.85.CPZUS	----K-----A-SM-G---S-	1000



B.FR.83.HXB2
01_AE.CF.90.90CF11697
01_AE.CN.05.F1054
01_AE.HK.x.HK001
01_AE.JP.93.93JP.NH1
01_AE.TH.01.01TH.R2184
01_AE.TH.02.OUR7691
01_AE.TH.90.CM240
01_AE.US.00.00US.MSC1164
02_AG.CM.02.02CM.4082STN
02_AG.EC.x.EC041
02_AG.FR.91.DJ264
02_AG.GH.03.GHNJ196
02_AG.NG.01.PL0710
02_AG.NG.x.IBNG
02_AG.SE.94.SE7812
02_AG.SN.98.MP1211
02_AG.UZ.02.02UZ710
03_AB.RU.97.KAL153_2
04_cpx.CY.94.CY032
05_DF.BE.x.VII310
06_cpx.AU.96.BFP90
06_cpx.RU.05.04RU001
07_BC.CN.97.CN54
08_BC.CN.97.97CNGX_6F
09_cpx.GH.96.96GH2911
10_CD.TZ.96.96TZ_BF061
11_cpx.GR.x.GR17
12_BF.AR.99.ARMA159
13_cpx.CM.96.1849
14_BG.DE.01.9196_01
14_BG.ES.99.X397
15_01B.TH.99.99TH_MU2079
16_A2D.KR.97.97KR004
18_cpx.CU.99.CU76
19_cpx.CU.99.CU7
20_BG.CU.03.CB471
21_A2D.KE.91.KNH1254
23_BG.CU.03.CB118
24_BG.CU.03.CB378
25_cpx.CM.01.101BA
27_cpx.CD.97.97CDKTB49
28_BF.BR.99.BREPM12609
29_BF.BR.02.BREPM119
31_BC.BR.02.110PA
33_01B.MY.05.05MY.KL007_1
34_01B.TH.99.OUR2478P
35_AD.AF.05.05AF095
36_cpx.CM.00.00CMNYU830
37_cpx.CM.00.00CMNYU926
42_BFLU.03.luBF_05_03
N.CM.02.DJ0131
N.CM.04.04CM_1015_04
N.CM.04.04CM_1131_03
N.CM.95.YBF30
N.CM.97.YBF106
O.BE.87.ANT70
O.CM.91.MVP5180
O.CM.96.96CMA.BB637
O.CM.98.98CMA104
O.CM.99.99CMU4122
O.SN.99.SEMP1299
O.US.99.99USTWLA
O.FR.92.VAU
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpz.CAM13
CPZ.CM.05.SIVcpz.EK505
CPZ.CM.05.SIVcpz.LB7
CPZ.CM.05.SIVcpz.MB66
CPZ.CM.05.SIVcpz.MT145
CPZ.GA.88.GAB1
CPZ.TZ.01.TAN1
CPZ.US.85.CPZUS

Vif start

MENRWQVMIVQVDRMIRITWKSLVKHHMYVSGKARGVFRYHHYSPHPRISSSEVHIPLGD...ARLVITTYWGLHTGERDWHLGGQVSIIEWRKKRYSQVDPPELADQLIHYDFDCFSDSAIRKALLGHIVSPCEYQAGHN.KVGSLOYLALALITPK...KIKPP
Y I K KN Q KV I
N I K K S Q KV Y E GK R O K H KORQ N I D 00 R I QV RS D PS K T TR
N K K Y O KV R O K O H ORN I D V R O A R I QV RYK PS K T IG GSR
N I I K KR O KV F E I E I I K R O K O H ORK I E D 0 R I QV RR PS KV T R R
N K K QN KV K R O K H ORE I D R H I QVDRRK PS S Q K T R R
N I K K S Q KV K R O O H ORN I D K IO S I QV RS PS K T R R
N I K K QN KV I R O O K H ORN I D K O G I QV RR PS K T R R
N Y R K F CK KA V R N H F KO K I D H H H E I QV K Q K V R
N Y K K R V R N H KOR I D H H H H E I V R K K V R K
N I R KN F R KV I VR R A Y H KORQ I D H T I QV S Q K V V
T K N K KD F R KA E VR N A H OR I D H E I V QV R S S K V A T RR
N I K Q X R V VR N A H O I D H H E I V EL R K E T
N K K F R KVC VR H KO I D N I E R N V A T RT
N Y K R KD S V VR H KOR I D H H E I Q R T K V T RT
N I K HR R KV T I VR QP H V QO I D H E I QV R K K V T RR
N I KR V R KV I VR QK H QO D H H E I E R D K VP T RKR
N I I K V K RN K K V A E N E N I R S R R R T
A K N K K R R KV E I VR QP Q H LR D D MH E I R S T
N Y H K NR C F RN K V R S V E N I R F T T
N Y HI K KR N D N K S E N K LR R I SM IH E R D T K E R
N Y HI K KN D R K C E N E LR K I S MH E I I R K K R
L K N RR N D R KV K K O R H LR T I E G M A D I I D N T K
L K N RR N D R KV E K O R H A LRS N I E G M A I V I D N T K R
N Y HT K KR D N K A E V K K L R S I H E I I DR T N R T
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
A N K L R V EE DM V E R R R MH AE I R P R K V RT
N Y H K KR F R V EK E QP OG R I G R I E I KI N T A T
N Y H K GR D N K K V K LR G G TH I R D T H
N R K F G KV K VR K H A QRS H D ET N I R S R T T
N I K I F G KV K VR K H OR S I D H EL V T A
N K K Q KV Q EE K Q K H ORG I R Q E R I QV RYS D PS K T R R
N K D C IN V G K I I V K P K OE I D GH H I K R K VGET RT
N I K KN R KV T V V KE H LR R D GH H E N I R T
N I K NK V RN KV K VK K H OR I D H H QV R K V T RSR
K N I K K G F RN V N K I V K K Q H KORE I D QH K I EQIR S K V I T
V N I R KD V Q R V EE K I VR S K H O G I D H H E I RL K S K V A RT
H X K I K C K TN E K I K K G Q E S I R S K V T
N I I K KD C F RN KV E K I VR K H KORE I D QH K EK R S H S K V T
NK Y H K K K N E AV E K H LR R G R MH T I R S K V A T
N Y I K R D N V V E LR R G R TH E I YR K R T
N I R K K V K T V K VI E D G R TH E N I R R D T T TK
N Y IHI K K L NR K EE K I Q OG G SH N E N I R RS T T L R
L X N X RR S X R KX K K O H LRS D MH A I R R D X T K
V N K I K L Q KV I KIA R O K H O K I D N R I QV RY PS KV T Q R R
N I K K Q KV E R O K Q H ORN I D R Q E R I QV RR PS K T R R
N I K T V K R KV K I VR K H L I DV H N VI QV S Q K V T
N HI KT KD R KV H KORQ I D H H I Q D P K V T R
N HI K K RD V E W R QP H LRK R I D H T I R R P S K VP T RKR
N K K Y T K X K X V T QP OGK M K C N X E N TI R 0 T X
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 0 T H I WVG
V I N H K K Y R KT V Q T V T OP H LRK K M K T Q V RO F P 0 DT T WGI
V K I N H K K Y TR K V Q V T OP H LRK K M K TA Q I RO F P 0 T T WGI
V K K N K Y TH K V Q TV T OS H LRK K M K H TA Q V RP L P K O T T WGA
V X K N K K Y TK KT X V X X V T QP OGK M K C N X E N TI R 0 T X
L I QKVK 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 QV LTK PT S O T L R VVKAR SR
O C M 91 M V P 5 1 8 0 Q K V K N Y K M K A N R RN KV A Y V A E D I V MP E E H O Y E K I T RM H T TE I QV LTK L S O T F K VVKY RR
O C M 96 96 C M A B B 6 3 7 L Q X V K N X Y K X K Q K Y TRN G Y V T N I V MP E H O Y K I T R I H T I G R I QV LTK L S O T L R VVKD RS
O C M 98 98 C M A 1 0 4 L Q K V K N Y K R R K D Y X RN V G Y V Q C I V MP E D O H O Y K R I T R M T E T QV LTK D P E S T L R VVVKT RN
O C M 99 99 C M U 4 1 2 2 L X O K V K N Y K H R T K E H K F F RN A G Y V M C V V MP E H O Y N T I T R I QH T TE I I QV LTK L S O T L R VVVKE RN
O S N 99 S E M P 1 2 9 9 L Q V K N Y K R R T D Y F RN V G Y V G P M I V MP E H Y K K I T R M I A TE I QV LTK L S O T L R VVVKE RN
O U S 99 99 U S T W L A L Q K V K N Y K R K K D Y TRN G Y V G M I V MP E H E N K I T R M H T T T QV LTK L S O T L R VVVGE TSR
O F R 92 V A U L Q K V K N Y K R R T D Y K IN V R Y V P C I M V MP E E Q H Q C R I I T K H T TE R T QV LTK L S O T L R VVKA RNR
CPZ.CD.90.ANT I K KIEMI NI IWETKVKP K ND KKG E PTLDPK V VF QC P H CGK I T M Q Q V Q I ERLITF H K K S O T F FCKILEFRGY PKGPRRQ
CPZ.CM.01.SIVcpz.CAM13 K KA N Y KTK KD H N V FK N SK I A T A H Q MGS V YV K SQ A R I E A R E R O T F K VKE RLR
CPZ.CM.05.SIVcpz.EK505 V V N K TDN K I E K V MP P H QGR I K TA Q V RP I K P K O T I WVGQ RR
CPZ.CM.05.SIVcpz.LB7 K L I KA N Y I R EN K I E K V TP A H KO K I E H M N A E I R V FP K T I LQRO PTR
CPZ.CM.05.SIVcpz.MB66 K N Y R N NN K T A H A QRS R I D AE I Q V ST K S PRR
CPZ.CM.05.SIVcpz.MT145 Q V K N Y R K K DGTN MA QIY TNO K I V MP E H A ERK I T R TAT N I QV LTK L S O T F W KO VKT NR
CPZ.GA.88.GAB1 K N Y I R K DH N KVA I FR Y SK I V A SP A H Q LGS V FT R SO A E T R I QV LTK L S O T F R SER RHR
CPZ.TZ.01.TAN1 Q MI L T IFTTKCKD K TDT KRAG I TER SK VLH AC P H IGL QGS I T R T AAG V Q I ERLITF HF S R O T F FRKVVESDQKQKPPRR
CPZ.US.85.CPZUS L MI VKA N Y R K K TN VN IDK KI I K QA H OGT K I M K TA V RP I K PT F K VGS RR



Vpr start	amphipathic α -helix oligomerization	Vpr end in HXB2 frameshift in HXB2 H(S/N)RIG motifs	Vpr end
B.FR.83.HXB2	MEQAPEDQGPQREPHNEWTLLELLEEKNEAVRHFRPIW.	LHLGLGQHIYETGDTWAGVEAIRILQQLLFHIXFRIGCRHSRIGVTR.	QRRAR. NGASRS*
A1.GE.99.99GEMZ011	-----M-----D-----H-----P-----	-----Y-----N-----E-----T-----V-----Q-----	IEQ. R-G. S-
A1.KE.00.KER2008	-----Y-----M-----D-----H-----GP-----	-----D-----H-----E-----ML-----V-----T-----	IIIP. R-C. D-SG
A1.KE.00.KNH1144	-----Y-----M-----D-----H-----P-----	-----D-----H-----E-----T-----V-----Q-----	II. G-C. SG
A1.KE.00.KSM4024	-----Y-----M-----D-----H-----P-----	-----D-----H-----E-----L-----T-----Q-----	IIIPGRR. G-V.
A1.KE.00.MSA4069	-----Y-----M-----D-----H-----P-----	-----D-----H-----E-----L-----T-----Q-----	II. G-V. D-SG
A1.KE.00.NKU3005	-----Y-----M-----D-----H-----P-----	-----D-----H-----E-----L-----T-----Q-----	II. G-V. D-SG
A1.RU.00.RU00051	-----Y-----M-----D-----H-----P-----	-----D-----H-----E-----L-----T-----Q-----	IN. R-V. S-
A1.RU.03.03RU_20_06_13	-----Y-----M-----D-----H-----XP-----	-----X-----N-----E-----T-----LV-----X-----	IGG. R-V. S-
A1.RW.93.93RW_024	-----Y-----M-----D-----H-----P-----	-----D-----H-----E-----V-----V-----Q-----	IIP. G-G.
A1.SE.95.SE8891	-----Y-----A-----V-----H-----P-----	-----N-----E-----E-----T-----V-----Q-----	II. R-V. D-S-
A1.SE.95.UGSE8131	-----Y-----M-----D-----H-----P-----	-----D-----H-----E-----V-----T-----V-----X-----	II. G-. D-S-
A1.TZ.01.A173	-----Y-----M-----D-----H-----P-----	-----D-----H-----E-----V-----V-----Q-----	II-G. G-V. S-
A1.UA.01.01UADN139	-V-----Y-----M-----D-----H-----P-----	-----N-----E-----E-----V-----V-----Q-----	IIVQ. R-G. S-
A1.UG.92.92UG037	-----Y-----M-----D-----H-----P-----	-----D-----H-----E-----T-----V-----Q-----	INTR. G-V. D-SG
A1.UG.99.99UGA07072	-----Y-----M-----D-----H-----P-----	-----D-----H-----E-----T-----V-----Q-----	I. R-. D-S-
A1.UZ.02.02UZ0659	-----P-----Y-----M-----D-----H-----P-----	-----L-----S-----E-----T-----V-----L-----	IIVQ. R-V. -SG-
A2.CD.97.97CDKS10	-----P-----Y-----M-----D-----H-----P-----	-----L-----S-----E-----T-----V-----V-----	ISK. KKSKD.
A2.CD.97.97CDKTB48	-----P-----Y-----M-----D-----H-----P-----	-----L-----S-----E-----T-----V-----V-----	IIP. R-V. D-P
A2.CY.94.94CY017_41	-----P-----Y-----M-----D-----H-----P-----	-----L-----S-----E-----T-----V-----V-----	II. R-V. D-P
B.AR.04.04AR151516	-----N-----Y-----M-----D-----H-----S-----X-----	-----Q-----S-----Y-----N-----E-----V-----Y-----	I-L.
B.AU.87.MBC925	-----P-----Y-----M-----D-----H-----P-----	-----D-----H-----E-----V-----V-----Q-----	I-Y. K.
B.BO.99.BOI0122	-----Q-----Y-----M-----D-----H-----S-----	-----N-----E-----E-----LL-----S-----L-----	II.
B.BR.03.BREP2012	-----Q-----Y-----M-----D-----H-----S-----	-----N-----E-----E-----LL-----S-----L-----	II.
B.CA.97.CANB3FULL	-----Y-----M-----D-----H-----S-----	-----N-----E-----E-----M-----T-----Q-----	I. R.
B.CN.05.05CNHB_hp3	-----Y-----M-----D-----H-----S-----	-----N-----E-----E-----M-----T-----Q-----	I.
B.CO.01.PCM001	-----T-----Y-----M-----D-----H-----S-----	-----N-----E-----E-----M-----T-----Q-----	I. G-.
B.GB.83.CAM1	-----T-----Y-----M-----D-----H-----S-----	-----N-----E-----E-----M-----T-----Q-----	I. S.
B.GB.86.GB8	-----T-----Y-----M-----D-----H-----S-----	-----N-----E-----E-----M-----T-----Q-----	I. S.
B.GE.03.03GEMZ010	-----T-----Y-----M-----D-----H-----S-----	-----N-----E-----E-----M-----T-----Q-----	I. S.
B.IT.05.SG1	-----T-----Y-----M-----D-----H-----S-----	-----N-----E-----E-----M-----T-----Q-----	I. H-.
B.JP.05.DR6538	-----T-----Y-----M-----D-----H-----S-----	-----N-----E-----E-----M-----T-----Q-----	IIP. R.
B.KR.05.05CSR3	-----T-----Y-----M-----D-----H-----S-----	-----N-----E-----E-----M-----T-----Q-----	II.
B.NL.00.671_00T36	-----N-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I. RG-V. -P
B.RU.04.04RU128005	-----N-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I.
B.TH.00.00TH_C3198	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I.
B.UA.01.01UAKV167	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II.
B.US.04.ES10_53	-----R-----N-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II.
B.US.99.PR959_03	-----R-----N-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II.
C.AR.01.ARG4006	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II.
C.BR.04.04BR013	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II.
C.BW.00.00BW07621	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	IM. T.
C.CN.98.YNRL9840	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	IW.
C.ET.02.02ET_288	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I.
C.GE.03.03GEMZ033	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II.
C.IL.99.99ET7	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	IMP. R.
C.IN.99.01NS565_10	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II.
C.KE.00.KER2010	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II.
C.MM.99.mIDU101_3	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	IV.
C.MW.93.93MW_965	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. T. PN
C.SN.90.90SE_364	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I.
C.SO.89.89SM1145	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II.
C.TZ.02.C0178	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I.
C.UY.01.TRA3011	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II.
C.YE.02.02YE511	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. H. N
C.ZA.04.04ZASK164B1	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. H.
C.ZA.05.05ZAPSK240B1	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I. N.
C.ZM.02.02ZM108	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II.
D.CD.83.ELI	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. S.
D.CM.01.01CM_0009BBY	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I. S.
D.KE.01.NKU3006	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. TG.
D.KR.04.04KBH8	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. S.
D.TD.99.MN011	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. TA. P
D.TZ.01.A280	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I. T. S-P
D.UG.99.99UGK09259	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II.
D.YE.01.01YE386	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. I.
D.YE.02.02YE516	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	ISP. S. P
D.ZA.90.R1	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. S.
F1.AR.02.ARE933	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I.
F1.BE.93.V1850	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I.
F1.BR.01.01BR125	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I.
F1.ES.x.P1146	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I.
F1.FL93.FIN9363	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. V.
F2.CM.02.02CM_0016BBY	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. I. NG
F2.CM.95.MP257	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. I. SG
F2.CM.97.CM53657	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. L. NG
G.BE.96.DRCBL	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	IIPQ. R-V. D-SG
G.CM.01.01CM_4049HAN	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I. V.
G.CU.x.Cu74	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	ISL. R-V. D-PG
G.ES.00.X558	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I. V.
G.ES.99.X138	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I. V. D-PG
G.GH.03.03GH175G	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I. V. D-P
G.KE.93.HH8793_12_1	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I. R-V. P
G.NG.01.01NGPLD669	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I. R-V. D-SG
G.PT.x.PT695	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I. R-V. D-PG
G.SE.93.SE6165	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I. R-V. D-PG
H.BE.93.V1991	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I. V. P
H.BE.93.V1997	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. V. P
H.CF.90.056	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	I. V. P
J.CD.97.J_97DC_KTB147	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. G. S-
J.SE.93.SE7887	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. G.
J.SE.94.SE7022	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	IIP. G.
K.CD.97.EQTB11C	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	II. GG.
K.CM.96.MP555	-----P-----Y-----M-----D-----H-----S-----	-----D-----H-----E-----X-----L-----M-----	IIP. G. S-

	Tat start	disulfide bonding	NLS	exon 1 end	exon 2 start	Tat end	
B.FR.83.HXB2	MEPVDPRLEPKHPGSPQKTA	CTNICYCKKCCFFHCQVCF	ITKALGISYGRKKRR	.QRRR	.AHONSQTHQASLSKQPTSOP	.RGD.PTGPKE+KKKVERETETDPFD*	100
01_AE.CF.90.90CF11697	-----S-----N-----T-----SK-----W-----L-----LK-----G-----K-----H-----G-----PPP-----G-----KD-----NPVP-----LPTT-----N-----D-----S-----E-----AKKAK-----A-----						102
01_AE.CN.05.FJ051	-----N-----N-----T-----N-----W-----L-----LK-----G-----H-----K-----H-----G-----TP-----S-----KD-----NPVP-----L-----LII-----N-----D-----S-----E-----ASKA-----*						101
01_AE.CN.06.FJ054	-----N-----N-----K-----T-----SK-----I-----W-----L-----LK-----G-----K-----H-----G-----TP-----S-----KD-----NLTR-----LPTI-----N-----D-----S-----E-----ASK-----*						101
01_AE.HK.x.HK001	-----N-----N-----K-----T-----SK-----I-----W-----L-----LK-----G-----K-----H-----G-----TP-----S-----KD-----NPVP-----LPTI-----N-----D-----S-----E-----ASK-----*						101
01_AE.JP.93.93JP.NH1	-----N-----N-----T-----SK-----W-----L-----LK-----G-----H-----K-----H-----G-----TP-----S-----KD-----TPIE-----LPII-----N-----D-----S-----E-----ASKA-----*						102
01_AE.TH.01.01TH.R2184	-----N-----N-----T-----SK-----I-----W-----L-----LK-----G-----K-----H-----G-----TP-----S-----KD-----NP-----PE-----LPII-----N-----D-----S-----E-----ASK-----*						101
01_AE.TH.02.OUR7691	-----N-----N-----K-----T-----SK-----I-----W-----L-----LK-----G-----H-----K-----H-----G-----P-----S-----ED-----DPVP-----LPII-----N-----D-----S-----E-----ASK-----*						102
01_AE.TH.90.CM240	-----L-----N-----N-----T-----SK-----W-----L-----LK-----G-----K-----H-----G-----TP-----S-----KD-----NPVP-----LPTI-----RN-----D-----S-----E-----ASKA-----QC-----						102
01_AE.US.00.00US_MSC1164	-----N-----N-----T-----S-----I-----W-----L-----LK-----G-----K-----H-----G-----TP-----S-----KS-----NPTR-----FPIT-----N-----D-----S-----E-----SKA-----*						102
02_AG.CM.02.02CM_4082STN	-----L-----L-----N-----N-----T-----ST-----W-----L-----LN-----G-----R-----G-----TP-----SH-----D-----SPVP-----LPTT-----N-----A-----E-----ST-----ASK-----*						101
02_AG.EC.x.EC041	-----M-----D-----N-----A-----NK-----M-----W-----L-----LN-----G-----R-----G-----TP-----SR-----D-----NPVP-----LPTT-----N-----S-----E-----ASK-----*						102
02_AG.FR.91.DJ264	-----S-----N-----T-----S-----I-----W-----L-----LN-----G-----X-----R-----G-----TP-----SR-----DR-----NPV-----LPTT-----N-----S-----RE-----SK-----K-----C-----						102
02_AG.GH.03.GHNJ196	-----DL-----N-----D-----N-----T-----SK-----I-----W-----LG-----LN-----G-----P-----G-----TP-----R-----DY-----NPAP-----LPTT-----I-----QADQ-----S-----E-----ASK-----*						102
02_AG.NG.01.PL0710	-----L-----L-----S-----N-----T-----SK-----I-----W-----L-----LK-----G-----R-----G-----TP-----SH-----D-----NPV-----LPTT-----N-----A-----S-----E-----ASK-----*						102
02_AG.NG.x.IBNG	-----L-----L-----S-----N-----T-----SK-----M-----X-----L-----LN-----G-----R-----G-----TP-----SR-----D-----NPVP-----LPTT-----N-----D-----S-----E-----SK-----K-----C-----						102
02_AG.SE.94.SE7812	-----L-----L-----S-----N-----T-----SK-----I-----W-----L-----LN-----G-----K-----R-----G-----TP-----SR-----DN-----DPVP-----LPTT-----N-----A-----S-----E-----AGK-----*						102
02_AG.SN.98.MP1211	-----L-----L-----K-----N-----T-----SK-----L-----W-----L-----LN-----G-----R-----G-----TP-----SR-----D-----NPVP-----LPTT-----E-----Q-----S-----E-----ASK-----*						102
02_AG.UZ.02.02UZ710	-----DL-----L-----S-----N-----T-----NK-----M-----W-----L-----LN-----G-----R-----G-----TS-----SR-----D-----NPVP-----LPTA-----N-----S-----E-----ASK-----*						102
03_AB.RU.97.KAL153_2	-----D-----D-----N-----T-----D-----NK-----F-----W-----L-----LK-----G-----K-----H-----G-----SL-----G-----KG-----NLIP-----L-----Q-----PN-----S-----E-----Q-----ASK-----*						102
04_cpx.CY.94.CY032	-----D-----D-----N-----N-----R-----R-----NQ-----M-----Y-----N-----G-----P-----PP-----G-----A-----DPVPE-----P-----N-----S-----E-----Q-----ASK-----*						103
05_DFBE.x.VII310	-----D-----D-----N-----N-----R-----R-----NQ-----M-----Y-----N-----G-----P-----PP-----G-----A-----DPVPE-----P-----N-----S-----E-----Q-----ASK-----*						102
06_cpx.AU.96.BFP90	-----K-----I-----NQ-----R-----K-----Y-----P-----LN-----G-----Q-----PPG-----KN-----DPV-----L-----T-----QRE-----Q-----EKS-----E-----SKA-----P-----R-----						102
06_cpx.RU.05.04RU001	-----K-----I-----NQ-----R-----P-----P-----Y-----PI-----LN-----G-----T-----PPG-----KN-----DPV-----V-----T-----QRE-----EKKQ-----EM-----SKATP-----R-----						102
07_BC.CN.97.CN54	-----N-----N-----N-----E-----N-----S-----TK-----G-----F-----S-----RS-----ED-----NPI-----LPRT-----Q-----SE-----S-----SK-----G-----						102
08_BC.CN.97.97CNGX_6F	-----N-----N-----N-----E-----N-----R-----SY-----L-----QK-----G-----S-----P-----S-----ED-----NLI-----SLPRT-----Q-----QA-----SE-----S-----SK-----KPA-----						102
09_cpx.GH.96.96GH2911	-----L-----L-----N-----N-----A-----NK-----R-----Y-----H-----LN-----G-----A-----SCG-----KSN-----DPVP-----LP-----T-----N-----S-----E-----SK-----KS-----Q-----						102
10_CD.TZ.96.96TZ_BF061	-----H-----N-----N-----N-----MK-----H-----A-----SY-----L-----Q-----S-----S-----PPS-----ED-----NLI-----S-----S-----A-----A-----						102
11_cpx.GR.x.GR17	-----G-----I-----NID-----NQ-----NQ-----R-----Y-----H-----LK-----G-----W-----T-----SRS-----KN-----DPIPE-----LP-----A-----SRN-----E-----P-----E-----SKA-----PA-----						102
12_BF.AR.99.ARMA159	-----L-----L-----N-----N-----K-----S-----P-----K-----R-----Y-----V-----G-----P-----G-----N-----I-----A-----S-----E-----SKAKS-----C-----						102
13_cpx.CM.96.1849	-----D-----D-----N-----N-----P-----K-----K-----G-----S-----SH-----N-----DLIP-----LPI-----N-----Q-----S-----E-----SK-----W-----						102
14_BG.DE.01.9196_01	-----D-----D-----N-----N-----W-----L-----LN-----G-----K-----R-----G-----PP-----S-----ED-----NPVP-----LPTT-----N-----E-----S-----SK-----K-----N-----						102
14_BG.ES.99.X397	-----D-----D-----N-----N-----K-----I-----W-----L-----LN-----G-----K-----H-----G-----PP-----S-----KD-----NPVP-----LPTT-----N-----E-----S-----SK-----K-----						102
15_01B.TH.99.99TH_MU2079	-----D-----D-----L-----N-----N-----R-----T-----N-----I-----W-----I-----LK-----G-----K-----H-----G-----TP-----S-----KD-----NPAP-----L-----IS-----K-----D-----KS-----E-----ASKA-----*						101
16_A2D.KR.97.97KR004	-----D-----D-----N-----N-----R-----NK-----R-----Y-----P-----LN-----G-----P-----PP-----GGTG-----NPIE-----SLPR-----ORI-----TE-----S-----TE-----SKA-----*						102
18_cpx.CU.99.CU76	-----D-----D-----N-----N-----R-----Y-----L-----LK-----G-----T-----TPYD-----KH-----DLPIE-----L-----RT-----SW-----Q-----E-----ASKA-----A-----QC-----						102
19_cpx.CU.99.CU7	-----D-----D-----N-----N-----N-----T-----L-----TK-----G-----G-----PP-----SH-----D-----NPVP-----LPDT-----E-----T-----S-----E-----SK-----K-----						101
20_BG.CU.03.CB471	-----D-----D-----N-----N-----N-----I-----W-----L-----LN-----G-----H-----G-----TPPS-----KD-----NPVP-----IPTT-----N-----S-----ASK-----P-----						102
21_A2D.KE.91.KNH1254	-----D-----D-----N-----N-----R-----P-----NK-----H-----Y-----T-----L-----G-----S-----PP-----SG-----P-----VHIP-----S-----S-----S-----						102
23_BG.CU.03.CB118	-----D-----D-----N-----D-----N-----NK-----M-----W-----L-----LK-----G-----K-----H-----G-----TP-----D-----KD-----IPVP-----LPTT-----N-----S-----E-----SK-----A-----						102
24_BG.CU.03.CB378	-----D-----D-----N-----N-----N-----NK-----V-----W-----L-----LK-----G-----K-----H-----G-----TPLS-----NKD-----NPVP-----LPTA-----N-----KS-----E-----SK-----KA-----						102
25_cpx.CM.01.101BA	-----D-----D-----NI-----NQ-----T-----N-----I-----W-----L-----LN-----G-----TPPG-----KD-----YPVP-----IPTT-----N-----Q-----S-----ASK-----A-----S-----						102
27_cpx.CD.97.97CDKTB49	-----D-----D-----NI-----NQ-----NQ-----S-----Y-----L-----LK-----G-----FH-----A-----TPYG-----KN-----DPIP-----SIP-----T-----QRV-----S-----E-----SE-----SKA-----RL-----W-----						102
28_BFBR.99.BREPM12609	-----L-----Q-----R-----NK-----L-----MK-----G-----P-----DN-----NN-----V-----A-----S-----N-----E-----S-----T-----A-----						102
29_BFBR.02.BREPM119	-----D-----D-----N-----N-----P-----N-----R-----K-----G-----SPRH-----PA-----P-----S-----S-----KA-----Q-----						102
31_BC.BR.02.110PA	-----L-----L-----N-----X-----N-----N-----SY-----L-----L-----G-----S-----PPX-----ED-----NLI-----LP-----X-----Q-----SE-----S-----K-----K-----						102
33_01B.MY.05.05MY_KL007_1	-----L-----L-----N-----N-----K-----T-----S-----R-----W-----L-----LK-----G-----K-----HKH-----RGTP-----S-----ED-----NP-----P-----L-----IS-----N-----DS-----S-----ASKA-----GLCA-----						104
34_01B.TH.99.OUR2478P	-----N-----N-----N-----T-----SK-----N-----W-----L-----LK-----G-----K-----H-----G-----TP-----S-----KD-----NPIE-----LPII-----N-----Q-----S-----E-----ASKA-----*						101
35_AD.AF.05.05AF095	-----DL-----L-----D-----D-----N-----T-----P-----NK-----F-----V-----SY-----LN-----G-----TP-----S-----D-----NP-----P-----P-----A-----Q-----I-----S-----E-----S-----SKA-----P-----R-----						102
36_cpx.CM.00.00CMNYU830	-----L-----L-----NI-----NI-----N-----R-----N-----F-----W-----LN-----G-----R-----G-----P-----SR-----D-----DPV-----LPAI-----N-----KN-----S-----E-----ASK-----N-----C-----						103
37_cpx.CM.00.00CMNYU926	-----D-----D-----N-----N-----N-----TK-----G-----T-----A-----SR-----DN-----H-----VP-----LPTT-----N-----S-----E-----ASKA-----CA-----						102
42_BFLU.03.luBF_05_03	-----I-----I-----K-----E-----R-----R-----G-----P-----N-----PPPD-----E-----AT-----X-----						102
N.CM.02.DJ00131	-----N-----N-----N-----NG-----Y-----Y-----MC-----TK-----G-----S-----TPKS-----KN-----DPIPE-----L-----L-----QOOP-----Q-----Q-----KO-----AL-----DKA-----*						103
N.CM.04.04CM_1015_04	-----N-----N-----N-----NK-----N-----Y-----V-----TK-----G-----S-----K-----TPNS-----KDY-----DPIPE-----L-----Q-----P-----Q-----Q-----Q-----AL-----SK-----A-----LC-----						102
N.CM.04.04CM_1131_03	-----N-----N-----N-----N-----Y-----Y-----TK-----G-----S-----TPNS-----KNY-----DPIPE-----L-----Q-----P-----Q-----Q-----Q-----AL-----XK-----A-----C-----						102
N.CM.95.YBF30	-----N-----N-----N-----N-----R-----Y-----LY-----TK-----G-----S-----TP-----S-----KS-----DLPIE-----L-----Q-----Q-----Q-----Q-----Q-----EAL-----SK-----A-----C-----						102
N.CM.97.YBF106	-----N-----N-----N-----NK-----Y-----MC-----TK-----G-----S-----PPKS-----KD-----DPIPE-----L-----RQ-----QP-----Q-----Q-----KO-----AL-----GK-----A-----C-----						103
O.BE.87.ANT70	-----D-----D-----EVP-----H-----QIP-----N-----R-----Y-----VR-----G-----G-----P-----AASHPD-----KDPVP-----SPTIT-----KPK-----QERQE-----OEEE-----KKAG-----GGYPR-----						100
O.CM.91.MVP5180	-----D-----D-----EMP-----H-----K-----Q-----P-----N-----R-----Y-----Y-----TK-----G-----H-----PA-----AASYPD-----NKDPVPE-----SL-----HT-----GRK-----OKROE-----OE-----K-----GPS-----GDC-----						101
O.CM.96.96CMAABB637	-----D-----D-----EVP-----H-----P-----P-----NA-----Y-----YL-----TK-----G-----H-----PA-----AASCS-----NNKDLVPE-----STST-----NRK-----OKROE-----OEE-----AK-----G-----GGD-----PC-----						101
O.CM.98.98CMA104	-----D-----D-----EMP-----H-----Q-----P-----N-----S-----Y-----Y-----LVR-----G-----P-----TASH-----PDNKDPVPE-----SP-----IT-----NRK-----E-----QOE-----OE-----K-----GPG-----YHC-----						98
O.CM.99.99CMU4122	-----D-----D-----EMP-----H-----Q-----P-----NK-----Y-----Y-----TS-----G-----P-----AAH-----PDNKDLVPE-----SPTIT-----NRK-----QERQE-----OE-----K-----GP-----*-----YPG-----						99
O.SN.99.SEMP1299	-----D-----D-----EMP-----H-----Q-----P-----NK-----A-----Y-----Y-----AS-----G-----PA-----AAR-----PDN-----DIVPE-----L-----YIT-----NRK-----OKROE-----OE-----N-----AC-----PR-----YPG-----						100
O.US.99.99USTWLA	-----D-----D-----HEMP-----H-----S-----Q-----P-----NT-----S-----Y-----I-----TK-----G-----PT-----TAER-----HDNKDPVPE-----SPTIT-----SRK-----E-----RO-----OE-----KKAG-----GG*-----PR-----						99
O.FR.92.VAU	-----D-----D-----I-----EMP-----H-----Q-----P-----NA-----Q-----Y-----YL-----TK-----G-----H-----PA-----AASYPD-----NKDPVPE-----SP-----NN-----KPK-----E-----QE-----OE-----TK-----GAG-----LHC-----						100
CPZ.CD.90.ANT	-----D-----D-----A-----ETP-----L-----PAT-----A-----P-----N-----C-----Y-----PL-----TK-----G-----R-----ARRN-----T-----TAES-----ENN-----DPV-----*-----SLPKT-----SRI-----OSSQ-----K-----E-----EK-----GSG-----GPC-----						103
CPZ.CM.01.SIVcpzCAM13	-----I-----I-----E-----N-----N-----N-----A-----M-----Y-----TK-----G-----A-----R-----T-----FAS-----KNN-----DPVR-----L-----L-----N-----QERSE-----KOE-----SQAAA-----*						101
CPZ.CM.05.SIVcpzEK505	-----I-----I-----E-----N-----N-----N-----H-----V-----Y-----Y-----TK-----G-----S-----TP-----SNKS-----DP-----P-----L-----R-----L-----Q-----Q-----Q-----TL-----S-----*-----TA-----CA-----						101
CPZ.CM.05.SIVcpzLB7	-----I-----I-----S-----D-----N-----P-----N-----F-----Y-----PF-----TK-----G-----W-----S-----TSKS-----PHY-----DPI-----FPFS-----SLS-----E-----Q-----SKAQ-----Q-----*-----C-----						101
CPZ.CM.05.SIVcpzMB66	-----I-----I-----N-----N-----N-----Q-----Y-----L-----TK-----G-----R-----K-----TSES-----N-----DPVP-----L-----YIT-----G-----I-----E-----Q-----KS-----E-----SQ-----TS-----Q-----A-----						102
CPZ.CM.05.SIVcpzMT145	-----I-----I-----D-----N-----N-----K-----R-----E-----NK-----F-----S-----Y-----IL-----TK-----G-----G-----R-----KD-----TAER-----HDNKDPVPE-----L-----YIT-----SRK-----Q-----SEKO-----E-----TS-----A-----OS-----W-----						101
CPZ.GA.88.GAB1	-----D-----D-----I-----D-----N-----R-----V-----N-----A-----Y-----I-----TK-----G-----TTR-----T-----PAG-----KNN-----D-----IP-----L-----S-----N-----KE-----SEKST-----E-----ASK-----A-----Q-----*						102
CPZ.TZ.01.TAN1	-----*-----IXQVA-----E-----AA-----E-----P-----P-----TK-----G-----R-----G-----SAVH-----TNN-----DPVR-----SLPKR-----SRI-----ONSQE-----SQEE-----A-----TS-----GRR-----PR-----						96
CPZ.US.85.CPZUS	-----D-----D-----NI-----Q-----N-----V-----TK-----G-----GRR-----T-----TAKS-----ADN-----DP-----R-----F-----Q-----L-----N-----Q-----Q-----Q-----AL-----KQ-----TS-----QR-----						103

	Rev start	exon 1 end	exon 2 start	NLS	Leu-rich effector domain	Rev end										
B.FR.83.HXB2	MAGRSGD	.SDEE	.LIRTVRLIKLLYQSNPPPN	.PEG	TRQARRNRRRWRERQRQTHSISERILGTYLGRSAEPVPLQLPPLRLTLDCN	.EDCGTSGT QGVGSPQILVESPTVLESQTK*										
01_AE.CF.90.90CF11697	N	T-D	LQA-I-I	Y-P	K-R	A-L	VACV-T	H-S	Q0SGGTET	G-SG-SVI-G	N					
01_AE.CN.05.FJ051	N	L-AA	T-I	Y-P	S-S	A-R	SSCV-CT	H-S	Q0SGGTET	R-FGK-SV-G	N					
01_AE.CN.06.FJ054	S	T	L-A	I-I	E	Y-S	S-C	GM	H-S	G-A00SGSET	R-SG-PSDI-P	N				
01_AE.HK.x.HK001	S	T	LGA-A-I	Y-P	S-S	K-A	RE	ST	S	H-S	G-SG-SVI-G	N				
01_AE.JP.93.93JP_NH1	S	T	L-A	I-I	F-S	S-S	T-K	A-RA-D	S-C	T	H-S	Q0SGTETP	R-SG-PSVI-P	N		
01_AE.TH.01.01TH_R2184	S	I	L-A	I-I	Y-S	S-S	T-K	A-GE	S-C	T	H-S	S-Q0SGGTET	R-SG-SVI-GP	N		
01_AE.TH.02.OUR7691	S	T	L-A	KI-RI	F-S	S-S	K-A	A-RA	SAC	T	H-S	Q0SGGTET	K-SG-SGI-GP	N		
01_AE.TH.90.CM240	S	T	L-A	I-I	F-S	S-S	T	A-SA	S-C	T	V-N	S-G	Q0SGGTET	R-SG-SVI-GP	N	
01_AE.US.00.00US_MSC1164	S	T	L-A	A-RI	E-F	S-S	S	A-RA	SAC	T	H-S	Q0SGGTET	R-SG-SVI-GP	EN		
02_AG.CM.02.ZCM_4082STN	A	T	L-AI	I-I	Y-P	G	K-A	VRA	S-C	P	I-C	S-R	Q0SGGTET	SG-CSII-G	N	
02_AG.EC.x.EC041	A	T	L-A	I-I	Y-P	G	K-A	RA	S-C	P	L-I	C-S	N-P	Q0SGGTET	SG-CSII-G	N
02_AG.FR.91.DJ264	X	A	L-V	IV-I	Y-P	G	K-G	A-RA-Q	S-C	P	I-C	S-G	Q0SGGTET	L-SG-SDI-GA	N	
02_AG.GH.03.GHNJ196	P	L	I-I	H-Y	Y-P	G	S-T	K-AS-S-RA	L-SSC	-PK	I	S-S	Q0QEPAT	S-SNI-G	N	
02_AG.NG.01.PL0710	A	L	AI	I-I	Y-P	G	K-A	A-GAL	S-C	P	SF	I-S	Q0SGTKT	VS-H-VI-G	N	
02_AG.NG.x.IBNG	A	L	A	I-I	Y-P	G	T-K	A-RA	S-C	P	I	N-S	QLSGTET	S-YII-G	N	
02_AG.SE.94.SE7812	N	A	L	A	T-I	Y-P	S	K-A	A-RA	S-C	P	I	R-S	S-SI-G	N	
02_AG.SN.98.MP1211	A	G	L	A	I-RI	Y-P	S	K-A	VRA	S-C	P	I	C-S	S-Q0SGGTET	S-TS-SVI-G	N
02_AG.UZ.02.02UZ710	V	G	L	A	I-I	Y-P	S	A	RAL-D-V	SSC	-PT	S-I	C-S	A-Q0SGGTAT	S-SNL-G	N
03_AB.RU.97.KAL153_2	L	K	I	F	S	S	H	Q	S	PE	L	H	S	D	N	
04_cpx.CY.94.CY032	N	I	D	FKAA	A	I	Y	N	T	A	K	L	A	P	K	N
05_DFBE.x.VII310	R	D	LKA	I	L	S	A	N	G	L	S	E	N	N	S	S
06_cpx.AU.96.BFP90	DR	LLA	I	I	Y-K	N	S	A	N	D	V	S	C	T	I	R
06_cpx.RU.05.04RU001	N	Q	LLA	T	I	Y-K	N	S	Q	A	H	D	Q	SP	T	I
07_BC.CN.97.CN54	E	A	EA	I	I	Y-E	R	K	KA	G	SAC	P	I	G	HISGS	GG
08_BC.CN.97.97CNGX_6F	A	L	A	I	I	Y-E	R	K	A	N	L	SAC	PT	I	G	HISGS
09_cpx.GH.96.96GH2911	N	Q	LAA	V	I	Y-K	N	S	A	N	RA	RAC	PE	N	H	N
10_CD.TZ.96.96TZ_BF061	A	LQA	I	I	S	S	A	GG	DA	PE	N	N	N	H	GH	S
11_cpx.GR.x.GR17	AQ	LAA	I	I	S	Y-K	A	A	N	L	Q	S	C	E	P	I
12_BF.AR.99.ARMA159	L	KAA	I	I	Y-K	N	S	A	S	RA	ID	PE	F	G	N	N
13_cpx.CM.96.1849	V	LT	I	I	N	Y	S	Q	K	A	GO	S	C	H	N	N
14_BG.DE.01.9196_01	S	A	D	LKA	I	I	Y-P	G	K	A	SACV	XE	H	S	S	L
14_BG.ES.99.X397	S	T	D	LKA	I	I	Y-P	G	K	A	SACV	T	S	H	N	S
15_01B.TH.99.99TH_MU2079	S	T	L	A	I	H	Y	S	A	S	T	A	KR	RE	S	Q0SGGTET
16_A2D.KR.97.97KR004	P	D	L	A	A	RI	S	D	R	S	K	A	RA	SAC	PT	H
18_cpx.CU.99.CU76	Q	LT	SI	I	Y-E	A	S	A	N	C	GP	N	S	F	Q0SGGTET	
19_cpx.CU.99.CU7	D	L	VI	I	I	Y-T	QR	A	K	A	S	C	E	F	H	N
20_BG.CU.03.CB471	G	T	LQA	I	RI	Y-P	G	K	A	S	R	T	C	PT	F	S
21_A2D.KE.91.KNH1254	A	D	LKV	L	SI	S	S	A	S	PE	N	N	S	I	T	G
23_BG.CU.03.CB118	S	T	L	I	I	Y-P	G	K	A	Q	ADC	PP	N	S	I	Q0SGGTET
24_BG.CU.03.CB378	S	T	L	AI	I	I	Y-P	G	K	A	V	T	C	P	L	S
25_cpx.CM.01.101BA	N	L	A	T	NI	R	Y	P	G	K	A	RQ	TSC	LE	HI	S
27_cpx.CD.97.97CDKTB49	EG	N	R	LTA	I	I	Y	K	R	SQ	K	A	D	G	S	C
28_BFBR.99.BREPM12609	D	K	I	I	F	S	S	A	L	RW	SNH	PE	F	I	C	Q0SGGTET
29_BFBR.02.BREPM119	D	LG	I	I	F	S	R	S	R	N	TW	S	R	P	H	N
31_BC.BR.02.110PA	R	A	LQX	I	I	Y	X	S	A	G	S	C	T	F	I	N
33_01B.MY.05.05MY_KL007_1	S	T	SDEE	LKA	I	I	Y	S	A	T	K	A	VCAL	D	SAC	T
34_01B.TH.99.OUR2478P	S	T	L	A	T	I	Y	S	S	K	A	K	RE	D	ISC	T
35_AD.AF.05.05AF095	L	A	I	I	H	K	R	S	K	A	N	D	Q	AC	P	H
36_cpx.CM.00.00CMNYU830	A	L	V	T	I	Y	P	S	R	K	Q	A	RA	S	C	FE
37_cpx.CM.00.00CMNYU926	Q	L	A	I	NI	Y-P	G	K	A	RAL	S	C	H	N	S	RD
42_BFLU.03.luBF_05_03	R	T	LQ	I	R	S	S	X	QTX	GW	N	PE	H	S	V	A
N.CM.02.DJ00131	V	N	L	A	I	I	Y	NNQ	X	T	RE	Q	D	C	GLQ	D
N.CM.04.04CM_1015_04	V	N	K	LTA	I	I	Y	NQ	X	A	CA	S	C	GLQ	GDPP	N
N.CM.04.04CM_1131_03	V	N	LTA	I	I	Y	NQ	X	XA	RA	X	S	C	GLQ	GDPP	X
N.CM.95.YBF30	V	N	L	A	V	I	Y	SK	A	RA	RA	S	C	GPP	D	P
N.CM.97.YBF106	V	N	D	LKA	KI	I	Y	D	SSQ	A	RA	A	GPO	D	P	G
O.BE.87.ANT70	E	Q	LQAIQI	I	I	Y	S	N	K	R	A	VDTLAA	V	A	VHGPONNNLVD	Q
O.CM.91.MVP5180	EE	Q0	LQAIQI	I	I	Y	S	N	K	R	A	VD	LAT	A	VHGPONNNLVD	Q
O.CM.96.96CMAAB637	Q	Q0	LQA	I	I	Y	S	N	K	R	A	VELTAA	V	A	VHGPONNNLVD	Q
O.CM.98.98CMA104	D	Q0	LQAI	I	I	Y	S	N	K	R	A	VDT	AA	A	VHGPONNNLVD	Q
O.CM.99.99CMU4122	D	Q0	VP	LH	I	I	Y	S	N	K	R	A	DTLAA	F	ASVHGPONNNLVD	Q
O.SN.99.SEMP1299	DG	Q0	LAIQI	I	I	Y	S	N	K	R	A	HAHVDTLAA	A	VHGPONNNLVD	Q	
O.US.99.99USTWLA	E	Q0	LQAIQI	I	I	Y	S	N	K	R	A	VD	LAT	A	VHGPONNNLVD	Q
O.FR.92.VAU	EE	Q0	LQAIQI	I	I	Y	S	N	K	R	A	VD	AA	V	ASVHGPONNNLVD	Q
CPZ.CD.90.ANT	EELEGT	Q	LKA	KI	I	Y	K	A	S	A	K	K	D	VEGLAA	V	R
CPZ.CM.01.SIVcpzCAM13	EH	D	AR	LQA	KI	I	E	Y	S	N	K	A	RE	N	V	AAH
CPZ.CM.05.SIVcpzEK505	V	N	L	AI	V	I	Y	DS	K	A	RAL	SSCV	GLQ	ST	P	D
CPZ.CM.05.SIVcpzLBT	R	L	A	I	I	Y	H	S	A	H	SE	AG	S	C	E	PEG
CPZ.CM.05.SIVcpzMB66	G	N	L	A	I	I	Y	S	S	A	N	L	S	C	E	PTG
CPZ.CM.05.SIVcpzMT145	EG	DENL	RAIR	RI	E	Y	S	Q	S	K	A	NA	G	V	ASC	S
CPZ.GA.88.GAB1	EPQD	AR	LQA	KI	I	Y	S	K	A	K	S	G	V	A	PPK	G
CPZ.TZ.01.TAN1	EE	AN	LY	I	I	D	Y	S	GA	S	T	Q	H	VDALAS	QYR	PH
CPZ.US.85.CPZUS	GDA	Q	L	A	I	I	A	F	NS	KN	NE	Q	S	C	P	D

	transmembrane domain	phos phos cytoplasmic domain			
	Vpu start	α-helix	α-helix	Vpu end	
B.FR.83.HXB2	TQPTPTVAIVAVLVAIIIAIVVSWIVIIIEYRKILRQRKIDRLIDRLIERAEDSGNESEG	...EISALV...EMGVEMGHAPWDVD	.DL*	
A1.GE.99.99GEMZ011	MT-LE-Y-VFV	-T-G-L-L-K	K-IR	-DAE-L-T-M	-V-NYVLL-DN-NV
A1.KE.00.KER2008	LSALE-C-AG-I-L	-T-G-L-R-L-K	E-IR	-D-DTE-LA-I	-NYDLG-AN
A1.KE.00.KNH1144	LP-L-S-G-I-A-L	-I-T-G-L-K	E-IR	-D-DTE-L-I	-NYDLGTDN-N
A1.KE.00.KSM4024	L-LE-W-TG-L-L	-I-T-G-CRKY-D	E-IR	-D-DTE-L-I	-NYDLG-AN
A1.KE.00.MSA4069	WTALQ-C-V-G-I-L	-T-G-KRL-K	E-IR	-D-DTE-L-T-I	-L-DYGLG-DN-N
A1.KE.00.NKU3005	LS-LE-W-G-L-L	-T-G-K-L-K	E-IR	-D-DTE-LA-	-NYDLGNDN-N
A1.RU.00.RU00051	PL-LE-Y-I-L	-T-G-L-K	E-IR	-DAE-V-T-M	-GV-ND-FL-DN-NV
A1.RU.03.03R120_06_13	MT-LE-Y-FV	-T-G-L-K	E-IR	-DAE-L-T-M	-V-NX-LXDN-NV
A1.RW.93.93RW_024	M-SLE-CT-GFT-L	-T-G-K-V	I-K-IR	-D-DTE-L	R-NYDLG-DN-N
A1.SE.95.SE8891	MT-LE-W-G-I-VS	-T-G-K-L-K	R-IR	-D-DTE-LAK	-NYDLG-NN-N
A1.SE.95.UGSE8131	MS-LE-W-G-I-L	-T-GL-K-L-K	E-IR	-D-DTE-L-T-I	-L-DYDLG-DN-N
A1.TZ.01.A173	MT-LE-W-G-L-V	-T-G-KRL-K	E-IR	-D-DTE-L-L	-NYDLGIDN-N
A1.UA.01.01UADN139	MT-LE-YT-VSVL-F-TA	-G-LRK	E-IR	-DAE-L-T-M	-GV-NV-L-DN-NV
A1.UG.92.92UG037	M-LLE-L-V-G-L	-T-G-K-L-K	E-IR	-D-DTE-L	D-DYDLG-DN-N
A1.UG.99.99UGA07072	MT-LE-S-G-L	-T-G-KRLRK	I-K-IR	-D-DAE-LAK	-YYDLG-DN-N
A1.UZ.02.02UZ0659	MTLLE-Y-VFV	-T-G-L-K	E-IR	-DAE-L-T-M	-V-NV-L-DN-NV
A2.CD.97.97CDKS10	MLSLE-A-G-L-A	-T-F-RE	W-Q-IR	-D-DTE-L	-NLNFG-AN-N
A2.CD.97.97CDKTB48	MS-LA-LS-G-S-L	-TV-F-KK	W-LE-IR	-D-DTE-L-KM	-G-NLGF-DN-NV
A2.CY.94.94CY017_41	ML-LV-L-G-I-L	-T-F-K-KK	W-K-IR	-D-DTE-L	-R-LDFG-N-NV
B.AR.04.04AR151516	MT-LQ-A-TA-IVG	-A-L	L-IL	DOE-LA	-N
B.AU.87.MBC925	M-SLV-LT-F	-T-F	IS	DOE-L	-VIN
B.BO.99.BOI0122	M-SLI-L-A-L	-T-F	IS	DOE-L-V	GMEH-D
B.BR.03.BREP2012	L-LV-S-A-VT	-TL-L	IA	DOE-L	D-N-ND
B.CA.97.CANB3FULL	MHSLO-V-I-A	-T-G	E-I	AOK-L	
B.CN.05.05CNHB_hp3	L-ALT-F-I-VA	-T-LL	IR	D-DOE-L	LI-RNA
B.CO.01.PCM001	MHALQ-S-TA	-T-GL-I-V	E-IA	DOE-LA	D-N-N-M
B.GB.83.CAM1	M-LQ-L-V-GS	-L-T-F	R-K	DOE-L-M	N-N
B.GB.86.GB8	YV-V-O-LT-G-V	-I-F-KK	XK-I	DOE-L	D
B.GE.03.03GEMZ010	M-LQVL-VG	-T-F-K	IA	DOE-L-M	R
B.IT.05.SG1	M-LG-L	-L-R	R-N	DOE-L	EMVV-D-VIN
B.JP.05.DR6538	M-LLE-S-VT	-I-T-F	IT	DOE-L-M	N
B.KR.05.05CSR3	M-SLE-L-A-L	-I-T-F-R-E	K-E-IR	D-DOE-L-K	R-L-I
B.NL.00.671.00T36	M-SLE-L-A-L	-I-T-F	R-E	DOE-L	
B.RU.04.04RU128005	M-SLX-A-G-X-X	-L-KXL	IR	DOE-L-G-I	-L-I
B.TH.00.00TH_C3198	M-TLT-L-A	-T-L	IR	D-DOE-L-FM	
B.UA.01.01UAKV167	QTAIVAL-I-I-VAV	-TL-L	I	DOE-L-T	D-L-NI
B.US.04.ES10_53	MLSLV-L-VA	-I	K-V	D-DLE-L	ERGOL
B.US.99.PRB959_03	M-LH-L-G-L	-L	IR	D-DQE-L	R-L
C.AR.01.ARG4006	KXVDYRLG-G-I-L	-I-T-AV-LV	KR-W-E-IR	DTE-LETM	D-LRLLH-X
C.BR.04.04BR013	RVVDYRLG-G-II-LVL	-I-T-AV-LV	KR-W-E-IR	DTE-LETM	D-NLRLH-P
C.BW.00.00BW07621	AAVDYRIGVA-II-L	-T-AY-L-K	K-K-I	D-DHE-L-TM	D-LRLL-AN-G
C.CN.98.YNRL9840	LGIDYRLG-G-I-L	-T-Y-RLV	K-IR	DTE-S-M	D-NLRL-N
C.ET.02.02ET_288	ERVDA-R-G-G-I-L	-I-T-AY-L	X-E-IR	DTE-L-TM	D-LRLL-H
C.GE.03.03GEMZ033	AKVDYRLG-G-I-L	-I-T-AY-L	KRL-W-E-IR	G-DTE-L-T	D-NLRL-AN-G
C.IL.99.99IT7	FVDYRLG-G-I-L	-I-T-Y-LV	R-WN-IR	D-DAE-L-TM	D-LRLL-AN
C.NL.99.01NS565_10	WNLDYKLVG-II-L	-I-T-Y-RLV	K-W-K-IR	D-DTE-L-TM	D-NLRL-N
C.KE.00.KER2010	ARVDYRLG-G-I-L	-I-T-Y-LV	W-E-IR	DTE-LATM	D-LRLL-N
C.MM.99.mIDU101_3	LEIDYRLG-G-I-L	-I-T-Y-LVK	W-E-IR	DTE-L-TM	D-LRLL-N-E
C.MW.93.93MW_965	AKVDYRLG-G-I-L	-I-T-Y-LV	W-VK-IR	D-DTE-LATI	D-LRL-AN
C.SN.90.90SE_364	AKVDYRLG-G-I-L	-I-T-AV-LV	K-L-W-VK-IR	D-DVE-L-TM	D-LRLL-AT
C.SD.89.89SD_145	AGVDYRT-G-FS	-L-T-Y-LV	W-K-IR	D-LRLL-LATM	D-LRLL-AN
C.TZ.02.CO178	FLEDYRLG-G-I-L	-I-T-AYL-L	R-QI-K-IR	D-DTE-LATM	D-LRLL-N
C.UY.01.TRA3011	EKIDYRL-VG-I-S	-V-I-Y-LV	R-W-VE-IR	DTE-LETM	D-LGLLA-N-G
C.YE.02.02YE511	AKVDYRIG-A-F-L	-I-T-AV-LVI	W-VK-IR	D-DTE-L-TM	D-LRLL-IN
C.ZA.04.04ZASK164B1	VT-SEVGV-A-FIL-L	-L-T-YL-VK	W-E-IR	DTE-LA-M	D-QLRLL-N-GV
C.ZA.05.05ZAPSK240B1	AGVDYRIG-G-L-L	-I-T-Y-LV	L-K-IR	DTE-L-TM	D-E-LRLLVDN-NVR
C.ZM.02.02ZM108	EKVDYRTVA-FII-L	-I-T-AV-LV	W-K-IR	D-DHE-L-TM	D-LRLL-AN
D.CD.83.ELI	M-LG-I-A	-L-L-T-F-R-KK	R-C-L-IT	DREK-L-K	I
D.CM.01.01CM_0009BBY	M-SLE-L-VK	-L-L-T-AV-VK	NQ-I-E	DRE-L-T-M	A-DNI
D.KE.01.NKU3006	M-SLA-L-G-I-L	-L-L-V-T-CRKYQ	K-IR	D-DRE-L	M
D.KR.04.04KBH8	M-V-VI-X-TL	-I-T-F-CCRRLK	N-IR	D-DEE-L-T-M	N
D.TD.99.MN011	M-LV-L-A-L	-I-T-Y-K	Y-I	D-DKE-L-T-M	D-P-A-V-D
D.TZ.01.A280	MT-LE-T-A-I-S-L	-I-T-Y-R-K	K-W-IR	D-DTE-L-TM	N-I
D.UG.99.99UGK09259	M-LE-G-L-L	-L-L-T-F-CRRLK	W-IR	D-DKE-L-T-M	N-M
D.YE.01.01YE386	M-SLV-L-L	-L-L-T-F-CRRLK	W-IA	D-DRE-L	D-N-M
D.YE.02.02YE516	M-TLE-LS-I-A	-I-T-Y-R	Q-IR	DEE-L-T-M	A-N-A-D
D.ZA.90.R1	PL-V-L-L	-L-L-T-F-R-RK	Y-IR	DGE-L-K	YD-N
F1.AR.02.ARE933	MSYLLAIG-T-I	-L-T-AY-K-LV	R-N-YK-IR	DAE-LA-G	PFI-R-I-N
F1.BE.93.V1850	MSYLLAIG-A-I	-L-T-Y-K-LV	NK-YK-IR	DAE-LA-G	PFI-G-IN-N
F1.BR.01.01BR125	MSDLLVIGFT-A-L	-I-Y-LV	K-N-YE-IR	DAE-LA-G	PLI-G-I-N
F1.ES.x.P1146	HW-LATV-FI-L	-I-T-Y-LV	KR-W-E-IR	D-DAE-LA-G	V-PFI-G-IN-N
F1.FL.93.FIN9363	MSDLLAIT-FT-L	-V-T-F-K-LV	N-YI-IR	DAE-LA-G	K-PFI-G-N-N
F2.CM.02.02CM_0016BBY	MSYLI-LV-FI-L	-A-I-T-Y-K-Q	KR-N-YE-IR	DAE-LA-G	V-LFI-GNIN-N
F2.CM.95.MP257	MSLSL-V-A-VI-VL	-L-I-T-Y-K	KR-N-YE-I	DAE-LA-G	V-PLI-G-IN-N
F2.CM.97.CM53657	MSSLLTI-YI-L	-I-T-Y-K-L	KR-NK-YK-IR	DAE-LA-G	PFI-G-IN-N
G.BE.96.DRCBL	M-LE-S-G-I-S-A	-I-T-F-RK	KR-EK-L-IR	DTE-LAT-M	L-DFD-VG-N
G.CM.01.01CM_4049HAN	M-LE-A-G-I-F-A	-I-T-F-RK	K-K-LE-IR	DTE-LA-M	NFD-VG-N
G.CU.x.C174	MKSLAIV-G-I-G-A	-I-T-F-Q-E-RK	E-IR	DTE-LAT	DFH-VG-N
G.ES.00.X558	M-LE-A-G-I-S-A	-I-X-F-Q-E-RK	K-Q-IL-IR	DTE-LAT	DFD-VG-N
G.ES.99.X138	M-LE-A-G-I-G-A	-I-I-F-Q-E-RK	K-Q-IL-IR	DTE-LAT	DFD-IG-N
G.GH.03.03GH175G	M-SLE-S-G-I-S-A	-I-T-F-RK	K-EK-IR	DTE-LA	DFD-VG-N
G.KE.93.HH8793_12_1	M-SLE-S-G-I-F-A	-I-T-LVQ-E-RK	VE-IR	DRE-LAT-M	DFD-VG-N
G.NG.01.01NGPLD669	M-ALE-S-G-G-A	-A-I-T-F-RK	KR-L-IR	DTE-LAT	D-V-FD-VG-N
G.PX.P12695	M-SLV-L-G-I-G-A	-I-T-F-Q-E-RK	N-QIL-IR	DTE-LAT	DFD-VG-N
G.SE.93.SE6165	M-SLV-L-G-I-F-A	-I-T-F-E-RK	KR-GK-L-IR	D-DTE-LVT	DFD-VG-N
H.BE.93.V1991	*MN-LGIG-G-F	-T-AY-LXK	E-IR	D-DTE-L-K	-LNLGY-A
H.BE.93.V1997	CMY-IGIG-G-I-F	-T-Y-LVK	K-Q-I-G	DE-L-TM	R-LTFGY-A
H.CF.90.056	*MY-LGLG-G-TF	-VI-T-Y-K-LV	K-E-IG	D-DTE-L-K-M	-LNLGY-A
J.CD.97.J_97DC_KTB147	PL-D-A-G-I-V	-VI-T-F-V	N-IR	D-DTE-LEK	P-DL-N-N
J.SE.93.SE7887	MT-LQ-A	-FI-FL-GM-T-Y	K-L	D-DTE-LAD	R-P-DL-N-N
J.SE.94.SE7022	MVSLQ	-I-FFL-C-T-Y	K-L	D-DTE-LAE	P-DL-N-N
K.CD.97.EQTB11C	MV-LT-G-I	-A-L-I-T-AYL	VV-KR-NM-F-IR	DTE-LA-G	T-LILG-IN-N
K.CM.96.MP555	MVSL-AIS	-L-L-I-T-Y-LV	KR-NM-IR	DAE-LADIG	-L-LILGNI-N

	transmembrane domain	phos	phos	cytoplasmic domain	
	Vpu start	α-helix	α-helix	Vpu end	
B.FR.83.HXB2	TQPIPTVAIVALVVAIIIAIVVWSVIIIEYRKILRQRKIDRLIDRLIERAEDSGNESEG	...EISALV	..EMGVEMGHAPWDVD	.DL*
01_AE.CF.90.90CF11697	MSALQ	---G-I-L-L-L---	T-F-K	E-IR	---DTN-LAT---
01_AE.CN.05.FJ051	MT-LE	S---G-I-L-L-L-I-T-A-IRK	---D	---VK-IR	---DTD-LAK---
01_AE.CN.06.FJ054	MT-LE	S---G-I-L-L-L-T-A-FK	---	---K-IR	---DTD-LAK---
01_AE.HK.x.HK001	MT-LE	S---G-I-L-L-L-T-GL-V	---	---K-IR	---DTD-LAK---
01_AE.JP.93.93JP_NH1	MS-LE	S---G-I-L-L-L-T-A-K	---	---E-IR	---DTD-LAK---
01_AE.TH.01.01TH_R2184	MT-LE	S---AG-I-L-L-L-T-G-FK	---	---K-IR	---DTDDLAK---
01_AE.TH.02.OUR7691	MS-LE	S---G-I-L-L-L-T-V-LK	---	---K-IQ	---DTD-LAT---
01_AE.TH.90.CM240	MT-LE	S---G-I-L-L-L-T-A-VK	---	---VK-IR	---DTD-LAK---
01_AE.US.00.00US_MSC1164	MT-LL	S---G-T-L-L-L-TL-A-F	---	---VK-IR	---E---DTD-LAK---
02_AG.CM.02.02CM_4082STN	M-LLA	A---G---G-A---T-L-RKK	---IR	---D-DTE-L-T-M---
02_AG.EC.x.EC141	M-SLE	A---G---F-A-L---L-RKK	---IR	---D-DTE-L-T-M---
02_AG.FR.91.DJ264	MKSLE	A---G---AG---T-V-RKK	---K-L-IR	---D-DTE-L-T-L---
02_AG.GH.03.GHNJ196	M-SLV	A---G---F-A---T-EY---R-RKKQ	---L-IR	---D-DTE-L-M---
02_AG.NG.01.PL0710	M-SLK	A---G---A---A-T-Y-RKKR	---K-IR	---G-D-DTE-L-T-M---
02_AG.NG.x.IBNG	M-LT	T---G---F-A---Y-RKK	---L-IR	---D-DTE-L-T-M---
02_AG.SE.94.SE7812	MNSLD	---G---F-A---T-Y-RKK	---L-IR	---D-DTE-L-T-M---
02_AG.SN.98.MP1211	M-LAVA	G---F-L-T-F-RKG	---K-L-IR	---D-DTE-L-T-L---
02_AG.UZ.02.02UZ710	M-SLE	---G---F-A---T-Y-RKK	---K-L-IR	---D-DIE-L-T-M---
03_AB.RU.97.KAL153_2	M-SLA	A---VG---G-F-F	---	---IR	---DQE_AL.M---
04_cpx.CY.94.CY032	MLFWE	W---G---L-V---TL-F-K-LRR-S	---Y-IR	---D-DAE-L-T---
05_DF.BE.x.VII1310	MSDLTI	V---I-L-L---I-Y---K-LVN	---YK-IR	---DAE-LA-G---
06_cpx.AU.96.BFP90	M-ALE	A---G---FLA---T-F-Q-E-RKK	---EK-L-IR	---DTD-LAT-M---
06_cpx.RU.05.04RU001	M-ALE	S---G-I-SLA---T-L-Q-E-RKEK	---L-IR	---DTD-LAT-M---
07_BC.CN.97.CN54	M-ALT	L---T---T---T---F	---	---IR	---GD-DQE-L-FM---
08_BC.CN.97.97CNGX_6F	MLDL	EL_VG-I-L-L-T-Y-RLVK	---W-K-IR	---DTE-L-TM---
09_cpx.GH.96.96GH29PI	M-SLQ	AS-GII-F-A-IA-T-F-V-RKKR	---IR	---DTE-LAK---
10_CD.TZ.96.96TZ_BF061	ARVYKLRV	G-I-L-L---IR-Y-Q-LVPET	---W-K-IR	---R-HD-DTE-L-TM---
11_cpx.GR.x.GR17	*LCEQLSIVALVSSIXL	-I-T-AG--RL-H	---	---IDY.Y-IRX	---DTE-LAQ---
12_BF.AR.99.ARMA159	M-SLVTL	---A-L-A-L---	---	---IR	---DAE-LA-FG---
13_cpx.CM.96.1849	MNAL	-IS---I-F-L-V---T-AY---LVHK	---EK-L-IR	---D-TDDLAK---
14_BG.DE.01.9196_01	MK-LE	S---G-I-S-A-I-I-F-Q-KE-RKK	---Q-IL-IRD	---DTE-L-T---
14_BG.ES.99.X397	M-SLE	S---G-I-G-A-I-I-F-Q-KE-RKK	---Q-IL-IR	---DTE-LAI---
15_01B.TH.99.99TH_MU2079	MS-LE	S---G-I---L-L---T-A-I-N-K	---	---VK-IR-E	---DTD-LAK---
16_A2D.KR.97.97KR004	MNSLQ	LS-G-I-F-L-V-T-F-K-KN	---W-K-IS	---DTE-L-T-M---
18_cpx.CU.99.CU76	MXFE	-G-I-I-F-V-C-T-Y---KK	---G-FE-IRD	---Y---DTE-L-K-M---
19_cpx.CU.99.CU7	M-ALQ	A---G-I-F-A---T-Y---RKEEK	---L-IR	---D-DTE-L-T-M---
20_BG.CU.03.CB471	M-SLE	A---G-I-G-A-I-A-F-Q-KE-RKILQ	---IRK	---DTD-LAT---
21_A2D.KE.91.KNH1254	M-LQ	---A---A---T-V-R-KD	---IR	---DEE-L-K---
23_BG.CU.03.CB118	MHSLE	A-V-G-I-G-A-I-A-F-L-KQ-RKK	---ILE-IR	---DTE-LA---
24_BG.CU.03.CB378	M-SLE	A---G-I-G-A-I-A-IF-Q-KE-RKILE	---IR	---DTD-LAT---
25_cpx.CM.01.101BA	M-LE	C---G-I-F-A-V-T-F-IK-RKK	---EQ-L-IRT	---DTE-L-T-M---
27_cpx.CD.97.97CDKTB49	M-ALE	S---I-F-VV---T-Y---K-L	---	---I-E-IRD	---DTE-LAR---
28_BFBR.99.BREPM12609	M-ALI	A---VA---T-F-F	---	---IR	---DAE-LE-G---
29_BFBR.02.BREPM119	ML-LQ	---S-VA---T-F-F	---	---R-IRD	---DQE-L---
31_BC.BR.02.110PA	DKIDYRX	VX-II-L---I-T-Y-LVKW	---VK-IR	---DTE-LATM---
33_01B.MY.05.05MYKL007_1	KT-LE	S---G-IA-L-S---T-A-LQ-K	---	---VK-IR	---DTD-LAK---
34_01B.TH.99.OUR2478P	PL	E-S---G-I-L-L-L-I-A-EE	---VK-IQ	---DTD-LAK---
35_AD.AF.05.05AF095	PL	H-L---G---L-L-T-G-LKRLI	---K-IS	---D-DTD-LAK---
36_cpx.CM.00.00CMNYU830	SL	Q-A-AG---F-A-GG-T-L-K-IK	---LE-IK-K	---NTE-L-T---
37_cpx.CM.00.00CMNYU926	PL	E-YS-G---F-A---T-L-K-RKN	---IR	---D-DTE-L-T-M---
42_BFLU.03.luBF_05_03	SL	V-L---XA---L---KA	---IS	---D-DHE-L-X---
N.CM.02.DJ00131	YMLEWGF	ALGVAI	---VII-VLLYK-K-LKLE	---EIQRO-IRD-T---
N.CM.04.04CM_1015_04	YMLS	LGFSALGVAI	---VII-VLLYK-KE-KLE	---EIQKO-VRD---
N.CM.04.04CM_1131_03	YMLE	LGFIAGVAI	---VII-VLLYK-KE-KLE	---EKIRO-IRD-KE-K---
N.CM.95.YBF30	YMLS	LGFIAGAA	S-VI-ALLYR-K-KLE	---KHIRO-IR-E---
N.CM.97.YBF106	YMLW	GFALGVAI	---VII-VLLYK-K-KLE	---KHIRO-IR-E---
O.BE.87.ANT70	MHRDL	L-IISALL-NVLI-GFILRK-LEQKDRKE	---E-LERLR-IR-IRD	---DY-N-EEEOEVRDL---
O.CM.91.MVP5180	MHOENLL	LI-SALCL-NVLI-LFNLR-LVQRKDRREO	---LERLR-IR-IRD	---DY-N-EEEOEVRDL---
O.CM.96.96CMAABB637	MNYKELL	SLIVS-LLLA--I-MFILKK-LEQKDRRE	---ELLKR-E-X-XRD	---DY-N-DEEOEVRDL---
O.CM.98.98CMA104	HYRDL	LAL-IVSALL-L-NVLL-FILRK-LEQKDRRE	---E-LERLR-IR-IRD	---DY-N-EEEOEVRDL---
O.CM.99.99CMU4122	QHKDL	LIL-ITSALL-L-NVLI-LF-LKQCLEQKTKRE	---E-I-RLR-IR-IED	---DY-N-EEEOEVRDL---
O.SN.99.SEMP1299	MHRDL	L-LITTSALLTNV-L-TEILRL-QKDKRE	---E-LERLR-IR-IED	---DY-N-TEEOEVRDL---
O.US.99.99USTWLA	QYKDL	LAYSALF-LITLAI-MF-LKLCLEQKGKR	---E-ROKIRO-IR-E	---DY-N-DEEVRDL---
O.FR.92.VAU	HHOGLL	L-LIALL-L-NVLL-MFNLR-LEQKDRTA	---E-LER-R-IR-VED	---DY-N-EEEOEVRDL---
CPZ.CD.90.ANT	CPZ	CM.01.SIVcpz	CAM13	CPZ	CM.05.SIVcpz
CPZ	CM.05.SIVcpz	EK505	CPZ	CM.05.SIVcpz	LB7
CPZ	CM.05.SIVcpz	MB66	CPZ	CM.05.SIVcpz	MT145
CPZ	GA.88.GAB1	CPZ	TZ.01.TAN1	CPZ	US.85.CPZUS

HIV-1/SIVcpz Proteins

glycosylation NDT VI

Table with columns: Env start, signal peptide gp120 start, glycosylation NVT, and protein ID. Lists various HIV/SIV proteins and their amino acid sequences.

	Env start	signal peptide_gp120 start	glycosylation NVT	glycosylation NDT	
B.FR.83.HXB2	MRVKE...KYQHLW...	RWGWRWGTMLLGLML...ICSATEK...LWVTYVYGVPMVKEATTTL*FCASDAKAYDTEVHNWATHACVPTDPNPQEVVLVNVTFENFMKMKNDMVEQMHEDIISLWDQ...SLKPKVKLT...PLCVS...LKCTDLKND...			137
01_AE.CF.90.90CF11697	--G--GTQMNWP	...K--LI--LVI--SDT	RD-D--N-QE--V--IH-E--D--N--Q-V--		136
01_AE.CN.05.FJ051	...--TQMNWP	...K--LI--LVI--SDN	RD-D--HK--MH-E--D--N--Q-V--		136
01_AE.CM.06.FJ054	...--TQMNWP	...K--LI--LVI--SDN	RD-D--R-HE--I-V--IH-E--D--N--Q-V--		136
01_AE.HK.X.HK001	...--TQMNWP	...K--LI--LVI--SDN	RD-D--HE--IK-GD--X-K-A--Q-V--		136
01_AE.JP.93.93JP_NH1	...--TQMNWP	...K--LI--LVI--SDN	RD-D--HE--IH-E--D--N--Q-V--		136
01_AE.TH.01.01TH_R2184	...--TQMTWPN	...K--LI--LVI--SDN	RD-D--HE-A--IH-E--D--N--Q-V--		136
01_AE.TH.02.OUR7691	...GTQMSWP	...TG-LII-LVI--S-N	RD-D--IH-E--D--N--Q-V--		136
01_AE.TH.90.CM240	...--TQMNWP	...K--LI--LVI--SDN	RD-D--R-HE--IH-E--D--N--Q-V--		136
01_AE.US.00.00US_MSC1164	...--TQMNWP	...K--LI--LVI--SD	RD-E--R-HE--IY-E--D--K-A--Q-V--		136
02_AG.CM.02.02CM_4082STN	--MGIOKN-PL	...--MIIFWIMI--N-KD	RD-E--S-MP-E--D--N--V--G-R--T-H--NFHRS		135
02_AG.EC.x.ECJ41	--MGIOKN-PL	...--MIIFWIMI--N-AN	RD-E--Q--S-ING--D--N--V--G-R--T-H--NFHRS		135
02_AG.FR.91.DJ264	--MGIORN-PLF	...K--MIIFWIMI--N-EK	RD-E--V--S-IH-K--K--N--E--T-D--YNVSSN		135
02_AG.GH.03.GHNJ196	--MGTKN-PC	...--IIIFWIML--KGED	RD-D--S--S-IY-E--N--E--T-E--HSVT-S		135
02_AG.NG.01.PL0710	G-M-TQKSCPP	...K--MIIFWIMI--N-D	RD-E--A--IH-E--K--K--N--T-N--RL-N		135
02_AG.NG.x.IBNG	--MGIOKN-PL	...--MIIFWIMI--N-EQ	RD-E--IH-E--K--N--N--T-D--HNFN-S		135
02_AG.SE.94.SE7812	--MGIOKN-PL	...--MIIFWIMI--N-GN	RD-E--IH-G--D--S--Q--T-H--Q-NLTS		135
02_AG.SN.98.MP1211	--MGILKSCPPF	...RWGML--WI-I--N-EN	RRD-E--E--S-MH-E--N--V--T-E--HNVT		136
02_AG.UZ.02.02UZ710	--MGILKN-PP	...--IIIFWIMI--K-QD	RD-E--I--S-IH-E--N--N--T-D--H-Y-SN		135
03_AB.RU.97.KAL153_2	--IRK-HL	...--LF--N	RD-E--SK--Y--S-IP-K--G--N--T-N--KE		133
04_cpx.CY.94.CY032	--MGMRN-P	...E--LI--LVI--SMN	RD-E--E--EK-I--A-I--ME-G-A-S--T-FT-INATTT		136
05_DF.BE.x.VII310	--RGMORNP-G	...K--LLF--I-I--D--F	RD-E--G--K--G--T--T--N--F-A-N		136
06_cpx.AU.96.BFP90	--GIOTSW	...K--LI--LVI--SKN-M	A-ED-D-I--SPDK--IS-K--N--D--E--T-T--NATLG		137
06_cpx.RU.05.04RU001	--RGTQMNWP	...K--LIF-LVI--SKD	A-ED-D-I--HS-K--D--IS-E--H--K--E--T-H--KNVT-N		136
07_BC.CN.97.CNS4	--TGIRKN-R	...S-VGN	A--M-E--E--N--Q-V--T-E--RNVSSN		136
08_BC.CN.97.97CNGX_6F	--RGTRRN-OW	...I--VLGFW--NVEGN	K--E--S-I-ME--N--N--VT--T-E--RNVSSN		136
09_cpx.GH.96.96GH29PI	--IQRRW-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-TGIQRNCROW	...I--ILGFW--N-R	T--KA-A-I--R-I-E--G-D--G-I--T-N--NGINAT		136
11_cpx.GR.x.GR17	--TORNWHN	...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	--RGMORNP-G	...N-FLF--I-I--N-N	RD-E--S-ER--E--D--N--T--T-N--SNANAT		136
13_cpx.CM.96.1849	--ARGIQMTW-N	...K--LI--LVI--SMN	RD-E--S--K--IN-G--E--N--T--I--S--T-N--NNITS		136
14_BG.DE.01.9196_01	KA-GIORNC-R	...K--VI--N-V-Q	T--R--S-L-E--N--D--N--NET		136
14_BG.ES.99.X397	KA-GTORNW-S	...K--LI--LVI--SDN	RD-E--A--A-E--E--N--D--Q--E--T-N--FNIT		136
15_01B.TH.99.99TH_MU2079	--TQMNWP	...K--LI--LVI--A-N	RD-E--A-G--S-EI-G--E--Q--N--R--T-N--R		136
16_A2D.KR.97.97KR004	--RGIORN	...K--ILI--SK--D	RD-E--A--IN-E--G--T-N--SRV-T		136
18_cpx.CU.99.CU76	--RGTORN-PS	...K--II--L-M--DH	RD-V--S-S-S--D-IT-G--N--E--T-N--NVXSS		136
19_cpx.CU.99.CU7	--MGTTRS-PP	...N-LI--LVI--FDN	RD-E--S-IV-E--E--N--A--T-D		130
20_BG.CU.03.CB471	--RGIPIWNL	...T--LII-LGI--SSN	ED-Y-P--VA-S--H-IP-Q--N--E--T-N--SVDON		136
21_A2D.KE.91.KNH1254	--R-MKRCN-N	...K--I--F--DN	R-NAS--TK-A-I--IE-E--N--E--T-N--AGSTTS		136
23_BG.CU.03.CB118	--GIQKNW-R	...T--LII-LGI--SDN	ED-Y-P--S-S--IP-X--H--E--T-N--KVOCE		136
24_BG.CU.03.CB378	--GIQMNW	...T--LII-LGI--SMN	RD-D-P--S-S--IP-E--N--E--T-N--T-MSN		136
25_cpx.CM.01.101BA	--R-IQRN-PL	...--LII-LGI--SDN	RD-E-X--SP-A--S-IE-T--N--E--Q--T-N--VT-NCS		140
27_cpx.CD.97.97CDKTB49	--GTQRNW-G	...L-I--A--K-D	RD-D--HT--IT-E--N--D--E--T-N--NANAT		136
28_BFBR.99.BREPM12609	--RGMQRNW-G	...T--LLF--I-I--N-NQ-S	RD-E--N-VK-A-I--I--K--N--E--I--T-N--NNVT		136
29_BFBR.02.BREPM119	--ARGTRKNC	...E--IL--EQ	RD-E--N--A--D-IP-K--Q--Q--T-N--S-G-N		135
31_CD.BR.02.110PA	--GIQRNW-OW	...I--ILGFW--YNVRGN	R-K--S-LD-G--E--D--Q--V--*--Q--T-N--S-ATSN		136
33_01B.MY.05.05MYKL007_1	--TRMNWP	...K--LI--LVI--M-SMN	RD-D--HE--IH-K-I--K-A--Q-V--Q--I--H--NANLT		136
34_01B.TH.99.OUR2478P	--TQMNWLN	...K--LI--LVI--SDN	RD-D--HE--Y--IH-E--N--Q-V--DQ--T-N--NAALTNTVTI		140
35_AD.AF.05.05AF095	T--MGTQMNW-N	...I--II--V-N	RD-E--E-M--IP-G--D--K--T--DQ--T--NATTNI		137
36_cpx.CM.00.00CMNYU830	--ARGTQRNL	...K--LII-LVI--SDL	RD-N--ID--N--Q--DQ--T-E--QNFQSNNH		139
37_cpx.CM.00.00CMNYU926	--MGIORN-PP	...N-LI--LVI--NNL	RD-D--T-VP--LPX--X-KINVG--KD--N-QI--GPKL--E-X--TX-NSSHKV-PQDRE		143
42_BFLU.03.luBF_05_03	--TGIMKN-X	...K--I--I--S	X--I--N--DQ--T-N--NXXXXXXXXXX		141
N.CM.02.DJ00131	V--MMGMOIGWPF	...C-MIS-T--G-K--Y-A	RDVE-V--HS-A-I--Q--P-D--P--E-K-A-VQ--E--T-M-N-SNSNGN		129
N.CM.04.04CM_1015_04	--MMGOSGWPF	...C--ISSI--G-EQ--H	RDTE-V--HS-A-I--Q--P-I--H-D--N-A-Q--E--X--T-MH-XNSNGX		129
N.CM.04.04CM_1131_03	K--MMGOSGWTRVLGMMGOWPFC	...C--ISSI--G-EQ--H	R-TD-V--HS-A-I--Q--H-I--K-D--N-A-Q--E--T-MH-N-SNGF		141
N.CM.95.YBF30	K--MMGOSGWMGM	...KSGWLLFYLLVSLIKV	R-E-V--HS-A-I--Q--L-P--K--E-K-AD--Q--E--T-ML-N-SYGE		138
N.CM.97.YBF106	R--MMGOSGWPF	...C--ISSI--T-G-T	R-E-V--HS-A-I--Q--L-P--K--E-K-A-Q--E--T-ML-N-SNSNGN		129
O.BE.87.ANT70	KAM-KRN-KLWTL	...YIAMLII-P-LSLRO	YA-A--ED-PV--NLTS-K-I-SO--T-Y-YP-H--DD-I--Y--Q--OM-F--Q-ME-NIAGT		133
O.CM.91.MVP5180	K-MKNNRKSWSL	...YIAMLII-P-LSYS-Q	YA-S--E-APV--NLTS-Q-I-SO--H-FP-G--D-DI--Y--D--E--M-F--Q-MN-V-OTN		134
O.CM.96.96CMA8B637	K-M-MKN-KLIL	...CI-MAL-I-P-LNSDQ	YA-FS--ED-P--NLTS-K-I-SO--S-H-YP-I--DK-I-E-Y-D--Q--E--OM-F--Q-MN--VNAT		133
O.CM.98.98CMA104	KAMKKKN-RLWIL	...YIYM-LIT-P-LSSSQ	FYA-A--EN-NPV--NLTS-Q-I-AO--L-HP-H--DR-DI-E-Y-D--Q--TD--E--M-FM--Q-ME-N-INTT		133
O.CM.99.99CMU4122	--AM-KRNRGLIL	...YIYMALII-P-LSYDQ	YA-S--E-KPX--VNLTS-K-I-SO--T-H-YP-K--DK-I-E-Y--Q--E--OM-F--Q-MN--NVNDT		133
O.SN.99.SEMPI299	K-M-ORNRKLGIL	...CIVMALII-P-LSYNO	HYA-A--E-PV--NLTS-Q-I-SO--S-Y-YP-TK--D-I--Y--D--E--OM-F--Q-MN--NYVGG		133
O.US.99.99USTWLA	K-M-KNRKSWEL	...CXIMALII-P-LNS-Q	YA-S--ED-SP--VNLTS-Q-I-SO--S-N-YD-K--DY--I--Y--D--E--OM-F--Q-MN--V-DN		133
O.FR.92.VAU	KAMKRNKRRLGIL	...CLII-ALII-P-LSCNQ	YA-S--ED-KP--NLTS-Q-I-Q--S-N-YE-K--DK-I-E-Y-D--Q--D--OM-F--Q-MN--I-S		133
CPZ.CD.90.ANT	--KPIHII	...WG-AL-I-OFIEKGTNEYD	F--RN-P--TN-SMTS--TS--I-D-IV-R-L-TSVN--AY-Y-S-T-MXQ-FQ--H--M-IK-MN--GYNG		127
CPZ.CM.01.SIVcpzCAM13	K-M-ERRKSFNFG	...YMTM-ICMAL-I-P-LTSD	E--RN-V-V--HS-A-I--H--D--N-Q--Q--S--T-M--SNVTRQ		137
CPZ.CM.05.SIVcpzEK505	K-T-MQKNWLI	...CC--IG-I-KIIGSE	RD-E-V--HS-A-I--Q--LIP--R-D--N-D--Q--E--T-S--SSWRSV		130
CPZ.CM.05.SIVcpzLB7	--GTWTSMPYY	...VWGI-LS-CSLII-SSTKN	N--EA--WQKA-A-I-SO--D-KAIE-S--T-A-N-QQ-V--I--T--H-STP-F		137
CPZ.CM.05.SIVcpzMB66	KAM-TQRNCR-T-S	...LKEIILCT-VLGI-GTIKCEDNM	R--D--QNP--SO--S-N--E-IE-T--Y-A-E-N--D-Q--VN--V--I--T-N-SLF-CI		145
CPZ.CM.05.SIVcpzMT145	K-MGKMYWRNS	...CRSISII-LIGW-A-CFGEENW	R-K--P--S-A-I--Q--T--L--P--Y--E--D--Q--E--OM-F--T-NPT-TIK		140
CPZ.GA.88.GAB1	K-M-KL-RDWS	...LSIITIIITLII	P-LTS--HD-PP--HS-A-I--Q--S--F-P--I-S--E--N--D--Q--E--OM-F--T-Q-SKANFS		136
CPZ.TZ.01.TAN1	*KMK-NLIGIT-I	...LIITI-G-GFS--YYT-F	OP--DP--DITSRDK-I--N-L--Y--T-A-SIR--EE-Y-Q-K-L-FQ--F--F-IK-MT--MTNT		128
CPZ.US.85.CPZUS	K-M-KK-RLW-S	...YCL-SSLII-PGLSS	A--RDVE--KQ-A-I--Q--H-P--K-D--E-N-A--Q--I--T-MT-LNPDSN		131

Table with columns for glycosylation sites (V1, V2, NVS, NGS, NFT) and corresponding amino acid sequences for various HIV-1/SIVcpz proteins. The table lists protein names on the left and their amino acid sequences on the right, with glycosylation sites indicated by asterisks and numbers.

	glycosylation NNT		glycosylation NCT		V3 tip		glycosylation NNT		glycosylation NKT		glycosylation NST		glycosylation NST		glycosylation NNT	
	glycosylation NFT	glycosylation NTS	glycosylation NFT	glycosylation NTS	V3	V3	glycosylation NNT	glycosylation NKT	CD4	glycosylation NST	glycosylation NST	V4	glycosylation NNT	glycosylation NNT	glycosylation NNT	glycosylation NNT
B.FR.83.HXB2	FTDNAKTIIVOLNTSVEINCTRPNNTRKRIR.IQ.RGPGRAVFT.IG.KIGM...
A1.GE.99.999GEMZ011	I--G-I--KP-N-T-I-G-I-TS--	I--I--OT-YG-T-DV-DI--	I--K-Y-V-N-TE--OK-STQ--NY-N--
A1.KE.00.KER2008	I--T-N--KOP--RGVH--	I--L--R-Y--T.OVV-DI--	I--Y-V-KS-DD-KVVYQ--KY-N--
A1.KE.00.KNH1144	I--N--N--VPR--ES--	I--I--Q-FA--A-DI-DI--	I--V--V--SQ--E--H-K--
A1.KE.00.RSM4024	I--N--N--DKP-K--SVY--	I--I--Q-FA--A-DI-DI--	I--V--V--SQ--E--H-K--
A1.KE.00.MSA4069	I--N--N--VNP-R--RD.G--	I--I--OT-YA--A-AI-DI--	I--Y-V--A--K-OEV-KO-OE-N--
A1.KE.00.NKU3005	IS--I--I--P--T--SV--	I--I--O-Y--TD-I-DI--	I--V--V--KSE--K--O-VIO--KHWN--
A1.RU.00.RU00051	I--T--T--SKP-N-T-I-G-TS--	I--I--OT-YA--T-DI-DI--	I--K-V--TE--G-ATEE-RKO-OTY-P--
A1.RU.03.03RU20_06_13	I--N--G--X--TEP-N-T-I-G-TS--	I--I--OT-YA--TXDV-DI--	I--K-Y-XV--A--KS-STO--KY-N--
A1.RW.93.93RW_024	I--N--N--L--VKP-R--Y--FR-V--	I--I--OT-YA--SDIV-DI--	I--K--V--SE--E--Q-VVQ--GY--
A1.SE.95.SE8891	I--N--I--N--F-K--I--I--S--	I--I--O-YA--T-DI-DI--	I--Y-DV-N-TE--EA-QKVVNO-KTH-K--
A1.SE.95.UGSE8131	IS--N--FTPK--I--S--	I--I--O-YG--M-DI-DI--	I--K-V--S--E--KV-IO--KYWN--
A1.TZ.01.A173	I--N--N--FTPE-Q-I--I--VH--	I--I--OT-YA--T-DI-DI--	I--Y-TV-N-TE--T-QKVGKO-DLYE-N--
A1.UA.01.01UADN139	I--G--N--TEP-N-T-I-G-TS--	I--I--OT-YA--T-EI-P--	I--K-V--VA--OK-SIO--KY-N--
A1.UG.99.99UG0037	I--N--V--N--E--T--R--S--A--	I--I--OT-YA--T-DI-DI--	I--V--G--GSO--R--VVGE--KYWN--
A1.UG.99.99UGA07072	I--N--N--KP-N-T-I--OSV--	I--I--O-YA--T-DI-DI--	I--Y-DV-N-TE--K-OEV-YO-GKH--
A1.UZ.02.02UZ0659	I--G--N--TEP-N-T-I-G-TS--	I--I--OT-YA--T-DV-DI--	I--K-Y-V--A--OK-STO--KYY-N--
A2.CD.97.97CDKS10	I--N--N--F-E--P--T--I--G--P--	I--I--O-Y--TSDI-DI--	I--Y-S--NKT--DAS--OKV-EQ--KH-P--
A2.CD.97.97CDKTB48	I--N--N--F-K-P--P--S--	I--F--O--Y--NNMI-DI--	I--ITE--A--KVVEQ--H-P--
A2.CY.94.94CY017_41	I--N--N--FTKP-L-T-I--S--	I--F--O--Y--TNEI-DI--	I--NKTL--D--QKV-EQ--K-PP--
B.AR.04.04AR151516	I--N--I--ET--S--SVH--	I--I--YA--T-EI-DI--	I--E--R--L--R--MT--
B.AU.87.MBC925	I--N--N--KOP--S--H--	I--I--YA--T-EI-DI--	I--Y-TL--N--R--D--E--G--K--
B.BO.99.BO10172	I--N--N--K--IS--S--H--	I--I--LVA--T-EI-DI--	I--ATE--E--VF--G--VN--
B.BR.03.BREP2M012	I--I--I--K--S--H--M-W--	I--I--E-T--I--	I--GTN--T--R--V--O--E--FN--
B.CA.97.CANB3FULL	I--T--ET--S--H--IA--	I--I--EIT-DI--	I--A--E--VK--G--K--
B.CO.05.05CNHB_hp3	I--I--K--M--V--L--Y--	I--I--OIV-DI--	I--K--L--PO--DD--R--VVG--
B.CO.01.PCM001	I--I--K--M--V--L--Y--	I--I--OIV-DI--	I--K--L--PO--DD--R--VVG--
B.GB.83.CAM1	I--N--N--KEP--L--G--H--L--K--YA--	I--I--S-DIV--I--	I--TL--N--T--EK--G--V--N--
B.GB.86.GB8	I--M--T--KEA--L--G--Y--I--R--Y--	I--I--TDR-DI--	I--L--ST--R--VT--K--YK--
B.GE.03.03GEMZ010	I--V--E--Q--G--S--H--I--YA--	I--I--T-EI-DI--	I--L--SDO--S--H--T--K--
B.IT.05.SG1	I--S--T--KT--I--G--S--H--M--W--	I--I--EIV-V--	I--L--TA--D--LVVE--K--PK--
B.JP.05.DR6538	I--N--N--KOP--GY--I--VS--	I--I--A-TKITE-DT--	I--KK--Y--TL--NKTO--D--R--V--E--
B.KR.05.05CSR3	I--N--N--EL--K--S--H--I--K--YA--	I--I--A-DI-T--	I--KS--TD--K--D--S--
B.NL.00.671.00T36	I--N--N--KE--L--S--H--I--K--YA--	I--I--T-DI--	I--NGM--E--GVGE--OKH-P--
B.RU.04.04RU128005	I--N--N--KOP--Q--S--S--I--X--YA--	I--I--T-DI--	I--X--L--G--D--TK--E--VK--X--YN--
B.TH.00.00TH_C3198	I--S--T--V--E--Q--I--G--H--I--O--Y--	I--I--T-DI--	I--K--L--STN--K--VK--R--
B.UA.01.01UAKV167	I--V--E--IT--G--H--I--O--Y--	I--I--T-EI-DI--	I--L--O--K--E--VI--GK--NS--
B.US.04.ES10_53	I--T--T--H--N--G--T--G--QS--H--IA--	I--I--AIT--I--	I--K--L--IVS--D--VVE--Q--H-P--
B.US.99.PRB959_03	I--N--T--H--E--T--ES--	I--I--OT-YA--T-DI-DI--	I--E--TTA--K--HEFK--KKL-P--
C.AR.01.ARG4006	IS--V--H--KEP--S--H--I--YA--	I--I--T-DI--	I--E--TTA--K--HEFK--KKL-P--
C.BW.00.00BW07621	I--N--H--E--V--G--RSV--	I--I--OT-YA--T-DI-DI--	I--KT--K--SR--GE--K--H-P--
C.CN.98.YNRL9840	I--N--V--H--O--V--TS--	I--I--OT-YA--T-EI-DI--	I--E--E--ORVGE--A--H-P--
C.ET.02.02ET_288	I--N--I--EP--QS--	I--I--OT-YA--T-EI-DI--	I--GDR--K--ENKVER--OK--P--
C.GE.03.03GEMZ033	I--N--I--H--V--G--QS--	I--I--OT-YA--R-DI-DI--	I--NEN--T--OKV--E--KKY-P--
C.IL.99.99CIT7	I--N--I--H--KH--V--SM--	I--I--OT-YA--T-SI-DI--	I--Y--T--NGM--E--GVGE--OKH-P--
C.HN.99.01NS65_10	I--N--V--H--K--I--V--I--G--SV--	I--I--OT-YA--T-DI-DI--	I--K--Y--KE--O--VRK--E--H-P--
C.KE.00.KER2010	I--HF--OP--E--S--	I--I--OT-YA--T-EI-DI--	I--KO--E--ORVGR--A--L-P--
C.MM.99.mIDU101_3	I--V--H--E--V--A--SV--	I--I--OT-YA--T-AI-DI--	I--TI--K--OGVEK--K--H-P--
C.SN.90.90SE_364	I--N--H--E--I--V--A--SM--	I--I--OT-YA--T-DI-DI--	I--SGN--A--EKVKG--Q--H-PG--
C.SD.89.89SDM145	I--S--H--P--Q--V--YA--SV--	I--I--OT-YA--T-DI-DI--	I--H--G--Q--V--TNDI-DI--
C.TZ.02.CO178	I--S--I--H--E--P--V--TS--	I--I--O-YA--T-DI--	I--EST--K--OWGI--E--HFN--
C.UY.01.TRA3011	I--T--I--H--E--TS--	I--I--OT-YA--T-EI-DI--	I--KKA--KE--OKVRN--A--H-P--
C.YE.02.02YE511	I--V--H--K--D--K--SM--	I--I--OT-YA--T-SI-DI--	I--EK--D--OE--GK--K--H-P--
C.ZA.04.04ZASK164B1	I--N--H--E--A--S--	I--I--OT-YA--T-EI-DI--	I--KKA--EI--OKVKE--K--H-SD--
C.ZA.05.05ZAPSK240B1	I--L--G--H--ETXK--V--A--S--	I--I--O-Y--A--HTD-DI--	I--TE--K--ERVKK--G--H-P--
C.ZM.02.02ZM108	I--S--I--H--P--Q--V--H--SV--	I--I--O-Y--T-EV-DI--	I--A--EGL--E--HNSVE--F--Y-P--
D.CD.83.EL1	I--N--N--AH--E--K--T--A--YO--Q--TP--	I--I--OSLY--TS--RSII--	I--G--O--SK--Q--V--R--GTLNKK--
D.CM.01.01CM_0009BBY	I--V--N--F--TIN--T--I--G--S--H--	I--I--O-LYA--ND-I-DI--	I--G--K--GKD--QEV-R--GNLYN--
D.KE.01.NKU3006	I--N--E--T--S--OS--	I--I--O-Y--TE-V-I--	I--L--F--E--NGTE--K--V--K--GNL-P--
D.KR.04.04KBH8	I--N--N--V--P--K--E--RG--H--IA--	I--I--O-LFA--TTNV-DI--	I--R--ST--E--O--L--K--GDL-MK--
D.TD.99.MN011	I--L--T--N--ETIN--T--I--RSVH--	I--I--O-LY--TN-V-DI--	I--G--O--QEV-R--GAALNT--
D.TZ.02.A280	I--N--V--ET--K--S--	I--I--O-V--OT-F--AE-VT-DI--	I--K--Y--G--E--DK--Q--V--T--GDLNKK--
D.UG.99.99UGK09259	I--N--I--H--E--V--Y--G--H--V--Q--Y--	I--I--TNIW-DI--	I--L--DE--K--OKVTNR--EKLS-I--
D.YE.01.01YE386	I--AN--V--N--E--T--Y--OGVH--	I--I--LY--TN--I--I--	I--K--TV--DE--K--OKVTNR--EKLS-I--
D.YE.02.02YE516	I--I--ETIN--T--I--OQ--H--IA--	I--I--O-LY--TN-V-DI--	I--GTO--OK--K--GKLYNK--
D.ZA.90.R1	I--N--N--A--Y--QNTQ--	I--I--OT-YA--SRMI-DI--	I--EE--KR--Q--V--I--GNLLN--
F1.AR.02.ARE933	IS--H--E--Q--S--H--L--O--YA--	I--I--T-DI--I--	I--K--V--GMQ--T--ERVKA--KPL-P--
F1.BE.93.V1850	ISN--H--E--Q--G--H--L--O--YA--	I--I--T-DI--I--	I--K--V--GTO--E--YVKA--KSH-P--
F1.BR.01.01BR125	IS--H--E--Q--L--S--H--I--O--YA--	I--I--T-DI--I--	I--K--V--G--O--K--GRVRA--KSH-PD--
F1.E.S.x.P1146	I--H--T--H--S--S--I--O--YA--	I--I--T-DI--I--	I--K--V--G--O--K--GRVRA--KSH-PD--
F1.FL.93.FIN9363	IS--H--E--Q--S--S--I--O--YA--	I--I--T-EI-DI--	I--K--V--GEO--K--DRVKA--KLH-N--
F2.CM.02.02CM_0016BBY	I--T--F--E--K--S--S--I--O--YA--	I--I--T-EI-DI--	I--K--T--NGTL--A--NRV--AEVKLNTN--
F2.CM.95.MP257	IS--T--FKNP--K--RS--H--I--YA--	I--I--T-EI-DI--	I--K--T--EKO--YD--TK--TEFKD--YN--
F2.CM.97.CM53657	IS--T--FR--A--T--I--RSM--	I--I--V--YA--T-VL-DI--	I--K--Y--T--NGTL--K--EGV--KEVOSH--LN--													

	CD4	CD4	glycosylation NIT	glycosylation NES	fusion peptide	gp120 end	gp41 start							
B.FR.83.HXB2	RIKQII	NMWQ	KVGKAMYAPP	ISGQIRCSNITGLLLTRDGG	NSNNESEIFRPPGGDMRDNRSLEYKYKVKVIEPLGVAPTKAKRRVV	OREKRA	VGIGALFLGFLGAAGSTMGAAAMLTVQ	AR	QLL	SGTVQQQNNLLR	AIEAQOH	564		
A1.GE.99.99GEMZ011	R	Q	K	N	E	I	Q	I	E	I	L	A	S	560
A1.GE.99.99KER2008	RT	Q	I	K	V	E	N	STVN	T	R	R	R	Q	562
A1.KE.00.KNH1144	RT	Q	I	V	E	E	N	STN	T	R	R	R	Q	562
A1.KE.00.KSM4024	R	Q	P	V	S	V	I	V	I	R	R	R	S	569
A1.KE.00.MSA4069	R	Q	P	I	D	D	T	T	T	R	R	R	S	564
A1.KE.00.NKU3005	R	Q	Q	V	E	S	T	E	T	R	R	R	S	560
A1.RU.00.RU00051	R	Q	K	I	E	NK	GS	GT	I	R	R	R	S	566
A1.RU.03.03RU20_06_13	R	Q	K	S	E	X	GGT	XXX	I	N	R	R	S	554
A1.RW.93.93RW_024	R	Q	Q	V	E	D	SSPD	T	I	Q	R	R	S	562
A1.SE.95.SE8891	RA	Q	I	Q	V	E	D	T	I	Q	R	R	S	563
A1.SE.95.UGSE8131	RT	Q	T	P	V	Q	R	VTN	TNN	T	L	Q	K	572
A1.TZ.01.A173	RA	Q	I	Q	V	K	V	I	E	RN	SLN	T	R	566
A1.UA.01.01UADN139	R	Q	R	N	S	V	T	SST	T	E	I	R	S	562
A1.UG.99.99UGA07072	R	Q	Q	V	K	E	I	VNS5	T	R	R	R	S	568
A1.UZ.02.02UZ0659	R	Q	K	S	K	GGT	SSN	T	I	N	K	I	S	560
A2.CD.97.97CDKS10	V	D	E	A	V	Y	T	II	S	N	SR	R	S	562
A2.CD.97.97CDKTB48	R	R	A	V	K	T	MI	K	STN	T	I	E	R	560
A2.CY.94.94CY017_41	R	R	A	T	K	T	II	N	GTN	T	L	R	S	566
B.AR.04.04AR151516	R	E	R	R	SG	ESKON	T	E	VTG	M	L	X	560	
B.AU.87.MBC925	R	E	I	R	K	PD	KDT	I	VTG	M	L	567		
B.BO.99.BOI0122	R	E	R	R	K	DN	ST	T	N	R	R	R	S	568
B.BR.03.BREP2012	V	R	R	N	T	Q	INQTN	T	A	R	R	R	S	568
B.CA.97.CANB3FULL	R	E	G	N	I	N	STN	T	K	QV	R	R	S	558
B.CN.05.05CNCB_hp3	V	R	E	I	A	P	N	K	INKTDT	T	R	R	S	567
B.CO.01.PCM001	R	R	R	R	V	V	NDT	NN	T	N	R	R	S	558
B.GB.83.CAM1	R	Q	R	T	S	RGE	T	VTG	M	L	564			
B.GB.86.GB8	V	R	E	R	A	H	RE	NT	R	R	R	R	S	557
B.GE.03.03GEMZ010	R	E	R	R	L	A	SADO	N	NIK	L	R	R	S	555
B.IT.05.SG1	V	E	R	W	N	D	D	N	DT	T	R	R	S	524
B.JP.05.DR6538	R	V	T	Y	K	T	KGN	TDT	V	I	R	R	S	572
B.KR.05.05CSR3	V	R	E	R	K	S	K	DN	TTGSMETKG	A	R	R	S	577
B.LL.00.67.L00T36	R	E	R	R	E	R	RE	NT	R	R	R	R	S	563
B.RU.04.04RU128005	R	E	R	R	K	K	SNSTNN	T	A	I	M	R	S	558
B.TH.00.00TH_C3198	V	R	E	R	A	N	S	I	W	N	KNGNKT	DT	T	560
B.UA.01.01UAKV167	R	E	R	E	I	L	NSSNN	T	N	K	S	PTL	R	573
B.US.04.ES10_53	R	E	R	P	YLN	I	V	HNT	TT	T	V	R	S	559
B.US.99.PRB959_03	R	E	R	R	YLN	I	V	HNT	TT	T	V	R	S	554
C.AR.01.ARG4006	R	E	R	R	M	T	R	I	T	GT	NT	X	N	564
C.BR.04.04BR013	K	V	R	G	Q	V	K	M	S	T	NT	T	A	573
C.BW.00.00BW07621	R	E	R	A	N	I	K	EGE	TT	T	A	R	S	551
C.CN.98.YNRL9840	R	E	R	K	K	T	K	V	TENN	TET	NT	T	A	569
C.ET.02.02ET_288	R	E	R	K	I	T	R	M	NT	NT	T	A	R	558
C.GE.03.03GEMZ033	R	E	R	A	Y	T	N	NG	S	CT	T	I	M	566
C.IL.99.99ET7	R	E	R	E	N	T	K	TD	ETI	T	I	R	S	566
C.IN.99.01INS565_10	R	E	R	A	N	T	K	I	V	KTQD	DT	A	R	562
C.KE.00.KER2010	R	E	R	A	N	T	K	I	Q	KT	KT	DT	N	563
C.MM.99.mIDU101_3	R	E	R	A	N	T	K	V	GTQ	T	N	K	I	566
C.MW.93.93MW_965	R	E	R	A	N	T	K	V	GTQ	T	N	K	I	566
C.SN.90.90SE_364	R	E	R	A	N	T	K	V	GTQ	T	N	K	I	553
C.SD.89.89SD145	R	E	R	A	N	T	K	V	GTQ	T	N	K	I	549
C.TZ.02.C0178	R	F	T	R	A	N	T	N	S	W	NDPKRT	ET	N	562
C.UY.01.TRA3011	V	R	E	R	E	N	T	R	S	W	SSKETS	TT	T	570
C.YE.02.02YE511	V	R	E	R	A	N	T	R	S	W	SSKETS	TT	T	559
C.ZA.04.04ZASK164B1	R	E	R	A	N	T	K	V	GTQ	T	N	K	I	563
C.ZA.05.05ZAPSK240B1	R	E	R	A	N	T	K	V	GTQ	T	N	K	I	572
C.ZM.02.02ZM108	K	V	R	E	R	A	N	T	K	V	GTQ	T	N	548
D.CD.83.ELI	K	XW	A	X	I	ERN	L	IN	STN	T	R	R	S	580
D.CM.01.01CM_0009BBY	E	R	A	K	T	NN	SGN	G	ST	N	K	I	R	562
D.KE.01.NKU3006	R	E	R	R	A	N	L	V	I	STN	T	R	R	577
D.KR.04.04KBH8	R	E	R	E	V	K	I	A	N	T	N	K	QV	550
D.TD.99.MN011	R	E	R	A	K	K	K	ASN	T	R	R	R	S	553
D.TZ.01.A280	R	E	R	A	L	T	NVN	SRE	T	I	R	R	S	560
D.UG.99.99UGK09259	R	E	R	E	L	T	A	N	SON	T	I	R	R	563
D.YE.01.01YE386	G	R	E	L	ENN	T	ENN	T	I	R	R	R	S	550
D.YE.02.02YE516	G	R	E	L	ENN	T	ENN	T	I	R	R	R	S	554
D.ZA.90.R1	R	G	R	E	L	D	SSH	N	T	I	R	R	S	566
F1.AR.02.ARE933	V	E	R	NNH	HA	N	T	G	GTNCTDGG	CT	T	N	K	553
F1.BE.93.V1850	V	G	R	TS	A	N	T	R	EN	NI	N	K	E	541
F1.BR.01.01BR125	K	V	R	E	R	A	E	T	R	RNI	TDOP	N	K	559
F1.EG.0.P146	K	V	R	E	R	A	E	T	R	Q	GR	N	K	569
F1.FL.93.FIN9363	FV	E	R	R	A	N	T	N	I	QSN	SD	T	N	550
F2.CM.02.02CM_0016BBY	V	V	R	R	A	K	Q	N	SKA	NTD	L	I	E	553
F2.CM.95.MP257	V	R	R	Q	A	K	Q	N	EG	TL	Q	K	R	558
F2.CM.97.CM53657	R	V	R	R	A	A	Q	N	S	I	KN	TN	DTL	557
G.BE.96.DRCBL	K	V	R	R	Q	A	N	T	R	I	DN	ST	E	555
G.CM.01.01CM_4049HAN	K	V	R	R	Q	A	N	T	R	I	NTRD	T	E	557
G.CU.0.C174	K	V	R	R	Q	A	N	T	R	I	NTRD	T	E	566
G.ES.00.X558	K	V	R	R	Q	A	N	T	R	I	TX	XT	GT	558
G.ES.99.X138	V	R	R	Q	A	N	T	R	I	X	TDI	GT	A	559
G.GH.03.03GH175G	K	V	R	R	Q	E	E	D	T	I	NS	NT	NT	563
G.KE.93.HH8793_12_1	GX	V	RL	R	Q	S	ARN	T	K	NA	NAS	T	A	551
G.NG.01.01NGPLD669	V	R	R	Q	A	D	T	R	V	EN	I	STN	T	557
G.PX.P12693	V	R	R	Q	E	N	M	EN	GT	T	R	R	R	584
G.SE.93.SE6165	K	V	R	R	Q	A	N	E	N	S	NN	NTNTS	T	567
H.BE.93.V1991	K	V	R	R	Q	R	N	T	I	I	F	R	N	570
H.BE.93.V1997	V	R	R	Q	K	S	T	V	I	Y	ED	KG	DNVT	563
H.CF.90.056	V	R	R	Q	N	M	V	I	E	E	ASA	NYT	I	556
J.CD.97.J_97DC_KTB147	K	R	V	R	R	Q	I	A	N	T	T	NT	KES	564
J.SE.93.SE7887	K	V	R	R	Q	A	N	T	R	I	NRNG	E	GT	559
J.SE.94.SE7022	K	V	R	R	Q	A	N	T	R	I	NRNG	E	GT	561
K.CD.97.EQTB11C	K	D	R	Q	I	TA	N	T	R	MI	ND	NTRTE	T	559
K.CM.96.MP535	K	D	R	Q	I	A	S	N	MI	ND	N	THN	T	550

	CD4	CD4	glycosylation NIT	glycosylation NES V5	fusion peptide gp120 end gp41 start											
B.FR.83.HXB2	RIKQII	.MMWQ	.KVGKAMYAPP	.ISGQIRCSNITGLLLTRDGG	.NSNNESEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKARRVV	.OREKRA.VGIGALFLGFLGAAGSTMGAAASMTLVQ.AR.QLL.SGTVQ00NNLLR.AIEAQOQH	564									
01_AE.CF.90.90CF11697	---VV	---E	---EAVN	---II	---.ATN	---.N-K	---.Q	---.I	---.R	---.R	---.E	---.MIF	---.L	---.V	---.S	564
01_AE.CN.05.FJ051	---E	---Q	---	---	---.GNTGN	---.NWK	---.N	---.Q	---.I	---	---	---	---	---	---	566
01_AE.CN.06.FJ054	*S.L	*TCGREWP	*QCLMLMYA	*P	*T	*TSGTN	*N-K	*Q	*L	*L	*R	*MIF	*I	*R	*Q	557
01_AE.HK.x.HK001	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	560
01_AE.JP93.93JP.NH1	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	558
01_AE.TH.01.01TH.R2184	K	---	---	---	---	---	---	---	---	---	---	---	---	---	---	564
01_AE.TH.02.OUR7691	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	558
01_AE.TH.90.CM240	KL	---	---	---	---	---	---	---	---	---	---	---	---	---	---	558
01_AE.US.00.00US_MSC1164	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	560
02_AG.CM.02.CM_4082STN	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	554
02_AG.EC.x.EC141	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	550
02_AG.FR.91.DJ264	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	556
02_AG.GH.03.GHNJ196	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	567
02_AG.NG.01.PL0710	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	549
02_AG.NG.x.IBNG	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	555
02_AG.SE.94.SE7812	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	567
02_AG.SN.98.MP1211	K	---	---	---	---	---	---	---	---	---	---	---	---	---	---	561
02_AG.UZ.02.02UZ710	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	561
03_AB.RU.97.KAL153_2	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	548
04_cpx.CY.94.CY032	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	562
05_DF.BE.x.VII310	K	---	---	---	---	---	---	---	---	---	---	---	---	---	---	569
06_cpx.AU.96.BFP90	K	---	---	---	---	---	---	---	---	---	---	---	---	---	---	567
06_cpx.RU.05.04RU001	K	---	---	---	---	---	---	---	---	---	---	---	---	---	---	549
07_BC.CN.97.CN54	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	560
08_BC.CN.97.97CNGX_6F	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	565
09_cpx.GH.96.96GH29PI	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	550
10_CD.TZ.96.96TZ_BF061	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	548
11_cpx.GR.x.GR17	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	550
12_BF.AR.99.ARMA159	K	---	---	---	---	---	---	---	---	---	---	---	---	---	---	550
13_cpx.CM.96.1849	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	568
14_BG.DE.01.9196_01	K	---	---	---	---	---	---	---	---	---	---	---	---	---	---	557
14_BG.ES.99.X397	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	566
15_01B.TH.99.99TH_MU2079	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	555
16_A2D.KR.97.97KR004	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	563
18_cpx.CU.99.CU76	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	567
19_cpx.CU.99.CU7	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	512
20_BG.CU.03.CB471	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	563
21_A2D.KE.91.KNH1254	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	570
23_BG.CU.03.CB118	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	556
24_BG.CU.03.CB378	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	565
25_cpx.CM.01.101BA	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	575
27_cpx.CD.97.97CDKTB49	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	560
28_BF.BR.99.BREPM12609	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	576
29_BF.BR.02.BREPM119	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	557
31_BC.BR.02.110PA	G	---	---	---	---	---	---	---	---	---	---	---	---	---	---	576
33_01B.MY.05.05MY_KL007_1	K	---	---	---	---	---	---	---	---	---	---	---	---	---	---	562
34_01B.TH.99.OUR2478P	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	551
35_AD.AF.05.05AF095	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	552
36_cpx.CM.00.00CMNYU830	K	---	---	---	---	---	---	---	---	---	---	---	---	---	---	557
37_cpx.CM.00.00CMNYU926	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	561
42_BFLU.03.luBF_05_03	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	539
N.CM.02.DJO0131	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	540
N.CM.04.04CM_1015_04	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	533
N.CM.04.04CM_1131_03	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	555
N.CM.95.YBF30	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	552
N.CM.97.YBF106	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	538
O.BE.87.ANT70	KLR	VV	RS	I	RGOSGL	---	---	---	---	---	---	---	---	---	---	562
O.CM.91.MVP5180	KLR	LV	RS	M	GESRI	---	---	---	---	---	---	---	---	---	---	572
O.CM.96.96CMABB637	K	---	---	---	---	---	---	---	---	---	---	---	---	---	---	578
O.CM.98.98CMA104	KLR	VV	RS	I	OG_SGL	---	---	---	---	---	---	---	---	---	---	580
O.CM.99.99CMU4122	K	---	---	---	---	---	---	---	---	---	---	---	---	---	---	571
O.SN.99.SEMPI299	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	575
O.US.99.99USTWLA	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	574
O.FR.92.VAU	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	573
CPZ.CD.90.ANT	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	557
CPZ.CM.01.SIVcpzCAM13	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	557
CPZ.CM.05.SIVcpzEK505	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	535
CPZ.CM.05.SIVcpzLB7	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	546
CPZ.CM.05.SIVcpzMB66	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	581
CPZ.CM.05.SIVcpzMT145	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	575
CPZ.GA.88.GAB1	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	554
CPZ.TZ.01.TAN1	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	557
CPZ.US.85.CPZUS	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	536

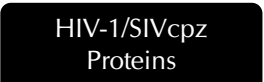
Table with columns for glycosylation sites (N, K, S, Y, E, D, Q, R, H, L, V, I, T, S, T, P, F, G, A, M, K, R, N, C, D, E, G, H, J, L, M, O, P, Q, U, V, W, X, Y, Z) and corresponding amino acid sequences for various HIV-1/SIVcpz proteins. Includes glycosylation regions: immunodominant region, glycosylation NAS, glycosylation NHT, glycosylation NKS, glycosylation NYT.

Table with columns for glycosylation (NGS, NAT), gp41 end, and Env end. Rows list protein sequences for various HIV-1 and SIVcpz strains, including B.FR.83.HXB2, A1.GE.99.99GEMZ011, and many others.

856
860
855
869
866
834
860
866
864
862
864
853
861
866
862
858
860
862
851
853
860
866
861
851
860
857
850
848
815
866
870
866
851
853
866
847
850
851
866
860
869
858
866
854
867
863
866
853
847
862
863
859
863
865
849
880
855
866
870
843
846
853
855
843
846
851
850
855
846
852
862
843
846
851
850
855
857
866
859
863
856
849
850
852
854
852
852
852

	myristoylation Nef start		acidic cluster	phos poly-P helix	phos	HXB2 premature Nef end																					
B.FR.83.HXB2		MGGKWSKSSVIGWPTVRERMRR	AEPADRVAASRDLEKHAITSNTAATAAACAWEAEQEEEE	.VGFVPTQVQLPRPMTYKAANDLSHLFKEKGGLEGLIHS	.QRRQDLDLWIYHTQGYF.PD+QNYTPGP	.GVRYPLTFGWICYKL	144																			
A1.GE.99.99GEMZ011		IV-Q-0-L	APA-PG-PV-Q-D-Y-V	-D	-D	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	-G	Y-RK-E	V	-F	-WH	I-F	143															
A1.KE.00.KNH1144		IV-EI-I	TPA-PG-V-Q-D	VNHPS-V-Y-D	-R	-G	Y-RK-E	V	-F	-WH	I-F	144															
A1.KE.00.KSM4024		S-R-TME-E-C	AS-PG-V-Q-D-V	VNHPS-V	-K	-G	F	D-Y-RK-E	V-N-F	-W	I	142															
A1.KE.00.MSA4069		IV-E-E-E	VP-SG-V-Q-D-Y-V	VNHPS-V-N	-R	-G	F-RK-E	V-N	-W	-I	FE	143															
A1.KE.00.NKU3005		S-IV-EI-I	TPA-NG-V-Q-R-X	TNHP5-V-D	-R	-G	L-X	D-Y-RK-E	V-N	-W	T-F	143															
A1.RU.00.RU00051		IV-E-E-I	APA-RG-PV-Q-D-X-XV-N	-N	-D	-G	X	D-Y-RK-XE	X-VH-X	-W	IX	146															
A1.RU.03.03RU20_06_13		IV-Q-0-IG	APA-XG-PV-Q-D-V	N-D	-R	-G	X	D-Y-RK-E	V-N	-W	I-F	146															
A1.RW.93.93RW_024		IV-KI-I	AP-PS-TG-V-Q-AR-V	TNHP5-V	E	-R	-G	Y-RK-E	V-N	-F	E	145															
A1.SE.95.SE8891		IV-KI-I	TP-KG-V-Q-D-V	INHPS-V	-R	-G	-D	Y-RK-E	V-N	-W	T-F	143															
A1.SE.95.UGSE8131		N-G-E-I-Q	ARAPAHTP-PT-TG-V-Q	INHPS-T	AQ-DEE	-R	-G	L	-Y-RK-E	V	-WH	I	149														
A1.TZ.01.A173		IV-EI-I	TP-VRGS-V-Q-D	N	-D	-R	-G	F-Y-N	-D	-W	C	I	134														
A1.UA.01.01UADN139		IV-QI-I	APAPAAPA-TG-PV-Q-D-V	N-D	-R	-G	F	Y-RK-E	V	-W	I	151															
A1.UG.92.92UG037		N-IV-EI-I-Q	TPTAARERTRO-PT-KG-V-Q-AR-V	VNHPS-V	-R	-G	F-GF	D-Y-RK-E	V	-W	I	154															
A1.UG.99.99UGA07072		IV-EI-I	TPA-KG-V-Q-D-Y	INHPS-V	D	-R	I	F	D-Y-RK-E	V	-W	T-F	143														
A1.UZ.02.02UZ0659		IV-Q-0-I	APA-RG-PV-Q-D-V	N-D	-R	-G	I	D-Y-RK-E	V	-W	I-F	146															
A2.CD.97.97CDKS10		N-TIV-AI-I	TPPAAERVGATRQTP-EG-V-TT-V-IG	Y-PDS-V	E	-K	V-F-G-F	D-Y-K	V-N	-W	I	130															
A2.CD.97.97CDKTB48		RTIV-EI-I	TPPAAEGVRRTP-EG-V-Q-AR-V	N-PD	-R	-A	F-G-F	D-Y-K	V-N	-W	A-F	155															
A2.CY.94.94CY017_41		R-TP-AI-I	TPPTAORTAETS-PG-V-Q-AT-V	-V	-R	-F	G-F-F	D-Y-K	M-V	-W	I	156															
B.AR.04.04AR151516		YIRG-A-I-L	T-G-R	-D	G-E	-R	-VW-K	VX-F	-WH	T	-C	F	146														
B.AU.87.MBC925		-TF	G	E	TN-D	-K	-D	-Y-K	-W	-I	-I	F	145														
B.BO.99.99BO10122		K-AA-K	ARPAEAERRRQ-G	AER	-D	H-D	-R	I	R	V-N	V-F	-W	143														
B.BR.03.BREPM2012		S-HIGE-A-D	ARPAEAERRRQ-G	AER	-D	H-D	-R	F-Y	Y-OR	-V	-F	-W	155														
B.CA.97.CANB3FULL		L-R-MRE-A-L	G-V-R	-D	-K	-R	-VY	K-K	V	-F	-W	T-F	144														
B.CO.05.05CNHB_hp3		L-GMA-S	G-V-GR	TA-D	-R	-R	-G-L	R	Y-QK-E	V	-F	-WH	I	143													
B.CO.01.PCM001		R-KAAE-AD	EG-V	-D	X-H-D	-R	R	I-M	D-Y-KK-E	V	-E	-W	C	I	145												
B.GB.83.CAM1		R-LG-SR-D	AEPR-EG-V	-D	-R	-R	-L	I	Q	D-Y-PK-E	V	-W	T	F	149												
B.GB.86.GB8		L-RM-SR-D	AEPR-EG-V	-D	-R	-R	-L	I	Q	D-Y-PK-E	V	-W	T	F	145												
B.GE.03.03GEMZ010		L-RLTG-A-EQ	A-EG	-Y-L	-S	-E	-R	-G-L	IVW-Q	V-N	-F	-W	G	I	145												
B.IT.05.SG1		Q						X-I-F-X						26													
B.JP.05.DR6538		KLG-I	T-G-V	Y-L	N-D	-R	F-E-F	Y-K-E	V	-I	-WH	I	F	145													
B.KR.05.05CSR3		R-KGG-A-E	G-PV	RY-L	N-PDL	-R	-G-L	VY-K	VH	-FL	-W	G	I	I	142												
B.NL.00.671.KO1236		B.RU.04.04RU128005	-X-R-XVE-A-EK	V-G	P-T	-D	A-E	-R	-R	-K	VH	-W	I	F	143												
B.TH.00.00TH_C3198		C-V-OR-K	G-V	-R	-E	-R	-I	I	D-Y	-E	V-N	-W	X	I	133												
B.UA.01.01UAKV167		R-TG-A-E	EG-V-R	VNS	-D	-V	-G-E	-R	-L	I	-I	-D	-Y	-E	V-N	-W	I	143									
B.US.04.ES10_53		MSE-SI	AEPRMR-AG	-G	-N	VN	-T	A	G-E	-R	-G	-L	-F	-W	I	150											
B.US.99.PRB959_03		C-MG-TA-K	TEP-AG	-V	-N	VN	-T	A	G-E	-R	-G	-L	-F	-W	I	147											
C.AR.01.ARG4006		N-L-A-I	APPGVGDXRRQDA-EG	P-Q	DRY-L	P-N	-D	-Q	E	-R	F-G-F	F	-Y	-Q	K	V-N	-W	I	154								
C.BR.04.04BR1		N-C-PV-AI-I	TD-EG	-Q	DR-L	P-N	-D	-D	-D	-R	-G	-L	-GF	-W	-F	-W	T	S	PF	145							
C.BW.00.00BW07621		N-KVE-AI	T-EG	-Q	D-Y-L-X	-R	-R	-G	-L	-F	-W	-I	-F	-W	-I	-F	45										
C.CN.98.YNRL9840		C-PR-S-I	TP-EG	-Q	L-T	PEN	-Q	E	-R	-R	-V	F-GF	F	-D	-Y	-K	KE	V-N	-F	-W	146						
C.ET.02.02ET_288		N-IV-A-I	TD-EG	-Q	D-L	EN	-D	K	-Q	E	-R	-G	-F	-F	-D	-Y	-K	KE	V-N	-F	-W	146					
C.GE.03.03GEMZ033		L-IV-A-D-IS	TD-EG	-Q	D-L	V	-D	-Q	E	-R	-G	-F	-F	-D	-Y	-K	KE	V-N	-F	-W	146						
C.IL.99.99IT7		N-AI	AEPA-EG	-Q	D-Y-L	DTA	-PRE	-D	E	-R	-G	-F	-F	-D	-Y	-K	KE	VH	-FL	-W	146						
C.IN.99.01INS565_10		IV-AI	EG	-Q	D-Y-L	GN	-D	-Q	E	-R	R-G	I	F	-D	-Y	-K	KE	V-N	-F	-W	146						
C.KE.00.KER2010		R-IV-AI	T-EG	-Q	D-Y-L	DT5	-D	-R	-E	IR	-R	-G	-F	-F	-D	-Y	-K	KE	V-N	-F	-W	146					
C.MM.99.mIDU101_3		I-D	TS-EG	-Q	D-L	VSN	-D	-Q	E	-R	-G	-F	-F	-D	-Y	-K	KE	V-N	-F	-W	146						
C.MW.93.93MW_965		N-IV-EI-I	TN-TG	-Q	D-Y-L	E	-Q	-Q	E	-R	-G	-F	-F	-D	-Y	-K	KE	V-N	-F	-W	146						
C.SN.90.90SE_364		X-RIV-ANI	TPA-EG	-Q	D-Y-L	PTN	-D	-Q	E	-R	-G	-F	-F	-D	-Y	-K	KE	V-N	-F	-W	146						
C.SD.89.89SD145		R-IV-EI-I	TRPAEAGV-EG	-Q	D-Y-L	P	-DS	-Q	S-E	-R	-V	S	-F	-D	-Y	-K	KE	V-N	-F	-W	146						
C.TZ.02.CO178		N-C-PV-UI-I	TD-EG	-P	-Q	D-Y-L	P-N	-D	-Q	E	-R	F-G	-F	-F	-Y	-K	KE	V-N	-F	-W	146						
C.UY.01.TRA3011		IV-V-D-I	TNPAAGVGAASQDS-AG	-Q	A-L	PTN	-Q	D-E	-R	-R	-G	-F	-W	-Y	-K	KE	V-N	-F	-W	T	F	158					
C.ZA.04.04ZASK164B1		X-IV-AI-I	TPAAEGAG-EG	T	-D	L-T	TN	-D	-D	-R	-G	-F	-GF	-D	-Y	-K	KE	V-N	-F	-W	PF	152					
C.ZA.05.05ZAPSK240B1		K-R-EI-I-Q	TRP-A-G	-Q	D	F-T	SN	-DR	-Q	-D	-R	-V	G	-F	-GF	-D	-Y	-K	KE	V-N	-F	-W	PF	146			
C.ZM.02.02ZM108		N-IV-AI-I	TRTGAT-EG	-Q	D	-T	-D	-V	-R	-KGG	E	-V	-F	-W	-S	-I	-W	147									
D.CD.83.ELI		IV-AI-I	TN-G	-V	-K	-S	-D	-S	-D	-E	-R	-E	-L	-W	-KK	E	V-N	-I	-W	-E	146						
D.CM.01.01CM_0009BBY		I-AI-KI	TNPI-EG	-V	-K	-D	-D	-D	-D	-E	-R	-G	-L	-VY	-K	E	V-N	-I	-W	-X	134						
D.KE.01.NKU3006		I-AI-I	ID-EG	-V	-R	-V	-Q	-D	-D	-E	-R	-G	-L	-W	-KK	E	V-N	-I	-W	-I	FE	146					
D.KR.04.04KBBH8		IV-AI-I	T-G	-V	-R	-R	-D	-D	-E	-R	-T	-L	-D	-VW	-K	E	V-N	-F	-WH	-I	FE	141					
D.TD.99.MN011		IV-AI-I	TD-EG	-V	-R	-T	-D	-G	-E	-R	-A	-F	-L	-D	-VW	-K	E	V-N	-F	-WH	-I	FE	145				
D.TZ.01.A280		IV-AI-I	TD-EG	-V	-R	-H	-D	-G	-E	-R	-A	-F	-L	-D	-VW	-K	E	V-N	-F	-WH	-I	FE	146				
D.UG.99.99UGK09259		KMG-SA-A-I	EG	-V	-Y	T	-Q	-PD	-E	-R	-G	-L	-W	-K	E	V-N	-F	-WH	-I	-W	146						
D.YE.01.01YE386		IV-AI-I	T-EG	-V	-Q	SD	-DNA	-R	-R	-G	-L	-W	-PK	-E	-V	-W	-I	-W	-I	-F	FE	146					
D.YE.02.02YE516		X-IV-AI-I	TD-EG	-V	-NNPA	-D	-R	-D	-E	-R	-F	-E	-V	-W	-I	-W	-I	-W	-I	-F	FE	143					
D.ZA.90.R1		IV-AI-I	TD-G	-V	-S	-D	-S	-E	-K	-R	-F	-E	-W	-KK	-E	-V	-N	-I	-W	-I	-F	146					
F1.AR.02.ARE933		IV-EI	TP-EG	-V	-Q	DRR	KNN	-PDL	-R	-R	-F	-G	-X	-R	-V	-G	-F	-X	-R	-V	-G	-T	-V	-W	-I	-F	91
F1.BE.93.V1850		IV-A-G	TPT-EG	-V	-Q	DRR	RT	-PDL	-R	-V	-G	-F	-X	-R	-V	-G	-F	-X	-R	-V	-G	-T	-V	-W	-I	-F	145
F1.BR.01.01BR125		N-R-IV-AI-I-Q	TPT-EG	-V	-Q	DRR	RDN	-PDL	-R	-V	-G	-F	-X	-R	-V	-G	-F	-X	-R	-V	-G	-T	-V	-W	-I	-F	145
F1.E8.S.P146		IV-AI-I	TPPTAERVERRRQ-EG	-V	-Q	RR	G	-PDL	-D	-D	-R	-V	-G	-F	-Q	XX	-D	-Y	-K	KE	V-N	-F	-W	X	-I	-W	148
F1.FL.93.FIN9363		IV-AI-I	PPPA-EG	-V	-Q	RR	G	-PDL	-D	-D	-R	-V	-G	-F	-Q	XX	-D	-Y	-K	KE	V-N	-F	-W	X	-I	-W	148
F2.CM.02.02CM_0016BBY		IV-K-Q	TPV-EG	-Q	D-R	N	-G	-DL	-R	-R	-F	-G	-F	-Q	XX	-D	-Y	-K	KE	V-N	-F	-W	X	-I	-W	132	
F2.CM.95.MP257		IV-AI-I	EG	-V	-Q	D-R	N	-G	-DL	-R	-R	-L	-F	-W	-K	E	V-N	-F	-WH	-I	-W	-P	-F	-W	140		
F2.CM.97.CM53657		KV-I-I-Q	TPA-PG	-V	-Q	DR	Q	-RD	-SDL	-D	-R	-G	-F	-F	-D	-Y	-K	KE	V-N	-F	-W	T	-F	-W	145		
G.BE.96.DRCBL		X-N-RK-A-E	HPC-EG	-V	-Q	DR	R	-G	-PD	-DSE	-R	-F	-F	-D	-Y	-K	KE	V-N	-F	-W	-T	-V	-W	146			
G.CM.01.01CM_4049HAN		IV-KI-I	TPIAAE-AG	-V	-Q	AR	N	-PDS	-DSE	-R	-F	-F	-D	-Y	-K	KE	V-N	-F	-WH	-I	-W	-C	-X	-I	-W	121	
G.CU.C674		IV-AI-I	TP-EG	-V	-Q	AR	N	-PD	-Q	-DSE	-R	-G	-F	-F													

Table with columns for protein names (e.g., B.FR.83.HXB2, 01_AE.CF.90.90CF11697), amino acid sequences, and accession numbers (143-149). Headers include 'myristoylation Nef start', 'acidic cluster', 'poly-P helix', 'phos', and 'HXB2 premature Nef end'.



normal Nef end

B.FR.83.HXB2
A1.GE.99.99GEMZ011
A1.KE.00.KER2008
A1.KE.00.KNH1144
A1.KE.00.KSM4024
A1.KE.00.MSA4069
A1.KE.00.NKU3005
A1.RU.00.RU00051
A1.RU.03.03RU20_06_13
A1.RW.93.93RW_024
A1.SE.95.SE8891
A1.SE.95.UGSE8131
A1.TZ.01.A173
A1.UA.01.01UADN139
A1.UG.92.92UG037
A1.UG.99.99UGA07072
A1.UZ.02.02UZ0659
A2.CD.97.97CDKS10
A2.CD.97.97CDKTB48
A2.CY.94.94CY017_41
B.AR.04.04AR151516
B.AU.87.MBC925
B.BO.99.BOL0122
B.BR.03.BREPM2012
B.CA.97.CANB3FULL
B.CN.05.05CNHB_hp3
B.CO.01.PCM001
B.GB.83.CAM1
B.GB.86.GB8
B.GE.03.03GEMZ010
B.IT.05.SG1
B.JP.05.DR6538
B.KR.05.05CSR3
B.NL.00.671_00T36
B.RU.04.04RU128005
B.TH.00.00TH_C3198
B.UA.01.01UAKV167
B.US.04.ES10_53
B.US.99.PRB959_03
C.AR.01.ARG4006
C.BR.04.04BR013
C.BW.00.00BW07621
C.CN.98.YNRL9840
C.ET.02.02ET_288
C.GE.03.03GEMZ033
C.IL.99.99ET7
C.IN.99.01INS565_10
C.KE.00.KER2010
C.MM.99.mIDU101_3
C.MW.93.93MW_965
C.SN.90.90SE_364
C.SO.89.89SM_145
C.TZ.02.C0178
C.UY.01.TRA3011
C.YE.02.02YE511
C.ZA.04.04ZASK164B1
C.ZA.05.05ZAPSK240B1
C.ZM.02.02ZM108
D.CD.83.ELI
D.CM.01.01CM_0009BBY
D.KE.01.NKU3006
D.KR.04.04KBH8
D.TD.99.MN011
D.TZ.01.A280
D.UG.99.99UGK09259
D.YE.01.01YE386
D.YE.02.02YE516
D.ZA.90.R1
F1.AR.02.ARE933
F1.BE.93.V1850
F1.BR.01.01BR125
F1.ES.x.P1146
F1.FL.93.FIN9363
F2.CM.02.02CM_0016BBY
F2.CM.95.MP257
F2.CM.97.CM53657
G.BE.96.DRCBL
G.CM.01.01CM_4049HAN
G.CU.x.Cu74
G.ES.00.X558
G.ES.99.X138
G.GH.03.03GH175G
G.KE.93.HH8793_12_1
G.NG.01.01NGPLD669
G.PX.P12695
G.SE.93.SE6165
H.BE.93.V1991
H.BE.93.V1997
H.CF.90.056
J.CD.97.J_97DC_KTB147
J.SE.93.SE7887
J.SE.94.SE7022
K.CD.97.EQTB11C
K.CM.96.MP535

VPVEPKIEE . ANKGEN . TSLHHPVSLHGMDPEREVLWRFDLSRLAFHHVARELHPPEYFKNC*

--D-SEV--TE--N--ICQ--E--M-K--LT-R--Y-D
--D-EV-R--TE--N--ICQ--E--K-Q--T-Q--I--Y-D
--D-EV-K--TE--N--ICQ--E--G--R-K--H--L--KK--FY-D
--D-EV-K--TE--N--ICQ--E--K-Q--K--H--RK-R-L--FY-D
--D-EV-K--TE--N--ICQ--E--E-K--K-K--LR-R--M--Y-D
--A-XEV-K--E--N--MCO--ED-T-M-K--LK-R-Q--FY-D
--D-XEV--TE--GENNC--ICQ--D-G--M-K--LK-R-Q--P-FY-D
--D-EEV--TE--N--X--ICQ--E--K--M-K--X--LX-R--FY-D
--D-EV-K--TE--N--ICQ--E--K--V--LK-R--FY-D
--D-EV-K--TE--N--ICQ--E--T-M-K--LT-R--FY-D
--D-EV-K--TE--N--MCO--E--T-M-K--PH--K-R-F--Y-*
--D-EEV-K--TE--R--I-Q--D--T-R-T--K--LK-I--DFY-D
--D-AEV--TE--N--ICQ--V--E-K--M-K--LT-I--K--FY-D
--DE-EV--TG--N--ICQ--E--K--T-R-K--S--RV-K--FY-D
--D-EV-K--TE--N--ICQ--E--K--T-R-K--L-R-L--FY-D
--D-EEV--TE--N--ICQ--E--K--M-K--LT-R--FY-D
--D-SEV--TE--N--ICQ--AE--K-K--LR-L-Q--FY-D
--SEV--TO--N--ICQ--V--R-E-RS--RR-R--Y-D
--DKEQV-K--E--NC--G--E-S-K--I-K--M--Y-D
--E--E--N--M--E--K--V-K--M--FY-D
--LD--QV-R--M-Q--K--V-K--M--FY-D*
--L-EEV-K--E--GENN--I--E-A-G-V--LR-L-K--FY-D*
--E--E--NV--M--E--G--V-K--M-K--Y-D
--D-EQV--ENNC--M-Q--T--M-K--M-K--FY-D
--EEV-K--E--N--M--K--Q-K--M--FY-D
--D-EV-K--E--R--A--M--M--K--M--FY-D
--EEV-K--E--NC--M-Q--IE--K--V-K-N--M--FY-D
--DQEQV-K--E--NC--I-Q--V--K--M-K--K--K--Y-*
--E-QV--NC--M--E--V-K--S--L--R--Q--FY-D
--X--HY--
--KE-V--VTER--I-Q--IE--M-K--M--D--Y-D
--D-QV-K--E--N--M-Q--I--K--V-K--M--T--Y-D
--DQE-V--E--N--M--K--M-K--I--K--Y-*
--D-QV-K--E--NC--M-Q--K--M-K--M--I--Y-D
--TE-V--E--GENN--I--E--K--Q-K--R--M--Y-D
--V--E--NI--I--E--K--Q-K--H--R-M--Y-D
--D-REX--XT--N--M--E--SDG--Q-K--L--RR-M--Y-D
--D-KEV--E--N--M-Q--RI--DH--M-K--S--RR--Y-D
--D-KEV--E--NC--M-Q--RI--DH--M-K--S--RR--Y-D
--D-SEV--E--NC--I--EDK--M-K--IS--HR-L--Y-D*
--D-GEV--D--NC--M-Q--V--DK--M-K--S--RR-L--FY-D
--D-KEV--E--NC--M--IE--ED--R-K--H--RR-M--DFY-D
--D-REV--E--NC--MCO--IE--GD--M-K--H--RK-M--FY-D
--D-REV--E--NC--M-Q--IE--E--R-Q--L--HR-L-Q--Y-D
--D-RDV--T--D--NC--CO--E--EHG--Q-K--Q--LR-R--FY-D
--D-REV--E--NC--M-Q--IE--ED--I--R-K--Q--RR-L-Q--FY-D
--D-REV--E--NC--M--E--EX--M-K--TO--HR-M--FY-D
--D-REV--D--N--M-Q--E--ED--R-K--L--RR-M--Y-D
--D-SEV--E--NC--M--E--ED--K-K--L--HR-M--Y-D
--D-REV--E--N--M-M--E--HK--Q-K--L--HR-M--Y-D
--D-REV--E--N--I--E--GDG--M-K--HR--FY-D
--D-REV--SS--T--NC--INQ--ED--V-K--HI-M--WY-D
--D-REV--E--D--NF--M--E--E-G--K-K--S--RR-L--I--Y-D
--D-GEV--SE--D--NC--ANQ--EH--K-K--L--RR-M--Y-D
--D-QEV--DTE--T--N--ICQ--E--Q--K--N--E--K--M--FY-*
--I--K--V--NTE--NC--MCO--E--M--N--E--K--TM--Y-D
--SEEV--E--NC--M-Q--IE--Q--K--N--K--E--K--M--FY-D
--D-KV--ETE--NC--MNO--G--K--V--N--E--K--K--F--D
--D-KEV--TTE--SC--I-Q--E--T--K--V--N--E--K--S--Y-D
--E-V--TE--NC--INQ--E--K--I--E--K--Q--YMKDR
--D-KEV--DTE--SC--MCO--E--Q--K--M--N--K--E--K--KM--FY-D
--D-KMV--E--E--D--KC--MHO--G--K--V--N--K--E--K--KT--F--D
--D-QEV-Q--TE--GETNC--MNO--E--K--K--N--K--E--K--Y-D

205
208
205
206
204
205
205
208
208
205
210
196
213
216
205
208
130
217
218
208
207
203
215
206
205
207
211
207
206
36
207
204
205
205
133
205
212
209
216
207
207
45
206
208
208
208
208
208
208
208
208
208
212
208
209
206
134
208
203
207
208
208
132
200
205
208
121
208
207
208
208
210
210
208
208
208
207
202
204
205
207
207
202
204
205
207
207
202
208

B.FR.83.HXB2
01_AE.CF.90.90CF11697
01_AE.CN.05.FJ051
01_AE.CN.06.FJ054
01_AE.HK.x.HK001
01_AE.JP.93.93JP.NH1
01_AE.TH.01.01TH.R2184
01_AE.TH.02.OUR7691
01_AE.TH.90.CM240
01_AE.US.00.00US_MSC1164
02_AG.CM.02.02CM_4082STN
02_AG.EC.x.EC041
02_AG.FR.91.DJ264
02_AG.GH.03.GHNJ196
02_AG.NG.01.PL0710
02_AG.NG.x.IBNG
02_AG.SE.94.SE7812
02_AG.SN.98.MP1211
02_AG.UZ.02.02UZ710
03_AB.RU.97.KAL153_2
04_cpx.CY.94.CY032
05_DF.BE.x.VII310
06_cpx.AU.96.BFP90
06_cpx.RU.05.04RU001
07_BC.CN.97.CN54
08_BC.CN.97.97CNGX_6F
09_cpx.GH.96.96GH2911
10_CD.TZ.96.96TZ_BF061
11_cpx.GR.x.GR17
12_BF.AR.99.ARMA159
13_cpx.CM.96.1849
14_BG.DE.01.9196_01
14_BG.ES.99.X397
15_01B.TH.99.99TH_MU2079
16_A2D.KR.97.97KR004
18_cpx.CU.99.CU76
19_cpx.CU.99.CU7
20_BG.CU.03.CB471
21_A2D.KE.91.KNH1254
23_BG.CU.03.CB118
24_BG.CU.03.CB378
25_cpx.CM.01.101BA
27_cpx.CD.97.97CDKTB49
28_BF.BR.99.BREPM12609
29_BF.BR.02.BREPM119
31_BC.BR.02.110PA
33_01B.MY.05.05MYKL007_1
34_01B.TH.99.09UR2478P
35_AD.AF.05.05AF095
36_cpx.CM.00.00CMNYU830
37_cpx.CM.00.00CMNYU926
42_BFLU.03.luBF_05_03
N.CM.02.DJO0131
N.CM.04.04CM_1015_04
N.CM.04.04CM_1131_03
N.CM.95.YBF30
N.CM.97.YBF106
O.BE.87.ANT70
O.CM.91.MVP5180
O.CM.96.96CMAABB637
O.CM.98.98CMA104
O.CM.99.99CMAU4122
O.SN.99.SEMP1299
O.US.99.99USTWLA
O.FR.92.VAU
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzLB7
CPZ.CM.05.SIVcpzMB66
CPZ.CM.05.SIVcpzMT145
CPZ.GA.88.GAB1
CPZ.TZ.01.TAN1
CPZ.US.85.CPZUS

normal Nef end

VPVEPKIEE.ANKGEN.TSLLHPVSLHGMDPEREVLWRFDSRLAFHHVARELHPPEYFKNC*
--D-KEV--D-D-S-NC---M-Q-V-E---M-K-S-RR-I---R---Y-D-
I-D-KEA--DT----SC---M-Q---G-EHG---M-K-S-RR-I---KY-QFY-D-
--D-KDV--NT----C---I-Q---I-E-E-G---M-K-A-RR-I---M---Y-D-
--D-REV--D-A-Q---I-E-E-G---M-K-A-RR-L---R---FY-D-
--D-KEV--D-R--NC---M-Q---E---M-K-A-RR-I---M---FY-D-
.....
--DQREV--D-----NC---M-Q---IE-E---M-K-A-RR---Q---Y-D-
.....
E-MD-AEV---N---ICO---E-ED---V---S-RR-I---R---FY-D-
--ID-AEV---E---NV---ICO---E-ED---V-K---T-T---FY-D-
--MD-AEV---TE---N---ICO---E-DH-Q-V-K-Q---R-H---Y-D-
--MD-AEV---E---N---ICO---E-DD---I---R-T---Y-D-
--MD-AD--K.DTE---N---ICO---E-ED---V---T-K---M---FY-D-
--MD-AEV---Q---N---ICO---E-EDK---V---R---Y-D-
--MD-AE-K--E---NA---ICO---LE-ED---V---K-K---FY-D-
--D-AEV---TE---N---ICO---E-K---M-K---LT-R---FY-D-
--D-QEV---TE---C---I-Q---E-E---K-K---YK---FY-D-
--N-EEV-K--E---D-NC---M---E-DD---Q-K-S-RR-L-I---R---FYDQ*
--D-EEV-LT---NC---ICO---AE-E---K-K-S-RR-I---K---FY-D-
--D-REV---E---D-NC---CO---E-DH---K-K-Q-HR-R---FY-D-
--D-REV---E---D-NC---CO---E-EH---K-K-Q-HR-R---FY-D-
--D-KEV---TG---N---MCO-V-S.GG-M---T-LKXI-K---FY-D-
--D-REV---NC---M---E-KHG---V-K---T-HK-I---Y-D-
--D-REV---E---NC---M-Q---IE-ED---R-K-S-RR-I---DFY-D*
--D-VEV-K--E---NC---M-Q---E-ED---I-K---LR-L---K---WY-D*
--ID-KEV---E---NC---M-Q---IE-A---M-K-S-RR-I---FY-D-
--D-AE---T---N---ICO---E-XDN---R-T---S-RR---DFY-D-
E-MD-AE---T---N---ICO---E-ADN---I---S-RR-I---NFY-D-
--D-SDV---N---NC---M-Q---E-E---K-Q-A-RR-I---Y-D-
--S-AEV---TE---T---N---ICO---IE---K-V-H-LV-K---FY-D-
--DSNEV---E---N---ICO---E-E---K-K-A-RR-I---Q---FY-D-
--D-KEA-K.DTE---NC---ACQ-----K-V---N---X-K-IK---FY-D-
--MD-AEV---N---ICO---E-A-G---V---S-RR-I---FY-D-
--D-SEV---ETQ---N---ICO---AE---R-E-Q-GLR---M---FY-D-
--D-ADV---SN---N---ICO---E-DDK---M---S-RR-I---R-DFY-D-
--MD-AEV---R---N---I-Q---E-D-G-V---S-RR-I---FY-D-
--MD-EE---ED---N---LCX---AH---V---S-RR-I---K---WY-D-
--D-QEV---S---ICO---E-GDG---K-V---LK-L---T---FY-D-
--D-EQV-A--E---N---M-Q---K---M-K---M---Y-D*
--D-E-V--TE---NC---N---E---K---K-K---L---Y-
--D-REV---E---N---M---E-EHG---Q-K-L-HK-I---Y-Y-
--D-KEV---E---NC---M-Q---I-ED---M-K-A-RR---Y-D-
--D-GEVK.....ENSC---M-Q---IE-E---M-K-A-RR-I---FY-D-
--D-VEV-K--E---GENN---MCO---E-G---M-K---LK-R-Q---FY-D-
.....
--X-QV-K--TE.GEN---M---E-T---V-K---K---H-M---Y-D-
--LSAEAV---E-D--NA---LCQ-V-GHK---Q-Q---S-RR-I---DFY-
--LSAEAV---E-D--NA---ICO-V-DHK---V---S-RR-I---DFY-
--LSAEAV---E-D--NA---ICO-V-EHK---V---S-RR-L---DFY-
--LSAEAV---E-D--NA---ICO-A-DHK---V---S-RR---FY-
--LSAEAV---E-D--NA---ICO-V-DHKQ---V---S-RR---K---DFY-
--SEEEA-RLG-TC-R.AM---ACA-FE-THK-I-M-K-RS-GNT---MIT---L-QKD-
--SAAEA-RLG-TN-D.A---ACN-AE-AHG-I-K-Q-RS-GLT-I-LQK---L-PK-
--SEAEA-RLG-TC-R.AM---ACN-S-AHG-IPK-Q-TS-RT---LIT---L-NKO-
--SEAEA-LG-H-R.AK---ACA-FG-AHG-I-K-Q-RS-GLT---KIT---L-PKD-
--SEAEA-LG-CVT.AK---ACN-SE-HG-I-K-Q-RS-GNI---KVT---L-LKO-
--SEAEA-LG-C.R.A---ACN-FE-NHGQI-K-Q-RS-GST---MVTN---L-NKD-
--TEDEA-OLG-D-G.AM---ACN-SE-HHG-M-K-Q-S-GLT---MVT---L-IKU-
--SEAEA-ALG-C.R.A---ACN-YE-QHK-I-K-Q-RS-NT---LIT---L-SKD-
--SP.PDD.....R.NI---ACT-DG-HK-I-R-E-AS-MRR-I---R---RD*
--LTTEOV-A--E-D-NC---ICO---E-ESK---I---LR-I---K---Y-D*
--LS-EEV---Q-D--NV---MCO---E-DK---V---S-RV-R---FYQ-
--LT-EEV---NL---ICO---E-EDK---I-KY-Q-LR-I---S-Y-
--L-EEV-R--E-D-NI---ICO-QE-EA---V-T---LK-R---Y-D-
--LP-EEV-K--D-NC---MCO---E-E---V-T---S-GLT---MVT---Y-RD-
--LTTEOV-Q--E-D-NC---ICO---E-EDK---V---LR-I---Q---Y-D-
--D-PEDD--K.....NI---ACS-TT-DG-T-I---S-RR-I---RY---*G-
--LTEEEV-Q-----T.NI---MCO---E-EHG---I-Q-TE-RR-R-K-----R-*

205
205
207
207
207
207
205
205
130
130
205
128
140
138
207
207
229
212
207
207
207
210
207
203
220
206
206
208
207
207
207
207
208
207
207
207
207
208
200
208
208
201
217
208
208
208
209
224
209
208
209
207
209
207
203
205
131
132
207
214
213
214
213
208
214
211
218
214
212
212
212
212
216
198
207
207
207
207
207
206
205
196
209

