

V

HIV-1/SIVcpz Proteins

Contents

V-1	Introduction	311
V-2	Annotated features	312
V-3	Sequences	314
V-4	Alignments	322
V-4.1	Gag	322
V-4.2	Pol	330
V-4.3	Vif	344
V-4.4	Vpr	348
V-4.5	Tat	350
V-4.6	Rev	352
V-4.7	Vpu	354
V-4.8	Env	356
V-4.9	Nef	370

V-1 Introduction

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

V-2 Annotated features

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

Feature	Protein	Location	Page
Gag start, p17 start	Gag	1	322
membrane binding	Gag	1-30	322
phosphorylation site	Gag	112	322
p17 end	Gag	132	324
p24 start	Gag	133	324
CyPA binding	Gag	205-241	324
major homology region	Gag	285-304	326
p24 end	Gag	364	326
p2 start	Gag	365	326
p2 end	Gag	377	326
p7 start	Gag	378	326
Zn motif	Gag	392-404	326
Zn motif	Gag	413-425	326
p7 end	Gag	432	326
p1 start	Gag	433	328
p1 end	Gag	448	328
p6 start	Gag	449	328
Vpr binding	Gag	455-460	328
Vpr binding	Gag	489-494	328
p6 end, Gag end	Gag	501	328
Pol start	Pol	1	330
Gag-Pol TF start	Pol	1	330
Gag-Pol TF end	Pol	56	330
protease start	Pol	57	330
protease end	Pol	155	332
p66, p51 RT start	Pol	156	332
M41L	Pol	196	332
D67N	Pol	222	332
K70R	Pol	225	332
D110 catalytic site	Pol	265	332
polymerase motif	Pol	337-342	334
T215Y	Pol	370	334
K219Q	Pol	374	334
p51 RT end	Pol	595	336
p15 RNase H start	Pol	596	336
p66 RT, p15 Rnase H end	Pol	715	338
p31 Integrase start	Pol	716	338
p31 Integrase end	Pol	1004	342
Pol end	Pol	1004	342
Vif start	Vif	1	344
Vif end	Vif	193	346
Vpr start	Vpr	1	348
oligomerization	Vpr	1-41	348
amphipathic α -helix	Vpr	17-33	348
H(S/N)RIG motifs	Vpr	71-83	348
frameshift in HXB2	Vpr	72	348
Vpr end in HXB2	Vpr	79	348
Vpr end	Vpr	98	348

Feature	Protein	Location	Page
Tat start	Tat	1	350
C-rich region	Tat	22-37	350
nuclear localization	Tat	49-57	350
exon 1 end	Tat	72	350
exon 2 start	Tat	73	350
Tat end	Tat	102	350
Rev start	Rev	1	352
exon 1 end	Rev	25	352
exon 2 start	Rev	26	352
NLS	Rev	34-49	352
Leu-rich effector domain	Rev	75-83	352
Rev end	Rev	117	352
Vpu start	Vpu	1	354
transmembrane domain	Vpu	1-27	354
cytoplasmic domain	Vpu	28-82	354
α -helix	Vpu	43-51	354
phos	Vpu	53	354
phos	Vpu	57	354
α -helix	Vpu	58-70	354
Vpu end	Vpu	83	354
Env start	Env	1	356
signal peptide end	Env	30	356
gp120 start	Env	31	356
glycosylation NVT	Env	88-90	356
CD4 binding	Env	124	356
V1	Env	131-156	356
glycosylation NDT	Env	136-138	356
glycosylation NSS	Env	141-143	356
glycosylation NCS	Env	156-158	358
V2	Env	158-196	358
glycosylation NIS	Env	160-162	358
glycosylation NDT	Env	186-188	358
glycosylation NTS	Env	197-199	358
CD4 binding	Env	196	358
glycosylation NKT	Env	230-232	358
glycosylation NGT	Env	234-236	358
glycosylation NVS	Env	241-243	358
glycosylation NGS	Env	262-264	360
glycosylation NFT	Env	276-278	360
CD4 binding	Env	279	360
glycosylation NTS	Env	289-291	360
glycosylation NCT	Env	295-297	360
V3	Env	296-331	360
glycosylation NNT	Env	301-303	360
V3 tip	Env	312-315	360
glycosylation NNT	Env	339-341	360
glycosylation NKT	Env	356-358	360
CD4 binding	Env	365	360
V4	Env	385-418	362
glycosylation NST	Env	386-388	362
glycosylation NST	Env	392-394	362
glycosylation NST	Env	397-399	362
glycosylation NNT	Env	406-408	362

Feature	Protein	Location	Page
CD4 binding	Env	425	362
glycosylation NIT	Env	448-450	362
CD4 binding	Env	455	362
V5	Env	460-471	362
glycosylation NES	Env	463-465	362
CD4 binding	Env	469	362
fusion peptide	Env	512-527	364
gp120 end	Env	511	364
gp41 start	Env	512	364
immunodominant region	Env	588-607	364
glycosylation NAS	Env	611-613	364
glycosylation NKS	Env	616-618	364
glycosylation NHT	Env	624-626	364
glycosylation NYT	Env	637-639	366
transmembrane domain	Env	685-704	366
gp41 cytoplasmic tail start	Env	705	366
glycosylation NGS	Env	750-752	366
glycosylation NAT	Env	816-818	368
cytoplasmic tail end	Env	857	368
gp41 end	Env	857	368
Env end	Env	857	368
Nef start	Nef	1	370
myristoylation	Nef	2-7	370
acidic cluster	Nef	62-65	370
poly-P helix	Nef	69-78	370
phosphorylation	Nef	77-81	370
HXB2 premature Nef end	Nef	124	370
normal Nef end	Nef	207	372

V-3 Sequences

Sequences included in the HIV-1 protein alignments.

Name	Accession	Proteins	Author	Reference
B.FR.83.HXB2	K03455	All	Wong-Staal, F.	<i>Nature</i> 313 (6000):277-284 (1985)
A1.CD.97.97CD_KCC2	AM000053	All	Vidal, N.	<i>ARHR</i> 22 (2):182-187 (2006)
A1.CM.08.886_24	KP718928	All	Luk, K.-C.	<i>PLoS One</i> 10 (11); e0141723 (2015)
A1.CY.08.CY236	JF683783	All	Kousiappa, I.	<i>ARHR</i> 27 (11); 1183-99 (2011)
A1.IN.09.NARI_FLS_IVC19_1	KT152839	All	Pandey, S.S.	<i>ARHR</i> 32 (5); 489-502 (2016)
A1.KE.11.DEMA111KE002	KF716474	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
A1.RU.11.11RU6950	JX500694	All	Baryshev, P.B.	<i>ARHR</i> 30 (6); 592-7 (2014)
A1.RW.11.DEMA111RW002	KF716472	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
A1.SN.01.DDI579	AY521629	All	Meloni, S.T.	<i>J Virol</i> 78 (22):12438-12445 (2004)

Name	Accession	Proteins	Author	Reference
A1.UG.11.DEMA110UG009	KF716486	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
A1.ZA.04.503_15344_T10_A1	KT183312	All	Hertz, T.	Unpublished
A2.CD.97.97CDKTB48	AF286238	All	Gao, F.	<i>ARHR</i> 17 (8):675-688 (2001)
A2.CM.01.01CM_1445MV	GU201516	All	Carr, J.K.	<i>Retrovirology</i> 2010 Apr 28;7:39 doi: 101186/1742-4690-7-39
A2.CY.94.94CY017_41	AF286237	All	Gao, F.	<i>ARHR</i> 17 (8):675-688 (2001)
B.BR.10.10BR_RJ032	KJ849801	All	Pessoa, R.	<i>Transfusion</i> 55 (5); 980-90 (2015)
B.CA.07.502_1191_03	JF320424	All	Rolland, M.	<i>Nat Med</i> 17 (3); 366-71 (2011)
B.CH.08.M2_0803101_NFLG8	KC797225	All	Castro, E.	<i>AIDS</i> 28 (12); 1840-4 (2014)
B.CN.12.DEMB12CN006	KP109511	All	Hora, B.	Unpublished
B.CU.14.14CU005	KR914676	All	Blanco, M.	Unpublished
B.ES.14.ARP1195	KT276255	All	Cuevas, M.T.	Unpublished
B.FR.11.DEMB11FR001	KF716496	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.HT.05.05HT_129389	EU839602	All	Nadai, Y.	<i>PLoS ONE</i> 4 (3):E4814 (2009)
B.JP.12.DEMB12JP001	KF716498	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.KR.07.HP_18_07JHS10_3909	KJ140263	All	Kim, B.-R.	<i>Haemophilia</i> 21 (1); e1-11 (2015)
B.RU.11.11RU21n	JX500708	All	Baryshev, P.B.	Unpublished
B.SE.12.SE600057	KP411828	All	Grossmann, S.	<i>J Int AIDS Soc</i> 2015 Jun 25;18:20035 doi: 107448/IAS18120035 eCollection 2015
B.TH.10.DEMB10TH002	KP109514	All	Hora, B.	Unpublished
B.US.13.RV_1	KT284371	All	Rassler, S.	Unpublished
B.ZA.09.DEMB09ZA022	KP109515	All	Hora, B.	Unpublished
C.BR.07.DEMC07BR003	JX140663	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
C.BW.00.00BW5031_1	AF443115	All	Novitsky, V.	<i>J Virol</i> 76 (11):5435-5451 (2002)
C.CN.10.YNFL19	KC870038	All	Wei, H.	Unpublished
C.CY.09.CY260	JF683803	All	Kousiappa, I.	<i>ARHR</i> 27 (11); 1183-99 (2011)
C.ES.14.ARP1198	KT276258	All	Cuevas, M.T.	Unpublished
C.ET.02.02ET_288	AY713417	All	Brown, B.K.	<i>J Virol</i> 79 (10):6089-6101 (2005)
C.IN.09.T125_2139	KC156210	All	Parrish, N.F.	<i>PNAS USA</i> 110 (17); 6626-33 (2013)
C.KE.05.05KE369195V4	KT022371	All	Billings, E.	<i>PLoS ONE</i> 10 (8):E0135124 (2015)
C.MW.09.703010256_CH256.w96	KC156214	All	Parrish, N.F.	<i>PNAS USA</i> 110 (17); 6626-33 (2013)
C.SE.13.SE600311	KP411835	All	Grossmann, S.	<i>J Int AIDS Soc</i> 2015 Jun 25;18:20035 doi: 107448/IAS18120035 eCollection 2015
C.TZ.08.707010457_CH457.w8	KC156220	All	Parrish, N.F.	<i>PNAS USA</i> 110 (17); 6626-33 (2013)

Name	Accession	Proteins	Author	Reference
C.US.11.17TB4_4G8	KF526226	All	Ho, Y.-C.	<i>Cell</i> 155 (3); 540-51 (2013)
C.YE.02.02YE511	AY795906	All	Saad, M.D.	<i>ARHR</i> 21 (7):644-648 (2005)
C.ZA.12.DEMC12ZA096	KP109517	All	Hora, B.	Unpublished
C.ZM.11.DEMC11ZM006	KF716467	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.CM.10.DEMD10CM009	JX140670	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.CY.06.CY163	FJ388945	All	Kousiappa, I.	<i>ARHR</i> 25 (8); 727-40 (2009)
D.KE.11.DEMD11KE003	KF716476	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.KR.04.04KBH8	DQ054367	All	Cho, Y.-K.	<i>ARHR</i> 29 (4); 738-43 (2013)
D.SN.90.SE365	AB485648	All	Takekawa, N.	Unpublished
D.TZ.01.A280	AY253311	All	Arroyo, M.A.	<i>ARHR</i> 20 (8):895-901 (2004)
D.UG.10.DEMD10UG004	KF716479	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.UG.11.DEMD11UG003	KF716480	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.YE.02.02YE516	AY795907	All	Saad, M.D.	<i>ARHR</i> 21 (7):644-648 (2005)
D.ZA.90.R1	EF633445	All	Jacobs, G.B.	<i>ARHR</i> 23 (12):1575-8 (2007)
F1.AO.06.AO_06_ANG32	FJ900266	All	Guimaraes, M.L.	<i>Retrovirology</i> 6 , 39 (2009)
F1.AR.02.ARE933	DQ189088	All	Aulicino, P.C.	<i>ARHR</i> 21 (2):158-164 (2005)
F1.BR.07.07BR844	FJ771010	All	Sanabani, S.S.	<i>Virol J</i> 2009 Jun 16;6:78
F1.BR.10.10BR_PE107	KJ849782	All	Pessoa, R.	<i>PLoS One</i> 9 (11); e112674 (2014)
F1.BR.10.10BR_RJ015	KJ849791	All	Pessoa, R.	<i>Transfusion</i> 55 (5); 980-90 (2015)
F1.CY.08.CY222	JF683771	All	Kousiappa, I.	<i>ARHR</i> 27 (11); 1183-99 (2011)
F1.ES.02.ES_X845_4	FJ670516	All	Fernandez-Garcia, A.	<i>ARHR</i> 25 (11):1187-1191 (2009)
F1.ES.11.VA0053_nfl	KJ883138	All	Delgado, E.	<i>PLoS ONE</i> 10 (11):E0143325 (2015)
F1.RO.96.BCI_R07	AB485658	All	Takekawa, N.	Unpublished
F1.RU.08.D88_845	GQ290462	All	Fernandez-Garcia, A.	<i>ARHR</i> 25 (11):1187-1191 (2009)
F2.CM.02.02CM_0016BBY	AY371158	All	Kijak, G.H.	<i>ARHR</i> 20 (5):521-530 (2004)
F2.CM.10.DEMF210CM001	JX140672	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
F2.CM.10.DEMF210CM007	JX140673	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
G.CM.07.920_49	KP718923	All	Luk, K.-C.	<i>PLoS One</i> 10 (11); e0141723 (2015)

Name	Accession	Proteins	Author	Reference
G.CM.10.DEMG10CM008	JX140676	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
G.CM.10.DEURF10CM020	KP109502	All	Hora, B.	Unpublished
G.CN.08.GX_2084_08	JN106043	All	Liu, W.	<i>Zhonghua Liu Xing Bing Xue Za Zhi</i> 34 (1); 53-6 (2013)
G.ES.09.X2634_2	GU362882	All	Cuevas, M.T.	<i>ARHR</i> 26 (9); 1019-25 (2010)
G.ES.14.ARP1201	KT276261	All	Cuevas, M.T.	Unpublished
G.GH.03.03GH175G	AB287004	All	Takekawa, N.	Unpublished
G.KE.09.DEMG09KE001	KF716477	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
G.NG.09.09NG_SC62	JN248593	All	Charurat, M.	<i>J Infect Dis</i> 205 (8); 1239-47 (2012)
G.PT.x.PT3306	FR846409	All	Freitas, F.B.	<i>ARHR</i> 2012 Sep 25
H.BE.93.VI991	AF190127	All	Janssens, W.	<i>AIDS</i> 14 (11):1533-1543 (2000)
H.BE.93.VI997	AF190128	All	Janssens, W.	<i>AIDS</i> 14 (11):1533-1543 (2000)
H.CF.90.056	AF005496	All	Gao, F.	<i>J Virol</i> 72 (7):5680-5698 (1998)
H.GB.00.00GBAC4001	FJ711703	All	Holzmayer, V.	<i>ARHR</i> 25 (7):721-726 (2009)
J.CD.97.J_97DC_KTB147	EF614151	All	Abecasis, A.B.	<i>J Virol</i> 81 (16):8543-8551 (2007)
J.SE.93.SE9280_7887	AF082394	All	Laukkanen, T.	<i>ARHR</i> 15 (3):293-297 (1999)
J.SE.94.SE9173_7022	AF082395	All	Laukkanen, T.	<i>ARHR</i> 15 (3):293-297 (1999)
K.CD.97.97ZR_EQTB11	AJ249235	All	Triques, K.	<i>ARHR</i> 16 (2):139-151 (2000)
K.CM.96.96CM_MP535	AJ249239	All	Triques, K.	<i>ARHR</i> 16 (2):139-151 (2000)
01_AE.AF.07.569M	GQ477441	All	Sanders-Buell, E.	<i>ARHR</i> 26 (5):605-608 (2010)
01_AE.CM.11.1156_26	KP718930	All	Luk, K.-C.	<i>PLoS One</i> 10 (11); e0141723 (2015)
01_AE.CN.12.DE00112CN011	KP109508	All	Hora, B.	Unpublished
01_AE.HK.04.HK001	DQ234790	All	Tsui, S.K.W.	Unpublished
01_AE.IR.10.10IR_THR48F	AB703616	All	Jahanbakhsh, F.	<i>ARHR</i> 29 (1); 198-203 (2013)
01_AE.JP.11.DE00111JP003	KF859741	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
01_AE.SE.11.SE601018	KP411841	All	Grossmann, S.	<i>J Int AIDS Soc</i> 2015 Jun 25;18:20035 doi: 107448/IAS18120035 eCollection 2015
01_AE.TH.10.DE00110TH001	KP109513	All	Hora, B.	Unpublished
01_AE.TH.90.CM240	U54771	All	Carr, J.K.	<i>J Virol</i> 70 (9):5935-5943 (1996)
01_AE.US.05.306163_FL	JX863920	All	Heipertz, R.A. Jr.	<i>ARHR</i> 29 (10):1310-1320 (2013)
02_AG.CM.10.DE00210CM013	KF859739	All	Sanchez, A.M.	<i>J Immunol Methods</i> 2014 Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
02_AG.ES.06.P1423	EU884501	All	Fernandez-Garcia, A.	<i>ARHR</i> 25 (1); 93-102 (2009)
02_AG.GW.05.CC_0048	FJ694792	All	Vinner, L.	<i>APMIS</i> 119 (8); 487-97 (2011)
02_AG.KR.12.12MHR9	KF561435	All	Cho, Y.-K.	Unpublished
02_AG.LR.x.POC44951	AB485636	All	Baesi, K.	<i>PLoS One</i> 9 (9); e105098 (2014)
02_AG.NG.09.09NG_SC61	JN248592	All	Charurat, M.	<i>J Infect Dis</i> 205 (8); 1239-47 (2012)
02_AG.NG.x.IBNG	L39106	All	Howard, T.M.	<i>ARHR</i> 10 (12):1755-1757 (1994)

Name	Accession	Proteins	Author	Reference
02_AG.SE.11.SE602024	KP411845	All	Grossmann, S.	<i>J Int AIDS Soc</i> 2015 Jun 25;18:20035 doi: 107448/IAS18120035 eCollection 2015
02_AG.SN.98.98SE_MP1211	AJ251056	All	Toure-Kane, C.	<i>ARHR</i> 16 (6):603-609 (2000)
02_AG.US.06.502_2696_FL01	JF320297	All	Rolland, M.	<i>Nat Med</i> 17 (3); 366-71 (2011)
03_AB.RU.97.KAL153_2	AF193276	All	Liitsola, K.	<i>ARHR</i> 16 (11):1047-1053 (2000)
04_cpx.CY.94.94CY032_3	AF049337	All	Gao, F.	<i>J Virol</i> 72 (12):10234-10241 (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, R.B.	<i>ARHR</i> 14 (16):1495-1500 (1998)
07_BC.CN.98.98CN009	AF286230	All	Rodenburg, C.M.	<i>ARHR</i> 17 (2):161-168 (2001)
08_BC.CN.97.97CNGX_6F	AY008715	All	Piyasirisilp, S.	<i>J Virol</i> 74 (23):11286-11295 (2000)
09_cpx.GH.96.96GH2911	AY093605	All	McCutchan, F.E.	<i>ARHR</i> 20 (8):819-826 (2004)
10_CD.TZ.96.96TZ_BF061	AF289548	All	Kouliniska, I.N.	<i>ARHR</i> 17 (5):423-431 (2001)
11_cpx.CM.95.95CM_1816	AF492624	All	Wilbe, K.	<i>ARHR</i> 18 (12):849-56 (2002)
12_BF.AR.99.ARMA159	AF385936	All	Carr, J.K.	<i>AIDS</i> 15 (15):F41-F47 (2001)
13_cpx.CM.96.96CM_1849	AF460972	All	Wilbe, K.	<i>ARHR</i> 18 (12):849-56 (2002)
14_BG.ES.05.X1870	FJ670522	All	Cuevas, M.T.	<i>ARHR</i> 26 (9); 1019-25 (2010)
15_01B.TH.99.99TH_MU2079	AF516184	All	Viputtijul, K.	<i>ARHR</i> 18 (16):1235-1237 (2002)
16_A2D.KR.97.97KR004	AF286239	All	Gao, F.	<i>ARHR</i> 17 (8):675-688 (2001)
17_BF.AR.99.ARMA038	AY037281	All	Carr, J.K.	<i>AIDS</i> 15 (15):F41-F47 (2001)
18_cpx.CU.99.CU76	AY586540	All	Thomson, M.M.	<i>AIDS</i> 19 (11):1155-63 (2005)
19_cpx.CU.99.CU7	AY894994	All	Casado, G.	<i>JAIDS</i> 40 (5):532-537 (2005)
20_BG.CU.99.Cu103	AY586545	All	Sierra, M.	<i>JAIDS</i> 45 (2):151-160 (2007)
21_A2D.KE.99.KER2003	AF457051	All	Dowling, W.E.	<i>AIDS</i> 16 (13):1809-1820 (2002)
22_01A1.CM.01.01CM_0001BBY	AY371159	All	Kijak, G.H.	<i>ARHR</i> 20 (5):521-530 (2004)
23_BG.CU.03.CB118	AY900571	All	Sierra, M.	<i>JAIDS</i> 45 (2):151-160 (2007)
24_BG.ES.08.X2456_2	FJ670526	All	Cuevas, M.T.	<i>ARHR</i> 26 (9); 1019-25 (2010)
25_cpx.CM.02.1918LE	AY371169	All	Kijak, G.H.	<i>ARHR</i> 20 (5):521-530 (2004)
26_AU.CD.02.02CD_MBTB047	FM877782	All	Vidal, N.	<i>ARHR</i> 25 (8):823-832 (2009)
27_cpx.FR.04.04CD_FR_KZS	AM851091	All	Vidal, N.	<i>ARHR</i> 24 (2):315-321 (2008)
28_BF.BR.99.BREPM12609	DQ085873	All	Sa Filho, D.J.	<i>ARHR</i> 22 (1):1-13 (2006)
29_BF.BR.01.BREPM16704	DQ085876	All	Sa Filho, D.J.	<i>ARHR</i> 22 (1):1-13 (2006)
31_BC.BR.04.04BR142	AY727527	All	Sanabani, S.	<i>ARHR</i> 22 (2):171-176 (2006)
32_06A1.EE.01.EE0369	AY535660	All	Adojaan, M.	<i>JAIDS</i> 39 (5):598-605 (2005)
33_01B.ID.07.JKT189_C	AB547463	All	SahBandar, I.N.	<i>ARHR</i> 27 (1); 97-102 (2011)
34_01B.TH.99.OUR2478P	EF165541	All	Tovanabutra, S.	<i>ARHR</i> 23 (6):829-833 (2007)
35_AD.AF.07.169H	GQ477446	All	Sanders-Buell, E.	<i>ARHR</i> 26 (5):605-608 (2010)
36_cpx.CM.00.00CMNYU830	EF087994	All	Powell, R.L.	<i>ARHR</i> 23 (8):1008-1019 (2007)
37_cpx.CM.00.00CMNYU926	EF116594	All	Powell, R.L.	<i>ARHR</i> 23 (7):923-933 (2007)
38_BF1.UY.03.UY03_3389	FJ213783	All	Ruchansky, D.	<i>ARHR</i> 25 (3); 351-6 (2009)
39_BF.BR.04.04BRRJ179	EU735535	All	Guimaraes, M.L.	<i>AIDS</i> 22 (3):433-435 (2008)
40_BF.BR.05.05BRRJ055	EU735537	All	Guimaraes, M.L.	<i>AIDS</i> 22 (3):433-435 (2008)
42_BF.LU.06.luBF_18_06	EU170139	All	Struck, D.	<i>ARHR</i> 31 (5); 554-8 (2015)
43_02G.SA.03.J11223	EU697904	All	Badreddine, S.	<i>ARHR</i> 23 (5):667-674 (2007)
44_BF.CL.00.CH80	FJ358521	All	Delgado, E.	<i>ARHR</i> 26 (7); 821-6 (2010)
45_cpx.FR.04.04FR_AUK	EU448295	All	Frangé, P.	<i>Retrovirology</i> 2008 Aug 1;5:69 doi: 101186/1742-4690-5-69
46_BF.BR.07.07BR_FPS625	HM026456	All	Sanabani, S.S.	<i>Virol J</i> 2010 Apr 16;7:74 doi: 101186/1743-422X-7-74
47_BF.ES.08.P1942	GQ372987	All	Fernandez-Garcia, A.	<i>ARHR</i> 26 (7); 827-32 (2010)
48_01B.MY.07.07MYKT021	GQ175883	All	Li, Y.	<i>JAIDS</i> 54 (2):129-136 (2010)

Name	Accession	Proteins	Author	Reference
49_cpx.GM.03.N26677	HQ385479	All	de Silva, T.I.	<i>Retrovirology</i> 7 (1):82 (2010)
50_A1D.GB.10.12792	JN417240	All	Foster, G.M.	<i>PLoS One</i> 9 (1); e83337 (2014)
51_01B.SG.11.11SG_HM021	JN029801	All	Ng, O.T.	<i>ARHR</i> 28 (5); 527-30 (2012)
52_01B.MY.03.03MYKL018_1	DQ366664	All	Tee, K.K.	<i>JAIDS</i> 43 (5):523-529 (2006)
53_01B.MY.11.11FIR164	JX390610	All	Chow, W.Z.	<i>J Virol</i> 86 (20):11398-11399 (2012)
54_01B.MY.09.09MYSB023	JX390976	All	Ng, K.T.	<i>J Virol</i> 86 (20):11405-11406 (2012)
55_01B.CN.10.HNCS102056	JX574661	All	Han, X.	<i>Genome Announc</i> 1 (1):E00050-12 (2013)
56_cpx.FR.10.URF5_patient_A	JN882655	All	Leoz, M.	<i>AIDS</i> 25 (11):1371-1377 (2011)
57_BC.CN.09.09YNLX19sg	KC899008	All	Han, X.	<i>PLoS ONE</i> 8 (5):E65337 (2013)
58_01B.MY.09.09MYPR37	KC522031	All	Chow, W.Z.	<i>PLoS ONE</i> 9 (1):E85250 (2014)
59_01B.CN.09.09LNA423	JX960635	All	An, M.	<i>J Virol</i> 86 (22); 12402-6 (2012)
60_BC.IT.11.BAV499	KC899079	All	Simonetti, F.R.	<i>Infect Genet Evol</i> 2014 Apr;23:176-81 doi: 101016/j.meegid.2014.02.007 Epub 2014 Mar 3
61_BC.CN.10.JL100010	KC990124	All	Li, X.	<i>Genome Announc</i> 2013 Jun 27;1(3) pii: e00326-13 doi: 101128/genomeA00326-13
62_BC.CN.10.YNFL13	KC870034	All	Wei, H.	<i>ARHR</i> 30 (4):380-383 (2014)
63_02A1.RU.10.10RU6637	JN230353	All	Baryshev, P.B.	<i>Arch Virol</i> 157 (12); 2335-41 (2012)
64_BC.CN.09.YNFL31	KC870042	All	Hsi, J.	<i>ARHR</i> 30 (4); 389-93 (2014)
65_cpx.CN.10.YNFL01	KC870027	All	Feng, Y.	<i>ARHR</i> 30 (6); 598-602 (2014)
67_01B.CN.11.ANHUI_HF115	KC183779	All	Wu, J.	<i>PLoS ONE</i> 8 (1):E54322 (2013)
68_01B.CN.11.ANHUI_WH73	KC183782	All	Wu, J.	<i>PLoS ONE</i> 8 (1):E54322 (2013)
69_01B.JP.05.05JPMYC113SP420	LC027100	All	Kusagawa, S.	<i>Genome Announc</i> 2015 May 28;3(3) pii: e00196-15 doi: 101128/genomeA00196-15
70_BF1.BR.10.10BR_PE004	KJ849758	All	Pessoa, R.	<i>PLoS One</i> 9 (11); e112674 (2014)
71_BF1.BR.10.10BR_PE008	KJ849759	All	Pessoa, R.	<i>PLoS One</i> 9 (11); e112674 (2014)
72_BF1.BR.10.10BR_MG002	KJ671534	All	Pessoa, R.	<i>Genome Announc</i> 2 (3):e00386-14 (2014)
74_01B.MY.10.10MYPR268	KR019771	All	Cheong, H.T.	<i>PLoS ONE</i> 10 (7):E0133883 (2015)
O.BE.87.ANT70	L20587	All	Vanden Haesevelde, M.	<i>J Virol</i> 68 (3):1586-1596 (1994)
O.CM.98.98CMA104	AY169802	All	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
O.CM.98.98CMABB141	AY169807	All	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
O.CM.98.98CMABB212	AY169804	All	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
O.CM.98.98CMU5337	AY169808	All	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
O.CM.99.99CMU4122	AY169815	All	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
O.FR.92.VAU	AF407418	All	Vartanian, J.P.	<i>J Gen Virol</i> 83 (Pt 4):801-805 (2002)
O.GA.11.11Gab6352	JX245015	All	Liegeois, F.	<i>ARHR</i> 29 (7); 1085-90 (2013)
O.SN.99.99SE_MP1299	AJ302646	All	Toure-Kane, C.	<i>ARHR</i> 17 (12):1211-1216 (2001)
O.US.10.LTNP	JN571034	All	Buckheit, R.W.3.	<i>ARHR</i> 30 (6); 511-513 (2014)
N.CM.02.DJO0131	AY532635	All	Bodelle, P.	<i>ARHR</i> 20 (8):902-908 (2004)
N.CM.02.SJGddd	GQ324959	All	Vallari, A.	<i>ARHR</i> 26 (1):109-115 (2010)
N.CM.04.04CM_1015_04	DQ017382	All	Yamaguchi, J.	<i>ARHR</i> 22 (1):83-92 (2006)
N.CM.06.U14296	GQ324962	All	Vallari, A.	<i>ARHR</i> 26 (1):109-115 (2010)
N.FR.11.N1_FR_2011	JN572926	All	Delaugerre, C.	<i>Lancet</i> 378 (9806); 1894 (2011)
P.CM.06.U14788	HQ179987	All	Vallari, A.	<i>J Virol</i> 85 (3); 1403-7 (2011)
P.FR.09.RBF168	GU111555	All	Plantier, J.-C.	<i>Nat Med</i> 15 (8); 871-2 (2009)

Name	Accession	Proteins	Author	Reference
CPZ.CD.90.ANT	U42720	All	Vanden Haesevelde, M.M.	<i>Virology</i> 221 (2):346-350 (1996)
CPZ.CM.05.SIVcpzLB7	DQ373064	All	Keele, B.F.	<i>Science</i> 313 (5786):523-526 (2006)
CPZ.CM.09.Ptt_09Cam155	FR686511	All	Etienne, L.	<i>Retrovirology</i> 8 (1):4 (2011)
CPZ.TZ.06.TAN5	JN091691	All	Rudicell, R.S.	<i>J Virol</i> 85 (19); 9918-28 (2011)
CPZ.US.85.US_Marilyn	AF103818	All	Gao, F.	<i>Nature</i> 397 (6718):436-441 (1999)
GOR.CM.07.SIVgor2139_287	FJ424866	All	Takehisa, J.	<i>J Virol</i> 83 (4):1635-1648 (2009)
GOR.CM.12.SIVgor_BQID2	KP004991	All	D'arc, M.	<i>PNAS USA</i> 112 (11); E1343-52 (2015)
GOR.CM.13.SIVgor_BPID15	KP004990	All	D'arc, M.	<i>PNAS USA</i> 112 (11); E1343-52 (2015)

	membrane binding										phosphorylation site																		
	Gag start, p17 start																												
	MGARASVLSGGEIWRKIRLRPGGKKYKLVHIVASRELERFAVNPGLLETSEGRQLTGLQPLSQTGSEELRSYNTVATLYCVHORIEIKDKEALDKIEEEQ.....NKS.....KKA.....QQA.....AADT.....GHS.....N.....126																												
02 AG.KR.12.12MHR9	K	A	R	L	L	A	Q	L	I	E	I	S	A	R	F	K	I	V	W	D	L	V	Q	T	A	I	125		
02 AG.LR.x.POC44951	K	A	R	L	L	A	Q	M	E	S	T	S	K	F	I	W	R	L	D	T	M	M	Q	T	A	S	125		
02 AG.NG.09.09NG SC61	#	K	A	R	L	L	A	Q	L	M	E	S	T	K	K	F	L	W	D	R	V	M	K	Q	T	A	S	123	
02 AG.NG.x.IBNG	K	A	R	L	L	A	Q	L	M	E	S	A	R	K	F	I	W	D	R	D	M	I	Q	T	A	S	125		
02 AG.SE.11.5602024	-	I	K	G	R	A	R	S	R	L	L	A	Q	L	E	S	T	K	K	F	I	W	R	D	T	A	S	124	
02 AG.SN.98.98SE MP1211	#	K	A	R	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	M	E	Q	T	A	S	124	
02 AG.US.06.502.2696 FL01	-	T	K	S	R	L	L	D	L	G	Q	M	E	T	R	K	F	V	L	V	L	I	Q	T	A	S	125		
03 AB.RU.97.KAL153.2	-	K	A	R	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126		
04 cpx.CY.94.94CY032.3	-	K	A	R	L	L	A	Q	L	M	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126	
05 DF.BE.x.VT110	-	K	A	R	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	127		
06 cpx.AU.96.BFP90	-	K	A	R	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	123		
07 BC.CN.98.98CN009	-	I	R	K	K	H	M	L	L	M	K	A	T	K	F	I	W	D	R	V	L	I	Q	T	A	S	125		
08 BC.CN.97.97CNX.6F	#	I	R	K	K	H	M	L	L	K	I	K	A	T	F	I	W	D	R	V	L	I	Q	T	A	S	121		
09 cpx.GH.96.96GH2911	-	K	A	R	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126		
10 CD.TZ.96.96TZ BF061	-	K	A	R	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126		
16 AZD.FR.97.97KR00	-	K	A	R	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126		
12 BF.AR.99.ARMA159	-	K	A	R	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	122		
13 cpx.CM.96.96CM 1849	-	K	K	R	M	L	L	T	Q	R	A	T	K	F	I	W	D	R	V	L	I	Q	T	A	S	126			
14 BG.ES.05.X1870	-	K	S	R	L	L	D	A	Q	M	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	125	
15 01B.TH.99.99TH MU2079	-	K	A	R	L	L	I	S	A	Q	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126	
17 BF.AR.99.ARMA038	-	K	A	R	L	L	I	S	A	Q	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126	
18 cpx.CU.99.CU76	-	K	K	R	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126		
19 cpx.CU.99.CU7	L	K	A	S	Q	R	M	L	I	S	A	Q	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	126	
20 BG.CU.99.Cu103	-	K	A	R	L	L	V	Q	I	N	A	K	T	K	F	I	W	D	R	V	L	I	Q	T	A	S	124		
21 AZD.KE.99.KER2003	-	K	A	R	L	L	V	Q	I	N	A	K	T	K	F	I	W	D	R	V	L	I	Q	T	A	S	124		
22 01A1.CM.01.01CM 0001LBY	R	A	K	A	Q	R	M	L	D	L	A	Q	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	124	
23 BG.CU.03.CB118	-	K	A	R	L	L	D	L	A	Q	M	S	A	K	T	K	F	I	W	D	R	V	L	I	Q	T	126		
24 BG.ES.08.X2456.2	-	K	A	R	L	L	D	L	A	Q	N	A	K	T	K	F	I	W	D	R	V	L	I	Q	T	A	S	126	
25 cpx.CM.02.1918LE	I	R	A	R	R	L	L	S	N	G	Q	S	A	K	T	K	F	I	W	D	R	V	L	I	Q	T	125		
26 AU.CD.02.02CD MBT047	I	Q	A	R	R	L	L	K	S	I	A	Q	A	V	K	K	F	I	W	D	R	V	L	I	Q	T	123		
27 cpx.FR.04.04FR K25	-	K	A	R	L	L	I	A	A	D	L	R	A	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126	
28 BF.BR.99.BREPM12609	-	K	A	R	L	L	I	A	A	D	L	R	A	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126	
29 BF.BR.01.BREPM16704	-	K	A	R	L	L	I	A	A	D	L	R	A	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	133	
31 BC.BR.04.04BR142	I	R	E	K	T	R	Q	M	L	L	D	G	K	I	Q	A	R	T	K	V	E	K	V	R	K	Q	T	125	
32 06A1.EE.01.EE0369	-	K	E	R	L	L	A	Q	V	V	A	R	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	125		
33 01B.ID.07.01R289.C	-	K	A	R	L	L	A	Q	I	D	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	124		
34 01B.TH.99.01R2478P	-	K	A	R	L	L	A	Q	I	D	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	124		
35 AD.AF.07.169H	I	R	K	T	R	L	L	S	T	Q	M	E	A	K	F	I	W	D	R	V	L	I	Q	T	A	S	125		
36 cpx.CM.00.00CMNYU830	-	K	A	R	L	L	D	S	A	Q	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	124	
37 cpx.CM.00.00CMNYU926	-	T	K	A	Q	R	M	L	L	A	Q	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	122
38 BF1.UY.03.UY03.3389	-	K	A	R	L	L	S	K	M	A	H	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126			
39 BF.BR.04.04BR3179	-	K	A	R	L	L	S	K	M	A	H	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126			
40 BF.BR.05.05BRJ055	-	K	A	R	L	L	S	A	A	H	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	127				
42 BF.LU.06.luBF 18.06	#	-	K	A	R	L	L	S	A	A	H	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126			
43 02G.SA.03.111223	-	K	A	R	L	L	I	S	A	Q	M	E	X	S	T	R	K	F	I	W	D	R	V	L	I	Q	T	125	
44 BF.CL.00.CH80	I	K	A	R	L	L	I	D	A	P	K	I	A	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	122	
45 cpx.FR.04.04FR AUK	-	K	A	R	L	L	D	A	Q	E	N	K	I	K	F	I	W	D	R	V	L	I	Q	T	A	S	126		
46 BF.BR.07.07BR FPS625	-	K	A	R	L	L	L	D	S	P	K	I	A	S	K	F	I	W	D	R	V	L	I	Q	T	A	S	122	
47 BF.ES.08.P1942	-	K	A	R	L	L	L	L	A	S	A	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	125			
48 01B.MY.07.07MYKT021	I	K	A	R	L	L	L	A	Q	L	I	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126
49 cpx.GM.03.N26677	I	K	K	R	L	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126	
50 A1D.GB.10.12732	-	K	T	R	L	L	D	A	L	M	D	I	A	R	T	X	F	I	W	D	R	V	L	I	Q	T	A	S	126
51 01B.SG.11.11SG HM021	#	-	K	R	R	L	L	D	A	L	M	D	I	A	R	T	X	F	I	W	D	R	V	L	I	Q	T	124	
52 01B.MY.03.03MYKL018.1	X	K	A	R	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126		
53 01B.MY.11.11FIR164	T	K	A	R	L	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126	
54 01B.MY.09.09MYSB023	I	Q	A	R	L	L	Y	L	S	K	E	A	K	G	F	I	W	D	R	V	L	I	Q	T	A	S	126		
55 01B.CN.10.HNCS102056	-	K	A	R	L	L	L	S	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126
56 cpx.FR.10.URF5 patient_A	-	K	A	R	L	L	D	L	A	Q	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126	
57 BC.CN.09.09YNLX1959	I	R	K	K	R	M	L	L	E	A	K	L	M	K	A	T	G	F	K	F	I	W	D	R	V	L	I	123	
58 01B.MY.09.09MYPR37	T	K	A	R	L	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	125	
59 01B.CN.09.09LNA423	I	T	K	A	R	L	L	K	Q	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	122		
60 BC.IT.11.BAV499	I	R	E	K	T	R	K	K	H	M	L	L	D	G	K	I	A	K	V	H	A	E	K	V	R	K	Q	T	122
61 BC.CN.10.1106010	I	R	K	K	H	M	L	L	L	K	I	A	T	F	I	W	D	R	V	L	I	Q	T	A	S	126			
62 BC.CN.10.YNFL13	I	R	K	K	H	M	L	L	D	K	L	M	K	A	T	F	I	W	D	R	V	L	I	Q	T	A	S	119	
63 02A1.RU.10.10RU6637	-	K	A	R	L	L	A	Q	L	M	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	125	
64 BC.CN.09.YNFL31	I	R	K	K	H	M	L	L	K	M	K	A	T	F	I	W	D	R	V	L	I	Q	T	A	S	123			
65 cpx.CN.10.YNFL01	T	K	A	R	L	L	I	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126	
67 01B.CN.11.ANHUI HF115	-	K	A	R	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126		
68 01B.CN.11.ANHUI WH73	-	K	A	R	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	126		
69 01B.JP.05.05JPMYC113SP420	-	K	A	R	L	L	A	Q	L	E	S	T	K	K	F	I	W	D	R	V	L	I	Q	T	A	S	122		
70 BF1.BR.10.10BR PE004	I	I	K	R	L	L	Q	R	V	A	A	R	A	K	F	I	W	D	R	V	L	I	Q	T	A	S	126		
71 BF1.BR.10.10BR PE008	-	K	K	R	L	L	L	A	A	R	A	K	F	I	W	D</													

	p17 end_p24 start	CyPa binding	
B. FR. 83. HXB2	... QV ...	SONYPIVQNIQGMVHQIAPSRPTLNANWVKVVEEKAFSPVIMPFMSALSEGATPQDLNMTLNTVGGHQAMQMLKETINEEAEDWRVHPVHAGPIAPGQMRPRGSDIAGTTSTLQEQIGWMT	... NNPPPI. PVGEIYKRWIIILGLNKIVRMYSPITSILD
02 AG. KR. 12. 12MHR9	... -A -T PM	... M I D	... S -A -V -V
02 AG. LR. x. POC44951	... -A -T S	... M I D	... V -V -V
02 AG. NG. 09. 09NG SC61	... -A -T PM	... M I D	... V -V -V
02 AG. NG. x. IBNG	... -AK -T SM	... I -G -T	... V -V -V
02 AG. SE. 11. 5602024	... -A -T S	... M I D	... D -V -V
02 AG. SN. 98. 98SE MP1211	... -F -A -S	... M I D	... D -V -V
02 AG. US. 06. 502 2696 FL01	... -F -A -WT NM	... M I D	... D -V -V
03 AB. RU. 97. KAL153 2	... -K -F -A -T SM	... M I D	... D -V -V
04 cpx. CY. 94. 94CY032 3	... -N -A -S	... M I D	... D -V -V
05 DF. BE. x. VT110	... -A -A -M	... M I D	... D -V -V
06 cpx. AU. 96. BFP90	... -NL -A -L	... I -D -T	... D -V -V
07 BC. CN. 98. 98CN009	... -K -L -L	... M I D	... D -V -V
08 BC. CN. 97. 97CNGX 6F	... -K -P -PL	... M I D	... D -V -V
09 cpx. GH. 96. 96GH2911	... -K -A -I	... M I D	... D -V -V
10 CD. TZ. 96. 96TZ BF061	... -K -L -I PLT	... M I D	... D -V -V
11 cpx. CR. 95. 95CM 1816	... -K -A -I	... M I D	... D -V -V
12 BF. AR. 99. ARMA159	... -G -L -PL	... M I D	... D -V -V
13 cpx. CM. 96. 96CM 1849	... -A -A -S	... M I D	... D -V -V
14 BG. ES. 05. X1870	... -A -A -P	... L -DA	... M -OQ
15 01B. TH. 99. 99TH MU2079	... -T -H -A -PV	... G -N	... M -I
16 AZD. FR. 97. 97KR00	... -K -L -I	... M I D	... D -V -V
17 BF. AR. 99. ARMA038	... -KG -L -TY ML	... M I D	... D -V -V
18 cpx. CU. 99. CU76	... -H -A -S	... M I D	... D -V -V
19 cpx. CU. 99. CU7	... -H -HKL -L	... I -N	... M I D
20 BG. CU. 99. CU103	... -A -A -PI	... M I D	... D -V -V
21 AZD. KE. 99. KER2003	... -A -A -W	... M I D	... D -V -V
22 01A1. CM. 01. 01CM 0001BBY	... -N -A -I L	... G -N	... T -T -M -II
23 BG. CU. 03. CB118	... -A -A -P	... M I D	... D -V -V
24 BG. ES. 08. X2456 2	... -A -A -PI	... M I D	... D -V -V
25 cpx. CM. 02. 1918LE	... -A -A -PI	... M I D	... D -V -V
26 AU. CD. 02. 02CD MBT047	... -K -A -L	... M I D	... D -V -V
27 cpx. FR. 04. 04CD K25	... -A -H -L	... M I D	... D -V -V
28 BF. BR. 99. BREPM12609	... -NL -L -L	... M I D	... D -V -V
29 BF. BR. 01. BREPM16704	... -GNSS -L -P A	... M I D	... D -V -V
31 BC. BR. 04. 04BR142	... -N -L -L	... M I D	... D -V -V
32 06A1. EE. 01. EE0369	... -T -F -A -T S	... M I D	... D -V -V
33 01B. ID. 07. JKT189 C	... -K -A -PV	... G -N	... M I D
34 01B. TH. 99. TH99 0UR2478P	... -K -A -PV	... G -N	... M I D
35 AD. AF. 07. 169H	... -K -A -I SL	... G -N	... M I D
36 cpx. CM. 00. 00CMNYU830	... -T -R -A -A	... G -N	... M I D
37 cpx. CM. 00. 00CMNYU926	... -KI -A -I SL	... T -D -H -A -V	... M I D
38 BF1. UY. 03. UY03 3389	... -G -L -L	... M I D	... D -V -V
39 BF. BR. 04. 04BRR3179	... -K -L -L	... M I D	... D -V -V
40 BF. BR. 05. 05BRRJ055	... -K -L -L	... M I D	... D -V -V
42 BF. LU. 06. LuBF 18 06	... -L -L -L	... M I D	... D -V -V
43 02G. SA. 03. J11223	... -G -A -SL	... M I D	... D -V -V
44 BF. CL. 00. CH80	... -G -L -L	... M I D	... D -V -V
45 cpx. FR. 04. 04FR AUK	... -G -L -I L	... M I D	... D -V -V
46 BF. BR. 07. 07BR FPS625	... -G -L -L	... M I D	... D -V -V
47 BF. ES. 08. P1942	... -A -H -I -P	... C	... M I D
48 01B. MY. 07. 07MYKT021	... -K -A -PV	... G -N	... M I D
49 cpx. GM. 03. N26677	... -SI -T -PM	... M I D	... D -V -V
50 A10. GB. 10. 12702	... -A -L -L	... M I D	... D -V -V
51 01B. SG. 11. 11SG HM021	... -K -A -L	... M I D	... D -V -V
52 01B. MY. 03. 03MYKL018 1	... -K -A -PV	... G -N	... M I D
53 01B. MY. 11. 11FIR164	... -K -M -A -X	... G -N	... M I D
54 01B. MY. 09. 09MYSB023	... -K -L -L	... M I D	... D -V -V
55 01B. CN. 10. HNC5102056	... -K -A -PL	... G -N	... M I D
56 cpx. FR. 10. URF5 patient_A	... -K -L -L	... M I D	... D -V -V
57 BC. CN. 09. 09YNLX195g	... -K -L -L	... M I D	... D -V -V
58 01B. MY. 09. 09MYPR37	... -K -A -PV	... G -N	... M I D
59 01B. CN. 09. 09LNA423	... -KI -A -T SL	... G -N	... M I D
60 BC. IT. 11. BAV499	... -K -I -L	... M I D	... D -V -V
61 BC. CN. 10. J1100010	... -K -L -L	... M I D	... D -V -V
62 BC. CN. 10. YNFL13	... -KG -L -L	... M I D	... D -V -V
63 02A1. RU. 10. 10RU6637	... -K -A -T SM	... M I D	... D -V -V
64 BC. CN. 09. YNFL31	... -K -L -L	... M I D	... D -V -V
65 cpx. CN. 10. YNFL01	... -K -L -L	... M I D	... D -V -V
67 01B. CN. 11. ANHUI HF115	... -K -A -PV	... G -N	... M I D
68 01B. CN. 11. ANHUI WH73	... -K -A -PV	... G -N	... M I D
69 01B. JP. 05. 05JPMYC113SP420	... -K -L -L	... M I D	... D -V -V
70 BF1. BR. 10. 10BR PE004	... -SSQSG -M -I	... I -K	... M I D
71 BF1. BR. 10. 10BR PE008	... -T -H -L -P	... T -D -S -A -I	... M I D
72 BF1. BR. 10. 10BR MG002	... -P -L -L	... M I D	... D -V -V
74 01B. MY. 10. 10MYPR268	... -K -L -L	... M I D	... D -V -V
0. BF. 87. ANT70	... -T -G -S -A	... A -N -I -M	... TSY -I -AI
0. CM. 98. 98CMA104	... -T -G -V -A	... A -N -I -M	... SV -I -AI
0. CM. 98. 98CMA1141	... -T -G -V -P -A	... I -N -I -M	... I -Y -I -AI
0. CM. 98. 98CMA1212	... -T -G -V -A -A	... X -N -I -M	... I -Y -I -AI
0. CM. 98. 98CMU5337	... -T -G -V -P -A	... H -L -L	... I -Y -I -AI
0. CM. 99. 99CMU4122	... -T -G -V -P -A	... H -L -L	... I -Y -I -AI
0. FR. -VAU -V -VAU	... -T -G -V -A -A	... P -L -L	... I -Y -I -AI
0. GA. 11. 11GAB6352	... -T -G -V -IP -A	... A -K -N -I -M	... I -Y -I -AI
0. SN. 99. 99SE MP1299	... -T -G -V -T -A	... A -N -I -M	... I -Y -I -AI
0. US. 10. LTNP	... -T -G -V -A -A	... T -VPL -L -L	... V -Y -I -AI
N. CM. 02. DJ00131	... -NI -R -L -A	... I -PLT -L	... I -N -M -A
N. CM. 02. SJGdd	... -NI -R -L -TA	... I -PLT -L	... I -N -M -S
N. CM. 04. 04CM 1015 04	... -NI -R -L -A	... H -PLT -L	... I -N -M -S
N. CM. 06. U14296	... -NI -R -L -A	... H -PLT -L	... I -N -M -S
N. FR. 11. N1 FR 2011	... -NI -R -L -A	... I -PLT -L	... I -N -M -S
P. CM. 06. U14788	... -GT -A -T -MPL	... A -N -I -M	... I -D -I -A
P. FR. 09. RBF168	... -GT -A -T -LPL	... A -N -I -M	... SI -D -I -A
CPZ. CD. 90. ANT	... RHLSGE -GR -IVDAG -IAR -PLT	... C -N -N -H	... A -D -G -V -V
CPZ. CM. 05. SIVcpzLB7	... -SG -R -T -A -P -L	... A -A -V -Q	... L -I -M -V
CPZ. CM. 09. Ptt 09Cam155	... -POIS -M -VI -A -P -L	... V -Q -L -V -Q	... AI -G -V -V
CPZ. TZ. 06. TAN5	... -SG -RL -VITDA -IAR -SP	... T -D -G -N	... Y -S -AI
CPZ. US. 98. US Mgr1lyn	... -G -R -VI -V -M	... A -M -A -D	... G -V -A -D
GOR. CM. 07. SIVgor2139 287	... -ET -G -V -A -A	... PL -L -L	... A -N -I -M
GOR. CM. 12. SIVgor_B01D2	... -GT -G -V -P -A	... T -PL -L	... A -N -I -M
GOR. CM. 13. SIVgor_BP1D15	... -GT -G -V -A -A	... I -PL -L	... A -N -I -M

	major homology region				p24 end_p2 start				p2 end_p7 start				Zn motif		Zn motif		p7 end
	IRQGPKEPFRDYVDRFYKTLRAEQASQEVKNMWTETLLVQNANPCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVT N . . . SAT . IMM. ORGNF . . . RNQ . RKIVKCFNCGKEGHTARNCRAPRKKGKWCQKGEHQMKDCT . . . E. ROAN																
B. FR. 83. HXB2	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432
A1. CD. 97. 97CD KCC2	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
A1. CM. 08. 886 24	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
A1. CY. 08. CY236	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	435
A1. TN. 09. NARI FLS TVC19 1	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
A1. KE. 11. DEM11KE002	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
A1. RU. 11. DEURU6950	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	426
A1. RW. 11. DEM11RW002	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	431
A1. SN. 01. DDI579	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	428
A1. UG. 11. DEM110UG009	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432
A1. ZA. 04. 50112344 T10 A1	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	437
A2. CD. 97. 97CDKTB48	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	436
A2. CM. 01. 01CM 1445MV	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	428
A2. CY. 94. 94CY017 41	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	428
B. BR. 10. 10BR RJ032	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	434
B. CA. 07. 502 1191 03	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	434
B. CH. 08. M2 0803101 NFLG8	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	431
B. CN. 12. DEM12CM006	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432
B. CU. 14. 14CU005	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432
B. ES. 14. ARP1195	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	433
B. FR. 11. DEMB11FR001	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432
B. HT. 05. 05HT 129389	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	429
B. JP. 12. DEMB12JP001	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432
B. KR. 07. HP 18 07JHS10 3909	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432
B. RU. 11. 11RU21n	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	437
B. SE. 12. SE600057	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	435
B. TH. 10. DEMB10TH002	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432
B. US. 13. RV 1	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	434
B. ZA. 09. DEMB09ZA022	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	431
C. BR. 02. 02YXC07BR003	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	426
C. BU. 00. 00BUB031 1	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	428
C. CN. 10. YNFI19	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	427
C. CY. 09. CY260	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	433
C. ES. 14. ARP1198	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	428
C. ET. 02. 02ET 268	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	427
C. IN. 09. T125 2139	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	427
C. KE. 05. 05KE369195V4	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	424
C. MW. 09. 703010256 CH256.w96	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	427
C. SE. 13. SE600311	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	428
C. TZ. 08. 707810457 CH457.w8	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	427
C. US. 11. 11784 468	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	429
C. YE. 02. 02YE511	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	426
C. ZA. 12. DEMC12ZA096	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	428
C. ZM. 11. DEMC11ZM006	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	429
D. CM. 10. DEMD10CM009	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
D. CY. 06. CY163	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432
D. KE. 11. DEMD11KE003	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	433
D. KR. 04. 04KR88	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432
D. SN. 90. SE365	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	433
D. TZ. 01. A280	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	427
D. UG. 10. DEMD10UG004	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	433
D. UG. 11. DEMD11UG003	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	431
D. YE. 02. 02YE516	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	431
D. ZA. 90. R1	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	433
F1. AO. 06. A0 06 ANG32	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	428
F1. AR. 02. ARE933	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	427
F1. BR. 07. 07BR844	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432
F1. BR. 10. 10BR PE107	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	431
F1. BR. 10. 10BR RJ015	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
F1. CY. 08. CY222	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	426
F1. ES. 02. ES X845 4	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	428
F1. ES. 11. VA0053 nfl	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	427
F1. RO. 96. BCI R07	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	427
F1. RU. 08. D88 845	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	433
F2. CM. 02. 02CM 6016BBY	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	425
F2. CM. 10. DEMF210CM001	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
F2. CM. 10. DEMF210CM007	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	427
G. CM. 07. 920 49	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	429
G. CM. 10. DEMG10CM008	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432
G. CM. 10. DEURF10CM020	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432
G. CN. 08. GX 2084 08	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	429
G. ES. 09. X2634 2	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
G. ES. 14. ARP1201	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	434
G. GH. 03. 03GH175G	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	431
G. KE. 09. DEMG09KE001	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	431
G. NG. 09. 09NG SC62	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	426
G. PT. x. PT3306	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
H. BE. 93. V1991	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	435
H. BE. 93. V1997	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	433
H. CF. 90. 056	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	433
H. GB. 00. 00GBAC4001	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	433
J. CD. 97. J 97DC KTB147	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	421
J. SE. 93. SE9280 7887	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	431
J. SE. 94. SE9173 7022	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	431
K. CD. 97. 97ZR E0TB11	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	427
K. CM. 96. 96CM MP535	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	427
01 AE. AF. 07. 569M	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
01 AE. CM. 11. 115 26	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
01 AE. CN. 12. DE0012CN011	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
01 AE. HK. 04. HK001	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
01 AE. IR. 10. 10IR. THR48F	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	431
01 AE. JP. 11. DE00111JP003	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
01 AE. SE. 11. SE601038	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	434
01 AE. TH. 10. DE00110TH001	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
01 AE. TH. 90. CM240	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	430
01 AE. US. 05. 306163 FL	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	427
02 AG. CM. 10. DE00210CM013	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432
02 AG. ES. 06. P1423	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	427
02 AG. GW. 05. CC 0048	-K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	432

Table with columns: major homology region, p24 end_p2 start, p2 end_p7 start, Zn motif, Zn motif, p7 end. Rows list various HIV-1 protein sequences and their amino acid alignments.

	p1 start	p1 end	p6 start	Vpr binding	Vpr binding	p6 end	Gag end															
B. FR. 83. HXB2	FLGKIWPS. YK	GRPGNFIQSRP. E	PT	APPE	ESFRSGVGE. TT	T	PP	Q	KQEPTI	D	K	E	LY	PLTSLRSFLGNDPSSQ*	500							
02 AG. KR. 12. 12MHR9	S	P	A	LGM	E	I	S	P	QREEP	R	D	G	P	K	499							
02 AG. LR. x. POC44951	S	P	A	GM	E	I	S	P	R	D	G	P	K	*	495							
02 AG. NG. 09. 09NG SC61	S	P	A	VGMRE	A	I	SS	P	G	D	P	K	*	495								
02 AG. NG. x. IBNG	S	P	A	GM	E	I	P	Q	R	D	P	K	*	495								
02 AG. SE. 11. 5602024	S	P	A	GM	E	I	SS	P	E	G	P	K	*	495								
02 AG. SN. 98. 98SE MP1211	S	P	A	GM	E	I	SS	P	E	G	P	K	*	495								
02 AG. US. 06. 502 2696 FL01	R	S	A	GM	E	V	SS	P	R	D	P	K	*	495								
03 AB. RU. 97. KAL153 2	R	S	A	N	GM	E	I	L	OK	R	QHP	SI	K	D	499							
04 cpx. CY. 94. 94CY032 3	RM	S	N	A	CLERKE	E	SS	L	R	D	P	K	*	501								
05 DF. BA. x. VT1110	V	H	S	A	GM	E	I	A	SS	P	OK	E	G	P	503							
06 cpx. AU. 96. BFP90	H	N	S	A	GM	E	I	A	SS	P	OK	E	E	E	500							
07 BC. CN. 98. 98CN009	N	S	A	F	EM	T	S	S	T	S	K	E	E	K	496							
08 BC. CN. 97. 97CNGX 6F	F	H	S	A	FE	E	A	P	K	R	G	S	L	*	490							
09 cpx. GH. 96. 96GH2911	N	P	S	A	GMRE	I	SS	X	XKD	R	XPP	X	K	L	497							
10 CD. TZ. 96. 96TZ BF061	N	S	A	GF	E	I	A	SS	OK	E	PH	A	K	L	501							
11 cpx. CM. 95. 95CM 1816	N	S	A	GF	E	I	A	SS	OK	E	PH	A	K	L	499							
12 BF. AR. 99. ARMA159	N	S	A	GF	E	I	A	SS	P	R	OK	E	E	G	494							
13 cpx. CM. 96. 96CM 1849	N	S	A	GF	E	I	A	SS	P	K	E	T	AA	S	508							
14 BG. ES. 05. X1870	R	S	A	GF	E	I	A	SS	P	K	E	M	A	K	497							
15 01B. TH. 99. 99TH MU2079	RL	S	A	NWGM	E	I	SL	L	OK	R	HTP	V	K	LL	498							
16 AZD. KR. 97. 97KR00	HS	P	A	D	GM	E	I	D	P	K	N	R	QHT	AI	501							
17 BF. AR. 99. ARMA038	N	S	A	LG	E	I	A	SS	P	SK	E	E	K	Y	497							
18 cpx. CU. 99. CU76	N	S	A	LG	E	I	A	SS	P	SK	E	E	K	Y	497							
19 cpx. CU. 99. CU7	H	S	A	GF	E	I	A	SS	P	DOK	E	D	A	K	501							
20 BG. CU. 99. CU103	L	H	S	A	GF	E	I	A	SS	P	GOK	E	D	T	500							
21 AZD. KE. 99. KER2003	S	S	A	H	GM	E	I	SS	OK	E	D	SV	K	LL	497							
22 02A1. CM. 01. 01CM 0001BBY	N	P	S	A	VGM	E	I	SS	OK	E	OPP	V	K	L	499							
23 BG. CU. 03. CB118	L	H	S	A	GF	E	I	A	SS	P	G	K	E	Q	498							
24 BG. ES. 08. X2456 2	L	H	S	A	LG	E	I	A	SS	P	GOP	E	D	I	497							
25 cpx. CM. 02. 1918LE	R	N	S	A	LG	E	I	A	SS	P	OK	E	MA	K	499							
26 AU. CD. 02. 02CD MBT047	RF	P	N	A	NLGM	E	EG	S	P	KEG	OSP	T	K	L	496							
27 cpx. FR. 04. 04CD FR K25	SP	N	S	A	LG	E	I	M	S	P	OK	E	E	G	499							
28 BF. BR. 99. BREPM12609	H	N	S	A	F	E	I	S	P	OK	E	E	G	P	499							
29 BF. BR. 01. BREPM16704	H	N	S	A	LG	E	R	S	P	OK	E	E	G	P	510							
31 BC. BR. 04. 04BR142	R	H	S	A	FE	E	A	SS	P	OK	R	G	K	P	506							
32 06A1. EE. 01. EE0369	H	S	A	FE	E	A	A	SS	S	K	K	SA	K	*	495							
33 01B. ID. 07. JKT189 C	L	H	S	A	FE	E	A	SS	S	Q	VA	K	*	497								
34 01B. TH. 99. QUR2478P	F	N	S	A	FE	E	A	SS	S	Q	VA	K	*	498								
35 AD. AF. 07. 169H	NN	P	S	A	GF	E	I	K	SS	OK	G	K	*	497								
36 cpx. CM. 00. 00CMNYU830	S	P	S	A	GF	E	I	S	SS	SR	G	P	K	*	498							
37 cpx. CM. 00. 00CMNYU926	R	S	A	EVKE	V	M	A	SS	L	G	E	G	P	A	493							
38 BF1. UY. 03. UY03 3389	H	S	A	GF	E	I	P	KQE	OK	E	G	M	P	A	499							
39 BF. BR. 04. 04BR13179	N	S	A	LG	E	I	A	SS	P	OK	N	E	G	P	502							
40 BF. BR. 05. 05BRR1055	P	N	S	A	GF	E	I	A	SS	P	OK	N	E	G	500							
42 BF. LU. 06. LuBF 18 06	N	N	S	A	GF	E	I	SS	L	OK	E	E	G	P	503							
43 02G. SA. 03. J11223	H	S	A	GF	E	I	A	SS	P	R	E	SP	A	K	498							
44 BF. CL. 08. CH80	N	P	N	A	GF	E	I	SS	P	OK	E	G	K	P	501							
45 cpx. FR. 04. 04FR AUK	SR	N	S	A	GF	E	I	SS	P	OK	E	G	K	P	501							
46 BF. BR. 07. 07BR FPS625	N	S	A	EIRE	E	I	S	P	OK	E	E	G	P	A	496							
47 BF. ES. 08. P1942	S	N	S	A	LG	E	I	S	P	OK	E	E	G	P	498							
48 01B. MY. 07. 07MYKT021	L	H	S	A	F	E	I	A	SS	P	OK	E	A	S	499							
49 cpx. GM. 03. N26677	R	H	S	A	GF	E	I	A	SS	P	OK	E	SA	K	496							
50 A10. GB. 10. 1272	H	S	A	GF	E	I	T	S	OK	E	SA	K	*	498								
51 01B. SG. 11. 11SG HM021	H	S	A	GF	E	I	T	S	OK	E	SA	K	*	499								
52 01B. MY. 03. 03MYK1018 1	N	S	A	NWGM	E	X	SS	L	OX	D	PPP	SV	K	500								
53 01B. MY. 11. 11FIR164	L	X	H	S	F	EG	S	S	L	OX	D	PPP	SV	K	499							
54 01B. MY. 09. 09MYSB023	H	S	A	L	E	E	S	S	L	OX	D	PPP	SV	K	501							
55 01B. CN. 10. HNC5102056	RL	N	S	A	DWGM	E	I	SS	P	OK	G	HRP	VV	K	499							
56 cpx. FR. 10. URF5 patient A	N	P	S	A	VGM	E	A	SS	P	R	SRD	K	P	S	500							
57 BC. CN. 09. 09YNLX1959	L	H	S	A	FE	E	A	SS	P	LK	R	G	K	L	492							
58 01B. MY. 09. 09MYPR37	F	P	N	S	FE	E	X	G	X	G	K	S	*	498								
59 01B. CN. 09. 09LNA423	N	S	A	NWGM	E	S	L	L	OK	R	OPH	SI	K	L	492							
60 BC. IT. 11. BAV499	R	H	S	A	FE	E	A	SS	P	OK	R	K	S	L	491							
61 BC. CN. 10. J1100010	H	S	A	FE	E	A	SS	P	OK	R	K	S	L	492								
62 BC. CN. 10. YNFL13	H	S	A	FE	E	A	SS	P	OK	R	K	S	L	488								
63 02A1. RU. 10. 10RU6637	RL	NN	P	A	VGM	E	I	SS	P	Q	R	G	P	A	495							
64 BC. CN. 09. YNFL31	H	S	A	F	E	T	S	P	GT	R	G	P	A	K	496							
65 cpx. CN. 10. YNFL01	H	S	A	FE	E	A	SS	P	R	SK	R	G	P	A	495							
67 01B. CN. 11. ANHUI HF115	N	S	A	DWGM	E	I	SS	P	OK	R	Q	SV	K	L	502							
68 01B. CN. 11. ANHUI WH73	N	S	A	DWGM	E	I	SS	P	OK	R	Q	SV	K	L	502							
69 01B. JP. 05. 05JPMYC113SP420	H	S	A	L	EET	L	D	S	OK	E	G	R	P	A	495							
70 BF1. BR. 10. 10BR PE004	R	H	S	A	F	E	M	P	KK	QE	OK	E	G	R	504							
71 BF1. BR. 10. 10BR PE008	RM	S	N	A	GF	E	M	S	OK	S	E	G	R	P	499							
72 BF1. BR. 10. 10BR MG002	L	H	G	N	A	GF	E	K	S	P	K	E	G	P	499							
74 01B. MY. 10. 10MYPR268	L	H	S	A	GF	E	I	A	T	N	SA	K	*	499								
0. BE. 87. ANT70	Y	P	GG	T	YV	RPA	H	S	M	EEVK	Q	N	E	KGG	P	N	FA	K	T	Q	L	0
0. CM. 98. 98CMA184	Y	P	GG	T	M	KQV	A	S	M	EEVK	Q	N	E	KGG	P	N	FA	K	T	Q	L	0
0. CM. 98. 98CMABB141	Y	P	GG	A	YM	KQV	P	S	M	EMMK	Q	N	E	KGD	T	T	FA	K	T	Q	L	0
0. CM. 98. 98CMABB212	Y	P	GG	T	Y	KQV	S	X	EIVK	Q	N	S	GXX	Q	X	FA	K	T	Q	L	0	
0. CM. 98. 98CMU5337	Y	P	GG	T	YA	ROV	S	M	EEVKRQ	N	E	KGD	Q	N	FA	K	T	Q	L	0		
0. CM. 99. 99CMU4122	Y	P	GG	T	YA	ROV	S	M	EEVKRQ	N	E	KGD	Q	N	FA	K	T	Q	L	0		
0. FR. 92. VAU	K	V	RG	T	VV	EIV	K	S	M	EMK	Q	N	E	KGD	P	N	FA	K	T	Q	L	0
0. GA. 11. 11Gab6352	Y	P	GG	T	YV	KQV	H	S	M	EEVOK	Q	N	E	KGD	P	N	FA	K	T	Q	L	0
0. SN. 99. 99SE MP1299	Y	P	GG	T	YA	ROV	S	M	TEEMK	Q	N	E	KED	Q	N	FA	K	T	Q	L	0	
0. US. 10. LTNP	Y	P	GG	T	YV	KQV	A	S	M	EEVNRQ	N	E	KGS	P	T	FA	K	T	Q	L	0	
N. CM. 02. DJ00131	G	SP	F	P	TTTT	RKE	L	NYG	QE	R	S	TO	G	ERQE	N	TENS	P	L	513			
N. CM. 02. SJGddd	RG	SP	F	P	TTA	RKE	L	Y	GE	R	S	TO	G	EMOE	N	QERTENS	P	L	508			
N. CM. 04. 04CM 1015 04	G	SP	F	P	TAT	RKE	L	NYG	QE	R	S	TO	G	EMOE	N	QERTENS	P	L	512			
N. CM. 06. U14296	RG	SP	F	P	TTT	RKE	L	NYG	QE	R	S	TO	G	EMOE	N	QERTENS	P	L	502			
N. FR. 11. N1 FR 2011	RG	SP	F	P	TTT	RKE	L	NYG	QE	R	S	TO	G	EMOE	N	QERTENS	P	L	513			
P. CM. 06. U14788	P	GG	K	YV	KQV	Q	M	FEEMTQK	XE	KE	E	K										

	protease end_p66, p51 RT start	M41L	K70R D67N	D110 catalytic site	
B.FR.83.HXB2	RVRQVDYDIIIEICGHKAIGTVLVGPTVNIIGRNLTLQIGCTLNFFIPISPIETVPVKLKPMDGPKVKQWPLTEEEKILALVEICTEMEKEGKISKIPENPYNTVPVFAIKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGL			.KKKKSVTVLVDVGDYFVSVPLDEDF	279
A1.CD.97.97CD.KCC2	--K--K--K--M--S--T--N--I--R--			.AN--	279
A1.CM.08.886.24	--K--E--K--M--T--T--T--I--			.H--	279
A1.CY.08.CY236	--K--P--K--M--R--X--T--I--			.Y--	276
A1.IN.09.NARI.FLS.TVC19.1	--K--K--K--M--E--X--T--I--			.H--	276
A1.KE.11.DEM111KE002	--K--K--K--M--T--T--R--I--			.Y--	278
A1.RU.11.DERRU6950	--K--K--K--M--L--T--T--K--			.S--	279
A1.RW.11.DEM111RW002	--K--K--K--M--L--T--T--A--			.G--	279
A1.SN.01.DDI579	--K--K--K--M--K--T--T--A--K--E--			.S--	278
A1.UG.11.DEM110UG009	--K--E--P--K--M--Q--T--I--			.S--	279
A1.ZA.04.502.12344.T10.A1	--K--P--K--M--T--T--I--			.H--	279
A2.CD.97.97CDKTB48	--V--K--K--M--V--L--T--K--			.D--	279
A2.CM.01.01CM.1445MV	--V--EKR--M--I--V--L--T--K--			.D--	279
A2.CY.94.94CY017.41	--A--KR--M--V--L--T--K--			.A--	279
B.BR.10.10BR.RJ032	--K--P--T--I--L--			.KE--	278
B.CA.07.502.1191.03	--K--P--T--I--L--			.Y--	279
B.CH.08.M2.0803101.NFLG8	R--K--P--V--I--VM--L--			.G--N--	283
B.CN.12.DEM12CM006	--P--K--K--M--V--			.I--	284
B.CU.14.14CU065	--VPV--E--I--R--			.S--Q--I--	279
B.ES.14.ARP1195	--QV--E--I--R--			.S--Q--I--	285
B.FR.11.DEMB11FR001	--VH--T--R--K--			.K--	278
B.HT.05.05HT.129389	--VH--T--L--			.X--	278
F1.JP.12.DEM12JP001	--K--A--T--L--			.K--	292
B.KR.07.HP.18.07JHS10_3909	--VA--T--I--			.N--NR--	279
B.RU.11.11RU21n	--VP--Q--I--R--			.R--N--	283
B.SE.12.SE600057	--VSV--EQ--I--			.X--	283
B.TH.10.DEMB10TH002	--S--T--A--I--			.E--	278
B.US.13.RV.1	--P--R--S--I--M--L--			.K--	283
B.ZA.09.DEMB09ZA022	--VS--R--S--I--M--L--			.N--	279
C.BR.09.DEMC09BR003	--P--K--K--M--L--			.N--R--R--	277
C.BW.00.00BW003.1	--M--Q--K--M--L--			.TA--D--R--	278
C.CN.10.YNFL19	--P--K--K--M--L--			.TA--E--R--	281
C.CY.09.CY260	--XX--K--S--M--L--			.XD--E--X--	279
C.ES.14.ARP1198	--P--K--K--I--MM--L--			.T--N--T--	282
C.ET.02.02ET.268	--H--K--K--M--L--			.T--N--T--	273
C.IN.09.T125.2139	--P--K--K--M--L--			.T--E--T--	281
C.KE.05.05KE369195V4	--P--K--K--S--M--L--			.I--E--T--	274
C.MW.09.703010256.CH256.w96	--K--S--K--M--L--			.I--E--T--	275
C.SE.13.SE600311	--K--V--K--S--M--L--			.S--	277
C.TZ.08.707010457.CH457.w8	--V--K--K--S--M--L--			.TA--E--T--	290
H.US.11.17784.468	--K--K--K--M--L--			.TA--ED--T--	275
C.YE.02.02YE511	--E--V--V--KR--M--L--			.TA--ED--T--	289
C.ZA.12.DEMC12ZA096	--EE--P--K--K--L--			.TA--E--T--	282
C.ZM.11.DEMC11ZM006	--E--P--K--K--L--			.TA--D--T--	280
D.CM.10.DEMD10CM009	--P--K--K--M--L--			.T--D--R--I--	284
D.CY.06.CY163	--G--Y--K--M--L--			.T--D--R--I--	278
D.KE.11.DEMD11KE003	--G--Y--K--M--L--			.T--D--R--I--	278
D.KR.04.04KR048	--K--N--M--M--L--			.T--D--R--I--	278
D.SN.90.SE365	--K--N--M--M--L--			.T--D--R--I--	284
D.TZ.01.A280	--K--H--SVD--I--			.V--T--R--E--	280
D.UG.10.DEMD10UG004	--K--H--SVD--I--			.V--T--R--E--	278
D.UG.11.DEMD11UG003	--E--V--V--PL--L--			.T--D--R--I--	280
D.YE.02.02YE516	--E--V--V--PL--L--			.T--D--R--I--	278
D.ZA.90.R1	--T--D--Y--T--I--M--			.T--E--	283
F1.A0.06.A0.06.AMG32	--K--N--D--TC--F--I--M--H--K--I--L--S--I--			.X--	279
F1.AR.02.ARE933	--K--N--D--TC--F--I--M--H--K--I--L--S--I--			.X--	285
F1.BR.07.07BR844	--K--N--D--TC--F--I--M--H--K--I--L--S--I--			.X--	279
F1.BR.10.10BR.PE107	--K--SV--Y--M--L--			.N--R--	281
F1.BR.10.10BR.RJ015	--K--N--I--D--T--M--L--			.N--R--	282
F1.CY.08.CY222	--K--T--D--R--M--L--			.N--R--	279
F1.ES.02.ES.X845.4	--K--T--D--R--M--L--			.N--R--	281
F1.ES.11.VA0053.nfl	--K--EN--Y--T--M--L--			.X--I--D--R--	279
F1.RO.96.BCI.R07	--K--T--D--R--M--L--			.X--I--D--R--	279
F1.RU.08.D88.845	--K--V--K--M--L--			.T--AD--	279
F2.CM.02.02CM.6016BBY	--VP--Q--I--M--L--			.S--C--	279
F2.CM.10.DEMF210CM001	--VH--Q--I--M--L--			.S--C--	284
F2.CM.10.DEMF210CM007	--VH--Q--I--M--L--			.S--C--	278
G.CM.07.920.49	--K--E--P--E--K--I--M--L--			.R--	278
G.CM.10.DEMG10CM008	--K--E--P--E--K--I--M--L--			.R--	278
G.CM.10.DEURF10CM020	--K--T--E--N--T--I--M--L--			.Q--	278
G.CN.08.GX.2084.08	--K--VP--E--K--I--M--L--			.R--	283
G.ES.09.X2634.2	--K--P--E--K--I--M--L--			.R--	278
G.ES.14.ARP1201	--K--P--E--K--I--M--L--			.R--	278
G.GH.03.03GH175G	--K--P--E--K--I--M--L--			.R--	279
G.KE.09.DEMG09KE001	--K--P--E--K--I--M--L--			.R--	277
G.NG.09.09NG.SC62	--PLK--E--K--R--I--M--L--			.K--E--T--M--E--	278
G.PT.x.PT3306	--E--VA--F--K--I--M--L--			.K--E--T--M--E--	278
H.BE.93.VI991	--E--VA--F--K--I--M--L--			.K--E--T--M--E--	277
H.BE.93.VI997	--E--VA--F--K--I--M--L--			.K--E--T--M--E--	277
H.CF.90.056	--E--VA--F--K--I--M--L--			.K--E--T--M--E--	277
H.GB.00.00GBAC4001	--E--VA--F--K--I--M--L--			.K--E--T--M--E--	277
J.CD.97.J.97CD.KTB147	--ND--E--K--I--M--L--			.R--E--T--E--	276
J.SE.93.SE9280.7887	--NEVP--E--K--I--M--L--			.T--O--A--L--E--RV--	277
J.SE.94.SE9173.7022	--NEVP--E--K--I--M--L--			.T--O--A--L--E--RV--	277
K.CD.97.97ZR.E0TB11	--VCM--Q--I--M--L--			.I--P--	279
K.CM.96.96CM.MP535	--V--Q--I--M--L--			.I--P--	279
01.AE.AF.07.569M	--T--K--K--M--L--			.N--T--	279
01.AE.CM.11.1156.26	--T--K--K--M--L--			.N--T--	278
01.AE.CN.12.DE00112CN011	--T--K--K--M--L--			.N--T--	277
01.AE.HK.04.HK001	--T--K--K--M--L--			.N--T--	276
01.AE.IR.10.10IR.THR48F	--T--K--K--M--L--			.N--T--	271
01.AE.JP.11.DE00111JP003	--T--K--K--M--L--			.N--T--	278
01.AE.SE.11.SE601018	--T--K--K--M--L--			.N--T--	279
01.AE.TH.10.DE00110TH001	--T--K--K--M--L--			.N--T--	279
01.AE.TH.90.CM240	--T--K--K--M--L--			.N--T--	284
01.AE.US.05.306163.FL	--T--K--K--M--L--			.N--T--	277
02.AG.CM.10.DE00210CM013	--M--K--K--M--L--			.R--I--	279
02.AG.ES.06.P1423	--K--K--K--M--L--			.R--I--	279
02.AG.GW.05.CC.0048	--K--K--K--M--L--			.R--I--	277

	protease end_p66, p51 RT start	M41L	K70R D67N	D110 catalytic site	
B. FR. 83. HXB2	KVRYQYDQILIEICGHKAIGTVLVGPTVNIIGRNL	LTIQIGTCLNFPISPIETVPVKLPGMDGPKVQWPLTEEEKILAI	VEICTEMEKEGKISKIGPENPYNTPVFAIKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGL	KKKKSVTVLVDVGDYFVSLPDEDF	279
02 AG. KR. 12. 12MHR9	-----	-----	-----	-----	283
02 AG. LR. x. POC44951	-----	-----	-----	-----	279
02 AG. NG. 09. 09NG SC61	-----	-----	-----	-----	279
02 AG. NG. x. IBNG	-----	-----	-----	-----	279
02 AG. SE. 11. 5602024	-----	-----	-----	-----	279
02 AG. SN. 98. 98SE MP1211	-----	-----	-----	-----	279
02 AG. US. 06. 502 2696 FL01	-----	-----	-----	-----	279
03 AB. RU. 97. KAL153 2	-----	-----	-----	-----	279
04 cpx. CY. 94. 94CY032 3	-----	-----	-----	-----	279
05 DF. BE. x. VT110	-----	-----	-----	-----	278
06 cpx. AU. 96. BFP90	-----	-----	-----	-----	279
07 BC. CN. 98. 98CN009	-----	-----	-----	-----	281
08 BC. CN. 97. 97CNX 6F	-----	-----	-----	-----	279
09 cpx. GH. 96. 96GH2911	-----	-----	-----	-----	278
10 CD. TZ. 96. 96TZ BF061	-----	-----	-----	-----	278
11 cpx. CM. 95. 95CM 1816	-----	-----	-----	-----	277
12 BF. AR. 99. ARMA159	-----	-----	-----	-----	279
13 cpx. CM. 96. 96CM 1849	-----	-----	-----	-----	279
14 BG. ES. 05. X1870	-----	-----	-----	-----	278
15 01B. TH. 99. 99TH MU2079	-----	-----	-----	-----	286
16 AZD. KR. 97. 97KR000	-----	-----	-----	-----	278
17 BF. AR. 99. ARMA038	-----	-----	-----	-----	278
18 cpx. CU. 99. CU76	-----	-----	-----	-----	279
19 cpx. CU. 99. CU7	-----	-----	-----	-----	278
20 BG. CU. 99. Cu103	-----	-----	-----	-----	278
21 AZD. KR. 99. KR2003	-----	-----	-----	-----	275
22 01A1. CM. 01. 01CM 0001BBY	-----	-----	-----	-----	279
23 BG. CU. 03. CB118	-----	-----	-----	-----	275
24 BG. ES. 08. X2456 2	-----	-----	-----	-----	275
25 cpx. CM. 02. 1918LE	-----	-----	-----	-----	275
26 AU. CD. 02. 02CD MBT047	-----	-----	-----	-----	278
27 cpx. FR. 04. 04FR K25	-----	-----	-----	-----	277
28 BF. BR. 99. BREPM12609	-----	-----	-----	-----	279
29 BF. BR. 01. BREPM16704	-----	-----	-----	-----	279
31 BC. BR. 04. 04BR142	-----	-----	-----	-----	279
32 06A1. EE. 01. EE0369	-----	-----	-----	-----	285
33 01B. ID. 07. JKT189 C	-----	-----	-----	-----	279
34 01B. TH. 99. 01R2478P	-----	-----	-----	-----	279
35 AD. AF. 07. 169H	-----	-----	-----	-----	278
36 cpx. CM. 00. 00CMNYU830	-----	-----	-----	-----	279
37 cpx. CM. 00. 00CMNYU26	-----	-----	-----	-----	277
38 BF1. UY. 03. UY03 3389	-----	-----	-----	-----	283
39 BF. BR. 04. 04BR3179	-----	-----	-----	-----	282
40 BF. BR. 05. 05BRR1055	-----	-----	-----	-----	279
42 BF. LU. 06. LuBF 18 06	-----	-----	-----	-----	283
43 02G. SA. 03. J11223	-----	-----	-----	-----	278
44 BF. CL. 08. CH80	-----	-----	-----	-----	279
45 cpx. FR. 04. 04FR AUK	-----	-----	-----	-----	279
46 BF. BR. 07. 07BR FPS625	-----	-----	-----	-----	279
47 BF. ES. 08. P1942	-----	-----	-----	-----	278
48 01B. MY. 07. 07MYKT021	-----	-----	-----	-----	279
49 cpx. GM. 03. N26677	-----	-----	-----	-----	277
50 A1D. GB. 10. 12792	-----	-----	-----	-----	279
51 01B. SG. 10. 1156 HM021	-----	-----	-----	-----	279
52 01B. MY. 03. 03MYKL018 1	-----	-----	-----	-----	279
53 01B. MY. 11. 11FIR164	-----	-----	-----	-----	279
54 01B. MY. 09. 09MYSB023	-----	-----	-----	-----	279
55 01B. CN. 10. HNC5102056	-----	-----	-----	-----	279
56 cpx. FR. 10. URF5 patient A	-----	-----	-----	-----	279
57 BC. CN. 09. 09YNLX195g	-----	-----	-----	-----	281
58 01B. MY. 09. 09MYPR37	-----	-----	-----	-----	275
59 01B. CN. 09. 09LNA423	-----	-----	-----	-----	279
60 BC. IT. 11. BAV499	-----	-----	-----	-----	276
61 BC. CN. 10. J1100010	-----	-----	-----	-----	274
62 BC. CN. 10. YNFL13	-----	-----	-----	-----	275
63 02A1. RU. 10. 10RU6637	-----	-----	-----	-----	279
64 BC. CN. 09. YNFL31	-----	-----	-----	-----	279
65 cpx. CN. 10. YNFL01	-----	-----	-----	-----	275
67 01B. CN. 11. ANHUI HF115	-----	-----	-----	-----	277
68 01B. CN. 11. ANHUI WH73	-----	-----	-----	-----	277
69 01B. JP. 05. 05JPMYC113SP420	-----	-----	-----	-----	277
70 BF1. BR. 10. 10BR PE004	-----	-----	-----	-----	281
71 BF1. BR. 10. 10BR PE008	-----	-----	-----	-----	279
72 BF1. BR. 10. 10BR MG002	-----	-----	-----	-----	280
74 01B. MY. 10. 10MYPR268	-----	-----	-----	-----	279
0. BF. 87. ANT70	-----	-----	-----	-----	275
0. CM. 98. 98CMA104	-----	-----	-----	-----	275
0. CM. 98. 98CMABB141	-----	-----	-----	-----	275
0. CM. 98. 98CMABB212	-----	-----	-----	-----	275
0. CM. 98. 98CMU5337	-----	-----	-----	-----	275
0. CM. 99. 99CMU4122	-----	-----	-----	-----	275
0. FR. 92. VAU	-----	-----	-----	-----	275
0. GA. 11. 11Gab6352	-----	-----	-----	-----	275
0. SN. 99. 99SE MP1299	-----	-----	-----	-----	275
0. US. 10. LTNP	-----	-----	-----	-----	275
N. CM. 02. DJ00131	-----	-----	-----	-----	284
N. CM. 02. SJGddd	-----	-----	-----	-----	281
N. CM. 04. 04CM 1015 04	-----	-----	-----	-----	286
N. CM. 06. U14296	-----	-----	-----	-----	286
N. FR. 11. N1 FR 2011	-----	-----	-----	-----	286
P. CM. 06. U14788	-----	-----	-----	-----	276
P. FR. 09. RBF168	-----	-----	-----	-----	275
CPZ. CD. 90. ANT	-----	-----	-----	-----	278
CPZ. CM. 05. SIVcpzLB7	-----	-----	-----	-----	282
CPZ. CM. 09. Ptt 09Cam155	-----	-----	-----	-----	278
CPZ. TZ. 06. TAN5	-----	-----	-----	-----	275
CPZ. US. 05. US Mar31yn	-----	-----	-----	-----	281
GOR. CM. 07. SIVgor2139 287	-----	-----	-----	-----	271
GOR. CM. 12. SIVgor B01D2	-----	-----	-----	-----	275
GOR. CM. 13. SIVgor BPID15	-----	-----	-----	-----	274

	polymerase motif	K219Q	T215Y	
B.FR.83.HXB2	SK E	A A K F		Q
A1.CD.97.97DC KCC2	T			
A1.CM.08.886 24	C			
A1.CY.08.CY236	C			
A1.IN.09.NARI FLS TVC19 1	Y A SK E I	K A A K F		Q
A1.KE.11.DEMA11KE002	D			
A1.RU.11.DRU6950	LK	V A S F		M D
A1.RW.11.DEMA11RW002	V	A A S F		
A1.SN.01.DDI579	Y	E I	N	
A1.UG.11.DEMA11UG009	T	SK E I	V	KK S F F
A1.ZA.04.5012344 T10 A1	T	SK E I		
A2.CD.97.97CDKTB48	D	AR EM	D	A A S F F
A2.CM.01.01CM 1445MV	I	TK E I		
A2.CY.94.94CY017 41	T	SK TELI	S	V A K FY
B.BR.10.10BR RJ032	T	EK		
B.CA.07.502 1191 03	A	Y	E	K F
B.CH.08.M2 0803101 NFLG8	T	A	E V	A K Y L I E
B.CN.12.DEMB12CN006	M	R	D	L A A E K I
B.CU.14.14CU065	C	D	L	A A E K M
B.ES.14.ARP1195	C	R	L	A A E K M
B.FR.11.DEMB11FR001				
B.HT.05.05HT 129389	L	C	X	Y E X F F
B.JP.12.DEMB12JP001	L	C	Y	E X R G F
B.KR.07.HP 18 07JHS10 3909				
B.RU.11.11RU21n	D	E		
B.SE.12.SE600057				
B.TH.10.DEMB10TH002	V	H	D	L I K F M
B.US.13.RV 1	V	H		
B.ZA.09.DEMB09ZA022				
C.BR.07.DENC07BR003	R	A	E	A E F K F
C.BW.00.00BW003 1	V			
C.CN.10.YNFL19	A	I	A	E A V K F
C.CY.09.CY260	A	C	I	A A E K F
C.ES.14.ARP1198				
C.ET.02.02ET 268	T	R	A	EL A K X K F
C.IN.09.T125 2139				
C.KE.05.05KE369195V4	V			
C.MW.09.703010256 CH256.w96	T	C		
C.SE.13.SE600311				
C.TZ.08.707010457 CH457.w8				
C.US.11.17784 468				
C.YE.02.02YE511	F			
C.ZA.12.DEMC12ZA096	Y	AK	A	E A D K F
C.ZM.11.DEMC11ZM006				
D.CM.10.DEMD10CM009	T	C	K	EL E K F
D.CY.06.CY163	C	K	K	EL E K F
D.KE.11.DEMD11KE003				
D.KR.04.04KR048	T	A	Q	NK E F F
D.SN.90.SE365				
D.TZ.01.A280				
D.UG.10.DEMD10UG004	A			
D.UG.11.DEMD11UG003				
D.YE.02.02YE516	T	C		
D.ZA.90.R1				
F1.A0.06.A0 06 ANG32	V	C	TR	E A A K F
F1.AR.02.ARE933	V	C	TR	M A K T F
F1.BR.07.07BR844	L	D	TK	M A D K F I E
F1.BR.10.10BR PE107				
F1.BR.10.10BR RJ015	T	C	D	T A V K S F F
F1.CY.08.CY222				
F1.ES.02.ES 4845 4	L	K		
F1.ES.11.VA0053 nFl	L	CX	D	AK GL E K F F
F1.R0.96.BCI R07	L	Y	AK	I L A V D K F
F1.RU.08.D88 845				
F2.CM.02.02CM 6016BBY	K	TK		
F2.CM.10.DEMF210CM001				
F2.CM.10.DEMF210CM007				
G.CM.07.920 49	V	C	R	TR EM A E F K F
G.CM.10.DEMG10CM008				
G.CM.10.DEURF10CM020	V	C	R	K EM A E F K F
G.CN.08.GX 2084 08				
G.ES.09.X2634 2	A	X	K	IK EM T A D K F
G.ES.14.ARP1201				
G.GH.03.03GH175G				
G.KE.09.DEMG09KE001	V	K	I	TK M A V E K F F
G.NG.09.09NG SC62				
G.PT.x.PT3306	L	G	TK	EL E A K S F
H.BE.93.VI991	T			
H.BF.93.VI997				
H.CF.90.056				
H.GB.00.00GBAC4001	T	N	A	E EM A K F
J.CD.97.J 97DC KTB147				
J.SE.93.SE9280 7887	C	K	ER	E E R K E K F
J.SE.94.SE9173 7022	C	K	ER	E E R K E K FY
K.CD.97.97ZR E0TB11				
K.CM.96.96CM MP535	V	H	CK	M L E F F
01 AE.AF.07.569M	T			
01 AE.CM.11.1156 26	T			
01 AE.CN.12.DEB0112CN011	T			
01 AE.HK.04.HK001				
01 AE.IR.10.10IR.THR48F	T	K		
01 AE.JP.11.DE00111JP003				
01 AE.SE.11.SE601038	V	A		
01 AE.TH.10.DEB0110TH001				
01 AE.TH.90.CM240				
01 AE.US.05.306163 FL				
02 AG.CM.10.DE00210CM013	V	A	A	AK E I A E K F F
02 AG.ES.06.P1423				
02 AG.GW.05.CC 0048	V	X	A	X TK E EK A E G K F

	polymerase motif										K219Q	T215Y
B. FR. 83. HXB2												
02 AG. KR. 12. 12MHR9												
02 AG. LR. x. POC44951												
02 AG. NG. 09. 09NG SC61												
02 AG. NG. x. IBNG												
02 AG. SE. 11. 5602024												
02 AG. SN. 98. 98SE MP1211												
02 AG. US. 06. 502 2696 FL01												
03 AB. RU. 97. KAL153 2												
04 cpx. CY. 94. 94CY032 3												
05 DF. BE. x. VT1116												
06 cpx. AU. 96. BFP90												
07 BC. CN. 98. 98CN09												
08 BC. CN. 97. 97CNGX 6F												
09 cpx. GH. 96. 96GH2911												
10 CD. TZ. 96. 96TZ BF061												
11 cpx. CM. 95. 95CM 1816												
12 BF. AR. 99. ARMA159												
13 cpx. CM. 96. 96CM 1849												
14 BG. ES. 05. X1870												
15 01B. TH. 99. 99TH MU2079												
16 AZD. KR. 97. 97KR06												
17 BF. AR. 99. ARMA038												
18 cpx. CU. 99. CU76												
19 cpx. CU. 99. CU7												
20 BG. CU. 99. Cu103												
21 AZD. KR. 99. KER2003												
22 01A1. KE. 01. 01CM 0001BBY												
23 BG. CU. 03. CB118												
24 BG. ES. 08. X2456 2												
25 cpx. CM. 02. 19181E												
26 AU. CD. 02. 02CD MBT047												
27 cpx. FR. 04. 04CD K25												
28 BF. BR. 99. BREPM12609												
29 BF. BR. 01. BREPM16704												
31 BC. BR. 04. 04BR142												
32 06A1. EE. 01. EE0369												
33 01B. ID. 07. JKT189 C												
34 01B. TH. 99. TH99R2478P												
35 AD. AF. 07. 169H												
36 cpx. CM. 00. 00CMNYU830												
37 cpx. CM. 00. 00CMNYU926												
38 BF1. UY. 03. UY03 3389												
39 BF. BR. 04. 04BRR3179												
40 BF. BR. 05. 05BRR1055												
42 BF. LU. 06. LuBF 18 06												
43 02G. SA. 03. J11223												
44 BF. CL. 08. CH80												
45 cpx. FR. 04. 04FR AUK												
46 BF. BR. 07. 07BR FPS625												
47 BF. ES. 08. P1942												
48 01B. MY. 07. 07MYKT021												
49 cpx. GM. 03. N26677												
50 A10. GB. 10. 12702												
51 01B. SG. 11. 11SG HM021												
52 01B. MY. 03. 03MYKL018 1												
53 01B. MY. 11. 11FIR164												
54 01B. MY. 09. 09MYSB023												
55 01B. CN. 10. HNC5102056												
56 cpx. FR. 10. URF5 patient A												
57 BC. CN. 09. 09YNLX19sq												
58 01B. MY. 09. 09MYPR37												
59 01B. CN. 09. 09LNA423												
60 BC. IT. 11. BAV499												
61 BC. CN. 10. J1100010												
62 BC. CN. 10. YNFL13												
63 02A1. RU. 10. 10RU6637												
64 BC. CN. 09. YNFL31												
65 cpx. CN. 10. YNFL01												
67 01B. CN. 11. ANHUI HF115												
68 01B. CN. 11. ANHUI WH73												
69 01B. JP. 05. 05JPMYC113SP420												
70 BF1. BR. 10. 10BR PE004												
71 BF1. BR. 10. 10BR PE008												
72 BF1. BR. 10. 10BR MG002												
74 01B. MY. 10. 10MYPR268												
0. BF. 87. ANT70												
0. CM. 98. 98CMA184												
0. CM. 98. 98CMABB141												
0. CM. 98. 98CMABB212												
0. CM. 98. 98CMU5337												
0. CM. 99. 99CMU4122												
0. FR. 92. VAI												
0. GA. 11. 11Gabb6352												
0. SN. 99. 99SE MP1299												
0. US. 10. LTNP												
N. CM. 02. DJ00131												
N. CM. 02. SJGddd												
N. CM. 04. 04CM 1015 04												
N. CM. 06. U14296												
N. FR. 11. N1 FR 2011												
P. CM. 06. U14788												
P. FR. 09. RBF168												
CPZ. CD. 90. ANT												
CPZ. CM. 05. SIVcpzLB7												
CPZ. CM. 09. Ptt 09Cam155												
CPZ. TZ. 06. TAN5												
CPZ. US. 05. US Mar3Lyn												
GOR. CM. 07. SIVgor2139 287												
GOR. CM. 12. SIVgor_B0ID2												
GOR. CM. 13. SIVgor_BPID15												

B. FR. 83. HXB2	VILVAVHVASGYIEAEVIPAETGQETAYFLLLKLAGRWPVKTIHTDNGSNFTGATVRAACIWAGKQKQEFGIPYNPQSGVGVSMNKLKIIQVRRDQAEHLKTAVQMAVFIHNFKRKGIGGYSAGERVVDIIATDIQTKELQKQTKIQNFRVYRDRNPLWKGP	AKL	956
A1. CD. 97. 97DC KCC2	-----I-----V-----STA-K-----S-Q-----E-----L-----R-----H-----M-----D-I-----	956	
A1. CM. 08. 886 24	-----L-----V-----VW-----STA-K-----Q-----I-----S-----R-----I-----M-----D-I-----	956	
A1. CY. 08. CY236	-----A-----V-----P-----S-A-K-----D-Q-----I-----S-----R-----I-----M-----D-I-----	953	
A1. IN. 09. NARI FLS TVC19 1	I-----I-----R-----V-----P-----S-A-K-----N-Q-----L-----E-----I-----N-----I-----D-I-----	953	
A1. KE. 11. DEMA11KE002	-----I-----R-----V-----P-----S-A-K-----N-Q-----L-----E-----I-----N-----I-----D-I-----	956	
A1. RU. 11. 11RU6950	-----I-----R-----V-----P-----S-A-K-----N-Q-----L-----E-----I-----N-----I-----D-I-----	956	
A1. RW. 11. DEMA11RW002	-----I-----R-----V-----P-----S-A-K-----N-Q-----L-----E-----I-----N-----I-----D-I-----	956	
A1. SN. 01. DDI579	-----I-----R-----V-----P-----S-A-K-----N-Q-----L-----E-----I-----N-----I-----D-I-----	955	
A1. UG. 11. DEMA110UG009	-----A-----V-----V-----G-----S-A-K-----DVO-----I-----L-----N-----K-----D-I-----	956	
A1. ZA. 04. 501 12344 T10 A1	-----A-----V-----V-----G-----S-A-K-----DVO-----I-----L-----N-----I-----D-I-----	953	
A2. CD. 97. 97CDKTB48	I-V-----I-----V-----P-----S-A-K-----D-----D-----I-----R-----I-----E-I-----	949	
A2. CM. 01. 01CM 1445MV	-----T-----D-----I-----V-----P-----S-A-K-----Q-----D-----I-----R-----I-----E-I-----	956	
A2. CY. 94. 94CY017 41	-----T-----D-----I-----V-----P-----S-A-K-----Q-----D-----I-----R-----I-----E-I-----	956	
B. BR. 10. 10BR RJ032	I-----I-----R-----V-----P-----S-K-----I-----I-----I-----D-----D-I-----	955	
B. CA. 07. 502 1191 03	-----I-----R-----V-----P-----S-K-----I-----I-----I-----D-----D-I-----	956	
B. CH. 08. M2 0803101 NFLG8	I-----I-----R-----K-----P-----S-N-K-----N-----E-----T-----N-----V-----D-----R-----	960	
B. CN. 12. DEMB12CM006	-----I-----R-----K-----P-----S-N-K-----N-----E-----T-----N-----V-----D-----R-----	961	
B. CU. 14. 14CU005	-----I-----R-----K-----P-----S-N-K-----N-----E-----T-----N-----V-----D-----R-----	956	
B. ES. 14. ARP1195	-----V-----A-----NV-----P-----STA-K-----Y-----N-----E-----I-----S-----F-----D-I-----	962	
B. FR. 11. DEMB11FR001	I-----I-----R-----V-----P-----S-K-----D-----R-----E-----I-----S-----I-----D-I-----	955	
B. HT. 05. 05HT 129389	I-----I-----R-----V-----P-----S-K-----D-----R-----E-----I-----S-----I-----D-I-----	955	
F1. JP. 12. DEMB12JP001	I-----I-----R-----V-----P-----S-K-----D-----R-----E-----I-----S-----I-----D-I-----	959	
B. BR. 07. HP 18 07H510 3909	I-----I-----T-----V-----STA-K-----V-----I-----M-----N-----O-----R-----Q-----N-D-I-----	955	
B. RU. 11. 11RU21n	I-----I-----T-----V-----STA-K-----V-----I-----M-----N-----O-----R-----Q-----N-D-I-----	960	
B. SE. 12. SE600057	I-----I-----V-----P-----SSADK-----L-----NO-----Q-----I-----Q-----H-----D-I-----	955	
B. TH. 10. DEMB10TH002	-----V-----N-----I-----V-----P-----S-A-K-----L-----NO-----Q-----I-----Q-----H-----D-I-----	960	
B. US. 13. RV 1	-----V-----N-----I-----V-----P-----S-A-K-----L-----NO-----Q-----I-----Q-----H-----D-I-----	956	
B. ZA. 09. DEMB09ZA022	-----V-----N-----I-----V-----P-----S-A-K-----L-----NO-----Q-----I-----Q-----H-----D-I-----	954	
C. BR. 07. DEMC07BR003	I-----I-----V-----I-----A-----P-----ISN-K-----Q-----N-----I-----I-----D-I-----	952	
C. BW. 00. 00BW003 1	I-----I-----V-----I-----A-----P-----ISN-K-----Q-----N-----I-----I-----D-I-----	958	
C. CN. 10. YNFL19	I-----I-----M-----V-----I-----V-----S-A-K-----L-----I-----R-----I-----D-I-----	956	
C. CY. 09. CY260	I-----I-----M-----V-----I-----V-----S-A-K-----L-----I-----R-----I-----D-I-----	959	
C. ES. 14. ARP1198	I-----I-----M-----YI-----V-----P-----S-A-K-----Q-----I-----M-----R-----LQ-----X-----D-I-----	950	
C. ET. 02. 02ET 268	I-----I-----M-----YI-----V-----P-----S-A-K-----Q-----I-----M-----R-----LQ-----X-----D-I-----	958	
C. IN. 09. T125 2139	I-----I-----M-----YI-----V-----P-----S-A-K-----Q-----I-----M-----R-----LQ-----X-----D-I-----	953	
C. KE. 05. 05KE369195V4	I-----I-----M-----YI-----V-----P-----S-A-K-----Q-----I-----M-----R-----LQ-----X-----D-I-----	952	
C. MW. 09. 703010256 CH256.w96	I-----I-----T-----YI-----V-----P-----S-A-K-----Q-----I-----M-----R-----LQ-----X-----D-I-----	954	
C. SE. 13. SE600311	I-----I-----T-----YI-----V-----P-----S-A-K-----Q-----I-----M-----R-----LQ-----X-----D-I-----	967	
C. TZ. 08. 707010457 CH457.w8	I-----I-----E-----YI-----V-----R-----SN-K-----E-----E-----R-----D-----T-----I-----S-----N-----L-----D-I-----	952	
C. US. 11. 11US 468	I-----I-----E-----YI-----V-----R-----SN-K-----E-----E-----R-----D-----T-----I-----S-----N-----L-----D-I-----	956	
C. YE. 02. 02YE511	I-----I-----E-----YI-----V-----R-----SN-K-----E-----E-----R-----D-----T-----I-----S-----N-----L-----D-I-----	950	
C. ZA. 12. DEMC12ZA096	I-----I-----M-----YI-----V-----SNA-K-----Q-----I-----I-----I-----D-I-----	959	
C. ZM. 11. DEMC11ZM006	I-----I-----M-----YI-----V-----SNA-K-----Q-----I-----I-----I-----D-I-----	957	
D. CM. 10. DEMD10CM009	I-----I-----K-----I-----V-----V-----S-K-----V-----R-----L-----I-----R-----D-I-----	961	
D. CY. 06. CY163	I-----I-----K-----I-----V-----V-----S-K-----V-----R-----L-----I-----R-----D-I-----	955	
D. KE. 11. DEMD11KE003	I-----I-----K-----I-----V-----V-----S-K-----V-----R-----L-----I-----R-----D-I-----	955	
D. KR. 04. 04KR88	I-----I-----K-----I-----V-----V-----S-K-----V-----R-----L-----I-----R-----D-I-----	953	
D. SN. 90. SE365	I-----I-----K-----I-----V-----V-----S-K-----V-----R-----L-----I-----R-----D-I-----	961	
D. TZ. 01. A280	I-----I-----K-----I-----V-----V-----S-K-----V-----R-----L-----I-----R-----D-I-----	957	
D. UG. 10. DEMD10UG004	I-----I-----K-----I-----V-----R-----STA-K-----E-----R-----I-----S-L-----T-----V-----D-I-----	955	
D. UG. 11. DEMD11UG003	I-----I-----K-----I-----V-----R-----STA-K-----E-----R-----I-----S-L-----T-----V-----D-I-----	957	
D. YE. 02. 02YE516	I-----I-----K-----I-----V-----R-----STA-K-----E-----R-----I-----S-L-----T-----V-----D-I-----	957	
D. ZA. 90. R1	I-----I-----K-----I-----V-----R-----STA-K-----E-----R-----I-----S-L-----T-----V-----D-I-----	955	
F1. AO. 06. AO 06 ANG32	-M-----L-----I-----I-----S-A-K-----S-----Q-----S-----R-----I-----N-D-V-----	960	
F1. AR. 02. ARE933	I-----L-----A-----P-----S-A-K-----S-----Q-----S-----R-----I-----N-D-V-----	956	
F1. BR. 07. 07BR844	I-----L-----A-----P-----S-A-K-----S-----Q-----S-----R-----I-----N-D-V-----	962	
F1. BR. 10. 10BR PE107	I-----L-----A-----P-----S-A-K-----S-----Q-----S-----R-----I-----N-D-V-----	956	
F1. BR. 10. 10BR RJ015	I-----L-----A-----P-----S-A-K-----S-----Q-----S-----R-----I-----N-D-V-----	959	
F1. CY. 08. CY222	I-----L-----A-----P-----S-A-K-----S-----Q-----S-----R-----I-----N-D-V-----	955	
F1. ES. 02. ES X845 4	I-----L-----A-----P-----S-A-K-----S-----Q-----S-----R-----I-----N-D-V-----	958	
F1. ES. 11. VA0053 nfl	I-----L-----A-----P-----S-A-K-----S-----Q-----S-----R-----I-----N-D-V-----	956	
F1. RO. 96. BCI R07	I-----L-----A-----P-----S-A-K-----S-----Q-----S-----R-----I-----N-D-V-----	956	
F1. RU. 08. D88 845	I-----L-----A-----P-----S-A-K-----S-----Q-----S-----R-----I-----N-D-V-----	956	
F2. CM. 02. 02CM 6016BBY	I-----I-----I-----I-----S-A-K-----Q-----D-----C-----S-----L-----A-----S-----K-----F-----D-I-----	956	
F2. CM. 10. DEMF210CM001	I-----I-----I-----I-----S-A-K-----Q-----D-----C-----S-----L-----A-----S-----K-----F-----D-I-----	961	
F2. CM. 10. DEMF210CM007	I-----I-----I-----I-----S-A-K-----Q-----D-----C-----S-----L-----A-----S-----K-----F-----D-I-----	955	
G. CM. 07. 920 49	I-----I-----M-----P-----A-----I-----V-----R-----SNA-K-----N-T-----E-----I-----PIN-----R-----I-----S-----Q-----D-I-----	955	
G. CM. 10. DEMG10CM008	I-----I-----M-----P-----A-----I-----V-----R-----SNA-K-----N-T-----E-----I-----PIN-----R-----I-----S-----Q-----D-I-----	955	
G. CM. 10. DEURF10CM020	I-----I-----M-----P-----A-----I-----V-----R-----SNA-K-----N-T-----E-----I-----PIN-----R-----I-----S-----Q-----D-I-----	955	
G. CN. 08. GX 2084 08	I-----I-----M-----P-----A-----I-----V-----R-----SNA-K-----N-T-----E-----I-----PIN-----R-----I-----S-----Q-----D-I-----	960	
G. ES. 09. X2634 2	I-----I-----M-----P-----A-----I-----V-----R-----SNA-K-----N-T-----E-----I-----PIN-----R-----I-----S-----Q-----D-I-----	960	
G. ES. 14. ARP1201	I-----I-----M-----P-----A-----I-----V-----R-----SNA-K-----N-T-----E-----I-----PIN-----R-----I-----S-----Q-----D-I-----	955	
G. GH. 03. 03GH175G	I-----I-----M-----P-----A-----I-----V-----R-----SNA-K-----N-T-----E-----I-----PIN-----R-----I-----S-----Q-----D-I-----	956	
G. KE. 09. DEMG09KE001	I-----I-----M-----P-----A-----I-----V-----R-----SNA-K-----N-T-----E-----I-----PIN-----R-----I-----S-----Q-----D-I-----	954	
G. NG. 09. 09NG SC62	I-----I-----M-----P-----A-----I-----V-----R-----SNA-K-----N-T-----E-----I-----PIN-----R-----I-----S-----Q-----D-I-----	955	
G. PT. x. PT3306	I-----I-----M-----P-----A-----I-----V-----R-----SNA-K-----N-T-----E-----I-----PIN-----R-----I-----S-----Q-----D-I-----	955	
H. BE. 93. VI991	I-----I-----XP-----I-----M-----T-----S-A-K-----D-----H-----D-----R-----I-----P-----S-----Q-----K-----D-I-----	954	
H. BE. 93. VI997	I-----I-----XP-----I-----M-----T-----S-A-K-----D-----H-----D-----R-----I-----P-----S-----Q-----K-----D-I-----	953	
H. CF. 90. 056	I-----I-----XP-----I-----M-----T-----S-A-K-----D-----H-----D-----R-----I-----P-----S-----Q-----K-----D-I-----	955	
H. GB. 00. 00GBAC4001	I-----I-----XP-----I-----M-----T-----S-A-K-----D-----H-----D-----R-----I-----P-----S-----Q-----K-----D-I-----	954	
J. CD. 97. J 97DC KTB147	I-----I-----I-----F-----I-----RV-----S-A-K-----V-----E-----I-----S-----R-----I-----D-I-----	953	
J. SE. 93. SE9280 7887	I-----I-----I-----F-----I-----RV-----S-A-K-----V-----E-----I-----S-----R-----I-----D-I-----	954	
J. SE. 94. SE9173 7022	I-----I-----I-----F-----I-----RV-----S-A-K-----V-----E-----I-----S-----R-----I-----D-I-----	954	
K. CD. 97. 97ZR E0TB11	I-----I-----I-----RV-----S-V-K-----D-----E-----I-----I-----D-I-----	956	
K. CM. 96. 96CM MP535	I-----I-----I-----RV-----S-V-K-----D-----E-----I-----I-----D-I-----	956	
01 AE. AF. 07. 569M	I-----I-----M-----S-----V-----RV-----S-A-K-----NVR-----E-----I-----M-----IN-----D-I-----	956	
01 AE. CM. 11. 116 26	I-----I-----M-----S-----V-----RV-----S-A-K-----NVR-----E-----I-----M-----IN-----D-I-----	954	
01 AE. CN. 12. DE0012CN011	I-----I-----M-----S-----V-----RV-----S-A-K-----NVR-----E-----I-----M-----IN-----D-I-----	954	
01 AE. HK. 04. HK001	I-----I-----M-----S-----V-----RV-----S-A-K-----NVR-----E-----I-----M-----IN-----D-I-----	953	
01 AE. IR. 10. 10IR. THR48F	I-----I-----M-----S-----V-----RV-----S-A-K-----NVR-----E-----I-----M-----IN-----D-I-----	948	
01 AE. JP. 11. DE0011JP003	I-----I-----M-----S-----V-----RV-----S-A-K-----NVR-----E-----I-----M-----IN-----D-I-----	954	
01 AE. SE. 11. SE601038	I-----I-----M-----S-----V-----RV-----S-A-K-----NVR-----E-----I-----M-----IN-----D-I-----	954	
01 AE. TH. 10. DE00110TH001	I-----I-----M-----S-----V-----RV-----S-A-K-----NVR-----E-----I-----M-----IN-----D-I-----	956	
01 AE. TH. 90. CM240	I-----I-----M-----S-----V-----RV-----S-A-K-----NVR-----E-----I-----M-----IN-----D-I-----	961	
01 AE. US. 05. 306163 FL	I-----I-----M-----S-----V-----RV-----S-A-K-----NVR-----E-----I-----M-----IN-----D-I-----	954	
02 AG. CM. 10. DE00210CM013	I-----I-----M-----S-----V-----RV-----S-A-K-----NVR-----E-----I-----M-----IN-----D-I-----	956	
02 AG. ES. 06. P1423	I-----I-----M-----S-----V-----RV-----S-A-K-----NVR-----E-----I-----M-----IN-----D-I-----	955	
02 AG. GW. 05. CC 0048	I-----I-----M-----S-----V-----RV-----S-A-K-----NVR-----E-----I-----M-----IN-----D-I-----	954	

Vif start

MENRWQVMIVWQVDMRRTKWSLVKHHMYVSGKARGVFRYHHYSEPHRISSEVHPL. .GD. AR. LVITTYWGLHTGERDWHLGQVSIIEWPKKRYSTQVDPPELADQLIHLHYDFCFSDSAIRKALLGHVTSRPRCEYQAGH. N. KVGSLQYLALAAALIT.PKKI 159

B. FR. 83. HXB2
A1. CD. 97. 97CD KCC2
A1. CM. 08. 886 244
A1. CY. 08. CY236
A1. IN. 09. NARI FLS TVC19 1
A1. KE. 11. DEM111KE002
A1. RU. 11. DEUR6950
A1. RW. 11. DEMA111RW002
A1. SN. 01. DDI579
A1. UG. 11. DEMA110UG009
A1. ZA. 04. 501 15344 T10 A1
A2. CD. 97. 97CDKTB48
A2. CM. 01. 01CM 1445MV
A2. CY. 94. 94CY017 41
B. BR. 10. 10BR RJ032
B. CA. 07. 502 1191 03
B. CH. 08. M2 0803101 NFLG8
B. CN. 12. DEMB12CN006
B. CU. 14. 14CU065
B. ES. 14. ARP1195
B. FR. 11. DEMB11FR001
B. HT. 05. 05HT 129389
B. JP. 12. DEMB12JP001
B. KR. 07. HP 07H510 3909
B. RU. 11. 11RUZ1n
B. SE. 12. SE600057
B. TH. 10. DEMB10TH002
B. US. 13. RV 1
B. ZA. 09. DEM089ZA022
C. BR. 07. DEMC07BR003
C. BU. 00. 00BU003 1
C. CN. 10. YNFI19
C. CY. 09. CY260
C. ES. 14. ARP1198
C. ET. 02. 02ET 268
C. IN. 09. TI25 2139
C. KE. 05. 05KE369195V4
C. MW. 09. 703010256 CH256.w96
C. SE. 13. SE600311
C. TZ. 08. 707010457 CH457.w8
C. US. 11. 11US 468
C. YE. 02. 02YE511
C. ZA. 12. DEMC12ZA006
C. ZM. 11. DEMC11ZM006
D. CM. 10. DEMD10CM009
D. CY. 06. CY163
D. KE. 11. DEMD11KE003
D. KR. 04. 04KR048
D. SN. 90. SE365
D. TZ. 01. A280
D. UG. 10. DEMD10UG004
D. UG. 11. DEMD11UG003
D. YE. 02. 02YE516
D. ZA. 90. R1
F1. AO. 06. A0 06 ANG32
F1. AR. 02. ARE933
F1. BR. 07. 07BR844
F1. BR. 10. 10BR PE107
F1. BR. 10. 10BR RJ015
F1. CY. 08. CY22 1
F1. ES. 02. ES 484 4
F1. ES. 11. VA0053 nFl
F1. RO. 96. BCI R07
F1. RU. 08. D88 845
F2. CM. 02. 02CM 6016BBY
F2. CM. 10. DEMF210CM001
F2. CM. 10. DEMF210CM007
G. CM. 07. 920 49
G. CM. 10. DEMG10CM008
G. CM. 10. DEURF10CM020
G. CN. 08. GX 2084 08
G. ES. 09. X2634 2
G. ES. 14. ARP1201
G. GH. 03. 03GH175G
G. KE. 09. DEMG09KE001
G. NG. 09. 09NG SC62
G. PT. x. PT3306
H. BE. 93. V1991
H. BE. 93. V1997
H. CF. 90. 056
H. GB. 00. 00GBAC4001
J. CD. 97. J 97DC KTB147
J. SE. 93. SE9280 7887
J. SE. 94. SE9173 7022
K. CD. 97. 97ZR E0TB11
K. CM. 96. 96CM MP535
01 AE. AF. 07. 569M
01 AE. CM. 11. 11 26 26
01 AE. CN. 12. DE00112CN011
01 AE. HK. 04. HK001
01 AE. IR. 10. 10IR. THR48F
01 AE. JP. 11. DE00111JP003
01 AE. SE. 11. SE601018
01 AF. TH. 10. DE00110TH001
01 AE. TH. 90. CM240
01 AE. US. 05. 306163 FL
02 AG. CM. 10. DE00210CM013
02 AG. ES. 06. P1423
02 AG. GW. 05. CC 0048

A N H K I L R K VK G H KLR I D H T E I R K K V TRA 159
N K K K L D R KV N Q K H LRK I D QV S A T H V K V TRT 159
K H KR KN R F R KV R T VK Q K Q H L I DM H N I QV S D H R VA T 159
H Y R R K N V F R V XX EVK X K H KLRG I D QV S T H K V T 159
H R E R G V E I V K Q K Q H QG I D V H K E IV GR N P F K V T 159
H I R KD V K CRN E T IIVK KE Q H L S I DM H N I QV S LS K V VA ART 159
N I K KD DCR A M VR P Q H O I D H N T R I O R T H R VA S T 159
N K K F R KV K I VR N O Q H LR I D H N T I QV S T H R VA PR 159
N I K KN F RN KVA K I VR N O Q H LR I D H N T I QV S T H R VA PR 159
N I R K V F R KV KE I VR P K H QG I D H H E R I E L K H R VA STRT 159
N I R K V R N E IIVR I K H QN H I D H H E R I E L K H R VA STRT 159
N Y IHX K KE G TN SK VI R R D H H A N I R R D Q T L L 159
I K N D K K A V NI R R D H H A N I R R D Q T L L 159
N IHT K I K TN K V A ERK I G DS Q H E N I NR R E V TA T T 159
N R K K VI S G I DS Q H E N I NR R E V TA T T 159
NA I K KR V D K V I R R D H H A N I R R D Q T L L 159
N I Y IHT K N V N V DK V P N E R D TNH H Q I YR S FP I V K X 159
K K I NN E N V ER T G KN I R S V K 159
A KR T SN G K I O N W I R S V K 159
NA I K KR V SN G K I O N W I R S V K 159
N I K K V K K I E N D K H E N I RL R E P S T K T 159
N IH K S V K RN K K K A E NI RR S R R T R 159
N I R V K RN K K K A E NI RR S R R T R 159
K H IH K Q I NT K I E K K R N G H E N I NR S R T 159
S Y I R K L S S M E NI K QH E N I O SS T 159
A IHT K A H H S M E NI K QH E N I O SS T 159
L K N I KR Y R K V V E K Q A H V R D D H A I R D T K RR 159
D GL K N RR N H R KV E K K O A H V LR N G MH A I R D D N Q T K RP 159
L K N RR N D R KV E K K O K H LRK N IE G M A A I R D N K R 159
L K N RR N R V E S R O E H W G MH A O I I K T K T 159
K GL K N I KR O S R K X E G I V LG T G IH T I I N P T E 159
L K A N I RR N L K RN T IVK O H LRG T G MH AE I OR I D T LK R 159
L K N RR V R K V E K I OP E H LR N G MH AE I Y I D P T 159
D L K N RR DCR KV I K Q E H L K G MH A I R D D T K 159
L K N RR D R K K O A H L R K I D MH A O I V D T K 159
L K A N I K K R R K MH AE I Q R N S T K T 159
L I K N RR N Y R KV E I K OP E H LR I G MH A I Y T D Q T VK 159
L K N RR N Y R KV E I K OP E H LR I G MH A I Y T D Q T VK 159
GL K N I RR R K V E K I I K OP E H LR N I G MH T I Y I D T T K R 159
L K N I R S V NR KV E K I K Q A H A LGK I D MH A O I F D T K 159
GL N I R S R R KV E K K O E H A LR G G MH A O I I D P T K R 159
V I I K TA D K K VI R R G M E I R T K OR 159
Y K K A DNH K K V N R R G M M E I R T A OR 159
I KQ KR W C N E K V N E L R I D TH A O I R IS N P T LA 159
N Y I K N V D N KL E R KV N K P L E R I D MH A O I I D T RA 159
N K Q C K D H K E V K L E R I D G I E O I I D T RA 159
R O L D H K E V V R R G M T A E O I I D T I K 159
K K W E V K O K G T A E O I I I N T K OT 159
S G I KQ N L DC N K E K I V N R G I AE I R F N T V 159
N Y H K K D K E V K K P R I G I E R D N K A S T 159
N Y IH K K V F R V RK DK N E LGK I SM E R I YR N H V 159
N Y H K K V F R V EE K I XGN R I G I S E G I RI N RL T A Q T 159
N Y I K R V FO R V NE E L OG N I G I H E E O I KI N T A 159
N Y H K TKN V F R I KE DK E LG I G I H E E O I YKI N T 159
N Y HI R K C R N I DK N I V E A OG R I G I E E I YR N ST T A 159
N Y HI K K F R K I D K V E A OG R I G I E E I YR N ST T A 159
N Y K Q NK V F R KV KE D A OG R I G I E E I YR N ST T A 159
N Y HI K K S F R V EE K I V OG R I G I E E I YR N ST T A 159
N Y K K S F R V EE K I V OG R I G I E E I YR N ST T A 159
N R R Q R V EWSK V K H LG R I SM IH E E I RI N T VA 159
X N R A Q R V A DST V K H Y R S I G K H E E T N I R N T V 159
L N K K Q R V K DSK V H Q X R G H E E N YR N T N RT 159
N I K QR R KV T VR P K Q H L I D R H E I QV R KV VA 159
N K K R KV I VR N KA H SLR I D SH E I O R K V TRR 159
N I R K R K E K V E K I VR K Q H OR I D TH H E I QL R K V 159
N K K R V L K I VR E H K OR H D N E N I QL R K V 159
K A N I K I VR E H K OR H D N XN X H K I QL R P S K VP 159
N I K K R KV T VR P K Q H L I D R H E I QV R KV VA 159
N K K R KV F R KV E T VK QA N T OR N I D H H E I E R K 159
N K K S F R KV E T VK N K Q H T OR I I D H N E II E R K 159
N K N K S F R V E K VR K Q H OR I I D H H E I QV R K V 159
V N I K K R KV K V E K V H OR H I D H H E I EL R F K VP 159
N K N I K KR V TN KT V H QE I D H H E I R I D R T S RT 159
N Y I K NR V D T E N E E G MH E E G I RV R N P K Q T T VA 159
N Y HI K KD F T K E I N E QR E D I R Y S T T K Q T T V 159
N Y HN K L D N KV C E V L R R G MH EA I Q D T K R 159
N Y N K L D N K E I V O QR R G MC I O D T R RR 159
N Y N K O L D R K E I V O QR R G MH I O D T K RR 159
N Y K NR D N K E V L E L K R G I E E I R C T Q T A 159
N Y HI K NR Y R K E V L E L K R G I E E I R C T Q T A 159
N K I K KL QN KV O I G I R O K Q H LRK N I D H H E I QV RH D PS K A 159
A I K I K QN KV O I G I R O K Q H LRK N I D H H E I QV RC PS K S 159
N I K N S O K V I I G I R O K Q H A ORN I D R IO E A E R I QV R PS 159
N K K Y O K K O H ORN I D R O A R I QV RYK PS R TG GS 159
N K K W KV I I G I R O K Q H ORN I D R O A ET G I KV R X PS K T 159
H K K R C O KV I I G I R O K Q H A ORN I D R O N E I QV RY PS 159
I K H I K TK K O KV T I R O K Q H A ORN I D R O N E I QV RY PS 159
N I K K V O KV I I G I R O K Q H ORN I D R O N E I QV RY PS 159
N I K K S O KV O I I G I R O K Q H ORN I D R O N E I QV RY PS 159
L K H I R K D S R V O I I G I R O K Q H ORN I D R O N E I QV RY PS 159
N K K R K E S T K IVR O H H M QES I D MH G I QV R PS K VP 159
A N R K TKD TR KV VK N KQ I D H H N E I QV R S Q K V TRK 159

	Tat start	C-rich region	nuclear localization	exon 1 end exon 2 start	Tat end
B.FR.83.HXB2	MEPVPDRLEPKWHPGSPQKACT.	..NCYCKKCCFHQCVCIFITKAL	..GISYGRKKRR.	.QRR.RAHONSSQTHOASLSKOPTS.QPR.	.G.DPTGPK*KKKVERETETDPFD
A1.CD.97.97CD.KCC2	---N---E---P-N-R	---Y---P---VN-G	---K---L---	TPRSNKD-NPIP-L-AQ-R	---E-S-E-SKA-P-R
A1.CM.08.886.24	-D-N-N-N-R-T-S	---S---Y---A-LR-G	---	GTPYS-KD-NP-PE-SLP-TO	---EW-E-E-SO-R
A1.CY.08.CY236	-D-N-N-N-N-T-S-K	---Y---D-LK-G	---	PP-S-ED-NPIPE-TP-TO	---VS-E-E-SK-A-R
A1.TN.09.NARI.FLS.TVC19.1	-D-N-N-N-N-K-T-P-S-K	---Y---X-LK-G	---K---	PPKS-ED-NPVQ-SIP-T	---V-E-S-E-MNKATAGRX
A1.KE.11.DEM111KE002	-D-N-N-N-N-T-P-N-K	---Y---PA-LN-G	---	GTP-S-KD-NPVP-TP-RAQ	---IS-E-E-SKA-R
A1.RU.11.DUR6950	-D-N-N-N-N-A-S	---W---L-LK-G	---H---	GTPHS-KD-NPI-LP-HTQ-R-Q	---E-SA-SKA-R-A
A1.RW.11.DEM111RW002	-D-N-N-N-N-A-S	---Y---P-LS-G	---V---	GTP-SNKD-NPVP-SIP-TK-N	---E-S-E-SKA-R
A1.SN.01.DDI1579	-D-N-N-N-N-H-S	---A---W-LN-G	---K---	GTP-SNKD-NPVP-LP-TP	---S-SK-R-A
A1.UG.11.DEM110UG009	-D-N-N-N-N-T-P-S-K	---E---Y-I-OH-G	---W---	T-G-KD-NPIP-LP-AP	---S-SKA-R
A1.ZA.04.503.15344.T10.A1	-D-N-N-N-N-T-P-N-K	---Y---EL-LN-G	---K---	GTP-SNKD-NPVP-LP-AP	---NS-E-E-SK-A-R-N
A2.CD.97.97CDKTB48	---K---N---R-N	---V---W-LN-G	---P---	GPD-GND-NPVP-SLP-TO-R-S	---E-S-SKA-R
A2.CM.01.01CM.1445MV	---K---N---R-N	---V---W-LN-G	---P---	PS-S-KD-NPVP-SLP-HAO-R-VS-Q	---S-E-SK-HQ
A2.CY.94.94CY017.41	---K---N---R-N	---Y---L-LN-G	---P---	KPPSPNKD-NPIP-SLP-AQ-R-V	---E-P-E-SKA-R
B.BR.10.10BR.RJ032	---S---E-P-N-K	---Y---T-G	---	SP-DG-N-D-Y	---S-K-AV
B.CA.07.502.1191.03	-I-S---N-N-K	---Y---T-G	---	SP-S-KTD-P-A-H	---P-E-HQVG
B.CH.08.M2.0803101.NFLG8	---S---P-D	---TK-G	---	SP-D	---P-E-IH-S
B.CN.12.DEM12CM006	---R---P	---G	---	TP-D-PPD-A-A-S	---P-E-V-L
B.CU.14.14CU005	---R---P	---S-W-Y-M-G	---	PP-D-E-D-AT-L	---S-E-V-R
B.ES.14.ARP1195	-D---	---Q---K-G	---	TP-Y-E-D	---Q-O-V-P-R
B.FR.11.DEMB11FR001	-D---	---N-K-R-Y-L-G	---	TP-D-P	---S-E-KA-H-DA
B.HT.05.05HT.129389	-L---	---R-X-X-Y-L-G	---	XS-X-A-A-L	---K-Q-S-Q-W
B.JP.12.DEM12JP001	-S---	---N-P-TK-G	---	PP-G-V-P-A-A	---Q-E-SK-VH
B.KR.07.HP.18.07JHS10_3909	---S---	---P	---	P-DNKN-V-P-A-A	---Q-S-E-K-M-VH
B.RU.11.11RU21n	-D---	---L-MK-G	---	SP-N-D-P-A-A	---S-E-K-H
B.SE.12.SE600057	-D---	---L-MK-G	---	SP-GD-D-P-A	---S-E-K-K-A
B.TH.10.DEMB10TH002	---N---	---TK-G	---	SP-D-N-Y-P-A-S	---Q-S-E-KH
B.US.13.RV.1	---N---	---P	---	P-D-S-V-A-S	---S-S-VS
B.ZA.09.DEMB09ZA022	---N---	---P	---	TPED-T-T-G	---N-S-V-D
C.BR.07.DEMC07BR003	-D-I-HN-N-N-P-N-K	R-SY-I-L-G	Y	S-PPS-ED-DLT-L-A	S-E-FK-SK-A
C.BW.00.00BW003.1	-L-I-N-N-N-D-N-K	R-SY-L-OK-G	---	S-P-S-ED-NPI-L-RTK	N-S-E-S-SK-K-WRTDS
C.CN.10.YNFL19	-L-I-N-N-N-N-K	Y-LA-Q-G	---	S-PR-ED-DLT-L-T	E-S-K-KNK-ETQLV
C.CY.09.CY260	-D-I-N-N-N-N-P-N-S	SY-P-L-G	Y	T-PPSNKD-NPVP-LP-T	N-E-S-E-SK-TA
C.ES.14.ARP1198	-D-N-N-N-N-N-D	F-SY-P-OK-G	---	S-PPS-ES-DP-LP-T	N-E-S-E-SK-A
C.ET.02.02ET.288	-D-N-N-N-N-N-D	F-SY-P-OK-G	---	S-PPS-ED-NPI-LP-RAQ	S-E-S-SK-A
C.IN.09.T125.2139	-D-N-N-N-N-N-P-N-K	SY-L-L-G	---	S-PPS-ED-NPI-LP-T	S-E-S-SK-A
C.KE.05.05KE369195V4	-D-N-N-N-N-N-P-N-K	SY-L-L-G	---	S-PPS-ED-NPI-LP-T	N-E-S-E-SO-W
C.MW.09.703010256.CH256.w96	-I-K-N-N-N-N-N-K	SY-L-L-G	---	S-PPS-KD-NP-P-L-T	N-E-S-E-SO-W
C.SE.13.SE600311	-D-N-N-N-N-N-N	F-SY-V-OK-G	GIS	S-PPS-ED-NPI-LP-PTO	R-R-RL-TKP-P-L-Y
C.TZ.08.707010457.CH457.w8	-Q-K-N-N-N-N-N-T	F-R-SY-L-L-G	---	S-PPS-ED-NPI-LP-T	N-E-S-E-SK-T-L
C.US.11.17784.468	-D-N-N-N-N-N-N	F-SY-L-L-G	---	S-PPS-ED-NPI-LP-T	N-E-S-E-SK-T-L
C.YE.02.02YE511	-I-K-N-N-N-R-N-K	A-Y-L-OK-G	---	S-PPD-ED-NLV-LP-T	A-S-E-S-SK-C-NWRTDS
C.ZA.12.DEMC12ZA096	-D-K-N-N-N-N-K	F-H-SY-L-OR-G	W	RG-PPS-KD-DP-P-LP-T	Q-S-E-SK-AT
C.ZM.11.DEMC11ZM006	-D-K-N-N-N-N-K	R-Y-L-Q-G	---	S-PPG-KD-NPI-SP-A-N	S-E-S-SK-KA-YA
D.CM.10.DEMD10CM009	-D-N-N-N-N-S-N-S	R-Y-R-S-G	G	KPPKNDT-VPVPD-S-H	E-SK-A-W
D.CY.06.CY163	-D-N-N-N-N-N-S	R-Y-C-G	G	PP-GNK-GPVPE-S-HP	P-P-NPS-E-SK-A-W
D.DE.11.DEMD11KE003	-D-I-S-N-N-N-R-P-N-K	Y-O-VS-G	G	P-GG-V-DPIP-P	E-SKA-AW
D.KR.04.04KR04	-D-N-N-N-N-N-N	Y-L-LK-G	---	PP-GD-A-VP-LP-P	Q-E-SK-ANO-W
D.NS.90.SE365	-D-N-N-N-N-N-P-N	Y-N-G	---	TP-GN-V-VP-S-H	SKA-Q-W
D.TZ.01.A280	-D-N-N-N-N-N-R-P-N	Y-N-G	---	TP-GD-A-DPIP-S	Q-SKA-A-R-W
D.UG.10.DEMD10UG004	-D-N-N-N-N-N-P-N-K	N-Y-N-G	---	PP-SG-A-DPIP-S	SKA-Q-CWEDYRHLGGTT
D.UG.11.DEMD11UG003	-L-N-D-N-N-N-R-P-N-K	Y-N-N-G	---	GTP-GG-A-VPVP-S-L	S-A-Y-W
D.YE.02.02YE516	-D-N-N-N-N-N-R-S-P-S	P-Y-L-G	G	SP-SN-VPVP-S-L	S-SOKAP-W
D.ZA.90.R1	-D-N-N-N-N-N-P-N-K	H-Y-L-S-G	---	KPPHSD-D-VPVPE-S	S-A-NO-W
F1.A0.06.A0.06.AMG32	-I-N-D-N-N-T-P-K	A-YY-G-G	---	TP-S-KN-NPVE-L-A	S-EM-SK-S-*
F1.AR.02.ARE933	-L-N-D-N-N-T-P-K	R-YR-LAR-G	---	H-TPHGT-I-DLVP-L-A	N-S-E-SKAK-C-*
F1.BR.07.07BR844	-I-N-D-N-N-T-P-K	R-YW-T-G	H	H-TP-SN-DLV-I-A	N-S-E-SKA-G-*
F1.BR.10.10BR.PE107	-L-N-D-N-N-T-P-K	Q-YW-T-G	---	O-PSLS-S-DLIP-I-A	N-S-E-SKA-CA-*
F1.BR.10.10BR.RJ015	-L-N-D-N-N-T-P-K	Q-YW-T-G	---	H-HR-L-DPVP-S-A	N-S-E-SK-CA-*
F1.CY.08.CY222	-L-N-D-N-N-T-P-K	Q-YW-T-G	---	H-TP-D-N-YW-T-G	DRV-N-E-SK-C-*
F1.ES.02.ES.X845.4	-V-N-D-N-N-T-P-K	T-V-YR-T-G	---	H-PP-SNKN-NPVE-L-A	N-S-E-SK-G-*
F1.ES.11.VA0053.nfl	-S-D-N-N-A-P-S	R-W-V-G	---	H-PP-S-PL-DPV-I-A	N-S-E-SKA-N-C-*
F1.RO.96.BCI.R07	-L-N-D-N-N-T-P-N-R	YH-T-G	D	H-TP-S-PT-NPVP-L-A	N-S-E-NK-A-*
F1.RU.08.D88.845	-L-N-D-N-N-A-P-K	R-TM-G	---	GTX-S-I-NPVE-S	Q-SKA-*
F2.CM.02.02CM.6016BBY	-H-K-D-N-N-NP-N-K	R-L-TS-G	---	S-SHS-KD-DPV-LP-T	N-S-E-SK-K-*
F2.CM.10.DEMF210CM001	-V-KID-N-N-E-P-N-K	R-L-T-G	---	TP-S-ET-NPVP-LP-T	E-S-SKAK
F2.CM.10.DEMF210CM007	-I-N-D-Q-Q-E-P-N	R-TS-G	---	SP-S-KV-DPVP-L-A	E-S-SK-K-G-*
G.CM.07.920.49	-D-N-N-N-N-N-N-K	V-Y-LN-G	K.H	G-PPG-KD-NPVE-LP-TT	N-S-E-SKA
G.CM.10.DEMG10CM008	-D-K-N-N-N-N-N-K	A-W-LN-G	K.H	GSSHG-KD-NPVE-SP-IS	N-Q-S-E-ASKA-Q
G.CM.10.DEURF10CM020	-L-K-N-N-N-T-N-K	F-W-LK-G	R	GTP-SR-N-NPVE-PP-TT	NT-P-E-SK-G-VN
G.CN.08.GX.2084.08	-D-N-N-N-N-N-K	V-W-LN-G	K.H	ATPSSNKD-NPVP-TP-TT	N-D-S-SK-K
G.ES.09.X2634.2	-D-N-N-N-N-N-N-K	V-W-LN-G	K.H	TP-NTO-NPVE-LP-TT	N-E-S-SK-N
G.ES.14.ARP1201	-D-N-N-N-N-N-N-K	V-W-LN-G	K.H	GXPXS-D-NPVP-PP-TT	N-E-S-SKA-A
G.GH.03.03GH175G	-D-N-N-N-N-T-P-N-K	V-W-LN-G	K	GTP-G-KD-NPVP-LP-IS	NQ-**-P-E-ASK-G-CA
G.KE.09.DEMG09KE001	-D-N-N-N-N-R-N-K	V-W-LN-G	K.H	SPG-KD-NPVE-LP-TT	N-A-S-E-S-K
G.NG.09.09NG.SC62	-D-N-KI-N-N-N-K	V-Y-LN-G	N.H	TPSS-N-NPVP-PP-TT	N-S-E-SK
G.PT.x.PT3306	-D-N-N-N-N-N-K	W-LN-G	K.H	GTPH-KD-NPVP-SSP-TT	NQ-DQ-S-E-SK-N
H.BE.93.VI991	-D-NQ-N-N-N-R-N	Y-L-LK-G	Y	GTPKSL-D-TLIP-L-RTS	EK-ASK-A-W
H.BF.93.VI997	-D-NQ-N-N-N-N-P-N	Y-L-LK-G	S.R	ATPSV-D-NPVP-LP-TT	N-S-E-SKAK-A-C
H.CF.90.056	-D-N-N-N-N-Q-N	Y-M-LK-G	---	H-TPASL-D-N-I-L-RTH	Q-E-ASK
H.GB.00.00GBAC4001	-D-K-N-N-N-N-G	N-Y-L-LK-G	---	H-GTPAGL-DN-DPIP-LP-RT	N-S-E-SKA-LH
J.CD.97.J.97CD.KTB147	-D-NI-NQ-NQ-N-Q	R-Y-PL-LK-G	H	A-PSG-K-DPIP-L-TO	R-ES-OEKS-E-SKA-P-R
J.SE.93.SE9280.7887	-NR-N-N-N-N-N	Y-LO-G	---	S-PPG-K-DLIP-L-TO	R-K-E-S-SKA-P-R
J.SE.94.SE9173.7022	-NR-N-N-N-N-N	Y-LO-G	---	S-PPG-KN-DLIPE-LF-TO	R-K-E-S-SKA-P-R
K.CD.97.97ZR.E0TB11	-D-NI-NQ-NQ-N-Q	Y-L-LO-G	C.E	T-TTPYA-KN-KDPIP-LP-A	S-E-SK-K-*
K.CM.96.96CM.MP535	-D-NI-NQ-NQ-N-Q	R-Y-I-LK-G	P	TTPY-EN-DP-R-L	EQ-D-S-SK-K-*
01.AE.AF.07.569M	-L-N-N-N-N-T-N-S	W-L-LK-G	K.H	GTP-S-KD-NPVP-LP-II	N-Q-S-E-ASKA-C
01.AE.CM.11.1156.26	-D-N-N-N-N-T-P-N-K	V-W-L-LK-G	K.H	GTPPG-KD-NPVP-LP-II	N-S-E-SASKA-C
01.AE.CN.12.DE0012CN011	-L-N-N-N-N-T-P-N-K	V-W-L-LK-G	K.H	GTP-S-KD-NPVP-LP-II	N-S-E-SASKA-C
01.AE.HK.04.HK001	-N-N-N-N-N-K-T	K-I-SW-L-LK-G	K.H	GTPRS-EG-NPVP-SL-TS	NQ-D-E-S-AGK-*
01.AE.IR.10.101R.THR48F	-N-N-N-N-N-P-V-S	K-I-W-L-LK-G	K.H	TP-SNKN-DPIPE-SLP-TS	N-D-S-E-ASK-G-CA
01.AE.JP.11.DE00111JP003	-N-N-N-N-N-T-P-N-K	W-L-LK-G	K.H	GTP-S-KD-NPVP-LP-II	N-D-S-E-ASKA-C
01.AE.SE.11.SE601038	-N-N-N-N-N-T-P-N-K	W-L-LK-G	H	GTP-S-KD-NPVP-LP-II	N-S-E-SASKA-C
01.AE.TH.10.DE00110TH001	-L-N-N-N-N-T-P-N-K	W-L-LK-G	K.H	GTP-S-KD-NPVP-LP-II	N-D-S-E-ASKA-C
01.AE.TH.90.CM240	-L-N-N-N-N-T-P-N-K	W-L-LK-G	K.H	GTP-S-KD-NPVP-LP-II	N-D-S-E-ASKA-C
01.AE.US.05.306163.FL	-L-N-N-N-N-T-P-N-K	W-L-LK-G	K.H	GTP-S-KD-NPVP-LP-II	N-D-S-E-ASKA-C
02.AG.CM.10.DE00210CM013	-S-N-N-N-N-A-S	W-L-LN-G	H	GTP-RRD-NPVP-LP-TT	N-Q-P-E-SK-Y
02.AG.ES.06.P1423	-K-S-N-N-N-N-S	W-L-LN-G	P	TP-SR-D-NPVP-LP-TT	N-S-E-SK-O
02.AG.GW.05.CC.0048	-S-N-N-N-N-T-S	K-I-W-L-LK-G	Q	GTPPS-D-NPVP-FT-TI	I-A-S-E-ASK-RC

Table with columns: Tat start, C-rich region, nuclear localization, exon 1 end, exon 2 start, Tat end. Rows list HIV-1 protein sequences with their corresponding Tat start and end positions.

Rev start	exon 1 end	exon 2 start	NLS	Leu-rich effector domain	Rev end																				
B. FR. 83. HXB2	MAGRSGD	.SDE	ELTRTVRLIKLLYQSNPP	P.NP.P	E.GTRQARRNRRRRERQRIHSISERILGTYLGRSAEPVPLQLPPLERLTLDCNEDCG	TSQGTQ	GVGSPQILVESPTVLESQTK	...	*															
A1.CD.97.97CD KCC2	S	A	LAAI	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	G	S	SG	EN		
A1.CM.08.886 24	S	A	LTA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
A1.CY.08.CY236	S	A	LAI	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
A1.IN.09.NARI FLS TVC19 1	N	D	LAI	I	I	N	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K
A1.KE.11.DEMA11KE002	N	D	LAI	I	I	N	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K
A1.RU.11.DEM11RW002	S	G	LAI	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
A1.SN.01.DD1579	N	D	LAI	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
A1.UG.11.DEMA11UG009	N	D	LAI	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
A1.ZA.04.5013 15344 T10 A1	S	P	LAI	I	I	N	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K
A2.CD.97.97CDKTB48	S	P	LAI	I	I	N	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K
A2.CM.01.01CM 1445MV	T	D	P	LAI	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K
A2.CY.94.94CY017 41	T	D	P	LAI	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K
B.BR.10.10BR RJ032	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
B.CA.07.502 1191 03	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
B.CH.08.M2 0803101 NFLG8	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
B.CN.12.DEMB12CN006	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
B.CU.14.14CU005	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
B.ES.14.ARP1195	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
B.FR.11.DEMB11FR001	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
B.HT.05.05HT 129389	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
B.JP.12.DEMB12JP001	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
B.KR.07.HP 18 07JHS10_3909	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
B.RU.11.11RUZ1n	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
B.SE.12.SE600057	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
B.TH.10.DEMB10TH002	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
B.US.13.RV 1	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
B.ZA.09.DEMB09ZA022	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.BR.07.07YK07BR003	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.BW.00.00BW003 1	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.CN.10.YNFI19	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.CY.09.CY260	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.ES.14.ARP1198	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.ET.02.02ET 288	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.IN.09.T125 2139	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.KE.05.05KE369195V4	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.MW.09.703010256 CH256.w96	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.SE.13.SE600311	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.TZ.08.707810457 CH457.w8	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.US.11.17784 468	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.YE.02.02YE511	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.ZA.12.DEMC12ZA096	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
C.ZM.11.DEMC11ZM006	N	A	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
D.CM.10.DEMD10CM009	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
D.CY.06.CY163	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
D.KE.11.DEMD11KE003	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
D.KR.04.04KR04	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
D.SN.90.SE365	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
D.TZ.01.A280	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
D.UG.10.DEMD10UG004	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
D.UG.11.DEMD11UG003	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
D.YE.02.02YE516	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
D.ZA.90.R1	R	D	LKA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
F1.A0.06.A0 06 ANG32	N	L	LTA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
F1.AR.02.ARE933	N	L	LTA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
F1.BR.07.07BR844	N	L	LTA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
F1.BR.10.10BR PE107	N	L	LTA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
F1.BR.10.10BR RJ015	N	L	LTA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
F1.CY.08.CY22	N	L	LTA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
F1.ES.02.ES X845 4	N	L	LTA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
F1.ES.11.VA0053 nfl	N	L	LTA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
F1.R0.96.BCI R07	N	L	LTA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
F1.RU.08.D88 845	N	L	LTA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
F2.CM.02.02CM 6016BBY	N	L	LTA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
F2.CM.10.DEMF210CM001	N	L	LTA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
F2.CM.10.DEMF210CM007	N	L	LTA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
G.CM.07.920 49	S	T	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
G.CM.10.DEMG10CM008	S	T	LQA	I	I	Y	K	R	K	A	N	D	R	C	E	FC	H	S	OSQGTET	R	AP	A	G	K	
G.CM.10																									

	transmembrane domain	phos phospho cytoplasmic domain		Vpu end
Vpu start		α -helix	α -helix	
B.FR.83.HXB2	TOPI...P.I....VAIVALVVAII.AIVVSWISIVIEIKYLRILR...KIDRLIDRLIERAE.DSGNESE...GEI.SALVEMG.V.EMGGH.APWD.VDDL...*			82
A1.CD.97.97DC.KCC2	MN.L...WS.I.I.F.L...G.I.AF...K.LRK...W.E.IR...GDT.E.L.T...M...NL.DFL.NNN...			81
A1.CM.08.886.24	MLSLGIWE...W.G.I.VL...T.V...CKRV.K.K...KI.A.IR...DGD.T.D.L.AR...NY.NLG.DN...			85
A1.CY.08.CY236	MT.L.E...S.G.VS...T.G.IK.V...K.IR...GDT.D.L.AL...NY.DLG.DNN...			83
A1.CE.11.DEM11KE002	ML.P.LQ...C.AG.L.L.L...A.T.G...C.LQ...E.IR...DGD.T.D.L.AT...R.NY.DLGN.NX...			83
A1.RU.11.RU6950	MT.L.E...Y.A.VF...L.A.I.G...IK.N...I.E.IR...DGD.R.E.L.K...I.D.DY.DLGY.DNI...			82
A1.RW.11.DEMA11RW002	M.L.L.V.V...W.IG.I.L.L...I.T.GL.K.L.K...LE.IR...DGD.T.E.L.W.T...NY.DLGN.DNN...			82
A1.SN.01.DD1579	MTSL.E...W.IG.I.L.L...T.G.K.V.K.K...R.E.IR...DGD.T.E.L.T...V.DD.I.LY.DNN...			82
A1.UG.11.DEMA11UG009	MR.L.E...W.G.I.L.L...T.G.KRLR.K...E.IK.E...GDT.E.L.AG...I.N.NY.DLGN.DNN...			82
A1.ZA.04.503.15.344.T10.A1	MTLL.E...W.TG...S.L...T.V.F...K.K...W.L.E.IS...GDA...I.W...T.D.NL.GY...			82
A2.CD.97.97CDKTB48	MS.L.A...LS.G.I.F.L...T.V.F...Q.KELKK...W.K.IS...DGD.T.E.L.KM...G.NL.GY.DN.V...			82
A2.CM.01.01CM.1445MV	ML.L.A...L.G.I.L.L...T.F.K.K.KK...W.K.IS...DGD.T.E.L.T...R.L.DFG.NNV...			82
A2.CY.94.94CY017.41	ML.L.V...L.G.I.L.L...T.F.K.K.KK...W.K.IS...DGD.T.E.L.T...R.L.DFG.NNV...			82
B.BR.10.10BR.RJ032	ML.L.Q...A.X.VA...T.F...L.I...GDE.EQL.A...M...N...			82
B.CA.07.502.1191.03	MT.L.A...L.A.A.L...T.G...N.L.TA...GDO.E.L...R.HL...N...			87
B.CH.08.M2.0803101.NFLG8	MT.L.Y...WS...A.L...T.AL...IL...IR...DGD...D...D...N...			82
B.CN.12.DEMB12CN006	M.L.I.V...L.A.A.L...T.F.L...R...IR.T...DGD...D...D...N...			82
B.CU.14.14CU005	M.L.V...G...I.A...T.F...E.IRD...GSGRI.QP...G...G...EC...			85
B.ES.14.ARP1195	M.L.V...L...A...X.F...X...N...ITD...XDE.EAL.A...DR.D...V...			82
B.FR.11.DEMB11FR001	M.SL.T...L...L...T.F...S...IR...EDQ...E...M...L...I...			79
B.HT.05.05HT.129389	M.SL.X.X...L...X.X...I.X.F...X.K.K...R...IR...DGD...E...M...D...L...N...			82
B.JP.12.DEMB12JP001	M.L.Q...L...A.L...T...G...N...IR...DGD...L...M...N...M...			82
B.KR.07.HP.18.07H510_3909	M.L.V...NYL...K.K...T...IR...DGD...T...M...L...N...M...			83
B.RU.11.11RUZ1n	M.L.V...AT...F.VG...L.V...A.I...DGD...E...M...L...L...M...			79
B.SE.12.SE600057	M.SL.V...L...IVG...I.T...L...IR...DGD...E...#...D...D...N...			81
B.TH.10.DEMB10TH002	MSSL.Y...LT...F.LVT...F...R...I...IR...DGD...E...M...GN...I...			82
B.US.13.RV.1	M.L.G...I...L...T...Y...S...IR...DGD...E...L...D...N...I...			81
B.ZA.09.DEMB09ZA022	M.L.E...G...L...T...Y...R...IR...DGD...E...L...D...N...I...			82
C.BR.07.DEM07BR003	MLT...DRVDYRLG.G.T.L...I.T.AY...LV...R...E.IR...DGD.T.E.L.ET...D.NL.RLL...HA...			87
C.BU.00.00BU0031.1	MFAL.F.E.VDRLT.G.FI.LFL...I.T.AY...LV...R.Q.VK...IR...DGD.T.E.L.TM...D.L.RLL...N...			86
C.CN.10.YNFL19	MVDL...DYK.TVG.FI.L.L...I.T.AY...LTK...W.K.IG...GDT.E.L.TM...D.RL.RLL...N...			83
C.CY.09.CY260	MTGL.I.EKVDYKI.V.FII.LV...I.YL...LV...E.IR...DGDH.E.L.KM...D.NL.RLL...AH...			87
C.ES.14.ARP1198	MLDL.L.TRVDYKI.VG.FII.LL...I.AY...LV...W.N.IR...GDI.E.L.TM...D.QL.RLL...N...			87
C.ET.02.02ET.288	MFSF.T.ERVDAR.G.G.I.L.L...I.T.AY...LV...E.IR...GDT.E.L.TM...D.TL.RLL...H...			87
F1.IN.09.T125.2139	M.L.V...NYL...K.K...T...IR...DGD...E...M...L...N...M...			83
C.KE.05.05KE369195V4	MLDL.L.AKVYRI.VG.FI.LLL...I.T.Y...LVK...W.K.IR...DGD.T.E.L.TM...D.L.RLL.AN...			87
C.MW.09.703010256.CH256.w96	MLDL.V.ARVYRI.VG.V.I.L...I.LAY...KRW.K.K...W.K.IR...GDT.E.L.TM...D.E.L.GLL...NG...			86
C.SE.13.SE600311	MVDL.T.AGVYRI.A.FAI.LL...I.IY...LV...N.W.VK...IR...DGD.T.E.L.TM...D.NL.RLL.AV...			87
C.TZ.08.707010457.CH457.w8	MLDL.L.ARVYRI.VG.A.LI.LL...I.T.AY...K.L.K...R.W.K.IR...DGD.I.E.V.ET...D.NL.RLL...IN*			86
F1.US.11.17784.468	M.L.E...A.ARVYRI.G.A.FII.LL...I.T.AY...LV...DGT.D.L.ATM...D.NL.RLL...D...			87
C.YE.02.02YE511	MLDL.L.AKVYRI.VG.FI.LL...I.T.AY...LVI...W.VK...IR...DGD.T.E.L.TM...D.L.RLL.IN...			86
C.ZA.12.DEM12ZA096	MVNW.L.EQVDYRLG.VG.IA.LL...T.AYL...LV...N.E.IR...DGD.T.E.L.TM...D.YT.RLL.GN.V...			87
C.ZM.11.DEM11ZM006	MISL.I.ERVDYRLG.VG.II.LL...T.AYL...L.K...W.E.VR...GDI.E.L.TM...D.L.RLL.GNE...			87
D.CM.10.DEM10CM009	M.L.V...IS...TL...T.Y...Q.K...W.IR.E...GDK.E.L.T...D...IVA...			83
D.CY.06.CY163	M.SL.E...L.A.L.L...T.Y...K...W.IG...GDK.E.L.Q...M...NVA.I...			83
D.KE.11.DEM11KE003	M.SL.Q...A...L.L...T.F...C.RLR...#...W.IR...GDE.E.L.F...G.Y...IN.I...			81
D.KR.04.04KR88	M.V.V...I.X.TL...I.T.F...C.RLR...#...W.IR...GDE.E.L.F...M...N...			81
D.SN.90.SE365	M.SL.E...L.A.I.L.L...T.F...V.R...W.IK...GDE.K.L.T...M...D...N...			82
D.TZ.01.A280	MT.L.E...T.A.I.S.L...T.Y...R.K.K.K...W.IR...DGD.T.E.L.T...D...N...I...			82
D.UG.10.DEM10UG004	M.L.V...L.LL...T.F...C.L.R.K...R.W.IR...GDE.E.L.KF...M...N.I.M...			82
D.UG.11.DEM11UG003	MNLL.V...IG...L...A.F...W.RLK...KI.N.TA...GDR.E.L.T...G.L...IG.I...			81
D.YE.02.02YE516	M.L.E...LS...I.L...I.T.Y...K...R.N.YE.IR...GDA.E.L.A...G...V.PF.I.GG.INN...			82
D.ZA.90.R1	M.L.V...L.L.L...T.F...R.RK...Y.IR...DGD.E.L.K...YD...N...			83
F1.A0.06.A0.06.AMG32	MSDL.L.A.I...L...T.AYL...VVK.K...R.N.YE.IR...DGD.A.E.L.AE...G.V.PF.I.GG.INN...			82
F1.AR.02.ARE933	MSYL.L.A.IG.T.I.L...T.AY...K.LV...R.N.YK.IR...GDA.E.L.A...G...PF.I.R.I.N...			82
F1.BR.07.07BR844	MPDL.L.A.S.A.I.F...I.T.AY...VV...N.YE.I...GDA.E.L.A...G.V.PF.I.G.INN...			82
F1.BR.10.10BR.PE107	MNYL.V.S.I.I...VV...T.AYL...LV.N...R.NS.YO.T...GDA.E.L.A...G...LF.I.GN.I.N...			81
F1.BR.10.10BR.RJ015	MTDL.L.V.I.S.A.I.L.V...T.YL...LV...N.YO.R...GDA.E.L.A...G...PF.L.G.INN...			82
F1.CY.08.CY22.2	MTLA.L.A.I.A.I.L.V...T.F...K.VI...N.YKIR...GDA.E.L.A...G.V.PF.I.G.INN...			82
F1.ES.02.ES.X845.4	MSNL.L.A.I.S.L...I.T.Y...LV...N.YE.IR...GDA.E.L.A...G.V.PF.I.G.INN...			82
F1.ES.11.VA0053.nfL	M.L.Y...IS...I.L.V...T.AY...LV...N.YE.IR...GDA.E.L.AD...G.V.PF.I.G.INN...			80
F1.R0.96.BCI.R07	MSHV.L.A.I...I.L...I.T.AY...LV...N.YE.IR...GDA.E.L.A...G...PF.I.G.INN...			82
F1.RU.08.D88.845	MTE.I...IG...I.L.V...T.F...LV...N.YE.IR...DGD.A.E.L.A...G...PF.V.G.NN...			82
F2.CM.02.02CM.6016BBY	MSVL.I...LV.FI.L.A...I.T.Y...K.Q.K...R.N.YE.IR...GDA.E.L.A...G.V.F.I.G.INN...			82
F2.CM.10.DEMF210CM001	MPYL.I...IL.FIAVL...I.T.Y...K.K...N.K.YK.IR...N.GDA.E.L.A...G.V.PF.I.G.INN...			81
F2.CM.10.DEMF210CM007	MSL.S.I...V.A.FI.VL...I.T...K.L.K...R.NK.YE.IR...GDA.E.L.A...G...PF.I.GN.INN...			82
G.CM.07.920.49	M.L.E...S.G.II.SLA...T.A...RK.N...R.KVLN.IR...GDT.E.L.AQ...D.VDF.DL.V.G.N...			82
G.CM.10.DEMG10CM008	M.AL.V...A.G.I.FLA...T.T.F...Q.E.RK.N...K.L.IR.I...GDT.E.L.A...D.V.DF.DL.I.GNN...			82
G.CM.10.DEURF10CM020	M.SL.Q...AS.G...F.A.TI...T.Y...RK.K...R.K.N.IR...GDT.E.L.T...D.VDF.D.V.G.N...			82
G.CN.08.GX.2084.08	M.L.E...A.G.I.F.A...T.Y...R.K...R.KIL.IR...GDT.E.L.AT...D.VDF.D.V.G.NM...			82
G.ES.09.X2634.2	M.L.E...A.G.I.GLG...I.T.F...LKE.RK.Q...R.Q.IL.IG...GDT.E.L.AK...M...D.V.DF.V.G.N...			83
G.ES.14.ARP1201	M.L.E...S.G.I.F.A...V.T...X...RK.K...R.KIL.IR...GDT.E.L.AT...D.VDF.D.V.G.N...			82
G.GH.03.03GH175G	M.SL.E...S.G.I.S.A...I...F...R.RK.K...EK...IR...GDT.D.L.A...D.F.D.V.G.N...			82
G.KE.09.DEMG09KE001	M.L.E...AS.G.I.FLA...TI...T.F...Q.E.RK...Y.IR...GDE.E.L.AT...GR.AF.D.V.G.N...			82
G.NG.09.09NG.SC62	M.SL.E...A.G.I.F.A...I.T.F...R.K.K...R.S.LN.IR...GDT.E.L.AT...D.VDF.D.V.GNN...			82
G.PT.X.PT3306	ML.L.Q...F.AG.I.L.G.A...I.T.F...Q.E.RK.K...EKTL.IR...GDT.D.L.A...NF.D.V.V.G.I...			82
H.BE.93.VI991	MM.N.L.G...IG.G...F...T.AY...L#K...E.IR...DGD.T.E.L.K...L.NL.GY.A...			80
H.BF.93.VI997	MY.T.G...LG.G...TF...VI...T.Y...K.LV.K...E.IG...DGD.T.E.L.KM...M...L.NL.GY.A...			78
H.CF.90.056	MY...L.G...IG.G...I.G.L.C.I...T.YL...LVQ.K...R...LE.IR...DGD.T.E.L.T...R.L.NL.GY.A.V...			81
H.GB.00.00GBAC4001	M.L.D...A...G.I.V...VI...T.F...V...N.IR...DGD.T.D.L.EK...P...DL.N.N...			82
J.CD.97.J.97DC.KTB147	MI.L.Q...A...FI...FL...GM...T.Y...K.L...K...N.IR...DGD.T.E.L.AD...R.P...DL.N.N...			82
J.SE.94.SE9173.7022	MVSL.Q...I...FFL...C...T.Y...K.L...K...N.IR...DGD.T.D.L.AE...P...DL.N.N...			82
K.CD.97.97ZR.E0T811	MV.L.T...G.I...A.L.L...I.T.AYL...VV...K...R.NW.F...IR...GDT.E.L.A...G.T.L.ILG.INN...			81
K.CM.96.96CM.MP535	MVSL.A...IS...L.L.L...I.T.Y...LVK.K...R.NW...IR...GDA.E.L.ADI...G.L.L.I.LGN.I.N...			81
01.AE.AF.07.569M	MT.L.E...S.G.I.L.S...T.A...AK.L...VK...IR...GDT.E.W.AK...DF.D.V.G.N...			82
01.AE.CM.11.1156.26	MNAL.Q...W...G.VL.V...A...K...R.K.VK...IR...GDT.D.L.AT...DF.DH.V.G.N...			82
01.AE.CN.12.DE00112CN011	MT.L.E...S...T.L...T.V...AL.I...I...K...IR.E...GDK.AK...NF.D.V.G.NM...			85
01.AE.HK.04.HK001	MT.L.O...S...G.I.L.L...T...GL.V...K...KIR.E...GDT.D.L.AK...DF.D.V.G.N...			82
01.AE.IR.10.101R.THR48F	MT.L.E...S...AG...L.L...T...A...N...K...IS...X...K...IR.E...GDT.E.L.AK...DF.D.V.G.N...			82
01.AE.JP.11.DE00111JP003	MT.L.O...S...#...R...F...K...K...IR.E...GDT.D.L.AK...NF.D.V.G.N...			81
01.AE.SE.11.SE601018	MNSL.E...S...AG...I.L.L...T...A...Q...I...Q...K...K...IR...NE...GDT.D.L.AK...DF.D.V.G.N...YCFR			85
01.AE.TH.10.DE00110TH001	MT.L.E...S...G...I.L.L...T...A...V...K...K...IR...GDT.D.L.AK...DF.D.V.G.N...			82
01.AE.TH.90.CM240	MT.L.E...S...G...I.L.L...T...A...V...K...K...IR...GDT.D.L.AK...DF.D.V.G.N...			81
01.AE.US.05.306163.FL	MS.L.E...S...G...I.A.S...T...A...VK...F...VK...IR...GDT.D.L.AK...DF.D.V.G.N...			82
02.AG.CM.10.DE00210CM013	M.L.E...S...G...F.A...T...Y...KK.K...L...IR...DGDE.E.L.T...M.I.YD.NL.NNN...			82
02.AG.ES.06.P1423	MT.L.E...A...G...F.A...T...Y...RK.K...K...L...IR...GDT.E.L.T...I.V.DY.NL.NNI...			82
02.AG.GW.05.CC.0048	M.SL.A...A...G...F.A...T...Y...RKE...SN.L...IR...DGD.T.E.L.A...M...DN.NIL.NNN...			82

Table with columns: Vpu start, transmembrane domain, cytoplasmic domain, and Vpu end. It lists amino acid sequences for various HIV-1 proteins and domains, including gp120, gp41, p17, p24, p55, p66, p75, p97, p100, p105, p110, p120, p130, p150, p160, p170, p180, p190, p200, p210, p220, p230, p240, p250, p260, p270, p280, p290, p300, p310, p320, p330, p340, p350, p360, p370, p380, p390, p400, p410, p420, p430, p440, p450, p460, p470, p480, p490, p500, p510, p520, p530, p540, p550, p560, p570, p580, p590, p600, p610, p620, p630, p640, p650, p660, p670, p680, p690, p700, p710, p720, p730, p740, p750, p760, p770, p780, p790, p800, p810, p820, p830, p840, p850, p860, p870, p880, p890, p900, p910, p920, p930, p940, p950, p960, p970, p980, p990, p1000.

Env start	signal peptide end	gp120 start	glycosylation NVT	glycosylation NDT	CD4 binding	glycosylation NNS
B.FR.83.HXB2	MRVK...EKY	QHLWRWG.W.RWGTM.LLGM.LMICS...AT...E..KLVWTVYYG...VPVWKEATTLTFCASDAKAYTEVHNWVATHACVPTDPNPQEVLLVNVTFENFMWKNOMVEQMHEDIISLWDSQLPCKVCLT.PLCVSLKCTDLKNDTNTNS..				142
A1.CM.08.886.24	-K-RGIO.RNS...	-L-IFWT.II...	-D-E...H-N-EK-Q-I...	-L-M-D-D-E...I-D-R-R-Q-R...	-T-D-OPVNST...	136
A1.CN.08.CY236	--MGTO.MNW-G...	-I-I...VA.G.N...	-RD-E...M...	-D-S-T-D-N...	-T-N-SIITN...	136
A1.TY.09.NART.FLS.TVC19.1	--RGIO.MNS.HC...	-L-I...II...	-D-E...T-A-I...	-E-LN-T...	-T-D-RHNVTH-NVTN...	142
A1.KE.11.DEMA11KE002	--MGIO.MNW-X...	-GV-I...X-Y...	-D-E...K...	-D-K-T-T...	-T-N-VTVN-TNSVNNK...	142
A1.RU.11.DEUR6950	--MTE.RNC-N...	-L-T...IIF-N...	-D-K...K...	-D-N-T-G...	-T-H-EKVINHTMKE...	142
A1.RW.11.DEMA11RW002	-K-KGIO.RN-GQ...	-I-II...I-N...	-N-E...K...	-D-IM-Q...	-T-H-AGSNSTSSN...	132
A1.SN.01.DDI579	--MGIO.RN-LP...	-MI-S-C-II-K...	-D-E...N...	-D-N-T...	-T-N-NNITVNG...	148
A1.UG.11.DEMA11UG009	--RGIO.RNW-G...	-L-T...I-I-N...	-D-E...E-K...	-D-H-LD...	-T-N-IPKNFTNT...	140
A1.ZA.04.501.1344.T10.A1	--GMO.RNS...	-L-T...I-I...	-D-E...K...	-D-H-MH...	-T-S-NA-SGSTG...	141
A2.CD.97.97CDKTB48	-T-MGTO.RNC-KW...	-E-L-IF-I-M-K...	-RD-D...K...	-D-N-A...	-T-D-H-W-SN-N...	140
A2.CM.01.01CM.1445MV	-K-MGTO.RNC...	-L-IF-I-K...	-RD-E...A...	-D-N-A...	-T-N-SEV-DN...	135
A2.CY.94.94CY017.41	--MGTO.RN...	-G-L-I-I-M-K...	-D-D-I...	-D-N-I-N...	-T-N-SNANTS-HS...	141
B.BR.10.10BR.RJ032	--A-GMR.NNW-R...	-K-I...TA.G.O...	-A-M...S-I-E...	-D-D...	-T-N-SNANVSETRYN...	141
B.CA.07.502.1191.03	-GIR.RN-S-G-I...	-I-I...S-A.D.OG...	-K-G-A...S-M-E...	-E-A-SN...	-T-N-NANVSET-YAS...	148
B.CH.08.M2.0803101.NFLG8	-K-GTK.KN...	-I-I...S-A.D.O...	-K-G-A...S-M-E...	-E-A-SN...	-T-N-DLKN-NIS...	142
B.CN.12.DEMB12CN006	--TGIR.KNC-W...	-L-T...I-I...	-D-N...E-SS...	-D-N-T...	-T-N-W-SN-N...	140
B.CU.14.14CU005	--GDQE.LSS-W...	-L-I-V...-W-K...	-VTWVKGL-N...	-E-L-RDKC...	-T-N-N...-V-CS...	141
B.ES.14.ARP1195	-K-RGMK.RHW...	-TG-M-L...-V-D.O...	-D-N...H-L...	-D-N-Q-D...	-T-EI-Y-S-V...	142
B.FR.11.DEMB11FR001	-K-KTR...K...	-L-I-I...A.O...	-N...N...	-D-N-Q-D...	-T-N-G-N-A-IT...	138
B.HT.05.05HT.129389	-GIR.KN-W...	-T-L...A.M...	-N...E-A...	-D-N-N...	-T-N-KPV-C-KEEN...	140
B.JP.12.DEMB12JP001	-K-GTR.KNC...	-K-I...V-KD.N...	-N...E-A...	-D-N-N...	-T-N-AG-T-EAGN...	143
B.KR.07.HP.07H510.3909	--MEIR.KN-W...	-L-I...W-G-EM...	-N...I...	-D-N-N...	-T-N-SN-R-E-IKN...	140
B.RU.11.11RU21n	--MEIR.KN-W...	-L-I...W-G-EM...	-N...I...	-D-N-N...	-T-N-NVASTN-SS...	137
B.SE.12.SE600057	-K-GTR.RS-R-#...	-I-M...K-ER-S...	-N...RV-SK-A...	-D-N-I...	-T-N-WEKNNTSTN...	142
B.TH.10.DEMB10TH002	-K-EIR.KNC.R...	-I-I...K-E-E...	-V...K...	-D-N-N...	-T-H-CK-NDTKCHDLN...	141
B.US.13.RV.1	--TGIR.KNC.R...	-I-I...A.D.O...	-N...K...	-D-N-N...	-T-N-WS-S-I...	141
B.ZA.09.DEMB09ZA022	-GMR.KN-N...	-M-I...A.Q.O...	-K...K...	-D-N-N...	-T-N-VGNGTMIT...	140
C.BR.02.02YCG07R003	--MGTO.RNC-QW...	-I-IL-GFW-F-L-N...	-D-K...H-MI...	-D-D-OQ...	-T-N-NATSTPLCVTP...	136
C.BW.00.00BW0031.1	--MGTO.RNC-OW...	-I-IL-GFW-YN...	-D-K...E-K...	-D-D-OQ...	-T-N-SAAN-T...	141
C.CN.10.YNFI19	--REIL.RN-QW...	-I-VL-GFW-VG-G-NR...	-D-K...G-EK...	-D-N-V...	-T-N-NVT-VNSNTI...	136
C.CY.09.CY260	--MGIM.RNC-RW...	-I-IL-GFW-N-VG-D...	-K-A-E-R-ER...	-D-D-E...	-T-N-S-Y-GTH-NTT...	141
C.ES.14.ARP1198	--RGRL.RN-POW...	-I-SL-GFW-N-VG-G-N...	-K-ER...	-D-V-Q...	-T-N-NVHTN-T...	138
C.ET.02.02ET.288	--RGIO.RNC-OW...	-I-LL-GFW-F-N-VG-G-N...	-D-SP...	-D-V-E...	-T-H-SSNVTSSNVT...	141
C.FI.09.T125.2139	--MEL.RN-OW...	-I-IL-GFW-VG-G-N...	-N...K...	-D-V-E...	-T-H-GNVSUFSV...	142
C.KE.05.05KE369195V4	--MGIL.RNC-OW...	-I-IL-GFW-T-N-VG-G-N...	-HE...I...	-D-Q...	-T-KNVTNST...	138
C.MW.09.703010256.CH256.w96	--RGIO.RN-QW...	-I-SIL-AFW-V-VR-G-N...	-K...K...	-D-Q...	-T-N-SHNTVNG-MG...	140
C.SE.13.SE600311	--MGTO.RN-LOW...	-I-IL-GFW-L-SN-VG-G-N...	-N...K-M...	-D-Q...	-T-D-SNNTNINV-TV...	139
C.TZ.08.707010457.CH457.w8	-K-MEIL.RN-W-K-W...	-I-IL-GFW-L-SN-VG-G-N...	-R-K...ER...	-D-Q...	-T-R-ENATYNDTN...	139
C.US.11.17784.468	--RGMK.RN-OW...	-I-IL-GFW-N-VG-G-N...	-NPS...	-D-Q...	-T-D-NYNTSNT...	142
C.YE.02.02YES11	--TGIR.RN-QW...	-I-IL-GFWI-N-VG-G-N...	-K-ES...	-D-Q...	-T-N-NNNTNINV-D...	141
C.ZA.12.DEMC12ZA096	--TGTR.RN-QW...	-I-IL-GFW-V-M-N-MH-G...	-K-G-E...	-D-V...	-T-N-S-VT-S-STN...	140
C.ZM.11.DEMC11ZM006	--RGIL.RN-POW...	-I-IL-GFW-V-S-VR-G-N...	-K...E-K...	-D-N-E...	-T-S-NFTNSTIIT...	141
D.CM.10.DEMD10CM009	--ARGIK.RN...	-T-I-LL-TY-A-D-N...	-K...TKA-A-I...	-D-P...	-T-N-ETHLNSTQSV...	142
D.CY.06.CY163	--RGIK.RN-HN...	-K-L-TY-A-D-N...	-KA...VTKA-A-I...	-D-N-E...	-T-D-I-VETNNK-TV...	143
D.KE.11.DEMD11KE003	-A-GMK.RN...	-G-I-TY-A-D-N...	-S-KE-A-I...	-D-N-E...	-T-N-EV-DNSTNSN...	141
D.KR.04.04KR048	--RGMK.RN...	-I-W-VA-G-N...	-NSLKA-A-I...	-D-E...	-T-N-TR-K-IN-N...	141
D.SN.90.SE365	--AREMK.RN...	-I-VA-G-N...	-S-E-K-I...	-D-N-E...	-T-N-R-ISS-ATS-T...	141
D.TZ.01.A280	--METO.RN...	-I-W-TY-VA-Q-I...	-R-N-SF-A-I...	-D-N-E...	-T-H-S-ANTTNSG-G...	141
D.UG.10.DEMD10UG004	--REMKN-K-MS...	-I-T-V-G-O-N...	-R-K...KP-A-I...	-D-N-E...	-T-N-KWTSN-TS...	141
D.UG.11.DEMD11UG003	--RGIE.KN-R...	-G-L-I-VTY-VA-Q-N...	-R...S-E...	-D-N-E...	-T-N-EWEPNATHAN...	141
D.YE.02.02YES16	--ARGME.RNC-N...	-K-M-VA-G-N...	-KA...SVK-A-I...	-D-N-E...	-T-N-I-Y-VD-N...	141
D.ZA.90.R1	--ARGME.RNC-N...	-K-M-VA-G-N...	-KA...GKK-A-I...	-D-N-E...	-T-N-AIFNSTSF...	141
F1.A0.06.A0.06.AMG32	--MGMO.RNW-N...	-G-K-L-F-V-I-N-A-N...	-G-ER...E-ER...	-D-N-T...	-T-H-K-ANTSSTE...	140
F1.AR.02.ARE933	--RGMO.RNW...	-G-K-PL-F-I-I-N-NL...	-S-ER...S-ER...	-D-N-T...	-T-X-AIRN...	136
F1.BR.07.07BR844	--RGMO.RNW...	-G-K-L-F-I-I-N-K-N...	-K-A...K-A...	-D-N-T...	-T-N-NYN-TOYN-T...	142
F1.BR.10.10BR.PE107	--ARGMO.RNW...	-D-K-FL-F-I-I-N-D-N...	-K-A...Y...	-D-N-T...	-T-N-NAOSNATAWA...	141
F1.BR.10.10BR.RJ015	-K-RGMO.RNW...	-G-K-L-F-F-I-I-N-E-D-N...	-K-K-A...K-K-A...	-D-N-T...	-T-N-NAARN...	136
F1.CY.08.CY222	--RGMO.RNW...	-G-K-L-F-I-I-N-A-D-N...	-N-G-ER...N-G-ER...	-D-N-T-V...	-T-N-ARKSTEOSS...	140
F1.ES.02.ES.E845.4	--RGMO.RNW...	-G-K-L-F-I-I-N-A-D-N...	-N-G-ER...N-G-ER...	-D-N-T-V...	-T-N-S-ARDANSO...	140
F1.ES.11.VA0053.nfL	--REMO.RNW.PIW...	-G-K-L-F-I-I-K-E-N...	-E-EE...N-F...	-D-N-T...	-T-N-VNVTITNA-S-D...	141
F1.R0.96.BCI.R07	TK-RVMO.KNW-W...	-G-K-L-F-I-I-D-N...	-G-EK...ID...	-D-N-E...	-T-N-NANGTLISA-Q...	141
F1.RU.08.D88.845	--MGMO.RNW-W...	-G-K-L-F-I-I-N-A-N...	-R...E-A...	-D-N-S...	-T-N-NATIPNGN-N...	141
F2.CM.02.02CM.6016BBY	--RGMO.RNW...	-G-K-FL-F-I-I-N-A-D-N...	-E-K...E-K...	-D-N-D...	-T-H-S-VNITA-NTWQ...	142
F2.CM.10.DEMF210CM001	--MRMO.RNW...	-G-L-F-I-I-N-A-D-N...	-R-N-V-G-ER...	-D-N-Q...	-T-N-NVTITN-TYA...	142
F2.CM.10.DEMF210CM007	--RGMO.RNW...	-G-K-L-F-I-I-N-A-D-N...	-R-N-V-E-EK-A-Y...	-D-N-T...	-T-S-ANDTSAV-T...	142
G.CM.07.920.49	-A-GIO.RNW.HN...	-T-W-L-TF-L-VI-N-S-N-N...	-ED-D-P...S-K-I...	-D-IS-E...	-T-N-SNNTISGNSTI...	142
G.CM.10.DEMG10CM008	-GIO.RNW...	-T-L-IF-L-VI-S-D-N...	-ED-TD...S-R...	-D-LF-E...	-T-H-NGNLNLYTDAN...	141
G.CM.10.DEURF10CM020	-K-GIO.KNC-L...	-T-W-L-I-L-VI-TS-D-N...	-ED-D-P...S-K-I...	-D-LF-E...	-I-T-NVTCANR-TI...	141
G.CN.08.GX.2084.08	-K-RGIO.RNW...	-I-WML-I-L-VI-TS-D-N...	-ED-D-P...R-S-K...	-D-P-N...	-T-N-AVTCANA-HGT...	142
G.ES.09.XZ634.2	-K-GTO.KSM-PN...	-K-L-I-L-VI-S-N-N...	-ED-D...R-S-K...	-D-P-N...	-T-N-NVTTSSTSTN...	141
G.ES.14.ARP1201	-K-RGIO.RNW...	-T-W-L-I-L-VI-S-N-NX...	-ED-D...R-S-K...	-D-P-N...	-T-N-SNGNSTSH...	142
G.GH.03.03GH175G	-GIO.TNW...	-K-L-I-L-VI-S-D-N...	-ED-D...S-S-K...	-D-P-N...	-T-N-IYVSN-RK...	141
G.HE.09.DEMG09KE001	-A-EMK.RNW-R...	-EG-LL-I-L-VI-S-I-N-A...	-ED-D...S-K-I...	-D-P-N...	-T-N-NVTSTGIE-N...	141
G.NG.09.09NG.SC62	-K-GIO.KNW.PP...	-T-W-L-I-L-VI-N-A-D-N...	-ED-E-P...S-K-I...	-D-P-N...	-T-N-VT-NCTDA-YI...	141
G.PT.X.PT3306	-K-GIO.MNW...	-L-I-L-VI-S-N-N...	-ED-D...S-I-VTF...	-D-P-N...	-T-N-NAYANVTS...	140
H.BE.93.VI991	-T-METO.RN-PS...	-L-I-L-VI-VV-G-N...	-K...K...	-D-V...	-T-D-SSVATNV-K...	141
H.BE.93.VI997	-T-METO.RN-POW...	-G-L-L-Y-A-G-N...	-K...EP-K...	-D-T...	-T-D-SNTRMDTI-N...	140
H.BF.90.056	-T-METO.RN-PS...	-L-I-L-VI-A-N-N...	-K...E-K...	-D-T...	-T-N-NVR-N-SNS...	141
H.GB.00.00GBAC4001	-T-METO.RN-P...	-L-SG-L-I-LM-TG-N...	-K...S-K...	-D-T...	-T-E-GEVK-K-SNSNS...	142
J.CD.97.J.97CD.KTB147	TK-METO.MNW.KN...	-K-L-IF-N-G...	-D-K...S-K...	-D-I...	-T-N-EATTNNS-DN...	141
J.SE.93.SE9280.7887	-T-METO.KNW-T...	-G-L-IF-K-K-D...	-D-K...S-K...	-D-O-V-E...	-T-N-SNITSNS-T...	141
J.SE.94.SE9173.7022	-T-METO.TSW.LS...	-L-IF-K-R-N...	-RD-K...S-K...	-D-O-V-E...	-T-N-S-VNSNNS-D-N...	142
K.CD.97.97ZR.E0TB11	--AREIO.RNW...	-G-KR-IL-F-I-I-A-N-N...	-P...E-K...	-D-N-T...	-T-T-NVT-NRTNAN...	142
K.CM.96.96CM.MP535	--RGMO.RNW-T...	-G-N-IL-F-I-I-NA-D-N...	-P...E-K...	-D-N-T...	-T-N-Y-GTNS-NA...	141
01.AE.AF.07.569M	-GTO.RNG.PN...	-K-L-I-L-VI-S-D-N...	-RD-D...HE-I-L-VI...	-D-N-Q-V...	-T-N-SNANLTNIN-T...	141
01.AE.CM.11.11E.26	-K-GTO.MNW-L...	-K-L-II-L-VI-S-D-N...	-RD-N...OES...	-D-N-Q-V-E...	-T-N-NATSTNS...	138
01.AE.CM.12.DE0012CN011	--ETO.MNW.PN...	-K-L-I-L-VI-S-N-N...	-RD-D...HE-I-L-VI...	-D-N-Q-V...	-T-N-EA-PMA-LTE...	141
01.AE.HK.04.HK001	--ETO.MNW.PN...	-K-L-I-L-VI-S-D-N...	-D-D...HE-I-L-VI...	-D-N-Q-V...	-T-N-OANWKS-NTT...	141
01.AE.IR.10.10IR.THR48F	-K-ETO.RNW.LNW...	-K-L-I-L-VT-S-D-N...	-RD-E...OE-A...	-D-N-Q-V...	-T-N-KANLTKLA-A...	141
01.AE.JP.11.DE0011JP003	-K-ETO.MNW.PN...	-K-L-I-L-V-S-N-N...	-RD-V...HE-I-L-VI...	-D-N-Q-V...	-T-M-NVANNTVANAT...	142
01.AE.SE.11.SE601038	-K-GTO.MNW-TS...	-K-L-I-L-V-I-S-N-Y...	-D-D...HN-T...	-D-N-Q-V-R...	-T-M-HATFOT-S-P-T...	141
01.AE.TH.10.DEM110TH001	--ETO.MNW.PN...	-K-L-I-L-VI-S-N-N...	-RD-D...HS-S...	-D-N-Q-V...	-T-N-SATVITV...	141
01.AE.TH.90.CM240	--ETO.MNW.PN...	-K-L-I-L-VI-S-D-N...	-RD-D...HE-I-L-VI...	-D-N-Q-V...	-T-N-NANLTNGSSK...	141
01.AE.US.05.306163.FL	-GTO.MNW.PN...	-K-L-I-L-AI-S-N-N...	-RD-D...HE-I-L-VI...	-D-N-Q-V...	-T-N-F-VP...	136
02.AG.CM.10.DE00210CM013	-K-RGMK.KN-PL...	-L-MI-IFWI-II-N-E-N-N...	-D-K...E-E...	-D-N-Q...	-T-N-S-IFR5SSINN...	141
02.AG.ES.06.P1423	--RGIO.RN-PP...	-L-II-IFWI-II-N-O-N...	-RD-E...H-A...	-D-N-N...	-T-R-H-N-TNI-I...	139
02.AG.GW.05.CC.0048	--MGIO.KNW-PF...	-K-II-IFWI-II-N-R-N...	-D-K...E-A...	-D-N-N...	-T-E-SNWRNVTNGT...	141

Env start	signal peptide end	p120 start	glycosylation NVT	glycosylation NNS
B. FR. 83. HXB2	MRVK	...EKY QHLWRWG.W.RWGTM.LLGM.LMICS...AT...E.LKLVTVYYG...VPVWKEATTLTFCASDAKAYDETVHNVWATHACVPTDPNPQEVVLVNVNTEFNFMKNDOMEQMHEDIISLWQSLKPKCVKLT.PLCVSKLCTDLKNDTNTNS.	...RD-E...N-T...A...IO-E...D...N...V...T-D...HNFSSNSSTFEI	142
02 AG. KR. 12. 12MHR9	...RGIO. KN. PL	...I IFWI MI-N...E...N...NM...RD-E...E...EK...I...Y...S IP...K...G...N...T-D...HNFSSNSSTFEI	143	
02 AG. LR. x. POC44951	...MGIO. KN. PL	...L IFWI MI-N...E...N...NM...RD-E...N...T...A...IO-E...D...N...V...T-D...HNFSSNSSTFEI	144	
02 AG. NG. 09. 09NG SC61	T-METO. RNW. LL	...G-LI MFWI MI-N...O...P...I...R-E...E...VK...S...IH-E...K...N...D...E...T-E...MNFSSN-ST-A	140	
02 AG. NG. x. IBNG	...MGIO. KN. PL	...N IFWI MI-N...E...O...P...RD-E...T...E...MH-T...I...K...N...D...E...T-D...HNFN-SYS-A	141	
02 AG. SE. 11. 56602024	...GTO. RS. SL	...MI IFWI MI-N...N...E...O...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
02 AG. SN. 98. 98SE MP1211	...MGIO. KS. PL	...I IFWI MI-N...E...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
02 AG. US. 06. 502. 2696 FL01	...EIR. K.	...L F I L VI...S...N...N...RD-E...E...EK...I...Y...S IP...K...G...N...T-D...HNFN-SYS-A	137	
03 AB. RU. 97. KAL153 2	...MGMO. RN. P	...E L I L VI...S...N...N...RD-E...E...EK...I...Y...S IP...K...G...N...T-D...HNFN-SYS-A	140	
04 cpd. CY. 94. 94CY032 3	...RGMO. RNW. P	...G K LL F I...I...N...D...F...A ED-D...I...G...K...NM...RD-E...D...I...G...K...NM...RD-E...D...I...G...K...NM	141	
05 DF. BE. x. VT1110	...GTO. RS. SL	...MI IFWI MI-N...N...E...O...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
06 cpd. AU. 96. BFP90	...TGIK. KN. R	...V G K N M...S...K...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
07 BC. CN. 98. 98CN009	...RGTR. RN. QW	...I VL GFW...N...VE...G...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
08 BC. CN. 97. 97CNGX 6F	...EIO. RNW. N	...L I F...N...A...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
09 cpd. GH. 96. 96GH2911	...TGIO. RNC. ROW	...I L GFW...N...R...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
11 cpd. CN. 95. 95CN 1816	...RGMO. RNW. P	...G N FL F I...I...N...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
10 CD. T2. 96. 96T2 BF061	...ARGIO. MTW. N	...K L I L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
12 BF. AR. 99. ARMA159	...KA-GTO. RNW. S	...AL I L VI-SG...S...N...D...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	140	
13 cpd. CM. 96. 96CM 1849	...ETO. MNS. PN	...K L I L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
14 BG. ES. 05. X1870	...RGMO. R-S	...K L I L VI...S...K...T...I...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
15 01B. TH. 99. 99TH MU2079	...RGTR. RN. PS	...K I I L LM...F...D...H...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
16 AZD. FR. 97. 97RKO	...MGTR. RS. PP	...N L I L VIM...F...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
17 BF. AR. 99. ARMA038	...GTO. RNC. C	...T L II L GI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
18 cpd. CU. 99. CU76	...REIE. RNC. N	...K I I L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
19 cpd. CU. 99. CU7	...RGTO. MTW. N	...K L I L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
20 BG. CU. 99. CU103	...GTO. RNC. N	...K I I L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
21 AZD. KE. 99. KE92003	...RGTO. MTW. N	...K L I L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
22 01A1. CM. 01. 01CM 0001BBY	...GTO. RNC. N	...K I I L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
23 BG. CU. 03. CB118	...GTO. RNC. N	...K I I L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
24 BG. ES. 08. X2456 2	...ARGIO. MNW. L	...T L II L GI...S...D...N...I...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
25 cpd. CM. 02. 1918LE	...K-RGIO. KN. PL	...I L M L VI...S...D...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
26 AU. CD. 02. 02CD MBT047	...MGTR. KNC. C	...L K I I L VI-F-N...S...D...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
16 AZD. FR. 97. 97RKO	...K-METO. RNW. P	...K L I L VI...S...K...T...I...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
28 BF. BR. 99. BREPM12609	...RGMO. RNW. K	...G T LL F I...I...N...Q...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
29 BF. BR. 01. BREPM16704	...A-ETM. KT. W-S	...M...S...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
31 BC. BR. 04. 04BR142	...GTO. RNW. KOW	...I L GFW FL YN...VR...G...D...P...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	139	
32 06A1. EE. 01. EE0369	...RGTO. MNW. S	...K L I L VI...S...K...D...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
33 01B. ID. 07. ID0189 C	...ETO. RNC. SN	...K L I L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
34 01B. TH. 99. 01BTH2478P	...RGTO. MNW. LN	...K L I L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
35 AD. AF. 07. 169H	...MGTO. TNW. N	...K L I L VI...S...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
36 cpd. CM. 00. 00CMNYU830	...ARGTO. RNC. L	...K L II L VI...S...D...L...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	137	
37 cpd. CM. 00. 00CMNYU926	...MGIO. RN. PP	...N L I L VI...S...N...L...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	139	
38 BFI. UY. 03. UY03 3389	...RGMO. RNW. P	...E KCSLL F I...I...N...E...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	139	
39 BF. BR. 04. 04BR31379	...K-MGTX. MNW. LN	...G I L F I...I...N...A...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
40 BF. BR. 05. 05BRR1055	...GIR. KNC. P	...I...N...A...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
42 BF. LU. 06. LuBF 18 06	...MGIM. KN. Y	...K I I...S...A...E...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
43 02G. SA. 03. J11223	...EIO. RNW. Y	...TMMXL I L XI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
44 BF. CL. 08. CH80	...RGMO. RNW. P	...G K LL F I...I...N...A...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
45 cpd. FR. 04. 04FR AUK	...RGTO. RNW. W	...G K LL F I...I...N...A...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
46 BF. BR. 07. 07BR FPS625	...RGMO. RNW. W	...G K LL F I...I...N...A...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
47 BF. ES. 08. P1942	...AREIK. KN. QW	...KG LL I...I...N...EKS...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
48 01B. MY. 07. 07MYKT021	...ETO. GNW. LNW	...K L I L VV...S...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
49 cpd. GM. 03. N26677	...METW. MIC. #NW	...G L I...VA...G...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
50 A1D. GB. 10. 1272	...REIE. RN. W	...H...VA...Q...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	140	
51 01B. SG. 10. 10SG HM021	...GTO. RNC. N	...G I L F I...I...N...A...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
52 01B. MY. 03. 03MYK1018 1	...ETO. MNW. P	...K L I L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
53 01B. MY. 11. 11FIR164	...ETO. MNW. PN	...K L I L VI...S...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
54 01B. MY. 09. 09MYB8023	...GTO. MTW. PS	...T L M L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
55 01B. CN. 10. 10CNCS102056	...K-RVTO. MSW. LN	...T L I L VIM...S...D...N...I...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
56 cpd. FR. 10. 10FR5 patient A	...K-MGTX. MNW. LN	...K L I L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
57 BC. CN. 09. 09YNLX1959	...RGTL. RNC. OW	...I SIL GFW...S...N...VG...G...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
58 01B. MY. 09. 09MYPR37	...ETO. MNW. PN	...K L I L VA...L...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
59 01B. CN. 09. 09LNA423	...AMGIR. KNC. S	...K V...TA...Q...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
60 BC. IT. 11. BAV499	...RMO. RNW. ITW	...I L GFW...MYN...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	139	
61 BC. CN. 10. J1106010	...RGTL. KN. POW	...I L GFW...GN...GG...K...G...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
62 BC. CN. 10. YNFL13	...MGTR. KTC	...I...I...A...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
63 02A1. RU. 10. 10RU6637	...MGIV. KN. R	...I...FW...IIM...K...A...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
64 BC. CN. 09. YNFL31	...MGIL. RNC. OW	...I L GFW...I...N...VG...G...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
65 cpd. CN. 10. YNFL01	...RGTO. MNW. PN	...I L I L VI...S...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
67 01B. CN. 11. ANHUI HF115	...GTO. MXW. LN	...K L II L VI...S...D...N...I...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
68 01B. CN. 11. ANHUI WH73	...K-MKKK. NR. K	...G I LYI...VM...L...I...S...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
69 01B. JP. 05. 05JPMY113SP420	...ETO. MNW. PD	...K L I L VI...S...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
70 BFI. BR. 10. 10BR PE004	...RGMO. RNW. W	...G T SFL F I...V...N...TA...D...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
71 BFI. BR. 10. 10BR PE008	...MGMO. RNWQ. O	...G K LW F RT...L...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
72 BFI. BR. 10. 10BR MG002	...KAMGIR. RN. W	...K L I L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	142	
74 01B. MY. 10. 10MYPR268	...K-MV. MNW. PNW	...K L I L VI...S...N...N...RD-E...RD-E...E...S...MH-E...N...V...T-D...HNFN-SYS-A	141	
0. BF. 87. ANT70	...KAHEKR. NR. K	...I LYL...AMAL...ITP...L...S...R...O...YA...A...ED-PPV...NLTLS...Q...I...SO...T...Y...YP...H...DD...I...Y...Q...OM...F...OME...NIAGT...133	133	
0. CM. 98. 98CMA184	...KAMKKK. NR. K	...I LYL...AMAL...ITP...L...S...R...O...YA...A...ED-NPV...NLTSP...Q...I...AQ...H...YP...H...DR...DI...E...Y...Q...TD...E...SM...FM...OME...N...TN...133	133	
0. CM. 98. 98CMA184	...K-MKKK. NR. K	...GILCI...AL...IP...L...SS...R...D...YA...A...ED-EPV...NLTSP...Q...I...SO...T...H...YP...H...DK...I...Y...D...Q...DP...E...OM...FM...IOMN...F...TN...134	134	
0. CM. 98. 98CMA184	...MEKK. NR. K	...I LYL...AMAL...ITP...L...SS...S...EQ...YA...S...ED-KPV...NLTSP...Q...I...SO...N...YT...T...DS...I...Y...D...Q...DP...E...OM...F...IOMN...P...I...STG...N...133	133	
0. CM. 98. 98CMU5337	...K-MKKK. NR. K	...E ILYI...VM...L...ITP...L...SY...N...MYA...A...ED-EPV...NLTSP...Q...I...SO...T...H...YP...H...DK...I...Y...D...Q...DP...E...OM...F...IOMN...ETEG...138	138	
0. CM. 99. 99CMU4122	...KAHEKR. NR. O	...G ILYI...VM...L...ITP...L...SY...D...O...YA...S...ED-EPV...NLTSP...Q...I...SO...T...H...YP...H...DK...I...Y...D...Q...DP...E...OM...F...IOMN...NVNDT...133	133	
0. FR. 92. VAU. 92VAU10	...K-MKKK. NR. P	...G I LYL...VM...L...ITP...L...SY...N...O...YA...S...ED-EPV...NLTSP...Q...I...SO...T...H...YP...H...DK...I...Y...D...Q...DP...E...OM...F...IOMN...S...IT...134	134	
0. GA. 11. 11GAB6352	...KAHEKR. NR. Q	...I CI...VM...F...ITP...L...SY...N...O...YA...S...ED-SPV...NLTSP...Q...I...SO...T...Y...YP...H...DK...DI...Y...D...Q...DP...E...TO...S...F...RMN...PN...N...134	134	
0. SN. 99. 99SE MP1299	...K-MEQR. NR. K	...G I LCI...VMAL...ITP...L...SY...N...O...YA...A...ED-EPV...NLTSP...Q...I...SO...T...Y...YP...TK...D...I...Y...D...Q...DP...E...OM...F...IOMN...NVQGN...135	135	
0. US. 10. LTNP	...KAMERK. NR. RS	...ILCI...VMVL...ITP...L...TS...K...O...YA...I...A...ED-APV...NLTSP...Q...I...AQ...S...N...YT...H...DK...DI...Y...D...Q...DP...E...OM...FM...IOMN...HNTT...133	133	
N. CM. 02. DJ00131	V-MMGM. IGW. PFF	...C LMLISL...TI...GS...K...Y...A...RD-EV...HS...A...I...O...P...D...P...E...K...A...VO...E...E...TMM...SNSNGRNT...DEK...135	135	
N. CM. 02. SJGdd	...MGIO. SYW. SLV	...C LLLVS...IKVI...GS...OH...A...RD-EI...HS...A...I...O...P...L...P...N...A...O...E...E...TMM...SNSNGEPVDDL...137	137	
N. CM. 04. 04CM 1015 04	...MGMO. SGW. PFF	...C LLLVS...II...GS...OH...A...RDTE...V...HS...A...I...O...P...T...H...D...N...A...O...E...E...TMM...XNSNGXIX...VDN...135	135	
N. CM. 06. U4293	...GIMK. RNL. PFF	...C LLLVS...II...Y...E01...GS...OH...A...RD-EI...HS...A...I...O...P...T...H...D...N...A...O...E...E...TMM...XNSNGXIX...VDN...135	135	
N. FR. 11. N1 FR 2011	...MGML. SYW. SLF	...Y LLLXL...IKVI...GS...OY...RD-EI...X...HS...A...I...O...P...D...P...E...K...A...O...V...E...E...TMM...SDG...GTNSTI...137	137	
P. CM. 06. U14788	...K-MKKK. KN. S	...G NLEI...C...AL...IMYFNALSYTS...G...TRY...RD-EI...A...SLTSX...Q...I...O...R...I...R...D...S...I...D...Y...T...Q...V...TMM...S...CSTVDC...NS...143	143	
P. FR. 09. RBF168	...AMKKK. KN. S	...G NLEI...C...AL...VMYFNALSYTS...G...IHY...RD-EI...A...SLTSX...Q...I...O...T...I...I...PI...S...I...Y...T...Q...V...TMM...SECHREASC...140	140	
CPZ. CD. 90. ANT	...KKMRKP. IHI. IWG	...L A LI OF-EK...G...N...EDY...F...RD-NP...TN...SMTS...TS...I...D...IV...R...TSWV...AY...Y...S...T...MXQ...FQ...H...M...IKMN...GYNGTPT...P...T...135	135	
CPZ. CM. 05. SIVcpzLB7	...GTW. TSC. MPY	...Y VV L I L CS...I...S...TKN...N...A...EA...OKA...A...I...SO...D...KAIE...S...T...A...N...Q...V...I...T...H...SIP...F...NSS...140	140	
CPZ. CM. 09. Ptt 09cam155	...AMERT. KRIS. WNI	...L SRSM...TL...CLTFN...VA...T...EY...A...D...E...PHSS...A...I...O...I...E...F...T...S...D...P...S...E...E...LVFL...TNS...STAEFKN...TN...142	142	
CPZ. T2. 06. TAN5	...KMRNL. IGI. T	...L T II...TTLGI...VF...S...QYIA...F...I...D...KPS...DVTSDRE...I...N...L...Y...P...A...M...VD...E...Y...QE...K...L...FQ...F...I...HPS...NLTSP...TVSPA...134	134	
CPZ. US. 05. US Mar15yn	...K-MEQR. RNL. SY	...CLLSI...LV...VA...S...A...RD-EI...HS...A...I...O...P...D...P...E...K...A...O...V...E...E...TMM...SDG...GTNSTI...137	137	
GOR. CM. 07. SIVgor 2139 287	...AMRKK. QL. N	...K VM...S...ILT...I...ES...R...THYA...RD-NV...A...NVYSK...O...I...O...R...T...FP...E...T...I...Y...D...Q...V...VM...T...N...O...NTNTT...OTE...138	138	
GOR. CM. 12. SIVgor B03D	...K-MKKK. KE. KS	...K LIXI...VMAL...ITP...L...SS...R...P...YA...RD-O...A...MASS...M...O...Q...I...LK...T...I...SG...D...VM...T...N...XRPXT...T...H...139	139	
GOR. CM. 13. SIVgor_BP1D15	...MRKK. KR. S	...G N L...S...AL...I...FNTGSVYM...G...TRY...RD-K...A...NLASK...O...I...O...T...IK...K...IS...PY...I...E...YI...K...QKN...LN...VI...TIN...SKF...PA...TNTTS...143	143	

glycosylation NDT

CD4 binding glycosylation NNS

Env

HIV-1/SIVcpz Proteins

	glycosylation NIS		glycosylation NCS		glycosylation NDT		glycosylation NTS		glycosylation NGT		252
	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2	
B. FR. 83. HXB2	S	S	S	S	S	S	S	S	S	S	252
02 AG. KR. 12. 12MHR9											254
02 AG. LR. x. POC44951											245
02 AG. NG. 09. 09NG SC61											252
02 AG. NG. x. IBNG											249
02 AG. SE. 11. 56602024											257
02 AG. SN. 98. 98SE MP1211											246
02 AG. US. 06. 502. 2696 FL01											247
03 AB. RU. 97. KAL153. 2											243
04 CPX. CY. 94. 94CY032. 3											257
05 OF. BE. x. VT119											249
06 CPX. AU. 96. BFP90											263
07 BC. CN. 98. 98CN009											256
08 BC. CN. 97. 97CN6X. 6F											256
09 CPX. GH. 96. 96GH2911											247
10 CD. TZ. 96. 96TZ. BF061											243
11 CPX. FR. 95. 95CM. 1816											263
12 BF. AR. 99. ARMA159											254
13 CPX. CM. 96. 96CM. 1849											255
14 BG. ES. 05. X1870											257
15 01B. TH. 99. 99TH. MU2079											243
16 AZD. KR. 97. 97KR00											252
17 BF. AR. 99. ARMA038											254
18 CPX. CU. 99. CU76											252
19 CPX. CU. 99. CU7											225
20 BG. CU. 99. CU103											250
21 AZD. KE. 94. KE2003											241
22 01A1. CM. 01. 01CM. 0001BBY											245
23 BG. CU. 03. CB118											258
24 BG. ES. 08. X2456. 2											262
25 CPX. CM. 02. 1918E											252
26 AU. CD. 02. 02CD. MBT047											252
27 CPX. FR. 04. 04FR. K25											250
28 BF. BR. 99. BREPM12609											255
29 BF. BR. 01. BREPM16704											258
31 BC. BR. 04. 04BR142											244
32 06A1. EE. 01. EE0369											261
33 01B. ID. 07. JKT189. C											257
34 01B. TH. 99. 99TH. BR2478P											248
35 AD. AF. 07. 169H											254
36 CPX. CM. 00. 00CMNYU830											250
37 CPX. CM. 00. 00CMNYU926											248
38 BF1. UY. 03. UY03. 3389											248
39 BF. BR. 04. 04BRR3179											254
40 BF. BR. 05. 05BRR1055											248
42 BF. LU. 06. LuBF. 18. 06											263
43 02G. SA. 03. J11223											250
44 BF. CL. 08. CH80											253
45 CPX. FR. 04. 04FR. AUK2											248
46 BF. BR. 07. 07BR. FPS625											247
47 BF. ES. 08. P1942											247
48 01B. MY. 07. 07MYKT021											266
49 CPX. GM. 03. N26677											255
50 A1D. GB. 10. 1272											251
51 01B. SG. 11. 11SG. HM021											254
52 01B. MY. 03. 03MYK1018. 1											264
53 01B. MY. 11. 11F1R164											261
54 01B. MY. 09. 09MYSR023											257
55 01B. CN. 10. HNC5102056											249
56 CPX. FR. 10. URF5. patient. A											248
57 BC. CN. 09. 09BNCX1395g											257
58 01B. MY. 09. 09MYPR37											266
59 01B. CN. 09. 09LNA423											251
60 BC. IT. 11. BAV499											256
61 BC. CN. 10. J1100010											249
62 BC. CN. 10. YNFL13											251
63 02A1. RU. 10. 10RU6637											246
64 BC. CN. 09. YNFL31											254
65 CPX. CN. 10. YNFL01											265
67 01B. CN. 11. ANHUI. HF115											252
68 01B. CN. 11. ANHUI. WF73											249
69 01B. JP. 05. 05JPMYC1135P420											262
70 BF1. BR. 10. 10BR. PE004											250
71 BF1. BR. 10. 10BR. PE008											254
72 BF1. BR. 10. 10BR. MG002											249
74 01B. MY. 10. 10MYPR268											259
0. BF. 87. AN770											244
0. CM. 98. 98CMA184											249
0. CM. 98. 98CMA8B141											267
0. CM. 98. 98CMA8B212											263
0. CM. 98. 98CMA5337											251
0. CM. 99. 99CMU4122											249
0. FR. 92. VAU. US. Mar3Lyn											257
0. GA. 11. 11GAb6352											251
0. SN. 99. 99SE. MP1299											253
0. US. 10. L2NP											238
N. CM. 02. DJ00131											248
N. CM. 02. SJGdd1											246
N. CM. 04. 04CM. 1015. 04											246
N. CM. 06. U429											246
N. FR. 11. N1. FR. 2011											251
P. CM. 06. U41788											261
P. FR. 09. RBF168											257
CPZ. CD. 90. ANT											254
CPZ. CM. 05. SIVcpzLB7											246
CPZ. CM. 09. Ptt. 09Cam155											257
CPZ. TZ. 06. TAN5											258
CPZ. US. 05. US. Mar3Lyn											247
GOR. CM. 07. SIVgor2139. 287											261
GOR. CM. 12. SIVgor. B01D2											267
GOR. CM. 13. SIVgor. BPID15											262

	glycosylation NGS	glycosylation NNT	glycosylation NTS	glycosylation NNT	glycosylation NNT	glycosylation NNT	glycosylation NKT	CD4 binding																													
B.FR.83.HXB2	PVVSTQLLLNGLAE.E	EVVIRSVNFTDNAAKTIIVQLNLS	VEINCTRPNMN	TRKRIRIQRGPGRAVFT	IG	KIGNMRQAHCNISRAKWNNTLQIASKLR	EOFG	NN	K	TIIFKQ	S.S	GGDPEIVTHSFNCGGE	381																								
A1.CD.97.97CD.KC2	G.GIK	A.ISY	N	DIP	K	S	TSV	QT	YA	T	DI	I	L	TA	D	YNSVKA	H	P				NK	L	VT	M	I		379									
A1.CM.08.886.22	Q.K.M	E	I	T	S	V	S	RS	05	YA	T	EI	DI	E	R	VNGTE	KA	QKVVI	H				K	NS	L	T	M	A	373								
A1.CY.08.CY236	N.RTM	E	I	N	N	TEP	N	T	I	S	S	0	YA	T	DI	I	K	EVNGTE	KA	GVVEO	TH	S		A	ND	L	T		385								
A1.NC.09.NARI.FLS.TVC19.1	N.G.T	K	E	I	N	TEP	N	T	I	S	TSV	QT	YA	T	DI	DI	V	SA	K	HDVSKO	TH	X	X	K	AN	L	T		R	393							
A1.KE.11.DEMA11KE002	G.K	E	I	N	N	REP	N	T	I	S	SV	Q	FA	T	DI	K	Y	V	GV	K	QKVVEK	K	Y	S	MM	A	L	T		384							
A1.RU.10.RUR06950	K.M	E	I	N	N	TE	S	T	G		TSV	QT	YA	T	DI	DVX	DI	R	Y	TE	S	QK	ST	R	H					385							
A1.RW.11.DEMA11RW002	K.I	E	I	N	N	VKP	IG	S			SV	Q	YA	T	EI	DI	E	FEA	N	SKO	G	K	R							382							
A1.SN.01.DD1579	K.M	E	I	N	N	VNP	Q	S			TSV	Q	YA	T	DI	DI	V	EN	TE	Q	V	KO	K	R						380							
A1.UG.11.DEMA11UG009	VK.G	R	R	K	E	ISE	H	N	V	TKP	S	R	I	T	TSV	Q	E	F							K	TG	V	L	T	376							
A1.ZA.04.5612344.T10.A1	G	TMV	E	I	N	I	D	A	S	I	G	SV	Q	YA	T	DI	K	Y	V	ED	T	QK	KO	S	K	Y	E			384							
A2.CD.97.97CD.KTB48	A	G	K	E	I	N	N	N	N		SV	Q	FA	T	DI	DI	K	Y	V	GV	K	QKVVEK	K	Y	S	MM	A	L	T		384						
A2.CM.01.01CM.1445MV	A	G	K	E	I	N	N	N	N		SV	Q	FA	T	DI	DI	K	Y	V	GV	K	QKVVEK	K	Y	S	MM	A	L	T		384						
A2.CY.94.94CY017.41	A	G	K	E	I	N	N	N	N		SV	Q	FA	T	DI	DI	K	Y	V	GV	K	QKVVEK	K	Y	S	MM	A	L	T		384						
B.BR.10.10BR.RJ032	D.I	E	I	N	NRN	KEP	IK				GVH	K	Y	T	EI	DI	L	K	D	E	K	Y	QVNV		E	A	D	P	A	M	393						
B.CA.07.502.1191.03	S	I	K	Y	SSN	N	H	E			G	S	L	IYA	T	DX	DI	G	O	R	R	K	V	I	N		D	P	H	T	M	390					
B.CH.08.M2.0803101.NFLG8	T	S									G	H	L	FA	T	DI	DI	T	NG	Q	A	Q	V	A	G	K	E	P	P	L	MM	403					
B.CN.12.DEMB12CM006	T	I	K	DI	V	E	N	K	T		QI	DI	V	AKN	T	DI	V	AKN	R	V	R	K	D	E		G	K	E	P	P	L	MM	385				
B.CU.14.14CU065	T	I	K	DI	V	E	N	K	T		QI	DI	V	AKN	T	DI	V	AKN	R	V	R	K	D	E		G	K	E	P	P	L	MM	385				
B.ES.14.ARP1195	V	LS									Q	N	A	WYA	AT	DI	DI	R	L	T		VVN	G	K	K	K		S	OP		I	T	380				
B.FR.11.DEMB11FR001	G	DTIL	E	ISN	E	H	E				M	G	H	YA	T	EI	DI	L	T	D		RS	I	K		R		N				367					
B.HT.05.05HT.129389	G	DTIL	E	ISN	E	H	E				M	G	H	YA	T	EI	DI	L	T	D		RS	I	K		R		N				368					
F1.JP.12.DEMB12JP000	H	K	L	K	N	E					RS	HL	OT	LYA	T	DI	DI	L	T	D		RS	I	K		R		N				390					
B.KR.07.HP.07H510.3909	Q										RS	HL	OT	LYA	T	DI	DI	L	T	D		RS	I	K		R		N				379					
B.RU.11.11RU21N	I	L	T								G	HM	K	YA	T	DI	DI	V	K	T	I	VVG		K		L	N	TK				369					
B.SE.12.SE600057	H	L	T								G	HM	K	YA	T	DI	DI	V	K	T	I	VVG		K		L	N	TK				369					
B.TH.10.DEMB10TH002	DI	L	K	S							G	HM	K	YA	T	DI	DI	V	K	T	I	VVG		K		L	N	TK				388					
B.US.13.RV.1	DI	L	K	S							G	HM	K	YA	T	DI	DI	V	K	T	I	VVG		K		L	N	TK				391					
B.ZA.09.DEMB09ZA022	L	E	N	V							S	PM	Y	T	DI	DI	L	S		V	K	Y	R		N	G	P				393						
C.BR.02.02YC07BR003	II	O	I	V							G	Y	T	AV	DI	DI	G	E	R	E	V	K	Y	R		N	G	P				371					
C.BW.00.00BW003.1	K	II	O	I	V						G	Y	T	AV	DI	DI	G	E	R	E	V	K	Y	R		N	G	P				403					
C.CN.10.YNFI19	G	II	O	I	V						G	Y	T	AV	DI	DI	G	E	R	E	V	K	Y	R		N	G	P				374					
C.CY.09.CY260	G	II	O	I	V						G	Y	T	AV	DI	DI	G	E	R	E	V	K	Y	R		N	G	P				389					
C.ES.14.ARP1198	G	II	O	I	V						G	Y	T	AV	DI	DI	G	E	R	E	V	K	Y	R		N	G	P				383					
C.ET.02.02ET.288	II	O	I	V							G	Y	T	AV	DI	DI	G	E	R	E	V	K	Y	R		N	G	P				368					
F1.FN.09.T125.2139	I	D	II	O	I	V					OS	OT	YA	T	EI	DI	G	E	R	E	V	K	Y	R		N	G	P				387					
C.KE.05.05KE369195V4	K	DI	A	L	N	V					OS	OT	YA	T	EI	DI	G	E	R	E	V	K	Y	R		N	G	P				388					
C.MW.09.703010256.CH256.w96	K	DI	A	L	N	V					OS	OT	YA	T	EI	DI	G	E	R	E	V	K	Y	R		N	G	P				382					
C.SE.13.SE600311	I	S	DI	A	L	N	V				RSV	L	OT	YA	T	DI	DI	D	N	VRG	K											380					
C.TZ.08.707810457.CH457.w8	I	S	DI	A	L	N	V				RSV	L	OT	YA	T	DI	DI	D	N	VRG	K											373					
C.US.11.17784.468	I	N	G	II	O	I	V				SM	OT	YA	T	EI	DI	K		S	E	ER	SN	E	H	P							381					
C.YE.02.02YE511	I	N	G	II	O	I	V				SM	OT	YA	T	EI	DI	K		S	E	ER	SN	E	H	P							380					
C.ZA.12.DEMC12ZA096	T	K	I	E	L	V					SV	Q	YA	T	DI	DI	Y	T	NG	O	A	Q	V	A	G	K	E	P	P	L	MM	380					
C.ZM.11.DEMC11ZM006	G	DI	E	M	T	N	H	E			E	V	SV	Q	YA	T	DI	DI	V	E	DE	HRV	R	Y	P	R	N	T	S	P	A	L	VT	R	380		
D.CM.10.DEMD10CM009	II	O	I	V							SS	H	SLY	T	DI	DI	N	Y	V	TN	H	V	E	G	K	L	L	N	K	T	K	EP	A	A	TA	372	
D.CY.06.CY163	D	II	O	I	V						DS	H	L	LYA	T	DI	DI	N	Y	V	TN	H	V	E	G	K	L	L	N	K	T	K	EP	A	A	TA	382
D.DE.11.DEMD11KE003	K	DI	E	ISS	N						SVHM	Q	LYA	T	DI	DI	N	Y	V	TN	H	V	E	G	K	L	L	N	K	T	K	EP	A	A	TA	385	
D.KR.04.04KR84	K	DI	E	ISS	N						SVHM	Q	LYA	T	DI	DI	N	Y	V	TN	H	V	E	G	K	L	L	N	K	T	K	EP	A	A	TA	371	
D.N9.S9.SE365	K	DI	E	ISS	N						SVHM	Q	LYA	T	DI	DI	N	Y	V	TN	H	V	E	G	K	L	L	N	K	T	K	EP	A	A	TA	384	
D.TZ.01.A280	G	II	O	I	V						Q	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	380			
D.UG.10.DEMD10UG004	M										Q	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	372				
D.UG.11.DEMD11UG003	I	IV	TK	L	T	I	KE				S	HM	A	Y	T	DI	DI	V	TA	Q	V	A	G	N	L	N	E	T	EP				379				
F2.CM.02.02CM.6016BB9	F										Q	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	378				
D.ZA.90.R	II	O	I	V							Q	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	385				
F1.A0.06.A0.06.AMG32	II	O																																			

Table of HIV-1 protein sequences with annotations for glycosylation sites (CD4 binding, glycosylation NNT, glycosylation NKT, V3 tip, V3) and other features. Includes columns for protein names (e.g., B_FR.83.HXB2, 02_AG.KR.12.12MHR9) and their corresponding amino acid sequences.

glycosylation NST

glycosylation NST

V4

glycosylation NST

glycosylation NNT

CD4 binding

glycosylation NIT

V5

glycosylation NES

CD4 binding

Table of HIV-1 protein sequences with annotations for glycosylation sites (NST, NNT, NIT, NES) and CD4 binding sites. The table lists various HIV-1 proteins and their amino acid sequences, with specific residues marked for glycosylation and CD4 binding. The proteins include Gag, Pol, and Env (Gp120, Gp41).

	gp120 end	gp41 start	fusion peptide	immunodominant region	glycosylation NAS	glycosylation NKS	glycosylation NHT	
B.FR.83.HXB2	AKRRVV	Q...REKRA...V.G.IGALFLGLGAAGSTMGAA.S.MTLTVQARQLLSGVQQQNLRLRAIEAQHLLQLTVWGIKQLQARLAVERYLKDQQLGIWGC	SGKLICTTAVPWNASMS.NKS.....	LEQIWNHTTWMEWDREIN	636		
A1.CD.97.97CD.KCC2	635	
A1.CM.08.886.24	635	
A1.CY.08.CY236	635	
A1.TN.09.NART.FLS.TVC19.1	635	
A1.KE.11.DEM111KE002	636	
A1.RU.11.DERU6950	636	
A1.RW.11.DEM111RW002	644	
A1.SN.01.DDI579	634	
A1.UG.11.DEM110UG009	627	
A1.ZA.04.501.12344.T10.A1	638	
A2.CD.97.97CDKTB48	635	
A2.CM.01.01CM.1445MV	635	
A2.CY.94.94CY017.41	638	
B.BR.10.10BR.RJ032	652	
B.CA.07.502.1191.03	642	
B.CH.08.M2.0803101.NFLG8	665	
B.CN.12.DEM12CN006	642	
B.CU.14.14CU005	635	
B.ES.14.ARP1195	644	
B.FR.11.DEM11FR001	620	
B.HT.05.05HT.129389	640	
B.JP.12.DEM12JP001	644	
B.KR.07.HP.13.07H510.3909	637	
B.RU.11.11RU21n	622	
B.SE.12.SE600057	643	
B.TH.10.DEM110TH002	646	
B.US.13.RV.1	646	
B.ZA.09.DEM09ZA022	628	
C.BR.07.07C07BR003	621	
C.BW.00.00BW0531.1	641	
C.CN.10.YNFL19	632	
C.CY.09.CY260	613	
C.ES.14.ARP1198	630	
C.ET.02.02ET.268	630	
C.IN.09.T125.2139	627	
C.KE.05.05KE369195V4	619	
C.MW.09.703010256.CH256.w96	626	
C.SE.13.SE600311	625	
C.TZ.08.707810457.CH457.w8	618	
C.US.11.11US.468	627	
C.YE.02.02YE511	635	
C.ZA.12.DEM12ZA096	628	
C.ZM.11.DEM11ZM006	636	
D.CM.10.DEM10CM009	624	
D.CY.06.CY163	638	
D.KE.11.DEM11KE003	630	
D.KR.04.04KR048	620	
D.SN.90.SE365	630	
D.TZ.01.A280	632	
D.UG.10.DEM10UG004	627	
D.UG.11.DEM11UG003	637	
D.YE.02.02YE516	636	
D.ZA.90.RJ.1	638	
F1.A0.06.A0.06.AMG32	620	
F1.AR.02.ARE933	625	
F1.BR.07.07BR844	629	
F1.BR.10.10BR.PE107	621	
F1.BR.10.10BR.RJ015	627	
F1.CY.08.CY222	624	
F1.ES.02.ES.X845.4	622	
F1.ES.11.VA0053.nfl	626	
F1.R0.96.BCI.R07	643	
F1.RU.08.D88.845	642	
F2.CM.02.02CM.6016BBY	625	
F2.CM.10.DEMF210CM001	635	
F2.CM.10.DEMF210CM007	627	
G.CM.07.920.49	621	
G.CM.10.DEMG10CM008	627	
G.CM.10.DEURF10CM020	646	
G.CN.08.GX.2084.08	631	
G.ES.09.X2634.2	651	
G.ES.14.ARP1201	651	
G.GH.03.03GH175G	635	
G.KE.09.DEMG09KE001	632	
G.NG.09.09NG.SC62	629	
G.PT.x.PT3306	645	
H.BE.93.VI991	642	
H.BS.93.VI997	625	
H.CF.90.056	628	
H.GB.00.00GBAC4001	635	
J.CD.97.J.97DC.KTB147	636	
J.SE.93.SE9280.7887	631	
J.SE.94.SE9173.7022	633	
K.CD.97.97ZR.E0TB11	631	
K.CM.96.96CM.MP535	622	
01.AE.AF.07.569M	631	
01.AE.CM.11.1156.26	632	
01.AE.CN.12.DE00112CN011	639	
01.AE.HK.04.HK001	632	
01.AE.IR.10.10IR.THR48F	640	
01.AE.JP.11.DE00111JP003	633	
01.AE.SE.11.SE601018	625	
01.AE.TH.10.DE00110TH001	637	
01.AE.TH.90.CM240	630	
01.AE.US.05.306163.FL	609	
02.AG.CM.10.DE00210CM013	657	
02.AG.ES.06.P1423	635	
02.AG.GW.05.CC.0048	640	

	gp120 end	gp41 start	fusion peptide	immunodominant region	glycosylation NAS	glycosylation NKS	glycosylation NHT											
B. FR. 83. HXB2	AKRRVVQREKRAV. G. IGALFLGLGAGSTMGAAAS. MTLTVQARQLLSGVQQQNLRLRAIEAQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGC	SGKLICTTAVPWASMS. NKSLEQIWNHTTWMEWDREIN	636										
02 AG. KR. 12. 12MHR9RGL-VLLSKV-L-G-RTY-N-GMM-LQ-K-S	642								
02 AG. LR. x. POC44951KGL-V-IISKRTYND-DMM-LQ-K-S	625								
02 AG. NG. 09. 09NG SC61RGL-VISQRTYND-DMM-LQ-K-S	637								
02 AG. NG. x. IENGSGL-VISKRTYND-DMM-LQ-K-S	627								
02 AG. SE. 11. S602024SSLKA. V. VVVFVW-RS-I-S-IMM-D-Y-CY-N-VK-SFF-SL-RSN-TYDD-QMM-LQ-K-S	634					
02 AG. SN. 98. 98SE MP1211RGL-VISKRTYDD-QMM-LQ-K-S	629								
02 AG. US. 06. 502 2696 FL01RGL-VISKRTYDD-QMM-LQ-K-S	629								
03 AB. RU. 97. KAL153 2RGL-VISKRTYDD-QMM-LQ-K-S	629								
04 cpx. CY. 94. 94CY032 3RGL-VISKRTYDD-QMM-LQ-K-S	629								
95 DF. BE. x. VT110RGL-VISKRTYDD-QMM-LQ-K-S	629								
06 cpx. AU. 96. BFP90RGL-VISKRTYDD-QMM-LQ-K-S	629								
07 BC. CN. 98. 98CN009RGL-VISKRTYDD-QMM-LQ-K-S	629								
08 BC. CN. 97. 97CNGX 6FRGL-VISKRTYDD-QMM-LQ-K-S	629								
09 cpx. GH. 96. 96GH2911RGL-VISKRTYDD-QMM-LQ-K-S	629								
10 CD. TZ. 96. 96TZ BF061RGL-VISKRTYDD-QMM-LQ-K-S	629								
11 cpx. FR. 95. 95CM 1816RGL-VISKRTYDD-QMM-LQ-K-S	629								
12 BF. AR. 99. ARMA159RGL-VISKRTYDD-QMM-LQ-K-S	629								
13 cpx. CM. 96. 96CM 1849RGL-VISKRTYDD-QMM-LQ-K-S	629								
14 BG. ES. 05. X1870RGL-VISKRTYDD-QMM-LQ-K-S	629								
15 01B. TH. 99. 99TH MU2079RGL-VISKRTYDD-QMM-LQ-K-S	629								
16 AZD. KR. 97. 97KR00RGL-VISKRTYDD-QMM-LQ-K-S	629								
17 BF. AR. 99. ARMA038RGL-VISKRTYDD-QMM-LQ-K-S	629								
18 cpx. CU. 99. CU76RGL-VISKRTYDD-QMM-LQ-K-S	629								
19 cpx. CU. 99. CU7RGL-VISKRTYDD-QMM-LQ-K-S	629								
20 BG. CU. 99. CU103RGL-VISKRTYDD-QMM-LQ-K-S	629								
21 AZD. KE. 99. KER2003RGL-VISKRTYDD-QMM-LQ-K-S	629								
22 02A1. CM. 01. 01CM 0001LBYRGL-VISKRTYDD-QMM-LQ-K-S	629								
23 BG. CU. 03. CB118RGL-VISKRTYDD-QMM-LQ-K-S	629								
24 BG. ES. 08. X2456 2RGL-VISKRTYDD-QMM-LQ-K-S	629								
25 cpx. CM. 02. 1918LERGL-VISKRTYDD-QMM-LQ-K-S	629								
26 AU. CD. 02. 02CD MBT047RGL-VISKRTYDD-QMM-LQ-K-S	629								
27 cpx. FR. 04. 04FR K25RGL-VISKRTYDD-QMM-LQ-K-S	629								
28 BF. BR. 99. BREPM12609RGL-VISKRTYDD-QMM-LQ-K-S	629								
29 BF. BR. 01. BREPM16704RGL-VISKRTYDD-QMM-LQ-K-S	629								
31 BC. BR. 04. 04BR142RGL-VISKRTYDD-QMM-LQ-K-S	629								
32 06A1. EE. 01. EE0369RGL-VISKRTYDD-QMM-LQ-K-S	629								
33 01B. ID. 07. JKT489 CRGL-VISKRTYDD-QMM-LQ-K-S	629								
34 01B. TH. 99. 01R2478PRGL-VISKRTYDD-QMM-LQ-K-S	629								
35 AD. AF. 07. 169HRGL-VISKRTYDD-QMM-LQ-K-S	629								
36 cpx. CM. 00. 00CMNYU830RGL-VISKRTYDD-QMM-LQ-K-S	629								
37 cpx. CM. 00. 00CMNYU926RGL-VISKRTYDD-QMM-LQ-K-S	629								
38 BFI. UY. 03. UY03 3389RGL-VISKRTYDD-QMM-LQ-K-S	629								
39 BF. BR. 04. 04BR3179RGL-VISKRTYDD-QMM-LQ-K-S	629								
40 BF. BR. 05. 05BRR1055RGL-VISKRTYDD-QMM-LQ-K-S	629								
42 BF. LU. 06. LuBF 18 06RGL-VISKRTYDD-QMM-LQ-K-S	629								
43 02G. SA. 03. J11223RGL-VISKRTYDD-QMM-LQ-K-S	629								
44 BF. CL. 08. CH80RGL-VISKRTYDD-QMM-LQ-K-S	629								
45 cpx. FR. 04. 04FR AUKRGL-VISKRTYDD-QMM-LQ-K-S	629								
46 BF. BR. 07. 07BR FPS625RGL-VISKRTYDD-QMM-LQ-K-S	629								
47 BF. ES. 08. P1942RGL-VISKRTYDD-QMM-LQ-K-S	629								
48 01B. MY. 07. 07MYKT021RGL-VISKRTYDD-QMM-LQ-K-S	629								
49 cpx. GM. 03. N26677RGL-VISKRTYDD-QMM-LQ-K-S	629								
50 A1D. GB. 10. 1272RGL-VISKRTYDD-QMM-LQ-K-S	629								
51 01B. SG. 10. 1156 HM021RGL-VISKRTYDD-QMM-LQ-K-S	629								
52 01B. MY. 03. 03MYKL018 1RGL-VISKRTYDD-QMM-LQ-K-S	629								
53 01B. MY. 11. 11FIR164RGL-VISKRTYDD-QMM-LQ-K-S	629								
54 01B. MY. 09. 09MYSB023RGL-VISKRTYDD-QMM-LQ-K-S	629								
55 01B. CN. 10. HNC5102056RGL-VISKRTYDD-QMM-LQ-K-S	629								
56 cpx. FR. 10. URFS patient_ARGL-VISKRTYDD-QMM-LQ-K-S	629								
57 BC. CN. 09. 09YNLX1959RGL-VISKRTYDD-QMM-LQ-K-S	629								
58 01B. MY. 09. 09MYPR37RGL-VISKRTYDD-QMM-LQ-K-S	629								
59 01B. CN. 09. 09LNA423RGL-VISKRTYDD-QMM-LQ-K-S	629								
60 BC. IT. 11. BAV499RGL-VISKRTYDD-QMM-LQ-K-S	629								
61 BC. CN. 10. J1100010RGL-VISKRTYDD-QMM-LQ-K-S	629								
62 BC. CN. 10. YNFL13RGL-VISKRTYDD-QMM-LQ-K-S	629								
63 02A1. RU. 10. 10RU6637RGL-VISKRTYDD-QMM-LQ-K-S	629								
64 BC. CN. 09. YNFL31RGL-VISKRTYDD-QMM-LQ-K-S	629								
65 cpx. CN. 10. YNFL01RGL-VISKRTYDD-QMM-LQ-K-S	629								
67 01B. CN. 11. ANHUI HF115RGL-VISKRTYDD-QMM-LQ-K-S	629								
68 01B. CN. 11. ANHUI WH73RGL-VISKRTYDD-QMM-LQ-K-S	629								
69 01B. JP. 05. 05JPMYC113SP420RGL-VISKRTYDD-QMM-LQ-K-S	629								
70 BFI. BR. 10. 10BR PE004RGL-VISKRTYDD-QMM-LQ-K-S	629								
71 BFI. BR. 10. 10BR PE008RGL-VISKRTYDD-QMM-LQ-K-S	629								
72 BFI. BR. 10. 10BR MG002RGL-VISKRTYDD-QMM-LQ-K-S	629								
74 01B. MY. 11. 10MYPR268RGL-VISKRTYDD-QMM-LQ-K-S	629								
0. BF. 87. ANT70	IA-P-ISMIFV-SA-T-ATHT-KDQ-Q-R-SXR-R-LL-TL-TLSL-KV-Y-S-KR-T-IGNSDTL-Q-Q	632	
0. CM. 98. 98CMA104	IA-P-ISMIFV-SA-T-ATHT-KDQ-Q-R-SXR-R-LL-TL-TLSL-KV-Y-S-KR-T-IGNSDTL-Q-Q	632	
0. CM. 98. 98CMABB141	IT-P-ISMIFV-SA-T-ATHT-KDQ-Q-R-SXR-R-LL-TL-TLSL-KV-Y-S-KR-T-IGNSDTL-Q-Q	632	
0. CM. 98. 98CMABB212	VS-PIISMIFF-V-SA-T-ARTHSV-RDQ-Q-R-SXR-R-LLQ-L-TLR-DL-KR-I-Y-SF-KT-GDYNDSGL-Q-Q	630	
0. CM. 98. 98CMU5337	IA-P-ISMIFV-SA-T-ATHS-MKDQ-Q-R-SXR-R-LLQ-L-TLTL-TLNL-KR-Y-SA-N-TKYNETNIDKGNL-Q-Q	639
0. CM. 99. 99CMU4122	IA-P-ISMIFV-SA-T-ATHT-MKDQ-Q-R-SXR-R-LLQ-L-TLTL-TLNL-KR-Y-SA-N-TKYNETNIDKGNL-Q-Q	639
0. FR. G. VAI	IA-PTISMIFV-SA-T-ARTOH-KDQ-Q-R-SXR-R-LLQ-L-TLTL-TLHNL-KR-Y-SK-T-GGDND-NGNL-Q-Q	644
0. GA. 11. 11Gab6352	IA-P-ISMIFV-SA-T-ATHT-KDQ-Q-R-SXR-R-LLQ-L-TLTL-TLNL-KR-Y-SK-T-GTNTD-NGNL-Q-Q	654
0. SN. 99. 99SE MP1299	IA-P-ISMIFV-SA-T-ATQS-MKDQ-Q-R-SXR-R-LLQ-L-TLTL-TLNL-KR-Y-SK-T-TCTNTNK-DDDKL-Q-Q	651
0. US. 10. LTNP	IA-P-ISMIFV-SA-T-AHT-MKDQ-Q-R-SXR-R-LLQ-L-TLSL-TLNL-KQV-Y-SK-DT.GNNK-DDDKL-Q-Q	651
N. CM. 02. DJ00131TSAFL-RITVL-HSIR-KV-IR-ISLT-Y-TETNTSYDTGNL-Q-Q	613
N. CM. 02. SJGddHTSAFL-RITLHSIR-KV-IR-ISLX-Y-TETSMTSYDTGNL-Q-Q	609
N. CM. 04. 04CM 1015 04TSAFL-RITVL-HSIR-KV-IR-ISLT-Y-TETSMTSYDTGNL-Q-Q	606
N. CM. 06. U14296HTSAFL-RITVL-HSIR-KV-IR-ISLT-Y-TETSMTSYDTGNL-Q-Q	631
N. FR. 11. N1 FR 2011HTSAFX-LIKHMSR-R-LI-T-RLT-Y-TETSMTSYDTGNL-Q-Q	617
P. CM. 06. U14788	IS-PTI																

Table with columns for protein ID, glycosylation NYT, transmembrane domain, gp41 cytoplasmic tail start, glycosylation NGS, and protein ID. Rows list various HIV-1 proteins such as B.FR.83.HXB2, A1.CD.97.97CD.KCC2, etc.

	glycosylation NAT	Env end gp41 end	cytoplasmic tail end
B.FR.83.HXB2	KYWNWNLQYV...SQELKNSAVS.....LLNATAIAVAEGTRDRIEVVQGAACRAIRHIPRRIRQGLERRILL...*		
A1.CD.97.97CD KCC2	-L-L-V-T-T-W-A-G-RIG-FLN-----F-A-----		
A1.CM.08.886 24	-HL-L-L-GR-N-N-DTI-GW-W-LG-RI-N-----A-----		
A1.CY.08.CY236	-LE-V-GR-I-IND-FDIT-GW-I-IG-IN-----A-----		
A1.IN.09.NARI FLS TVC19 1	-L-L-I-GR-X-IX-XDITV-V-GW-I-XG-IGG-L-----A-----		
A1.KE.11.DEM11KE002	-L-G-L-G-R-IND-DNT-GW-IG-R-----A-----		
A1.RU.11.RU6950	-LG-G-G-R-IND-VDIT-V-GW-IG-----CN-----A-A-O		
A1.RW.11.DEM11RW002	-LG-T-I-I-I-I-FDIT-T-GW-I-RI-LN-T-----F-A-----		
A1.SN.01.DD1579	-L-V-L-GR-S-I-I-V-W-IG-R-G-LL-V-----A-----		
A1.UG.11.DEM110UG009	-L-L-L-GR-T-ID-F-T-GW-I-IG-RIG-LNV-T-----A-----		
A1.ZA.04.503 15344 T10 A1	-HL-R-GR-T-IND-DTI-V-GW-IG-RVG-LN-T-----F-A-----		
A2.CD.97.97CDKT848	-L-G-R-GR-K-IND-DTI-V-W-IG-R-R-----A-----		
A2.CM.01.01CM 1445MV	-L-L-V-GR-N-IND-VD-I-VT-LG-RI-LN-T-----F-A-----		
A2.CY.94.94CY017 41	-NL-L-L-GR-I-IND-FDIT-V-W-IG-R-F-LN-----A-----		
B.BR.10.10BR RJ032	-----G-----I-IA-RIG-G-L-T-----A-----		
B.CA.07.502 1191 03	-----G-----F-----L-R-G-VL-V-----F-S-----		
B.CH.08.M2 0803101 NFLG8	-L-----TN-TI-----LA-RGF-FL-----A-O-X		
B.CN.12.DEMB12CN006	-L-----G-K-TN-V-VA-ATG-FLN-----A-TA-O		
B.CU.14.14CU005	-C-----I-----I-----I-R-FL-----A-----		
B.ES.14.ARP1195	-----I-----I-----I-A-RFF-G-L-A-----A-----		
B.FR.11.DEMB11FR001	-----I-----I-----I-RI-I-V-----A-----		
B.HT.05.05HT 129389	-----K-I-----D-V-G-I-IRRV-L-V-T-----A-----		
B.JP.12.DEMB12JP001	-----K-I-----I-----I-RL-L-V-----A-----		
B.KR.07.HP 18 07JHS10_3909	-----K-I-----H-R-I-A-IL-R-Y-----F-A-----		
B.RU.11.11RU21n	-----I-----I-----F-----RTF-VL-----F-A-----		
B.SE.12.SE600057	-----S-----I-----I-----I-RT-----A-----		
B.TH.10.DEMB10TH002	-----S-----I-----I-----I-A-V-G-----#-A-----E-A-O		
B.US.13.RV 1	-----I-----I-----I-----I-----L-----L-----		
B.ZA.09.DEM09ZA022	-L-----S-----F-I-TS-----R-GG-VL-----A-O		
C.BR.07.D2YC07BR003	-LGS-V-GL-KR-I-FDTV-I-II-IW-CN-----F-AA-O		
C.BW.00.00BW003 1	-LGS-V-GL-K-I-DTI-----I-RLY-LN-----F-AA-O		
C.CN.10.YNFL19	-LGG-V-GL-K-I-DT-----I-ATX-R-W-----T-G-D-P-VA-O		
C.CY.09.CY260	YFVS-VVL-----L-K-VS-DDT-TT-A-GAA-IV-LLRRISGI-CS-R-----RC-AAO		
C.ES.14.ARP1198	-LGS-V-GL-K-I-D-I-----I-GSI-RI-LN-----F-AT-----		
C.ET.02.02ET 288	-LGSIV-GL-K-I-FDTI-----I-FL-RV-CNA-----F-AA-----		
C.IN.09.T125 2139	-LGS-V-GL-KG-I-DT-----I-FL-RL-YN-----F-AA-O		
C.KE.05.05KE369195V4	-LGSIV-GL-K-I-DT-----I-FL-RL-YN-----F-AA-O		
C.MW.09.703010256 CH256.w96	-LGS-V-GL-K-I-DT-----I-FL-RL-Y-----F-AA-O		
C.SE.13.SE600311	RFLGT-V-G-L-R-VII-F-TPP-V-VKPF-EMKLF-RVG-V-SN-L-ELS-----IVFO		
C.TZ.08.707810457 CH457.w8	-LGS-V-GL-K-I-DTI-----I-RL-CNVT-----AA-O		
C.US.11.17784 468	-LGS-V-GL-K-IND-FDTI-----I-FL-RTW-CN-----F-AA-O		
C.YE.02.02YE511	-LGS-V-GL-K-IND-VDIT-I-RS-LCN-----F-AA-O		
C.ZA.12.DEMC12ZA096	-LGS-V-GL-K-I-DTI-----I-DF-LRI-T-N-----F-AA-X		
C.ZM.11.DEMC11ZM006	-LGS-I-GL-K-I-DTI-RT-----I-L-LNI-LS-----F-AA-O		
D.CM.10.DEMD10CM009	-L-----R-I-FD-----A-LR-F-L-----A-----		
D.CY.06.CY163	-L-----I-IND-D-RT-VLR-----F-A-----		
D.KE.11.DEMD11KE003	-L-----I-IND-FDIT-VI-----L-IGLRIG-IND-----F-A-----		
D.KR.04.04KR88	-LG-----I-IR-FD-IR-F-VL-----F-A-I		
D.SN.90.SE365	-L-S-----I-IND-DTI-----I-D-R-----GA-----		
D.TZ.01.A280	-L-----I-IND-I-V-I-II-R-F-VL-V-----A-----		
D.UG.10.DEMD10UG004	R-L-----I-V-I-FGT-V-----I-L-RIG-LN-T-----A-AF-----		
D.UG.11.DEMD11UG003	-L-----I-IND-F-T-V-----L-R-G-VLN-----F-A-----		
D.YE.02.02YE516	-L-----R-I-IND-I-IND-F-LN-----A-----		
D.ZA.90.R1	-L-----I-IND-DTI-T-----I-R-VLN-V-----A-----		
F1.A0.06.A0 06 ANG32	-L-V-----R-IND-DT-V-----L-R-G-VL-----A-----		
F1.AR.02.ARE933	-LLG-TL-----RI-I-F-T-V-----A-R-V-VLN-----AF-----		
F1.BR.07.07BR844	-LG-A-----I-F-T-VV-----IL-AL-R-G-VLN-----A-----		
F1.BR.10.10BR PE107	-LLG-IT-----I-----T-----KAL-R-GAV-C-----F-A-----		
F1.BR.10.10BR RJ015	-LG-IIL-----K-----V-G-I-AL-RFG-LN-----F-A-I		
F1.CY.08.CY22	-L-----R-IND-DTI-V-----I-AL-R-G-VLN-V-----K-----V-----A-----		
F1.ES.02.ES X845 4	-L-----R-IND-I-----I-----I-AL-RTG-V-----V-----A-----		
F1.ES.11.VA0053 nfl	-LL-V-----R-IND-I-----I-----I-V-LL-L-G-VLN-----L-----		
F1.R0.96.BCI R07	-L-T-----I-IND-I-----I-----I-VL-LN-----A-----		
F1.RU.08.D88 845	-L-A-----GR-I-IND-T-V-W-----L-AL-R-G-VL-----L-N-X		
F2.CM.02.02CM 6016BBY	-L-A-----GR-I-IND-T-V-IL-R-G-L-----F-A-----		
F2.CM.10.DEMF210CM001	-L-TLH-G-K-I-FDTI-----I-IL-R-G-L-----F-A-----		
F2.CM.10.DEMF210CM007	-L-A-H-----G-K-I-DT-V-----I-D-L-RIG-VL-----F-A-----		
G.CM.07.920 49	-L-----V-G-IND-DT-NW-I-A-RV-FLN-----F-A-----		
G.CM.10.DEMG10CM008	-V-----L-G-IND-DT-DW-----A-RVG-FLN-----F-A-----		
G.CM.10.DEURF10CM020	-L-----L-GG-IND-FDTV-V-NW-----A-RVG-VLN-----A-A-----		
G.CN.08.GX 2084 08	-L-----L-GR-IND-DTI-NW-----AHR-CN-----A-V-----		
G.ES.09.X2634 2	-L-----L-GR-S-IND-DTV-NW-----I-RVG-LN-----A-----		
G.ES.14.ARP1201	-L-----L-GR-IND-VDITL-NW-----I-RTG-LN-----A-A-O		
G.GH.03.03GH175G	-L-----L-V-IND-DTV-NW-----I-R-G-LN-T-----F-A-----		
G.KE.09.DEMG09KE001	-L-----V-G-IND-DTV-NW-----I-R-VL-----F-SA-----		
G.NG.09.09NG SC62	-L-----L-G-IND-VDITL-V-GNW-----A-R-G-FL-V-----A-----		
G.PT.x.PT3306	-L-----L-G-S-IND-DTV-V-NW-----G-RVG-FLN-V-----A-----		
H.BE.93.VI991	-LLG-L-----G-I-----T-----T-----L-R-W-L-----F-A-----		
H.BE.93.VI997	-L-----L-G-IND-I-----T-V-----I-R-W-VL-----F-S-----		
H.CF.90.056	-L-----L-G-IND-ID-----GI-VI-R-W-VL-----F-S-----		
H.GB.00.00GBAC4001	-L-----L-G-N-----T-V-----I-L-R-G-I-----F-A-----		
J.CD.97.J 97DC KTB147	-LG-VC-----G-I-IND-I-----A-I-IF-L-----X-----		
J.SE.93.SE9280 7887	-LV-VW-----G-I-IND-T-----I-IA-R-F-L-----A-----		
J.SE.94.SE9173 7022	-LV-VW-----G-I-IND-T-----I-IA-R-F-L-----A-----		
K.CD.97.97ZR E0TB11	-L-IL-----G-I-IND-T-----I-IND-YR-F-LL-----F-L-----		
K.CM.96.96CM MP535	-L-----V-I-IND-T-----G-----I-IG-R-F-LL-----A-----		
01 AE.AF.07.569M	-LG-L-----G-RI-I-D-I-GW-----I-W-FI-----F-A-----		
01 AE.CM.11.1156 26	-LG-S-----G-T-TT-D-I-GW-----I-R-VL-----F-A-----		
01 AE.CN.12.DE00112CN011	-LG-L-----G-I-IND-FD-V-W-I-AA-R-G-L-L-----A-O		
01 AE.HK.04.HK001	-LG-L-----G-I-IND-T-----W-----A-R-W-L-----T-V-----		
01 AE.IR.10.10IR.THR48F	-LG-A-----G-I-S-D-GW-----A-R-W-LL-----A-----		
01 AE.JP.11.DE00111JP003	-LG-I-----G-I-IND-D-GW-----A-R-W-L-T-----A-----		
01 AE.SE.11.SE601038	-LG-RML-----I-T-I-SD-S-GW-----LA-T-W-T-L-R-----HV-----		
01 AE.TH.10.DE00110TH001	-L-----L-G-IND-L-----A-W-L-----A-----		
01 AE.TH.90.CM240	-LG-L-----G-I-IND-D-A-GW-----A-W-L-----T-----		
01 AE.US.05.306163 FL	-LG-V-----G-T-I-D-T-W-----A-RTG-LN-----A-I-----		
02 AG.CM.10.DE00210CM013	-L-----S-G-IND-DTI-NW-----I-R-YN-----A-----		
02 AG.ES.06.P1423	-FLG-S-----GR-IND-DTI-NW-----IA-R-G-FL-----F-A-----		
02 AG.GW.05.CC 0048	-L-----L-G-IND-TI-V-NW-----IG-S-CN-T-----L-----		

		glycosylation NAT	Env end gp41 end cytoplasmic tail end
B. FR. 83. HXB2	KYWNWLLQYW...SQELKNSAVS.....	LLNATAIAVAEGTRDRIEVVQGA*	RAIRHIPRRIRQGLERRILL...*
02 AG. KR. 12. 12MHR9	--HL-S-A...G...L.....	FDTI...NW...IG-R-G...LN...	F-A.....
02 AG. LR. x. POC44951	--L-S...S...G...I...N.....	DTI...NW...IG-R-G...N-T...	KT.....
02 AG. NG. 09. 09NG SC61	--L-S...S...G...I...N.....	I...NW...I...IG-R-G...N...
02 AG. NG. x. IBNG	--L-S...V...G...I...N.....	TI...V...NW...A...IG-RV...N...	F-A.....
02 AG. SE. 11. SE602024	--L-S...A...G...I...N.....	A...T...GW...L...R...G...CN...	F-S.....
02 AG. SN. 98. 98SE MP1211	--HL-S...G...I...N.....	DT...V...NW...I...RTG...CN...	F-A-Q...
02 AG. US. 06. 502 2696 FL01	--VLG-S...G...I...N.....	DTV...NW...T...IGRR...G...LN-T...	F-A-Q...
03 AB. RU. 97. KAL153 2	--L-S...I...S...I...N.....	IGTI...GW...IG-RF...M-N...	A-KA-Q...
04 cpx. CY. 94. 94CY032 3	--L-F-L...G...I...N.....	F-T...V...I...A...R...CN...	F-A.....
05 DF. BE. x. VT1110	--L-S-P...R...G...I...N.....	T...V...V...A...AL...R...LN...	F-A.....
06 cpx. AU. 96. BFP90	--LG-IC...G...Q...I...N.....	FD-A...NW...A...RTF...FLNV...	F-A.....
07 BC. CN. 98. 98CN009	--LGS-V...G...K-TI...I...N.....	VDTI...I...L-L...YN...HNV...	F-AA-Q...
08 BC. CN. 97. 97CNGX 6F	--LGS-V...GL...K-TI...I...N.....	VDI...I...NR...I...HNV...	F-AA-Q...
09 cpx. GH. 96. 96GH2911	--LG-A...G...I...I...N.....	T...V...I...R...G...L...LN-T...	F-A.....
10 CD. TZ. 96. 96TZ BF061	--LG-T...I...I...I...N.....	DTI...E...GW...I...R...V...LN-T...	F-A.....
11 cpx. CM. 95. 95CM 1816	--L-L...R...I...I...N.....	T...V...G...I...IHRXL...VLN-V...	F-A.....
12 BF. AR. 99. ARMA159	--L-L...V...GR...I...IT.....	DS...W...I...TG-RIG...G-L...	F-A-Q...
13 cpx. CM. 96. 96CM 1849	--L-L...L...GR...I...I...N.....	DTI...NW...A-X...VG...LNV-V...	F-A.....
14 BG. ES. 05. X1870	--LG-L...G...I...I...N.....	DT...GW...A...R...W...LR...	F-A.....
15 01B. TH. 99. 99TH MU2079	--L-L...L...GR...I...I...N.....	F-T...V...W...I...AL-R-G...VLN...	F-A.....
16 AZD. KR. 97. 97KR008	--L-L...L...GR...I...I...N.....	F-T...V...W...I...AL-R-G...VLN...	F-A.....
17 BF. AR. 99. ARMA038	--L-L...V...G...I-S...IN.....	DTV...T...NW...I...RTG...LN...	A-A-Q...
18 cpx. CU. 99. CU76	--L-L...L...G...T...I...I...N.....	T...V...GW...A...II-R-W...VL...	F-A.....
19 cpx. CU. 99. CU7	--L-L...G...G...I...N.....	DTV...NW...I...I...RGF...FL...	F-A.....
20 BG. CU. 99. Cu103	--L-L...I...GR...R...I...N.....	Y...I...R...R...W...FL...T...	F-A.....
21 AZD. KE. 99. KER2003	--L-L...I...GR...R...I...N.....	D...T...GW...I...R...W...FL...T...	F-A.....
22 01A1. CM. 01. 01CM 0001BBY	--L-L...L...GR...I...N.....	DTV...N...I...I...RGF...FL...	F-A.....
23 BG. CU. 03. CB118	--L-L...L...GR...I...N.....	YDTV...N...I...I...RGF...FL...	F-A.....
24 BG. ES. 08. X2456 2	--L-L...L...GR...I...N.....	DTI...T...DW...IA...R...G...LNV-V...	F-A-Q...
25 cpx. CM. 02. 1918LE	--L-L...A...G...I...N.....	D...V...C...IA...RGG...L...LN...	F-A-Q...
26 AU. CD. 02. 02CD MBT047	--L-L...G...R...G...I...N.....	DTV...T...NW...IL...RI...N-T...	F-A-Q...
27 cpx. FR. 04. 04CD FR K25	--L-L...L...K...I...I...N.....	DTI...I...I...IG...YN...	F-AA-Q...
28 BF. BR. 99. BREPM12609	--LKS-V...GL...K-I...I...N.....	IDTI...DW...G...RVF...FLN...	IA.....
29 BF. BR. 01. BREPM16704	--L-L...S...I...I...I...N.....	IG...I...I...VW...L...	F-S.....
31 BC. BR. 04. 04BR142	R-L-L...L...G...T...I...I...N.....	R-Y...L...T...	A-I.....
32 06A1. EE. 01. EE0369	--L-L...L...G...T...IT.....	IDTI...GW...DIG...RFG...L...	GA.....
33 01B. ID. 07. JKT189 C	--L-L...V...G...T...IT.....	D...VT...GW...A...RTG...LNV-T...	A.....
34 01B. TH. 99. QUR2478P	--L-L...V...G...T...I...N.....	DT...V...GW...IA...R...L...	#-A-Q...
35 AD. AF. 07. 169H	Q-L-L...I...I...I...I...N.....	F-T...V...I...AL...R...G...VLN...	A-I.....
36 cpx. CM. 00. 00CMNYU830	--LG-L...L...G...I...I...N.....	HL-S...G...K-I...I...N.....	S.....
37 cpx. CM. 00. 00CMNYU926	--L-L...K...I...I...I...N.....	TL...I...A...RTG...G-L...	S.....
38 BF1. UY. 03. UY03 3389	--L-L...R...I...I...I...N.....	R...I...I...L...RTG...VL-T...	F-A.....
39 BF. BR. 04. 04BR3179	--L-L...L...R...I...I...N.....	VDTI...GW...A...R...IN...	F-A.....
40 BF. BR. 05. 05BRR1055	--LG-A...G...I...I...N.....	T...V...GW...L...R...G...VLN...	F-A.....
42 BF. LU. 06. LuBF 18 06	--LLG-AW...G...I...I...N.....	D...V...GW...T...SL...FLN...	F-A.....
43 02G. SA. 03. J11223	--L-L...A...G...I...I...N.....	T...V...GW...I...AL-R-G...VLN...	F-A.....
44 BF. CL. 00. CH80	--LG-I...A...I...I...I...N.....	T...V...I...AL-S-G...VL...	F-A.....
45 cpx. FR. 04. 04FR AUK	--LG-VL...GK...I...I...I...N.....	FD-I...V...SW...A...G...RTG...L...	A-A.....
46 BF. BR. 07. 07BR FPS625	--L-L...I...I...I...I...N.....	FDTI...V...SW...A...G...RTG...L...	A-A.....
47 BF. ES. 08. P1942	--L-L...I...I...I...I...N.....	D...T...I...RW...FL...L...	F-A.....
48 01B. MY. 07. 07MYKT021	R-LGX-S...G...T...I...I...N.....	DT...V...GW...I...R...Y...L...T...	F-A.....
49 cpx. GM. 03. N26677	--LG-V...G...I...I...I...N.....	D-A...GW...R...W...L...V...	A-Q...
50 A1D. GB. 10. 12792	--LG-L...G...I...IA...I...N.....	DT...V...GW...R...W...L...V...	A-Q...
51 01B. SG. 11. 11SG HM021	--LG-L...G...I...IA...I...N.....	TI...V...GW...AA...G...L...	S.....
52 01B. MY. 03. 03MYKL018 1	--HL-S...G...K-I...I...N.....	HL-S...G...K-I...I...N.....	F-A.....
53 01B. MY. 11. 11FIR164	--LGSIV...LK...K-T...I...N.....	YDSI...NW...I...FI...RI...S...LN-K...	FKAA-Q...
54 01B. MY. 09. 09MYSB023	--LG-L...GK...I...I...I...N.....	D...SIW...L...L...L...V...	A-Q...
55 01B. CN. 10. HNC5102056	--LGS-V...GL...K-TI...I...N.....	DTL...GW...LA...R...W...LL-V...	A-Q...
56 cpx. FR. 10. URFS patient A	--LGS-V...GL...K-TI...I...N.....	T...R...KI...L...RF...K...L...AK...D...	F-A-Q...
57 BC. CN. 09. 09YNLX195g	--LGSIV...GL...K...IN.....	YDTI...I...TI...R...Y...CN...	F-AA-Q...
58 01B. MY. 09. 09MYPR37	--L-L...A...G...I...N.....	FDTI...NW...G...RVG...LL...	F-A.....
59 01B. CN. 09. 09LNA423	--LGS-V...GL...K-I...I...N.....	DTI...I...R...L...L...V...	F-AA-Q...
60 BC. IT. 11. BAV499	--LG-L...G...I...I...I...N.....	D...VT...GW...A...W...L...V...	F-A.....
61 BC. CN. 10. J1100010	--LG-L...G...I...I...I...N.....	FD...X...GW...AA...R...W...FL...	F-A-V...
62 BC. CN. 10. YNFL13	--LG-L...G...I...I...I...N.....	D...V...T...GW...AA...R...W...FL...	F-A.....
63 02A1. RU. 10. 10RU6637	--LG-S...G...T...T...I...N.....	FD-V...V...GW...A...W...VL...	F-A.....
64 BC. CN. 09. YNFL31	--S.....R.....T.....	T...V...V...AL...R...G...VLNV...	F-A.....
65 cpx. CN. 10. YNFL01	--G-I...L...G...I...I...N.....	T...V...GAL...SIG...FL...	F-A.....
67 01B. CN. 11. ANHUI HF115	--LG-L...G...I...I...I...N.....	FD-A...GW...L...AL...TLG...I...	F-A-Q...
68 01B. CN. 11. ANHUI WH73	--L-L...R...I...I...I...N.....	DTL...V...NW...GI...AGI...RICTG-N...	S.....
69 01B. JP. 05. 05JPMYC113SP420	--L-L...R...I...I...I...N.....	DTI...V...GKW...DI...LGI...RIG...G-LN...	S.....
70 BF1. BR. 10. 10BR PE004	--L-L...R...I...I...I...N.....	VDTI...V...NW...NI...LGI...RLG...G-WN-T...	F-S.....
71 BF1. BR. 10. 10BR PE008	--L-L...R...I...I...I...N.....	FDTI...V-A...NW...NI...LGI...RIG...GXLN...	A-X-X...
72 BF1. BR. 10. 10BR MG002	--L-L...R...I...I...I...N.....	DTI...V-T...NW...SI...LGI...RIG...G-YN...	S.....
74 01B. MY. 10. 10MYPR268	--L-L...R...I...I...I...N.....	DTL...V...NR...SIFLGI...RIG...G-LN...	QA-V...
0. BF. 87. ANT70	--L-L...R...I...I...I...N.....	DTL...V...NW...SI...LGI...RIG...G-LN-T...	S.....
0. CM. 98. 98CMA104	--L-L...R...I...I...I...N.....	DTI...V...NW...VTI...LGI...RIG...G-LN...	S.....
0. CM. 98. 98CMABB141	--L-L...R...I...I...I...N.....	DTF...V...NW...GI...SGI...RIG...G-N...	S.....
0. CM. 98. 98CMABB212	TRL-GIA...GK...D...I...I...N.....	TI...V...L...LA...RIG...G-L...	A.....
0. CM. 98. 98CMU5337	TRL-GI-A...GK...RD...I...I...N.....	T...V...L...LA...RIG...G-L...	A-A-V...
0. CM. 99. 99CMU4122	--RL-GI-A...GK...D...I...I...N.....	T...V...G...L...LA...RIG...G-L...	T.....
0. FR. 92. VAU	--L-L...R...I...I...I...N.....	I-L-GI-A...GK...D...I...I...N.....	A-I.....
0. GA. 11. 11Gabb6352	--L-L...R...I...I...I...N.....	RDCAACG...T...Q...T.....	TV-S-N-W-Q-A-G-QIG-GFLN-L-S...
0. SN. 99. 99SE MP1299	--L-L...R...I...I...I...N.....	RDCAACG...T...Q...T.....	DTV-S-GW-Q-I-G-QIG-GFLN-I-S...X.
0. US. 10. LTNP	STL.....WTIIRTEIK.....	NIDRL-W-G-K-SILLAL-TIV-I-EV.....	IA-N.....
N. CM. 02. DJ00131	--L-LOGI...GK...T...T...I...N.....	DT...I...I...RIG...G-L...	F-A.....
N. CM. 02. SJGddd	--RLC-GV...TR...T...I...I...N.....	D-I-R...QT...FAA...RTG...GFLN-G...	L-A.....
N. CM. 04. 04CM 1015 04	--L-L...R...I...I...I...N.....	VIHRSSIVGVRQAL-WSGNTYASIRASLIQAIDRL-EFTGWW-I-A-VYTA-G-N...	IA-N.....
N. CM. 06. U14296	--L-L...R...I...I...I...N.....	AR...V...I...L...TRRL...I...S...X.	S.....
N. FR. 11. N1 FR 2011	--L-L...R...I...I...I...N.....	RDCAFIA...G...Q...I...I...N.....	DCV-VWT-NW-Q-ATA-RIG-G-LN-S...
P. CM. 06. U14788	--L-L...R...I...I...I...N.....	TLKATA...L...QR...XTN.....	DTV-V-NW-SI-LG-RFG-G-LN-LS...
P. FR. 09. RBF168	--L-L...R...I...I...I...N.....	RDCLAVCG...A...QO...T.....	DTV-VR-DW-Q-L-G-RIG-G-LN-L-S...
CPZ. CD. 90. ANT			
CPZ. CM. 05. SIVcpzLB7			
CPZ. CM. 09. Ptt 09Cam155			
CPZ. TZ. 06. TAN5			
CPZ. US. 05. US Mar1lyn			
GOR. CM. 07. SIVgor2139 287			
GOR. CM. 12. SIVgor B0ID2			
GOR. CM. 13. SIVgor BPID15			

	myristoylation Nef start				acidic cluster	phosphorylation poly-P helix		HXB2 premature Nef end
B.FR.83.HXB2	MGGKWSKSSVIGWPTVRERMR
A1.CD.97.97CD.KCC2
A1.CM.08.886.24
A1.CY.08.CY236
A1.IN.09.NARI.FLS.TVC19.1
A1.KE.11.DEM111KE002
A1.RU.11.DEM111RU002
A1.RW.11.DEM111RW002
A1.SN.01.DDI579
A1.UG.11.DEM110UG009
A1.ZA.04.561.15344.T10.A1
A2.CD.97.97CD.KDKT848
A2.CM.01.01CM.1445MV
A2.CY.94.94CY017.41
B.BR.10.10BR.RJ032
B.CA.07.502.1191.03
B.CH.08.M2.0803101.NFLG8
B.CN.12.DEM12CN006
B.CU.14.14CU005
B.ES.14.ARP1195
B.FR.11.DEM111FR001
B.HT.05.05HT.129389
F1.JP.12.DEM12JP001
B.BR.07.HP.07HS10.3909
B.RU.11.11RU21n
B.SE.12.SE600057
B.TH.10.DEM10TH002
B.US.13.RV.1
B.ZA.09.DEM09ZA022
C.BR.07.07YC07BR003
C.BW.00.00BW003.1
C.CN.10.YNFI19
C.CY.09.CY260
C.ES.14.ARP1198
C.ET.02.02ET.288
F1.IN.09.T125.2139
C.KE.05.05KE36915V4
C.MW.09.703010256.CH256.w96
C.SE.13.SE600311
C.TZ.08.707010457.CH457.w8
C.US.11.17784.468
C.YE.02.02YE511
C.ZA.12.DEM12ZA096
C.ZM.11.DEM11ZM006
D.CM.10.DEM10CM009
D.CY.06.CY163
D.KE.11.DEM11KE003
D.KR.04.04KR04
D.SN.90.SE365
D.TZ.01.A280
D.UG.10.DEM10UG004
D.UG.11.DEM11UG003
D.YE.02.02YE516
D.ZA.90.RJ.1
F1.A0.06.A0.06.AMG32
F1.AR.02.ARE933
F1.BR.07.07BR844
F1.BR.10.10BR.PE107
F1.BR.10.10BR.RJ015
F1.CY.08.CY222
F1.ES.02.ES.E845.4
F1.ES.11.VA0053.nfl
F1.RO.96.BCI.R07
F1.RU.08.D88.845
F2.CM.02.02CM.6016BBY
F2.CM.10.DEMF210CM001
F2.CM.10.DEMF210CM007
G.CM.07.920.49
G.CM.10.DEM10CM008
G.CM.10.DEURF10CM020
G.CN.08.GX.2084.08
G.ES.09.X2634.2
G.ES.14.ARP1201
G.GH.03.03GH175G
G.KE.09.DEMG09KE001
G.NG.09.09NG.SC62
G.PT.x.PT3306
H.BE.93.VI991
H.BE.93.VI997
H.CF.90.056
H.GB.00.00GBAC4001
J.CD.97.J.97DC.KTB147
J.SE.93.SE9280.7887
J.SE.94.SE9173.7022
K.CD.97.97ZR.E0TB11
K.CM.96.96CM.MP535
01.AE.AF.07.569M
01.AE.CM.11.156.26
01.AE.CN.12.DEM012CN011
01.AE.HK.04.HK001
01.AE.IR.10.10IR.THR48F
01.AE.JP.11.DEM011JP003
01.AE.SE.11.SE601018
01.AE.TH.10.DEM0110TH001
01.AE.TH.90.CM240
01.AE.US.05.306163.FL
02.AG.CM.10.DEM0210CM013
02.AG.ES.06.P1423
02.AG.GW.05.CC.0048

B.FR.83.HXB2
 A1.CD.97.97CD.KCC2
 A1.CM.08.886.24
 A1.CY.08.CY236
 A1.IN.09.NARI.FLS.TVC19.1
 A1.KE.11.DEM11KE002
 A1.RU.11.DRU6950
 A1.RW.11.DEM11RW002
 A1.SN.01.DD1579
 A1.UG.11.DEM110UG009
 A1.ZA.04.503.15344.T10.A1
 A2.CD.97.97CDKTB48
 A2.CM.01.01CM.1445MV
 A2.CY.94.94CY017.41
 B.BR.10.10BR.RJ032
 B.CA.07.502.1191.03
 B.CH.08.M2.0803101.NFLG8
 B.CN.12.DEMB12CN006
 B.CU.14.14CU065
 B.ES.14.ARP1195
 B.FR.11.DEMB11FR001
 B.HT.05.05HT.129389
 B.JP.12.DEMB12JP001
 B.KR.07.HP.18.07HS10_3909
 B.RU.11.11RU21n
 B.SE.12.SE600057
 B.TH.10.DEMB10TH002
 B.US.13.RV.1
 B.ZA.09.DEMB09ZA022
 C.BR.07.DEMC07BR003
 C.BW.00.08BW003.1
 C.CN.10.YNFL19
 C.CY.09.CY260
 C.ES.14.ARP1198
 C.ET.02.02ET.268
 C.IN.09.T125.2139
 C.KE.05.05KE369195V4
 C.MW.09.703010256.CH256.w96
 C.SE.13.SE600311
 C.TZ.08.707010457.CH457.w8
 C.US.11.17784.468
 C.YE.02.02YE511
 C.ZA.12.DEMC12ZA096
 C.ZM.11.DEMC11ZM006
 D.CM.10.DEMD10CM009
 D.CY.06.CY163
 D.KE.11.DEMD11KE003
 D.KR.04.04KR048
 D.SN.90.SE365
 D.TZ.01.A280
 D.UG.10.DEMD10UG004
 D.UG.11.DEMD11UG003
 D.YE.02.02YE516
 D.ZA.90.R1
 F1.A0.06.A0.06.AMG32
 F1.AR.02.ARE933
 F1.BR.07.07BR844
 F1.BR.10.10BR.PE107
 F1.BR.10.10BR.RJ015
 F1.CY.08.CY222
 F1.ES.02.ES.X845.4
 F1.ES.11.VA0053.nfl
 F1.RO.96.BCI.R07
 F1.RU.08.D88.845
 F2.CM.02.02CM.6016BBY
 F2.CM.10.DEMF210CM001
 F2.CM.10.DEMF210CM007
 G.CM.07.920.49
 G.CM.10.DEMG10CM008
 G.CM.10.DEURF10CM020
 G.CN.08.GX.2084.08
 G.ES.09.X2634.2
 G.ES.14.ARP1201
 G.GH.03.03GH175G
 G.KE.09.DEMG09KE001
 G.NG.09.09NG.SC62
 G.PT.x.PT3306
 H.BE.93.VI991
 H.BE.93.VI997
 H.CF.90.056
 H.GB.00.00GBAC4001
 J.CD.97.J.97DC.KTB147
 J.SE.93.SE9280.7887
 J.SE.94.SE9173.7022
 K.CD.97.97ZR.E0TB11
 K.CM.96.96CM.MP535
 01.AE.AF.07.569M
 01.AE.CM.11.1156.26
 01.AE.CN.12.DE0012CN011
 01.AE.HK.04.HK001
 01.AE.IR.10.10IR.THR48F
 01.AE.JP.11.DE0011JP003
 01.AE.SE.11.SE601038
 01.AE.TH.10.DE00110TH001
 01.AE.TH.90.CM240
 01.AE.US.05.306163.FL
 02.AG.CM.10.DE00210CM013
 02.AG.ES.06.P1423
 02.AG.GW.05.CC.0048

PGPGVRYPLTFGWCKYLPVPEPKIEE.ANKGENTSLHPVSLHGMD.D.PEREVLEWRFDLSRLAFHHVARELHPEFY.K.....N.....C*

---I-F-----F-----EAV--TG--N---TCQ-----M-K-----LK-R-K-----FY.....D.....
 ---I-F-----F-----EEV-R-TE--N---MCO-----D--T-I-K--H--K-I--M--FY.....D.....
 ---T-----F-----D-EV-K-E-E-NX---TCQ---E-E-K---Q-K-X-LK-R-X-M--FY.....D.....
 ---T-----F-----F-ADV-K-ETE--N---TCQ---E-E-F-K---K-Q---LT-R--M--Y.....D.....
 ---T-----F-----N-EEV-R-E-E-N---TCQ---E-E-G-T-M---LR--E--FY-DC*QEQ#0.D.....
 ---T-----F-----DQASV--TE--N---CO---K-K-M---LK-T--M--FY.....D.....
 ---T-----F-----E-EV-K-E-E-N---TCQ---E-T-K-K---K-L--DFY.....D.....
 ---A-F-----F-----M-SEV-K-AE-DNC---INO---D--T-V---K-K---LR-L-Q--FY.....D.....
 ---F-----F-----N-SEV--TE--N---TCQ-I---K--H--LR-R--Y-DFY.....D.....
 ---I-----F-----N-SEV--TO--N---TCQ-V---R-E-RS-RR-R--Y.....D.....
 ---T-----F-----QV---IN-M-Q---K-M-K---K--Y--DWO#--D.....
 ---T-----F-----EQV--TV-TNN-G---E-K-V--K---K-K-F-R--D.....R-
 ---T-----C-----F-----D-V--T-IN-M-Q---E-K-V-K---M--FY.....D.....
 ---I-----F-----LD-EQV-K-TE--NC-MNQ---K-V---L-KD-Y-D-Y-G*#--D.....
 ---EX-----F-----EQV-Q---E-IT---G---K-V-K-L--N--Y.....D.....
 ---K-I---C-----F-----DKE--Q--NC-M-Q---K-V-K-L-VRM-K-FY.....D.....
 ---T-----F-----D-QV-K-D-E-IN-M-Q---M-K-H-LQ---K--Y.....D.....
 ---T-F-----F-----E-V--TV-NC-MN---G--M-K--H-LQ---K--Y.....D.....
 ---T-----F-----E-E-K-TT-NC-ANM---V-K---Y--K-FY.....D.....
 ---T-----F-----D-EEV-R-TV-NC-ANV--N---M-K---Y--M--Y-L#--D.....
 ---T-W-----F-----N-SE-TN--A--E--G-V-K---L--VR-M--Y-Q-Y-E-#--D.....
 ---I-----C-----F-----E-V--E-N-I---V-K-L-VR-M--Y-Q-Y-E-#--D.....
 ---I-----F-----D-REV--T-N-M-Q---O-K---T--Y-N#--D.....
 ---I-----F-----D-KEV-TE-NC-TCQ---E-ED--Q-K-RL-RR-M--R-FY.....D.....
 ---F-----F-----D-KEV--E-NC-CO-E-ED--K-Q-IQ-HR-M--FY.....DC*HRRT...VT-
 ---F-----D-KEV-N-T-DNC-M-Q-E-A--O-K-S-RR--Y.....D.....
 ---F-----F-----D-SEV--E-NC-I---EDK-M-K-IS-HR-L-Q--FY.....D**HRRD...F#-
 ---F-----F-----I-D-KEV-E-NC-M-Q-E-GD--O-K-O-RR-I--FY.....D.....
 ---F-----F-----D-REV--E-NC-M-Q---DH--K-K-O-HK-M--FY.....D.....
 ---F-----F-----D-OTV--E-NC-I---E-ED--M-K-IQ-HK-M--FY.....D.....
 ---L-----L-----D-KEV-ETE--N-M---E-EH--K-K-M-RR-M--K--Y.....D.....
 ---T-----F-----D-SEV--E-NC-M-Q---E-ED--K-S-RR-M--Y.....D.....
 ---T-----F-----D-REV--E-N-I---E-GDG--M-K-S-HR--FY.....D.....
 ---F-----F-----N-OEV--E-N-I---I-E-E--K-E-M-RR-M--Y.....D.....
 ---PF-----D-QEA-H-RDDSC-M-M-Q-IE-EH-I-M-K-TQ-RR-L--WY.....D.....
 ---T-----FE---D-RVV-G-ETE--C---Q--E-T--V-K-N--E-K-QK--F--D.....
 ---I-----FE---L-EEV-K-TE--NC-M-Q---Q---I-N--E-K-Q--Y.....D.....
 ---I-----FE---D-SEV--E-NC-M-Q---E-EH--Q-K-N-K-E-K-M--FY.....D.....
 ---T---C-----F-----D-EEV--TE-DNC-LCO-IE-E-Q-V-N-K-E-K-K--FY.....D.....
 ---T-----FE---D-KEV-TTE-SC-I-O-E-T-K-V-N--E-K-S--Y.....D.....
 ---T-----FE---D-KEV-DTE-NC-MCO-E-Q--K-N--E-K-KIQ--Y.....D.....
 ---I-----F-----D-AKEA-DTE-NC-ACO-E-T--M-N--E-K-V--DFY.....D.....
 ---I-H-----F-----D-KWV--E-E-DK---M-H---K-V--N-K-KT--F--D.....
 ---I-----F-----D-QEV-Q-TE-TNC-MNQ-E--K-K---E-K--Y.....D.....
 ---F-----F-----N-EEV-R-E-NC-I-O-E-ED--M-K-S-RR-I-K-FY-AE.....D.....
 ---F-----F-----D-EDV-K-E-NC-M-Q-E-E-K-Q---LK-I-E-FY.RDW.....D.....X
 ---I-----L-----D-EEV-K-E-NC-M-Q---ADG--K-K-T-LR-I-RF--Y-D*#--D.....
 ---T-----L-----F-----D-EEV-K-E-NC-M-Q---EDG--K-Q---LK-I-R--Y-OD*#--D.....
 ---T-----L-----E-----D-EEV-K-E-NC-M-Q-E-ED--K-Q---K-E-R-FY-OD*DL*HRD.....D.....
 ---L-----L-----F-----D-EEV-R-E-C-M-Q---EDG--R-K---LR-L-K-FY-D*VVDI#--R.....
 ---X-----I-----F-----S-EEV-K-E-NC-M-Q-E-EDG--M-K-G-RR-I-K-FY-D*#--D.....
 ---T-F-----F-----D-EEV-K-E-TE-NC-I-O-IE-ED--M-K-T-MR-I-K-K-FY-D*#--D.....
 ---T-F-----F-----D-EEV-K-E-NC-TCQ-E-ED--K-T-S-RR-I-K-FY.....D.....
 ---T-----F-----MD-AE--T-R-NC-TCQ-E-ED--V-S-RR-I-K-FY.....D.....
 ---T-F-----F-----D-ADV-I-R--N---I-O-E-GDN--V-K-S-RR-I--Y.....D.....
 ---T-F-----F-----MD-AEV--E-N---TCQ-Q---ED--V-S-RR-R--FY.....D.....
 ---T-F-----F-----E-MD-AEV--TR--N---TC-AE-GOK-R-K-S-RR-L--FY.....D.....
 ---K-----F-----D-EEV-K-E-N---TCQ-IE-ED--V-S-RR-I--DFY.....D.....
 ---T-L-----F-----ID-AEV--TN---TCQ-IE-EDK--V-K-S-RR-L--Y.....D.....
 ---I-----F-----D-AE--TE-RNT-TCQ-E-SD--K-K-S-RR-I--Y.....D.....
 ---I-----PF-----MN-AE--N---I-O-E-DDK--V-K-S-RR-I--D-Y.....D.....*
 ---I-----F-----M-EE--TI--N---I-O-E-ADK--V-S-RR-I--DFY.....D.....
 ---E-----F-----D-ODV-K-E-N---MCO-IE-E--M-K---LR-R-K--FY.....D.....
 ---E-----F-----I-D-REV--E-NC-Y-O-T-IE-ED--M-K-S-RR-I-K-FY.....D.....
 ---E-F-----F-----M-OEV-Q---E-N-M---E-DG--M-K---LT-L-VK--Y.....D.....
 ---C-----C-----F-----D-OEV-K-E-NC-MCO-E-E-K-V--H--R-R-Q--FY.....D.....
 ---I-----I-----D-GEV-R-E-DNR-TCQ-E-ED--M-K-TS-RR-M--FY-NC*...OD.....
 ---I-----D-SEV--E-NC-TCQ-IE-E--O-K-S-RR-I--FY.....D.....
 ---TX-----F-----D-SEV--E-NC-ACO-IE-E--K-K-S-RR-I--FY.....D.....
 ---I-----D-REV--TE-NC-NO-E-EH--K-K-S-RK--M--Y.....D.....
 ---I-----D-AEV-TTE-DNC-INO-E-EH-I-M-K-S-RR--D-Y.....D.....
 ---T-F-----C-----F-----D-STV--D--NC-M-Q-IE-E--I-K-A-RK-I--Y.....D.....
 ---I-----C-----F-----D-REV--D--NC-Y-O-T-IE-ED--M-K-A-RK-I--Y-D*#--D.....
 ---I-----C-----F-----D-REV--D--SC-INO-IE-E--M-K-A-RK-I--R--Y.....D.....
 ---I-----C-----F-----D-REV--D--NC-A-O-I-E-E--M-K-A-RK-L--FY.....D.....
 ---I-----C-----F-----DROEV--NC-M-Q-OE-E--R-K-X-RK--Y.....D.....
 ---I-----C-----F-----D-REV-DT-E-NC-M-Q-V--DK--M-K-A-RK-I-K-FY.....D.....
 ---I-----C-----F-----D-O--N---M-Q-AE-E-I-K-S-RK-I--R-FY-RD*#--D.....
 ---I-----F-----I-ID-REV--E-NC-M-Q-IE-E--M-K-A-RK-I--Y.....D.....
 ---I-F-----F-----DQREV--D--NC-M-Q-IE-E--M-K-A-RK--Q--Y.....D.....
 ---T-F-----C-----F-----D-REV--NC-L-O-IE-E--M-K-A-RK--D--Y.....D.....
 ---T-F-----F-----LD-VE--R-E-N---TCQ-ED--V--T--K-K--FY.....D*#RCC*QKTADTGDG*#-
 ---T-F-----F-----D-SEV-K-E-NC-TCQ-ED--V--L-K-T--Y.....D.....
 ---I-----F-----MD-TEV--L-E-N---Q-E-ED--V--X-X.....D.....

normal Nef end

205
207
205
45
205
205
210
215
204
205
213
217
204
218
220
207
214
208
105
207
224
126
208
207
207
208
208
205
207
208
208
208
213
54
203
213
208
208
208
220
209
206
211
45
207
203
208
208
208
218
220
209
206
211
45
207
203
208
208
208
218
221
132
214
208
208
213
208
208
207
209
215
208
219
203
213
208
208
207
207
208
205
204
205
207
216
225
206
207
207
207
205
206
206
205
199
229
207
192

B. FR. 83. HXB2
 02 AG. KR. 12. 12MHR9
 02 AG. LR. x. POC44951
 02 AG. NG. 09. 09NG SC61
 02 AG. NG. x. 1BNG
 02 AG. SE. 11. SE602024
 02 AG. SN. 98. 98SE MP1211
 02 AG. US. 06. 502 2696 FL01
 03 AB. RU. 97. KAL153 2
 04 cpX. CY. 94. 94CY032 3
 95 DF. BE. x. VT1110
 07 cpX. AU. 96. BFP90
 07 BC. CN. 98. 98CN009
 08 BC. CN. 97. 97CN6X 6F
 09 cpX. GH. 96. 96GH2911
 10 CD. TZ. 96. 96TZ BF061
 11 cpX. CM. 95. 95CM 1816
 12 BF. AR. 99. ARMA159
 13 cpX. CM. 96. 96CM 1849
 14 BG. ES. 05. X1870
 15 01B. TH. 99. 99TH MU2079
 16 AZD. FR. 97. 97FR000
 17 BF. AR. 99. ARMA038
 18 cpX. CU. 99. CU76
 19 cpX. CU. 99. CU7
 20 BG. CU. 99. CU103
 21 AZD. XE. 99. KER2003
 22 02A1. CM. 01. 01CM 0001BBY
 23 BG. CU. 03. CB118
 24 BG. ES. 08. X2456 2
 25 cpX. CM. 02. 1918LE
 26 AU. CD. 02. 02CD MBT047
 27 cpX. FR. 04. 04CD FR K25
 28 BF. BR. 99. BREPM12609
 29 BF. BR. 01. BREPM16704
 31 BC. BR. 04. 04BR142
 32 06A1. EE. 01. EE0369
 33 01B. ID. 07. JKT189 C
 34 01B. TH. 99. 01B2478P
 35 AD. AF. 07. 169H
 36 cpX. CM. 00. 00CMNYU830
 37 cpX. CM. 00. 00CMNYU926
 38 BF1. UY. 03. UY03 3389
 39 BF. BR. 04. 04BR13179
 40 BF. BR. 05. 05BRR1055
 42 BF. LU. 06. LuBF 18 06
 43 02G. SA. 03. J11223
 44 BF. CL. 08. CH80
 45 cpX. FR. 04. 04FR AUK
 46 BF. BR. 07. 07BR FPS625
 47 BF. ES. 08. P1942
 48 01B. MY. 07. 07MYKT021
 49 cpX. GM. 03. N26677
 50 A1D. GB. 10. 12792
 51 01B. SG. 11. 11SG HM021
 52 01B. MY. 03. 03MYK1018 1
 53 01B. MY. 11. 11FIR164
 54 01B. MY. 09. 09MYSB023
 55 01B. CN. 10. HNC5102056
 56 cpX. FR. 10. URFS patient_A
 57 BC. CN. 09. 09YNLX19sg
 58 01B. MY. 09. 09MYPR37
 59 01B. CN. 09. 09LNA423
 60 BC. IT. 11. BAV499
 61 BC. CN. 10. J1190010
 62 BC. CN. 10. YNFL13
 63 02A1. RU. 10. 10RU6637
 64 BC. CN. 09. YNFL31
 65 cpX. CN. 10. YNFL01
 67 01B. CN. 11. ANHUI HF115
 68 01B. CN. 11. ANHUI WH73
 69 01B. JP. 05. 05JPMYC113SP420
 70 BF1. BR. 10. 10BR PE004
 71 BF1. BR. 10. 10BR PE008
 72 BF1. BR. 10. 10BR MG002
 74 01B. MY. 10. 10MYPR268
 O. BE. 87. ANT70
 O. CM. 98. 98CMA104
 O. CM. 98. 98CMABB141
 O. CM. 98. 98CMABB212
 O. CM. 98. 98CMU5337
 O. CM. 99. 99CMU4122
 O. FR. 92. VAU
 O. GA. 11. 11GAB6352
 O. SN. 99. 99SE MP1299
 O. US. 10. LTNP
 N. CM. 02. DJ00131
 N. CM. 02. SJGddd
 N. CM. 04. 04CM 1015 04
 N. CM. 06. U14296
 N. FR. 11. N1 FR 2011
 P. CM. 06. U14788
 P. FR. 09. RBF168
 CPZ. CD. 90. ANT
 CPZ. CM. 05. SIVcpzLB7
 CPZ. CM. 09. Ptt 09Cam155
 CPZ. TZ. 06. TAN5
 CPZ. US. 05. US Mar1lyn
 GOR. CM. 07. SIVgor2139 287
 GOR. CM. 12. SIVgor B0ID2
 GOR. CM. 13. SIVgor BPID15

PGPGVRYPLTFGWCKLVPVDPDKIEE. ANKGENTSLLHPVSLHGMD. PEREVLWRFDLSRLAFHHVARELHPEFY. K. N. C*

```

    ---I---F---LD-AEV---E-N---CO-E-ED---V-G---T-K-K---FY---D---
    ---T---F---MD-AEV---E-N---TCO-IE-TD-Q-V---S-K-R-Q---FY---D---
    ---T---F---MD-AD---E-N---TCO-IE-E-V---S-RR-I---Y---D---
    ---T---F---MD-AEV---E-N---TCO-IE-DD-J---S-RR-I---Y---D---
    ---I---F---MD-AEV---E-N---TCO-IE-ED-V-S---YT-Q---QV---D---
    ---T---F---LD-AEV---E-N---TCO-IE-AD-D-V---T-LT-R---FY---D---
    ---I---F---D-AEV---TE-N---TCO-IE-E-K-M-K---LT-R---FY---D---
    ---E-F-C---F-D-OEV---TE-C-I-Q---E-E-K-K---YK---FY---D---
    ---T---L---F-N-EEV-K-LE-DNC-TCO-AEV-ED-D-Q-K-S-LR-T-R-FY-OD*-D-
    ---T---F---D-REV---E-N---CO-IE-EH-K-K-O-HR-R---FY---D---
    ---F---F---D-REV---E-N---CO-IE-EH-K-K-O-HR-R---FY---D---
    ---F---F---D-REV---E-N---CO-IE-EH-K-K-O-HR-R---FY---D---
    ---F---F---D-REV---E-N---CO-IE-EH-K-K-O-HR-R---FY---D---
    ---T---C---F-D-REV---A-NC-M-Q-T-E-M-K-S-RR-I---Y---D---
    ---L---C---F-D-VEV-K-E-NC-M-Q-IE-ED-I-K-LR-L-K-WY-D*-D-
    ---C-E-F---D-KEV---E-NC-M-Q-IE-A-M-K-S-RK-I---FY---D---
    ---I---C---F-E-D-ADV---TG-N---TCO-IE-ADK-R-K-G-RR-I---R-DFY---D-
    ---T---C---F-D-SDV---N-NC-M-Q-IE-E-K-Q-A-RK-I---FY---D---
    ---I---L---F-S-AEV---TE-TN---TCO-IE-E-K-V-H-LV-K---Y---D---
    ---I---L---F-S-EEV-K-E-NC-M-Q-IE-ED-K-K-S-R-R-FY-QN*-D-
    ---F---L---F-DSNEV---E-N---TCO-IE-E-K-K-A-RR-I-Q-FY---D-
    ---I---F---FE-D-KEA-K-DTE-NC-ACO---K-V-N---E-K-IK-FY---D-
    ---T---F---F-MD-AEV---E-N---TCO-IE-N-I---S-RR-L-Q-FY---D-
    ---T---F---F-D-SEV---DT-N---MCQQ-I---R-K-PH-LK-R-Q-K-FY---D-
    ---X---F---D-ADV---SN-N---TCO-IE-DDK-M-S-RR-I---R-DFY---D-
    ---T---F---D-TEV---S-TN---TCO-IE-SDK-I-K-S-RR-I---DFY---D-
    ---X---F---D-AEV---E-N---AM-E-S-G-M-K-Q-RK-T---Y---D-
    ---E-F-L---F-E-EEV---E-N---M-Q-IE-E-K-M-K-LR-L-K-FY---D-
    ---T---F---F-D-EQV-A-E-N---M-Q-IE-E-K-M-K-M-F-Y-D*-D-
    ---F---F---D-QV-E-N---M-Q-IE-E-K-M-K-M-F-Y---D-
    ---T---F---F-D-REV---A-N---TCO-IE-EH-K-K-L-RR-M-FY---D-
    ---I---C---F-D-REV---T-NC-JC-AE-E-G-M-S-RR-I-K-FY---D-
    ---T---C---F-D-REV---E-N---M-Q-IE-ED-K-Q-A-RK-I-Q-FY---D-
    ---T---F---C---F-D-GEVK---E-NC-M-Q-IE-ED-K-Q-A-RK-I-Q-FY---D-
    ---I---C---F-D-AEV-K-D-E-N---MCQ-IE-E-T-M-K-T-K-FY---D-
    ---I---M---F---D-EV-K-E-NC-M-Q-A-EDK-V-K-K-LR-I-R-FY-OD*DCCHE-
    ---V---F---D-EEV---E-N---M-Q-IE-E-K-Q-AT---FH---D-
    ---I---C---F-L-Q-QV-K-TE---I---E-T-V-K-K-H-FY---D-
    ---I---F---MD-ADV---E-N---TCO-IE-ED-V-S-RR-I---Y---D-
    ---I---L---F-D-EEV-K-E-G-M-Q-ADK-M-K-S-LR-I-RF-FY-ODD*-D-
    ---T---C---F-D-KEV---E-N---M-Q-IE-ED-K-Q-A-RK-I-Q-FY---D-
    ---I---F---N-EEV-R-TE-NC-J-Q-IE-ED-K-Q-S-R-DFY-NCCI-EL-
    ---T---C---F-D-EEV-K-E-NC-M-Q-IE-ED-M-K-RH-SK---Y---D-
    ---T---F---C---F-D-OEV---D-SC-M-Q-IE-E-M-K-S-RR-I---R-Y---D-
    ---I---F---D-OEV-K-TE-G-NC---TCO-IE-E-K-Q-K-S-RK-I---FY---D-
    ---E-F---FX-D-QV-K-E-N---M-Q-IE-E-K-M-K-LR-M-K-K-D-Y---D-
    ---I---C---F-D-KEV---D-SC-I-Q-OE-E-G-M-K-A-RT-T-FY---D-
    ---T---C---F-D-AEV---E-NC-M-Q-IE-D-I-K-S-RR-I---Y---D-
    ---T---F---C---F-D-REV---E-NC-M-Q-AE-E-K-M-K-A-RT-M-FY---D-
    ---I---C---F-D-REV---TT-E-NC-M-Q-IE-E-K-M-K-A-RK-R-FY---D-
    ---I---F---D-RE---E-NC-CO-IE-EH-R-K-Q-HR-R---FY-R---D-
    ---I---C---F-D-REV---E-NC-M-Q-IE-N-M-K-A-RK-I-K-Y---D-
    X.
    ---R-I---C---F---TK---DIEE-SSC-MOO-RVE---M-K-S-M-K---D-
    ---I---F---F---KEV---TE-NC-CO-IE-EH-M-K-H-S-K-MY---D-
    ---F---F---D-REV---E-NC-M-Q-IE-EHG-R-K-Q-HR-R---FY---D-
    ---T---F---F---MD-AE---TE-N---TCO-LE-ED-V-K-D-R-FY---D-
    ---I---F---F---D-REV---E-NC-CO-IE-EH-M-K-Q-HR-R---FY---D-
    ---F---F---D-REV---S-E-NC-M-Q-IE-EHG-M-K-M-HR-M-Q-FY---D-
    ---I---C---F-D-REV---D-KC-M-Q-IE-EH-M-K-A-RK-I-X-Y---D-
    ---I---F---C---F-D-RDV---D-R-DS-C-I-Q-IE-EH-M-K-A-RK-I-D-Y---D-
    ---T---F---C---F-D-REV---ET---NC-MN---E-E-G-R-E-S-RK-I---Y---D-
    ---T---W---F---D-E-V---K-S-M---E-A-K-V-K-S-G---Y---D-
    ---I---C---F---L-EEV-K-SE#-NC-AN---PG-V-M-K-K-M-K---Y-N#-D-
    ---I---C---F---A-NC---M-Q-V---M-K-K---D-
    ---T---F---LF---SEEA-RLG-TC-RAN-ACA-FE-THK-I-M-K-RS-GNT-MIT-L-Q-D#-
    ---T---F---MF---SEAEA-LG-H-RAK-ACN-FE-AHG-I-K-Q-RS-GLT-KIT-L-P-D-
    ---T---F---Y-L---TREEA-RLG-EC-RA-ACN-FE-HG-I-K-Q-RS-GNT-KIT-X-X-D-
    ---I---F---L---TAEAA-RLG-TR-DA-ACN-AE-AHG-I-K-Q-RT-GLT-LQK-L-P-#-
    ---T---F---MF---SEEAADXLG-H-RAK-ACN-FE-HG-I-K-Q-RS-GGT-MIT-HL-N-D-
    ---T---F---LF---SEAEA-LG-CVTAK-ACN-SE-HG-I-K-Q-RS-GNI-KVT-L-L-D-
    ---I---F---LF---SEAEA-LG-C-RA-ACN-YE-EHG-I-K-Q-RS-MT-LIT-D-
    ---T---F---LF---TE-EAKRLG-ECVRAH-ACA-FE-HG-I-K-K-S-GLT-SKVI-L-T-D-
    ---P---F---LF---SEAEA-LG-C-RA-ACN-FE-NHGQI-K-Q-RS-GST-MVTN-L-N-D-
    ---T---F---F---LSAEAV---E-D-NA-LCQ-V-GHK-Q-Q-S-RR-I-DFY-
    ---F---F---LSAEAV---E-D-NA-TCQ-V-DHK-V-S-RR-I-DFY-
    ---T---F---F---LSAEAV---TE-N---TCO-V-DHK-V-S-RR-K-F---D-
    ---E---LF---SEVEA-MGD-Q-KAK-ACTY-FS-HK-I-V-K-S-GRE-LQK-L-I-D-#-
    ---E-F---LF---SEVEV-MGDEQ-KAK-ACTY-DS-HH-I-M-K-RS-GER-LQK-L-T-D-#-
    EE---CR---F---S---PPDD-RNI-ACT-DG-HK-I-R-E-AS-MRR-I-R-RD*-D-
    ---I---I---G-F---LT-EEV---NL---TCO-IE-EDK-I-KY-Q-LR-I-S-Y---D-
    A---CR---F---LS-EEV---N-E-NI-MCT-E-ED-T-K-LR-I-Q-Y---D-
    ---I---CR---LF---ED-RNL---ACS-K---G-L-I-K-RR-I-R-#-D-
    ---Y---F---LT-EEV---Q---TNEE---E-HG-I-Q-TE-RR-I-R-#-D-
    ---I---LF---SEAEA-MGTSN-KAM-ACS-R-YG-I-V---RA-GTT-LQK-L-KRD-#-
    ---X---Y---LF---SEAEA-MG-N-RAK-ACN-YG-EHGQI-K-Q-RS-GNN-LA-N-#-
    ---W---MF---SEAEA-YMG-D-RAK-ACTY-FS-HK-I-V-KY-RS-GTQ-LMK-L-I-D-#-
  
```

normal Nef end

205
221
210
207
207
212
210
208
203
220
209
206
207
207
208
208
200
208
201
217
206
207
207
204
208
130
209
224
137
208
208
212
212
223
207
205
203
206
131
132
208
213
207
215
208
213
207
206
99
205
211
205
206
213
206
84
208
206
129
202
207
213
208
208
212
208
208
208
97
205
211
209
216
210
215
211
212
214
218
216
214
103
214
133
213
93
212
219
211
201
207
211
198
210
213
213
213