

VII

PLV Proteins

Contents

VII-1	Introduction	405
VII-2	Sequences	406
VII-3	Alignments	416
	VII-3.1 Gag	416
	VII-3.2 Pol	420
	VII-3.3 Vif	427
	VII-3.4 Vpx	429
	VII-3.5 Vpr	430
	VII-3.6 Tat	431
	VII-3.7 Rev	433
	VII-3.8 Vpu	435
	VII-3.9 Env	436
	VII-3.10Nef	443

VII-1 Introduction

The selection of Primate Lentivirus Protein sequences for the following alignments was based on the sequences in the complete genome alignment as a starting alignment, and complete or nearly complete genes from other isolates were added if they increased the diversity of samples represented.

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

VII-2 Sequences

Sequences included in the PLV protein alignments.

Name	Accession	Proteins	Author	Reference
H1B.FR.83.HXB2	K03455	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Wong-Staal, F.	<i>Nature</i> 313 (6000):277-284 (1985)
H1A1.UG.85.U455_U455A	M62320	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Oram, J.D.	<i>ARHR</i> 6 (9):1073-1078 (1990)
H1C.ET.86.ETH2220	U46016	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Salminen, M.O.	<i>ARHR</i> 12 (14):1329-1339 (1996)
H1D.CD.84.84ZR085	U88822	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Gao, F.	<i>J Virol</i> 72 (7):5680-5698 (1998)
H1F1.BE.93.VI850	AF077336	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Laukkanen, T.	<i>Virology</i> 269 (1):95-104 (2000)
H1G.SE.93.SE6165_G6165	AF061642	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Carr, J.K.	<i>Virology</i> 247 (1):22-31 (1998)
H1H.CF.90.056	AF005496	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Gao, F.	<i>J Virol</i> 72 (7):5680-5698 (1998)
H1J.SE.93.SE9280_7887	AF082394	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Laukkanen, T.	<i>ARHR</i> 15 (3):293-297 (1999)
H1K.CM.96.96CM_MP535	AJ249239	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Triques, K.	<i>ARHR</i> 16 (2):139-151 (2000)
H1O.BE.87.ANT70	L20587	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Vanden Haesevelde, M.	<i>J Virol</i> 68 (3):1586-1596 (1994)
H1O.CM.91.MVP5180	L20571	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Gurtler, L.G.	<i>J Virol</i> 68 (3):1581-1585 (1994)
H1O.CM.94.BCF06	AB485666	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Takekawa	Unpublished

Name	Accession	Proteins	Author	Reference
H1O.CM.96.96CMA102	AY169803	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H1O.CM.98.98CMA104	AY169802	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H1O.CM.99.99CMU4122	AY169815	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H1O.CM.x.pCMO2_3	AY618998	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Tebit, D.M.	<i>Virology</i> 326 (2):329-339 (2004)
H1O.FR.92.VAU	AF407418	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Vartanian, J.P.	<i>J Gen Virol</i> 83 (Pt 4):801-805 (2002)
H1O.SN.99.99SE_MP1299	AJ302646	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Toure-Kane, C.	<i>ARHR</i> 17 (12):1211-1216 (2001)
H1O.US.99.99USTWLA	AY169814	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H1O.US.x.I_2478B	AB485668	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Takekawa	Unpublished
H1N.CM.02.DJO0131	AY532635	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Bodelle, P.	<i>ARHR</i> 20 (8):902-908 (2004)
H1N.CM.95.YBF30	AJ006022	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Simon, F.	<i>Nat Med</i> 4 (9):1032-1037 (1998)
H1N.CM.97.YBF106	AJ271370	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Roques, P.	<i>AIDS</i> 18 (10):1371-1381 (2004)
H1P.CM.06.U14788	HQ179987	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Vallari, A.	<i>J Virol</i> 85 (3); 1403-7 (2011)
H1P.FR.06.RBF168	GQ328744	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Plantier, J.-C.	<i>Nat Med</i> 15 (8); 871-2 (2009)

Name	Accession	Proteins	Author	Reference
CPZ.CD.90.ANT	U42720	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Vanden Haesevelde, M.M.	<i>Virology</i> 221 (2):346-350 (1996)
CPZ.CD.06.BF1167	JQ866001	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Li, Y.	<i>J Virol</i> 86 (19):10776-10791 (2012)
CPZ.CM.05.SIVcpzDP943	EF535993	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Van Heuverswyn, F.	<i>Virology</i> 368 (1):155-171 (2007)
CPZ.CM.05.SIVcpzEK505	DQ373065	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Keele, B.F.	<i>Science</i> 313 (5786):523-526 (2006)
CPZ.CM.05.SIVcpzMB897	EF535994	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Van Heuverswyn, F.	<i>Virology</i> 368 (1):155-171 (2007)
CPZ.CM.05.SIVcpzMT145	DQ373066	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Keele, B.F.	<i>Science</i> 313 (5786):523-526 (2006)
CPZ.CM.98.CAM3	AF115393	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Corbet, S.	<i>J Virol</i> 74 (1):529-534 (2000)
CPZ.CM.98.CAM5	AJ271369	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Muller-Trutwin, M.C.	<i>J Med Primatol</i> 29 (3-4); 166-72 (2000)
CPZ.GA.88.GAB1	X52154	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Huet, T.	<i>Nature</i> 345 (6273):356-359 (1990)
CPZ.GA.88.SIVcpzGAB2	AF382828	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Bibollet-Ruche, F.	<i>ARHR</i> 20 (12):1377-1381 (2004)
CPZ.TZ.00.TAN1	AF447763	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Santiago, M.L.	<i>J Virol</i> 77 (3):2233-2242 (2003)
CPZ.TZ.01.TAN2	EF394357	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Takehisa, J.	<i>J Virol</i> 81 (14):7463-7475 (2007)
CPZ.TZ.09.UG38	JN091690	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Rudicell, R.S.	<i>J Virol</i> 85 (19); 9918-28 (2011)

Name	Accession	Proteins	Author	Reference
CPZ.TZ.06.SIVcpzTAN13	JQ768416	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Takehisa	Unpublished
CPZ.US.85.US_Marilyn	AF103818	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Gao, F.	<i>Nature</i> 397 (6718):436-441 (1999)
MAC.US.x.239	M33262	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Kestler, H.	<i>Science</i> 248 (4959):1109-1112 (1990)
SMM.CI.79.SIVsmCI2	JX860430	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.LR.89.SIVsmLIB1	JX860431	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.SL.92.SIVsmSL92A	JX860432	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
STM.US.89.STM_37_16	M83293	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Novembre, F.J.	<i>Virology</i> 186 (2):783-787 (1992)
H2A.DE.x.BEN	M30502	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Kirchhoff, F.	<i>Virology</i> 177 (1):305-311 (1990)
H2A.PT.x.ALI	AF082339	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Reeves, J.D.	<i>J Virol</i> 73 (9); 7795-804 (1999)
H2A.SN.x.ST_HIV_2_ST	M31113	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Kumar, P.	<i>J Virol</i> 64 (2):890-901 (1990)
H2B.CI.x.EHO	U27200	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Rey-Cuille, M.A.	<i>Virology</i> 202 (1):471-476 (1994)

Name	Accession	Proteins	Author	Reference
H2B.GH.86.D205_ALT	X61240	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Dietrich, U.	<i>Nature</i> 342 (6252):948-950 (1989)
H2G.CI.92.Abt96	AF208027	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Brennan, C.A.	<i>ARHR</i> 13 (5):401-404 (1997)
H2U.FR.96.12034	AY530889	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Damond, F.	<i>ARHR</i> 20 (6):666-672 (2004)
GOR.CM.04.SIVgorCP684con	FJ424871	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Takehisa, J.	<i>J Virol</i> 83 (4):1635-1648 (2009)
GOR.CM.07.SIVgorCP2135con	FJ424863	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Takehisa, J.	<i>J Virol</i> 83 (4):1635-1648 (2009)
VER.DE.x.AGM3	M30931	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Baier, M.	<i>Virology</i> 176 (1):216-221 (1990)
VER.KE.x.9063	L40990	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Hirsch, V.M.	<i>J Virol</i> 69 (2):955-967 (1995)
VER.KE.x.AGM155	M29975	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Johnson, P.R.	<i>J Virol</i> 64 (3):1086-1092 (1990)
VER.KE.x.TYO1_patent	DJ048201	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Omori, T.	Patent: WO 2007049749-A 13 03-MAY-2007; Dनावेक CORPORATION
SAB.SN.x.SAB1	U04005	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Jin, M.J.	<i>EMBO J</i> 13 (12):2935-2947 (1994)
GRV.ET.x.GRI_677_gri_1	M66437	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Fomsgaard, A.	<i>Virology</i> 182 (1):397-402 (1991)
TAN.UG.x.SIVagmTAN1	U58991	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Soares, M.A.	<i>Virology</i> 228 (2):394-399 (1997)

Name	Accession	Proteins	Author	Reference
DRL.x.x.FAO	AY159321	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Hu, J.	<i>J Virol</i> 77 (8):4867-4880 (2003)
MND-2.GA.x.M14	AF328295	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Souquiere, S.	<i>J Virol</i> 75 (15):7086-7096 (2001)
MND-2.CM.98.CM16	AF367411	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Takehisa, J.	<i>ARHR</i> 17 (12):1143-1154 (2001)
MND-2.x.x.5440	AY159322	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Hu, J.	<i>J Virol</i> 77 (8):4867-4880 (2003)
MNE.US.x.MNE027	U79412	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Kimata, J.T.	<i>J Virol</i> 72 (1):245-256 (1998)
RCM.CM.00.SIVagi_00CM312	HM803690	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Ahuka-Mundeke, S.	<i>J Gen Virol</i> 2010 Aug 25
RCM.CM.02.SIVrcm_02CM8081	HM803689	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Ahuka-Mundeke, S.	<i>J Gen Virol</i> 2010 Aug 25
RCM.GA.x.SIVRCMGAB1	AF382829	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Bibollet-Ruche, F.	<i>J Virol</i> 78 (14):7748-7762 (2004)
RCM.NG.x.NG411	AF349680	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	<i>J Virol</i> 75 (24):12014-12027 (2001)
SYK.KE.x.KE51	AY523867	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Bibollet-Ruche, F.	<i>J Virol</i> 78 (14):7748-7762 (2004)
SYK.KE.x.SYK173_COMGNM	L06042	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Hirsch, V.M.	<i>J Virol</i> 67 (3):1517-1528 (1993)

Name	Accession	Proteins	Author	Reference
LST.CD.88.SIVlhoest447	AF188114	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	<i>J Virol</i> 74 (8):3892-3898 (2000)
SUN.GA.98.L14	AF131870	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	<i>J Virol</i> 73 (9):7734-7744 (1999)
MND-1.GA.x.MNDGB1	M27470	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Tsujimoto, H.	<i>Nature</i> 341 (6242):539-541 (1989)
LST.CD.88.SIVlhoest485	AF188115	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	<i>J Virol</i> 74 (8):3892-3898 (2000)
LST.CD.88.SIVlhoest524	AF188116	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	<i>J Virol</i> 74 (8):3892-3898 (2000)
LST.KE.x.lho7	AF075269	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Hirsch, V.M.	<i>J Virol</i> 73 (2):1036-1045 (1999)
DEB.CM.04.SIVdeb04CMPF3061	FJ919724	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Aghokeng, A.F.	<i>Infect Genet Evol</i> 10 (3); 386-96 (2010)
DEB.CM.99.CM40	AY523865	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Bibollet-Ruche, F.	<i>J Virol</i> 78 (14):7748-7762 (2004)
DEB.CM.99.CM5	AY523866	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Bibollet-Ruche, F.	<i>J Virol</i> 78 (14):7748-7762 (2004)
DEN.CD.x.CD1_CMO580407	AJ580407	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Dazza, M.C.	<i>J Virol</i> 79 (13):8560-8571 (2005)
GSN.CM.99.CN166	AF468659	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Cournaud, V.	<i>J Virol</i> 76 (16):8298-8309 (2002)
GSN.CM.99.CN71	AF468658	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Cournaud, V.	<i>J Virol</i> 76 (16):8298-8309 (2002)
MON.CM.99.L1_99CML1	AY340701	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Cournaud, V.	<i>J Virol</i> 77 (23):12523-12534 (2003)

Name	Accession	Proteins	Author	Reference
MON.NG.x.NG1	AJ549283	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Barlow, K.L.	<i>J Virol</i> 77 (12):6879-6888 (2003)
MUS-1.CM.01.CM1239	EF070330	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Aghokeng, A.F.	<i>Virology</i> 360 (2):407-418 (2007)
MUS- 1.CM.01.SIVmus_01CM1085	AY340700	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Courgnaud, V.	<i>J Virol</i> 77 (23):12523-12534 (2003)
MUS-2.CM.01.CM1246	EF070329	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Aghokeng, A.F.	<i>Virology</i> 360 (2):407-418 (2007)
MUS-2.CM.01.CM2500	EF070331	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Aghokeng, A.F.	<i>Virology</i> 360 (2):407-418 (2007)
TAL.CM.00.266	AY655744	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	<i>Virology</i> 349 (1):55-65 (2006)
TAL.CM.01.8023	AM182197	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	<i>Virology</i> 349 (1):55-65 (2006)
OLC.CI.97.97CI12	FM165200	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	<i>J Virol</i> 83 (1):428-439 (2009)
WRC.CI.97.97CI14	AM745105	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	<i>J Virol</i> 83 (1):428-439 (2009)
WRC.CI.98.98CI04	AM713177	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	<i>J Virol</i> 83 (1):428-439 (2009)
WRC.GM.05.Pbt_05GM_X02	AM937062	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Locatelli, S.	<i>Virology</i> 376 (1):90-100 (2008)
COL.CM.x.CGU1	AF301156	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Courgnaud, V.	<i>J Virol</i> 75 (2):857-866 (2001)
MAC.US.x.251_1A11	M76764	Vpx	Marthas, M.L.	<i>J Med Primatol</i> 18 (3-4):311-9 (1989)
MAC.US.x.251_BK28	M19499	Vpx	Hirsch, V.	<i>Cell</i> 49 (3):307-319 (1987)
SMM.SL.92.SIVsmSL92B	JX860433	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.SL.92.SL92B	AF334679	Vpx	Chen, Z.	<i>J Virol</i> 70 (6):3617-3627 (1996)

Name	Accession	Proteins	Author	Reference
SMM.US.04.G078	JX860415	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.04.G932	JX860416	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.04.M919	JX860417	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.04.M922	JX860418	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.04.M923	JX860419	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.04.M926	JX860420	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.04.M934	JX860421	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.04.M935	JX860422	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.04.M940	JX860423	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.04.M946	JX860424	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.04.M947	JX860425	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.04.M949	JX860426	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.04.M950	JX860427	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.04.M951	JX860428	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.04.M952	JX860429	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.05.D215	JX860413	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.06.FTq	JX860414	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.11.SIVsmE660_FL10	JQ864084	Vpx	Wu, F.	<i>J Virol</i> 86 (16): 8835-47 (2012)
SMM.US.11.SIVsmE660_FL14	JQ864087	Vpx	Wu, F.	<i>J Virol</i> 86 (16): 8835-47 (2012)
SMM.US.11.SIVsmE660_FL6	JQ864085	Vpx	Wu, F.	<i>J Virol</i> 86 (16): 8835-47 (2012)
SMM.US.11.SIVsmE660_FL8	JQ864086	Vpx	Wu, F.	<i>J Virol</i> 86 (16): 8835-47 (2012)
SMM.US.86.CFU212	JX860407	Vpx	Fischer, W.	<i>J Virol</i> 86 (24):13217-13231 (2012)
SMM.US.x.H9	M80194	Vpx	Courgnaud, V.	<i>J Virol</i> 66 (1):414-419 (1992)
SMM.US.x.PGM53	AF077017	Vpx	Novembre, F.J.	<i>J Virol</i> 72 (11):8841-8851 (1998)
SMM.US.x.SIVsmH635F_L3	DQ201172	Vpx	Kuwata, T.	<i>J Virol</i> 80 (3):1463-75 (2006)
SMM.x.x.pE660.CG7G	JX648292	Vpx	Lopker	Unpublished
SMM.x.x.pE660.CG7V	JX648291	Vpx	Lopker	Unpublished
H1B.US.90.WEAU160_GHOSH	U21135	Vpu	Wei, X.	<i>Nature</i> 422 (6929):307-312 (2003)
H101_AE.TH.90.CM240	U54771	Vpu	Carr, J.K.	<i>J Virol</i> 70 (9):5935-5943 (1996)
H102_AG.NG.x.IBNG	L39106	Vpu	Howard, T.M.	<i>ARHR</i> 10 (12):1755-1757 (1994)
H103_AB.RU.97.KAL153_2	AF193276	Vpu	Liitsola, K.	<i>ARHR</i> 16 (11):1047-1053 (2000)
H104_cpx.CY.94.94CY032_3	AF049337	Vpu	Gao, F.	<i>J Virol</i> 72 (12):10234-10241 (1998)
H1O.CM.96.96CMABB009	AY169806	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H1O.CM.96.96CMABB637	AY169810	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H1O.CM.97.97CMABB447	AY169813	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H1O.CM.97.97CMABB497	AY169809	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H1O.CM.98.98CMA105	AY169816	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H1O.CM.98.98CMABB141	AY169807	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H1O.CM.98.98CMABB197	AY169811	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H1O.CM.98.98CMABB212	AY169804	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H1O.CM.98.98CMU2901	AY169812	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H1O.CM.98.98CMU5337	AY169808	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H1O.CM.x.pCMO2_5	AY623602	Vpu	Tebit, D.M.	<i>Virology</i> 326 (2):329-339 (2004)
H1O.SN.99.99SE_MP1300	AJ302647	Vpu	Toure-Kane, C.	<i>ARHR</i> 17 (12):1211-1216 (2001)
H1N.CM.02.SJGddd	GQ324959	Vpu	Vallari, A.	<i>ARHR</i> 26 (1):109-115 (2010)
H1N.CM.04.04CM_1015_04	DQ017382	Vpu	Yamaguchi, J.	<i>ARHR</i> 22 (1):83-92 (2006)
H1N.CM.04.04CM_1131_03	DQ017383	Vpu	Yamaguchi, J.	<i>ARHR</i> 22 (1):83-92 (2006)
H1N.CM.06.U14296	GQ324962	Vpu	Vallari, A.	<i>ARHR</i> 26 (1):109-115 (2010)
H1N.CM.06.U14842	GQ324958	Vpu	Vallari, A.	<i>ARHR</i> 26 (1):109-115 (2010)
CPZ.CM.01.SIVcpzCAM13	AY169968	Vpu	Nerrienet, E.	<i>J Virol</i> 79 (2):1312-9 (2005)
CPZ.CM.05.SIVcpzEK505.c2	JN835460	Vpu	Bibollet-Ruche, F.	<i>J Clin Invest</i> 122 (5):1644-1652 (2012)
CPZ.CM.05.SIVcpzLB7	DQ373064	Vpu	Keele, B.F.	<i>Science</i> 313 (5786):523-526 (2006)

Name	Accession	Proteins	Author	Reference
CPZ.CM.05.SIVcpzMB66	DQ373063	Vpu	Keele, B.F.	<i>Science</i> 313 (5786):523-526 (2006)
CPZ.CM.05.SIVcpzMB897.c2	JN835461	Vpu	Bibollet-Ruche, F.	<i>J Clin Invest</i> 122 (5):1644-1652 (2012)
CPZ.CM.05.SIVcpzMT145.c2	JN835462	Vpu	Bibollet-Ruche, F.	<i>J Clin Invest</i> 122 (5):1644-1652 (2012)
CPZ.TZ.02.TAN3_1	DQ374658	Vpu	Takehisa, J.	<i>J Virol</i> 81 (14):7463-7475 (2007)
CPZ.TZ.06.TAN5	JN091691	Vpu	Rudicell, R.S.	<i>J Virol</i> 85 (19); 9918-28 (2011)
GOR.CM.07.SIVgor2139_287	FJ424866	Vpu	Takehisa, J.	<i>J Virol</i> 83 (4):1635-1648 (2009)

H1B.FR.83.HXB2
H1A1.UG.85.U455 U455A
H1C.ET.86.ETH2220
H1D.CD.84.84Z8685
H1E.BI.93.18759
H1G.SE.93.SE6165 G6165
H1H.CF.90.056
H1J.SE.93.SE9280 7887
H1K.CM.96.96CM MP535
H1O.RE.87.ANT78
H10.CM.91.MVP5180
H10.CM.94.BCF06
H10.CM.96.96CMA102
H10.CM.98.98CMA104
H10.CM.99.99CMU4122
H10.CM.x.CM02 3
H10.FR.92.VAU
H10.SN.99.99SE MP1299
H10.US.99.99USTWLA
H10.US.x.1.24788
H1N.CM.02.DJ00131
H1N.CM.95.YBF30
H1N.CM.97.YBF106
H1P.CM.06.U14788
H1P.FR.06.RBF168
CPZ.CD.90.ANT
CPZ.CD.06.BF1167
CPZ.CM.05.SIVcpzDP943
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzMB897
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CAM5
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.SIVcpzGAB2
CPZ.TZ.00.TAN1
CPZ.TZ.01.TAN2
CPZ.TZ.09.UG38
CPZ.TZ.06.SIVcpzTAN13
CPZ.US.85.US Marilynn

Table with columns for sequence identifiers, amino acid sequences, and accession numbers. The table is organized into sections: CyPA binding (lines 1-410), major homology region (lines 411-810), and PLV Proteins (lines 811-870). Each row represents a different HIV sequence variant.

	p24 end_p2 start	p2 end_p7 start	Zn motif	Zn motif	
H1B.FR.83.HXB2	RAEQASQEVKNMWTETLLVQNANPDCKITLKALGPAATLEEMMTACQGVGPGHKARVLAEAMSVNTSAT	...IMMORGNFRNOR	KIVKCFNCGKEGTARNCRAPRKK	...GCWKCGKEGHQMKD	425
H1A1.UG.85.U455.U455A	T-D	S-R	G-G	QOTS	GP
H1C.ET.86.ETH2220	T-D	D	R	G-S	A
H1D.CD.84.84Z8085				A-SASAA	KS-KGPK
H1F1.BE.93.VI850	D-G-D				KS-KGT
H1G.SE.93.SE6165.G6165	D-G-D	R-OG-S	S	ASGA-AA	S-KGP
H1H.CF.90.056	T-D	D	R-OG-SI	TN-A	K-KG
H1J.SE.93.SE9280.7887	T-D	D		TN	K-DHK
H1K.CM.96.96CM.MP535	T-D	D		PV	V-K-KGH
H10.BE.87.ANT70		S	0	S	V
H10.CM.91.MVP5180		S	0	S	V
H10.CM.94.BCF06		S	0	S	V
H10.CM.96.96CMA102		S	0	S	V
H10.CM.98.98CMA104		S	0	S	V
H10.CM.99.99CMU4122		S	0	S	V
H10.CM.x.CM02.3		S	0	S	V
H10.FR.92.VAU		S	0	S	V
H10.SN.99.99SE.MP1299		S	0	S	V
H10.US.99.99USTWLA		S	0	S	V
H10.US.x.I.24788		S	0	S	V
H1N.CM.02.DJ00131		S	0	S	V
H1N.CM.95.YBF30		S	0	S	V
H1N.CM.97.YBF106		S	0	S	V
H1P.CM.06.U14788		S	0	S	V
H1P.FR.06.RBF168		S	0	S	V
CPZ.CD.90.ANT70	P-A	D	I	H	T-S
CPZ.CD.06.BF1167	P-A	D	I	H	T-S
CPZ.CM.05.SIVcpzDP943	T-T-D-X	N-R	G-S	A	A
CPZ.CM.05.SIVcpzEK505	T-D	N-R	G-S	A	A
CPZ.CM.05.SIVcpzMB897	T-D	N-R	G-S	A	A
CPZ.CM.05.SIVcpzMT145	T-D	N-R	G-S	A	A
CPZ.CM.98.CAM5	T-D	N-R	G-S	A	A
CPZ.GA.88.GAB1	D	R	G	S	A
CPZ.GA.88.SIVcpzGAB2	D	R	G	S	A
CPZ.TZ.00.TAN1	P-T	H	OG	L	S-KI
CPZ.TZ.01.TAN2	P-T	H	OG	L	S-KI
CPZ.TZ.09.CG38	P-T	H	OG	L	S-KI
CPZ.TZ.06.SIVcpzTAN13	P-T	H	OG	L	S-KI
CPZ.US.85.US.Marilyn	P-T	H	OG	L	S-KI

	p27 end_p2 start	p2 end_p8 start			
MAC.US.x.239	---TDA--	0-I	LV-G-VNP	L---	0-LM-LKEALAPVPIP
SMM.CI.79.SIVsmCI2	---TDA--	0-I	LV-G-MNP	L-I	0-LM-LREAFOPGLP
SMM.LR.89.SIVsmL1B1	---TDA--	0-I	LV-G-MNP	L	0-LM-LKEAFOP.SLP
SMM.SL.92.SIVsmSL92A	---TDA--	0-I	LV-G-MNP	L	0-LM-LKEALGPTLP
STM.US.89.STM.37.16	---DPS--	R-I	LV-G-MNP	L	0-LM-LKE.FOPDPLP
H2A.DE.x.ANT70	---TDA--	0-I	LV-G-MNP	L	0-LM-LKEAM.PIP
H2A.PT.x.ALI	---TDA--	0-I	LV-G-MNP	L	0-LM-LKEAMP.PIP
H2A.SN.x.ST.HIV.2.ST	---TDA--	0-I	LV-G-MNP	L	0-LM-LKEAMP.PIP
H2B.CI.x.EH0	---TDA--	0-I	LV-G-MNP	L	0-LM-LKEALTPS.NP
H2B.GH.86.D205.ALT	---TDA--	0-I	LV-G-MNP	L	0-LM-LKEAL.PIP
H2B.CI.92.Abt.96	---TDA--	0-I	LV-G-MNP	L	0-LM-LKEAL.PTALP
H2U.FR.96.12034	---TDA--	0-I	LV-G-MNP	L	0-LM-LKEALOPMPTP
GOR.CM.04.SIVgorCP684con	---D	I	G	N	I
GOR.CM.07.SIVgorCP2135con	---D	I	G	N	I
VER.DE.x.AGM3	---G	0	S	I	V
VER.KE.x.9063	---G	0	S	I	V
VER.KE.x.AGM155	---G	0	S	I	V
VER.KE.x.TY01.patent	---G	0	S	I	V
SAB.SN.x.SAB1	---TDA--	0S	I	V	G
GRV.ET.x.GRI.677.gri.1	---P	D	0	I	L
TAN.UG.x.SIVagmTAN1	---TP	D	0	I	L
DRL.x.x.FA07	---HTDA--	0	I	LV	G-IHP
MND.2.GA.x.M14	---RTDPO	0	I	S	G-MNPS
MND.2.CM.98.CM16	---RTDPO	0	I	S	G-MNPS
MND.2.x.x.5440	---RTDPO	0	I	AT	G-MNP
MNE.US.x.MNE027	---TDA--	0	I	LV	G-MNP
RCM.CM.00.SIVacm.00CM312	---DPA	0S	I	V	G-MNP
RCM.CM.02.SIVacm.02CM0801	---DPA	0S	I	V	G-MNP
RCM.GA.x.SIVRCMGA1	---DPA	0S	I	V	G-MNP
RCM.NG.x.NG411	---DPA	0S	I	V	G-MNP
SVK.KE.x.KE51	---DPA	0S	I	V	G-MNP
SVK.KE.x.SYK173.COMGNM	---DPS	G	L	0	I
LST.CD.88.SIVlhoest447	---GG-H	E-K	KM	RMVI	EG-S
SUN.GA.98.L14	---SG-E	M-QOH	IE	E	0-RS
MND.1.GA.x.MNDGB1	---GG-H	E-K	KM	RMVI	EG-S
LST.CD.88.SIVlhoest485	---GG-H	E-K	KM	RMVI	EG-S
LST.CD.88.SIVlhoest524	---GG-H	E-K	KM	RMVI	EG-S
DEB.CM.04.SIVdeb04CMPF3061	---D	A	T	MI	R-LH
DEB.CM.99.CM40	---D	A	T	MI	R-LH
DEB.CM.99.CM5	---D	A	T	MI	R-LH
DEN.CD.x.CD1.CM0580407	---D-S	L	TS	0	E-NP
GSN.CM.99.CM166	---D	A	0S	I	V
GSN.CM.99.CM71	---D	A	0S	I	V
MON.CM.99.L1.99CML1	---TD	A	0	S	D
MON.NG.x.NG1	---L	M	G	S	LQ
MUS.1.CM.01.CM1239	---SD	A	S	0	L
MUS.1.CM.01.SIVmus.01CM1085	---SD	A	S	0	L
MUS.2.CM.01.CM1246	---D	A	0S	I	V
MUS.2.CM.01.CM2500	---D	A	0S	I	V
TAL.CM.00.266	---TD	A	0	S	LQ
TAL.CM.01.8023	---TD	A	0	S	LQ
OLC.CI.97.97CI12	---G	T	K	I	GV
WRC.CI.97.97CI14	---GAP	0	DK	I	G
WRC.CI.98.98CI04	---GAP	0	DK	I	G
WRC.GM.05.Pbt.05GM.X02	---GVP	0	DK	I	G
COL.CM.x.CGU1	---P	AGP	I	A	R

Table with columns: p7 end p1 start, p1 end p6 start, Vpr binding, Vpr binding, p6 end, Gag end. Rows include various HIV sequences and motifs like ERQANFLGKIW, PSYKGRPGNLF, QSRPE, PTAPPPEESFRSG, VETTTTPOKOE, PIDKELY, PLT, SLRSLFGNDPSSQ, etc.

Table with columns for HIV sequence identifiers (e.g., H1B, H1A1, H1C), protease end p66, p51 RT start, M41L, K70R D67N, and D110 catalytic site. It shows amino acid alignments for various HIV strains.

Table with columns for HIV sequence identifiers (e.g., MAC, SMM, SMM.LR), protease end p51 RT start, and D catalytic site. It shows amino acid alignments for various HIV strains.

Table with 2 columns: protein ID and amino acid sequence. Includes entries for H1B, H1A1, H1C, etc., up to CPZ.US.85.US.Marilyn.

Table with 2 columns: protein ID and amino acid sequence. Includes entries for MAC.US.x.239, SMM.CI.79.SIVsmC12, etc., up to COL.CM.x.CGU1.

Sequence alignment for polymerase motif and DD catalytic site. Header includes K219Q and T215Y. Sequences are aligned across various protein variants.

Table with columns for accession numbers (e.g., H1B, H1A1, H1C), amino acid sequences, and p51 RT end_p15 RNase H start coordinates (601-608).

Table with columns for accession numbers (e.g., MAC, SMM, SMM.LR), amino acid sequences, and p51 RT end_p15 RNase H start coordinates (612-619).

H1B.FR.83.HXB2
H1A1.UG.85.U455.U455A
H1C.ET.86.ETH2220
H1D.CD.84.84Z8085
H1F1.BE.93.VI850
H1G.SE.93.SE6165.G6165
H1H.CF.90.056
H1J.DE.93.SE9280.7887
H1K.CM.96.96CM.MP535
H1L.BE.87.A0770
H1O.CM.91.MVP5180
H1O.CM.94.BCF06
H1O.CM.96.96CMA102
H1O.CM.98.98CMA104
H1O.CM.99.99CMU4122
H1O.CM.x.CM02.3
H1O.FR.92.VAU
H1O.SN.99.99SE.MP1299
H1O.US.99.99USTWLA
H1O.US.x.I.24788
H1N.CM.02.DJ00131
H1N.CM.95.YBF30
H1N.CM.97.YBF106
H1P.CM.06.U14788
H1P.FR.86.RBF168
CPZ.CD.90.A0770
CPZ.CD.06.BF1167
CPZ.CM.05.SIVcpzDP943
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzMB897
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CM02.3
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.SIVcpzGAB2
CPZ.TZ.00.TAN1
CPZ.TZ.01.TAN2
CPZ.TZ.09.UG38
CPZ.TZ.06.SIVcpzTAN13
CPZ.US.85.US.Marilyn

NRETlKLGAGYVTRGRQKVVTLTDDTNNQKTELQAIYLAQ05GLSEVNIIVTDSQYALGIIQA0PQDSESELVNIQIEQLIKKKEYVLWPAHKIGIGGNEQVDKLVSAQ.IR.KVLFLDGID.KAQDEHEK.YHNSWRAMASDFNLPPVVAKEIVASCDKCOLK.GEAM
H1B.FR.83.HXB2
H1A1.UG.85.U455.U455A
H1C.ET.86.ETH2220
H1D.CD.84.84Z8085
H1F1.BE.93.VI850
H1G.SE.93.SE6165.G6165
H1H.CF.90.056
H1J.DE.93.SE9280.7887
H1K.CM.96.96CM.MP535
H1L.BE.87.A0770
H1O.CM.91.MVP5180
H1O.CM.94.BCF06
H1O.CM.96.96CMA102
H1O.CM.98.98CMA104
H1O.CM.99.99CMU4122
H1O.CM.x.CM02.3
H1O.FR.92.VAU
H1O.SN.99.99SE.MP1299
H1O.US.99.99USTWLA
H1O.US.x.I.24788
H1N.CM.02.DJ00131
H1N.CM.95.YBF30
H1N.CM.97.YBF106
H1P.CM.06.U14788
H1P.FR.86.RBF168
CPZ.CD.90.A0770
CPZ.CD.06.BF1167
CPZ.CM.05.SIVcpzDP943
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzMB897
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CM02.3
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.SIVcpzGAB2
CPZ.TZ.00.TAN1
CPZ.TZ.01.TAN2
CPZ.TZ.09.UG38
CPZ.TZ.06.SIVcpzTAN13
CPZ.US.85.US.Marilyn

MAC.US.x.239
SMM.CI.79.SIVsMCI2
SMM.LR.89.SIVsMLB1
SMM.SL.92.SIVsML92A
STM.US.89.89STM.37.16
H2A.DE.x.AL1
H2A.PT.x.AL1
H2A.SN.x.ST.HIV.2.ST
H2B.CI.x.EH0
H2B.GH.86.D205.ALT
H2B.CI.92.A0770
H2U.FR.96.12034
GOR.CM.04.SIVgorCP684con
GOR.CM.07.SIVgorCP2135con
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TY01.patent
SAB.SN.x.SAB1
GRV.ET.x.GRI.677.gri.1
TAN.UG.x.SIVagmTAN1
DRL.x.x.FA07
MND.2.GA.x.M14
MND.2.CM.98.CM16
MND.2.x.x.5440
MNE.US.x.MNE027
RCM.CM.00.SIVag1.00CM312
RCM.CM.02.SIVrcm.02CM8081
RCM.CM.x.SIVRCMAGB1
RCM.NG.x.NG411
SVK.KE.x.KE51
SVK.KE.x.SYK173.COMGNM
LST.CD.88.SIVlhoest447
SUN.GA.L14
MND.1.GA.x.MNDGB1
LST.CD.88.SIVlhoest485
LST.CD.88.SIVlhoest524
LST.CD.88.SIVlhoest524
DEB.CM.04.SIVdeb04CMPF3061
DEB.CM.99.CM40
DEB.CM.99.CM5
DEN.CD.x.CD1.CM0580407
G5N.CM.99.CM166
G5N.CM.99.CM167
MON.CM.99.L1.99CML1
MON.NG.x.NG1
MUS.1.CM.01.CM1239
MUS.1.CM.01.SIVmus.01CM1085
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
TAL.CM.00.266
TAL.CM.01.80213
OLC.CI.97.97CI12
WRC.CI.97.97CI14
WRC.CI.98.98CI04
WRC.GM.05.Pbt.05GM.X02
COL.CM.x.CG1

MAC.US.x.239
SMM.CI.79.SIVsMCI2
SMM.LR.89.SIVsMLB1
SMM.SL.92.SIVsML92A
STM.US.89.89STM.37.16
H2A.DE.x.AL1
H2A.PT.x.AL1
H2A.SN.x.ST.HIV.2.ST
H2B.CI.x.EH0
H2B.GH.86.D205.ALT
H2B.CI.92.A0770
H2U.FR.96.12034
GOR.CM.04.SIVgorCP684con
GOR.CM.07.SIVgorCP2135con
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TY01.patent
SAB.SN.x.SAB1
GRV.ET.x.GRI.677.gri.1
TAN.UG.x.SIVagmTAN1
DRL.x.x.FA07
MND.2.GA.x.M14
MND.2.CM.98.CM16
MND.2.x.x.5440
MNE.US.x.MNE027
RCM.CM.00.SIVag1.00CM312
RCM.CM.02.SIVrcm.02CM8081
RCM.CM.x.SIVRCMAGB1
RCM.NG.x.NG411
SVK.KE.x.KE51
SVK.KE.x.SYK173.COMGNM
LST.CD.88.SIVlhoest447
SUN.GA.L14
MND.1.GA.x.MNDGB1
LST.CD.88.SIVlhoest485
LST.CD.88.SIVlhoest524
LST.CD.88.SIVlhoest524
DEB.CM.04.SIVdeb04CMPF3061
DEB.CM.99.CM40
DEB.CM.99.CM5
DEN.CD.x.CD1.CM0580407
G5N.CM.99.CM166
G5N.CM.99.CM167
MON.CM.99.L1.99CML1
MON.NG.x.NG1
MUS.1.CM.01.CM1239
MUS.1.CM.01.SIVmus.01CM1085
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
TAL.CM.00.266
TAL.CM.01.80213
OLC.CI.97.97CI12
WRC.CI.97.97CI14
WRC.CI.98.98CI04
WRC.GM.05.Pbt.05GM.X02
COL.CM.x.CG1

p66 RT, p15 RNase H end, p31 Integrase start

p15 RNase H end, p31 Integrase start

H1B_FR.83.HXB2	HGQVDCSPGIWLDCTHLEGGVILVAVHVASGYIEAEVIPAETGQETAYFLLLKLAGRPVPKTI	HTDNGSNFTGATVRAACWAGIKOEFGIPYNPQSQGVSMNKLKIIIGQVRDQAEHLKTAVMQAVIHFNRKKGIGGYSAGER	IVDIATDIQTKELQK0I	932
H1A1.UG.85.U455.U455A				931
H1C.ET.86.ETH2220				930
H1D.CD.84.84ZRB85				929
H1E.BE.93.VB59				928
H1G.SE.93.SE6165.G6165				927
H1H.CF.90.056				926
H1J.SE.93.SE9280.7887				925
H1K.CM.96.96CM.MP535				924
H1O.RE.87.ANT70				923
H10.CM.91.MVP5180				922
H10.CM.94.BCF06				921
H10.CM.96.CMA102				920
H10.CM.98.98CMA104				919
H10.CM.99.99CMU4122				918
H10.CM.x.DCM02.3				917
H10.FR.92.VAU				916
H10.SN.99.99SE.MP1299				915
H10.US.99.99USTWLA				914
H10.US.x.1.24788				913
H1N.CM.02.DJ00131				912
H1N.CM.95.YBF30				911
H1N.CM.97.YBF106				910
H1P.CM.06.U14788				909
H1P.FR.06.RBF168				908
CPZ.CD.90.ANT7				907
CPZ.CD.06.BF1167				906
CPZ.CM.05.SIVcpzDP943				905
CPZ.CM.05.SIVcpzEK505				904
CPZ.CM.05.SIVcpzMB897				903
CPZ.CM.05.SIVcpzMT145				902
CPZ.CM.98.CAM5				901
CPZ.GA.88.GAB1				900
CPZ.GA.88.SIVcpzGAB2				999
CPZ.TZ.00.TAN1				998
CPZ.TZ.01.TAN2				997
CPZ.TZ.09.UG38				996
CPZ.TZ.06.SIVcpzTAN13				995
CPZ.US.85.US.Marilyn				994
MAC.US.x.239				943
SMM.CI.79.SIVsmCI2				942
SMM.LR.89.SIVsmLR1B				941
SMM.SL.92.SIVsmSL92A				940
STM.US.89.STM.37.16				939
H2A.DE.x.BEM				938
H2A.PT.x.ALI				937
H2A.SN.x.ST.HIV.2.ST				936
H2B.CI.x.EH0				935
H2B.HG.86.D205.ALT				934
H2G.CI.92.Abt9E				933
H2U.FR.96.12034				932
GOR.CM.04.SIVgorCP684con				931
GOR.CM.07.SIVgorCP2135con				930
VER.DE.x.AGM3				929
VER.KE.x.9063				928
VER.KE.x.AGM155				927
VER.KE.x.TY01.patent				926
SAB.SN.x.SAB1				925
GRV.ET.x.GRI.677.gri.1				924
TAN.UG.x.SIVagMTAN1				923
DRL.x.x.FAO.7				922
MND.2.GA.x.M14				921
MND.2.CM.98.CM16				920
MND.2.x.x.5440				919
MNE.US.x.MNE027				918
RCM.CM.00.SIVagi.00CM312				917
RCM.CM.02.SIVrcm.02CM8081				916
RCM.GA.x.SIVRCMGAB1				915
RCM.NG.x.NG411				914
SVK.KE.x.KE51				913
SVK.KE.x.SYK173.COMGNM				912
LST.CD.88.SIVlhoest447				911
SUN.GA.98.L14				910
MND.1.GA.x.MNDGB1				909
LST.CD.88.SIVlhoest485				908
LST.CD.88.SIVlhoest524				907
LST.KE.x.01				906
DEB.CM.04.SIVdeb04CMPF3061				905
DEB.CM.99.CM40				904
DEB.CM.99.CM5				903
DEN.CD.x.CD1.CM0580407				902
GSN.CM.99.CM166				901
GSN.CM.99.CM71				900
MON.CM.99.L1.99CML1				999
MON.NG.x.NG1				998
MUS.1.CM.01.CM1239				997
MUS.1.CM.01.SIVmus.01CM1085				996
MUS.2.CM.01.CM1246				995
MUS.2.CM.01.CM2500				994
TAL.CM.00.266				993
TAL.CM.01.8023				992
OLC.CI.97.97C112				991
WRC.CI.97.97C114				990
WRC.CI.98.98C104				989
WRC.GM.05.Pbt.05GM.X02				988
COL.CM.x.CGU1				987

H1B.FR.83.HXB2
H1A1.UG.85.U455.U455A
H1C.ET.86.ETH2220
H1D.CD.84.84ZR085
H1F1.BE.93.VI850
H1G.SE.93.SE6165_G6165
H1H.CF.90.056
H1J.SE.93.SE9280_7887
H1K.CM.96.96CM_MP535
H1O.BE.87.ANT70
H1O.CM.91.MVP5180
H1O.CM.94.BCF06
H1O.CM.96.96CMA102
H1O.CM.98.98CMA104
H1O.CM.99.99CMU4122
H1O.CM.x.x.CM12_3
H1O.FR.92.VAU
H1O.SN.99.99SE_MP1299
H1O.US.99.99USTWLA
H1O.US.x.I.24788
H1N.CM.02.DJ00131
H1N.CM.95.YBF30
H1N.CM.97.YBF106
H1P.CM.06.U14788
H1P.FR.06.RBF168
CPZ.CD.90.ANT
CPZ.CD.06.BF1167
CPZ.CM.05.SIVcpzDP943
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzMB897
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CAM3
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.SIVcpzGAB2
CPZ.TZ.00.TAN1
CPZ.TZ.01.TAN2
CPZ.TZ.09.UG38
CPZ.TZ.06.SIVcpzTAN13
CPZ.US.85.US_Marilyn

TKIQNFRVYRSDRNPLWKGPAKLLWKGEGAVVIQDNS.DIKVPPRRKAKIIRDYDGKQMGDDCVASRODED...
S-----D-I-----M-G-----*
L-----D-I-----A-G-----
L-----D-I-----E-I-----V-G-----
SN-K-----D-V-----N-E-----E-G-----
LN-K-----D-I-----P-----E-----G-----
L-XHK-----D-I-----Q-----KG-----E-----T-SM-G-T-SESVVEOPSEIP
L-N-----D-I-----Q-----KG-----N-----V-T-S-----T-SESMVEOPGETP
L-----D-I-----Q-----KG-----X-----H-----A-SX-G-X-SESXE-PGETP
L-----D-I-----Q-----KG-----H-----H-----A-SM-G-T-SESVVEOPGETS
F-K-Q-----D-I-----Q-----T-SM-----N-----T-SM-G-X-SESVVEOPGETP
F-K-Q-----D-I-----Q-----KG-----H-----A-S-----G-T-GEVVEOPSEIP
L-----D-I-----Q-----KG-----H-----T-SM-G-T-SESVVEOPSEIP
L-----D-I-----Q-----KG-E-----T-M-----G-T-SENMEQXSEIS
L-----D-I-----Q-----G-----T-SM-----G-A-NODMA-----T-SESVVEOPGETP
L-V-----D-I-----G-----G-----G-----NOEME
L-V-X-----D-I-----G-----N-----GX-NODME
L-----D-I-----T-----KG-----DT-R-SESLE05G
L-----D-I-----T-----KG-----KE-R-KIEDR-DL-G-N-----DT-R-SESLE
F-Q-H-----D-V-----Q-----K-Q-F-E-----KE-R-KIEDR-DL-G-N-----DT-R-SESLE
F-L-----D-V-----R-----KEGE-E-----K-----V-AGG-D-N-----DT-R-SESLE
L-V-----D-I-----R-----K-KE-EV-----GSM-G-SONLE
L-V-----E-T-----R-----G-----G-----SODME
S-V-K-----E-I-----L-KE-E-----VE-N-L-D-S
L-V-Q-----K-D-----I-----R-----KERE-EV-I-----A-S-----SOGLE
L-V-----D-I-----R-----KE-E-EV-----SM-GG-----SOGLE
L-V-K-----D-I-----T-----QG-EL-----GSM-G-N-QNLE
L-K-----D-I-----S-----KH-V-----L-GG-----NONME
L-V-----A-D-I-----R-----KEGE-----KE-----AGGMD-D-N-T
L-W-----A-D-I-----R-----KEGE-----K-----AGGMD-N-T
L-----D-V-----R-----KEGE-----V-----K-----SAGSMD-D-N-AONME
L-V-----A-D-I-----R-----KEGE-----KE-----AGS-DD-N-A
L-V-----D-I-----KE-E-EV-----K-----A-SM-G-S

p31 Integrase end
Pol end

1003
1003
1011
1011
1004
1005
1001
1003
1002
1004
1010
1010
1010
1010
1010
1010
1010
1008
1011
1010
1010
1013
1010
1013
1015
1015
1008
1004
1004
1004
1007
1015
1014
1012
1003
1013
1009
1002
1011
1000
1001
999
999
1006

p31 Integrase end
Pol end

MAC.US.x.239
SMM.CI.79.SIVsmCI2
SMM.LR.89.SIVsmLIB1
SMM.SL.92.SIVsmSL92A
STM.US.99.STM_37_16
H2A.DE.x.BEM
H2A.PT.x.ALI
H2A.SN.x.ST_HIV_2_ST
H2B.CI.x.EH0
H2B.GH.86.D205_ALT
H2G.CI.92.Abt95
H2U.FR.96.12034
GOR.CM.04.SIVgorCP684con
GOR.CM.07.SIVgorCP2135con
VER.DE.x.AGM3
VER.KE.x.9063
VER.KZ.x.AGM155
VER.KE.x.TY01_patent
SAB.SN.x.SAB1
GRV.ET.x.GRI_677_gri_1
TAN.UG.x.SIVagmTAN1
DRL.x.x.FAO7
MND-2.GA.x.M14
MND-2.CM.98.CM16
MND-2.x.x.5440
MNE.US.x.MNE027
RCM.CM.00.SIVagI_00CM312
RCM.CM.02.SIVrcm_02CM8081
RCM.GA.x.SIVRCMGAB1
RCM.NG.x.NG411
SVK.KE.x.KE51
SVK.KE.x.SVK173_COMGNM
LST.CD.88.SIVlhoest447
SUN.GA.98.L14
MND-1.GA.x.MNDGB1
LST.CD.88.SIVlhoest485
LST.CD.88.SIVlhoest524
LST.KE.x.Lho7
DEB.CM.04.SIVdeb04CMPF3061
DEB.CM.99.CM40
DEB.CM.99.CM5
DEN.CD.x.CD1_CM0580407
GSM.CM.99.CM166
GSM.CM.99.CM71
MON.CM.99.L1_99CML1
MON.NG.x.NG1
MUS-1.CM.01.CM1239
MUS-1.CM.01.SIVmus_01CM1085
MUS-2.CM.01.CM1246
MUS-2.CM.01.CM2500
TAL.CM.00.266
TAL.CM.01.8023
OLC.CI.97.97CI12
OLC.CI.97.97CI14
WRC.CI.98.98CI04
WRC.GM.05.Pbt_05GM_X02
COL.CM.x.CGU1

S-FK-----EG-DQ-----GE-----ILKVGIT-----K-----GGKEV-SSSHMEDTGEAREVA
S-FK-----EG-DQ-----N-----I-KVGT-----I-----K-----GGKELGNSPYLENP-EDGKMAQPD
S-FK-----EG-DQ-----GE-----ILKVGIT-----K-----GGKEV-SSSHLEDTREIREMA
S-FKK-----EG-DQ-----GE-----I-KVG-----I-----K-----GGKEM-SSADMENTTQVRETAQLSEISKV
S-FK-----EG-DQ-----GE-----IVKVGIT-----K-----GGKEV-SSSHLEDTREAREVA
S-FK-----EG-DQ-----GE-----IVKVGIT-----K-----GROEL-SGPHLEGAREDEGEMACPQVPEIQNKRRPGGALCSPPOGGMG
L-FK-P-----EG-DQ-----GE-D-----IVKVGIT-----K-----GROEL-SGPHLEGAREDEGEMACPQVPEIQNKRRPGGALCSPPOGGMG
S-L-----F-EG-DQ-----GE-D-----IVKVGIT-----II-----K-----GROEM-SGSNLEGAREDEGEMACPQVPEIQNKRRPGGALCSPPOGGMG
L-F-----EG-DQ-----GD-----I-KVGT-E-----I-----N-----GGKEL-CSADVEDTMOAREVAQSN
F-F-Q-----EG-DQ-----GE-----I-KVGT-E-----H-----GGKEL-CSADVEDTMOAREVAQSN
S-FK-----EG-DQ-----D-----L-KVGT-----I-----K-----GGKEV-SSSHMEDTGEAREVA
L-FK-----EG-DQ-----D-----L-KVGT-----I-----K-----GGKEL-SSSLEQSS
L-----D-I-----T-----KG-----DT-G-SESLEQSS
L-----D-I-----T-----KG-----DT-G-SESLEQSS
O-L-----EG-D-V-----GO-I-----KGGV-EL-EY-----K-----EPRKRMG-ESNLEGAGGADN
O-L-----EG-D-V-----O-I-----KEGE-EL-I-----K-----EPRKRMG-EGNMEGLRGPDMQARN5Q-ILDD
VER.KE.x.AGM155
O-L-----EG-D-V-----O-I-----LK-DGS-L-----K-----EPKRVGNEGDVEGTRGSDN
S-----EG-D-V-----I-----EQG-EL-TI-----K-----L-SOAPLENGRTAGEVD
O-L-K-----EG-D-V-----R-I-----KEGE-----K-----ERKTM-SEGSMGVREANKQMEGSDLODQE
O-S-K-----EG-E-V-----T-----KEGE-EL-----K-----ERKTVG-KTNMEG
L-LK-K-----EG-DQ-----E-V-----KEGT-L-FS-----K-----NVDOSHIME
S-----EG-DQ-----E-I-----KEGT-L-----K-----TVDSNPHMED-ETA
O-----EG-DQ-----E-I-----KEGT-L-----K-----TVDS-PN-EA
S-----Q-----EG-DQ-----E-I-----KEGV-L-----K-----TVDSNPHMGR
S-FK-----EG-DQ-----GE-----ILKVGIT-----K-----GGKEV-SSSHMEDTGEAREVA
S-F-K-----EG-DQ-----T-I-----GO-L-T-----C-VK-----RKNVDSEINMEGGO-EN
S-----F-EG-DQ-----T-I-----GO-L-T-----C-K-----RKVDSEINMEGRO-KD
S-----EG-DQ-----R-----I-E-----EETG-L-----KE-----RKVDSEANLAGRO-EN
S-----F-EG-DQ-----I-E-----OE-L-----C-K-----RKEVDRETNMEGRO-ES
O-FS-K-----HGT-SD-Q-S-----TPDQOQVIA-----SSD-ERVDSGTHLEITSKSN
P-FS-----OGK-NF-R-----V-TEEG-FA-----T-H-ER-DSGSH-ENDPKT
S-F-K-----EG-D-Q-----V-----VKEGD-T-FA-----LVK-----EGPKDSEGLHNN
S-LK-K-----EG-D-Q-----K-----KOGE-N-L-----LVK-----GESSVEM-G
S-NL-K-----F-EG-DQ-----GI-----LKYQE-E-I-----C-K-----E-SGKNSQ-NLESV
S-S-LK-K-----OG-D-Q-----OI-----VKEGD-T-FS-----LVK-----EGPKDSEGLSDNN
S-LK-K-----OG-D-Q-R-----O-----VKEGE-N-FSI-----LVK-----EGPKDSEGLSDNN
S-FK-----EG-D-Q-----T-----I-----K-----KTKEGE-VIT
S-FSG-K-----EGAD-S-----T-----KLDTG-L-----V-K-----DVGSKIDDPDTHS
S-LLG-K-----EGAD-T-----T-----CKTEVG-----V-KE-----D-DSKVNTQNSNE
S-LLGYK-----EGAD-T-----T-----CKTEVG-----V-K-----GVDSKTDTQNSPE
P-FS-K-----EGA-Q-Q-E-----KTOAG-PP-----KP-----NAE-SRKGQNKGM-SETD1
S-F-K-----OGA-Q-Q-V-----V-TOAGE-IT-----KP-----EAKTE-VGG-THHN
S-F-K-----OGA-H-Q-V-----V-TOAGE-IT-----KP-----EAK-E-VGS-AHTSN-REEGRMAD
F-FSK-----F-TGAD-Q-R-----KTEEG-LT-----KP-----TTENVGGDINOYLRKQDGLAN
F-FSK-----F-TGAD-S-H-----L-KTDOGEVIT-----P-----SKEDVGSKPSAH-TREVGMAD
S-F-----TGSD-S-----A-----VKTEGQVIT-----KP-----AKENVGSKSNTGDHRRKEDGLDN
Y-F-----TGSD-S-----V-----R-----L-KTDOGEVIT-----KP-----AKEDVGSKSDTGLRKEDGLDN
F-FSK-----RGA-Y-LR-R-----L-KTDOGEVIT-----K-----TRQNV-SEPD-VHVRKEDGLAD
F-FSK-----RGA-S-O-Q-----I-----KTDOGE-IT-----K-----TKENV-SDPNPVY-RKEDGLAD
QNFSD-K-----KGA-Q-Q-I-----LRTDEGEVIT-----KP-----QALGNKIDLE-SKEQ-AEMGRDN
HNFSD-K-----KGA-Q-Q-H-V-----LRTDEGEVIT-----KP-----QA-GNKTDLEGSKEQ-AEMGRDN
L-FK-SGGRV-YKNKEGO-C-E-----LLKEGD-KYFS-L-----VKP-----VDSANV
L-FK-Q-----TGA-Q-Q-G-P-----L-ETPE-G-IT-----L-KVWNGEGDRSSSDKN
L-FK-Q-----TGA-Q-Q-G-----L-VETPE-G-IT-----K-----KVWNGEGDRSSSDKN
F-FK-Q-----TGA-Q-Q-G-V-----L-V-TPEG-LLT-----K-----KTWDGKGMDSSTTN
O-FK-KA-WKEH.TGE-Q-GE-V-----RNSQTLF-K-----V-T-LQYGEDV-SENLL-NGQKEAETVKGM

1020
1028
1010
1010
1023
1016
1016
1059
1033
1033
1036
1035
1036
1055
1003
1003
1023
1034
1032
1026
1040
1041
1020
1013
1017
1011
1012
1016
1018
1024
1019
1023
1041
1037
1020
1017
1010
1020
1021
1020
1020
1017
1011
1008
1008
1032
1032
1017
1018
1018
1021
1012
1012
1017
1014
1020
1021
995
1023
1022
1013
991

Table with columns for accession numbers (e.g., H1B, H1A1, H1C, etc.), amino acid sequences, and identifiers. It includes a 'Vif start' section and a 'Vif end' section.

Table with columns for accession numbers (e.g., MAC.US.x.239, H2A.DE.x.BEN, etc.), amino acid sequences, and identifiers. It includes a 'Vif start' section and a 'Vif end' section.

	Vpx start	Vpx end
MAC.US.x.239	MSD..PRERIPPGNSGEETIG.EAF..EWLNRVTEEINREAVNHLPRELIFQVWORSWEYWHDEQGMSPSYVKYRYLCLIQKALFMHCKKGCRCLEGEHGAGGW.RPGPPPPPPGLA*	112
H2A.DE.x.BEN	-T.....V.....E..I..AL.....R.....A..T.....M..I..T..F..R..T..W..DM..RE..LEDO.....V.....	114
H2A.PT.x.ALI	-AN...TV.....D...AL.....R.....Q..T.....M..MYT..FM..T..G..P.....S.....V.....	113
H2A.SN.x.STI_HIV_2_ST	-AG...T.....D...AL.....R.....I..T.....M..M..I..S..R..T..G..P.....S.....V.....	113
H2B.CI.x.EH0V.....D...V.....A..E..T..L..L..V.....K..A..RE..I..T.....M..M..I..FA..G..R..P.....S.....V.....	112
H2B.GH.86.D295.ALTV.....D...V.....A..E..IT..L..V.....A..RE..I..T.....L..M..M..V..YT.....Q..P.....S.....	112
H2G.CI.92.Abt96X..E..D.....E..XT..I.....R.....A..T.....L..M..M..V.....T..QK..P.....Q.....	112
H2U.FR.96.12034	-G.....E..V.....S..E..RD..A..Q.....Q..R.....E..T.....Q.....P.....S.....I.....	112
DRL.x.x.FA0	-AERQSV..A..AEPMGAGEV..LE..Q..SLLR..Q..RL..FHP..FL..RL..NTCM..HY..ALQL..FT..S.....L..L..M.....FQO..S..OGR..PPPLRPAGDRL.....PP..	113
MAC.US.x.751.1A11E.....M.....	114
MAC.US.x.751.BK28E.....M.....	113
MND.2.CM.98.CM16	-AE..GAPE..E..A..V.....DLN..T..E..SL..K..Q..RL..FHP..FL..RL..NACI..H..RHQR..L..A.....L..MN..M..T..MQOE..P..RSG..P.....MV.....	100
MND.2.GA.x.M14	-AE..RAPEA..Q..A..V.....GLE..Q..E..SL..QV..QL..FHP..FL..RL..NTCV..H..RLRRTL..A.....L..M..M..V..MQO..P..RSG..S.....	100
MND.2.x.x.5440	-AE..RAPEA..E..A..V.....GLE..Q..ETS..L..R.....RL..FHP..FL..RL..NTCV..H..RHQR..LD..A.....L..MH..MYT..MQO..P..RNG..RPR.....M.....	100
MNE.US.x.MNE027E.....M.....	113
RCM.CM.00.SIVagi_00CM312	-AE.....A..E..PT..A.....EVE..QPWLARMLY..INO..RL..F..G..F..HL..RTCV.....RTL..E..AG..L..M.....T..MRS..KLR.....DPPR..OR..ERV..IL..MO.....	106
RCM.CM.02.SIVrcm_02CM8081	-AE.....A..EVPT..A.....EVE..QPWLARMLY..INO..RL..FHP..F..RL..RTCV..H.....L..R..LE..AG..L..M.....T..MRS..LR.....DPPR..OR..ERV..IL..MO.....	106
RCM.GA.x.SIVRCMGAB1	-AE.....A..EVPT..A.....EAE..QPWLDRMLEKVN..L..RL..FHP..F..RL..RTCV..H.....VHQR..LE..AA..L..M.....I..QT..SOR.....PNPR..AV..ERITIL..M.....	105
RCM.NG.x.NG411	-AE..G.....V..EAPT..A.....DVE..APWLHRMLT..VNL..RL..FHP..F..RL..RTCV..H.....RL..R..LE..AG..L..M.....I..QS..SOR.....Q..QAREA..ERITIL..M.....	109
SMM.CI.79.SIVsmCI2	-GS..A.....T.....E.....YS..AV..G..K.....R.....F.....QA..TQ.....L..M.....C..V.....T..R.....S.....A.....	112
SMM.LR.89.SIVsmLIB1E.....Q.....R.....F.....A..T.....M..L.....S.....S.....	112
SMM.SL.92.SIVsmSL92AQ.....D.....D..E.....V..T.....R.....AA..T.....P.....F.....	111
SMM.SL.92.SIVsmSL92B	-T.....HN.....AL..QT..Q.....S.....R..C.....V.....Y.....A.....VQ..M..M..Q..FR..T..R.....SQ.....T.....	113
SMM.US.04.G078	-T.....HN.....AL..QT..Q.....R..C.....V.....Y.....A.....VQ..M..M..Q..FR..T..R.....SQ.....T.....	113
SMM.US.04.G932H..A.....S.....	113
SMM.US.04.M919D..D..A.....R.....KI..V..T.....M.....R.....E..E.....S.....	113
SMM.US.04.M922L..D..D.....R.....M..E..T.....L..M.....R.....E.....S.....	113
SMM.US.04.M923E.....A.....R.....I..V..T.....GEY.....S.....	113
SMM.US.04.M926	-A.....D.....T.....R.....V.....R.....YE..Q..Q.....S.....	113
SMM.US.04.M934Q..A.....R.....YE..Q..Q.....S.....	113
SMM.US.04.M935V.....D..D.....A.....R.....M..E..T.....M.....R.....E.....S.....	113
SMM.US.04.M940Q.....A.....R.....VA..T.....W..M.....R.....	113
SMM.US.04.M946Q..A.....R.....YE..PR..Q.....R.....	113
SMM.US.04.M947V.....D..E..D..A.....R.....I..V..T.....GEY.....S.....	113
SMM.US.04.M949V.....D..E..M.....R.....E..T.....W.....R.....	113
SMM.US.04.M950Q..D.....R.....YE..Q..Q.....W..M.....	113
SMM.US.04.M951V.....H..D.....R.....YE..R..E..T.....W..M.....	113
SMM.US.04.M952AA..VL..H..V.....R.....YE..Q..E.....I.....	113
SMM.US.05.D215	-N.....E.....D..V.....R.....T.....YI.....D.....	113
SMM.US.06.FTQE.....D..A.....R.....R..A..T.....V.....R.....T.....	113
SMM.US.11.SIVsmE660_FL10E.....G..A.....R.....M..E..T.....V.....T.....	113
SMM.US.11.SIVsmE660_FL14E.....G..A.....R.....M..E..T.....V.....E.....T.....	113
SMM.US.11.SIVsmE660_FL6E.....G..A.....R.....M..E..T.....V.....T.....	113
SMM.US.11.SIVsmE660_FL8E.....G..A.....R.....M..E..T.....V.....T.....	113
SMM.US.86.CFU212	-A.....X.....XX..X..D..H..X.....A.....X.....R..X.....M..V..T..X.....M..M..R..N..R..E..Q.....	113
SMM.US.x.H9	-XX..X.....XX..X..D..H..X.....A.....X.....R..X.....M..V..T..X.....M..M..R..N..R..E..Q.....	113
SMM.US.x.PG53E.....G..A.....R.....M..E..T.....V.....R.....E.....S.....	113
SMM.US.x.SIVsmH635F_L3E.....G..A.....R.....M..E..T.....V.....T.....	113
SMM.x.x.pE660.CG7GE.....G..A.....R.....M..E..T.....V.....T.....	113
SMM.x.x.pE660.CG7VE.....G..A.....R.....M..E..T.....V.....T.....	113
STM.US.89.STM_37_16A..E.....H..D.....R.....PG.....S..R.....	113

Vpr start	oligomerization	amphiphatic α-helix	H(S/N)RIG motifs	Vpr end in HXB2	Vpr end
H1B.FR.83.HXB2					
H1A1.UG.85.U455.U455A					
H1C.ET.86.ETH2220					
H1D.CO.84.84Z8685					
H1F1.93.93.V1850					
H1G.SE.93.SE6165.G6165					
H1H.CF.90.056					
H1J.SE.93.SE9280.7887					
H1K.CM.96.96CM.MP535					
H1O.BE.87.ANT70					
H10.CM.91.MVP5180					
H10.CM.94.BCF06					
H10.CM.96.96CMA102					
H10.CM.98.98CMA104					
H10.CM.99.99CMU4122					
H10.CM.x.CM02.3					
H10.FR.92.VAU					
H10.SN.99.99SE.MP1299					
H10.US.99.99USTWLA					
H10.US.x.I.24788					
H1N.CM.02.DJ00131					
H1N.CM.95.YBF30					
H1N.CM.97.YBF106					
H1P.CM.06.U14788					
H1P.FR.06.RBF168					
CPZ.CO.90.ANT					
CPZ.CO.06.BF1167					
CPZ.CO.05.SIVcpzDP943					
CPZ.CO.05.SIVcpzEK505					
CPZ.CO.05.SIVcpzMB897					
CPZ.CO.05.SIVcpzMT145					
CPZ.CO.98.CAM5					
CPZ.CO.98.CAM5					
CPZ.GA.88.GAB1					
CPZ.GA.88.SIVcpzGAB2					
CPZ.TZ.00.TAN1					
CPZ.TZ.01.TAN2					
CPZ.TZ.09.UG38					
CPZ.TZ.06.SIVcpzTAN13					
CPZ.US.85.US.Marilyn					
GOR.CO.04.SIVgorCP684con					
GOR.CO.07.SIVgorCP2135con					

Vpr start	H(S/N)RIG motifs	Vpr end
MAC.US.x.239		
SM.MI.79.SIVsmCI2		
SM.MI.89.SIVsmL1B1		
SM.MI.92.SIVsmL92A		
STM.US.89.STM.37.16		
H2A.DE.x.BEN		
H2A.PT.x.ALI		
H2A.SN.x.ST.HIV.2.ST		
H2B.VI.01.HIV.1		
H2B.GI.86.D205.ALT		
H2G.CI.92.Abt96		
H2U.FR.96.12034		
VER.DE.x.AGM3		
VER.KE.x.9063		
VER.KE.x.AG055		
VER.KE.x.TY01.patent		
SAB.SN.x.SAB1		
GRV.ET.x.GRI.677.gri.1		
TAN.UG.x.SIVagmTAN1		
DRL.x.x.FAO		
MND.2.CM.98.CM16		
MND.2.GA.x.M14		
MND.2.x.x.5440		
MNE.US.x.MNE027		
RCM.CO.00.SIVcm.00CM312		
RCM.CO.02.SIVcm.02CM8081		
RCM.CO.x.SIVRCM8AB1		
RCM.NG.x.NG411		
SVK.KE.x.KE51		
SVK.KE.x.SYK173.CMGNM		
LST.CO.88.SIVLhoest447		
LST.CO.88.SIVLhoest485		
LST.CO.88.SIVLhoest524		
LST.KE.x.lho7		
SUN.GA.98.L		
MND.2.GA.x.MNDB1		
DEB.CO.04.SIVdeb04CMF3061		
DEB.CO.99.CM40		
DEB.CO.99.CM40		
DEN.CO.x.C01.CM0580407		
GSN.CO.99.CM166		
GSN.CO.99.CM171		
MON.CO.99.L1.99CM11		
MON.NG.x.NG1		
MUS.1.CM.01.CM1239		
MUS.1.CM.01.SIVmus.01CM1085		
MUS.2.CM.01.CM1246		
MUS.2.CM.01.CM2500		
TAL.CO.00.266		
TAL.CO.01.8023		
OLC.CI.97.97CI12		
WRC.CI.97.97CI14		
WRC.CI.98.CI04		
WRC.CO.05.Pbt.05GM.X02		
COL.CO.x.CGU1		

Table with columns: Tat start, C-rich region, nuclear localization, exon 1 end, exon 2 start. Rows include sequences for various HIV strains like H1B, H1A1, H1C, etc.

Table with columns: Tat start, exon 1 end, exon 2 start. Rows include sequences for various HIV strains like MAC, H2A, H2B, etc.

		Tat end	
H1B. FR. 83. HXB2	D.....*		100
H1A1.UG.85.U455.U455A	A.....*		102
H1C.ET.86.ETH2220	A.....*		101
H1D.CD.84.84ZR085	A.....*		87
H1F1.BE.93.VI850	A.....*		102
H1G.SE.93.SE6165_G6165*		102
H1H.CF.90.056*		100
H1J.SE.93.SE9280_7887*		102
H1K.CM.96.96CM_MP535*		102
H1O.BE.87.ANT70	PRRKGSCHCCTRTSE0		115
H1O.CM.91.MVP5180	PCHQDSCNSCTRISG0		116
H1O.CM.94.BCF06*		93
H1O.CM.96.96CMA102*		79
H1O.CM.98.98CMA104*		79
H1O.CM.99.99CMU4122*		97
H1O.CM.x.CM02_3*		98
H1O.FR.92.VAU*		80
H1O.SN.99.99SE_MP1299*		98
H1O.US.99.99USTWLA*		80
H1O.US.x.I_24788*		98
H1N.CM.02.DJ00131*		102
H1N.CM.95.YBF30*		102
H1N.CM.97.YBF106*		103
H1P.CM.06.U14788*		72
H1P.FR.06.RBF168*		72
CPZ.CD.06.BF1167	-IQ.....*		101
CPZ.CD.98.ANT*		74
CPZ.CM.05.SIVcpzDP943	A.....*		103
CPZ.CM.05.SIVcpzEK505*		94
CPZ.CM.05.SIVcpzMB897	A.....*		101
CPZ.CM.05.SIVcpzMT145	-W.....*		102
CPZ.CM.98.CAM3	A.....*		103
CPZ.CM.98.CAM5	-WSAGS.....*		108
CPZ.GA.88.GAB1*		101
CPZ.GA.88.SIVcpzGAB2*		100
CPZ.TZ.00.TAN1	PROODSVSSGRTSGTSSSGYTRPKTSSGSGSACKH		137
CPZ.TZ.01.TAN2	PROODSVSSGRTSGTSSSGTRPVETSSGSGRSCKH		137
CPZ.TZ.06.SIVcpzTAN13	PRPEDSSASSGRTSGTSSSGSTRPVSTSSGCGPYSPK		137
CPZ.TZ.09.UG38	PCTEDSSLSSGRTSGPSSCGSTRPVKTSAGSSGRS		134
CPZ.US.85.US_Marilyn*		103
		Tat end	
MAC.US.x.239	LGR.....*		131
H2A.DE.x.BEN	LGO.....*		131
H2A.PT.x.ALI	PGRSHIYIS.....*		137
H2A.SN.x.ST_HIV_2_ST	PGR.....*		131
H2B.CI.x.EH0	PGRSNTSTSRFAN.....*		139
H2B.GH.86.D205_ALT	PGR.....*		129
H2G.CI.92.Abt96*		113
H2U.FR.96.12034*		112
COL.CM.x.CG01*		80
DEB.CM.04.SIVdeb04CMPF3061*		97
DEB.CM.99.CM40	VSSGRISNNC.....*		111
DEB.CM.99.CM5*		97
DEN.CD.x.CD1_CM0580407*		99
DRL.x.x.FA0	PPS0DPPKLGAKKQ.....*		135
GOR.CM.04.SIVgorCP684con*		83
GOR.CM.07.SIVgorCP2135con*		83
GRV.ET.x.GRI_677_gri_1*		95
GSN.CM.99.CN166	GP.....*		113
GSN.CM.99.CN71	VSCGTT.....*		120
LST.CD.88.SIVlhoest447*		101
LST.CD.88.SIVlhoest485*		101
LST.CD.88.SIVlhoest524*		112
LST.KE.x.lho7*		112
MND-1.GA.x.MNDGB1*		116
MND-2.CM.98.CM16	APK.....*		123
MND-2.GA.x.M14*		88
MND-2.x.x.5440*		111
MNE.US.x.MNE027	LGR.....*		132
MON.CM.99.L1_99CML1	HI.....*		112
MON.NG.x.NG1	SMREDF.....*		120
MUS-1.CM.01.CM1239	-I.....*		110
MUS-1.CM.01.SIVmus_01CM1085	RCFKTSTT.....*		107
MUS-2.CM.01.CM1246*		117
MUS-2.CM.01.CM2500	AI.....*		110
OLC.CI.97.97CI12*		108
RCM.CM.00.SIVagi_00CM312	-LSIS.....*		116
RCM.CM.02.SIVrcm_02CM8081	-ISTP.....*		113
RCM.GA.x.SIVRCMGAB1*		106
RCM.NG.x.NG411	-IFIS.....*		116
SAB.SN.x.SAB1*		94
SMM.CI.79.SIVsmCI2	LGK.....*		133
SMM.LR.89.SIVsmLIB1*		114
SMM.SL.92.SIVsmSL92A*		103
STM.US.89.STM_37_16	LGK.....*		132
SUN.GA.98.L14*		103
SYK.KE.x.KE51*		99
SYK.KE.x.SYK173_COMGNM	NL.....*		111
TAL.CM.00.266	PLSGRTSTASSATVTRP.....*		111
TAL.CM.01.8023	PO.....*		128
TAN.UG.x.SIVaqmTAN1	GRKNLA00SGRATGASD.....*		102
VER.DE.x.AGM3	GRKDLERDKREAVGAN.....*		120
VER.KE.x.9063	SRKNLGDETRGPVGAGN.....*		120
VER.KE.x.AGM155*		101
VER.KE.x.TY01_patent*		101
WRC.CI.97.97CI14*		125
WRC.CI.98.98CI04*		132
WRC.GM.05.Pbt_05GM_X02*		132

H1B.FR.83.HXB2
H1A1.UG.85.U455 U455A
H1C.ET.86.ETH2220
H1D.CO.84.84Z8685
H1E.BE.93.VT850
H1G.SE.93.SE6165 G6165
H1H.CF.90.056
H1J.SE.93.SE9280 7887
H1K.CM.96.96CM MP353
H1L.RE.87.AUT78
H1O.CM.91.MVPS180
H10.CM.94.BCF06
H10.CM.96.96CMA102
H10.CM.98.98CMA104
H10.CM.99.99CMU4122
H10.CM.x.DCM02_3
H10.FR.92.VAU
H10.SN.99.99SE MP1299
H10.US.99.99USTWLA
H10.US.x.1.24788
H1N.CM.02.DJ00131
H1N.CM.95.YBF30
H1N.CM.97.YBF106
H1P.CM.06.U14788
H1P.FR.06.RBF168
H2B.CM.06.BT167
CPZ.CO.98.NIT
CPZ.CM.05.SIVcpzDP943
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzMB897
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CAM5
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.SIVcpzGAB2
CPZ.TZ.00.TAN1
CPZ.TZ.01.TAN2
CPZ.TZ.06.SIVcpzTAN13
CPZ.TZ.09.UG38
CPZ.US.85.US_Marilyn

MAC.US.x.239
H2A.DE.x.BEN
H2A.PT.x.ALI
H2A.SN.x.ST_HIV_2_ST
H2B.CI.x.EH_0205 ALT
H2C.DI.06.D205 ALT
H2G.CI.92.Abt96
H2U.FR.96.12034
COL.CM.x.CGUI
DEB.CM.04.SIVdeb04CMP3061
DEB.CM.99.CM40
DEB.CM.99.CM5
DEN.CO.x.CD1 CM0580407
DRL.x.x.FAO
GOR.CM.04.SIVgorCP684con
GOR.CM.07.SIVgorCP2135con
GRV.TX.x.GR677_gri_1
GSN.CM.99.CN166
GSN.CM.99.CN71
LST.CD.88.SIVlhoest447
LST.CD.88.SIVlhoest485
LST.CD.88.SIVlhoest524
LST.KE.x.lho7
MND.1.GA.x.MNDG81
MND.2.GA.x.CM16
MND.2.GA.x.M14
MND.2.x.x.5449
MNE.US.x.MNE027
MON.CM.99.L1 99CML1
MON.NG.x.NG1
MUS.1.CM.01.CM1239
MUS.1.CM.01.SIVmus_01CM1085
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
OLC.CI.97.97C112
RCM.CM.00.SIVag1_00CM312
RCM.CM.02.SIVrcm_02CM8081
RCM.GA.x.SIVCMGAB1
RCM.NG.x.NG411
SAB.SN.x.SAB1
SMM.CI.79.SIVsMC12
SMM.LR.89.SIVsmlB1
SMM.SI.92.SIVsmlS92A
STM.US.89.SIM_37_16
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173_COMGNM
TAL.CM.00.266
TAL.CM.01.8023
TAN.UG.x.SIVaomTAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TYR1patent
WRC.CI.97.97C114
WRC.CI.98.98C104
WRC.GM.05.05GM_X02

Table with columns: Rev start, exon 1 end, exon 2 start, NLS, Leu-rich effector domain. Contains amino acid sequences for various HIV proteins and domains.

Table with columns: Rev start, exon 1 end, exon 2 start. Contains amino acid sequences for various HIV proteins and domains.

Table with 2 columns: Env start, signal peptide end_gp120 start. Contains protein entries such as H1B, H1A1, H1C, etc., and their corresponding amino acid sequences.

Table with 2 columns: Env start, signal peptide end_gp120 start. Contains protein entries such as MAC, H2A, H2B, etc., and their corresponding amino acid sequences.

glycosylation NIS
glycosylation NCS

CD4 binding

glycosylation NVS

Table with columns V1, glycosylation NDT, glycosylation NTS, glycosylation NGT, glycosylation NKT. Rows include protein names like H1B, H1A1, H1A, etc., and their corresponding amino acid sequences.

Table with columns V1, V2. Rows include protein names like MAC, H2A, H2B, etc., and their corresponding amino acid sequences.

Table with columns for glycosylation NST, glycosylation NNT, glycosylation NIT, glycosylation NES, and fusion peptide start. Rows include sequences for HIV strains like H1B, H1A1, H1C, etc., with amino acid positions and domain labels like CD4 binding and V4.

Table with columns for glycosylation NST, glycosylation NNT, glycosylation NIT, glycosylation NES, and fusion peptide start. Rows include sequences for HIV strains like MAC, H2A, H2B, etc., with amino acid positions and domain labels like V4 and V5.

Table with 4 columns: accession ID, fusion peptide, glycosylation NAS, glycosylation NYT. Rows include H1B.FR.83.HXB2, H1A1.UG.85.U455 U455A, H1C.ET.86.ETH2220, etc.

Table with 4 columns: accession ID, fusion peptide, glycosylation NAS, glycosylation NYT. Rows include MAC.US.x.239, H2A.DE.x.BEN, H2A.PT.x.ALI, etc.

Table with columns for protein names (e.g., H1B, H1A1, H1C), accession numbers, and amino acid sequences. The sequences are aligned under two headers: 'transmembrane domain' and 'glycosylation NGS'. The table contains multiple rows of data, each representing a different HIV sequence variant.

H1B.FR.83.HXB2
H1A1.UG.85.U455.U455A
H1C.ET.86.ETH2220
H1D.CD.84.84Z8685
H1F1.BE.93.VI850
H1G.SE.93.SE6165_G6165
H1H.CF.90.056
H1J.SE.93.SE9280_7887
H1K.CM.96.96CM.MP535
H1O.BE.87.ANT78
H1O.CM.91.MVP5180
H1O.CM.94.BCF06
H1O.CM.96.96CMA102
H1O.CM.98.98CMA104
H1O.CM.99.99CMU4122
H1O.CM.x.CM02_3
H1O.FR.92.VAU
H1O.SN.99.99SE.MP1299
H1O.US.99.99USTWLA
H1O.US.x.I.24788
H1N.CM.02.DJ00131
H1N.CM.95.YBF30
H1N.CM.97.YBF106
H1P.CM.06.U14788
H1P.FR.86.RBF168
CPZ.CD.06.BF1167
CPZ.CD.98.ANT
CPZ.CM.05.SIVcpzDP943
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzMB897
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CAM3
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.SIVcpzGAB2
CPZ.TZ.00.TAN1
CPZ.TZ.01.TAN2
CPZ.TZ.06.SIVcpzTAN13
CPZ.TZ.09.UG38
CPZ.US.85.US_Marilyn

glycosylation NAT
AVSLLNATAIAVAEGTRDRIEIVV...
-IT-D-V-V--GWI--IG--TIG--LN--A-
-IN--T--V-G--F-LI--RIW--FCN--A-Q-
-I--VD--I--DI--RR--K-VL--T--A--A-
-I-F-T-V--R-G-VLN--T--A--A-
-I--DTV--NW--A-R--LN-T--A--A-
-ID--T--GI-VI--R-W-L--F-S--
-I--T--I-IA--R-F-L--A--
-I--T--G--I-IG--R-F-L--A--
-T-DTG-V-NW-GI-AGI--RTIG-N--S-
-TN-DTI-VS-NW-GI-LGI--RIGQGF--A-V-
-TN-DTI-V-GNW-SI-LGI--RIG-GFLN--A-A-N-
-TN-DTF-V-GKW-NI-LGI--RIG-G-LN--XX-S-
-T--DTI-V-GKW-DI-LGI--RIG-G-LN--X-S-
-TN-XDTL-V--NR-SIFGI--RIG-G-LN--O-A-V-
-TN-XDTL-V--NR-SIFGI--RIG-G-LN--O-A-V-
-TN-DTV-V--NW-SI-LGI--SIG-G-LN--L-
-T--DTI-V--NW-VTI-LGI--RIG-G-LN--S-
-T--DTV-V--GSW-NI-SG--RIG-G-LN--A-X-
-TN-VDTI-V--NW-SI-LG--RIG-G-WNV--F-S-
-I--T--V--I-LA--RIG-G-L--A-I-
-I--T--V--F-LA--RIG-G-L--A-
-T-X-DTV-S-NW-Q-A-G-OIG-GFLN--L-S-
-T--DTV-S-GW-O-I-G-OIG-GFLN--I-S-
CIVCWTLEFTGHW-G-AL-RVVDI-A-N-
IIKNIDRL-W-G-K-SILLAL--TIV-I-EV--I-A-N-
-T--DT--I--RIG-G-L--S-
-IN-DT-V--I-L--RIG-G-L--A-
-I--D--X-IA--RVG-G-L--A-
-I--D--G--IL-AI--TRLG-G-L--A-
-I--D--IL-A--IIG-G-L--S-
-T--DT--N--I-F--IVG-G-L--S-
-I--D--I-AF--VTL-I-N--A-
-I--DT--I--R--TI--A--TRIG-G-L--A-E-
LIQAIDRL-NFTGHW--LI-G--VVIAG-G-N--L-A-N-
LIIM-DFTGHW--LI-GI--TVIYG-G-N--A-N-
FIRN-DRL-NFT-WW--IL-GA--VNIF-G-N--T--I-A-N-
LIQAIDRL-NFTGHW--L-AG--AFVAQG--N--I-A-N-
-T--DT--I--LT--RRLFLG-I--S--

Env end
gp41 end
cytoplasmic tail end

856
856
852
857
853
867
846
849
843
864
873
878
879
881
872
869
874
879
872
899
843
853
840
883
879
885
859
850
836
842
853
860
866
855
857
872
872
871
865
837

MAC.US.x.239
H2A.DE.x.BEN
H2A.PT.x.ALI
H2A.SN.x.ST_HIV_2_ST
H2B.CI.x.EH0
H2B.GH.86.D295_ALT
H2G.CI.92.Abt96
H2U.FR.96.12034
COL.CM.x.CGU1
DEB.CM.04.SIVdeb04CMPF3061
DEB.CM.99.CM49
DEB.CM.99.CM5
DEN.CD.x.CD1_CM0580407
DRL.x.x.FA0
GOR.CM.04.SIVgorCP684con
GOR.CM.07.SIVgorCP2135con
GRV.ET.x.GRI_677_gri_1
GSN.CM.99.CN166
GSN.CM.99.CN71
LST.CD.88.SIVlhoest447
LST.CD.88.SIVlhoest485
LST.CD.88.SIVlhoest524
LST.KE.x.lho7
MND-1.GA.x.MNDGB1
MND-2.CM.x.CM16
MND-2.GA.x.M14
MND-2.x.x.5449
MNE.US.x.MNE027
MON.CM.99.L1_99CML1
MON.NG.x.NG1
MUS-1.CM.01.CM1239
MUS-1.CM.01.SIVmus_01CM1085
MUS-2.CM.01.CM1246
MUS-2.CM.01.CM2500
OLC.CI.97.97CI12
RCM.CM.00.SIVagi_00CM312
RCM.CM.02.SIVrcm_02CM8081
RCM.GA.x.SIVRCMGAB1
RCM.NG.x.NG411
SAB.SN.x.SAB1
SHM.CI.79.SIVsmCI2
SHM.LR.89.SIVsmL1B1
SHM.SI.92.SIVsmL92A
STM.US.89.5TM_57_16
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173_COMGNM
TAL.CM.00.266
TAL.CM.01.8923
TAN.UG.x.SIVagmTAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TY01_patent
WRC.CI.97.97CI14
WRC.CI.98.98CI04
WRC.CM.05.Pbt_05GM_X02

VOAVRSATETL-GAWGDLW-TL...
FOAFART-RETL-GAWGWLW-AA...
FRAFARIARETLTNTWRDLWGA...
FRAFAR-GETLTNAWRGFWGTL...
JQAAAR-AGETL-SARTSWG-L...
LRAAAR-GETL-SAGETLW-AL...
CTAATR-AOETLSTWRALWKTL...
FOAFGK-ARETLSTRGRELW-TL...
LVPGRGSLQTL-TLQ-WLRSA...
I-WAGGELGOTFORWAVALO-L...
ITWAGGELGOTFORWAVALO-L...
IIVAGQETIWEWLOGIAOVA-OGL...
LRD-GTSAIQGRATAEV-LAAL...
LOEVAQRIWREGROLGLSSARAL...
-I--DCV-VWT-DW-O-AIA--RIG-G-LN--S-
-I--DCV-VWT-DW-O-AIA--RIG-G-LN--S-
-TG-AQ-L-R-AR-AWG-LGAI--RS-Y-VINS--V-K-A-G-
-TGI-DSL--FTWNW-EA-LHAC--RRVW-EFLA--A-I-LFN-
VSGI-DSL-VFTWNW-EPLVOT--GRVW-EFLA--A-I-L-N-
WSWNKIRNRNRDISAAHARNLRL...
WSWNKIRNRNRDISAAHARDIRL...
WIRNKSRRNRGELSK-NCTNIRL...
WLREKRRNRGOLSS--KNIQL...
-GYTGWLO-NRIFT-CREA-IAA...
QGYAQLWRSRGHQL-LSTAACFR...
QSFARLWRSRGHQL-LSVIRGAA...
OGFAQLWRSRGHQL-LSG-RGAA...
VOAVRSATETL-GAWGDLW-TL...
VT--D-L-RFTIIV--ALLHAG--GRLW--VA--A-I-F-N-
CTGAD-L-RFTIIV--A-LRLG--GRLW-GLVAV--A-I-L-N-
GTAA-DOL--FTWNW-EPILOIG--RRVW-EFLA--A-I-L-N-
STAV-DOL-TFTWNW-ET-LOAG--RRVW-EFLA--V-A-I-L-N-
VSKG-D-L-FTWNW-EPILOIG--KIK-FLA--A-I-L-N-
VTAADIGL--WT-NW-A-LO-S-KOL-K-FLA-K--A-I-L-N-
RTLKQKRO...
-W-GEMA-ENAYTYWRGLOSLA...
SGW-AEMAOQNAVYFVWRGLCVTA...
SGW-AEMAOQNAVYFVWRGLCA...
SSW-AEMARTNAYYWRGLCA-C...
-A--TV-VVQ-ATSAS-S-RHAC...
COEAAR-ARETL-SAGRDLW-TL...
LOAAWR-ARETL-SAMRDLWOTL...
LOEAIA-AROTLVSRTGRTWOTL...
-QAAWR-ARETL-SAGRDLW-TL...
RIRNACR-SRERVSSQKARSRTFSLGRKW-PKWNRT-GS-IP-TTETT...
-I--GREV-HQMVAIWQALLAYA...
-A--GREIRDW--AIWQAIYAAT...
WRDGTI-L-RFTWNW-NG-LAAA...
WGDCTI-L-RFTWNW-SG-LAAA...
-ARFEL-L-S-AOSARTLWNAAC...
-OEAVV-L-RLAQNAGYQIWLAC...
-OEARL-V-RFAQNAHQIWLAC...
-OEAV-HL-SFARNAHQIWLAC...
-DEI-OTL-GVAQNAHQIWLAC...
KKIAKSRKENTQ--VSATGSRKT...
KKITKSRKETTEK--ISENSNRET...
EKIAARRQKREK-IYTPGSRET...
SRSSSIRGITA-LCTNLNPCRGGQ...

Env end
gp41 end

880
861
863
861
853
860
885
868
838
847
851
851
865
896
859
859
855
882
887
916
915
917
913
879
867
868
891
883
870
884
892
898
873
883
879
844
855
862
851
858
868
887
879
878
883
920
841
833
873
891
880
878
878
878
879
885
877

Table with columns: myristoylation, Nef start, acidic cluster, phosphorylation, poly-P helix. Rows list HIV sequences and their corresponding motifs across these categories.

HXB2 premature Nef end		normal Nef end	
H1B.FR.83.HXB2	LDLWIYHTOGYFPD*QNYTPGPGV.....RYPLTFGVCKYLVPEVPDKIEEA.NKGENTSLHPVSLHGMDDP...EREVLEWRFSRLAFHHVARELHPEYF.....KNC*		
H1A1.UG.85.U455A	-V-F-W-I.....F-D-AEV-TG-N-ICO-V-E-K-M-K-T-LK-R-Y.....FY.....D.....D.....		205
H1C.ET.86.ETH2220	-V-N-F-W-I.....F-D-SEV-I-E-NC-A-F-E-D-K-K-H-RR-M-Y.....D.....D.....		206
H1D.CO.84.84Z8085	-V-F-W-I.....F-D-EV-K-TE-DNC-ICO-F-E-K-V-N-E-K-K-KY-Y.....D.....D.....		216
H1F1.BE.93.VT850	-V-F-W-I.....F-L-FE-D-EEV-K-T-N-M-Q-D-D-V-K-S-L-R-R-Y.....D.....D.....		218
H1G.SE.93.SE6165_G6165	-V-N-W-T-I.....F-F-MD-AEV-V-N-ICO-F-E-D-V-S-RR-I-Y.....D.....D.....		208
H1H.CF.90.056	-V-N-W-E-I.....F-F-N-OEV-Q-E-N-M-E-D-G-M-K-LT-L-VK.....D.....D.....		207
H1J.SE.93.SE9280_7887	-VHN-W-I.....F-D-SEV-E-NC-ICO-TE-E-O-K-S-RR-I.....FY.....D.....		204
H1K.CM.96.96CM_MP535	-V-F-W-I.....F-D-AEV-T-TE-DNC-INO-E-E-H-I-M-K-S-RR-I.....D-V-Y.....D.....		216
H1O.BE.87.ANT70	-V-N-F-W-I.....F-LF-SAEAA-RLG-TN-DA-ACN-AE-A-HG-I-K-O-RS-GLT-MIT-L-Q.....D.....D.....		212
H1O.CM.91.MVP5180	-V-N-F-W-C-P.....F-LF-SAEAA-RLG-TN-DA-ACN-AE-A-HG-I-K-O-RS-GLT-I-LQK-L-Q.....D.....D.....		212
H1O.CM.94.BCF06	-VH-F-W-P.....F-LF-SAEAA-RLG-SS-DA-ACN-Y-DE-E-HK-I-K-O-RS-GLT-LQK-L-P.....D.....D.....		211
H1O.CM.96.96CMA102	-V-N-FL-W-T.....P-F-SAEAA-LG-SDRAK-ACN-FG-R-HG-I-K-O-RS-IT-KVTN-L-L.....D.....D.....		215
H1O.CM.98.98CMA104	-VHN-FX-W-T.....F-MF-SAEAA-LG-H-RAK-ACA-FG-A-HG-I-K-O-RS-GLT-KIT-L-P.....D.....D.....		214
H1O.CM.99.99CMU4122	-V-N-F-W-T.....F-LF-SAEAA-LG-CVTKA-ACN-SE-HG-I-K-O-RS-GLT-MIT-DL-H.....D.....D.....		212
H1O.CM.x.CM02_3	-V-N-F-W-T.....F-LF-SAEAA-LG-C-RAK-ACN-FG-A-HG-I-K-O-RS-GLT-MIT-DL-H.....D.....D.....		202
H1O.FR.92.VAU	-VH-F-W-T.....F-LF-SAEAA-ALG-C-RA-ACN-YE-Q-HK-I-K-O-RS-NT-LIT-L-S.....D.....D.....		216
H1O.SN.99.99SE_MP1299	-V-F-W-G-P.....F-LF-SAEAA-LG-C-RA-ACN-FE-N-HG01-K-O-RS-GST-MVTN-L-N.....D.....D.....		212
H1O.US.99.99USTWLA	-V-F-W-P.....F-LF-TEQEA-QLG-D-GAM-ACN-SE-H-HG-M-K-O-S-GIT-MVT-L-I.....D.....D.....		212
H1O.US.x.1_24788	-H-F-W-T.....F-LF-TE-EAKRLG-VC-RAH-ACN-FE-HG-I-K-O-RS-GNT-KIT-L-P.....D.....D.....		212
H1N.CM.02.DJ00131	-V-F-W-T.....F-LF-LSAEAV-E-N-WA-LCO-V-G-HK-Q-Q-S-RR-I.....DFY.....D.....		214
H1N.CM.95.YBF30	-M-IL-WH-I.....V-F-LSAEAV-E-D-NA-ICO-A-D-HK-V-S-RR-I.....FY.....D.....		213
H1N.CM.97.YBF106	-V-F-W-I.....XC-F-LSEAEV-E-D-NA-ICO-V-D-HK0-V-S-RR-K-D-FY.....D.....D.....		208
H1P.CM.06.U14788	-V-F-W-E.....F-LF-SEVEA-MGD-Q-KAK-ACTY-FS-HK-I-V-K-S-GRE-LQK-L.....IKD.....D.....		218
H1P.FR.06.RBF168	-T-MFN-I-W.....CR-LF-AD-PEED-R.....NI-M-ACS-Y-R-NG-L-V-K-E-RR-I-R.....D.....D.....		103
CPZ.CO.06.BF1167	-M-MFN-I-W-EE.....CR-F-SPPDD-R.....NI-ACT-DG-HK-I-R-E-AS-MRR-I-R.....RD.....D.....		199
CPZ.CO.05.SIVcpzDP943	-V-F-W-I.....F-LT-EEV-K-E-DTNI-ICO-EIE-E-HG-V-Y-LR-I-Q-Y.....D.....D.....		208
CPZ.CO.05.SIVcpzEK505	-V-I-W-I.....F-LS-EEV-Q-D-NV-MCO-E-E-DK-V-S-RV-R-FY.....D.....D.....		207
CPZ.CO.05.SIVcpzMB897	-V-F-W-I.....V-F-DTDV-E-D-NV-MCO-O-E-E-H-M-O-H-LT-R-K-Y.....D.....D.....		207
CPZ.CO.05.SIVcpzMT145	-V-F-W-I.....F-LT-EEV-K-E-DTNI-MCO-E-E-HG-V-T-K-LR-K-Y.....D.....D.....		206
CPZ.CO.98.CAM3	X-HGY-F-W-I.....Y-G-F-LT-EEV-Q-E-DNI-ICO-F-E-HG-V-LR-I-Q-Y.....D.....D.....		218
CPZ.CO.98.CAM5	-V-F-W-I.....F-LT-EEV-K-E-DTNI-ICO-E-E-HG-V-LK-L-Q-YGDHNPAPAQ.....D.....D.....		214
CPZ.GA.88.GAB1	-V-F-W-T-T.....F-C-F-LTEEQV-Q-E-D-NC-ICO-E-E-DK-V-LR-I-Q-Y.....D.....D.....		206
CPZ.GA.88.SIVcpzGAB2	-V-N-WN-I.....Y-F-S-EAV-DNL-LCT-FE-E-K-I-K-Q-LR-L-K-Y.....RD.....D.....		206
CPZ.TZ.00.TAN1	-T-M-N-V-W-I.....CR-LF-D-PEDD-K.....NI-ACS-DG-T-I-S-RR-I-RV.....D.....D.....		196
CPZ.TZ.01.TAN2	-T-MFN-I-W.....Y-F-L-ADDPYENDER-NI-DAHOGQ-EDPHK-R-V-K-S-YCYK-GVAEOR-EHRRCM-FP.....RK.....D.....		196
CPZ.TZ.06.SIVcpzTAN13	-T-M-N-V-W-I.....CR-LF-T-PEDD-R.....NI-ANS-RE-DGDG-T-M-Q-TS-RR-I-R.....D.....D.....		197
CPZ.TZ.09.UG38	-M-M-N-I-W-I.....CR-LF-D-PEDD-R.....NI-YCS-RE-AG-G-N-I-C-S-RR-I-RY.....D.....D.....		198
CPZ.US.85.US_Marilyn	-V-I-W-I.....Y-F-LTEEEV-Q.....TNI-MCO-E-E-HG-I-Q-TE-RR-R-K.....D.....D.....		210

max HIV-1 similarity		normal Nef end	
MAC.US.x.239	--TYLEKEE-II-W-D-S-I.....K-LW---NVS-D-OEDE.EHY-M-AQTSOW--PWG--A-K-PT-YTYE-VYRY-E-GSKSGLSEEEVRRRLTARGLL...NMADKKETR.....		263
H2A.DE.x.BEN	--TYLEKEE-II-W-H-I.....-MY-LW-S-LS-Q-EEDE.ANC-V-AQTSRH...EHG-T-V-Q-M-YNYK-FY-Y-E-GHKSGLPEKEWKARLKARGIP...YSE.....		258
H2A.PT.x.ALI	--TYLEKEE-II-W-H-I.....-MF-LW-DVP-Q-GEDE.THC---QTSRH---THG-T-V---PK-HDYK-FI---E-GYKSLPEDEWKARLKARGIP...FSKNRNS.....		263
H2A.SN.x.ST_HIV_2_ST	--TYLEKEE-II-G-W-H.....-KF-LW-DVP-Q-GDDE.THC-V-AQTSRF---PHG-T-V---PT-SYE-FIRY-E-GYKSLPEDEWKARLKARGIP...FS.....		256
H2B.CI.x.EH0205_ALT	--TYLENEE-IVSGW-H.....-KF-LW-IMH1-A-P-EDE-THC-W-AQTSAW---PHE-T-V-Q-L-YDY-FRFE-E-GYKSLPEDEWKARLKARGIP...TE.....		259
H2B.GH.06.D205_ALT	--TYLENEE-IVSGW-H.....-TY-FE-L-ADDPYENDER-NI-DAHOGQ-EDPHK-R-V-K-S-YCYK-GVAEOR-EHRRCM-FP.....RK.....D.....		240
H2G.CI.92.Abt96	--TYMEKE-II-W-A-I.....-X-K-LW-TVS-D-OEDE.THC-V-AQTSFP...P-T-A-K-PT-YDYR-FI---E-RWKSGLPEAVWKEKLRQGLP...IE.....		257
H2U.FR.96.12034	--TYLENEE-II-W-S-T-I.....-W-MC-LW-E.DAN.-DK-QEDE.RLY-VGSAQTSCEE...HWG-A-V-K-S-YSYQ-FIKC-E-GSKSGLSEEEVRRRLTARGLP...V.....		224
COL.CM.x.CU1	IEQYA-TEW-CLKGWLE-EDLE-EDGALKED-K-VA-LW-YI-QL.....GEY-AYSVDLSLLSVTSRKKKKKQOQVATEMVD.....		232
DEB.CM.04.SIVdeb04CMP3061	--E-HAQNEW-II-GWLO-E-I.....K-Y-FLF-IADDPYENDER-NI-DAHOGQ-EDPYK-R-V-K-Q-YCYK-GHEAHTKETHRRCM-FP.....RK.....		252
DEB.CM.99.CM40	--E-HAQNEW-II-GWLO-E-I.....TY-FLF-IADDPYENDER-NI-DAHOGQ-EDPHK-R-V-K-S-YCYK-GVAEOR-EHRRCM-FP.....RK.....		231
DEB.CM.99.CM5	--EIAHAH-EW-II-GWLK-E-P.....E-FLW-TIEE-YDNKEE-NC-DRYEGQQ-ADPW-T-V---PE-WCYK-GHKKLOTSOHNNYCO...RAK.....		249
DEN.CO.x.CD1_CM0580407	--YAOQNEH-FITGW-T-D-I.....E-FLW-TIEE-YDNKEE-NC-DRYEGQQ-ADPW-T-V---PE-WCYK-GHKKLOTSOHNNYCO...RAK.....		249
DRL.x.x.FA0	--YAOQNEH-FITGW-S-K-I.....-KM-FLF---V-EDLV-NQOCCR-N-S-QM-PI-D-P-G-R-M-Q-PG-YTFQ-CIH---E-GHVKSIKS...VEEPGCKRKKWQFKPTTEGGCHDNLQK...ORD.....		269
GOR.CM.04.SIVgorCP684con	--V-W-W-I.....Y-FLF-SAEAA-MGTAS-RAI-ACS-R---HG-T-V-RA-GTT-LQK-L.....D.....D.....		211
GOR.CM.07.SIVgorCP2135con	--V-W-W-I.....Y-FLF-SAEAA-MGTAS-RAI-ACS-R---HG-T-V-RA-GTT-LQK-L.....D.....D.....		211
GRV.ET.x.GRI_677_gri_1	-N-YALNEW-IID-W-A-S-I.....-RV-F-LF-SAEAA-MGTAS-RAI-ACS-R---HG-T-V-RA-GTT-LQK-L.....D.....D.....		224
GSN.CM.99.CN166	EQYA-LEW-LV-GWLSF-T-D-I.....TIP-F-IC-R-ATTEDS-PGDD-FL-T-A-YOGRS-EDOHK-FVFS-C-K---IKSGIOLDLOQOEEERKR...LTAN.....		219
GSN.CM.99.CN71	-ESYQ-YEW-IV-GWLQW-I.....T-MP-F-LW-R-AMTDS-PGDD-FL-T-A-YOGRS-EDOHK-FVFS-C-K---IKSGIOLDLOQOEEERKR...LTAN.....		218
LST.CO.88.SIVLhoest447	-C-YCENEW-LIG-FMT-D-T.....F-A-LWO-E-AC-EYTPD-SDYROC-S-QL-VOED-PWG-R-I-H-NPM-VDY-LRKO-KSQIQAASIGL...ORR.....		206
LST.CO.88.SIVLhoest485	-C-YCENEW-LIG-FMT-D-T.....F-LWO-E-AC-EYTPD-SDASOC-S-QL-VOED-PWG-R-I-H-NPM-VDF-LRKO-KSQIQAASISLN...C.....		208
LST.CO.98.SIVLhoest524	-T-YCENEW-LIG-FM-D-T.....A-LWO-E-AC-EFKAP-SD-ROC-S-QL-VLED-PWG-R-I-H-NPM-VDY-LRKO-ASVQATAMRFN...C.....		207
LST.KE.x.lho7	-I-YCENEW-LIG-FM-D-T.....I-LWO-E-AC-EYKDP-SD-TOC-S-QL-VLED-PWG-R-I-H-NPM-VDF-I-LKKQ-AKIQNTAFAFD...C.....		208
MND-1.GA.x.MNDG61	--YAEQNEW-FE-GW-Q-T-T.....K-FLF-E-SRAJGD---YAA-NH-S-QLCPQED-P-G-T-M-SGTLI-PMTLQH.....		215
MND-2.CM.98.CM16	--YAOQNEW-FI-TW-S-D-I.....K-RFLF-C-AVPPDD---NN-CNK-Q-S-QL-IQEE#PWG-R-V-K-G-YTFY-PIKR-GE-KHVQISISHIAYAKEHKPECCCKRKKWQF...C.....		231
MND-2.GA.x.M14	-N-YAOQNEW-FI-W-E-S-I.....KR-FLF-I-QVPPDD---NQ-CNR-N-S-QL-IQEE#PWG-R-M-K-PE-YTFY-PIKC-KEYRVHTSLSVTAYQ...EKSDCCKRKKWQF...C.....		264
MND-2.x.x.5449	--YAOQNEW-LI-NW-K-I.....K-FLF-AVSPPL-ED-CNR-N-S-QM-IQED-P-G-R-M-K-G-YTFY-PIKR-EVKCVTSLSYEAYKKEKPKDCKRKKWQF...C.....		251
MNE.US.x.MNE027	--TYLEKEE-IV-W-D-S-P.....K-LW-NVS-D-OEGE.ENY-AQTSOW--PWG-V-K-PT-YTYE-VYRY-E-GSKSGLSEEEVRRRLTARGLL...KMADKKETS.....		233
MON.CM.99.L1_99CML1	IEQYA-LEW-LI-GWLO-E-I.....T-MP-FLWC-R-ATTEDS-EGDED-FL-T-A-YOGR-E-PH-QF-VFS-C-K---VKSQROLAOLQOEEERKR...LAANRIL.....		225
MON.NG.x.NG1	EQYA-VMW-LV-GWLOF-X.....TIP-FLWC-R-AIQEDS-DGXDE-FL-T-A-YOGRDE-PHKOF-VXS-C-K-G-VKSQPOLDLQOEEERKR...LTANRIL.....		229
MUS-1.CM.01.CM1239	ETYT-VDW-IV-GWLOF-E-T.....TLP-FLWC-RE-ATTEDS-EGDDE-YL-T-AAYOGKEE-PHKOF-VFS-C---MKSQROLDOQOEEERKR...LTANRIL.....		218
MUS-1.CM.01.SIVmus_01CM1085	ETI-A-VEW-II-GWLM-D-T.....KTP-FLWC-RE-AMTDS-EGDDE-YL-D-AFOGREE-HHKOF-VFS-C---LKSGLOLDNMOQOEEERKR...LTANRIL.....		218
MUS-2.CM.01.CM1240	-TYA-VKSKIV-RSLI-D-T.....RQF-FLWC-R-AMTDS-EGDD-YL-T-AVYGRDE-HHK-FVFS-C---LKSQROLNMOQOEEERKR...LTANRIL.....		218
MUS-2.CM.01.CM2500	ETYA-VEW-LI-GWLM-E-I.....MP-VLFC-R-HITENS-LGDDE-HL-C-AFIGREE-PHK-FVFS-C---LKSQROLNMOQOEEERKR...LAANRILSQSP.....		222
OLC.CI.97.97CI12	IY-YALNCW-IIGNF-ATEVEEETGL.....R-K-P-LWD-RR---LPELLEERG.....		152
RCM.CM.00.SIVagi_00CM312	--TYLENEH-II-W-P.....TF-LWM-DVSDAEEK-D-EH---AES-IE---PWG-T-A-K-NPM-VDYIGYR-DF-GERKN.....TO.....		225
RCM.CM.02.SIVrcm_02CM8081	--TYLENEH-II-W-K-I.....MM-LW-DVIDEAK-D-EH---AETC-GYR-FE-E-D-GERKN.....TO.....		225
RCM.GA.x.SIVRCMGAB1	--YMENEH-II-W-S-I.....TM-LWO-DVSDATEE...N-RHC-AETS---PWG-A-K-KPL-VYAGYR---F-GERKN.....TK.....		225
RCM.NG.x.NG411	--YALNEW-IVDGV---D-T.....KM-F-LF-DLS---K-SE-HC-AQVAYE---AWK-T-V-K-PL-VDY-WR---QVPSAOG.....		227
SAB.SN.x.SAB1	--MYLENEE-II-W-S-I.....MM-LW-NVS-D-OEDE.RHC-V-AQISSL---PWG-A-K-PQ-YNYE-FIRY-E-GSKSGLSEDEVKRRLTARGLL...KMADKKENS.....		263
SMM.CI.79.SIVsmCI2	--TYLEKEE-IV-W-S-I.....MF-LW-MDVS-D-OEDE.THC-V-AQTSH---PWG-A-S-PK-YTYE-VYRY-E-GSKSGLSEDEVKRRLTARGLL...KMADKKETR.....		262
SMM.LR.89.SIVsmLB1	--SYL-NEE-II-W-R-I.....KO-LW-DVS-D-DEDDGTHY-V-AQT-QV---PWG-T-A-K-NPM-YTYE-FVRY-E-GWKSGLTEVKKRLASRKKPNWKKHQHAEQES...C.....		240
SMM.SL.92.SIVsmSL92A	--MYLENEE-IV-W-A-S-I.....K-LW-DVS-D-DEDDGTHY-V-AQT-QV---PWG-V-K-PL-HTYE-FVR-E-GSKSGLPEEVEVRRLTARGLL...KMADKKETS.....		265
STM.US.x.97_16	--YAEQNEW-FITGW-D-K-I.....KA-LW-A-TI-EDDRP-NHPCOA---S-Q0-VNED-PWG-R-I-T-PT-YDYR-IQK---E-KHVTSLOW.EV.....		209
SUN.GA.98.L14	-Q-YLKNEH-VID-DIT-S-T.....I-LWE-A-N-IEGLVDEEDT...LMM-AAGV-ASED-PH-N-M-N-NPH---YTPGWEARQOLEROTGKR.....		229
SYK.KE.x.KES1	-H-YLQNEH-IID.RI-S-T.....I-LWE-N-IEGL-YEETH...L-L-A-GO-SS-SMG-PHVE-LOPPPG...YTPGWEARLQEROTGKPOELOSALSKNIS.....		241
SYK.KE.x.SYK173_COMGNM	-N-YAHNEW-IL-W-E-T.....C-TLW-C-IHEDDEDE...HL-P-AVDGOAE---PWG-S-V-V-EK---YTPGAKMA-WDRLEKRLMLAPPOTASS.....		231
TAL.CM.01.8023	-N-YAHNEW-IL-W-S-E.....C-TLW-C-IHEDDEDE...HL-M-AVDGOAE---PWG-A-V-V-EK---YTPGAKMA-WDRLEKRLMLAPPOTASS.....		231
TAN.UG.x.SIVagntAN1	-N-YALNEW-IID-W-A-S-I.....K-C-F-F-DVS-Q-OD-E.RHC-AQIENES---PWK-T-V-K-NPL-VQYNP-DSFKDMH...GLVKKR.....		223
VER.DE.x.AGM3	-N-YALNEW-IID-WNAWSE-I.....RC-F-F-DLH---ETCE.RHC-V-AQVREDP-GINHG-V-K-PM-VQYDPNRKYLDMH...DLGKRK.....		230
VER.KE.x.9063	-N-YALNEW-IID-WNAWST-I.....RC-F-F-DLH---ETCE.RHC-V-AQIREDP-GISHG-T-V-K-PM-VQYDPNRKYF-DMH...ATVKKR.....		230
VER.KE.x.AGM155	-N-YALNEW-IID-WNAWSK-I.....RC-F-F-ALH---ETCE.RHC-V-AQ-EDP-GINHG-I-A-K-PM-VQYDPNSREVFYDLS...TVGNTG.....		231
VER.KE.x.TY01 patent	-N-YALNEW-IID-WNAWSK-I.....F-TL-F-DLH---OTCE.RHC-V-AQSGDP-GISHG-I-V-K-PM-IQYDPNSREVFYDLS...GLVKKR.....		234
WRC.CI.97.97CI14	WN-YAHNEW-LIDGW-A-E-I.....W-K-W-HFSEESNDTPYERKN-LHPACAEPE---PWG-H-V-K-PK-CDF-GR-PGQSATLQKAEATESLYKNK_KOCK.....		234
WRC.CI.98.PBC104	WN-YAFNEW-LIDGW-A-E-I.....W-K-W-HFSEESNDSPYERKN-LHPGCOESE---PWG-H-V-K-PK-CDF-SR-PGHIAAQGEDALETLHKELKRRK.....		234
WRC.GM.05.Pbt_05GM_X02	WN-YA-EW-IVEGW-A-D.....K-W-HFSEESNPSDQ-FEKN-LHPAAESE---PWK-V-V-K-PK-CDF-GR-ADGOIATG...IETLYEMKKRNK.....		231