

VII

PLV Proteins

Contents

VII-1	Introduction	411
VII-2	Sequences	412
VII-3	Alignments	422
	VII-3.1 Gag	422
	VII-3.2 Pol	426
	VII-3.3 Vif	433
	VII-3.4 Vpx	436
	VII-3.5 Vpr	437
	VII-3.6 Tat	438
	VII-3.7 Rev	440
	VII-3.8 Vpu	442
	VII-3.9 Env	443
	VII-3.10Nef	450

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)
H1B.US.90.WEAU160_GHOSH	U21135	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Wei, X.	<i>Nature</i> 422 (6929):307-312 (2003)
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)
H101_AE.TH.90.CM240	U54771	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Carr, J.K.	<i>J Virol</i> 70 (9):5935-5943 (1996)
H102_AG.NG.x.IBNG	L39106	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Howard, T.M.	<i>ARHR</i> 10 (12):1755-1757 (1994)

Name	Accession	Proteins	Author	Reference
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
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.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.02.SJGddd	GQ324959	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Vallari, A.	<i>ARHR</i> 26 (1):109-115 (2010)
H1N.CM.04.04CM_1131_03	DQ017383	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Yamaguchi, J.	<i>ARHR</i> 22 (1):83-92 (2006)
H1N.CM.06.U14296	GQ324962	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Vallari, A.	<i>ARHR</i> 26 (1):109-115 (2010)
H1N.CM.06.U14842	GQ324958	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Vallari, A.	<i>ARHR</i> 26 (1):109-115 (2010)

Name	Accession	Proteins	Author	Reference
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)
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)
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.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.CM.01.SIVcpzCAM13	AY169968	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Nerrienet, E.	<i>J Virol</i> 79 (2):1312-9 (2005)
CPZ.CM.05.LB715	JX178450	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Takehisa	Unpublished
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.06.SIVcpzTAN13	JQ768416	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Takehisa	Unpublished
CPZ.TZ.06.TAN5	JN091691	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Rudicell, R.S.	<i>J Virol</i> 85 (19); 9918-28 (2011)
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)
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)

Name	Accession	Proteins	Author	Reference
MAC.US.x.239	M33262	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Kestler, H.	<i>Science</i> 248 (4959):1109-1112 (1990)
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.86.ST_JSP4_27	M31113	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Kumar, P.	<i>J Virol</i> 64 (2):890-901 (1990)
H2B.CL.x.EHO	U27200	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Rey-Cuille, M.A.	<i>Virology</i> 202 (1):471-476 (1994)
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.CI.07.07IC_TNP3	KC693505	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Ayouba, A.	<i>AIDS</i> 27 (15):2488-2491 (2013)
H2U.FR.96.12034	AY530889	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Damond, F.	<i>ARHR</i> 20 (6):666-672 (2004)
COL.CM.x.CGU1	AF301156	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Courgnaud, V.	<i>J Virol</i> 75 (2):857-866 (2001)
COL.UG.10.BWC01	KF214240	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Lauck, M.	<i>Retrovirology</i> 10 (1):107 (2013)

Name	Accession	Proteins	Author	Reference
COL.UG.10.BWC07	KF214241	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Lauck, M.	<i>Retrovirology</i> 10 (1):107 (2013)
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)
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)
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.SIVgor2139_287	FJ424866	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Takehisa, J.	<i>J Virol</i> 83 (4):1635-1648 (2009)
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)
GSN.CM.99.CN166	AF468659	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Courgnaud, V.	<i>J Virol</i> 76 (16):8298-8309 (2002)
GSN.CM.99.CN71	AF468658	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Courgnaud, V.	<i>J Virol</i> 76 (16):8298-8309 (2002)
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)

Name	Accession	Proteins	Author	Reference
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)
MAC.US.x.EMBL_3	Y00295	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Franchini, G.	<i>Nature</i> 328 (6130):539-543 (1987)
MND-1.GA.x.MNDGB1	M27470	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Tsujimoto, H.	<i>Nature</i> 341 (6242):539-541 (1989)
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.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.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)
MON.CM.99.L1_99CML1	AY340701	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Courgnaud, V.	<i>J Virol</i> 77 (23):12523-12534 (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)

Name	Accession	Proteins	Author	Reference
OLC.CI.97.97CI12	FM165200	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	<i>J Virol</i> 83 (1):428-439 (2009)
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)
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)
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)
SMM.SL.92.SL92B	AF334679	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Chen, Z.	<i>J Virol</i> 70 (6):3617-3627 (1996)
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)

Name	Accession	Proteins	Author	Reference
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)
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)
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)
TAN.UG.x.SIVagmTAN1	U58991	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Soares, M.A.	<i>Virology</i> 228 (2):394-399 (1997)
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; DNAVEC CORPORATION
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)

Name	Accession	Proteins	Author	Reference
MON.NG.x.NG1	AJ549283	Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Barlow, K.L.	<i>J Virol</i> 77 (12):6879-6888 (2003)
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.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
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)
H10.CM.96.96CMABB009	AY169806	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H10.CM.96.96CMABB637	AY169810	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H10.CM.97.97CMABB447	AY169813	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H10.CM.97.97CMABB497	AY169809	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H10.CM.98.98CMA104	AY169802	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H10.CM.98.98CMA105	AY169816	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H10.CM.98.98CMABB141	AY169807	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H10.CM.98.98CMABB197	AY169811	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H10.CM.98.98CMABB212	AY169804	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H10.CM.98.98CMU2901	AY169812	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H10.CM.98.98CMU5337	AY169808	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H10.CM.99.99CMU4122	AY169815	Vpu	Yamaguchi, J.	<i>ARHR</i> 19 (11):979-988 (2003)
H10.CM.x.pCMO2_3	AY618998	Vpu	Tebit, D.M.	<i>Virology</i> 326 (2):329-339 (2004)
H10.CM.x.pCMO2_5	AY623602	Vpu	Tebit, D.M.	<i>Virology</i> 326 (2):329-339 (2004)
H10.SN.99.99SE_MP1300	AJ302647	Vpu	Toure-Kane, C.	<i>ARHR</i> 17 (12):1211-1216 (2001)
H1N.CM.04.04CM_1015_04	DQ017382	Vpu	Yamaguchi, J.	<i>ARHR</i> 22 (1):83-92 (2006)

Name	Accession	Proteins	Author	Reference
H1N.CM.97.YBF106	AJ271370	Vpu	Roques, P.	<i>AIDS</i> 18 (10):1371-1381 (2004)
CPZ.CM.05.SIVcpzDP943	EF535993	Vpu	Van Heuverswyn, F.	<i>Virology</i> 368 (1):155-171 (2007)
CPZ.CM.05.SIVcpzEK505.c2	JN835460	Vpu	Bibollet-Ruche, F.	<i>J Clin Invest</i> 122 (5):1644-1652 (2012)
CPZ.CM.05.SIVcpzEK505	DQ373065	Vpu	Keele, B.F.	<i>Science</i> 313 (5786):523-526 (2006)
CPZ.CM.05.SIVcpzLB7	DQ373064	Vpu	Keele, B.F.	<i>Science</i> 313 (5786):523-526 (2006)
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.SIVcpzMB897	EF535994	Vpu	Van Heuverswyn, F.	<i>Virology</i> 368 (1):155-171 (2007)
CPZ.CM.05.SIVcpzMT145.c2	JN835462	Vpu	Bibollet-Ruche, F.	<i>J Clin Invest</i> 122 (5):1644-1652 (2012)
CPZ.CM.05.SIVcpzMT145	DQ373066	Vpu	Keele, B.F.	<i>Science</i> 313 (5786):523-526 (2006)
CPZ.CM.98.CAM3	AF115393	Vpu	Corbet, S.	<i>J Virol</i> 74 (1):529-534 (2000)
CPZ.CM.98.CAM5	AJ271369	Vpu	Muller-Trutwin, M.C.	<i>J Med Primatol</i> 29 (3-4); 166-72 (2000)
CPZ.GA.88.GAB1	X52154	Vpu	Huet, T.	<i>Nature</i> 345 (6273):356-359 (1990)
CPZ.TZ.01.TAN2	EF394357	Vpu	Takehisa, J.	<i>J Virol</i> 81 (14):7463-7475 (2007)
CPZ.TZ.02.TAN3_1	DQ374658	Vpu	Takehisa, J.	<i>J Virol</i> 81 (14):7463-7475 (2007)
GOR.CM.07.SIVgorCP2135con	FJ424863	Vpu	Takehisa, J.	<i>J Virol</i> 83 (4):1635-1648 (2009)

H1B.FR.83.HXB2
H1A1.UG.85.U455 U455A
H1B.US.90.WEAU160 GHOSH
H1C.ET.86.ETH2220
H1K.CD.84.842885
H1F1.BE.93.VI850
H1G.SE.93.SE6165 G6165
H1H.CF.90.056
H1J.SE.93.SE9280 7887
H1K.CM.96.96CM MP535
H101.AE.H.90.CM240
H102.AG.NG.x.IBNG
H10.BE.87.ANT70
H10.CM.91.MVP5180
H10.CM.94.BCF06
H10.CM.96.96CMA102
H10.FR.92.VAU
H10.SN.99.99SE MP1299
H10.US.99.99USTWLA
H10.US.x.1.24788
H1N.CM.02.UJ00131
H1N.CM.02.SJGddd
H1N.CM.04.04CM 1131_03
H1N.CM.06.U14296
H1N.CM.06.U14842
H1N.CM.05.Y0F30 PLT3
H1P.CM.06.U14788
H1P.FR.06.RBF168
CPZ.CD.06.BF1167
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.H1B15
CPZ.GA.88.SIVcpzGAB2
CPZ.TZ.00.TAN1
CPZ.TZ.06.SIVcpzTAN13
CPZ.TZ.06.TAN5
CPZ.TZ.09.UJ038
CPZ.US.85.US.Marilyn

Table with columns for protein names, amino acid sequences, and major homology regions. The table lists various HIV sequences and their corresponding amino acid alignments, with major homology regions indicated by bold text.

MAC.US.x.239
H2A.DE.x.BEN
H2A.TI.x.AL1
H2A.SN.86.ST JSP4 27
H2B.CI.x.EHO
H2B.GH.86.D205 ALT
H2C.CI.92.Abt96
H2U.CI.07.07TC TNP3
H2U.FR.96.L2034
COL.CM.x.CGU1
COL.UG.10.BWC01
COL.UG.10.BWC07
DEB.CM.04.SIVdeb04CMFP3061
DEB.CM.99.CM40
DEB.CM.99.CM5
DEN.CD.x.CD1 CM0580407
DRL.x.x.FAO
GOR.CM.84.SIVgorCP684con
GOR.CM.07.SIVgor2139 287
GRV.ET.x.GRI 677 gri_1
GSN.CM.99.CN166
GSN.CM.99.CN71
LST.CD.88.SIVlhoest485
LST.CD.88.SIVlhoest524
LST.KE.x.lho7
MAC.US.x.EMBL 3
MND.1.GA.x.MNDG81
MND.2.CM.98.CM16
MND.2.GA.x.H14
MND.2.x.x.5440
MNE.US.x.MNE027
MON.CM.99.L1 99CML1
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.NG41
SAB.SN.x.SAB1
SMM.CI.79.SIVsmC12
SMM.LR.89.SIVsmL1B1
SMM.SL.92.SIVsmSL92A
SMM.TP.92.SIVsmTP92
STM.US.89.STM 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.SIVagtAN1
VER.DE.x.AGM3
VER.KE.x.9663
VER.KE.x.AGM155
VER.KE.x.901 patent
WRC.CI.97.97CI14
WRC.CI.98.P8C104
WRC.CM.05.PBt 05GM_X02

Table with columns for protein names and amino acid sequences. Rows include HIV proteins (H1B, H1A1, H1B, H1C, etc.) and MAC proteins (MAC.US.x.239, H2A, H2B, etc.).

Table with columns: p7 end p1 start, p1 end p6 start, Vpr binding, Vpr binding p6 end, Gag end. Rows include sequences for HIV strains like H1B.FR.83.HXB2, H1A1.UG.85.U455A, etc.

Table with columns: p8 end p1 start, p1 end p6 start, PTAP motif, PSAP in HIV-2 B. U, p6 end Gag end. Rows include sequences for MAC.US.x.239, H2A.CD.x.BE01, H2A.PT.x.ALI, etc.

Table with columns for HIV sequence identifiers (e.g., H1B, H1A1, H1A), protease end p66, p51 RT start, M41L, K70R D67N, D110 catalytic site, and Pol. The table contains multiple rows of sequence alignments and their corresponding amino acid positions.

H18.FR.83.HXB2
H1A1.UG.85.U455.U455A
H1A.US.90.WEAU160.GHOSH
H1C.ET.86.ETH2220
H1D.CD.84.84Z085
H1F1.BE.93.VI850
H1G.SE.93.SE6165.G6165
H1H.CF.90.056
H1J.SE.93.SE9280.7887
H1K.CM.96.H617.MP85
H101.AE.H.90.CM240
H102.AG.NG.x.IBNG
H10.BE.87.ANT70
H10.CM.91.MVP5180
H10.CM.94.BCF06
H10.CM.96.HCM102
H10.FR.92.VAU
H10.SN.99.99SE.MP1299
H10.US.99.99USTWLA
H10.US.x.I.24788
H1N.CM.02.DJ00131
H1N.CM.02.SJGddd
H1N.CM.04.04CM.1131_03
H1N.CM.06.U14296
H1N.CM.06.U14842
H1N.CM.09.YM1387
H1P.FR.06.U14788
H1P.FR.06.RBF168
CPZ.CD.06.BF1167
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.IB715
CPZ.GA.88.SIVcpzGAB2
CPZ.TZ.00.TAN1
CPZ.TZ.06.SIVcpzTAN13
CPZ.TZ.06.TAN5
CPZ.TZ.09.US638
CPZ.US.85.US_Marilyn

Table of amino acid sequences for HIV-1 p66 RT, p15 Rnase H end, and p31 Integrase start. The table contains three columns of sequences, with the first column aligned under the first two columns of the second. The sequences are separated by line breaks and contain various amino acid symbols and gaps (-).

MAC.US.x.ZEN
H2A.DE.x.B39
H2A.IT.x.AL1
H2A.SN.86.ST.J5P4.27
H2B.CI.x.EH0
H2B.GH.86.D205.ALT
H2B.CI.92.Abt196
H2C.CI.07.071C.TMP3
H2U.FR.96.L2034
COL.CM.x.CGU1
COL.UG.10.BWC01
COL.UG.10.BWC07
INSS.EEEI.V.HEGS
INSS.EEEEI.V.HEGS
ISEM.O4.SIVc06b04CMPF3061
DEB.CM.99.CM40
DEB.CM.99.CM5
DRL.x.x.FAO
GOR.CM.04.SIVc0rCP684cn
GOR.CM.99.SIVc0r2139.287
GRV.ET.x.GRI.677.gri_1
GSN.CM.99.CN166
GSN.CM.99.CN171
LST.CD.88.SIV1hoest485
LST.CD.88.SIV1hoest524
LST.KE.x.lh07
MAC.US.x.EMBL.3
MND.1.GA.x.MNDG81
MND.2.CM.98.CM16
MND.2.GA.x.M14
MND.2.x.x.5440
MNE.US.x.MNE027
MON.CM.99.L1.99CML1
MON.NG.x.NG1
MUS.1.CM.01.CM1239
MUS.1.CM.01.SIVmus1C1CM1085
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
OLC.CI.97.97CI12
RCM.CM.00.SIVag1_09CM312
RCM.CI.02.SIVc0r02CM0801
RCM.GA.x.SIVRCMGAB1
RCM.NG.x.NG411
SAB.SN.x.SAB1
SHM.CI.79.SIVsm1C2
SHM.FR.89.SIVsm1TB1
SHM.SI.89.SIVsm1S92A
SHM.SL.92.SL92B
STM.US.89.STM.37.16
SUN.GA.98.L14
SVK.KE.x.KE51
SVK.KE.x.SVK173_COMGNM
TAL.CM.00.266
TAL.CM.01.8023
TAN.UG.x.SIVagmTAN1
VER.DE.x.AGM3
VER.KE.x.9863
VER.KE.x.AGM155
VER.KE.x.TY01.patent
WRC.CI.97.97CI14
WRC.CI.98.P8C104
WRC.GM.05.Pbt.05GM_X02

Table of amino acid sequences for HIV-1 p15 Rnase H end and p31 Integrase start. The table contains three columns of sequences, with the first column aligned under the first two columns of the second. The sequences are separated by line breaks and contain various amino acid symbols and gaps (-).

Table with columns for accession numbers (e.g., H1B.FR.83.HXB2), amino acid sequences, and protein names (e.g., QLDCTHLEKGVLLVAHVHVS...).

H1B.FR.83.HXB2
 H1A1.UG.85.U455.U455A
 H1B.US.90.WEAU160.GHOSH
 H1C.ET.86.ETH2220
 H1D.CD.84.84Z085
 H1F1.BE.93.VI850
 H1G.SE.93.SE6165.G6165
 H1H.CF.90.056
 H1J.SE.93.SE9280.7887
 H1K.CM.96.96CM.MP35
 H101.AE.H.90.CM240
 H102.AG.NG.x.IBNG
 H10.BE.87.ANT70
 H10.CM.91.MVP5180
 H10.CM.94.BCF06
 H10.CM.96.96CM102
 H10.FR.92.VAU
 H10.SN.99.99SE.MP1299
 H10.US.99.99USTWLA
 H10.US.x.I.24788
 H1N.CM.02.DJ00131
 H1N.CM.02.SJGddd
 H1N.CM.04.04CM.1131_03
 H1N.CM.06.U14296
 H1N.CM.06.U14842
 H1N.CM.06.YBF30
 H1P.CM.06.U14788
 H1P.FR.06.RBF168
 CPZ.CD.06.BF1167
 CPZ.CD.90.ANT
 CPZ.CM.01.SIVcpzCAM13
 CPZ.CM.05.LB715
 CPZ.GA.88.SIVcpzGAB2
 CPZ.TZ.00.TAN1
 CPZ.TZ.06.SIVcpzTAN13
 CPZ.TZ.06.TAN5
 CPZ.TZ.09.UG38
 CPZ.US.85.US_Marilyn

SRNPLWKGPALKLWKGEVAVIQDNS.DIKVVPRKAKIIRDYQKMGDDCVASRODED.....
 -D-I-----M-G-----
 -D-----A-G-----
 -D-I-----V-G-----
 -D-V-----N-E-E-----E-----
 -D-I-----P-----E-----E-----
 -D-I-----E-----G-----
 -D-I-----G-----G-----
 -D-I-----G-----G-----
 -D-I-----Q-----K-G-----E-----T-SM-G-T-SESVEOPSEIP
 -D-I-----Q-----K-G-----T-SM-N-T-SESMEOPGEIP
 -D-I-----Q-----K-G-----N-V-T-S-----SESMEOPGETP
 -D-I-----Q-----K-G-----X-----A-S-X-G-X-SESSE+PGEIP
 -D-I-----Q-----K-G-----X-----A-S-----G-T-SESVEOPSEIP
 -D-I-----Q-----K-G-----H-----T-SM-G-T-SENVEOPGEIP
 -D-I-----Q-----K-G-E-----T-M-G-T-SENMEQXSEIS
 -D-I-----Q-----K-G-----T-SM-----T-SESVEOPGEIP
 -D-I-----G-----G-----A-NOOMA
 -D-I-----G-----G-----NOEME
 -D-I-----G-----G-----NOOME
 -D-I-----G-----G-----NODMD
 -D-I-----G-----G-----NODME
 -D-I-----G-----G-----NOEME
 -D-I-----T-----K-G-----N-----DT-R-SESLEQSG
 -D-I-----T-----K-G-----DI-R-SESLE
 -D-V-----R-----KEGE-E-----K-V-AGG-D-N-----
 -D-V-----Q-----K-GE-E-----K-E-R-KIEDR-DL-G-N-----
 -D-I-----T-----K-GE-----L-OG-----V-N-M-GG-----SQGME
 -D-I-----S-----L-GE-E-----KH-V-L-GG-----NONME
 A-D-I-----R-----KEGE-----KE-----AGGMD-N-T-----
 A-D-I-----R-----KEGE-----KE-----AGS-DD-N-A-----
 A-D-V-----R-----KEGE-----KE-----AGSMD-N-T-----
 G-D-I-----R-----K-AGGMD-N-NOEME
 -D-I-----R-----KE-E-EV-----K-----A-SM-G-----S-----

p31 Integrase end
 Pol end

p31 Integrase end
 Pol end

1003
 1003
 1004
 1011
 1004
 1005
 1001
 1003
 1002
 1004
 1009
 1004
 1010
 1010
 1010
 1011
 1011
 1010
 1013
 1010
 1013
 1010
 1010
 1015
 1015
 1016
 1015
 1008
 1004
 1007
 1004
 1003
 1013
 1011
 1013
 999
 1000
 999
 1006

MAC.US.x.239
 H2A.DE.x.BEN
 H2A.PT.x.ALI
 H2A.SN.86.ST.J5P4_27
 H2B.CI.x.EHO
 H2B.GH.86.D205.ALT
 H2G.CI.92.Abt96
 H2U.CI.07.071C.TNP3
 H2U.FR.96.12034
 COL.CM.x.CGU1
 COL.UG.10.BWC01
 COL.UG.10.BWC07
 DEB.CM.04.SIVdeb04CMPF3061
 DEB.CM.99.CM40
 DEB.CM.99.CM5
 DEN.CD.x.CD1.CM0580407
 DRL.x.x.FA0
 GOR.CM.04.SIVgorCP684con
 GOR.CM.07.SIVgor2139.287
 GRV.ET.x.GRI.677.gri_1
 GSN.CM.99.CN166
 GSN.CM.99.CN71
 LST.CD.88.SIVlhoest485
 LST.CD.98.SIVlhoest524
 LST.KE.x.lho7
 MAC.US.x.EMBL_3
 MND-1.GA.x.MNDG81
 MND-2.CM.98.CM16
 MND-2.GA.x.M14
 MND-2.x.x.5440
 MNE.US.x.MNE027
 MON.CM.99.L1.99CML1
 MON.NG.x.NG1
 MUS-1.CM.01.CM1239
 MUS-1.GM.01.SIVmus_01CM1085
 MUS-2.CM.01.CM1246
 MUS-2.CM.01.CM2500
 OLC.CI.97.97CI12
 RCM.CM.00.SIVag1_00CM312
 RCM.CM.02.SIVcm_02CM8081
 RCM.GA.x.SIVRCMGAB1
 RCM.NG.x.NG411
 SAB.SN.x.SAB1
 SHM.CI.79.SIVsmC12
 SHM.LR.89.SIVsmL181
 SHM.SL.92.SIVsmSL92A
 SHM.SL.92.SL92B
 STM.US.89.STM.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.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.GM.05.Pbt_05GM.X02

G-DQ---GE-----ILKVGTL-----K---GGKEV-SSSHMEDTGEAREVA
 H2A.DE.x.BEN-----IVKVGTL-----GROEL-SGPHLEGAREDEGEVA
 G-DQ---GE---D---IVKVGTL-----GROEL-SGPHLEGAREDEGEVA
 G-DQ---GE---D---IVKVGTA---II-----K---GROEM-SGSNLEGAREDEGEVA
 G-DQ---GD---D---IKVGT-E-I-----N---GGKEL-CSADVEDTMOAREVAQSN
 G-DQ---GD---D---IKVGT-E-I-----H---GGKGL-CSADMEDTROAREMAQSD
 G-DQ---GE---D---IKVGT-E-I-----K---GGKEV-SSTMEDTROAREVA
 G-DQ---GE---D---IKVGT-E-I-----K---GGKEV-SGSNLEGAREDEGEVAALGEP
 G-DQ---D---L-KVGT-I-----GGKEL-SSSLEGA-KVREMLPDPTPEYNTKELQKVCYVPHKVGWAMWTCSRVTPL
 H.TGE-Q-GE-V-----RNSQGLF-K-V-T-LQYGEDV-SENLL-NGQKEATVKGMD
 H.TGE-Q-GQ-V-----KNQEGSLF-K-V-TKVNYSGERMVKALLPNEQ-EAGEVKEVD
 H.TGE-Q-GQ-V-----RSTEGTLF-K-V-S-ITHGT00-VSLEDLPMGKPKGETETLD
 GAD-T-----T-----KLDTG-L-----V-K---DVGSKDPDTE
 GAD-T-----T-----CKTEVG-----V-K---D-DSKVNTONSNE
 GAD-T-----T-----CKTEVG-----V-K---GVDSKTDTONPSE
 GA-Q-Q-E-----KTOAG-FP-----KP-NAE-SRKGEQNKGM-SETDI
 G-DQ---E-V-----KEGT-L-I-----K---NVDSTNME
 -D-I-----T-----K-G-----DT-G-SESLEQSS
 -D-I-----T-----K-G-----DT-G-SESLEQSS
 G-D-V-----R-I-----KEGE-----K---ERKTM-SEGSMEGVREANKQMEGSDLDQDE
 GA-Q-Q-V-----V-TQAGE-IT-----KP-EAKTE-VGG-THTNN
 GA-H-Q-V-----V-TQAGE-IT-----KP-EAK-E-VGS-AHTSN-REEGRMAD
 G-D-Q---QI-----VKEGE-T-FS-----LVK---EGPKQSEGLDNN
 G-D-Q---Q-----VKEGE-N-FSI-----LVK---EGPKQSSMDN-KD
 G-D-Q---Q-----VKEGE-N-FS-----LVK---EGPKQSESSLDNN
 G-DQ---GE-----ILKVGTL-----K---GGKEV-SSSHMEDTGEAREVA
 G-DQ---GI-----LKYE-E-I-----C---K---E-SGKNSO-NLESV
 G-DQ0---E-I-----KEGT-L-----TVDS-PN-EA
 G-EQ---E-I-----KEGT-L-----TVDSNPHMED-ETA
 G-DQ0---E-I-----KEGV-L-----K---TVDSNPHMGR
 G-DQ---GE-----ILKVGTL-----K---GGKEV-SSSHMEDTGEAREVA
 GAD-Q-R-----KTEEG-LT-----KP-TTENVGDDTNOYLRKQDGLAN
 GAD-S-H-----L-KTDQGEVIT-----P-SKEDVGSKPSAH-TREVDMAD
 GSD-S-A-----VKTFQGOVIT-----KP-AKENVGSKSNITGDHRKEDGLDN
 GAD-S-R-----KTD-GVLDN-----K-KTD-GVLDN
 GA-Y-LR-R-----L-KTKEGE-VT-----K-TRONV-SEPD-VHVRKEDGLAD
 GA-S-Q-Q-----I-KTDQGE-IT-----K-TKENV-SDPNPZY-RKEDGLAD
 NKEGO-C-E-----LLKEGD.KYFS-L-----VKP-VDSSANV
 G-DQ0---T-I-----GQ-L-----C-VK---RKNVSETNMEGG0-EN
 G-EQ---G-----GQ-L-----C-K---RKNVSETNMEGG0-EN
 G-DQ-R-I-----EETG-L-----K---RKDVSSEANLAGRO-EN
 G-DQ0---I-E-----QE-L-----C-K---RKEVDRETNMEGRQ-ES
 G-D-V-----I-----EQG.EL-TI-----K---L-SOAPLEGNRTAGEVD
 G-DQ---N-----I-KVGT-I-----K---GGKELGNSPYLENP-EDGKMAQPD
 G-DQ---GE-----ILKVGTL-----K---GGKEV-SGSHLEDTREAREVA
 G-DQ---E-----I-KVGT-I-----K---GGKEM-SSADMENTOVRETAQLSEISKV
 G-DQ---E-----I-KVGT-E-I-----K---GROEMGSSASMEDQ
 G-DQ---GE-----IVKVGTL-----K---GGKEV-SGSHLEDTREAREVA
 G-D-Q-----KQGE-N-L-----LVK---GESSVEM-G
 GT-SD-Q-----TPDQOVIA-----SSD-ERVDSGTHLEITTSKN
 GK-NE-----R-----V-TEEG-FA-----I-H-ER-DGSH-ENDPKT
 GA-Q-Q-H-I-----LRTDEGEVIT-----KP-QALGNKIDLE-SKEQ-AEMGRDN
 GA-Q-Q-H-V-----LRTDEGEVIT-----KP-QA-GNKTDLEGSKEQ-AEMGRDN
 G-E-V-----T-----KEGE.EL-----K---ERKTVG-KTNMEG
 G-D-V-----GO-I-----KGV.EL-EY-----K---EPRKRMG-ESNLEGAGGADN
 G-D-V-----R-I-----LKEGE.EL-----K---EPRKRMG-ESNLEGAGGADN
 G-D-V-----Q-I-----LKEGE.EL-----K---EPRKRMG-ESNLEGAGGADN
 G-D-V-----Q-I-----LKEGE.EL-----K---EPRKRMG-ESNLEGAGGADN
 G-D-V-----Q-I-----LKEGE.EL-----K---EPRKRMG-ESNLEGAGGADN
 G-D-V-----Q-I-----LKEGE.EL-----K---EPRKRMG-ESNLEGAGGADN
 GA-Q-Q-G-P-----L-ETPE-G-IT-----L-KVNWINGEMDRSSDKN
 GA-Q-Q-G-----L-VETPE-G-IT-----K-KVNWINGEM#
 GASQ-Q-G-V-----L-V-TPEG-LLT-----K-TWDGKGM#

1020
 1070
 1033
 1033
 1036
 1035
 1016
 1022
 1066
 991
 993
 995
 1011
 1008
 1008
 1032
 1013
 1003
 1003
 1041
 1016
 1017
 1020
 1021
 1020
 1014
 1010
 1011
 1017
 1012
 1016
 1018
 1018
 1012
 1017
 1017
 1014
 995
 1018
 1018
 1019
 1023
 1040
 1028
 1010
 1023
 1006
 1016
 1017
 1041
 1037
 1020
 1021
 1020
 1021
 1020
 1023
 1034
 1032
 1026
 1023
 1014
 1007

H1B.FR.83.HXB2
H1A1.UG.85.U455.U455A
H1B.US.90.WEAU160.GHOSH
H1C.ET.86.ETH2220
H1D.CD.84.842085
H1F1.BE.93.V1850
H1G.SE.93.SE6165.G6165
H1H.CF.90.056
H1J.SE.93.SE9280.7887
H1K.CM.96.96C.MP35
H101.AE.H.90.CM240
H102.AG.NG.x.IBNG
H10.BE.87.ANT70
H10.CM.91.MVP5180
H10.CM.94.BCF06
H10.CM.96.96CMA102
H10.FR.92.VAU
H10.SN.99.99SE.MP1299
H10.US.99.99USTWLA
H10.US.x.I.24788
H1N.CM.02.DJ00131
H1N.CM.02.SJ6dd
H1N.CM.04.04CM.1131_03
H1N.CM.06.U14296
H1N.CM.06.U14842
H1N.CM.05.YB36
H1P.CM.06.U14788
H1P.FR.06.RBF168
CPZ.CD.06.BF1167
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.LB15
CPZ.GA.88.SIVcpzGAB2
CPZ.TZ.00.TAN1
CPZ.TZ.06.SIVcpzTAN13
CPZ.TZ.06.TAN5
CPZ.TZ.09.U6388
CPZ.US.85.US_Marilyn

Vif start

Table with 100 columns and 100 rows of amino acid sequences for Vif start region, including accession numbers like ENRW, QVMIVWQVDRMR, and various amino acid abbreviations.

MAC.US.x.239
H2A.DE.x.BEN
H2A.PT.x.AL1
H2A.SN.86.ST.J5P4.27
H2B.CI.x.EH0
H2B.GH.86.D205.ALT
H2C.CI.92.Abt96
H2U.CI.07.0716.TMP3
H2U.FR.96.L2034
COL.CM.x.CGU1
COL.UG.10.BWC01
COL.UG.10.BWC07
DEB.CM.04.SIVdeb04CMPF3061
DEB.CM.99.CM40
DEB.CM.99.CM5
DEN.CD.x.CD1.CM0580407
DRL.x.x.FA0
GOR.CM.04.SIVgorCP684c0n
GOR.CM.07.SIVgor2139.287
GRV.ET.x.GRI.677.gri_1
GSN.CM.99.CN166
GSN.CM.99.CN171
LST.CD.88.SIVlhoest485
LST.CD.88.SIVlhoest524
LST.KE.x.lho7
MAC.US.x.EMBL.3
MND.1.GA.x.MNDGB1
MND.2.CM.98.CM16
MND.2.GA.x.M14
MND.2.x.x.5440
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.97.97C12
RCM.CM.00.SIVag1.00C312
RCM.CM.02.SIVcm.02CM0801
RCM.GA.x.SIVRCMGAB1
RCM.NG.x.NG411
SAB.SN.x.SAB1
SHM.CI.79.SIVsmc12
SHM.FR.89.SIVsm1TB1
SHM.LI.92.SIVsm1S192A
SHM.SL.92.SL92B
STM.US.89.STM.37.16
SUN.GA.98.L14
SVK.KE.x.KE51
SVK.KE.x.SYK173.COMGNM
TAL.CM.00.266
TAL.CM.01.8023
TAN.UG.x.SIVagmTAN1
VER.DE.x.AGM3
VER.KE.x.908E
VER.KE.x.AGM155
VER.KE.x.TY01.patent
WRC.CI.97.97C114
WRC.CI.98.PBC104
WRC.CM.05.PB.05GM.X02

Vif start

Table with 100 columns and 100 rows of amino acid sequences for Vif start region, including accession numbers like MAC.US.x.239, H2A.DE.x.BEN, and various amino acid abbreviations.

H1B.FR.83.HXB2	Vif end	192
H1A1.UG.85.U455_U455A	...	193
H1B.US.90.WEAU160_GHOSH	...	193
H1C.ET.86.ETH2220	...	193
H1D.CD.84.84ZRO85	...	193
H1F1.BE.93.VI850	...	193
H1G.SE.93.SE6165_G6165	...	193
H1H.CF.90.056	...	193
H1J.SE.93.SE9280_7887	...	193
H1K.CM.96.96CM_MP35	...	193
H101.AE.TH.90_CM240	...	190
H102.AG.NG.x.IBNG	...	193
H10.BE.87.ANT70	...	193
H10.CM.91.MVP5180	...	193
H10.CM.94.BCF06	...	193
H10.CM.96.96CMA102	...	193
H10.FR.92.VAU	...	193
H10.SN.99.99SE_MP1299	...	193
H10.US.99.99USTWLA	...	196
H10.US.x.I_24788	...	193
H1N.CM.02.DJ00131	...	193
H1N.CM.02.SJGddd	...	193
H1N.CM.04.04CM_1131_03	...	193
H1N.CM.06.U14296	...	193
H1N.CM.06.U14842	...	193
H1N.CM.95.YBF30	...	193
H1P.CM.06.U14788	...	196
H1P.FR.06.RBF168	...	196
CPZ.CD.06.BF1167	...	197
CPZ.CD.90.ANT	...	193
CPZ.CM.01.SIVcpzCAM13	...	194
CPZ.CM.05.LBT15	...	193
CPZ.GA.88.SIVcpzGAB2	...	193
CPZ.TZ.00.TAN1	...	199
CPZ.TZ.06.SIVcpzTAN13	...	199
CPZ.TZ.06.TAN5	...	199
CPZ.TZ.09.UG38	...	199
CPZ.US.85.US_Marilyn	...	193
Vif end		
MAC.US.x.239	...	215
H2A.DE.x.BEN	...	210
H2A.PT.x.ALI	...	216
H2A.SN.86.ST_JSP4_27	...	216
H2B.CI.x.EH0	...	217
H2B.GH.86.D205_ALT	...	217
H2G.CI.92.Abt96	...	216
H2U.CI.07.071C_TNP3	...	215
H2U.FR.96.12034	...	214
COL.CM.x.CGU1	...	172
COL.UG.10.BWC01	...	173
COL.UG.10.BWC07	...	172
DEB.CM.04.SIVdeb04CMPF3061_VLS	...	251
DEB.CM.99.CM40	...	231
DEB.CM.99.CM5	...	231
DEN.CD.x.CD1_CM0580407	...	259
DRL.x.x.FA0	...	218
GOR.CM.04.SIVgorCP684con	...	193
GOR.CM.07.SIVgor2139_287	...	193
GRV.ET.x.GRI_677_gri_1	...	220
GSN.CM.99.CN166	...	238
GSN.CM.99.CN71	...	238
LST.CD.88.SIVlhoest485	...	228
LST.CD.88.SIVlhoest524	...	227
LST.KE.x.lho7	...	227
MAC.US.x.EMBL_3	...	215
MND-1.GA.x.MNDGB1	...	173
MND-2.CM.98.CM16	...	213
MND-2.GA.x.M14	...	213
MND-2.x.x.5440	...	214
MNE.US.x.MNE027	...	215
MON.CM.99.L1_99CML1	...	256
MON.NG.x.NG1	...	258
MUS-1.CM.01.CM1239	...	236
MUS-1.CM.01.SIVmus_01CM1085	...	239
MUS-2.CM.01.CM1246	...	248
MUS-2.CM.01.CM2500	...	204
OLC.CI.97.97CI12	...	221
RCM.CM.00.SIVag1_00CM312	...	221
RCM.CM.02.SIVrcm_02CM8081	...	221
RCM.GA.x.SIVRCMGAB1	...	217
RCM.NG.x.NG411	...	223
SAB.SN.x.SAB1	...	235
SMM.CI.79.SIVsmCI2	...	217
SMM.LR.89.SIVsmLFB1	...	215
SMM.SL.92.SIVsmSL92A	...	214
SMM.SL.92.SL92B	...	218
STM.US.89.STM_37_16	...	215
SUN.GA.98.L14	...	242
SYK.KE.x.KE51	...	224
SYK.KE.x.SYK173_COMGNM	...	222
TAL.CM.00.266	...	228
TAL.CM.01.8023	...	228
TAN.UG.x.SIVagmTAN1	...	239
VER.DE.x.AGM3	...	233
VER.KE.x.9063	...	232
VER.KE.x.AGM155	...	233
VER.KE.x.TY01_patent	...	236
WRC.CI.97.97CI14	...	208
WRC.CI.98.98CI04	...	208
WRC.GM.05.Pbt_05GM_X02	...	209

	Vpx start	Vpx end
MAC.US.x.239	MSD..PRERIPPGNSGEETIGEAFEWLNRTVEEINREAVNHLPRELIFQVWORSWEYWHDEQGMSPSYVYKRYLCLIQKALFMHCKKGCRCLEGEHGAGGW.RPGPPPPPPGLA*	112
H2A.DE.x.BEN	-T.....V.....E-I-AL.....R.....A-T.....M.....I-T-F-R-T-W-DM-RE-LEDQ.....V.....	114
H2A.PT.x.ALI	-AN.....TV.....D.....AL.....R.....Q-T.....M.....MYT-FM-T-G-P.....S.....V.....	113
H2A.SN.86.ST.JSP4_27	-AG.....T.....D.....A.....R.....I-T.....M.....M-I-S-R-T-G-P.....S.....V.....	113
H2B.CI.x.EH0V.....V.....ET-L-HL-V.....K-A-RE.....I-T.....M.....M-I-FA-G-R-P.....S.....	113
H2B.GH.86.D295.ALTV.....D.....V.....A-IT-L-V.....-A-RE.....I-T.....L-M-M-V-YT-Q-P.....S.....	112
H2G.CI.92.Abt196X-E-D.....E-XT-I.....R.....A-T.....L-M-M-V-T-QK-P.....Q.....	112
H2U.CI.07.071C.TNP3D-VE.....E-S-TA-Q.....G.....T.....Y.....R-Q-R-P.....T.....	113
H2U.FR.96.12034	-G.....E-V-Q.....S-E-RD-A-Q.....Q-R.....E-T.....Y.....R-Q-R-P.....T.....	113
DRL.x.x.FA0	-AERQSV-A-AEPM-AGEVELE-Q-SLLR-Q-RL-FHP-FL-RL-NTCM-HY-ALQL-FT-S-L-L-M-FQ-QGR-PPPLRPAGDRL-PP-	114
MAC.US.x.251.IA11Q.....M.....	113
MAC.US.x.251.BK28Q.....M.....	113
MND-2.CM.98.CM16	-AE....GA-EIPE-AGEVDLN.T-E-SL-K-Q-RL-FHP-FL-RL-NACI-H-RHOR-L-A-L-MN-M-T-MQOE-P-RSG-P.....MV-	100
MND-2.GA.x.M14	-AE....A-EAPO-AGEV-LE-Q-E-SL-QV-QL-FHP-FL-RL-NTCV-H-RLRRTLN-A-L-M-M-V-MQO-P-RSG-S.....	100
MND-2.x.x.5440	-AE....A-EAPE-AGEV-LE-Q-ETSL-R-RL-FHP-FL-RL-NTCV-H-RHOR-LD-A-L-MH-MYT-MQO-P-RNG-RPR.....M-	100
MNE.US.x.MNE627Q.....M.....	113
RCM.CM.00.SIVagi_00CM312	-AE....A-E-PT-AGEVEFOPWLARMY-INQEA.RL-F-G-F-HL-RTCV-RTLE-AG-L-M-T-MRS-KLR...DPPR-OR-ERV-IL-MQ-	106
RCM.CM.02.SIVrcm_02CM8081	-AE....A-EVPT-AGEVEFOPWLARMY-INQEA.RL-FHP-F-RL-RTCV-H-L-R-LE-AG-L-M-T-MRS-LR...DPPR-OR-ERV-IL-MQ-	106
RCM.GA.x.SIVRCMGAB1	-AE....A-EVPT-AGEAEFOPWLRDMLKVNLEA.RL-FHP-F-RL-RTCV-H-VHOR-LE-AA-L-M-I-QT-SQR...PNPR-AV-ERITIL-M-	105
RCM.NG.x.NG411	-AE.G-V-EAPT-AGDVEF-PWLRHMLT-VNLEA.RL-FHP-F-RL-RTCV-H-RL-R-LE-AG-L-M-I-QS-SQR...Q-QAREA-ERIQIL-M-	109
SMM.CI.79.SIVsmCI2	-GS.A-T-Y-S-AV-G-K-R-F-QA-TQ-L-M-C-V-T-R-S-A.....	113
SMM.LR.89.SIVsmLIB1E-Q.....R-F-A-T-M-L-S-S.....	112
SMM.SL.92.SIVsmSL92AQ-D-D-E-V-T-Q-R-F-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.SL.92.SL92B	-T.....HN-AL-QT-Q-R-C-V-Y-A-VQ-M-M-Q-FR-T-R-SQ-T.....	113
SMM.US.04.G078H-A-R-Y-R-E-T-W-M-R.....	113
SMM.US.04.G932D.....S.....M-EI.....	113
SMM.US.04.M919D-D-A-R-KI-V-T-R-E-E-S.....	113
SMM.US.04.M922L-D-D-A-R-M-E-T-M-R-E-S.....	113
SMM.US.04.M923E-A-R-I-V-T-GEY.....	113
SMM.US.04.M926	-A.....D-T-R-V-R-YE-Q-Q.....S.....	113
SMM.US.04.M934V-D-D-Q-A-R-M-E-T-M-R-E-S.....	113
SMM.US.04.M935V-D-D-Q-A-R-M-E-T-M-R-E-S.....	113
SMM.US.04.M940V-D-Q-A-R-YE-PR-Q-R-GEY-S.....	113
SMM.US.04.M946V-D-E-M-A-R-I-V-T-W-R-GEY-S.....	113
SMM.US.04.M947V-D-E-M-A-R-I-V-T-W-R-GEY-S.....	113
SMM.US.04.M949V-D-E-M-A-R-I-V-T-W-R-GEY-S.....	113
SMM.US.04.M950V-D-Q-A-R-YE-Q-Q.....S.....	113
SMM.US.04.M951V-H-D-R-YE-R-E-T-W-M.....	113
SMM.US.04.M952AA-VL-H-D-R-YE-Q-E-I.....	113
SMM.US.05.D215	-N.....AA-VL-H-D-R-YE-Q-E-I.....	113
SMM.US.06.FTQE-D-V-R-T-YI-D.....	113
SMM.US.11.SIVsmE660_FL10E-D-A-R-R-A-T-V-R-T.....	113
SMM.US.11.SIVsmE660_FL14E-D-A-R-R-A-T-V-R-T.....	113
SMM.US.11.SIVsmE660_FL6E-D-A-R-R-M-E-T-V-E-T.....	113
SMM.US.11.SIVsmE660_FL8E-D-A-R-R-M-E-T-V-E-T.....	113
SMM.US.86.CFU212	-A.....E-G-A-R-R-C-F-M-A-T-M-M-R-N-R-E-Q.....	113
SMM.US.x.H9	-XX.X-X-D-H-X-L-V-R-X-M-V-T-X-M-M-R-N-R-E-Q.....	113
SMM.US.x.PGM53E-G-A-R-R-M-E-T-V-R-E-S.....	113
SMM.US.x.SIVsmH635F_L3E-G-A-R-R-M-E-T-V-R-E-S.....	113
SMM.x.x.pE660.CG7GE-G-A-R-R-M-E-T-V-R-E-S.....	113
SMM.x.x.pE660.CG7VE-G-A-R-R-M-E-T-V-R-E-S.....	113
STM.US.89.STM_37_16A-E-H-D-R-NT-CM-HY-ALQL-FT-S-L-L-M-FQ-QGR-PPPLRPAGDRL-PP-	113

Table with columns: Vpr start, oligomerization, amphipathic alpha-helix, H(S/N)RIG motifs, Vpr end in HXB2, frameshift in HXB2, Vpr end. Rows include sequences like MEQA, H1A1.UG.85.U455.U455A, H1B.US.90.WEAU160.GHOSH, etc.

Table with columns: Vpr start, oligomerization, amphipathic alpha-helix, H(S/N)RIG motifs, Vpr end in HXB2, frameshift in HXB2, Vpr end. Rows include sequences like MAC.US.x.239, H2A.DE.x.BEN, H2A.CI.x.ALI, etc.

H1B.FR.83.HXB2
H1A1.UG.85.U455.U455A
H1B.US.90.WEAU160.GHOSH
H1C.ET.86.ETH2220
H1D.CD.84.84Z085
H1F1.BE.93.VI850
H1G.SE.93.SE6165.G6165
H1H.CF.90.056
H1J.SE.93.SE9280.7887
H1K.CM.96.96CF06
H101.AE.H.90.CM240
H102.AG.NG.x.IBNG
H10.BE.87.ANT70
H10.CM.91.MVP5180
H10.CM.94.BCF06
H10.CM.96.96CMA102
H10.FR.92.VAU
H10.SN.99.99SE.MP1299
H10.US.99.99USTWLA
H10.US.x.I.24788
H1N.CM.02.DJ00131
H1N.CM.02.SJGddd
H1N.CM.04.04CM.1131.03
H1N.CM.06.U14296
H1N.CM.06.U14842
H1N.CM.06.Y14736
H1P.CM.06.U14788
H1P.FR.06.RBF168
CPZ.CD.06.BF1167
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.LB715
CPZ.CM.88.SIVcpzGAB2
CPZ.TZ.00.TAN1
CPZ.TZ.06.SIVcpzTAN13
CPZ.TZ.06.TAN5
CPZ.TZ.09.U63011
CPZ.US.85.US_Marilyn

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

MAC.US.x.239
H2A.DE.x.BEN
H2A.CI.x.AL1
H2A.SN.86.ST.J5P4.27
H2B.CI.x.EH0
H2B.GH.86.D205.ALT
H2C.CI.92.Abt96
H2U.CI.07.0716.TMP3
H2U.FR.96.L2034
COL.CM.x.CGU1
COL.UG.10.BWC01
COL.UG.10.BWC07
DEB.CM.04.SIVdeb04CMPF3061
DEB.CM.99.CM40
DEB.CM.99.CM5
DEN.CD.x.CD1.CM0580407
DRL.x.x.FA0
GOR.CM.04.SIVgorCP684con
GOR.CM.92.SIVgor2139.287
GRV.ET.x.GRI.677.gri_1
GSN.CM.99.CN166
GSN.CM.99.CN71
LST.CD.88.SIVLhoest4
LST.CD.88.SIVLhoest524
LST.KE.x.lho7
MAC.US.x.EMBL.3
MND.1.GA.x.MNDG81
MND.2.CM.98.CM16
MND.2.GA.x.M14
MND.2.x.x.4440
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.97.97C12
RCM.CM.00.SIVag1.00CM312
RCM.CI.92.SIVrcm.02CM0801
RCM.GA.x.SIVRCMGAB1
RCM.NG.x.NG411
SAB.SN.x.SAB1
SHM.CI.79.SIVsmC12
SHM.FR.89.SIVsm1TB1
SHM.US.99.SIVsm1S92A
SHM.SL.92.SL92B
STM.US.89.STM.37.16
SUN.GA.98.L14
SVK.KE.x.KE51
SVK.KE.x.SYK173.COMGNM
TAL.CM.00.266
TAL.CM.01.8023
TAN.UG.x.SIVagmTAN1
VER.DE.x.AGM3
VER.KE.x.9863
VER.KE.x.AGM155
VER.KE.x.TY01.patent
WRC.CI.97.97C14
WRC.CI.98.98C104
WRC.GM.05.PB05M.X02

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

	Tat end	
H1B.FR.83.HXB2	100
H1A1.UG.85.U455_U455A	102
H1B.US.90.WEAU160_GHOSH	102
H1C.ET.86.ETH2220	101
H1D.CD.84.84Z085	87
H1F1.BE.93.VI850	102
H1G.SE.93.SE6165_G6165	102
H1H.CF.90.056	100
H1J.SE.93.SE9280_7887	102
H1K.CM.96.96CM_MP35	102
H101.AE.TH.90.CM240	102
H102.AG.NG.x.IBNG	101
H10.BE.87.ANT70	115
H10.CM.91.MVP5180	116
H10.CM.94.BCF06	93
H10.CM.96.96CMA102	79
H10.FR.92.VAU	80
H10.SN.99.99SE_MP1299	98
H10.US.99.99USTWLA	80
H10.US.x.I_24788	98
H1N.CM.02.DJ00131	102
H1N.CM.02.SJGddd	102
H1N.CM.04.04CM_1131_03	102
H1N.CM.06.U14296	102
H1N.CM.06.U14842	102
H1N.CM.95.YBf30	102
H1P.CM.06.U14788	72
H1P.FR.06.RBF168	72
CPZ.CD.06.BF1167	101
CPZ.CD.90.ANT	74
CPZ.CM.01.SIVcpzCAM13	100
CPZ.CM.05.LB715	103
CPZ.GA.88.SIVcpzGAB2	100
CPZ.TZ.00.TAN1	137
CPZ.TZ.06.SIVcpzTAN13	137
CPZ.TZ.06.TAN5	137
CPZ.TZ.09.UG38	134
CPZ.US.85.US_Marilyn	103
	Tat end	
MAC.US.x.239	131
H2A.DE.x.BEN	131
H2A.PT.x.ALI	137
H2A.SN.86.ST_JSP4_27	131
H2B.CI.x.EH0	139
H2B.GH.86.D205_ALT	129
H2G.CI.92.Abt96	113
H2U.CI.07.071C_TNP3	122
H2U.FR.96.12034	112
COL.CM.x.CGU1	80
COL.UG.10.BWC01	56
COL.UG.10.BWC07	57
DEB.CM.04.SIVdeb04CMPF3061	97
DEB.CM.99.CM40	111
DEB.CM.99.CM5	97
DEN.CD.x.CD1_CM0580407	99
DRL.x.x.FA0	135
GOR.CM.04.SIVgorCP684con	83
GOR.CM.07.SIVgor2139_287	83
GRV.ET.x.GRI_677_gri_1	95
GSN.CM.99.CN166	113
GSN.CM.99.CN71	120
LST.CD.88.SIVlhoest485	101
LST.CD.88.SIVlhoest524	112
LST.KE.x.lho7	112
MAC.US.x.EMBL_3	118
MND-1.GA.x.MNDGB1	116
MND-2.CM.98.CM16	123
MND-2.GA.x.M14	88
MND-2.x.x.5440	111
MNE.US.x.MNE027	132
MON.CM.99.L1_99CML1	112
MON.NG.x.NG1	120
MUS-1.CM.01.CM1239	110
MUS-1.CM.01.SIVmus_01CM1085	107
MUS-2.CM.01.CM1246	117
MUS-2.CM.01.CM2500	110
OLC.CI.97.97CI12	108
RCM.CM.00.SIVag1_00CM312	116
RCM.CM.02.SIVrcm_02CM8081	113
RCM.GA.x.SIVRCMGAB1	106
RCM.NG.x.NG411	116
SAB.SN.x.SAB1	94
SMM.CI.79.SIVsmCI2	133
SMM.LR.89.SIVsmLB1	114
SMM.SL.92.SIVsmSL92A	103
SMM.SL.92.SL92B	105
STM.US.89.STM_37_16	132
SUN.GA.98.L14	103
SYK.KE.x.KE51	99
SYK.KE.x.SYK173_COMGNM	111
TAL.CM.00.266	111
TAL.CM.01.8023	128
TAN.UG.x.SIVagmTAN1	102
VER.DE.x.AGM3	120
VER.KE.x.9063	120
VER.KE.x.AGM155	120
VER.KE.x.TY01_patent	101
WRC.CI.97.97CI14	125
WRC.CI.98.98CI04	132
WRC.GM.05.Pbt_05GM_X02	112

Table with 4 columns: Accession ID, NLS, Leu-rich effector domain, and protein name. Rows include H1B.FR.83.HXB2, H1A1.UG.85.U455.U455A, H1B.US.90.WEAU160.GHOSH, etc.

Table with 4 columns: Accession ID, Rev start, exon 1 end, exon 2 start, and protein name. Rows include MAC.US.x.239, H2A.DE.x.BEN, H2A.LT.x.AL1, H2A.SN.86.ST.J5P4.27, etc.

	Rev end	
H1B.FR.83.HXB2*	
H1A1.UG.85.U455 U455A*	
H1B.US.90.WEAU160 GHOSH*	
H1C.ET.86.ETH2220*	
H1D.CD.84.84Z085*	
H1F1.BE.93.VI850*	
H1G.SE.93.SE6165 G6165*	
H1H.CF.90.056*	
H1J.SE.93.SE9280 7887*	
H1K.CM.96.96CM MP35*	
H101.AE.TH.90.CM240*	
H102.AG.NG.x.IBNG*	
H10.BE.87.ANT70*	
H10.CM.91.MVP5180*	
H10.CM.94.BCF06*	
H10.CM.96.96CMA102*	
H10.FR.92.VAU*	
H10.SN.99.99SE MP1299*	
H10.US.99.99USTWLA*	
H10.US.x.I.24788*	
H1N.CM.02.DJ00131*	
H1N.CM.02.SJGddd*	
H1N.CM.04.04CM 1131_03*	
H1N.CM.06.U14296*	
H1N.CM.06.U14842*	
H1N.CM.95.YB36*	
H1P.CM.06.U14788*	
H1P.FR.06.RBF168*	
CPZ.CD.06.BF1167*	
CPZ.CD.90.ANT*	
CPZ.CM.01.SIVcpzCAM13*	
CPZ.CM.05.LB715*	
CPZ.GA.88.SIVcpzGAB2*	
CPZ.TZ.00.TAN1*	
CPZ.TZ.06.SIVcpzTAN13*	
CPZ.TZ.06.TAN5*	
CPZ.TZ.09.UG38*	
CPZ.US.85.US_Marilyn*	

	Rev end	
MAC.US.x.239*	
H2A.DE.x.BEN*	
H2A.PT.x.ALI*	
H2A.SN.86.ST JSP4_27*	
H2B.CI.x.EH0	RDHCNTQEDOTRG.....*	
H2B.GH.86.D205 ALT	SNHRDPOENQTRT.....*	
H2G.CI.92.Abt96*	
H2U.CI.07.071C TNP3*	
H2U.FR.96.12034*	
COL.CM.x.CGU1*	
COL.UG.10.BWC01*	
COL.UG.10.BWC07*	
DEB.CM.04.SIVdeb04CMPF3061*	
DEB.CM.99.CM40*	
DEB.CM.99.CM5*	
DEN.CD.x.CD1 CM0580407*	
DRL.x.x.FA0*	
GOR.CM.04.SIVgorCP684con*	
GOR.CM.07.SIVgor2139 287*	
GRV.ET.x.GRI 677 gri_1	KKSAARP.....*	
GSN.CM.99.CN166*	
GSN.CM.99.CN71	SSSCERHOTGTGKEMEIIOIREKWVSQ.....*	
LST.CD.88.SIVlhoest485	KN.....*	
LST.CD.88.SIVlhoest524	EH.....*	
LST.KE.x.lho7*	
MAC.US.x.EMBL 3*	
MND-1.GA.x.MNDGB1*	
MND-2.CM.98.CM16*	
MND-2.GA.x.M14*	
MND-2.x.x.5440*	
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	TSIKQIW.....*	
RCM.CM.00.SIVag1 00CM312*	
RCM.CM.02.SIVrcm 02CM8081*	
RCM.GA.x.SIVRCMGAB1*	
RCM.NG.x.NG411*	
SAB.SN.x.SAB1*	
SMM.CI.79.SIVsmCI2*	
SMM.LR.89.SIVsmLB1*	
SMM.SL.92.SIVsmSL92A*	
SMM.SL.92.SL92B*	
STM.US.89.STM 37_16*	
SUN.GA.98.L14	IVSEGKK.....*	
SVK.KE.x.KE51*	
SVK.KE.x.SVK173_COMGNM*	
TAL.CM.00.266*	
TAL.CM.01.8023*	
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.GM.05.Pbt 05GM_X02*	

116
124
117
108
117
117
117
117
117
124
124
103
104
104
103
104
104
104
104
107
105
105
105
105
105
105
103
105
126
119
124
111
124
106
96
106
111
106

108
104
108
108
151
154
101
96
99
109
108
127
93
86
83
85
102
105
105
148
125
121
174
149
149
109
96
101
101
100
109
121
125
115
119
108
126
158
108
105
94
108
99
104
103
104
100
100
151
85
78
126
112
104
90
90
90
85
136
127
149

Table with columns: Env start, signal peptide end_gp120 start, glycosylation NVT, glycosylation V1, CD4 binding glycosylation NSS, glycosylation NNTT, glycosylation V1. Rows include HIV strains like H1B_FR.83.HXB2, H1A1.UG.85.U455.U455A, etc.

Table with columns: Env start, signal peptide end_gp120 start, glycosylation NVT, glycosylation V1, CD4 binding glycosylation NSS, glycosylation NNTT, glycosylation V1. Rows include HIV strains like MAC.US.x.239, H2A.PT.x.ALI, H2A.SN.86.S2.J5P4.27, etc.

glycosylation NIS

glycosylation NCS

CD4 binding

glycosylation NGT

Table with columns for V1, glycosylation NDT, glycosylation NTS, glycosylation NGT, and V2. Lists protein sequences with their respective amino acid residues and positions.

Table with columns for V1, glycosylation NCS, glycosylation NDT, glycosylation NTS, glycosylation NGT, and V2. Lists protein sequences with their respective amino acid residues and positions.

glycosylation NVS

glycosylation NGS

glycosylation NFT

glycosylation NCT

V3 tip

glycosylation NNT

glycosylation NKT

CD4 binding

Table of HIV sequence alignments for glycosylation sites. Columns include sequence identifiers (e.g., H1B.FR.83.HXB2), amino acid sequences, and alignment positions (1-349). Annotations above the sequences indicate glycosylation sites: CD4 binding, glycosylation NNT, glycosylation NCT, V3 tip, glycosylation NKT, and CD4 binding.

Table of HIV sequence alignments for V3 tip region. Columns include sequence identifiers (e.g., MAC.US.x.239), amino acid sequences, and alignment positions (1-404). Annotations above the sequences indicate glycosylation sites: glycosylation NNT, glycosylation NCT, V3 tip, glycosylation NKT, and CD4 binding.

Table with columns for accession numbers (e.g., H1B.FR.83.HXB2), protein regions (fusion peptide, gp120 end, gp41 start, immunodominant region, glycosylation NAS, glycosylation NHT), and amino acid sequences. Includes a reference sequence (Q.....REKRAV...GIGALFLGLGAAGSTMGAAASMTLVQARQLLSGIVQQNNLL...LEQIWNHTTWMEWDREINNYTSLTHLSEESNQ00) and various other sequences.

Table with columns for accession numbers (e.g., MAC.US.x.239), protein regions (gp120 end, gp41 start), and amino acid sequences. Includes a reference sequence (GG...TS-N-G-FVL...AT-A-L-A-S-T-A...00-DVVKR-E-R...T-N-T-VT-I-K...AQ-NA-AFROV-H-T-PNAS) and various other sequences.

H1B.FR.83.HXB2
H1A1.UG.85.U455.U455A
H1B.US.90.WEAU160.GHOSH
H1C.ET.86.ETH2220
H1D.CD.84.84ZRO85
H1F1.BE.93.V17850
H1G.SE.93.SE6165.G6165
H1H.CF.90.056
H1J.SE.93.SE9280.7887
H1K.CM.96.96CM.MP35
H101.AE.TH.90.CM240
H102.AG.NG.x.IBNG
H10.BE.87.ANT70
H10.CM.91.MVP5180
H10.CM.94.BCF06
H10.CM.96.96CM102
H10.FR.92.VAU
H10.SN.99.99SE.MP1299
H10.US.99.99USTWLA
H10.US.x.I.24788
H1N.CM.02.DJ00131
H1N.CM.02.SJGddd
H1N.CM.04.04CM.1131_03
H1N.CM.06.U14296
H1N.CM.06.U14842
H1N.CM.95.YBF30
H1P.CM.06.U14788
H1P.FR.06.RBF168
CPZ.CD.06.BF1167
CPZ.CD.90.ANT
CPZ.CD.91.SIVcpzCAM13
CPZ.CM.05.BT15
CPZ.GA.88.SIVcpzGAB2
CPZ.TZ.00.TAN1
CPZ.TZ.06.SIVcpzTAN13
CPZ.TZ.06.TAN5
CPZ.TZ.09.UG38
CPZ.US.85.US_Marilyn

MAC.US.x.239
H2A.CD.x.BEN
H2A.PT.x.ALI
H2A.SN.86.ST.J5P4_27
H2B.CI.x.EHO
H2B.GH.86.D205.ALT
H2G.CI.92.Abt96
H2U.CI.97.071C.TNP3
H2U.FR.96.12034
COL.CM.x.CGU1
COL.UG.10.BWC01
COL.UG.10.BWC07
DEB.CM.04.SIVcb04CMF3061
DEB.CM.99.CM40
DEB.CM.99.CM5
DEN.CD.x.CD1.CM0580407
DRL.x.x.FA0
GOR.CM.84.SIVgorCP684con
GRI.CD.97.SIVgor2139.287
GRV.ET.x.GRI.677.gri_1
GSN.CM.99.CN166
GSN.CM.99.CN171
LST.CD.88.SIVlhoest485
LST.CD.98.SIVlhoest524
LST.KE.x.lho7
MAC.US.x.EMBL_3
MND.1.GA.x.MNDG81
MND.2.GA.x.CM16
MND.2.GA.x.M14
MND.2.x.x.5440
MNE.US.x.MNE627
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.SIVag1_00CM312
RCM.CM.02.SIVcm_02CM8081
RCM.GA.x.SIVRCMGAB1
RCM.NG.x.NG411
SAB.SN.x.SAB1
SHM.CI.79.SIVsmC12
SHM.LR.89.SIVsmL1B1
SHM.SL.92.SIVsmSL92A
SHM.SL.92.SL92B
STM.US.89.STM_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.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.GM.05.Pbt_05GM_X02

Env end
gp41 end
glycosylation NAT
cytoplasmic tail end
WNLLQYWSQELKNSAVSLLNATAIAVAEGTRVIEVV...OGACRAIRHIPRRIRQGLER...ILL...
---L-GR---I-IT-D-V-V---GWI---IG...TIG-LN-----A...
GS-V--G--K--IN--T--V-G---F-LT--RW--FCN-----A-A-Q...
G--TR--R--I--VD--T--DI--RR--K-VL--T-----A...
G--TR--R--I--F-T-V---I-L--R-G-VLN-----A-A...
---L-GR---I-DTV---NW--A--R--LN-T-----A...
G--ID--T---GI-VI--R-W-L-----F-S...
V-VW-G---I--T---I-IG--R-F-L-----A...
---V--T---G---I-IG--R-F-L-----A...
G--L--G--I--D---A-GW---W-L-----A...
---IS-V---IN--TI--V--NW--A--IG--RVG--N-----F-A...
AAVT--L-Q--T--DTL-V--NW-GI-AGI--RIGTG--N-----S...
GAVM--L---TN--DTI--VS--NW-GI-LGL--RIGOGFL-----A--V...
GAVT--L-RA--TN--DTI--V--GNW-SI-LGI--RIG-GFLN-----A-N...
NAVI--L--G--I--D---A-GW---W-L-----A...
KAI1--L--QT--TN--DTV--V--NW-SI-LGI--SIG-G-LN-----L...
IAVI--L-Q--T--DTI--V--NW-VTI-LGI--RIG-G-LN-----S...
AAVT--L--I--T--DTV--V--GSW-NI-SG--RIG-G-LN-----A--X...
AAVI--L-Q--T--TN-VTI--V--NW-SI-LG--RIG-G-WNV-----F-S...
-GTIIA-GK--D--I--TI--V---L-LA--RIG-G-L-----A...
-GI-A-GG-XRD--I--T--V---L-LA--RIG-G-L-----A-A-V...
-GX-A-GK--D--I--XTX---X-LA--RIG-G-L-V-----T-V...
-GTIIA-GK--D--I--XT---L-LA--IG-G-L-----A...
-GTIIA-XK--XD--I--T--V---L-LA--RIG-G-L-----V...
-GI-G--K--RD--T--V---L-LA--RIG-G-L-----S-I...
FAACG--T--X--T-X-TV-S--NW-Q--A-G--OIG-GFLN--L-----S...
FAACG--T--Q--T--DTV--S--GW-Q--I-G--OIG-GFLN--I-----S...
LH-ET-GLLRHCIVCIDTL-EFTGWM--G--AL--RV-VDI-----T--I-A-N...
N--STL--TIIRTEIKNIDRL--W-G-K-SILLAL--TIV-I-EV-----I-A-N...
RGI1--GR--RT--T--D--I--G--I-RA--TIV-I-GN-----T...
KGV1--GRX--V--DT-----A--REG-G-LN-T-----S...
KGIG--L--RT--I--DT--I--R-TI--A--TRIG-G-L-----A-E...
IEWSSNTYAS-RVLLIQADRL-NFTGWM--LI-G--VYIA-G-N-----L-A-N...
RDW-ANAYVIV-T-FIRN-DRL-NFT-WW-IL-GA--VNI-F-G-N-T-----I-A-N...
LEWSSNTYASIRA-LIQADRL-EFTGWM--I--A--VYIA-G-N-----I-A-N...
FMZ-SNTYSL-RT-LIQADRL-NFTGWM--L-A-A--RIG-G-L-----I-A-N...
-GVIA--AR--V--T--DT-----I-LT--RRLFLG-I-----S

Env end
gp41 end
LTY--GWSYFHEAVQAVWRSATETL-GAWGDLW-TL--RRGG-W-LA-----L-L-T...
AAD-GCEWIOEAFQAFART-RET-L-GAWGDLW-LA--RRIG-G-LAV-----A-I-A...
AAY--GCEWIOEAFRAFARIARETLNTNRDLWGA--WVG-R-LAV-----A-I-A...
TAY--GGEWIOEAFRAFAR-GETLTNAWRGFWGTL--GOIG-G-LAV-----A-I-A...
Y-R-GISWFOEAIQAAAAR-AGETL-SAARTSWG-L-RR-AGE-IA-----A-L-A...
VAY--GISWFOEALRAAAR-GETL-SAGETLW-L-RR-A--IA-----L-L-T...
AAVFS-GFRW-DEACTATR-AGETLTSWRALWKTL--GRVG-G-LA-----L-L-T...
GAVX--GC-WFEALQAGWK-XRETLSANRWLW-TL--GRVG-G-LA-----F-L-A...
IAY--GW-YF-EAFOAFGK-ARETLSRTGRELW-TL--GRVG-WL-A-----F-L-A...
DKRRA--T-RYL--LVPRGPSLQIL-TLQ-WLRSA--ARGW-RAPEYL-GWIDYRPO-GPA...
...-W-L-LVPRCPDLSTP-TLS-WLRSA--TRRW-RASEYL-GWLFDRPE-GPA...
...-MPL-A-WRL--LVPRGSANLTPOTLQ-WLRSA--TS-G-RALESCG-WIR-DPO-GPA...
-GY--GWK-F-DTI-WAGGEGOTFORWAVALQ--A-A-L-F-N...
GGY--GWK-F-DTIWAGGELQWLOGIAEMALQGL--VR-GG-LRV-A-----A-L-N...
-GYC--GWK-F-DTIWAGGEIWEWLOGIAQVA-OGL--VWGG-NLLA-A-----A-L-L-N...
YF--GL--QTGLRD-GTSAIQGRATAEV-LAAL--TR-A-EVVA-----I-V-N...
-QEAGRGC-TLLQEAQRIWREGROLGLSSARAL--RALAQEVAA-----A-V-LFN...
FAEIA--G--Q--I--DCV-VFT-DW--Q-AIA--RIG-G-LN-----S...
LST.CD.98.SIVlhoest524
LST.KE.x.lho7
MAC.US.x.EMBL_3
MND.1.GA.x.MNDG81
MND.2.GA.x.CM16
MND.2.GA.x.M14
MND.2.x.x.5440
MNE.US.x.MNE627
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.SIVag1_00CM312
RCM.CM.02.SIVcm_02CM8081
RCM.GA.x.SIVRCMGAB1
RCM.NG.x.NG411
SAB.SN.x.SAB1
SHM.CI.79.SIVsmC12
SHM.LR.89.SIVsmL1B1
SHM.SL.92.SIVsmSL92A
SHM.SL.92.SL92B
STM.US.89.STM_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.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.GM.05.Pbt_05GM_X02

Table with 6 columns: protein ID, myristoylation, Nef start, acidic cluster, phosphorylation, poly-P helix. Lists various HIV-1 proteins like H1B, H1A1, H1A2, etc., with their corresponding amino acid sequences.

Table with 6 columns: protein ID, Nef start, R17Y mutation, premature stop in Mac239. Lists HIV-1 proteins like MAC.US.x.239, H2A.DE.x.BEN, etc., with their corresponding amino acid sequences and mutations.

Table with columns for HIV sequence identifiers (e.g., H1B.FR.83.HXB2), amino acid sequences, and accession numbers (205-210). The sequences are aligned and include a 'normal Nef end' label.

Table with columns for HIV-1 similarity identifiers (e.g., MAC.US.x.239), amino acid sequences, and accession numbers (263-321). The sequences are aligned and include a 'normal Nef end' label.