

IV

PLV Complete Genomes

Contents

IV-1	Introduction	221
IV-2	Sequences	222
IV-3	Alignments	224

Gag-Pol precursor polyproteins have been experimentally determined for HIV-1 and at least one strain of HIV-2, the study of analogous cleavages in SIV polyproteins have not been published. Two representative genomes have been translated; HIV-1 subtype B strain HXB2, and SMM-239. The translations are provided as a visual aid for finding landmarks in the genomes.

IV-1 Introduction

Complete genomes of HIV-1, HIV-2, and SIV sequences are represented in the alignment. For the selection, one of each of the major subtypes and a few CRF s of the HIV-1 M group was chosen, plus one N group and 2 O group sequences. Most of the SIV-CPZ s were included, including new sequences from wild caught chimpanzees. Three HIV-2 A group and 2 HIV-2 B group were included along with the single genomes sequenced to date for HIV-2 G and HIV-2 U. For the rest of the SIVs, all complete genomes were included. The principle in deciding which sequences to include in this alignment was to provide a representative sampling of all primate lentiviruses which have had complete or nearly complete genomes sequenced. Many HIV-1, HIV-2, SIVsmm and SIVmac genomes were not included, but they can be found in the HIV-1/SIVcpz and HIV-2/SIVsmm alignments. Also, many shorter sequences from PLVs are included in the PLV protein sequence alignments later in this compendium. Because of the complicated phylogeny, most likely muddled by recombination and many cross species transmissions, we do not show a tree for the PLV alignment this year. Please refer to the Sequence Compendium of 2005 for a more comprehensive evaluation of the PLV phylogeny.

The alignment was based on the previous version published last year, by a hidden Markov model using HMMER and manual editing using the programs BioEdit and Se-Al. It is codon-aligned, meaning that the correct translation reading frame has been maintained as much as possible; in the case of overlapping frames, obviously one will not translate correctly. We have tried to “reset” the alignment so that each new gene starts in the first frame; this means in some places empty columns have been inserted. The HIV-2/SIVsmm *vpx* gene is postulated to be a duplication of the *vpr* gene and thus there may be two alternative alignments of this region of the genome, as there are for the duplicated stem-loops of the TAR element.

Like last year, the annotation is based on known protein coding regions in HIV-1 and on annotations found in SIV sequence database entries. The protein cleavage sites that create Gag p17, Gag p24 and other mature peptides from the Gag and

IV-2 Sequences

Primate lentivirus (PLV) Sequences in the complete genome alignments.

Name	Accession	Country	Author	Reference
H1B.FR.83.HXB2	K03455	France	Wong-Staal, F	<i>Nature</i> 313 (6000); 277-84 (1985)
H101_AE.TH.90.CM240	U54771	Thailand	Laukkanen, T	<i>J Virol</i> 70 (9); 5935-43 (1996)
H102_AG.NG.x.IBNG	L39106	Nigeria	Howard, TM	<i>ARHR</i> 10 (12); 1755-7 (1994)
H103_AB.RU.97.KAL153_2	AF193276	Russia	Liitsola, K	<i>AIDS</i> 12 (14); 1907-19 (1998)
H104_cpx.CY.94.CY032	AF049337	Cyprus	Gao, F	<i>J Virol</i> 72 (12); 10234-41 (1998)
H1A1.UG.85.U455	M62320	Uganda	Oram, JD	<i>ARHR</i> 6 (9); 1073-8 (1990)
H1B.US.90.WEAU160	U21135	United States	Tozser, J	<i>FEBS Lett</i> 281 (1-2); 77-80 (1991)
H1C.ET.86.ETH2220	U46016	Ethiopia	Salminen, MO	<i>ARHR</i> 12 (14); 1329-39 (1996)
H1D.CD.84.84ZR085	U88822	D.R.C.	Gao, F	<i>J Virol</i> 72 (7); 5680-98 (1998)
H1F1.BE.93.VI850	AF077336	Belgium	Laukkanen, T	<i>Virology</i> 269 (1); 95-104 (2000)
H1G.SE.93.SE6165	AF061642	Sweden	Laukkanen, T	<i>Virology</i> 247 (1); 22-31 (1998)
H1H.CF.90.056	AF005496	C.A.R.	Murphy, E	<i>ARHR</i> 9 (10); 997-1006 (1993)
H1J.SE.93.SE7887	AF082394	Sweden	Laukkanen, T	<i>ARHR</i> 15 (3); 293-7 (1999)
H1K.CM.96.MP535	AJ249239	Cameroon	Peeters, M	<i>ARHR</i> 16 (2); 139-51 (2000)
H1N.CM.95.YBF30	AJ006022	Cameroon	Simon, F	<i>Nat Med</i> 4 (9); 1032-7 (1998)
H1O.BE.87.ANT70	L20587	Belgium	Vanden Haesevelde, M	<i>J Virol</i> 68 (3); 1586-96 (1994)
H1O.CM.91.MVP5180	L20571	Cameroon	Gurtler, LG	<i>J Virol</i> 68 (3); 1581-5 (1994)
CPZ.CD.90.ANT	U42720	D.R.C.	Vanden Haesevelde, MM	<i>Virology</i> 221 (2); 346-50 (1996)
CPZ.CM.01.SIVcpzCAM13	AY169968	Cameroon	Nerrienet, E	<i>J Virol</i> 79 (2); 1312-9 (2005)
CPZ.CM.05.SIVcpzEK505	DQ373065	Cameroon	Keele, BF	<i>Science</i> 2006 May 25;
CPZ.CM.05.SIVcpzLB7	DQ373064	Cameroon	Keele, BF	<i>Science</i> 2006 May 25;
CPZ.CM.05.SIVcpzMB66	DQ373063	Cameroon	Keele, BF	<i>Science</i> 2006 May 25;
CPZ.CM.05.SIVcpzMT145	DQ373066	Cameroon	Keele, BF	<i>Science</i> 2006 May 25;
CPZ.CM.98.CAM3	AF115393	Cameroon	Corbet, S	<i>J Virol</i> 74 (1); 529-34 (2000)
CPZ.CM.98.CAM5	AJ271369	Cameroon	Muller-Trutwin, MC	<i>J Med Primatol</i> 29 (3-4); 166-72 (2000)
CPZ.GA.88.GAB1	X52154	Gabon	Huet, T	<i>Nature</i> 345 (6273); 356-9 (1990)
CPZ.GA.88.GAB2	AF382828	Gabon	Gao, F	<i>ARHR</i> 20 (12); 1377-81 (2004)
CPZ.TZ.01.TAN1	AF447763	Tanzania	Santiago, ML	<i>J Virol</i> 77 (3);2233-2242 (2003)
CPZ.US.85.CPZUS	AF103818	United States	Gao, F	<i>Nature</i> 397 (6718); 436-41 (1999)
H2A.DE.x.BEN	M30502	Germany	Kirchhoff, F	<i>Virology</i> 177 (1); 305-11 (1990)
H2A.GW.x.ALI	AF082339	Guinea-Bissau	Azevedo-Pereira, JM	Unpublished (1998)
H2A.SN.x.ST	M31113	Senegal	Kumar, P	<i>J Virol</i> 64 (2); 890-901 (1990)
H2B.CI.x.EHO	U27200	Cote D'Ivoire	Rey-Cuille, MA	<i>Virology</i> 202 (1); 471-6 (1994)
H2B.GH.86.D205	X61240	Ghana	Kreutz, R	<i>ARHR</i> 8 (9); 1619-29 (1992)
H2G.CI.x.ABT96	AF208027	Cote D'Ivoire	Brennan, CA	<i>ARHR</i> 13 (5); 401-4 (1997)
H2U.FR.96.12034	AY530889	France	Damond, F	<i>ARHR</i> 20 (6); 666-72 (2004)
MAC.US.x.239	M33262	United States	Kestler, H	<i>Science</i> 248 (4959); 1109-12 (1990)
MAC.US.x.251_1A11	M76764	United States	Planelles, V	<i>ARHR</i> 7 (11); 889-98 (1991)
MAC.US.x.251_BK28	M19499	United States	Franchini, G	<i>Nature</i> 328 (6130); 539-43 (1987)
MAC.US.x.EMBL_3	Y00295	United States	Franchini, G	<i>Nature</i> 328 (6130):539-543 (1987)
SMM.SL.92.SL92B	AF334679	Sierra Leone	Chen, Z	<i>J Virol</i> 70 (6); 3617-27 (1996)
SMM.US.x.H9	M80194	United States	Courgnaud, V	<i>J Virol</i> 66 (1); 414-9 (1992)
SMM.US.x.PGM53	AF077017	United States	Novembre, FJ	<i>J Virol</i> 72 (11); 8841-51 (1998)

PLV Complete Genomes

Name	Accession	Country	Author	Reference
SMM.US.x.SIVsmH635F_L3	DQ201172	United States	Kuwata, T	<i>J Virol</i> 80 (3); 1463-75 (2006)
STM.US.x.STM	M83293	United States	Novembre, FJ	<i>Virology</i> 186 (2); 783-7 (1992)
SAB.SN.x.SAB1C	U04005	Senegal	Jin, MJ	<i>EMBO J</i> 13 (12); 2935-47 (1994)
TAN.UG.x.TAN1	U58991	Uganda	Stivahtis, GL	<i>Virology</i> 228 (2); 394-9 (1997)
VER.DE.x.AGM3	M30931	Germany	Baier, M	<i>Virology</i> 176 (1); 216-21 (1990)
VER.KE.x.9063	L40990	Kenya	Hirsch, VM	<i>J Virol</i> 69 (2); 955-67 (1995)
VER.KE.x.AGM155	M29975	Kenya	Johnson, PR	<i>J Virol</i> 64 (3); 1086-92 (1990)
VER.KE.x.TYO1	DJ048201	Kenya	Omori, T	Patent: WO 2007049749-A 13 03-MAY-2007; Dनावेक CORPORATION
COL.CM.x.CGU1	AF301156	Cameroon	Courgnaud, V	<i>J Virol</i> 75 (2); 857-66 (2001)
GRV.ET.x.GRI_677	M66437	Ethiopia	Fomsgaard, A	<i>Virology</i> 182 (1); 397-402 (1991)
MND_2.CM.98.CM16	AF367411	Cameroon	Takehisa, J	<i>ARHR</i> 17 (12); 1143-54 (2001)
DRL.x.x.FAO	AY159321		Hu, J	<i>J Virol</i> 77 (8):4867-4880 (2003)
RCM.GA.x.GAB1	AF382829	Gabon	Gao, F	<i>Science</i> 300 (5626); 1713 (2003)
RCM.NG.x.NG411	AF349680	Nigeria	Beer, BE	<i>J Virol</i> 75 (24); 12014-27 (2001)
MND_1.GA.x.MNDGB1	M27470	Gabon	Tsujimoto, H	<i>Nature</i> 341 (6242); 539-41 (1989)
MND_2.GA.x.M14	AF328295	Gabon	Souquiere, S	<i>J Virol</i> 75 (15); 7086-96 (2001)
MND_2.x.x.5440	AY159322		Hu, J	<i>J Virol</i> 77 (8):4867-4880 (2003)
MNE.US.x.MNE027	U79412	United States	Kimata, JT	<i>J Virol</i> 72 (1); 245-56 (1998)
LST.CD.88.447	AF188114	D.R.C.	Beer, BE	<i>J Virol</i> 74 (8); 3892-8 (2000)
LST.CD.88.485	AF188115	D.R.C.	Beer, BE	<i>J Virol</i> 74 (8); 3892-8 (2000)
LST.CD.88.524	AF188116	D.R.C.	Beer, BE	<i>J Virol</i> 74 (8); 3892-8 (2000)
LST.KE.x.lho7	AF075269	Kenya	Hirsch, VM	<i>J Virol</i> 73 (2); 1036-45 (1999)
GSN.CM.99.CN166	AF468659	Cameroon	Courgnaud, V	<i>J Virol</i> 76 (16); 8298-309 (2002)
GSN.CM.99.CN71	AF468658	Cameroon	Courgnaud, V	<i>J Virol</i> 76 (16); 8298-309 (2002)
MON.CM.99.L1	AY340701	Cameroon	Courgnaud, V	<i>J Virol</i> 77 (23):12523-12534 (2003)
MON.NG.x.NG1	AJ549283	Nigeria	Barlow, KL	<i>J Virol</i> 77 (12); 6879-88 (2003)
MUS_1.CM.01.1085	AY340700	Cameroon	Courgnaud, V	<i>J Virol</i> 77 (23):12523-12534 (2003)
MUS_1.CM.01.CM1239	EF070330	Cameroon	Aghokeng, AF	(<i>er</i>) <i>Virology</i> 360 (2):407-418 (2007)
MUS_2.CM.01.CM1246	EF070329	Cameroon	Aghokeng, AF	(<i>er</i>) <i>Virology</i> 360 (2):407-418 (2007)
MUS_2.CM.01.CM2500	EF070331	Cameroon	Aghokeng, AF	(<i>er</i>) <i>Virology</i> 360 (2):407-418 (2007)
DEN.CD.x.CD1	AJ580407	D.R.C.	Dazza, MC	<i>J Virol</i> 79 (13); 8560-71 (2005)
DEB.CM.99.CM40	AY523865	Cameroon	Bibollet-Ruche, F	<i>J Virol</i> 78 (14); 7748-62 (2004)
DEB.CM.99.CM5	AY523866	Cameroon	Bibollet-Ruche, F	<i>J Virol</i> 78 (14); 7748-62 (2004)
TAL.CM.00.266	AY655744	Cameroon	Liegeois, F	<i>Virology</i> 349 (1):55-65 (2006)
TAL.CM.01.8023	AM182197	Cameroon	Liegeois, F	<i>Virology</i> 349 (1); 55-65 (2006)
SUN.GA.98.L14	AF131870	Gabon	Beer, BE	<i>J Virol</i> 73 (9); 7734-44 (1999)
SYK.KE.x.KE51	AY523867	Kenya	Bibollet-Ruche, F	<i>J Virol</i> 78 (14); 7748-62 (2004)
SYK.KE.x.SYK173	L06042	Kenya	Hirsch, VM	<i>J Virol</i> 67 (3); 1517-28 (1993)

	5' LTR U3 start	
H1B.FR.83.HXB2TGGAAAGGGCTAATTCACCTCCCAACGAAGACAAGATATCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGAGAAGTT	169
H101.AE.TH.90.CM240	0
H102.AC.NG.x.IB.NG	0
H103.AB.RU.97.KALI.153.2	0
H104.cpx.CY.94.CY032	0
H1A1.UG.85.U455	0
H1B.US.90.WEAU160T.....AA.....G.....T.....T.....G.....CT.....C.....TG.....T.....G.....A.....G.....	169
H1C.ET.86.ETH220	0
H1D.CD.84.84Z.R085	0
H1F1.BE.93.VI850	0
H1G.SE.93.SE6165	0
H1H.CF.90.056	0
H1J.SE.93.SE7887	0
H1K.CM.96.MP535	0
H1N.CM.95.YBF30	0
H1O.BE.87.AN170	0
H1O.CM.91.MVP5180C.....T.....T.....T.....TAA.....GC.....A.....G.....T.....G.....G.....TA.....T.....T.....G.....A.....T.....G.....G.....TGT.....G.....A.....A.....C.....G.....TC.....A.....TG.....TT.....A.....G.....GTCAGA.....AG.....GC	170
CPZ.CD.90.ANTC.....T.....T.....T.....TAA.....GC.....A.....G.....T.....G.....G.....TA.....T.....T.....G.....A.....T.....G.....G.....TGT.....G.....A.....ACCT.....TC.....A.....TG.....TT.....A.....G.....GTCAG.....AG.....GC	170
CPZ.CM.01.SIVcpzCAM13	0
CPZ.CM.05.SIVcpzEK505	0
CPZ.CM.05.SIVcpzLB7	0
CPZ.CM.05.SIVcpzMB66	0
CPZ.CM.05.SIVcpzMT145	0
CPZ.CM.98.CAM3	0
CPZ.CM.98.CAB1	0
CPZ.GA.88.GAB1T.....G.....T.....AGGA.....G.....C.....C.....G.....T.....T.....C.....G.....A.....A.....A.....CA.....TC.....CTGT.....T.....A.....G.....CC.....GACAGAG.....C.....G.....	169
CPZ.GA.88.GAB2	0
CPZ.TZ.01.TAN1	0
CPZ.US.85.CPZUST.....G.....T.....T.....AG.....A.....G.....A.....C.....G.....T.....T.....AAT.....A.....C.....G.....T.....C.....T.....AG.....A.....C.....A.....A.....A.....GC.....TT.....T.....C.....TC.....CACAGA.....G.....AG.....	169
H2A.DE.x.BENC.....A.....GT.....T.....AGTAGGA.....G.....TAGA.....T.....A.....CA.....A.....ACC.....AG.....AA.....AGAGG.....GATAA.....A.....A.....G.....T.....T.....T.....AT.....AG.....A.....G.....C.....A.....TA.....C.....G.....CTG.....GG.....T.....A.....A.....CTC.....C	163
H2A.CW.x.ALIA.....GT.....T.....AGTG.....GA.....TAGA.....T.....A.....CA.....A.....ACT.....AG.....AA.....GGA.....GATAA.....T.....A.....G.....T.....T.....AT.....AG.....A.....G.....C.....GA.....TT.....G.....CTG.....GG.....A.....T.....GTCCC	163
H2A.SN.x.ALI	0
H2B.CL.x.EHOA.....TT.....CT.....TAGTG.....GA.....G.....TA.....A.....A.....CACAA.....CT.....AG.....AA.....TGA.....G.....ATTG.....GT.....GA.....G.....A.....T.....AT.....AG.....A.....CAA.....TT.....C.....CTA.....GG.....G.....A.....AA.....T.....T.....A.....	163
H2B.GH.86.D205A.....GTA.....T.....TAGTG.....GA.....CAGA.....AT.....A.....CACAA.....ATT.....TG.....GA.....TGA.....G.....ATTG.....GT.....GC.....G.....A.....T.....T.....AT.....A.....GC.....CAAATA.....T.....CTG.....GG.....G.....A.....GT.....CC	163
H2G.CL.x.AB196	0
H2U.FR.96.12034	0
	5' LTR U3 start	
MAC.US.x.239A.....TTA.....T.....AGTGC.....A.....TAGA.....T.....A.....CA.....A.....ACT.....AG.....AA.....GGA.....G.....AT.....A.....A.....A.....G.....G.....T.....CT.....A.....A.....T.....C.....AA.....A.....C.....CTA.....GG.....AT.....C.....T.....AA.....T.....GTA.....C	163
MAC.US.x.251.1A11A.....TTA.....T.....TAGTGC.....A.....TAGA.....T.....A.....CA.....A.....ACT.....AG.....AA.....GGA.....G.....AT.....A.....A.....A.....G.....G.....T.....CT.....A.....A.....T.....C.....AA.....A.....C.....CTA.....GG.....AT.....C.....T.....AA.....T.....GTA.....C	163
MAC.US.x.251.BK28A.....TTA.....T.....AGTGC.....A.....TAGA.....T.....A.....CA.....A.....ACT.....AG.....AA.....GGA.....G.....AT.....A.....A.....A.....G.....G.....T.....CT.....A.....A.....T.....C.....AA.....A.....C.....CTA.....GG.....AT.....C.....T.....AA.....T.....GTA.....C	163
MAC.US.x.EMB1.3	0
SMM.SL.92.SL92B	0
SMM.US.x.H9	0
SMM.US.x.PGM53	0
SMM.US.x.SIVsmH635F_L3A.....TTA.....T.....AATG.....GAA.....TAGA.....AT.....A.....A.....AC.....GG.....AA.....GGA.....G.....AATAA.....A.....A.....G.....A.....T.....T.....T.....A.....CT.....C.....TA.....TT.....C.....CTA.....GG.....AT.....G.....C.....AAGT.....GTC.....C	89
STM.US.x.8TMA.....TTA.....T.....AGTG.....	22
SAB.SN.x.SAB1CA.....TTA.....T.....AGTG.....A.....G.....A.....GA.....G.....T.....G.....C.....ATGCTCTTA.....TGAGTGG.....ATAG.....AGA.....GA.....G.....G.....A.....T.....T.....T.....GAT.....T.....C.....G.....C.....AA.....TG.....TT.....A.....G.....T.....G.....C.....CT.....C	163
TAN.UG.x.TAN1G.....CA.....TA.....TGG.....A.....C.....AAG.....G.....GC.....A.....G.....A.....C.....ATGCTCTTA.....TGAGTGG.....ATAG.....AGA.....GA.....G.....G.....A.....T.....T.....GAT.....T.....C.....G.....C.....AA.....TG.....C.....CT.....TT.....TTG.....T.....G.....G.....A.....T.....GTC.....C	163
VER.DE.x.AGM3	0
VER.KE.x.9063T.....A.....TTA.....T.....G.....A.....G.....GA.....TAGA.....CA.....C.....ATGCTCTTA.....TGA.....TGG.....AAT.....A.....TGA.....G.....A.....CGC.....GGT.....A.....A.....CG.....A.....C.....GCTG.....C.....TT.....T.....G.....A.....T.....CA	163
VER.KE.x.AGM155T.....A.....TTA.....T.....G.....TA.....G.....A.....TA.....G.....A.....ATGC.....CT.....A.....TGA.....TGG.....AAT.....A.....TGA.....G.....A.....CGCA.....GGT.....AA.....A.....T.....A.....A.....C.....GAG.....TG.....C.....TC.....T.....G.....G.....CC.....CT.....CA	163
VER.KE.x.TY01T.....A.....TTA.....T.....G.....TA.....G.....A.....TA.....G.....GA.....ATGCCTCTTA.....TGA.....TGG.....GATAA.....TGA.....G.....A.....TGC.....GGT.....GAAG.....A.....A.....A.....TC.....TAAATG.....G.....TC.....TT.....G.....G.....C.....TTACA	163
COL.CM.x.CGU1	0
GRV.ET.x.GRI.677T.....A.....TA.....T.....TG.....A.....G.....A.....G.....GA.....T.....ATGC.....TGA.....GAGTGG.....AATAA.....AGA.....G.....AGCT.....T.....C.....G.....A.....G.....C.....GAGAGT.....C.....TC.....TT.....C.....G.....C.....CT.....CA	163
MND.2.CM.98.CM16	0
DRL.x.x.FAO	0
RCM.GA.x.GAB1	0
RCM.NG.x.NG411	0
MND.1.GA.x.MNDGB1	0
MND.2.GA.x.M14	0
MND.2.x.x.5440	0
MNE.US.x.MNE027	0
LST.CD.88.447	0
LST.CD.88.485	0
LST.CD.88.524	0
LST.KE.x.lho7TA.....TGG.....GA.....G.....C.....T.....AAT.....C.....ACTGTG.....GA.....TGAGTGG.....TCT.....A.....TGA.....C.....TCAT.....GAT.....A.....A.....CA.....C.....T.....A.....TG.....GGC.....T.....A.....G.....CT.....TGTG.....TGA	166
GSN.CM.99.CN166	0
GSN.CM.99.CN71	0
MON.CM.99.L1	0
MON.NG.x.NG1	0
MUS.1.CM.01.1085	0
MUS.1.CM.01.CM1239	0
MUS.2.CM.01.CM1246	0
MUS.2.CM.01.CM2500	0
DEN.CD.x.CD1	0
DEB.CM.99.CM40	0
DEB.CM.99.CM5	0
TAL.CM.00.266	0
TAL.CM.01.8023CTGGT.....AG.....ACC.....G.....A.....T.....G.....T.....ATGCAG.....AA.....TGAGTGG.....T.....A.....AA.....GC.....G.....AG.....T.....CAA.....T.....AG.....G.....AA.....G.....C.....CTA.....GG.....C.....CACC.....ATTG.....TGA	166
SYK.NG.98.L14	0
SYK.KE.x.KF51	0
SYK.KE.x.SYK173A.....C.....C.....G.....C.....GT.....A.....T.....TA.....GGG.....A.....T.....AA.....TAGA	60

H1B.FR.83.HXB2
H101.AE.TH.90.CM240
H102.AC.NG.x.IBNG
H103.AB.RU.97.KALI.153.2
H104.cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WEAU160
H1C.ET.86.ETH220
H1D.CD.84.847R085
H1F1.BE.93.VI850
H1G.SE.93.SE6165
H1H.CF.90.056
H1J.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YB30
H1O.BE.87.AN170
H1O.CM.91.MVP5180
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzLB7
CPZ.CM.05.SIVcpzMB66
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CAM3
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.GAB2
CPZ.TZ.01.TAN1
CPZ.US.85.CPZUS
H2A.DE.x.BEN
H2A.GW.x.ALI
H2A.SN.x.
H2B.CL.x.EHO
H2B.GH.86.D205
H2G.CL.x.AB196
H2U.FR.96.12034

AGAAAGAA...GCCAACAAAGGAGAGAACCACCGCTGTTACACCTGTGAGC...CTGCATGGAATGGATGACCCGGAGAGAGAAGTGTAGAGTGGAGTTTGACAGCCGCCATAGCATTTATCACATGCCCCGAGAGCTGCATCCGGAGTACTTCAAGA
...T-TG...A...G...AT...A...A...AT...A...AAA...G-A...T-AG
...-AG-CTA-GAG-T-C-T-T---GGG-T-ATC-CC-G--T-A-CAT-T...GCC---T-T-A-TA-AC-T-A--A-AC-GAT--A---T-ATCT--GCAACACC-TG-T-TATGATAACT--C-A--CT--C-
...GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC---T-A-CTT-T...AAT---GCT--G-TG-AC-CG-G--GA-AC-A-A--CA---T-ATCAT--GC-AAACA-T-A---TG-AAA--C-A--CT---CCC-
...C---TG---T---A-T-C---G-T-CA-AT-TCAGCATGGGATGGAAG--AGGAC...A-G-C-G-TC-C-C---A-G-G-C-AAGA-T-T-A-A-AA---A-AG
...GCCG---T---C-A-TAC-C-G---CA-T--CAGCATGGAATGGAAG--A-CAT...G-C---C-ATC-CA---CTGAA-G-TCGGAGA--GA-TAA---A---CG-
...C---G---AG-GG--AT--GC-A-T--AG---A-CACAACAAGCAGACATGATG--AGCAT...G-G-GACA---T--CA---TC-ATG-G-C-A-A-CT--A---TTCACT--T-C-A--AG-G--TGG-C
...GGG-GG-CACT--CTCA-T-C-C---A--ACAAACAAGCAGCATGATG-CACGCAT...G-G-GACA---T--A---CCTAAG-G-TCA-G-T--AA---TTTTATT--A-C-A--AG-G--TGG-T
...C---C-AG-GG-C-AG--A-C-CATT-TC--G-G-T-A-CACAGACTCCGCATGGGATG-CCCCAC...GA--GACCC-T-TC-CA---TC-T---A-G-CT-TG---TTCAGCAG-TTC--A--G-G--TGG-T
...C-CGACCCGAG-GG-G-AG--A-C-CATT-C-AA-G---G-CACAGATCTCCTCATGGGATG-CATCCAT...G-G-GACTC-TATC---TTC-T---G--A-G-T-TG---TTTTCA-TAG-TT--A-AG-G--GG-T

MAC.US.x.239
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMB1_3
SMM.SL.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F_L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.UG.x.TAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TY01
COL.CM.x.CGU1
GRV.ET.x.GRI_677
MND_2.CM.98.CM16
DRL.x.x.FAO
RCM.GA.x.GAB1
RCM.NG.x.NG411
MND_1.GA.x.MNDGB1
MND_2.GA.x.M14
MND_2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.lho7
GSN.CM.99.CN166
GSN.CM.99.CN71
MON.CM.99.L1
MON.NG.x.NG1
MUS_1.CM.01.1085
MUS_1.CM.01.CM1239
MUS_2.CM.01.CM1246
MUS_2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.14
SYK.KE.x.KE51
SYK.KE.x.SYK173

...T-G...-AC-GG-G-AT--G-GCATTAT-AA-G-T-A-CTCAAACCTCCAGTGGGATG-CCCTTG...G---G-TC--CA--A---TCCAAC-T-G-C-ACAC-T-TGA--ATAT-TTAGAT-C-A--AG-G--TGG-
...T-G...-AC-GG-G-AT--G-GCATTAT-AA-G-T-A-CTCAAACCTCCAGTGGGATG-CCCTTG...G---G-TCC-CA--A---CCAAC-T-G-C-ACAC-T-TGA--ATAT-TTAGAT-C-A--AG-G--TGG-
...T-G...-AC-GG-G-AT--GGCATTAT-AA-G-G-A-CTCAAACCTCCAGTGGGATG-CCCTTG...G---G-TC-C--A---TCCAAC-T-G-C-ACAC-T-TGA--ATAT-CTAGAT-C-A--AG-G--TGG-
...-T---TC-AG--AC---CACATTA--AG-G-T-A-CTCAGATATCCAGTGGGATG-CCCTG...G---G-TC-G-CA--A---TTCA-AAT--C-ACAGAT-TGA--TTTTATTAGA-C-A--AG-G--TGGT-
...-C-GG--AC---CACATT-C-G-G-T-A-CACAGACTATCAGTGGGATG-CCCTG...G---G-AC-G-CA--A---TCCAGAA-T-A-AGCT-T-A--ATTTATTAA-T-C-A--AG-G--TGGT-
...G-G---T-G-TA-T---CA-T-C-C-G-T---CTCAAGTGGCTATGAGGATG-C-CCTG...A--GACCC-G-T--A---CCA-TA---T-GACT--GTT-CT-GGAGGCT-C
CC-G---C-GG-T-AG---GACA-T-T--C-C---CCAGATTTAATGGGAAGCG-CCCTG...A--GACA-G-T---AA--A-TCCA-T---TT-A-T-CAACCC-...G-C-
T---G...-AG-G-CCT-T---GACATTA---GCG-T---CACAATAAAA...GA-G-CCCCGAT-GCAT-AGTC-TG---GACC-G-TC--A---CCTATGT---T-A-T--GACCCAA--AGAC--T
T-G---AG-A-C-T-T-A-GGCATT---G---A-CACAACCTGCAT...GAAG-CCCTGAT-G-ATAAATC-TG---A-A-G-CA--A---TCCAATGT-G-T---TT-A-T-CGACCCC-CAAGGG--T
T-G---AC-A-C-T-T-A-GCATT--C-AG-C-T-A-C-CAGATGGGA...GAAG-CC-GAT-G-AT-AGCC-TG---GA-C-G-T---A---TCCTATGT-G---ATA-A-T-CGACCCCA-TCGGG--T
T-G-G---ACG--CT-T---GACA-T-TC--A-G-T-A-CACAGATGGG...GAAG-CCTGAT-GAATAGATC-TG---C-G-TC--A---CCGAAGT-G-G---T-GA-T-C-GCCC...G-C-
GT-TA---ATCC-TC--AT---CACAAT--C-T-G--TT-AAGCCAGCTGGGGTCTGGAAG-CCCTG...G---GAG-C-TATC--CAC-CA-TCCAATGT-G---T-GACTT-ATA-CCCT-AAA--C
...CAGG...ATCCT--TCATCCTTGTCAAGC-C-AC-G--TT-CAGCCAGCAAGGGTAAATGAAAG-CCCTTG...G---GAGAC-TATC--CA---TCC-ACT--G-C...T-TGATTT-AG--CCATACAG--C
G-G-TGC...TTGG-GT-T-AG---T-ACT-C-CC-C-T-A-CCTCT...GGGACAGGGGA-TTC-A-CAT...G-G-GAACCC-CAT--AC-CA-CC-A--G-C...

TCF-1 alpha

TCF-1 alpha

	TCF-1 alpha		NF-k-B-II	
H1B.FR.83.HXB2	ACTGC.....TGACA.....		TCGAGCTTGCT...ACA...AGGGACTTCCG.....	360
H101.AE.TH.90.CM240				0
H102.AC.NG.1.BNG				0
H103.AB.RU.97.KALI53.2				0
H104.cpx.CY.94.CY032				0
H1A1.UG.85.U455				0
H1B.US.90.WEAU160			-T-T.....A.....	360
H1C.ET.86.ETH220				0
H1D.CD.84.847R085				0
H1F1.BE.93.VI850				0
H1G.SE.93.SE6165				0
H1H.CF.90.056				0
H1J.SE.93.SE7887				0
H1K.CM.96.MP535				0
H1N.CM.95.YBF30				0
H1O.BE.87.ANT170	-G.....G-CTAAAACTGC.....TGACCTGAAGATTGCTGACACTG.....-A.....-A...GCAAAGACTG			388
H1O.CM.91.MVP5180	-G.....T-ACTGACACTG.....			359
CPZ.CD.90.ANT				0
CPZ.CM.01.SIVcpzCAM13				0
CPZ.CM.05.SIVcpzEK505				0
CPZ.CM.05.SIVcpzLB7				0
CPZ.CM.05.SIVcpzMB66				0
CPZ.CM.05.SIVcpzMT145				0
CPZ.CM.98.CAM3				0
CPZ.CM.98.CAB1				0
CPZ.GA.88.GAB1				0
CPZ.GA.88.GAB2				0
CPZ.TZ.01.TANI				0
CPZ.US.85.CPZUS				0
H2A.DE.x.BEN	--AAG...TCAGGATTGCCAGAGAAAAGAA...-GA-GGCAAACTGAAAGCAAGAGGGATACCATATAGTGAATAACAGGAACAACATACT...TGATGGCTGTAACCCGGCAGGCGC.AATA-AAC--G...-TG...A...TA.A...GGGACTTT			387
H2A.CW.x.ALI	--AAG...TCAGGCTGCCAGAGATGAG...-GA-GGCAAGACTGAAAGCAAGAGGGATACCATTTAGTAAGAACAGGAACAGCTGATT...TGCTCAAGGCAGGAAGTAGCTA.CTAAGAA-C.AGCTGA..GGCTGC...-A...GACTG			460
H2A.SN.x.ALI				457
H2B.CL.x.EHO	-TCAG...TCAGGAATGCCAGAGAAAAGAG...-GA-GGCTAACTGAGAGCAAGAGGAATACCTACAGATTAGACAGGAATAACAGCACAAGA.GGAACTAGCTAACACTGCATAGAGAAGGAA-CTAGCTGA..TACTGC...-A...GACTG			464
H2B.GH.86.D205	--CAG...TCAGGGCTACCAGAGGAGGAG...-GA-GGCTAGACTAAAAGCAAGAGGGATACCTACAGATTAGGCAAGAGCAGCAGCATAAACA.GGAACTGACTGACACTGCACAA.GAAGGAA-CTAGCAGA..CACTGC...-A...GACTG			467
H2G.CL.x.ABT96				0
H2U.FR.96.12034				0
	TCF-1 alpha		NF-k-B-II	
MAC.US.x.239	G-AAG...TCAGGCTGTGTCAGAGGAAGAG...GTTAGAAGAAGGCTAACCGCAAGAGGCCTTTCTTAACATGGCTGACAAGAAGGAAACTCGCTGA.....-ACAGC...-A.....			425
MAC.US.x.251_1A11	G-AAG...TCAGGCTGTGTCAGAGGAAGAG...GTTAGAAGAAGGCTAACCGCAAGAGGCCTTTCTTAACATGGCTGACAAGAAGGAAACTAGATTGA.....GACAGC...-A.....			425
MAC.US.x.251_BK28	G-AAG...TCAGGCTGTGTCAGAGGAAGAG...GTTAGAAGAAGGCTAACCGCAAGAGGCCTTTCTTAACATGGCTGACAAGAAGGAAACTGCTGA.....GATAGC...-A.....			424
MAC.US.x.EMB1_3				0
SMM.SL.92.SL92B				0
SMM.US.x.H9				0
SMM.US.x.PGM53	GTAAG...TCAGGCTGTGTCAGAGGAAGAG...GT-A-GAGAAGGCTAACCGCAAGAGGCCTTTTAAAAATGGCTGACAAGAAGGAAACAGCTGA.....GACAGC...-A.....			351
SMM.US.x.SIVsmH635F_L3	GTAAG...TCAGGCTGTGTCAGAGGAAGAG...GT-A-GAGAAGGCTAACCGCAAGAGGCCTTTTAAAAATGGCTGACAAGAAGGAAACAGCTGA.....GACAGC...-A...CAAGGGACTTT			436
STM.US.x.STM				98
SAB.SN.x.SAB1C	--CCAGAGCAGGTGCCCTCAGCA...CA-GGCTAACCAAAACCATCTCA...GAGAAGGCTAACCGCAAGAGGCCTTTCAAAGATGGCTGACAAGAAGGAAACAGCTGA.....TG.GAGTTGTCATGGTATGACAT-AAGAAC...-G...-TGAG			403
TAN.UG.x.TANI	G--TTAAGGACATGCATGGCTCT...GTTA-GAGGAAGTAGCTAACCGCAGGCTTGGC...GTTAAGCTAACTAGCCGTTGCTAGGAGATGACATAAAGAAG--G...-TCAGC...-A.....			413
VER.DE.x.AGM3				0
VER.KE.x.9063	-T-TTGAAGACATGCATGCAATA...GTGA-GAGGAAGTAG...CTAACCGCAGGCTTG...TG.GTTAAGCCGTTGCCAGGGAGATGACAT-T--AAC--GAC-G...-A...AGGGACTTT			424
VER.KE.x.AGM155	--TTACAGACTTATATCAACA...GTTGGTACAGGAACTAGCCGACCACAGGCTTG...CG.GTTTCTCGGTTGCCTAGGAGATGACAT-AAGAAC--GAC...-A.GCAGGGACTTT			428
VER.KE.x.TY01	--TTACTGACATGCATGGGCTG...GTGA-GAGGAAGTAGCCA..GACCGCAAGCCTG...CG.GTTAGAACCTACCATTGGAGATGACAT-AA-AAC--GAC...-AGCGAAGGGACTTT			427
COL.CM.x.COU1				0
GRV.ET.x.GRI_677	TG-TTAAGGACATGCACGAACAT...GCAAAAGCGC-A-T-TCA-CTTTTGGCGTT...CCAGGGACTTT			375
MND.2.CM.98.CM16				0
DRL.x.x.FAO				0
RCM.GA.x.GAB1				0
RCM.NG.x.NG411				0
MND.1.GA.x.MNDGB1				0
MND.2.GA.x.M14				0
MND.2.x.x.5440				0
MNE.US.x.MNE027				0
LST.CD.88.447				0
LST.CD.88.485				0
LST.CD.88.524				0
LST.KE.x.lho7	-GCCTGCCAAGATACAAAATACT...GCT-TCGCTTTTGACTGCAAGAGGAAGTAGCCAAACCGCAAACCGCAGCACATCTCTTGACGCCGGTTGCTAAGGCAACCGGGCTAGCCGATGCG-A-TGGCTTGCTG--GAGC...-A.....			454
GSN.CM.99.CN166				0
GSN.CM.99.CN71				0
MON.CM.99.L1				0
MON.NG.x.NG1				0
MUS.1.CM.01.1085				0
MUS.1.CM.01.CM1239				0
MUS.2.CM.01.CM1246				0
MUS.2.CM.01.CM2500				0
DEN.CD.x.CD1				0
DEB.CM.99.CM40				0
DEB.CM.99.CM5				0
TAL.CM.00.266				0
TAL.CM.01.8023	--CCTGAAGAGTTCAAGCATGTT...ACTAGTCTGCAGTGGGAGGTGTAGCTTAACCGCAAACCG...CATCCTCTTGCATCGCTAGGCAACGGGGCTAGCCGATGCGCGC...TAGAGT			427
UN.GA.98.14				0
SYK.KE.x.KE51	G--TGCAAGTTAGAGAGACAACA...GGAAAAACCAAGAGGCTGCAGTCAGCACTTAGTAAGAACTATCTTAGGAGACTCATGTGTGACAAGCTCGGCCAGGGGA...C-GGGC			309
SYK.KE.x.SYK173				0

Table with columns for sequence identifiers (e.g., H1B.FR.83.HXB2, H101.AE.TH.90.CM240), regulatory elements (NF-k-B-1, Sp1-III, Sp1-II, Sp1-I, TATA Box, TAR element start), and sequence alignments. Includes a column for alignment scores on the right.

Table with columns for sequence identifiers (e.g., MAC.US.x.239, MAC.US.x.251_1A11), regulatory elements (NF-k-B-1, TATA Box, TAR element start, 5' LTR U3 end, 5' LTR R repeat begin), and sequence alignments. Includes a column for alignment scores on the right.

	5' LTR U3 end, 5' LTR R repeat begin	+1 mRNA start site	TAR element end	
H1B.FR.83.HXB2GGGTCTCTCTGGTTAGACCAGA..TCTGAGCCTGGGAG..CTCTCTGGCTA..ACTAGGG..AACCCTACTGCTTA.....AGCCT.....			525
H101_AE.TH.90.CM240T.....G.....C.....G.....A.....A.....			72
H102_AC.NG.x.IB.NGT.....G.....GGA.....A.....			54
H103_AB.RU.97.KAL153_2T.....G.....GGA.....A.....			0
H104_cpx.CY.94.CY032T.....G.....GGA.....A.....			0
H1A1.UG.85.U455T.....G.....GGA.....A.....			0
H1B.US.90.WEAU160T.....G.....GGA.....A.....			525
H1C_ET.86.ETH2220T.....G.....GGA.....A.....			0
H1D_CD.84.RZR085T.....G.....GGA.....A.....			39
H1F1_BE.93.VI850T.....G.....GGA.....A.....			0
H1G_SE.93.SE6165T.....G.....GGA.....A.....			0
H1H_CF.90.056T.....G.....GGA.....A.....			0
H1J_SE.93.SE7887T.....G.....GGA.....A.....			0
H1K_CM.96.MP535T.....G.....GGA.....A.....			85
H1N_CM.95.YBF30T.....G.....GGA.....A.....			560
H1O_BE.87.ANT170T.....G.....GGA.....A.....			3
H1O_CM.91.MVP5180T.....G.....GGA.....A.....			533
CPZ_CD.90.ANTT.....G.....GGA.....A.....			0
CPZ_CM.01.SIVcpzCAM13T.....G.....GGA.....A.....			67
CPZ_CM.05.SIVcpzEK505T.....G.....GGA.....A.....			69
CPZ_CM.05.SIVcpzLB7T.....G.....GGA.....A.....			71
CPZ_CM.05.SIVcpzMB66T.....G.....GGA.....A.....			70
CPZ_CM.05.SIVcpzMT145T.....G.....GGA.....A.....			70
CPZ_CM.98.CAM3T.....G.....GGA.....A.....			10
CPZ_CM.98.CM145T.....G.....GGA.....A.....			195
CPZ_GA.88.GAB1T.....G.....GGA.....A.....			539
CPZ_GA.88.GAB2T.....G.....GGA.....A.....			0
CPZ_TZ.01.TANIT.....G.....GGA.....A.....			70
CPZ_US.85.CPZUST.....G.....GGA.....A.....			545
H2A_DE.x.BENT.....G.....GGA.....A.....			694
H2A_GW.x.AL1T.....G.....GGA.....A.....			690
H2A_SN.x.STT.....G.....GGA.....A.....			139
H2B_CL.x.EHOT.....G.....GGA.....A.....			697
H2B_GH.86.D205T.....G.....GGA.....A.....			700
H2G_CL.x.AB796T.....G.....GGA.....A.....			62
H2U_FR.96.12034T.....G.....GGA.....A.....			191
T.....G.....GGA.....A.....			545
T.....G.....GGA.....A.....			694
T.....G.....GGA.....A.....			139
T.....G.....GGA.....A.....			697
T.....G.....GGA.....A.....			700
T.....G.....GGA.....A.....			62
T.....G.....GGA.....A.....			191
T.....G.....GGA.....A.....			658
T.....G.....GGA.....A.....			657
T.....G.....GGA.....A.....			646
T.....G.....GGA.....A.....			139
T.....G.....GGA.....A.....			139
T.....G.....GGA.....A.....			141
T.....G.....GGA.....A.....			584
T.....G.....GGA.....A.....			671
T.....G.....GGA.....A.....			331
T.....G.....GGA.....A.....			651
T.....G.....GGA.....A.....			581
T.....G.....GGA.....A.....			89
T.....G.....GGA.....A.....			593
T.....G.....GGA.....A.....			597
T.....G.....GGA.....A.....			596
T.....G.....GGA.....A.....			146
T.....G.....GGA.....A.....			568
T.....G.....GGA.....A.....			292
T.....G.....GGA.....A.....			0
T.....G.....GGA.....A.....			0
T.....G.....GGA.....A.....			0
T.....G.....GGA.....A.....			143
T.....G.....GGA.....A.....			217
T.....G.....GGA.....A.....			0
T.....G.....GGA.....A.....			140
T.....G.....GGA.....A.....			0
T.....G.....GGA.....A.....			0
T.....G.....GGA.....A.....			649
T.....G.....GGA.....A.....			93
T.....G.....GGA.....A.....			94
T.....G.....GGA.....A.....			104
T.....G.....GGA.....A.....			0
T.....G.....GGA.....A.....			92
T.....G.....GGA.....A.....			93
T.....G.....GGA.....A.....			94
T.....G.....GGA.....A.....			95
T.....G.....GGA.....A.....			107
T.....G.....GGA.....A.....			125
T.....G.....GGA.....A.....			127
T.....G.....GGA.....A.....			122
T.....G.....GGA.....A.....			620
T.....G.....GGA.....A.....			137
T.....G.....GGA.....A.....			496

Table with columns for sequence identifiers, Poly-A signal, 5' LTR R repeat end, 5' LTR U5 start, Extensive secondary structure, and alignment positions. The table lists various HIV sequences and their corresponding genomic features and alignment coordinates.

	5' LTR U5 end	Lys tRNA primer binding site	
H1B.FR.83.HXB2AAAA.TCCTCTAGCAG.....TGGCGCCCGAACAGGG.....ACCTGA.....	657
H101_AE.TH.90.CM240C.....C.....TC.....	204
H102.AC.NG.x.IB.NGA.....T.....	185
H103_AB.RU.97.KAL153_2	0
H104_cpx.CY.94.CY032T.....	4
H1A1.UG.85.U455TC.....	126
H1B.US.90.WEAU160C.....	657
H1C.ET.86.ETH220G.....	36
H1D.CD.84.R47R083	171
H1F1.BE.93.VI850T.....	8
H1G.SE.93.SE6165GGCTT.....T.....	40
H1H.CF.90.056T.....	4
H1J.SE.93.SE7887	0
H1K.CM.96.MP535	0
H1N.CM.95.YBF30T.....	215
H1O.BE.87.ANT70T.....	687
H1O.CM.91.MVP5180GGC.....	659
CPZ.CD.90.ANT	9
CPZ.CM.01.SIVcpzC.AM13G.....	198
CPZ.CM.05.SIVcpzEK505T.....	198
CPZ.CM.05.SIVcpzLB7T.....	202
CPZ.CM.05.SIVcpzMB66T.....	204
CPZ.CM.05.SIVcpzMT145T.....	201
CPZ.CM.98.CAM3T.....	26
CPZ.CM.98.CAM5GTG.....T.....	331
CPZ.GA.88.GAB1AGG.....T.....	676
CPZ.GA.88.GAB2T.....	4
CPZ.TZ.01.TAN1GTGGCTAAGTA.....C.....	206
CPZ.US.85.CPZUSGGTG.....T.....	680
H2A.DE.x.BENATCCAAGGCAGG.....C.....G.....T.....	880
H2A.CW.x.ALIAACGAGGCAGG.....C.....G.....T.....	875
H2A.SN.x.AACCAAGGCAGG.....C.....G.....T.....	326
H2B.CL.x.EHOAACCAAGGCAGG.....C.....A.....T.....	876
H2B.GH.86.D205AACCGAAGCGGG.....C.....A.....T.....G.....	877
H2G.CL.x.ABT96AACCGAAGCGGG.....C.....A.....T.....	248
H2U.FR.96.12034AACCGAGCAGG.....C.....G.....T.....	373
	5' LTR U5 end	Lys tRNA primer binding site	
MAC.US.x.239AACCGAAGCAGG.....C.....A.....T.....T.....	843
MAC.US.x.251_1A11AACCGAAGCAGG.....C.....A.....T.....	842
MAC.US.x.251_BK28AACCGAAGCAGG.....C.....A.....T.....	830
MAC.US.x.EMB1_3AACCGAAGCAGG.....C.....TGAT.....	338
SMM.SL.92.SL92BGCAACTCTCTGAGTG.....A.....G.....T.....T.....	321
SMM.US.x.H9AACCGAAGCAGG.....C.....G.....T.....	326
SMM.US.x.PGM53AACCGAAGCAGG.....C.....A.....T.....	769
SMM.US.x.SIVsmH635F_L3AACCGAAGCAGG.....C.....G.....T.....	857
STM.US.x.STMTACTGAGGCAGA.....C.....A.....T.....	307
SAB.SN.x.SAB1CTCCCTGATAGAA.....CT.....	801
TAN.UG.x.TAN1T.....CCAGGCTGAG.....C.....	728
VER.DE.x.AGM3AACCCAGGCAGG.....G.....A.....C.....	241
VER.KE.x.9063AACCCAGGCAGG.....G.....A.....C.....	746
VER.KE.x.AGM155AACCCAGGCAGG.....G.....A.....C.....	745
VER.KE.x.TY01AACCCAGGCAGG.....G.....A.....C.....	748
COL.CM.x.CCU1TCCGGGAAGG.....GGC.....T.....G.....	391
GRV.ET.x.GRI_677TACCAGGTGGG.....G.....ACTC.....	710
MND_2.CM.98.CM16GGC.....AGTAGAGAA.....C.....	536
DRL.x.x.FAO	92
RCM.GA.x.GAB1	22
RCM.NG.x.NG411	25
MND_1.GA.x.MNDGB1AGGACTTCTAGTTA.....CCCTAGAAAGCCTTTTCTAG.....	295
MND_2.GA.x.M14GAG.....AGTAGAGAA.....C.....	470
MND_2.x.x.5440	97
MNE.US.x.MNE027AACCGAAGCAGG.....C.....A.....T.....	324
LST.CD.88.447	0
LST.CD.88.485	0
LST.CD.88.524	0
LST.KE.x.lho7GATCCGGG.....C.....GGG.....TC.....	867
GSN.CM.99.CN166CCG.....GGTGGACACC.....G.....CA.....C.....	248
GSN.CM.99.CN71CGC.....GGTGGATACC.....G.....CA.....C.....	257
MON.CM.99.L1GAG.....GGGACAGCAG.....G.....C.....TCTCCA.....G.....	243
MON.NG.x.NG1	0
MUS_1.CM.01.1085GTG.....GTGTGAGTCTGG.....AACACCT.....GCA.....G.....	246
MUS_1.CM.01.CM1239GTG.....GTGTGAGTCTGG.....AACACCT.....GCA.....G.....	248
MUS_2.CM.01.CM1246GGGA.....T.....CTC.....	320
MUS_2.CM.01.CM2500GATCCAGGGCT.....CA.....ACCT.....CCA.....G.....	276
MEN.CD.x.CD1CTTACCGAGGTGA.....G.....C.....C.....	278
DEB.CM.99.CM40CTAGGCTCAAG.....CG.....G.....GTC.....	269
DEB.CM.99.CM5CTAGGCTCTGGCAGG.....G.....AAC.....	273
TAL.CM.00.266GGGCCAAGCGGGCTTCGGGGGA.....CGCC.....GTTTTGGGTC.....CAAA.....G.....A.....AG.....AACCCGAGGGGTCGTGAGAGGACGACAGTTCCGGAGTCGAGCCCGGATCGCGCAAGACGACGCGCTCGCAAAGTAAAGTAAGGTAAGGACCGATCGGTCACCCGGCGAGTA.....	462
TAL.CM.01.8023GTGATCCGGGGCTG.....C.....GGC.....	862
SUN.GA.98.114G.....G.....C.....TT.....	265
SYK.KE.x.KE51G.....C.....TT.....	624
SYK.KE.x.SYK173C.....TT.....	624

H1B.FR.83.HXB2
H101.AE.TH.90.CM240
H102.AC.NG.x.IBNG3
H103.AB.RU.97.KALI153_2
H104.cbpx.XI.94.CY032
H1A1.UG.85.U455
H1B.US.90.WEAU160
H1C.ET.86.ETH220
H1D.CD.84.RZOR85
H1F1.BE.93.VI850
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP553
H1N.CM.95.YBF30
H1O.BE.87.AN170
H1O.CM.91.MVP5180
CPZ.CD.90.ANT
CPZ.CM.01.SIV.cpzCAM13
CPZ.CM.05.SIV.cpzEK505
CPZ.CM.05.SIV.cpzLB7
CPZ.CM.05.SIV.cpzMB66
CPZ.CM.05.SIV.cpzMT145
CPZ.CM.98.CAM3
CPZ.CM.98.CAM3
CPZ.GA.88.GAB1
CPZ.GA.88.GAB2
CPZ.TZ.01.TANI
CPZ.US.85.CPZUS
H2A.DE.x.BEN
H2A.GW.x.AL1
H2A.SN.x.ST
H2B.CH1.EHO
H2B.GH.86.D205
H2G.CL.x.AB19E
H2U.FR.96.12034
MAC.US.x.239
MAC.US.x.251_1A11
MAC.US.x.251_1BK28
MAC.US.x.EMB1_3
SMM.SL.92.S192B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIV.smmH635F_L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.UG.x.TANI
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TY01
COL.CM.x.CGU1
GRV.ET.x.GRI_677
MND_2.CM.98.CM16
DRL.x.x.FAO
RCM.GA.x.GAB1
RCM.NG.x.NG411
MND_1.GA.x.MNDGB1
MND_2.GA.x.M14
MND_2.x.x.5440
MINE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.lho7
GSN.CM.99.CN166
GSN.CM.99.CN172
MON.CM.99.L1
MUS_1.CM.01.1085
MUS_1.CM.01.CM1239
MUS_2.CM.01.CM1250
MUS_2.CM.01.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.4F51
SYK.KE.x.SYK173

Packaging loops begin

759
324
299
0
149
226
159
276
114
162
106
19
336
810
785
320
121
318
327
329
148
354
793
322
800
1041
1036
487
1034
1033
402
522
1000
999
987
508
425
989
1013
967
872
400
912
900
561
866
706
181
419
630
476
0
0
1037
428
414
6
416
498
448
437
446
632
118
1032
401
764



H1B.FR.83.HXB2
Gag
H101_AE.TH.90.CM240
H102_AG.NG.x.IBNG
H103_AB.RU.97.KAL153_2
H104_cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WE.AU160
H1C.ET.86.ETH.2320
H1D.CD.84.84ZR085
H1F1.BE.93.VI850
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.ANT170
H1O.CM.91.MVP5180
CPZ_CD.90.ANT
CPZ_CM.01.SIVcpz.CAM13
CPZ_CM.05.SIVcpz.EK505
CPZ_CM.05.SIVcpz.LB7
CPZ_CM.05.SIVcpz.MB66
CPZ_CM.05.SIVcpz.MT145
CPZ_CM.98.CAM3
CPZ_CM.98.CAM5
CPZ_GA.88.GAB1
CPZ_GA.88.GAB2
CPZ_TZ.01.TANI
CPZ_US.85.CPZUS
H2A.DE.x.BE1
H2A.GE.x.ALI
H2A.SN.x.ST
H2B.CL.x.EHO
H2B.GH.86.D205
H2G.CL.x.AB196
H2U.FR.96.12034

Table with columns for packaging loops end, Gag and Gag-Pol start, and sequence alignment. Includes sequences for H1B.FR.83.HXB2 and MAC.US.x.239.

MAC.US.x.239
Gag
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMBL_3
SMM.SL.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F.L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.UG.x.TAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TYO1
COL.CM.x.CGU1
GRV.ET.x.GRI_677
MND_2.CM.98.CM16
DRL.x.FAO
RCLM.GA.x.GAB1
RCLM.NG.x.NG411
MND_1.GA.x.MNDGB1
MND_2.GA.x.M14
MND_2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.Jho7
GSM.CM.99.CM166
GSM.CM.99.CM171
MON.CM.99.L1
MON.NG.x.NG1
MUS_1.CM.01.1085
MUS_1.CM.01.CM1239
MUS_2.CM.01.CM1246
MUS_2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.265
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

Table with columns for packaging loops end, Gag and Gag-Pol start, and sequence alignment. Includes sequences for MAC.US.x.239 and HIV Sequence Compendium 2007.

H1B.FR.83.HXB2
Gag
H101_AE.TH.90.CM240
H102_AG.NG.x.IBNG
H103_AB.RU.97.KAL153.2
H104_cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WEAU160
H1C.ET.86.ETH
H1D.CD.84.84ZR085
H1F1.BE.93.VI850
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.ANT170
H1O.CM.91.MVP5180
CPZ_CD.90.ANT
CPZ_CM.01.SIVcpzCAM13
CPZ_CM.05.SIVcpzEK505
CPZ_CM.05.SIVcpzLB7
CPZ_CM.05.SIVcpzMB66
CPZ_CM.05.SIVcpzMT145
CPZ_CM.98.CAM3
CPZ_CM.98.CAM3
CPZ_GA.88.GAB1
CPZ_GA.88.GAB2
CPZ_TZ.01.TANI
CPZ_US.85.CPZUS
H2A.DE.x.HBN
H2A.GW.x.ALI
H2A.SN.x.ST
H2B.CL.x.EHO
H2B.HG.86.D205
H2G.CL.x.AB796
H2U.FR.96.12034

MAC.US.x.239
Gag
MAC.US.x.251.1A11
MAC.US.x.251.BK28
MAC.US.x.EMBL_3
SMM.SL.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F.L3
STM.US.8TM
SAB.SN.x.SAB1C
TAN.UG.x.TANI
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.ACM155
VER.KE.x.TYO1
COL.CM.x.CGU1
GRV.ET.x.GRI.677
MND.2.CM.98.CM16
DRL.x.FAO
RCLM.GA.x.GAB1
RCM.NG.x.NG411
MND.1.GA.x.MNDGB1
MND.2.GA.x.M14
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.lho7
GSN.CM.99.CM166
GSN.CM.99.CM71
MON.CM.99.L1
MON.NG.x.NG1
MUS.1.CM.01.1085
MUS.1.CM.01.CM1239
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

TATGGGCAAGCAGGAGCTAGAACGATTGCGAGTTAATCGTGGCTGTAGAAACATCAGAAGGCTGTAGACAAACTGCGACAGCTACAACCTCCCTTCAGACAGGATCAGAAGAAGCTAGATCATTATATAACAGTAGCAACCCCTTATTGTGCATCAAAAGG 1062
V W A S R E L E R F A V N P G L E T S E G C R Q I L G Q L O P S L Q T G S E E L R S L Y N T V A T L Y C V H Q R
... (The rest of the sequence alignment follows in a similar format, with gaps represented by dashes and letters representing nucleotides) ...

H1B.FR.83.HXB2	ATAGAGATAAAAGACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAAGAAAAAGCACAGCAAGCAGCAGTGCACAGAA	1158
Gag	I E T K D T K E A L D K I E E E Q	N K S K K K A Q Q A A A D T G		
H101_AE.TH.90.CM240G.....T.....A.....TA.....T.....CC.....GG.....GA.....G.....G.....		723
H102_AG.NG.x.IBNGC.....T.....A.....G.....ATA.....T.....G.....C.....C.....G.....T.....GA.....C.....		698
H103_AB.RU.97.KAL153_2C.....T.....A.....G.....ATA.....T.....G.....C.....C.....GA.....C.....A.....G.....		390
H104_cpX.CY.94.CY032TG.....C.....T.....A.....ATA.....GT.....G.....C.....C.....GA.....G.....G.....C.....G.....		524
H1A1.UG.85.U455TG.....C.....A.....T.....A.....AT.....T.....G.....AC.....C.....GGA.....G.....G.....A.....		625
H1B.US.90.WEAU160G.....T.....C.....T.....A.....A.....G.....C.....C.....G.....G.....G.....A.....CTGACA.....A.....		1157
H1C.ET.86.ETHC.....T.....A.....A.....A.....G.....C.....C.....GG.....G.....G.....G.....		675
H1D.CD.84.84ZR085G.....C.....A.....C.....AG.....G.....AA.....GATA.....G.....C.....C.....G.....AT.....G.....G.....ATG.....T.....A.....G.....		513
H1F1.BE.93.VI850TG.....G.....T.....G.....T.....ATA.....G.....AC.....AC.....GC.....G.....A.....A.....ATG.....T.....AG.....A.....		561
H1G.SE.93.SE6165G.....T.....C.....G.....AC.....A.....G.....AC.....AC.....GC.....G.....A.....A.....A.....AA.....		505
H1H.CF.90.056G.....C.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		481
H1I.SE.93.SE7887	TATA A TGG T GC C G AA C A T A T A TAATGG.....G.....GC.....GG.....GT.....TGC.....G.....C.....TAAG.....AGACAC.....AGC.....		360
H1N.CM.96.MP535	T T C C CG T AC C G AA C A T A TAATGG.....G.....GC.....GG.....GT.....TGC.....G.....C.....TAAG.....AGACAC.....AGC.....		728
H1N.CM.95.YBF30G.....A.....C.....AG.....G.....AA.....GATA.....G.....C.....G.....AT.....G.....G.....ATG.....T.....A.....G.....		1207
H1O.CM.91.MVP5180TG.....G.....T.....G.....T.....ATA.....G.....AC.....AC.....GC.....G.....A.....A.....ATG.....T.....AG.....A.....		1182
CPZ_CD.90.ANTG.....T.....C.....G.....AC.....A.....G.....AC.....AC.....GC.....G.....A.....A.....A.....AA.....		570
CPZ_CM.01.SIVcpzCAM13C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		678
CPZ_CM.05.SIVcpzEBK505C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		706
CPZ_CM.05.SIVcpzLB7C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		704
CPZ_CM.05.SIVcpzMB66C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		704
CPZ_CM.05.SIVcpzMT145C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		700
CPZ_CM.98.CAM3C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		854
CPZ_CM.98.CAM3C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1222
CPZ_GA.88.GAB1C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		535
CPZ_GA.88.GAB2C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		789
CPZ_TZ.01.TANIC.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1209
CPZ_US.85.CPZUSC.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1479
H2A.DE.x.BRNC.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		924
H2A.GW.x.ALIC.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1447
H2A.SN.x.STC.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1446
H2B.CL.x.EHOC.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		842
H2B.GH.86.D205C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		958
H2G.CL.x.AB796C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		
H2U.FR.96.12034C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		
MAC.US.x.239C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1429
Gag	E K V K H T E A K Q I V O R H L V E	T G T T E T M P K T S R P A		
MAC.US.x.251_1A11C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1428
MAC.US.x.251_BK28C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1417
MAC.US.x.EMBL_3C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		915
SMM.SL.92.SL92BC.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		869
SMM.US.x.H9C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		911
SMM.US.x.PGM53C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1355
SMM.US.x.SIVsmH635F_L3C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1441
STM.US.x.STMC.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1085
SAB.SN.x.SAB1CC.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1426
TAN.UG.x.TANIC.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1299
VER.DE.x.AGM3C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		857
VER.KE.x.9063C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1339
VER.KE.x.ACMI155C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1328
VER.KE.x.TYO1C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1336
COL.CM.x.CGU1C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		975
GRV.ET.x.GRI_677C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1273
MND_2.CM.98.CM16C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1116
DRL.x.FAOC.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		670
RCLM.GA.x.GAB1C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		588
RCLM.NG.x.NG411C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		588
MND_1.GA.x.MNDGB1C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		800
MND_2.GA.x.M14C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1040
MND_2.x.x.5440C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		673
MNE.US.x.MNE027C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		909
LST.CD.88.447C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		333
LST.CD.88.485C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		333
LST.CD.88.524C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		330
LST.KE.x.lho7C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1438
GSN.CM.99.CM166C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		834
GSN.CM.99.CM71C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		840
MON.CM.99.L1C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		821
MON.NG.x.NG1C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		0
MUS_1.CM.01.1085C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		809
MUS_1.CM.01.CM1239C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		808
MUS_2.CM.01.CM1246C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		888
MUS_2.CM.01.CM2500C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		850
DEN.CD.x.CD1C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		924
DEB.CM.99.CM40C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		930
DEB.CM.99.CM5C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		925
TAL.CM.00.266C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1045
TAL.CM.01.8023C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		543
SUN.GA.98.L14C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1410
SYK.KE.x.KE51C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		865
SYK.KE.x.SYK173C.....C.....G.....CA.....AC.....G.....A.....GA.....A.....G.....G.....C.....CG.....GA.....A.....A.....A.....CT.....		1237

H1B.FR.83.HXB2
Gag
H101 AE.TH.90.CM240
H102 AG.NG.x.IBNG
H103 AB.RU.97.KAL153.2
H104 cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WE.AU160
H1C.ET.86.ETH2220
H1D.CD.84.84ZR085
H1F1.BE.93.VI850
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.ANT170
H1O.CM.91.MVP5180
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzLB7
CPZ.CM.05.SIVcpzMB66
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CAM3
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.GAB2
CPZ.TZ.01.TANI
CPZ.US.85.CPZUS
H2A.DE.x.BF
H2A.GW.87.AL1
H2A.SN.x.STI
H2B.CL.x.EHO
H2B.GH.86.D205
H2G.CL.x.AB796
H2U.FR.96.12034

Table with 20 columns and 100 rows of HIV sequence alignments. Header: Gag p17 Matrix end, Gag p24 Capsid start. Row 1: .CACAGCAATCAGGTCAGCCAAAATACCTATAGTCGAGAACATCCAGGGCCAAATGGTACATCAGGCCATACCTAGAACTTTAAATGCATGGTAAAGAGTAGTGAAGAAGAGGCTTCAGCCAGAAGTATACCCATGTTCAGCATTATCAGAAGAGGCCA 1327

MAC.US.x.239
Gag
MAC.US.x.251.1A11
MAC.US.x.251.BK28
MAC.US.x.EMBL_3
SMM.SL.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F.L3
STM.US.x.STM
SAB.N.S.x.SAB1C
TAN.UG.x.TANI
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.ACM155
VER.KE.x.TYO1
COL.CM.x.CGU1
GRV.ET.x.GRI_677
MND_2.CM.98.CM16
DRL.x.FAO
RCLM.GA.x.GAB1
RCLM.NG.x.NG411
MND_1.GA.x.MNDG1
MND_2.GA.x.M14
MND_2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.lho7
GSM.CM.99.CM166
GSM.CM.99.CM71
MON.CM.99.L1
MON.NG.x.NG1
MUS_1.CM.01.1085
MUS_1.CM.01.CM1239
MUS_2.CM.01.CM1246
MUS_2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

Table with 20 columns and 100 rows of HIV sequence alignments. Header: Gag p17 Matrix end, Gag p24 Capsid start. Row 1: A-CATCT-GCGCAGAG-AGG ---AG-CAA-A-TAGGT ---TA-CTAT--C-C-T-C-AT-AGC-G---A---G---A---AAA-TG-AG---AG-G-AGGA-CAG--C-G---TTG- 1596

H1B.FR.83.HXB2
H101.AE.TH.90.CM240
H102.AG.NG.x.IBNG
H103.AB.RU.97.KAL153.2
H104.cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WE.AU.160
H1C.ET.86.ETH.220
H1D.CD.84.84ZR085
H1F1.BE.93.VI850
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.AN170
H1O.CM.91.MVP5180
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzLB7
CPZ.CM.05.SIVcpzMB66
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CAM5
CPZ.CM.98.CAM3
CPZ.GA.88.GAB1
CPZ.GA.88.GAB2
CPZ.TZ.01.TANI
CPZ.US.85.CPZUS
H2A.DE.x.IB.N
H2A.GW.x.ALI
H2A.SN.x.ST
H2B.CL.x.EHO
H2B.GH.86.D205
H2G.CL.x.AB796
H2U.FR.96.12034
MAC.US.x.239
Gag
MAC.US.x.251.1A11
MAC.US.x.251.BK28
MAC.US.x.EMBL.3
SMM.SL.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F.L3
STM.US.x.STM
SAB.SN.x.SAB1
TAN.UG.x.TANI
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.ACM155
VER.KE.x.TYO1
COL.CM.x.CGU1
GRV.ET.x.GRI.677
MND.2.CM.98.CM16
DRL.x.x.FAO
RCLM.GA.x.GAB1
RCM.NG.x.NG411
MND.1.GA.x.MNDGB1
MND.2.GA.x.M14
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.lho7
GSN.CM.99.CM166
GSN.CM.99.CM71
MON.CM.99.L1
MON.NG.x.NG1
MUS.1.CM.01.1085
MUS.1.CM.01.CM1239
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

Table with 4 columns: Reference sequence (HXB2), Accession ID, Nucleotide sequence, and Position. The table lists various HIV-1 sequences and their corresponding positions in the genome.

H1B.FR.83.HXB2	AGGACCAAAGAACCTTTAGAGACTATGAGACCGTTCCTAAAACTCTAAGAGCCGAGCAAGCTTCACAGGAGTAAAAAATGGATGACAGAAACCTTTGGTCCAAAATGCGAACCCAGATTTGAAGCATTTTAAAAAGATTGGGACCAGCGCTACACTAG	1819
H101.AE.TH.90.CM240	---	1884
H102.AG.NG.x.IBNG	---	1350
H103.AB.RU.97.KAL153.2	---	1051
H104.cpx.CY.94.CY032	---	1185
H1A1.UG.85.U455	---	1274
H1B.US.90.WE.AU.160	---	1818
H1C.ET.86.ETH.220	---	1193
H1D.CD.84.84ZR085	---	1336
H1F1.BE.93.VI850	---	1162
H1G.SE.93.SE6165	---	1222
H1H.CF.90.056	---	1166
H1I.SE.93.SE.7887	---	1142
H1K.CM.96.MP535	---	1018
H1N.CM.95.YBF30	---	1383
H1O.BE.87.ANT170	---	1868
H1O.CM.91.MVP5180	---	1843
CPZ.CD.90.ANT	---	1240
CPZ.CM.01.SIVcpzCAM13	---	1396
CPZ.CM.05.SIVcpzEK505	---	1367
CPZ.CM.05.SIVcpzLB7	---	1365
CPZ.CM.05.SIVcpzMB66	---	1362
CPZ.CM.05.SIVcpzMT145	---	1358
CPZ.CM.98.CAM3	---	1209
CPZ.CM.98.CAM3	---	1512
CPZ.GA.88.GAB1	---	1880
CPZ.GA.88.GAB2	---	1197
CPZ.TZ.01.TANI	---	1460
CPZ.US.85.CPZUS	---	1867
H2A.DE.x.BRN	---	2138
H2A.GW.x.ALI	---	2131
H2A.SN.x.ST	---	1583
H2B.CL.x.EHO	---	2100
H2B.GH.86.D205	---	2099
H2G.CL.x.AB196	---	1501
H2U.FR.96.12034	---	1617
MAC.US.x.239	---	2085
Gag	---	
MAC.US.x.251.1A11	---	2084
MAC.US.x.251.BK28	---	2073
MAC.US.x.EMBL.3	---	1571
SMM.SL.92.SL92B	---	1528
SMM.US.x.H9	---	1570
SMM.US.x.PGM53	---	2014
SMM.US.x.SIVsmH635F.L3	---	2100
STM.US.x.STM	---	1747
SAB.SN.x.SAB1C	---	2088
TAN.UG.x.TANI	---	1958
VER.DE.x.AGM3	---	1499
VER.KE.x.9063	---	1998
VER.KE.x.ACM155	---	1987
VER.KE.x.TYO1	---	1995
COL.CM.x.CGU1	---	1606
GRV.ET.x.GRI.677	---	1935
MND.2.CM.98.CM16	---	1775
DRL.x.FAO	---	1329
RCLM.GA.x.GAB1	---	1247
RCLM.NG.x.NG411	---	1247
MND.1.GA.x.MNDGB1	---	1449
MND.2.GA.x.M14	---	1699
MND.2.x.x.5440	---	1332
MNE.US.x.MNE027	---	1365
LST.CD.88.447	---	991
LST.CD.88.485	---	991
LST.CD.88.524	---	985
LST.KE.x.lho7	---	2075
GSN.CM.99.CM166	---	1496
GSN.CM.99.CM171	---	502
MON.CM.99.L1	---	1483
MON.NG.x.NG1	---	49
MUS.1.CM.01.1085	---	1467
MUS.1.CM.01.CM11239	---	1469
MUS.2.CM.01.CM11246	---	1550
MUS.2.CM.01.CM2500	---	512
DEN.CD.x.CD1	---	1592
DEB.CM.99.CM40	---	1598
DEB.CM.99.CM5	---	1589
TAL.CM.00.266	---	1707
TAL.CM.01.8023	---	1205
SUN.GA.98.L14	---	2062
SYK.KE.x.KE51	---	1527
SYK.KE.x.SYK173	---	1899

	Gag p24 Capsid end	Gag p2 start	Gag p2 end	Gag p7 nucleocapsid start	Gag-Pol fusion TF protein start	
H1B.FR.83.HXB2	AAGAAATGATGACAGCATGTCCAGGAGTAGGAGGACCCGCCCAAGGCAAGAGT	TTTGGCTGAAGCAATGAGCCAAAGTAAATAATTCAGTACC	ATAATGATGCAGAGAGCAATTTTAGGAAACCAAAA	1509
Gag	E E M M T A C Q G V G G P G H K A R V L A E A M S O V C A T N S A T	I M M Q R G N F R N G R	1475
H102_AG.NG.x.IBNG	1176
H103_AB.RU.97.KAL153_2	1319
H104_cpx.CY.94.CY032	1399
H1AT.UG.85.U455	1949
H1B.US.90.WU.AU160	1321
H1C.ET.86.ETH2220	1470
H1D.CD.84.84ZR085	1287
H1F1_BE.93.VI850	1356
H1G.SE.93.SE6165	1300
H1H.CF.90.056	1270
H1I.SE.93.SE7887	1146
H1N.CM.95.MP535	1517
H1O.CM.91.MVP5180	2014
CPZ_CD.90.ANT	1989
CPZ_CM.01.SIVcpzCAM13	1380
CPZ_CM.05.SIVcpzEK505	1533
CPZ_CM.05.SIVcpzLB7	1498
CPZ_CM.05.SIVcpzMB66	1496
CPZ_CM.05.SIVcpzMT145	1484
CPZ_CM.98.CAM3	1489
CPZ_CM.98.CAM3	1490
CPZ_CM.98.CAM3	1643
CPZ_GA.88.GAB1	2017
CPZ_GA.88.GAB2	1328
CPZ_TZ.01.TANI	1588
CPZ_US.85.CPZUS	1998
H2A.DE.x.IBEM1	2257
H2A.GW.x.ALI	2250
H2A.SN.x.ST	1702
H2B.CL.x.EHO	2228
H2B.GH.86.D205	2230
H2U.CL.AB.T96	1632
H2V.FR.96.12034	1748
	Gag p24 Capsid end	Gag p2 Spacer start	Gag p2 Spacer end	Gag p8 Nucleocapsid start	
MAC.US.x.239	2213
Gag	E E M L T A C Q G V G G P G K A R L M A E A K A G A P V P P	F A A A Q G	2212
MAC.US.x.251_1A11	2201
MAC.US.x.251_BK28	1699
MAC.US.x.EMBL_3	1659
SMM.SL.92.SL92B	1698
SMM.US.x.H9	2142
SMM.US.x.PGM53	2228
SMM.US.x.SIVsmH635F_L3	1872
STM.US.x.STM	2228
SAB.SN.x.SAB1C	2083
TAN.UG.x.TANI	1624
VER.DE.x.AGM3	2122
VER.KE.x.9063	2117
VER.KE.x.ACM155	1746
VER.KE.x.TYO1	2054
COL.CM.x.CGU1	1900
GRV.ET.x.GRI_677	1454
MND_2.CM.98.CM16	360
DRL.x.FAO	1372
RCM.GA.x.GAB1	1574
RCM.NG.x.NG411	1824
MND_1.GA.x.MNDGB1	1457
MND_2.GA.x.M14	693
MND_2.x.x.5440	1128
MNE.US.x.MNE027	1128
LST.CD.88.447	1128
LST.CD.88.485	2212
LST.CD.88.524	1630
LST.KE.x.Iho7	633
GSN.CM.99.CN166	1593
GSN.CM.99.CN71	174
MON.CM.99.L1	1592
MON.NG.x.NG1	1585
MUS_1.CM.01.1085	1684
MUS_1.CM.01.CM11239	1646
MUS_2.CM.01.CM11236	1732
MUS_2.CM.01.CM2500	1735
DEN.CD.x.CD1	1726
DEB.CM.99.CM40	1844
DEB.CM.99.CM5	1339
TAL.CM.00.269	2202
TAL.CM.01.8023	1655
SUN.GA.98.L14	2018
SYK.KE.x.KE51	
SYK.KE.x.SYK173	

	Gag p7 nucleocapsid end															Gag p1 start															Gag p1 end															Gag p6 start														
	Gag-Pol -1 ribosomal slip site															Gag-Pol -1 ribosomal slip site															Gag-Pol -1 ribosomal slip site															Gag-Pol -1 ribosomal slip site														
	G C W K C G K E G H Q M K D C T															E R Q A N F L G K T W P S															L Q G K A															Y K G R P														
H1B.FR.83.HXB2	.GGCTGTTGAAATGTGAAAGGAGGACACCAATGAAAGTTGACT.....GAGACAGGCTAATTTTTAGGGAAGTCTGGCCTTC.....CTACAGGGAAGGCC.....AGGGAATTTTCTTC 2137																																																											
GagG C W K C G K E G H Q M K D C T E R Q A N F L G K T W P S L Q G K A Y K G R P																																																											
H101_AE.TH.90.CM240T.....C G.....T.....C.....C.....A.....T.....A.....G.....G.....A.....C.....1696																																																											
H102_AG.NG.x.IBNGG.....A.....C.....T.....C.....C.....A.....T.....A.....G.....G.....A.....C.....1662																																																											
H103_AB.RU.97.KAL153.2C.....G.....A.....G.....T.....C.....C.....A.....T.....A.....G.....G.....A.....C.....1363																																																											
H104_cpx.CY.94.CY032C.....G.....C.....G.....A.....G.....T.....C.....C.....A.....T.....A.....G.....G.....A.....C.....1506																																																											
H1A_FLG.85.185C.....G.....C.....G.....A.....G.....T.....C.....C.....A.....T.....A.....G.....G.....A.....C.....1586																																																											
H1B.US.90.WEAU160C.....G.....C.....G.....A.....G.....T.....C.....C.....A.....T.....A.....G.....G.....A.....C.....1536																																																											
H1C.ET.86.ETH2220C.....G.....C.....G.....A.....G.....T.....C.....C.....A.....T.....A.....G.....G.....A.....C.....1508																																																											
H1D.CD.84.84ZR085C.....G.....C.....G.....A.....G.....T.....C.....C.....A.....T.....A.....G.....G.....A.....C.....1657																																																											
H1F1_BE.93.VI850G.....A.....C.....C.....C.....G.....A.....T.....A.....T.....A.....G.....G.....A.....C.....1474																																																											
H1G.SE.93.SE6165G.....A.....C.....C.....C.....G.....A.....T.....A.....T.....A.....G.....G.....A.....C.....1543																																																											
H1H_CF.90.056G.....A.....C.....C.....C.....G.....A.....T.....A.....T.....A.....G.....G.....A.....C.....1487																																																											
H1I_SE.93.SE7887G.....A.....C.....C.....C.....G.....A.....T.....A.....T.....A.....G.....G.....A.....C.....1457																																																											
H1K_CM.96.MP535G.....G.....C.....A.....T.....G.....C.....C.....A.....T.....A.....G.....G.....A.....C.....1333																																																											
H1N_CM.95.YBF30G.....G.....C.....A.....T.....T.....G.....C.....C.....A.....T.....A.....G.....G.....A.....C.....1710																																																											
H1O_BE.87.ANT70T.....C.....G.....C.....A.....T.....T.....C.....G.....A.....A.....A.....A.....G.....G.....A.....C.....2204																																																											
H1O_CM.91.MVP5180T.....C.....G.....C.....A.....T.....T.....C.....G.....A.....A.....A.....A.....G.....G.....A.....C.....2179																																																											
CPZ_CD.90.ANTC.....G.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....1579																																																											
CPZ_CM.01.SIVcpzCAM13C.....G.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....1720																																																											
CPZ_CM.05.SIVcpzEK505C.....G.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....1691																																																											
CPZ_CM.05.SIVcpzLB7C.....G.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....1683																																																											
CPZ_CM.05.SIVcpzMB66C.....G.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....1674																																																											
CPZ_CM.05.SIVcpzMT145C.....G.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....1842																																																											
CPZ_CM.98.CAM5A.....C.....G.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....1845																																																											
CPZ_CM.98.CAM5A.....C.....G.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....1845																																																											
CPZ_GA.88.GAB1A.....C.....G.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....2204																																																											
CPZ_GA.88.GAB2A.....C.....G.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....1521																																																											
CPZ_TZ.01.TAN1T.....C.....G.....G.....C.....A.....G.....T.....G.....C.....C.....C.....A.....A.....A.....G.....G.....A.....C.....1793																																																											
CPZ_US.85.CPZUSG.....C.....G.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....2203																																																											
H2A.DE.x.BENC.....G.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....2431																																																											
H2A.GW.x.ALIC.....C.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....2444																																																											
H2A.SN.x.STC.....C.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....2444																																																											
H2B.CL.x.EHOA.....C.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....1886																																																											
H2B.GH.86.D205A.....C.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....2414																																																											
H2C.CL.x.AB796G.....C.....G.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....1819																																																											
H2U.FR.96.12034A.....C.....A.....C.....C.....T.....G.....C.....A.....C.....A.....A.....A.....G.....G.....A.....C.....1932																																																											

	Gag p8 Nucleocapsid end															Gag p1 Spacer start															Gag p1 spacer end															Gag p6 start														
	A A T G A C T G T T G C C A A															G G G F L G F L G P W G															A A A G A															C C C C A														
	G C W K C G K M D H V M A K C P															D R Q A G F L G G P W G															G K P															R N Q P														
MAC.US.x.239A.....C.....A.....A.....T.....G.....A.....C.....T.....G.....T.....G.....C.....C.....A.....A.....2397																																																											
PolG.....C.....W.....K.....C.....G.....K.....M.....D.....H.....V.....M.....A.....K.....C.....P.....D.....R.....Q.....A.....G.....F.....L.....G.....G.....P.....W.....G.....G.....K.....P.....R.....N.....Q.....P.....																																																											
MAC.US.x.251.1A11A.....C.....A.....A.....T.....G.....A.....C.....T.....G.....T.....G.....C.....C.....A.....A.....2396																																																											
MAC.US.x.251.BK28A.....C.....A.....A.....T.....G.....A.....C.....T.....G.....T.....G.....C.....C.....A.....A.....2385																																																											
MAC.US.x.EMBL_3A.....A.....C.....A.....A.....T.....G.....A.....C.....T.....G.....T.....G.....C.....C.....A.....A.....1883																																																											
SMM.US.x.SL92BA.....C.....G.....G.....A.....C.....C.....T.....G.....T.....G.....C.....C.....A.....A.....2412																																																											
SMM.US.x.H9C.....G.....G.....C.....A.....A.....C.....T.....G.....T.....G.....C.....C.....A.....A.....2412																																																											
SMM.US.x.PGM53C.....G.....G.....C.....A.....A.....C.....T.....G.....T.....G.....C.....C.....A.....A.....2412																																																											
SMM.US.x.SIVsmH635F.L3C.....G.....G.....C.....A.....A.....C.....T.....G.....T.....G.....C.....C.....A.....A.....2412																																																											
STM.US.x.STMT.....C.....G.....C.....A.....C.....G.....T.....G.....C.....C.....A.....A.....2412																																																											
SAB.SN.x.SAB1CA.....A.....C.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....2439																																																											
TAN.UG.x.TAN1A.....A.....C.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....2439																																																											
VER.DE.x.AGM3A.....A.....C.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1811																																																											
VER.KE.x.9063A.....A.....C.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....2316																																																											
VER.KE.x.AGM155A.....A.....C.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....2299																																																											
VER.KE.x.TYO1A.....A.....C.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....2304																																																											
COL.CM.x.CGU1A.....A.....C.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1964																																																											
GRV.ET.x.GRI.677A.....A.....C.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....2244																																																											
MND.2.CM.98.CM16A.....A.....C.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....2097																																																											
DRL.x.x.FAOA.....C.....T.....G.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1659																																																											
RCM.GA.x.GAB1A.....C.....T.....G.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1574																																																											
RCM.NG.x.NG411A.....C.....T.....G.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1577																																																											
MND.1.GA.x.MNDGB1A.....C.....T.....G.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1792																																																											
MND.2.GA.x.M14A.....C.....T.....G.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....2021																																																											
MND.2.x.x.5440A.....C.....T.....G.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1656																																																											
MNE.US.x.MNE027A.....C.....T.....G.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1877																																																											
LST.CD.88.447A.....C.....T.....G.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1354																																																											
LST.CD.88.485A.....C.....T.....G.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1354																																																											
LST.CD.88.524A.....C.....T.....G.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1348																																																											
LST.KE.x.ho7A.....C.....T.....G.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....2441																																																											
GSN.CM.99.CN166A.....A.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1829																																																											
GSN.CM.99.CN71A.....A.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1829																																																											
MON.CM.99.L1A.....A.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1816																																																											
MON.NG.x.NG1A.....A.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1816																																																											
MUS.1.CM.01.1085A.....A.....A.....A.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....397																																																											
MUS.1.CM.01.CM1239A.....A.....A.....A.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1824																																																											
MUS.2.CM.01.CM1246A.....A.....A.....A.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1814																																																											
MUS.2.CM.01.CM2500A.....A.....A.....A.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1907																																																											
DEN.CD.x.CD1G.....A.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1866																																																											
DEB.CM.99.CM40A.....A.....A.....A.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1949																																																											
DEB.CM.99.CM5A.....A.....A.....A.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1824																																																											
TAL.CM.00.266G.....G.....G.....C.....C.....A.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....2076																																																											
TAL.CM.01.8023G.....G.....G.....C.....C.....A.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1571																																																											
SUN.GA.98.L14C.....C.....C.....C.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....1872																																																											
SYK.KE.x.KE51A.....A.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....2437																																																											
SYK.KE.x.SYK173A.....A.....G.....C.....T.....G.....C.....A.....A.....A.....A.....G.....G.....A.....C.....2235																																																											

H1B.FR.83.HXB2	AGAGCAGACC.....AGAGCCAAACAGCCCCACAGAAAGAGAGCTTCAGGTTGCGGGTAGAGACAAC.....	1757
H101_AE.TH.90.CM240	Q E Q T P R A N S P T R E S F R S G V E D N T	1720
H102_AG.NG.x.IBNGG A T T C A A G G G A T G A T T	1421
H103_AB.RU.97.KAL153_2T C A A T G A T G A T T	1567
H104_cpx.CY.94.CY032	..A G.....G C C G A T A G A A G G A A A G	1644
H1A1.UG.85.U455C A T T G A T G A T T	2197
H1B.US.90.WB.AU160T A.....	1575
H1C.ET.86.ETH2220G A G T C T A G C A G A C C A A C C C C C A C C A G A	1718
H1D.CD.84.84ZR085C A G C T C G G T T G A G G A T A C	1535
H1F1.BE.93.VI850C G T C A A A G T	1604
H1G.SE.93.SE6165	..A G A.....A T C A C G T C A A G T G	1548
H1H.CF.90.056G.....	1515
H1I.SE.93.SE7887C C G C T C A A G T	1394
H1K.CM.96.MP535G T C A A G T	1771
H1N.CM.95.YBF30	..C A C A A A G G A A.....C G C T T A T G T C A A G	2232
H1O.BE.87.AN170A C C G.....	2207
H1Q.CM.91.MVP5180	..A A C A G T.....G T C C T T	1637
CPZ_CD.90.ANT	..A G A G G A A G T A G T G.....T C A T T A C A A G C A	1775
CPZ_CM.01.SIVcpz.CAM13	..A G A.....C T A T G A C C A A A G A	1746
CPZ_CM.05.SIVcpz.EK505	..C A C A A A A G G A A.....T A T T A T G A C A	1738
CPZ_CM.05.SIVcpz.LB7C G A T T G T C A C C A	1729
CPZ_CM.05.SIVcpz.MB66	..A G.....A A.....T A T A T G A T C G A T C A	1731
CPZ_CM.05.SIVcpz.MT145	..A G.....G A A.....G A C T T G C T T A T G T A C A G G	603
CPZ_CM.98.CAM5	..G G A A.....G A C T T A T G T C A A A C A T	1906
CPZ_CM.98.CAM5	..G G A A.....G A C T T A T G T C A A A C A T	2262
CPZ.GA.88.GAB1	..A A.....A G T A T G A C C A A G A G A	1582
CPZ.GA.88.GAB2	..A G A G G T.....G A G T C C C G A C A A G C C A	1815
CPZ.TZ.01.TAN1	..A C C.....A G A G A G G A A.....	2264
CPZ.US.85.CPZUS	..A A G A A.....G C T T A T G A T G A C A A	2551
H2A.DE.x.BEN	..T C C A.....T A G G C C T A C C A C A G A C T C C	2544
H2A.GW.x.ALI	..T T G A G G T T A C C A C A G A G G G A A A A G C A G A G A G A G A G	1996
H2A.SN.x.ST	..T G C C A.....A T G A G G C T C A C C A C A G A C C C C C	2522
H2B.Cl.x.EHO	..T C C A G G C.....C A G G G A T T C C A C T G C C C C	1878
H2B.GH.86.D205	..T C C A.....T G T A G G T C A C C A T C T G C	1906
H2G.Cl.x.AB196	..T C C A.....T T A G G T C T C A C C A G A C C C	2262
H2U.FR.96.I2034	..T G C C A.....T G A G G C T C C C A C A G C C C	1582
MAC.US.x.239	..T G C T C A.....T G A T A G G C C T G C C A C T G C C C C C	1815
MAC.US.x.251_1A11	..H G.....S H Q S G A D N C	2456
MAC.US.x.251_BK28	..T G C T C A.....T G A T A G G C C T G C C A C T G C C C C C	2455
MAC.US.x.EMBL_3	..T G C T C A.....T G A T A G G C C T G C C A C T G C C C C C	2444
SMM.SL.92.SL92B	..T G C T C A G T G C A T C A G G.....T G C A A T G T C C C C C	1942
SMM.US.x.H9	..T C A G C A A A T C A T T.....G C G T G C C C C C	1899
SMM.US.x.PGM53	..T G C C A.....T G A T G A T G A G G C T G C C A C T G C C C C C	1941
SMM.US.x.SIVsmH635F_L3	..T G C C A.....G A T G T A G G C C T C A C C A C T G C C C C C	2385
STM.US.x.STM	..T G C C A.....A T A T A G G C C T G C C A C T G C C C T C C	2471
SAB.SN.x.SAB1C	..C C T C T C A G A C C A A C C T.....C C A T G G A G G T T C A A G	2115
TAN.UG.x.TAN1	..C A A G C G G A G C C G T.....T C G A G G T C C C G A T C C T G C G G A	2570
VER.DE.x.AGM3	..C T C T T G G G G T G A G C C A A T G C.....T C C G A G T C C T G C G G A C C A A G	2350
VER.KE.x.9063	..C C C T C A T G G A G T G G A C C A G G C G C.....T C C G A C C C C	1879
VER.KE.x.AGM155	..C C C T C T T G G G G C G G A C C A A G T G C.....T C C G A C C A	2384
VER.KE.x.TYO1	..C T C T C T T G G A G C G G A C C A G T G C T.....T C C G A G G C A C	2376
COL.CM.x.CGU1	..C A G.....G A C A C A G T T G T T G G A C A A.....G C C C C C C	2381
GRV.ET.x.GRI_677G G G A T G C C T T A C C A A G G C.....	1964
MND_2.CM.98.CM16	..G C T C.....T G C C A T G C C C C A T G C C C C A G G A T A C C C A G G A G A T G C T C A G G A G A T G C T C	2321
DRL.x.x.FAO	..C T T C.....T T G A T C C C T C T C T.....G G G A T G G A G A C C A T T C G A C C	1730
RCM.GA.x.GAB1	..C T T C.....T T G A G A C C T A C.....T G T C C G	1645
RCM.NG.x.NG411T G C C A T G C C A T G.....G G	1660
MND_1.GA.x.MNDGB1T G C C A T G C C A T G.....G G	1844
MND_2.GA.x.M14T G C C A T G C C A T G.....G G	2094
MND_2.x.x.5440T G C C A T G C C A T G.....G G	1727
MNE.US.x.MNE027	..T G C C A.....T G C C A T G C C A T G.....G G	1936
LST.CD.88.447	..T G G G G G A A T G C.....G A G T G T C A T G T C C C C C	1425
LST.CD.88.485	..T G G G G G A A T G C.....G A G T G T C A T G T C C C C C	1425
LST.CD.88.524	..T G G G G G C T G C.....G A G T G T T G G C C C	1422
LST.KE.x.lho7	..T G A G G.....T G G G A G.....A T T C C A T C A G C T C C C C	2506
GSN.CM.99.CN166	..T A C T T C.....T G C T G C C A C A G C C G G A	1860
GSN.CM.99.CN71	..T A C T T C.....T G C T G C C C T C A G C C G G A	1860
MON.CM.99.L1	..T T C T G T G C A G C C C A G C G C.....C G A T T G.....	1854
MON.NG.x.NG1	..T T C C A G T A C T C C T T G C.....C C A C.....	462
MUS_1.CM.01.1085	..C T C T C T T G C C A C T G T C G C A G A A G.....G G G A A G G T C A A C T C T C T C C A G G T	1891
MUS_1.CM.01.CM1239	..C T G C T T T T A C C A C G C C G C A G A G A G.....G G A G A C G G T C A A C T C T C C C C C C G T	1881
MUS_2.CM.01.CM1246	..T T C A T C G T C A T G C C T A C G.....C C A G A T A.....C C G G G A C T C T C C C C C A C T G T A G G G C C A A G G A	1983
MUS_2.CM.01.CM2500	..T T C A T C G T C A T G C C T C A G T C C A G A T A.....C C G G G A C T C T C C C C C G T G T G G T C C A C G G A A G G A G A A	1948
DEN.CD.x.CD1	..C C C T C A G A T G A A T C A G.....T T T G C T G G A G C.....T G A G G T	2026
DEB.CM.99.CM40	..A G A T T G G A A G T A C A T C.....G G A G A G A G.....T C G T A A T C A	1977
DEB.CM.99.CM5	..A G G T T G G A A A T A C A T C.....G A G A G A A G.....T C G T T T A T C A A	1971
TAL.CM.00.266	..G A T T T T.....A G T G G T G C A.....T G A G A G.....A T.....	2132
TAL.CM.01.8023	..G A T T C.....A G T G G T G C A.....C C A A G G A T.....	1627
SUN.GA.98.L14	..T C T G C C T T.....T G C T C C G C C.....	1677
SYK.KE.x.KE51	..T T G.....G G A G A C C C A C A G C C C.....C A G G G A A A T T A T G C T C A C C C A A C A G A A G C A C A T A G C A A T G C C C C C G C T C A G C C A G C T C C	1997
SYK.KE.x.SYK173	..T G G.....G G A G C C A T C T C T C C.....T C T G G A	2339

PLV Complete Genomes

Table with 10 columns: Accession ID, Nucleotide sequence, Accession ID, Nucleotide sequence, Accession ID, Nucleotide sequence, Accession ID, Nucleotide sequence, Accession ID, Nucleotide sequence. Rows include various HIV-1 strains such as H1B.FR.83.HXB2, MAC.US.x.239, and MND.2.CM.98.CM16.

H1B.FR.83.HXB2
H101_AE.TH.90.CM240
H102_AG.NG.x.IBNG
H103_AB.RU.97.KAL153_2
H104_cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WE.AU160
H1C.ET.86.ETH.220
H1D.CD.84.84ZR085
H1F1.BE.93.VI850
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.ANT170
H1P.CM.91.MVP5180
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzLB7
CPZ.CM.05.SIVcpzMB66
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CAM3
CPZ.CM.98.CAM3
CPZ.GA.88.GAB1
CPZ.GA.88.GAB2
CPZ.TZ.01.TANI
CPZ.US.85.CPZUS
H2A.DE.x.BRN
H2A.GW.x.ALI
H2A.SN.x.ST
H2B.CL.x.EHO
H2B.GH.86.D205
H2G.CL.x.ABT96
H2U.FR.96.12034
MAC.US.x.239
Pol
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMBL_3
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F_L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.UG.x.TANI
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.ACM1155
VER.KE.x.TYO1
COL.CM.x.CGU1
GRV.ET.x.GRI_677
MND_2.CM.98.CM16
DRL.x.FAO
RCLM.GA.x.GAB1
RCLM.NG.x.NG411
MND_1.GA.x.MNDGB1
MND_2.GA.x.M14
MND_2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.Iho7
GSN.CM.99.CM166
GSN.CM.99.CM71
MON.CM.99.L1
MON.NG.x.NG1
MUS_1.CM.01.1085
MUS_1.CM.01.CM1239
MUS_2.CM.01.CM1246
MUS_2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

Genome alignment data showing nucleotide sequences (A, C, G, T) for various HIV-1 strains. The sequences are aligned in columns, with gaps represented by dashes. The alignment shows high conservation across different strains, with some minor variations in specific regions.

H1B.FR.83.HXB2	GCAATGGCTAGTATTTAACTGCCACTGTAGTACAAAGAAATAGTAGCCAGCTGTGATAAATGCAGCTAAAAGGAGAACCCATGCATGGACAAGTAGACTGTAGTCAGGAATATGGCAACTAGATTGTACACATTAGAAGGAAAATTTCTGGTAGCAGT	4459
Pol	A M A S D F N L P P V V A K E I V A S C D K C Q L K G E A M H G Q V D C S P G I W Q L D C T H L E G K V I L V A V	
H101_AE.TH.90.CM240	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4033
H102_AG.NG.x.IBNG	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3984
H103_AB.RU.97.KAL153_2	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3685
H104_cpx.CY.94.CY032	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3825
H1A1.UG.85.U455	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3905
H1B.US.90.WEAU160	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4458
H1C.ET.86.ETH220	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3841
H1D_CD.84.84ZR085	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3981
H1F1_BE.93.VI850	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3799
H1G_SE.93.SE6165	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3856
H1H_CF.90.056	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3806
H1I_SE.93.SE7887	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3773
H1K_CM.96.MP535	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3655
H1N_CM.95.YBF30	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4053
H1O_BE.87.ANT170	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4514
H1O_CM.91.MVP5180	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4489
CPZ_CD.90.ANT	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3898
CPZ_CM.91.SIVcpzCAM13	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4039
CPZ_CM.05.SIVcpzEK505	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4031
CPZ_CM.05.SIVcpzLB7	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4014
CPZ_CM.05.SIVcpzMB66	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4005
CPZ_CM.05.SIVcpzMT145	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4001
CPZ_CM.98.CAM3	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3873
CPZ_CM.98.CAM3	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4161
CPZ_GA.88.GAB1	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4520
CPZ_GA.88.GAB2	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3852
CPZ_TZ.01.TANI	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4103
CPZ_US.85.CPZUS	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4519
H2A.DE.x.BRN	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4841
H2A.GW.x.ALI	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4834
H2A.SN.x.ST	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4286
H2B.CL.x.EHO	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4812
H2B.GH.86.D205	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4814
H2G.CL.x.AB796	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4168
H2U.FR.96.12034	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4520
MAC.US.x.239	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4758
Pol	K E L V F K F G L P R I V A R Q I V D T C D K C H Q O K G E A I H G Q A T A G S D L G C T W Q M D C T H L E G K A I T I V A V	
MAC.US.x.251_1A11	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4757
MAC.US.x.251_BK28	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4734
MAC.US.x.EMBL_3	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4226
SMM.SL.92.SL92B	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4189
SMM.US.x.H9	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4231
SMM.US.x.PGM53	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4675
SMM.US.x.SIVsmH635F_L3	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4761
STM.US.x.8TM	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4405
SAB.SN.x.SAB1C	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4857
TAN.UG.x.TANI	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4658
VER.DE.x.AGM3	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4181
VER.KE.x.9063	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4686
VER.KE.x.ACM155	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4678
VER.KE.x.TYO1	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4683
COL_CM.x.CGU1	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4222
GRV.ET.x.GRI_677	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4632
MND_2_CM.98.CM16	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4460
DRL.x.x.FAO	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4026
RCLM.GA.x.GAB1	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3938
RCLM.NG.x.NG411	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3953
MND_1_GA.x.MNDGB1	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4146
MND_2_GA.x.M14	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4387
MNE.US.x.MNE027	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4020
LST.CD.88.447	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4226
LST.CD.88.485	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3736
LST.CD.88.524	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3736
LST.KE.x.lho7	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4820
GSN_CM.99.CM166	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4187
GSN_CM.99.CM71	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4169
MON_CM.99.L1	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4162
MON.NG.x.NG1	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	2755
MUS_1_CM.01.1085	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4149
MUS_1_CM.01.CM1239	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4139
MUS_2_CM.01.CM11246	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4247
MUS_2_CM.01.CM2500	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4206
DEN_CD.x.CD1	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4349
DEB_CM.99.CM40	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4265
DEB_CM.99.CM5	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4259
TAL_CM.00.266	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4425
TAL_CM.01.8023	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	3923
SUN.GA.98.L14	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4819
SYK.KE.x.KE51	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4314
SYK.KE.x.SYK173	-----T-----T-----A-----G-----A-----AT-----G-----G-----G-----GT-----C-----T-----A-----A-----	4659

H1B.FR.83.HXB2
Pol
H101_AE.TH.90.CM240
H102_AG.NG.x.IBNG
H103_AB.RU.97.KAL153_2
H104_cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WEAU160
H1C.ET.86.ETH2220
H1D.CD.84.84ZR085
H1F1.BE.93.VI850
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.ANT170
H1O.CM.91.MVP5180
CPZ_CD.90.ANT
CPZ_CM.01.SIVcpzCAM13
CPZ_CM.05.SIVcpzEK505
CPZ_CM.05.SIVcpzLB7
CPZ_CM.05.SIVcpzMB66
CPZ_CM.05.SIVcpzMT145
CPZ_CM.98.CAM3
CPZ_CM.98.CAM3
CPZ_GA.88.GAB1
CPZ_GA.88.GAB2
CPZ_TZ.01.TANI
CPZ.US.85.CPZUS
H2A.DE.x.HBN
H2A.GW.x.ALI
H2A.SN.x.ST
H2B.CL.x.EHO
H2B.GH.86.D205
H2G.CL.x.ABT96
H2U.FR.96.12034

MAC.US.x.239
Pol
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMBL_3
SMM.SL.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F_L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.UG.x.TANI
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.ACM155
VER.KE.x.TYO1
COL.CM.x.CGU1
GRV.ET.x.GRI_677
MND_2.CM.98.CM16
DRL.x.x.FAO
RCLM.GA.x.GAB1
RCM.NG.x.NG411
MND_1.GA.x.MNDGB1
MND_2.GA.x.M14
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.lho7
GSN.CM.99.CM166
GSN.CM.99.CM71
MON.CM.99.L1
MON.NG.x.NG1
MUS_1.CM.01.1085
MUS_1.CM.01.CM1239
MUS_2.CM.01.CM1246
MUS_2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

Table of HIV sequence alignments. Header: TCATGTAGCCAGTGGATATATAGAACGAGAAGTTATCCAGCAGAAACAGGCGAGAAACAGCATATTTTCTTTAAATAGCAGGAAGATGGCCAGTAAAAACAATA... CATACTGACAATGGCAGCAATTTCCACGGTGCTACGGTTAGGCCCGCTGTGGTGGG 4626. Rows list various HIV strains and their corresponding nucleotide sequences aligned to a reference sequence.

	GGGATTGGGGGTACAGTCAGGGGAAAAG...	..ATAGTAGACATAATAGCAACAGACATACAAACTAAGAATTCACAAAACAAATTT..	..ACAAAAATTCAAAATTTTCGGGTTTATTACAGGCACAGCAAAATCCACTTTGGAAAGGACCAGCAAAAGCTCTCTGGAA	4960
H1B.FR.83.HXB2	G I G G Y S A G E R	I V D I I A T D I Q T K E L Q K Q I	T K I Q N F R V Y Y R D S R N P L W K G P A K L W K	4534
H101_AE.TH.90.CM240		A - T - T - T - G - - T -	G - C - A - - - - - - - - -	4534
H102_AG.NG.x.IBNG		- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4185
H103_AB.RU.97.KAL153_2		A - - - - - - - - - - - - - - -	A - G - A - - - - - - - - -	4186
H104_cpx.CY.94.CY032		- - - - - - - - - - - - - - -	G - A - A - G - - - - - - - -	4326
H1A1.UG.85.U455	A - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	G - C - A - - - - - - - - -	4406
H1B_US.90.WEAU160		A - T - - - - - - - - - - - - -	G - C - - - - - - - - - - - -	4499
H1C_ET.86.ETH.2220		- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4352
H1D_CD.84.84ZR085		A - T - T - T - G - C - C - C -	A - - - - - - - - - - - - -	4482
H1F1_BE.93.VI850		- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4300
H1G_SE.93.SE6165	A - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4357
H1H_CF.90.056		A - - - - - - - - - - - - - - -	G - C - G - - - - - - - - -	4307
H1I_SE.93.SE7887	A - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	G - C - A - - - - - - - - -	4274
H1K_CM.96.MP535		- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4156
H1N_CM.95.YBF30	C - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	G - A - - - - - - - - - - -	4554
H1O_BE.87.ANT70	C - A - G - G - - - - - - - -	A - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	5015
H1O_CM.91.MVP5180	C - A - G - C - - - - - - - -	A - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4990
H1O_CD.90.ANT	CAC - T - AC - G - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4539
CPZ_CM.01.SIVcpzCAML3	C - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4540
CPZ_CM.05.SIVcpzEK505	A - C - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4532
CPZ_CM.05.SIVcpzLB7	A - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4515
CPZ_CM.05.SIVcpzMB66	A - C - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4506
CPZ_CM.05.SIVcpzMT145	C - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4502
CPZ_CM.98.CAM5	C - A - A - G - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4662
CPZ_CM.98.CAM5	C - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	5021
CPZ_GA.88.GAB1	C - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4353
CPZ_GA.88.GAB2	C - A - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	4604
CPZ_TZ.01.TANI	A - C - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	5020
CPZ_US.85.CPZUS	A - T - C - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	5342
H2A.DE.x.BBN	A - A - ATATG - CCC - CA -	C - A - CA - G - TA - C - ACA - G	TAC - TTCC - C - G - A - A - T -	T - A - C - C - - - - - - - - -
H2A.GW.x.ALI	A - A - ATATG - CCC - CA -	C - CA - CA - T - G - CT - C - ACA - G	TAC - TTCC - C - C - A - A - TTG -	T - A - - - - - - - - - - - -
H2A.SN.x.ST	A - A - ATATG - CCC - CA -	C - A - CA - T - GG - CA - TG - ACAGG -	TAC - TTCC - C - GC - A - A - T -	T - A - - - - - - - - - - - -
H2B.CL.x.EHO	A - A - ATATG - CCC - T - CA -	C - CA - G - A - T - ACA - G	TAC - TTCC - C - C - A - A - T -	T - C - - - - - - - - - - - -
H2B.GH.86.D205	A - A - ATATG - CCC - T - CA -	C - TA - G - A - C - GCA - G	TAC - GTTC - C - GC - A - A - TT -	T - A - C - A - - - - - - - -
H2G.CL.x.ABT96	A - A - ATATG - CCC - T - CA -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
H2U.FR.96.12034	A - A - ATATG - C - C - T - CA -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
MAC.US.x.239	A - A - ATATG - C - C - CA -	R - A - TA - G - CA - T - ACA - G	TAC - TTCA - TC - A - AC -	T - - - - - - - - - - - - -
Pol	G G I G D M T P A E	R L A I N M I T	E I Q O S K	N S K F K N F R V Y Y R E G R D Q L W K G P G E L L W
MAC.US.x.251_1A11	A - CA - ATATG - C - CA -	T - A - TA - G - CA - T - ACA - G	TAC - TTCA - TC - A - AC -	T - - - - - - - - - - - - -
MAC.US.x.251_BK28	A - A - ATATG - C - CA -	T - A - TA - G - CA - T - ACA - G	TAC - TTCA - TC - A - AC -	T - - - - - - - - - - - - -
MAC.US.x.EMBL_3	A - A - ATATG - C - C - CA -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
SMM.SL.92.SL92B	A - A - ATATG - C - C - T - CA -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
SMM.US.x.H9	A - A - ATATG - C - C - CA -	T - CA - T - G - TA - C - ACA - G	TAC - TTCCA - TC - A - A - T -	T - A - - - - - - - - - - - -
SMM.US.x.PGM53	A - A - ATATG - C - C - CA -	T - CA - T - G - CA - C - ACA - G	TAC - TTCCA - TC - A - A - T -	T - A - - - - - - - - - - - -
SMM.US.x.SIVsmH635F_L3	A - A - ATATG - C - C - CA -	T - CA - T - G - CA - C - ACA - G	TAC - TTCCA - TC - A - A - T -	T - A - - - - - - - - - - - -
STM.US.x.STM	A - A - ATATG - CCC - CA -	T - TA - T - G - CA - C - ACA - G	TAC - TTCCA - TC - A - A - T -	T - CA - - - - - - - - - - - -
SAB.SN.x.SABIC	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
TAN.UG.x.TANI	A - A - TATG - CCC - CA - G - G -	C - A - A - T - G - A - C - C - AC - G -	TAC - C - CA - C - CAA -	CA - TC - AA - C - C - A - AG - A - G - G - G -
VER.DE.x.AGM3	A - A - TA - CAC - G - CA - G -	T - A - CA - T - G - TA - T - AT - G -	TTAC - C - CC - CCA -	CA - TT - A - A - C - C - A - AG - G - G - TG - C -
VER.KE.x.9063	A - A - TA - CAC - CA - G -	C - A - TA - T - G - A - C - AT - G -	CTAC - AC - CCA -	CA - TT - A - A - C - C - A - AG - G - G - G -
VER.KE.x.ACM155	A - A - CTA - CA - T - CA - G -	C - A - A - T - G - A - C - AT - G -	TC - CACTC - CCA -	CA - TTG - A - A - C - C - A - AG - G - G - G -
VER.KE.x.TYO1	A - A - AC - G - C - T - CA - G -	C - A - TA - T - G - C - AT - G -	TAC - C - T - CCA -	CA - TT - A - A - C - C - A - AG - G - G - G -
COL.CM.x.CGU1	AT - A - - - - - - - - - - - - -	CA - - - - - - - - - - - - -	A - GC - AT - T - AT - G - T - AC - A - CC - AC -	TT - CAAAATCA - T - A - AA - CC - CGG - AA - GCA - CAGG - GAG - C -
GRV.ET.x.GRI_677	A - A - TAT - TC - AGT - CA -	T - G - TA - T - GC - A - C - AC - G -	CTA - TACTC - C - C - CA -	TTG - AA - C - C - A - AG - T - G - G - G - G -
MND_2_CM.98.CM16	A - A - AA - C - CA -	ATAG - ACATT - GC - A - GA - C - AC - C -	TAC - C - C - T - CA -	- - - - - - - - - - - - - - -
DRL.x.FAO	A - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
RCLM.GA.x.GAB1	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
RCLM.NG.x.NG411	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
MND_1_GA.x.MNDGB1	AC - A - ATG - CA - A - G -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
MND_2_GA.x.M14	A - A - A - C - T - A -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
MNE.US.x.MNE027	A - A - ATATG - C - CA -	G - T - A - T - G - CA - T - ACA - G	TAC - TTCCA - TC - A - AC -	T - - - - - - - - - - - - -
LST.CD.88.447	AC - A - CT - TCAC - CA - G -	T - TA - TA - T - G - TAATG -	T - G - AC - T - T -	T - - - - - - - - - - - - -
LST.CD.88.485	AC - A - CTTTCAC - CA - G -	T - TA - TA - T - G - TAATG -	T - G - AC - T - T -	T - A - - - - - - - - - - - -
LST.CD.88.524	AC - G - TATCAC - CA - G -	T - A - CA - T - G - CAATG -	T - T - G - AC - T - CA -	TT - A - - - - - - - - - - - -
LST.KE.x.ho7	AC - A - CTT - CAC - CA - G -	T - TA - TA - T - G - TAATG -	T - G - AC - T - TC -	TT - A - - - - - - - - - - - -
GSN.CM.99.CM166	A - A - ATATGT - CC - TACA -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
GSN.CM.99.CM71	A - A - ATATGT - CC - TACA -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
MON.CM.99.L1	A - A - ATATGGCCC - T - TA - C -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
MON.NG.x.NG1	A - A - ATAT - - - - C - NG1 -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
MUS_1_CM.01.1085	AG - G - ACCTATCAAAT - CT - G - G -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
MUS_1_CM.01.CM1239	A - A - ACCTAGCA - T - CT - G -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
MUS_2_CM.01.CM1246	A - A - ATATG - CC - TACA - G -	T - TA - T - GC - TAC - AC - G -	TAC - C - C - TC - AA -	T - C - - - - - - - - - - - -
MUS_2_CM.01.CM2500	A - A - ATATG - C - CACA - G -	T - CA - GC - TAT - A - G -	TAC - C - A - CAC - AA -	- - - - - - - - - - - - - - -
DEN.CD.x.CD1	AC - A - ATATG - CA - CA -	TAC - TAACATGC - ATA - CATTAGA - AT -	CAAC - T - ACC -	C - - - - - - - - - - - - -
DEB.CM.99.CM40	A - A - ATATGTC - C - GATA -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
DEB.CM.99.CM5	A - A - ATATGGCA - ATA - G -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
TAL.CM.00.266	AT - A - ATATGT - CC - CA - GC -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -	- - - - - - - - - - - - - - -
TAL.CM.01.8023	AT - A - ATATGT - C - CA - GC -	C - A - A - G - CTAC - AC - G -	TAC - C - C - ACC -	C - CA - TT - TCG - AA - C -
SUN.GA.98.L14	AC - A - AATG - CA - CA - G -	G - T - TA - CA - G - TAATG -	T - G - AC - T - C - G -	AT - A - T - - - - - - - - - - -
SYK.KE.x.KE51	A - A - ACT - CAC - CA - T - G -	TATA - TA - GC - TAC - AC -	CTAC - A - TAC - C - C -	CA - T - TCG - AAA - A -
SYK.KE.x.SYK173	A - A - ATATCAC - CA - T - G -	TACA - TA - T - GC - TAT - AC -	TTAC - C - AATAC - CT -	C - T - TCG - A - A - C -

H1B.FR.83.HXB2
Vif
H101.AE.TH.90.CM240
H102.AG.NG.x.IBNG
H103.AB.RU.97.KAL153_2
H104.cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WE.AU160
H1C.ET.86.ETH2220
H1D.CD.84.84ZR085
H1F1.BE.93.V1850
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.AN170
H1O.CM.91.MVP5180
CPZ.CD.90.ANT
CPZ.CM.01.SIV.cpz.CAM13
CPZ.CM.05.SIV.cpz.EK505
CPZ.CM.05.SIV.cpz.LB7
CPZ.CM.05.SIV.cpz.MB66
CPZ.CM.05.SIV.cpz.MT145
CPZ.CM.98.CAM5
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.GAB2
CPZ.TZ.01.TAN1
CPZ.US.85.CPZ.US
H2A.DE.x.BEN
H2A.GW.x.ALI
H2A.SN.x.ST
H2B.CL.x.EHO
H2B.GH.86.DP05
H2C.KI.x.BT96
H2U.FR.96.12034

AGGTGAAGGGGCAGTAGTAATACAAGATAAGT...GACATAAAAGTAGTCCCAAGAAGAAAAGCAAAGATCATTAGGGATTATGGAAACACAGATGGCAGGTGATGATTGTGGCAAGTAGACAGGATGAGGATTAG
Vif start Pol p31 integrase end
G E G A V V I Q D N S D I K V P R R K A K I I R D Y G K Q M A G D D C V A S R Q D E D *
...
H2U.FR.96.12034

MAC.US.x.239
Vif
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMBL_3
SMM.SL.92.SL92B
SMM.US.x.TAN1
SMM.US.x.PGM53
SMM.US.x.SIV.smH635F_L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.US.x.TAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TYO1
COL.CM.x.CGU1
GRV.ET.x.GRI.677
MND.2.CM.98.CM16
DRL.x.x.FAO
RCM.GA.x.GAB1
RCM.NG.x.NG411
MND.1.GA.x.MNDGB1
MND.2.GA.x.M14
MND.2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KI.x.Iho3
GSN.CM.99.CM166
GSN.CM.99.CN71
MON.CM.99.L1
M.NG.x.NG1
MUS.1.CM.01.1085
MUS.1.CM.01.CM1239
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

MAC.US.x.239
Vif
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMBL_3
SMM.SL.92.SL92B
SMM.US.x.TAN1
SMM.US.x.PGM53
SMM.US.x.SIV.smH635F_L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.US.x.TAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TYO1
COL.CM.x.CGU1
GRV.ET.x.GRI.677
MND.2.CM.98.CM16
DRL.x.x.FAO
RCM.GA.x.GAB1
RCM.NG.x.NG411
MND.1.GA.x.MNDGB1
MND.2.GA.x.M14
MND.2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KI.x.Iho3
GSN.CM.99.CM166
GSN.CM.99.CN71
MON.CM.99.L1
M.NG.x.NG1
MUS.1.CM.01.1085
MUS.1.CM.01.CM1239
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

Table with columns for sample names (e.g., H1B.FR.83.HXB2, MAC.US.x.239) and their corresponding nucleotide sequences. The sequences are aligned to a reference sequence starting with 'AACATGGAAAGTTT...'. The table includes various HIV strains and sequences from different regions and time points.

H1B.FR.83.HXB2	5558
Vif	
H101_AE.TH.90.CM240	5129
H102_AG.NG.x.IBNG	5083
H103_AB.RU.97.KAL153_2	4784
H104_cpx.CY.94.CY032	4924
H1A1.UG.85.U455	5004
H1B.US.90.WEAU160	5557
H1C.ET.86.ETH220	4950
H1D.CD.84.84ZR085	5080
H1F1.BE.93.VI850	4898
H1G.SE.93.SE6165	4955
H1H.CF.90.056	4905
H1J.SE.93.SE7887	4872
H1K.CM.96.MP535	4754
H1N.CM.95.YBF30	5152
H1O.BE.87.ANT70	5613
H1O.CM.91.MVP5180	5588
CPZ_CD.90.ANT	5015
CPZ_CM.01.SIVcpzCAM13	5141
CPZ_CM.05.SIVcpzEK505	5130
CPZ_CM.05.SIVcpzLB7	5113
CPZ_CM.05.SIVcpzMB66	5104
CPZ_CM.05.SIVcpzMT145	5100
CPZ_CM.98.CAM5	4972
CPZ_CM.98.CAM5	5260
CPZ_GA.88.GAB1	5622
CPZ_GA.88.GAB2	4951
CPZ_TZ.01.TANI	5220
CPZ_US.85.CPZUS	5618
H2A.DE.x.HB3	6071
H2A.GW.x.ALI	6064
H2A.SN.x.ST	5516
H2B.CL.x.EHO	6045
H2B.GH.86.D205	6047
H2G.CL.x.AB796	5398
H2U.FR.96.L2034	5546
MAC.US.x.239	
Vpx	
MAC.US.x.251_1A11	5984
MAC.US.x.251_BK28	5961
MAC.US.x.EMBL_3	5453
SMM.SL.92.SL92B	5422
SMM.US.x.H9	5458
SMM.US.x.PM153	5906
SMM.US.x.SIVsmH635F_L3	5988
STM.US.x.STM	5632
SAB.SN.x.SAB1C	5910
TAN.UG.x.TANI	5738
VER.DE.x.AGM3	5243
VER.KE.x.9063	5745
VER.KE.x.AGM155	5740
VER.KE.x.TYO1	5745
COL.CM.x.CGU1	5280
GRV.ET.x.GRI_677	5682
MND_2.CM.98.CM16	5681
DRL.x.x.FAO	5262
RCM.GA.x.GAB1	5174
RCM.NG.x.NG411	5204
MND_1.GA.x.MNDGB1	5226
MND_2.GA.x.M14	5608
MND_2.x.x.5440	5244
MNE.US.x.MNE027	5453
LST.CD.88.447	4777
LST.CD.88.485	4777
LST.CD.88.524	4774
LST.KE.x.lho7	5858
GSN.CM.99.CN166	5336
GSN.CM.99.CN71	5318
MON.CM.99.L1	5389
MON.NG.x.NG1	3999
MUS_1.CM.01.1085	5292
MUS_1.CM.01.CM1239	5314
MUS_2.CM.01.CM1246	5431
MUS_2.CM.01.CM2500	5384
DEN.CD.x.CD1	5495
DEB.CM.99.CM40	5327
DEB.CM.99.CM5	5321
TAL.CM.00.266	5517
TAL.CM.01.8023	5015
SUN.GA.98.L14	5908
SYK.KE.x.KE51	5357
SYK.KE.x.SYK173	5700

H1B.FR.83.HXB2	5558
Vit	
H101_AE.TH.90.CM240	5129
H102_AG.NG.x.IBNG	5083
H103_AB.RU.97.KAL153_2	4784
H104_cpx.CY.94.CY032	4924
H1A1.UG.85.U455	5004
H1B.US.90.WEAU160	5557
H1C.ET.86.ETH220	4950
H1D.CD.84.84ZR085	5080
H1F1.BE.93.VI850	4898
H1G.SE.93.SE6165	4955
H1H.CF.90.056	4905
H1J.SE.93.SE7887	4872
H1K.CM.96.MP535	4754
H1N.CM.95.YBF30	5152
H1O.BE.87.ANT70	5613
H1O.CM.91.MVP5180	5588
CPZ_CD.90.ANT	5015
CPZ_CM.01.SIVcpzCAM13	5141
CPZ_CM.05.SIVcpzEK505	5130
CPZ_CM.05.SIVcpzLB7	5113
CPZ_CM.05.SIVcpzMB66	5104
CPZ_CM.05.SIVcpzMT145	5100
CPZ_CM.98.CAM3	4972
CPZ_CM.98.CAM5	5260
CPZ_GA.88.GAB1	5622
CPZ_GA.88.GAB2	5622
CPZ_TZ.01.TANI	4951
CPZ_US.85.CPZUS	5220
H2A.DE.x.BRN	5618
H2A.GW.x.ALI	6238
H2A.SN.x.ST	5680
H2B.CL.x.EHO	6209
H2B.GH.86.D205	6211
H2G.CL.x.AB796	5363
H2U.FR.96.12034	5710
Vpx end	
MAC.US.x.239	6150
Vpx	
MAC.US.x.251_1A11	
MAC.US.x.251_BK28	6149
MAC.US.x.EMBL_3	6126
SMM.SL.92.SL92B	5467
SMM.US.x.H9	5586
SMM.US.x.PGM53	5624
SMM.US.x.SIVcmH635F_L3	6069
STM.US.x.STM	6154
SAB.SN.x.SAB1C	5798
TAN.UG.x.TANI	5910
VER.DE.x.AGM3	5738
VER.KE.x.9063	5243
VER.KE.x.ACM155	5745
VER.KE.x.TYO1	5740
COL.CM.x.CGU1	5745
GRV.ET.x.GRI_677	5280
MND_2.CM.98.CM16	5682
DRL.x.FAO	5819
RCM.GA.x.GAB1	5427
RCM.NG.x.NG411	5327
MND_1.GA.x.MNDGB1	5360
MND_2.GA.x.M14	5226
MND_2.x.x.5440	5746
MNE.US.x.MNE027	5382
LST.CD.88.447	5618
LST.CD.88.485	4777
LST.CD.88.524	4777
LST.KE.x.Jho7	4774
GSN.CM.99.CM166	5858
GMN.CM.99.CM71	5336
MON.CM.99.L1	5318
MON.NG.x.NG1	5389
MUS_1.CM.01.1085	3999
MUS_1.CM.01.CM1239	5292
MUS_2.CM.01.CM1246	5314
MUS_2.CM.01.CM2500	5431
DEN.CD.x.CD1	5384
DEB.CM.99.CM40	5495
DEB.CM.99.CM5	5327
TAL.CM.00.2023	5321
TAL.CM.01.8023	5517
SUN.GA.98.L14	5015
SYK.KE.x.KE51	5908
SYK.KE.x.SYK173	G 5358
	5700

Vpr start
 ATGGAACAAGCC.....CCAGAAGACCAAGGCCACAGGGAGCCACACATGAATGGACACTAGAGCTTTTAGAGGAGCTTAAGAATGAAGCTTTAG 5653
 Vif end
 W N K P E D Q G P Q R G S H T M N G H * E L L E L K N E A V R
 H101_AE.TH.90.CM240
 H102_AG.NG.x.IBNG
 H103_AB.RU.97.KAL153_2
 H104_cpx.CY.94.CY032
 H1A1.UG.85.U455
 H1B.US.90.WE.AU160
 H1C.ET.86.ETH2220
 H1D.CD.84.84ZR085
 H1F1.BE.93.VI850
 H1G.SE.93.SE6165
 H1H.CF.90.056
 H1I.SE.93.SE7887
 H1K.CM.96.MP535
 H1N.CM.95.YBF30
 H1O.BE.87.ANT170
 H1P.CM.91.MVP5180
 CPZ_CD.90.ANT
 CPZ_CM.01.SIVcpzC.AM13
 CPZ_CM.05.SIVcpzE.K505
 CPZ_CM.05.SIVcpzL.B7
 CPZ_CM.05.SIVcpzM.B66
 CPZ_CM.05.SIVcpzMT.145
 CPZ_CM.98.C.AM5
 CPZ_CM.98.C.AM5
 CPZ_GA.88.GAB1
 CPZ_GA.88.GAB2
 CPZ_TZ.01.TAN1
 CPZ_US.85.CPZ.US
 H2A.DE.x.BEN
 H2A.GW.x.ALI
 H2A.SN.x.ST
 H2B.CL.x.EHO
 H2B.GH.86.DP05
 H2C.L.x.AB.T06
 H2U.FR.96.12034

MAC.US.x.239
 Vpr start
 M.....GAAGAAAGACCT...-A-TG--A--A--A--TGGG--GT-G-G--G-C-G-A-A-G-AG-A--T-A-A 6248
 P E A P R E P V E W R D E V I E K E A
 MAC.US.x.251_1A11
 MAC.US.x.251_BK28
 MAC.US.x.EMBL_3
 SMM.SL.92.SL92B
 SMM.US.x.H9
 SMM.US.x.BGM53
 SMM.US.x.SIVsmH635F_L3
 STM.US.x.STM
 SAB.SN.x.SAB1C
 TAN.UG.x.TAN1
 VER.DE.x.AGM3
 VER.KE.x.9063
 VER.KE.x.AGM155
 VER.KE.x.TYO1
 COL.CM.x.CGU1
 GRV.ET.x.GRI_677
 MND_2.CM.98.CM16
 DRL.x.x.FAO
 RCM.GA.x.GAB1
 RCM.NG.x.NG411
 MND_1.GA.x.MNDGB1
 MND_2.GA.x.M14
 MND_2.x.x.5440
 MNE.US.x.MNE027
 LST.CD.88.447
 LST.CD.88.485
 LST.CD.88.524
 LST.KE.x.Jho7
 GSN.CM.99.CM166
 GSN.CM.99.CM71
 MON.CM.99.L1
 MON.NG.x.NG1
 MUS_1.CM.01.1085
 MUS_1.CM.01.CM1239
 MUS_2.CM.01.CM1246
 MUS_2.CM.01.CM2500
 DEN.CD.x.CD1
 DEB.CM.99.CM40
 DEB.CM.99.CM5
 TAL.CM.00.269
 TAL.CM.01.8023
 SUN.GA.98.L14
 SYK.KE.x.KE51
 SYK.KE.x.SYK173

frameshift insert in HXB2
Vpr premature end (HXB2 only)

H1B.FR.83.HXB2
Vpr (frameshifted)
H101_AE.TH.90.CM240
H102_AG.NG.x.IBNG
H103_AB.RU.97.KAL153.2
H104_cpx.CY.94.CY032
H1A_FLUG.85.2545
H1B.US.90.WEAU160
H1C.ET.86.ETH2220
H1D.CD.84.84ZR085
H1F1.BE.93.V1850
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.AN170
H1O.CM.91.MVP5180
CPZ_CD.90.ANT
CPZ_CM.01.SIVcpzEAM13
CPZ_CM.05.SIVcpzEK505
CPZ_CM.05.SIVcpzLB7
CPZ_CM.05.SIVcpzMB66
CPZ_CM.05.SIVcpzMT145
CPZ_CM.98.CAM3
CPZ_CM.98.CAM5
CPZ_GA.88.GAB1
CPZ_GA.88.GAB2
CPZ_TZ.01.TAN1
CPZ_US.85.CPZUS
H2A.DE.x.BEN
H2A.GW.x.ALI
H2A.SN.x.ST
H2B.CL.x.EHO
H2B.GH.86.DM5
H2G.CL.x.AB196
H2U.FR.96.12034

ACATTTTCCTAGGATTGGCTCCATGGCTAGGGCAACATCTATGAACTTATGGGGAT.....ACTTGGCCAGGAGTGGAGCCATAAAGAATTCTGCAACAACCTGCTTTATCCATTTTCAGAAATTGGGTGCACATAGCAGA 5799
H F P R I W L H G L G Q H I Y E T Y G D T W A G V E A I I R I L Q Q L L F I H F Q N W V S T S R
H1B.FR.83.HXB2
Vpr (frameshifted)
H101_AE.TH.90.CM240
H102_AG.NG.x.IBNG
H103_AB.RU.97.KAL153.2
H104_cpx.CY.94.CY032
H1A_FLUG.85.2545
H1B.US.90.WEAU160
H1C.ET.86.ETH2220
H1D.CD.84.84ZR085
H1F1.BE.93.V1850
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.AN170
H1O.CM.91.MVP5180
CPZ_CD.90.ANT
CPZ_CM.01.SIVcpzEAM13
CPZ_CM.05.SIVcpzEK505
CPZ_CM.05.SIVcpzLB7
CPZ_CM.05.SIVcpzMB66
CPZ_CM.05.SIVcpzMT145
CPZ_CM.98.CAM3
CPZ_CM.98.CAM5
CPZ_GA.88.GAB1
CPZ_GA.88.GAB2
CPZ_TZ.01.TAN1
CPZ_US.85.CPZUS
H2A.DE.x.BEN
H2A.GW.x.ALI
H2A.SN.x.ST
H2B.CL.x.EHO
H2B.GH.86.DM5
H2G.CL.x.AB196
H2U.FR.96.12034

Tat exon 1 start

MAC.US.x.239
Vpr
Tat exon 1
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMBL_3
SMM.US.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F.L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.US.x.TAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TY01
COL.CM.x.CGU1
GRV.ET.x.GRI.677
MND.2.CM.98.CM16
DRL.x.x.FAO
RCM.GA.x.GAB1
RCM.NG.x.NG411
MND.1.GA.x.MNDGB1
MND.2.GA.x.M14
MND.2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.ho7
GSN.CM.99.CN166
GSN.CM.99.CN71
MON.CM.99.L1
MON.NG.x.NG1
MUS.1.CM.01.1085
MUS.1.CM.01.CM1239
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM45
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

MAC.US.x.239
Vpr
Tat exon 1
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMBL_3
SMM.US.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F.L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.US.x.TAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TY01
COL.CM.x.CGU1
GRV.ET.x.GRI.677
MND.2.CM.98.CM16
DRL.x.x.FAO
RCM.GA.x.GAB1
RCM.NG.x.NG411
MND.1.GA.x.MNDGB1
MND.2.GA.x.M14
MND.2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.ho7
GSN.CM.99.CN166
GSN.CM.99.CN71
MON.CM.99.L1
MON.NG.x.NG1
MUS.1.CM.01.1085
MUS.1.CM.01.CM1239
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM45
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

MAC.US.x.239
Vpr
Tat exon 1
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMBL_3
SMM.US.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F.L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.US.x.TAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TY01
COL.CM.x.CGU1
GRV.ET.x.GRI.677
MND.2.CM.98.CM16
DRL.x.x.FAO
RCM.GA.x.GAB1
RCM.NG.x.NG411
MND.1.GA.x.MNDGB1
MND.2.GA.x.M14
MND.2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.ho7
GSN.CM.99.CN166
GSN.CM.99.CN71
MON.CM.99.L1
MON.NG.x.NG1
MUS.1.CM.01.1085
MUS.1.CM.01.CM1239
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM45
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

.....ACTTGGCCAGGAGTGGAGCCATAAAGAATTCTGCAACAACCTGCTTTATCCATTTTCAGAAATTGGGTGCACATAGCAGA 5799
H F P R I W L H G L G Q H I Y E T Y G D T W A G V E A I I R I L Q Q L L F I H F Q N W V S T S R
MAC.US.x.239
Vpr
Tat exon 1
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMBL_3
SMM.US.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F.L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.US.x.TAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TY01
COL.CM.x.CGU1
GRV.ET.x.GRI.677
MND.2.CM.98.CM16
DRL.x.x.FAO
RCM.GA.x.GAB1
RCM.NG.x.NG411
MND.1.GA.x.MNDGB1
MND.2.GA.x.M14
MND.2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.ho7
GSN.CM.99.CN166
GSN.CM.99.CN71
MON.CM.99.L1
MON.NG.x.NG1
MUS.1.CM.01.1085
MUS.1.CM.01.CM1239
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM45
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

5366
5323
5024
5164
5145
5190
5320
5138
5195
5145
5112
4994
5392
5853
5828
5828
5385
5381
5370
5353
5344
5340
5212
5500
5862
5191
5460
5828
6493
6483
5935
6464
6485
5809
5953
6393
6392
6369
5518
5826
5867
6312
6397
6041
6234
6008
6027
5498
3952
6073
5692
5589
5624
5481
6000
5638
5861
5041
5041
5038
6122
5612
5597
4234
5771
5361
5678
5729
5606
5600
5296
5678
6157
5667
6003

H1B.FR.83.HXB2
Vpr (frameshifted)
Tat exon 1
H101.AE.TH.90.CM240
H102.AG.NG.x.IBNG
H103.AB.RU.97.KAL153_2
H104.cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WE.AU160
H1C.ET.86.ETH2220
H1D.CD.84.84ZR085
H1F1.BE.93.VI850
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.AN170
H1P.CM.91.MVP5180
CPZ.CD.90.ANT
CPZ.CM.01.SIV.cpz.CAM13
CPZ.CM.05.SIV.cpz.EK505
CPZ.CM.05.SIV.cpz.LB7
CPZ.CM.05.SIV.cpz.MB66
CPZ.CM.05.SIV.cpz.MT145
CPZ.CM.98.CAM3
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.GAB2
CPZ.TZ.01.TAN1
VPZ.US.85.CPZ.US
H2A.DE.x.BEN
H2A.GW.x.ALI
H2A.SN.x.ST
H2B.CL.x.EHO
H2B.GH.86.DP05
H2G.CL.x.AB196
H2U.FR.96.12034

Tat exon 1 start Vpr end
ATAGGCGTTACTCGA.....CAGAGGAGCAAGAAATGGAGCCAGTAGATCTAG.....ACTAGAGCC 5859
I G V T R Q R R A R N N G A S R S
H E P V D P R##
L E P
---A---TG-C---GG-A---GC-G---A-TG---A---C---5426
---T---A---GG-A---GC-G---TG---C---5383
---A---T---A---G---G---TG---C---5084
---A---CT---A---AG-G-C-GG---TGG---CCAG-AG---ATCCTGAC---5228
---A---T---C---GG-A---GC-G---A---A---5304
---A---TTA---A---A---A---A---C---A---5250
---A---C---A---A---T---A---CA---T---5380
---A---GT-C---A---A---T---GG-C-G---A---CT---T---5198
---A---A-C---G---A---G-T---GG---C-G---A---C---5255
---AA---A---A---A---GG-G---C---A---A---5205
---A---TA-C---A---A---GG-G---C---A---CAG---5172
---A---TA-C---A---A---GG-G---T---A---CA---5054
---GA-C---CT---A---GAG-G---A---T---T---5449
---AA---AC-C---AGAGGA---AG-AG---T---GA---GG-GCCC---5916
---AA---CT-C---TCTAACACAAGAGGA---AG-AG---T---GA-GCCC---5291
---ACA---GGA-C---A---AG-GC---C---CG-GA---AC-CCT---5441
---TA-A-TA-C---A---AG-GC---A---C-GA---GT-G---5427
---AA-C---CT---A---GAG-G---A---T---5413
---A---A-C---AATCAG---A---AG---A---A---TT---5410
---A---A-AC-CT---A---AGG-G---A---A---G-GA---CT---A---5400
---G-C-T-C-CT---GG---GAG---A---A---Y-A---5269
---GA-CT-YT---A---AGG---A---Y-A---5557
---A-CCT-C---A---A-AG-TCC---T---A---GA---C-G---5922
---GA-GAA---AG-G-GA---ATGGATCCTATAGATCCTAGT---T---5254
---ATCA---TGAATCCCATAGATCCTAGT---CA---5497
GA-T-CTTGGCAAGGAGAATCTCT-A-G-GAG-G-T-A-5803
---T-TAA-AA-GAGAAGA---ACTCCTT-CC-GCTGCACCGA-CC---G-AT-CACTAA---6556
G-T-CAA-AA-GGGAAGA---A-TCCTCTCT-GCT-TACCGA-CC---AA-AT-CAATA---6546
---T-TCAGC-CA-GGACGA---A-TCCTTT-T-GCT-TACC-A-CC---GG-AT-CGATAA---5998
---T-ACAAC-GG-GGAGGA---A-TCCTCTCT-GCT-TACCGC-TC-TA-GG-AT-CAGTAA---6527
---T-ACAAT-AG-GGAGGA---A-TCCTCTCT-CT-TACCGC-CC-TA-GG-AT-CGATAA---6529
H2G.CL.x.AB196---T-A-GG-AGC-GGCGGA---A-TCCTCTCT-TC-TAGG-AT-CGATAA---5872
---C---AGC-GCA-GGAGGA---A-TCCTCTCA-CTGTACCGC-TC-T-GG-GT-CTCTAA---6016

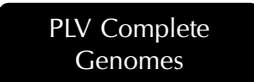
MAC.US.x.239
Vpr
Tat exon 1
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMBL_3
SMM.US.x.H192B
SMM.US.x.H192B
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F_L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.UG.x.TAN1
VER.DE.x.AGM13
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TYO1
COL.CM.x.CGU1
GRV.ET.x.GRI_677
MND_2.CM.98.CM16
DRL.x.x.FAO
RCM.GA.x.GAB1
RCM.NG.x.NG411
MND_1.GA.x.MNDGB1
MND_2.GA.x.M14
MND_2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.Iho7
GSN.CM.99.CN166
GSM.CM.99.CN71
MON.CM.99.L1
MON.NG.x.NG1
MUS_1.CM.01.1085
MUS_1.CM.01.CM1239
MUS_2.CM.01.CM1246
MUS_2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

--C---CAAC-G-GGAGGA.....A-TCCTCTCT-GCT-TACCGC-TC---AG-AT-CTATAA.....6456
I G V T R Q P L S I L S Q L Y R P L E A C Y N##
S A N L G E E N I L S Q L Y R P L E A C Y N##
---T---CAAC-G-GGGAAGA.....A-TCCTCTCT-GCT-TACCGC-TC---AG-AT-CTATAA.....6455
---C---CAAC-G-GGAGGA.....A-TCCTCTCT-CT-TACCGC-TC-T-GG-GT-CTATAA.....6432
---C---CAAC-G-GGAGGA.....A-TCCTCTCT-CT-TACCGC-TC-T-GG-GT-CTATAA.....5889
---T---ACAGC-AA-GGAGCA.....A-TCCCTG-G-TCT-TACGCT-GC---GA-GT-TTATAA.....8930
---C---CAGT-G-GGAGGA.....A-TCCTCTCT-CT-TACCGC-TC---GG-GT-TTATAA.....6375
---C---CAA-G-A-GGAGGA.....A-TCCTCTGT-CT-TACCGC-TC---GG-GT-CTATAA.....6460
---C---CAAT-A-GGAGGA.....A-TCCTCTCT-CT-TACCGC-TC---G-AT-CTATAA.....6104
---T---CAAC-AG-GGAGGA.....A-TCCTTTG---CT-TACCGC---C---GGAGT-TGTGAA.....6333
CA-A-CTACCCCTGCGTAT-GACAG-G-AGGAGGCCGCCCCAGGTCTGGGAGGAGCTTCAGGAGAGTGCATCGGCCCTCCAGGCCCTGTGA.6101
AGGAGAAAGCACTCTTT-AGCCCT-C-A-AA-G---AATGAGTCGGAGGGGGACGGGATGGCAGAGCCCTCCAGGACTTGCATAG.5803
CCCTTT-AA-ATACGAAGAAGGGATGGCAAG-G-G-G-GA-CGAA-C-GCTC-ACCAGGACTTGATTG.6105
CCCTTT-AGC-ATACGAGGAGGAGAGAAATGGCAAG-G-G-G---GACC-GAC-GTCC-ACCAGGACTTGATTG.6100
CCCTTT-AA-ATACGAGGAGGAGAGATGGCAAG-G-G-G-C-GA-C-AATC-GTCC-ACCAGGACTTGATTG.6105
TCTAGCAA-AC-CTTAG.....AGAAGGAGAGGACCTTTTT-CCCTT-C-A-AG-G---AATGGCAAGGAGGAGGA.....GCCCCACCCCTCTCCAGGACTTGCATAG.5516
---C---ACAAGAAG-GGGGAGT.....T-CCCTT-CG-TCCTT-CCGAGTGGGACAA-CCCTGTGA.....6039
C-T-A---C-AGC-GGACGCTG.....A-CCATTTCG-CCCT-CC-AGGGCG-ACCAAC-ATTGTAA.....5755
---T---AACAT-CA-GGAAGT.....A-TCCTCT-CG-TCC-T-CCGCAA-CG---AA-ATCATGTAA.....5652
---T---A-GAG-GA-GGAGGA.....A-TCCTCTTAG-TCC-T-CTTCA-TGC---AA-ATCTGTGAA.....5687
...A-G-AA-AA---AGACCACAA-T-CCCCCTT-G-TTCA-GC-A-AG-GATAGATTATAA.....A 5544
---T---ACAAGAAG-GGCAAA.....TATA-CCCTTAG-TCCCT-CC-AGGCC-ACAA-CCCTGTGA.....6066
---C---ACAAGAAG-GGAAGATAT.....A-CCCACTTCG-TCCTT-CCGAG-CGG-ACAA-CCCTGTGA.....5704
---C---CAAT-G-GGAGGA.....A-TCCTCTCT-CT-TACCGC-TC---AG-AT-CTATAA.....5924
GC-CAAGA.....GCAGGAACAGCT-CCC-C-A-GC-GCACC-AGACC-CTAAAAGAGACATACAAAAGAGCCGTTACTAA.5122
GC-ACAAGA.....GCAGGAACAGCT-CCC-C-A-GC-GCACC-AAACCC-CTGAAAAGATATACAAAAGAGCCGTTACTAA.5122
GC-C-AGA.....GCAGGAACAGCA-CCCA-C-A-GC-GCACCAGCA-CTGAAAGATATACAAAAGAGCCATTAAG.5119
---CA-CC-GA.....GCAGGAACAGCA-TCCA-C-A-GC-GCACC-AGACC-ATAAGAGACATACAAAAGAGCCATTAAG.5119
---ATTATCA.....GGGAGAT-CCAG-ATT-GCCCT-AGGG-C-CGG-CCACAGCAGTAATCTGTCTCTATGCCGATCTGAGCAGCTTAGGAGCCCTCAAGATATCGCATGGATGAATAG.5765
-ATA-AA.....GTCAAGAGT-CCAGT-ATC-GCCCTT-AGGG---CCG-CCACAGATAGTAATCTGTCTCTATGCTGATCTGAGCAGCCCAAGAGCCCTCAAGATACCGCATGGATGAATAG.5753
GC-AGCAA-TAA-AGATACCCCTGTTCTTAG-CA-TCAG-G-G-C-CTC-G-CCC-A-AGTTCGTCTCCAGGCCGATCCAGCAATCCAGAGAGCCCTCCAGATATAGAAATGGATGAATGA.5777
GCC-C-A-GRT---CAGGTACCCCTGRCCTG-GACACT-AG-G-C-C-CCC-AG-ACC-A-AGTTCGTCTCCCTATGAGCCAGAGAGCCCTTGCACCCGTCAGATACCCAGCAGTGAATAG.4390
---ACA-A-A-TGA---AGATAT-TCAGCATAG-CCCTT-AG-G-C-CACA-C-CC-GCT-TA-TTCTTAA.....TGGATCCATCAGTAGAGGAGT---5725
TCTAA-ACA-TCA-GAGATACCCCAATAGACCCCTTAG-G-C-CACA---TCC-CCT-TA-TTCAATGCTAATGCTATCTCTCTCTTGGACCCCTCAGATATAGAGTGGATGAGTAA.....TGGACCTCTCAGTAGA-AACCTT---5715
GCTCAAA-CA-TCA-GAGATACCCCAAGCTTGG-CCCTT-AG-G-C-GCAGC-CCAGAC-TA-TTCAATGCTAATGCTATGAGCCCAATCTTGGACCCCTCAGATATAGAGTGGATGAATAG.....TGGACCTCTCAGTGA-GGCCCT---5832
-GTCAAG---T-A-AGATACCCCAAGCTTGG-CCCTT-AG-G-G-CGCA-C-TCCAGAC-AA-TTCAACACCAATGTTGACCCCTTCTCCAGCAAGACCCCTCAGGATCCGATGGATGAGTAA.....TGGATCCAGAGTAGAGAAATTT---5785
GC---AAAGAGT-A-CACTATCCAAACATAAGACCTT-CGCCCTCGCA---AAGAGCC-GA.....ACGCG---CAG 5802
-G-AG-C-CA-GGG---TATCTAACATT---CC-TTG-C-GGAA-AAT-AGA-GTAG-GAGGGAATAA.....CCTTTCACCTTTGCTTAG-A-G---AG 3708
GG-AA-C---A-GAGA---TACCCCAATT---ACCCTT-C-GGAA-AA---GGAGT-AG-GATGGAGATAA.....CCTCTCTTACATCATAG-A-G---AGA 3702
CA-A-AAGA---GCCACTACCCCAATATC-GACCCCTTT-G-GA-G-AG-G-CCATGC-ATAA.....5865
CA-A-AAGA---GCCACTACCCCAATATC-GACCCCTTT-CG-C-GGA-G---AC-CCATGC-ATAA.....5360
---TCTT-C-TC---ATC---G---T-AG-A-ATGT-A-CCCAAGGCCCAACAAGACCAAGGAGCCATAGAAAGGCCATACAAAGACTAA.6259
-AGTA-CG-GA-ACAGGGATACCCCAT-CTTCTAG.....A 5707
-G-TAT-CC-GA-ACAGAGGGTACCCCAT-CTTCTAG.....A 6043

H1B.FR.83.HXB2
Tat exon 1
Vpu
Rev exon 1
H101.AE.TH.90.CM240
H102.AG.NG.x.IBNG
H103.AB.RU.97.KAL153.2
H104.cpx.CY.94.CY032
H1A.LG.85.LI455
H1B.US.90.WEAU160
H1C.ET.86.ETH2220
H1D.CD.84.84ZR085
H1F1.BE.93.VI850
H1G.SE.93.SE6165
H1H.CF.90.056
H1J.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.ANT70
H1P.CM.91.MVP180
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzLB7
CPZ.CM.05.SIVcpzMB66
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CAM3
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.GAB2
CPZ.TZ.01.TAN1
CPZ.US.85.CPZUS
H2A.DE.x.BEN
H2A.GW.x.ALI
H2A.SN.x.ST
H2B.CL.x.EHO
H2B.GH.86.D205
H2G.CL.x.ABT96
H2U.FR.96.12034

MAC.US.x.239
Tat exon 1
Rev exon 1
MAC.US.x.251.1A11
MAC.US.x.251.BK28
MAC.US.x.EMBL_3
SMM.US.x.92.S192B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F.L3
STM.US.x.STM
SAB.US.x.SAB1C
PAN.US.x.PAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TYO1
COL.CM.x.CCU1
GRV.ET.x.GRI.677
MND.2.CM.98.CM16
DRL.x.x.FAO
RCM.GA.x.GAB1
RCM.NG.x.NG411
MND.1.GA.x.MNDGB1
MND.2.GA.x.M14
MND.2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.lho7
GSN.CM.99.CN166
GSN.CM.99.CN71
MON.CM.99.L1
MON.NG.x.NG11
MUS.1.CM.01.1085
MUS.1.CM.01.CM1239
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
DEN.CM.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KK51
SYK.KE.x.SYK173

6151
6557
5722
5684
5376
5784
5600
6149
5563
5672
5500
5577
5494
5473
5473
5735
6199
6186
5589
5723
5721
5716
5715
5678
5567
5855
6206
5550
6477
6228
6703
6693
6145
6668
6640
6016
6157
6603
6602
6579
6038
6030
6071
6522
6601
6245
6477
6266
5777
6279
6274
6277
5637
6201
6293
5909
5797
5841
6220
6220
5858
6071
5287
5287
5284
6368
6064
6046
6038
4621
5946
5936
6056
6006
6114
6043
6033
6019
5514
6416
5894
6218



<p>H1B.FR.83.HXB2 Vpu Env H101.AE.TH.90.CM240 H102.AG.NG.x.IBNG H103.AB.RU.97.KAL153.2 H104.cpx.CY.94.CY032 H1A1.UJ.85.U455 H1B.US.90.WE.AU160 H1C.ET.86.ETH2220 H1D.CD.84.84ZR085 H1F1.BE.93.VI850 H1G.SE.93.SE6165 H1H.CF.90.056 H1I.SE.93.SE7887 H1K.CM.96.MP535 H1N.CM.95.YB30 H1O.BE.87.ANT70 H1O.CM.91.MP5180 CPZ.CD.90.ANT CPZ.CM.01.SIV.cpz.CAM13 CPZ.CM.05.SIV.cpz.EK505 CPZ.CM.05.SIV.cpz.LB7 CPZ.CM.05.SIV.cpz.MB66 CPZ.CM.05.SIV.cpz.MT145 CPZ.CM.98.CAM3 CPZ.CM.98.CAM5 CPZ.GA.88.GAB1 CPZ.GA.88.GAB2 CPZ.TZ.01.TANI CPZ.US.85.CPZ.US H2A.DE.x.BEN H2A.GW.x.ALI H2A.SN.x.ST H2B.Cl.x.EHO H2B.GH.86.D205 H2I.CL.x.AB196 H2U.FR.96.12034</p>	<pre> AGGAAAAATTAAGACAAAGAAAAATAGACAGTGTAAATTGATAGACTAATAGAAAGAGCAGAAGACAGTGGC.....A..ATGAGA..GTGAAGGAGAAATATCAGCACTTGTGGAGATGGGG 6265 R K I L R O R K I D R L I D R L I E R A E D S G#N E S E E T S A L V R M G M R V K E K Y Q H L W F M G A --- C - G - - - - - G - A - G - A - G - - - - - A - - - - - GT - A - G - GACA - - - ATGAAT - - - CC - AACCT 5833 AGG - A - - G - AG - - - - - A - G - - - - - T - - - - - GT - AT - GGGATA - - - A - GAAT - ATCC - CTCTT 5795 A - - - - - T - GAG - G - - G - G - - - - C - GTACA - A - G - - - - - GT - AT - GGGATG - - - AGGAAT - ATCC - CACTT 5635 AA - - - - - AA - T - GC - AG - - - - - A - - - - - T - AA - C - A - G - - - - - GT - AT - GGGATA - - - AGGAAT - ATCTT - CTT 5713 A - - - - - T - G - - - - - G - - - - - T - - - - - A - A - ACT - GG - - - - - A - - - - - GT - AT - GGGATG - - - AGGAAT - ATCAGCACTT 6360 C - T - G - AA - G - - - - - A - T - AA - TA - A - A - A - G - - - - - A - - - - - GT - A - GGGATAA - - - AGGAAC - ATCA - CCCCT 5783 A - - - - - C - GG - G - - - - - A - T - AA - TA - A - A - A - G - - - - - A - - - - - GT - AG - GGGATG - - - AGGAAT - - - CAGCACTT 5611 A - - - - - G - - - - - G - A - G - - - - - GA - A - C - - - - - A - G - - - - - A - - - - - GT - AC - GGGATA - - - AGGAAT - - - TA - CACTT 5668 A - - - - - A - C - G - - - - - G - - - - - T - A - - - - - A - - - - - A - - - - - AC - GA - TG - - - AT - G - GACA - - - AGGAAT - ATCC - A - CTT 5605 A - - - - - C - G - - - - - G - - - - - T - A - - - - - A - - - - - A - - - - - AC - GA - TG - - - AT - G - GACA - - - A - GAAT - CAGCACTT 5584 A - - - - - T - GG - A - - - - - A - G - - - - - A - T - - - - - A - - - - - A - - - - - ATGAGAGT - AG - GGGATG - - - AGGAAT - - - CAGCACTT 5464 A - - - - - A - - - - - AA - TTG - GGA - - - - - A - ACACA - - - GAC - G - A - G - - - - - A - - - - - TG - AAGTGATG - G - T - C - - - - - GTGG - TG - ATGG - AT - AGA - T - 5855 TTAGAACAAAAGGAAAC - GACAG - AGG - G - - - - - CTTGA - AGGT - AAGA - A - G - - - - - TTAGG - T - T - A - TATGAAAGC TTAGTGCAAAGA - AAC - GATAG - - - - - GG - GCAGG - - - - - CTTGAAAG - T - AAGG - A - - - - - AG - TCAGG - T - A - A - TATGAAAGT AAGCAGCAGCAATAGATAAT - A - G - ATCA - - - - - ATC - T - G - GTTA - TAG - GACT - TA - GATAGTGC - AT - - - - - A - - - - - G - - - - - AT - AGGAA - CC - T - CATAT 5697 A - - - - - AGATTATG - AAC - ATAAGCTG - GT - G - - - - - T - G - ACC - TGC - ITA - - - - - C - - - - - T - AA - - - - - TGAAA - - - - - T - G - GAGGA - - - AGAAG - TC - AGC - CTT 5846 A - - - - - G - AA - GT - G - GA - - - - - TCTCA - C - GAC - G - A - G - - - - - C - - - - - T - - - - - TGA - AGT - ACGGA - AT - CAGAA - AA 5820 A - - - - - AAG - G - ACA - - - - - AGGC - A - T - AA - - - - - G - GA - G - - - - - GG - - - - - A - - - - - ATGAGAGT - A - GGGCATG - ACGAGC - C - TGCCATA 5827 AAACTCTGT - AAG - GGATAGG - G - - - - - G - - - - - T - A - - - - - CA - A - G - - - - - AATGAAAGC AAAGAGTAT - AA - GT - CA - CA - GT - - - - - GCG - - - - - A - AC - TCAG - C - A - T - G - AGT - - - - - A - AATGAAAGT AGAAAGGAAAAGAG - GA - AAC - - - - - GC - CA - - - - - TAGG - CT - - - - - A - A - - - - - TC - CAG - T - - - - - A - - - - - ATG - G - T - A - AT - T - TG - TGGAG - AATC - CA - 5807 AGAGAAATAAAGAAAG - GGAAC - ACCCA - G - - - - - TT - - - - - AT - GCA - A - A - G - AT - CAG - T - - - - - AA - - - - - GTGAT - G - GAA - ATC - A - G - CA - 5681 AGAGAAATAAAGAAAG - GGAAC - ACCCA - G - - - - - TT - - - - - AT - GCA - A - A - G - AT - CAG - T - - - - - AA - - - - - GCAAT - G - GAG - AA - AGCTCAA 5969 TATAGAAGTATAAAGACAT - GGC - TGAG - C - G - - - - - TTGAGAGACTAAC - AATATTGAG - - - - - T - - - - - GAAA - - - - - A - TA - TGG - GA - A - GAA - A - AGAC - - - - - AA 6326 TATAAAAGGTGAAAAGACAC - GA - GAGCA - AG - TC - TTGATTG - TTA - - - - - AAATC - C - GA - - - - - T - - - - - T - A - AGTGATGGG - A - GACAA - - - - - C - CT - 5667 GGTATAAGAAGA - - - - - GAG - - - - - GA - AGGA - GGCA - C - TCAAAG - G - ATTAGA - AGGC - - - - - CT - - - - - AG - TT - AG - TA - A - - - - - AGTGAGTGAAGAAG AGACAAATAAAA - TA - - - - - GA - AAT - CT - - - - - CA - G - G - TC - A - ATC - TT - A - A - - - - - A - C - G - AT - - - - - A - - - - - T - AA - - - - - ATGA - A - T - AAT - 5892 AGCAAAATAAAA - TA - - - - - GA - AAT - CT - - - - - CA - G - G - TC - A - ATC - TT - A - A - - - - - A - C - G - AT - - - - - A - - - - - T - AA - - - - - ATGA - AGTG - GA - AGAA - AA 6336 AGCAAAATAAAA - TA - - - - - GA - AAT - CT - - - - - CA - G - G - TC - A - ATC - TT - A - A - - - - - A - C - G - AT - - - - - A - - - - - T - AA - - - - - ATGA - AGTG - GA - AGAA - AA 6336 ATGAT - TCTAGTA - A - ATCA - CT 6716 GATG - - - - - TGTGTA - - - - - ATCACT 6168 ATGGCA - TG - TAAT - AT - ACCT 6691 ATGGCAT - T - TA - C - CC - CCT 6693 ATGACAT - TC - TG - C - ATCA - CT 6319 ATGGCATGC - TG - C - ATCA - CT 6180 Env signal peptide start M G C L G N ATGGGATGC - TG - - - - - ATCA - CT 6626 ATGGGATGC - TG - - - - - ATCA - CT 6602 ATGGGATGC - TG - - - - - ATCA - CT 5751 ATGGGATGC - TG - - - - - ATCA - CT 6053 ATGGGATGC - TG - - - - - ATCA - CT 6094 ATGGGATGC - TG - - - - - ATCA - CT 6545 ATGGGATGC - TG - - - - - ATCA - CT 6624 ATGGCCTG - CCTG - A - ATCACT 6268 GTATGA - - - - - TTC - TACAGT - CT - CT 6503 ATGGGA - CA - AA - - - - - G - AAA - - - - - 6289 ATGA - - - - - TGACA - TACTGATA 5800 ATGA - A - CTC - T - TAG - TATA - T 6302 ATGACAA - G - C - TAG - - - - - ATTTT 6297 GTATGAG - T - TACAATA - T - ACCTT 6302 ATG - TTAGA - ATCTCTTTATAAAA 5660 ATGGG - AGA - - - - - CTT - T - AAAAT 6224 GT - - - - - TAAAGACTAGT - CTTA - TA 6319 GT - A - AAA - AT - TTA - ACA 5929 ATGA - - - - - TGTCCAG - TCTT - TTCT 5759 GT - ATAATAGTAAGTAT - CTTA - ACA 6245 GT - - - - - AACAGTAAGTAT - TT - A - ATA 5883 ATGGGATGC - TG - - - - - ATCA - CT 6094 ATGA - CTG - CCTG - ACT - ATACT 5310 ATGG - CTGCT - TAA - CTCCTACT 5310 ATGGCATGTCAG - T - ACATATT 5307 GGATGGCATGTCAG - ATT - G - A - - - - - 6392 A - AG - A - - - - - TGAGGA - ACAG - GAATCTAT - AT 6151 TG - A - - - - - TGATGAG - A - AG - G - ATTT - T - AA 6136 AG - A - - - - - CATGAG - A - A - - - - - 6187 GGATGAG - ATGAA - ACC - T - ATA - T 4790 GATG - C - AGGA - - - - - ACAT - GCT - A - 6109 ATGAGAT - T - GAGCA - - - - - ATGAT - C - T - AAT 6093 AGTAACAGCTAATTATCTGTATTGCT - TAGT - C - TTTATCT - GCCTATCAAAGG - G - TGTC - ACCTA - CC - CCAATAG - AGTAA - TCCTCCGCTCCGGAAAGGAGACTCAGACAGGC - CATCTTTG - G - AT - CTTGGAT - G - GAG - AC - - - - - GAITCT 6225 AGTAGTAACCTGCATATAATTTGTAGACTTTAGT - C - TTTG - CTT - CA - TCA - CCGTGG - GCC - C - CTC - A - AGG - CAG - A - AGGTCAACGTATACCCGCTCCTAGGAAAGG - GATACC - TAGCG - T - TCTTTGAGGAC - GAAA - AT - GC - GT - AT 6172 AATAGAAATCGTTTTGCTGAT - TTTG - GT - GTTATT - CCTGG - T - CGAG - CATAGTG - AGGGT - TCTTGCTT - TAGAGCCT - T - A - AGTTATAGGAAAGCTTAGTTATAAAGACTAAGT - TGTTG - GTATA - TA - T - TATGAGT - - - - - AG - CA - - - - - GACCCT 6283 GTACAGTCTGGGATAGTCC - AAGTTGAT - GTTCTGT - C - GCC - ACTCATCAGAGTCA - GAT - GGC - G - GA - - - - - CAGCATGCTGCCAC - AGTCTTAGTGAGAGAAAGTCCCA - - - - - ATG - CAA - TG - ACCCA - AACCT 6182 CCACTGTTGGCTTATCTGTACA - CT - - - - - G - TCAGA - TCAGG - TGGCTGT - ACAGCA - C - T - C - AC - GC - T - GT - - - - - TAGAGAAACAAGTCCC - - - - - ATG - CTA - TG - - - - - CCCAG - ATTT 6151 TATATGTTTAA - CAGC - ATG - TA - TG - - - - - AGCAAT - - - - - 6057 GTTACATGTTTAA - CC - C - ATGA - ATG - - - - - ACAAGTCTAT - 5555 T - - - - - GATG - G - TG - - - - - 6430 GC - C - ATGA - A - TT - TAAT - TCATAA - 5922 T - - - - - ATGGCAGCT - TA - A - CT - ACAT 6247 </pre>
<p>MAC.US.x.239 Env MAC.US.x.251.1A11 MAC.US.x.251.BK28 MAC.US.x.EMBL.3 SMM.SL.92.SL92B SMM.US.x.H5 SMM.US.x.RCM53 SMM.US.x.SIV.smH635F.L3 STM.US.x.STM SAB.SN.x.SAB1C TAN.UG.x.TANI VER.DE.x.AGM3 VER.KE.x.9063 VER.KE.x.AGM155 VER.KE.x.TYO1 COL.CM.x.CGU1 GRV.ET.x.GRI.677 MND.2.CM.98.CM16 DRL.x.x.FAO RCM.GA.x.GAB1 RCM.NG.x.NG411 MND.1.GA.x.MNDGB1 MND.2.GA.x.M14 MND.2.x.x.5440 MNE.US.x.MNE027 LST.CD.88.447 LST.CD.88.485 LST.CD.88.524 LST.KE.x.Iho7 GSN.CM.99.CN166 GSN.CM.99.CN71 MON.CM.99.L1 MON.NG.x.NG1 MUS.1.CM.01.1085 MUS.1.CM.01.CM1239 MUS.2.CM.01.CM1246 MUS.2.CM.01.CM2500 DEN.CD.x.CD1 DEB.CM.99.CM40 DEB.CM.99.CM5 TAL.CM.00.266 TAL.CM.01.8023 SUN.GA.98.L14 SYK.KE.x.KE51 SYK.KE.x.SYK173</p>	<pre> ATGGGATGC - TG - - - - - ATCA - CT 6626 M G C L G N ATGGGATGC - TG - - - - - ATCA - CT 6602 ATGGGATGC - TG - - - - - ATCA - CT 5751 ATGGGATGC - TG - - - - - ATCA - CT 6053 ATGGGATGC - TG - - - - - ATCA - CT 6094 ATGGGATGC - TG - - - - - ATCA - CT 6545 ATGGGATGC - TG - - - - - ATCA - CT 6624 ATGGCCTG - CCTG - A - ATCACT 6268 GTATGA - - - - - TTC - TACAGT - CT - CT 6503 ATGGGA - CA - AA - - - - - G - AAA - - - - - 6289 ATGA - - - - - TGACA - TACTGATA 5800 ATGA - A - CTC - T - TAG - TATA - T 6302 ATGACAA - G - C - TAG - - - - - ATTTT 6297 GTATGAG - T - TACAATA - T - ACCTT 6302 ATG - TTAGA - ATCTCTTTATAAAA 5660 ATGGG - AGA - - - - - CTT - T - AAAAT 6224 GT - - - - - TAAAGACTAGT - CTTA - TA 6319 GT - A - AAA - AT - TTA - ACA 5929 ATGA - - - - - TGTCCAG - TCTT - TTCT 5759 GT - ATAATAGTAAGTAT - CTTA - ACA 6245 GT - - - - - AACAGTAAGTAT - TT - A - ATA 5883 ATGGGATGC - TG - - - - - ATCA - CT 6094 ATGA - CTG - CCTG - ACT - ATACT 5310 ATGG - CTGCT - TAA - CTCCTACT 5310 ATGGCATGTCAG - T - ACATATT 5307 GGATGGCATGTCAG - ATT - G - A - - - - - 6392 A - AG - A - - - - - TGAGGA - ACAG - GAATCTAT - AT 6151 TG - A - - - - - TGATGAG - A - AG - G - ATTT - T - AA 6136 AG - A - - - - - CATGAG - A - A - - - - - 6187 GGATGAG - ATGAA - ACC - T - ATA - T 4790 GATG - C - AGGA - - - - - ACAT - GCT - A - 6109 ATGAGAT - T - GAGCA - - - - - ATGAT - C - T - AAT 6093 AGTAACAGCTAATTATCTGTATTGCT - TAGT - C - TTTATCT - GCCTATCAAAGG - G - TGTC - ACCTA - CC - CCAATAG - AGTAA - TCCTCCGCTCCGGAAAGGAGACTCAGACAGGC - CATCTTTG - G - AT - CTTGGAT - G - GAG - AC - - - - - GAITCT 6225 AGTAGTAACCTGCATATAATTTGTAGACTTTAGT - C - TTTG - CTT - CA - TCA - CCGTGG - GCC - C - CTC - A - AGG - CAG - A - AGGTCAACGTATACCCGCTCCTAGGAAAGG - GATACC - TAGCG - T - TCTTTGAGGAC - GAAA - AT - GC - GT - AT 6172 AATAGAAATCGTTTTGCTGAT - TTTG - GT - GTTATT - CCTGG - T - CGAG - CATAGTG - AGGGT - TCTTGCTT - TAGAGCCT - T - A - AGTTATAGGAAAGCTTAGTTATAAAGACTAAGT - TGTTG - GTATA - TA - T - TATGAGT - - - - - AG - CA - - - - - GACCCT 6283 GTACAGTCTGGGATAGTCC - AAGTTGAT - GTTCTGT - C - GCC - ACTCATCAGAGTCA - GAT - GGC - G - GA - - - - - CAGCATGCTGCCAC - AGTCTTAGTGAGAGAAAGTCCCA - - - - - ATG - CAA - TG - ACCCA - AACCT 6182 CCACTGTTGGCTTATCTGTACA - CT - - - - - G - TCAGA - TCAGG - TGGCTGT - ACAGCA - C - T - C - AC - GC - T - GT - - - - - TAGAGAAACAAGTCCC - - - - - ATG - CTA - TG - - - - - CCCAG - ATTT 6151 TATATGTTTAA - CAGC - ATG - TA - TG - - - - - AGCAAT - - - - - 6057 GTTACATGTTTAA - CC - C - ATGA - ATG - - - - - ACAAGTCTAT - 5555 T - - - - - GATG - G - TG - - - - - 6430 GC - C - ATGA - A - TT - TAAT - TCATAA - 5922 T - - - - - ATGGCAGCT - TA - A - CT - ACAT 6247 </pre>

	Vpu end	Env signal peptide end	Env gp120 start		V1 loop start
H1B.FR.83.HXB2	VTGGAGATGGGGCACCCTCTCTGGG...ATGTTGATGATCTGTAGTCTACAGAAAATGTTGGGTCACAGCTATTATGGGGTACTGTTGGAAGGAAGCAACC.....ACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATA			6423	
Vpu	V E M G H H A P V D M D C S A T E K L W V T V Y G V P V W K E A T				
H101_AE.TH.90.CM240	W A W G T T A L G M L M T C S A T E K L W V T V Y G V P V W K E A T				
H102_AG.NG.x.IBNG	A A G T T A T T A A A A T A A G T C C A G T T A G A T G A T G A T				
H103_AB.RU.97.KAL153.2	A A G T T A T T A A A A T A A G T C C A G T T A G A T G A T G A T				
H104_cpx.CY.94.CY032	A GA A T T A T T A A A A T A A G T C C A G T T A G A T G A T G A T				
H1A_FLU.85.CPZ155	A A T T A T T A T T G T G A A A T A A G T C C A G T T A G A T G A T G A T				
H1B.US.90.WEAU160	A A T T A T T A T T G T G A A A T A A G T C C A G T T A G A T G A T G A T				
H1C.ET.86.ETH2220	A T A T T A T T A G G T T A A A A T A A G T C C A G T T A G A T G A T G A T				
H1D.CD.84.84ZR085	A A T T A T T A T T A G T G A A A T A A G T C C A G T T A G A T G A T G A T				
H1F1_BE.93.VI850	G A A C T T T A T T G A A A A T A A G T C C A G T T A G A T G A T G A T				
H1G.SE.93.SE6165	A A G T T A T T A T T G T G A A A T A A G T C C A G T T A G A T G A T G A T				
H1H.CF.90.056	A G T T A T T A T T A T T A A A T A A G T C C A G T T A G A T G A T G A T				
H1I.SE.93.SE7887	A G T T A T T A T T A T T A A A T A A G T C C A G T T A G A T G A T G A T				
H1K.CM.96.MP535	G A C T T A T T A T T A T T A A A T A A G T C C A G T T A G A T G A T G A T				
H1N.CM.95.YBF30	T T T C T T T A T T A T T A A A T A A G T C C A G T T A G A T G A T G A T				
H1O_BE.87.AN170	A C C T T A T T A T T A T T A A A T A A G T C C A G T T A G A T G A T G A T				
H1O.CM.91.MP5180	A C C T T A T T A T T A T T A A A T A A G T C C A G T T A G A T G A T G A T				
CPZ_CD.90.ANT	T A T T T A T T A T T A T T A A A T A A G T C C A G T T A G A T G A T G A T				
CPZ.CM.01.SIVcpzEK13	C A C T G A C C A T G A A A T G A A C C T T G T G C A A C T G T G A A				
CPZ.CM.05.SIVcpzEK505	T C T A T C T G C C G A A A A T T G C C A A A T T G C A A A T T A				
CPZ.CM.05.SIVcpzLB7	C A C G T G T G T C A G T G T C T C A T A A T G T A C C A A A T T A				
CPZ.CM.05.SIVcpzMB66	G A T A T C T T A T G C A T G G A T A G G G A A A G G A A A T A G T G A				
CPZ.CM.05.SIVcpzMT145	C C T C A C A T A G T A A G A T T G G A T G A C C A T T G A G T T G				
CPZ.CM.98.CAM5	C T T G A A A A T T A T A G T T T T G A A C C A T G T G C G G T C				
CPZ.GA.88.GAB1	T A T A T T A T T G A C A G A A T G G T T A C A C C C A T T G A G C				
CPZ.GA.88.GAB2	C A C T T C A T A T T C A A A A A C A A C T T A C A C C A T T G A G				
CPZ.TZ.01.TANI	A A T C T T A T T G A C C T C A T A T A C C A G A A T A G T T A G C T				
CPZ.US.85.CPZUS	A A C T C C T A G T T A T T G C T C A G C T A T A T C A G A T T G C T				
H2A.DE.x.BEN	T T G T G C A T T T A C A G C A G T G T C T T A T A T A T T G T G C C A				
H2A.GW.x.ALI	C T G T A C A T C T T A C A G C T G T C T T A T A T A T T G T G C C A				
H2A.SN.x.ST	A T T G C C A C T T G C A G C T G T C T T A T A T A T T G T G C C A A T				
H2B.CL.x.EHO	C T G T A C A C T C T G C T A A G T G T A G G T A G G T T G T G T G T				
H2B.GH.86.DB5	C C T A T G C A C T C T G C T A A G T G T A G G T G T G T G T G T				
H2G.CL.x.AB196	C T A G C A A C C T T T G C A A G T C T A C T G A T A T A C T G C A A T				
H2U.FR.96.12034	C T A G C C T A T T G T G T A G T C A G C G T A T A C T G C A A T T				
MAC.US.x.239	C T A G C C A C T T G C T A G T G A G G A T C T A T T G T C T C A A T				
MAC.US.x.251.1A11	C T A G C C A C T T G C T A G T G A G G A T C T A T T G T C T C A A T				
MAC.US.x.251.BK28	C T A G C C A C T T G C T A G T G A G G A T C T A T T G T C T C A A T				
MAC.US.x.EMBL_3	C T A G C C A C T T G C T A G T G A G G A T C T A T T G T C T C A A T				
SMM.SL.92.SL92B	C T A A G A C A C T T G C T A G T G G T A G G A C C T G G T G C C A A T				
SMM.US.x.H9	C T A G C G C T C T G C A Y A G T G C T C A G G A T T A T T G T T C A A A T				
SMM.US.x.RGM53	C T A G C C A C T T G C T A G T G A G G A T C T A T T G T C T C A A A T				
SMM.US.x.SIVsmH635F.L3	C T A G C G C T C T G C A G A G T G T A G A G A T T G T G T G T C A A A T				
STM.US.x.STM	C T A G C A T C T T G C A A A G T G C T T A T T G G C T T C A A A T				
SAB.SN.x.SAB1C	T T T A T T A G A A T T G G A T A A G C T A A A G C T G T T A T A T G C C A				
TAN.UG.x.TANI	A A A T T A A T A A G G A A A G G A A G A G C T A A A G G C C A A				
VER.DE.x.AGM3	T T T A T T A G A A T T G G A T A A G C T A A A G C T A A A G G C C A A				
VER.KE.x.9063	A A T A A A A T A A G A A G A A G G A A G G T A A G A A G A A A				
VER.KE.x.AGM155	A A T A A A A T A A G A A G A A G G A A G G T A A G A A G A A A				
VER.KE.x.TYO1	A A T A A A A T A A G A A G A A G G A A G G T A A G A A G A A A				
COL.CM.x.CGU1	T T T G C A G A T A T A G G T A T A G T T A G T T A A A A A A A				
DRV.ET.x.GRI.677	A A A A A A C A A T A G G A A A G T A G A T A G A C C A C A T G				
MND_2.CM.98.CM16	A A A A A A C A A T A G G A A A G T A G A T A G A C C A C A A T				
DRL.x.x.FAO	T A T A A T T A G G C T A T G C A T T T A G G A A A G A A T A T A T G				
RCM.GA.x.GAB1	A T A G T A T A T A T T A A G A A T A G G G A A C A G A C A A G A A T A				
RCM.NG.x.NG411	T T T G C A A T A A T A G G A T A G G T A A T T A T A G G G C A A A T				
MND_1.GA.x.MNDGB1	T T A A T A T T A G C C A A T A A T A A G T A G G T A G T A T A T A T				
MND_2.GA.x.M14	T T T A T T A T T A G C C A A A T A A T A A G T A G G T A G T A T A T				
MND_2.x.x.5440	T T T C T T G T A T A C C T T A G G A A A A G A C T G T A T A T A T				
MNE.US.x.MNE027	C T A G C C A C T T G T G C A G T G C A A G A T C T A T T G C T C A A A				
LST.CD.88.447	T T A T T A T T A G C C A A A G A A G A A G G A G A A A A G G C A A A				
LST.CD.88.485	T T T G C T C T T T A G G A G C A A A G G G A G A A A G G C A A A				
LST.CD.88.524	C A T G C T C T T G T G G A C A C A G C A G A A A A G G C A A A T A				
LST.KE.x.1ho7	A C T C T T T C T T G G A T A T A T T G G G C A A A T A A A T A A				
GSN.CM.99.CM166	A T C T G C A T G G T A A G G C T T A A T G G T A A G A A C C G G T T A				
GSN.CM.99.CN71	A T C T G T A T G G C A A C C A T G C T T C A C C C G G T C G A G T				
MON.CM.99.L1	C T T G G C T T G T C A A C C C T G T G A T T G T T G T G T A C C A C				
MON.NG.x.NG1	T T C A T C C T T G T T A G A G A C C C A C C T G G G C R A T A G G A C				
MUS_1.CM.01.1085	A G C C A C A C G C T T G A A A C C T G C T G C G T A A T T G G G C A A				
MUS_1.CM.01.CM1239	G T T G C A C G C T T A A C C T G C T G C G T A A T T G G G C A A				
MUS_2.CM.01.CM1246	C A C C G G C T T T G C T A C C A G T A G C C A A T A A A A C A T C C				
MUS_2.CM.01.CM2500	A C C A C C T G G T T A A C C C T G C T G C T A A T A T T G G G C A				
DEN.CD.x.CD1	T A A T C T C C T A C C T T G G C A A T G G A A A T A A A G C A A C				
DEB.CM.99.CM40	T T				
DEB.CM.99.CM45	T T				
TAL.CM.01.8023	T T				
SUN.GA.98.L14	T T				
SVK.KE.x.KE51	T T				
SVK.KE.x.SYK173	T T				

Table with columns for sequence identifiers (e.g., H1B.FR.83.HXB2, Env), sequence alignments, and accession numbers (e.g., 6662, 6233, 6189).

Table with columns for sequence identifiers (e.g., MAC.US.x.239, Env), sequence alignments, and accession numbers (e.g., 7043, 7039, 7025).

Genomes
PLV Complete

Table with columns for sequence identifiers (e.g., H1B.FR.83.HXB2, MAC.US.x.239), V1 loop end, V2 loop start, and V3 loop start. The table contains multiple rows of sequence alignments and identifiers, with some entries including accession numbers and dates.

H1B.FR.83.HXB2
Env
H101_AE.TH.90.CM240
H102_AG.NG.x.IBNG
H103_AB.RU.97.KAL153_2
H104_cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WE.AU160
H1C.ET.86.ETH.220
H1D.CD.84.84ZR085
H1F1.BE.93.VI805
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1J.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.AN170
H1O.CM.91.MVP5180
CPZ_CD.90.ANT
CPZ_CM.01.SIVcpzCAM13
CPZ_CM.05.SIVcpzEK505
CPZ_CM.05.SIVcpzLB7
CPZ_CM.05.SIVcpzMB66
CPZ_CM.05.SIVcpzMT145
CPZ_CM.98.CAM3
CPZ_GA.88.GAB1
CPZ_GA.88.GAB2
CPZ_TZ.01.TANI
CPZ_US.85.CPZUS
H2A.DE.x.BRN
H2A.GW.x.ALI
H2A.SN.x.ST
H2B.CL.x.EHO
H2B.GH.86.D205
H2G.CL.x.ABT96
H2U.FR.96.12034
MAC.US.x.239
Env
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMBL_3
SMM.SL.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F_L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.UG.x.TAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.ACM155
VER.KE.x.TYO1
COL.CM.x.CGU1
GRV.ET.x.GRI_677
MND.2.CM.98.CM16
DRL.x.x.FAO
RCLM.GA.x.GAB1
RCM.NG.x.NG411
MND_1.GA.x.MNDGB1
MND_2.GA.x.M14
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.lho7
GSN.CM.99.CM166
GSN.CM.99.CM71
MON.CM.99.L1
MON.NG.x.NG1
MUS_1.CM.01.1085
MUS_1.CM.01.CM1239
MUS_2.CM.01.CM1246
MUS_2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

AATGGAATAACACTTTAAACAGATAGCTACGAAATTAAGAGAACAATTTGGAATAATAAAACAATAATC.....TTTAAAGCACTCTCA 7319
K W N N T L K Q A K L R E Q F G N N K T I I F K Q S S
-----AGT-----G-A-GAA-----A-G-C-AAT-G-A-C-----ATAATC-----C-A-Q-S-S-----6890
-----A-----C-T-G-T-CAC-----AGACGT-C-AAG-C-CC-C-----ATAATC-----GCTA-CC-----T-6837
-----A-----A-----T-T-----A-----G-----A-C-----ATAGTC-----T-----6511
-T-----G-C-----GT-----AG-GAAG-----G-A-AG-TC-CCCT-----A-C-T-A-----GCT-C-C-GT-----6701
-C-----GA-A-C-----G-----GAAC-----AGA-A-G-AAT-C-A-C-----ATAATC-----GCTAGC-----6764
GT-----A-C-C-A-G-----T-GAA-----CA-A-G-AT-AAACA-TT-----G-T-A-C-A-----A-----GTT-----A-----7329
-----A-C-C-A-G-----AAGGAA-----CA-A-G-C-CCCT-----ATAATC-----A-----7116
-----A-C-C-G-----G-----A-----G-A-C-TT-GAACC-A-CA-----ATAATT-----A-C-----6855
-----C-GG-GT-TG-AAGGCG-----AGTCG-T-CCCT-----ATAAAA-----C-----6638
-----AG-GG-C-----C-GA-TG-C-GCAG-C-G-GA-ATC-AAT-G-G-C-GCAG-AC-----ATAACC-----CTC-----TG-----6749
-C-----G-----C-C-G-----T-CAC-----G-AT-C-GAAC-----GA-C-----ATAAGC-----C-AA-----6644
-T-----G-----C-----CA-----C-----A-----A-----ATAAAC-----CATC-----6623
-T-----C-TG-----GCAGAG-AAGGAA-GC-A-----T-----AAG-GGA-C-T-----ATAATTTAAACCA-C-AA-C-----6524
-T-----G-ACCA-TG-GG-TAGA-C-AGAGAG-----AGA-ATCC-G-G-A-C-C-----ATAACCTTCAGGG-CGAG-GAGGAT 6900
-T-----GGA-A-TA-----A-C-----GAA-GG-ATTT-----T-G-AAAC-----CAGGT-----AGTATAACATGACA-C-T-CAG-AGC 7361
-T-----G-A-TG-CC-----A-C-----TA-GG-ATTT-A-TT-G-AAAC-C-AG-G-AT-----GTTACCAATAA-C-GCAG-A-TAGT 7369
-T-----G-AC-AG-ACGT-----CA-A-C-AGCA-GTTC-TGGG-G-T-GGAA-----AGTAG-C-AC-A-----C-AAT-----GCGAAAACAA-ATG-AC-T-CA-6781
-C-----G-A-AC-GC-A-AG-TAGAGAA-GC-----T-A-GG-A-AAA-GCA-----GC-ATGG-----A-----AATATAAGC-G-T-C-AG-AGT 6939
-----G-AC-A-TA-GG-TACA-C-AAACA-T-A-----T-A-CA-----A-----AATATTA-----TA-CCAATC-----6841
-T-----T-----A-TAC-----C-G-GAAA-AG-GC-----CCACGGTC-CAATG-ATC-----C-A-----ACTGTACAC-----GCAA-CAGT-----6851
-T-----C-----T-A-----A-AG-AGGCA-GCC-----AT-----AC-A-AAGGGC-----T-TCTCGG-ATGAA-----ACTATACA-----GGA-C-AG-A-G 6967
-----G-GCA-AG-CC-C-TG-A-C-AAAGAGTTC-----GA-CAT-C-G-A-G-----A-----AACATAAG-----T-GTCC-T-----6831
TC-----G-G-A-GAC-----TG-A-CT-G-CAAGC-----G-G-TT-----AC-GA-----GC-TAGC-A-G-C-----AAC-----CAGT-AGTC 6747
-G-----T-G-A-AG-CG-----CC-----TAGCA-----GC-A-GT-GA-AA-----GC-TAGC-A-G-C-----ACC-----CA-C-AC-C 7023
-C-----CGA-GG-GG-G-AG-AAA-AGCCT-----GC-ACCTC-CAATCGA-CAGC-G-C-AC-A-----ACAC-C-CAG-G-----T 7380
-T-----T-C-----CAGA-----AAACAA-----GA-T-A-A-CAGCAACCTGGC-C-GGGGGAC-----ACA-C-CAA-GCATC-----6733
-C-----T-C-GG-AC-----G-TT-GCTATGGCTGCCA-----ACAC-ACCA-GA-CGG-CA-----TGA-CGG-A-----CAACAAAG-----C-ATGGCAGAG 6966
-C-----G-G-ACAGAG-----T-C-CTTTAGAG-T-A-----A-A-GG-GCAACCTG-----C-GGT-----GAG-A-TT-C-AATG-----7348
-G-----GG-AG-CA-GC-GG-G-GAAGCA-----CC-TGT-C-----TCCA-T-C-AGG-----T-ATGATACA-----G-G-AGC-A-G-AGC-A-G-AGGAG-----7834
-----G-----AG-CA-GC-GG-G-GAAGGAA-CCC-TGT-A-----TCCA-GT-----AGG-----C-ATGA-ACA-----AACCAAATTAACCTTACA-A-C-GGAAG-----7815
-----G-----GG-AG-CA-GC-GG-G-GAAGTCA-CCC-TGC-A-----TCCA-GT-----AGG-----C-ACGA-ACA-----GAAAAATTCGTITTA-AGC-C-GGAG-----7261
-C-----CAG-AG-CA-C-GG-G-GAAGAG-CCA-T-A-A-T-----TCCA-T-----TCAGG-----C-ATATC-----TCACAGTAAGTTAGCAG-----CG-GAG-----7793
-C-----CAG-AG-CA-----GG-G-GAA-T-G-CCA-----T-A-A-----TCCA-GT-----AGG-GGTGC-----AAAATATCACA-----ACGCTAAGTTAGT-CAG-A-TGGAAA-----7807
-----GG-AG-CA-GC-GG-G-AAGGAA-CCG-GTCA-G-CCCCA-CT-----AGG-----ATGA-ACA-----AAGAAATAACTTTTACA-CA-C-GGAA-----7165
-T-----AGGGG-AA-CC-GG-G-GAAGGAG-CCA-TGCCA-C-----TCCA-GT-----C-GGT-C-AG-ATATT-----AGTCAGATAACTTCGAC-G-A-C-G-AGG-----7294

H1B.FR.83.HXB2
Env
H11 AE.TH.90.CM240
H102.AG.NG.x.IBNG
H103.AB.RU.97.KAL153_2
H104.cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WE.AU160
H1C.ET.86.ETH2220
H1D.CD.84.84ZR085
H1F1.BE.93.VI850
H1G.SE.93.SE1615
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.ANT170
H1O.CM.91.MV5180
CPZ.CD.90.ANT1
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzLB7
CPZ.CM.05.SIVcpzMB66
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CAM3
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.GAB2
CPZ.TZ.01.TANI1
CPZ.US.x.85.CPZUS
H2A.DE.x.BB1
H2A.GW.x.AL1
H2A.SN.x.ST
H2B.CL.x.EHO
H2B.GH.86.D205
H2G.CL.x.ABT96
H2U.FR.96.12034

GGAGGGGCCAGCAAATTTGTAACCGCACAGTTTAAATGTGGAGGGGAAATTTTCTACTGTAATTCAACACAAGTCTGTTAATAGTACTGGTGGTAAGTACTGAAGGGTCAAATAACACT...
G D P E I V T H S F N C G G E F Y C N S T Q L F N S T W S T W S E G S N M T ... GAAGGAAG 7456
...-A-T-T-G-G-ACCTTCATG-GACA--CA--A-G-CC--AATG-ATTGGT-TC-A--T-T-GGTAGAA-GG-A--AGCT-AC--CAGAGC AAAAG--CAGCA 7877

MAC.US.x.239
Env
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMBL_3
SMM.SL.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F_L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.UG.x.TAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.YO1
COL.CM.x.CGU1
GRV.ET.x.GRI_677
MND_2.CM.98.CM16
DRL.x.FAO
RCLM.GA.x.GAB1
RCLM.NG.x.NG411
MND_1.GA.x.MNDG1
MND_2.GA.x.M14
MND_2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.Jho7
GSN.CM.99.CM166
GSN.CM.99.CM71
MON.CM.99.L1
MON.NG.x.NG11
MUS_1.CM.01.1085
MUS_1.CM.01.CM1239
MUS_2.CM.01.CM1246
MUS_2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

...-A-T-T-G-G-ACCTTCATG-GACA--CA--A-G-CC--AATG-ATTGGT-TC-A--T-T-GGTAGAA-GG-A--AGCT-AC--CAGAGC AAAAG--CAGCA 7877
...G D P E I V T H S F N C R G E F L Y C K H N W F L N W V E E D R N N A A T G ... CAGAGC AAAAG--CAGCA 7877
...-A-T-T-G-G-ACCTTCATG-GACA--CA--A-G-CC--AATG-ATTGGT-TC-A--T-T-GGTAGAA-GG-A--AGCT-AC--CAGAGC AAAAG--CAGCA 7877

H1B.FR.83.HXB2
Env
H11 AE.TH.90.CM240
H102.AG.NG.x.IBNG
H103.AB.RU.97.KAL153.2
H104.cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WE.AU160
H1C.ET.86.ETH2220
H1D.CD.84.84ZR085
H1F1.BE.93.VI850
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP335
H1N.CM.95.YBF30
H1O.BE.87.ANT170
H1O.CM.91.MVP5180
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzLB7
CPZ.CM.05.SIVcpzMB66
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CAM3
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.GAB2
CPZ.TZ.01.TANI
CPZ.US.85.CPZUS
H2A.DE.x.ALI
H2A.GW.x.STI
H2A.SN.x.SL
H2B.CL.x.EHO
H2B.GH.86.D205
H2G.CL.x.AB796
H2U.FR.96.12034

MAC.US.x.239
Env
MAC.US.x.251.1A11
MAC.US.x.251.BK28
MAC.US.x.EMBL_3
SMM.SL.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.SIVsmH635F.L3
STM.US.x.STM
SAB.SN.x.SABIC
TAN.UG.x.TANI
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TYO1
COL.CM.x.CU01
GRV.ET.x.GRI_677
MND.2.CM.98.CM16
DRL.x.FAO
RCLM.GA.x.GAB1
RCLM.NG.x.NG411
MND.1.GA.x.MNDGB1
MND.2.GA.x.M14
MND.2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.Iho7
GSM.CM.99.CM166
GSM.CM.99.CM71
MON.CM.99.L1
MON.NG.x.NC91
MUS.1.CM.01.1085
MUS.1.CM.01.CM1239
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

Table with 4 columns: Sequence alignment, V5 loop end, Rev Responsive Element (RRE) region, and sequence number. The table contains multiple rows of sequence data with alignment markers and specific region labels.

	Env gp120 end	Env gp41 start		Env gp120 end	Env gp41 start	
H1B.FR.83.HXB2	AGAGAAAAAGAGCAGTG.....	GGAATAGGACGCTTTGTCCTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCCTCAATGACGGTACAGGGCAGACAAATATTGTCTGGTATAGTCGACGACGAGCAAAATTTGCTG...	AGGCGTAT 7900
Env	R E K R A V	G I G A F L G F L G A A G S T M G A A S M T L T V Q A R Q L L S G I V Q Q Q N N L L R A I				
H101_AE.TH.90.CM240	G.....	G.....	7453
H102_AG.NG.x.IBNG	G.....	G.....	7406
H103_AB.RU.97.KAL153_2	G.....	G.....	7077
H104_cpx.CY.94.CY032	G.....	G.....	7270
H1A1.UG.85.U455	G.....	G.....	7330
H1B.US.90.WE.AU160	G.....	G.....	7377
H1C.ET.86.ETH2220	G.....	G.....	7276
H1D.CD.84.84ZR085	G.....	G.....	7421
H1F1.BE.93.VI850	G.....	G.....	7177
H1G.SE.93.SE6165	G.....	G.....	7318
H1H.CF.90.056	G.....	G.....	7219
H1I.SE.93.SE7887	G.....	G.....	7207
H1K.CM.96.MP535	G.....	G.....	7060
H1N.CM.95.YBF30	G.....	G.....	7448
H1O.BE.87.ANT170	G.....	G.....	7954
H1O.CM.91.MVP5180	G.....	G.....	7971
CPZ_CD.90.ANT	G.....	G.....	7332
CPZ_CM.01.SIVcpzCAM13	G.....	G.....	7484
CPZ_CM.05.SIVcpzEK505	G.....	G.....	7383
CPZ_CM.05.SIVcpzLB7	G.....	G.....	7411
CPZ_CM.05.SIVcpzMB66	G.....	G.....	7524
CPZ_CM.05.SIVcpzMT145	G.....	G.....	7409
CPZ_CM.98.CAM3	G.....	G.....	7313
CPZ_CM.98.CAM5	G.....	G.....	7619
CPZ_GA.88.GAB1	G.....	G.....	7937
CPZ_GA.88.GAB2	G.....	G.....	7284
CPZ_TZ.01.TANI	G.....	G.....	7526
CPZ_US.85.CPZUS	G.....	G.....	7905
H2A.DE.x.BE1	G.....	G.....	8373
H2A.GW.x.ALI	G.....	G.....	8357
H2A.SN.x.ST	G.....	G.....	7803
H2B.CL.x.EHO	G.....	G.....	8335
H2B.HG.86.D205	G.....	G.....	8349
H2G.CL.x.AB196	G.....	G.....	7734
H2U.FR.96.12034	G.....	G.....	7857
MAC.US.x.239	G.....	G.....	8318
M1C.US.x.251_1A11	G.....	G.....	8314
M1C.US.x.251_BK28	G.....	G.....	8300
M1C.US.x.EMBL_3	G.....	G.....	7449
SMM.SL.92.SL92B	G.....	G.....	7766
SMM.US.x.H9	G.....	G.....	7807
SMM.US.x.PGM53	G.....	G.....	8264
SMM.US.x.SIVsmH635F_L3	G.....	G.....	8340
STM.US.x.STM	G.....	G.....	7969
SAB.SN.x.SAB1C	G.....	G.....	8201
TAN.UG.x.TANI	G.....	G.....	8005
VER.DE.x.AGM3	G.....	G.....	7546
VER.KE.x.9063	G.....	G.....	8048
VER.KE.x.AGM155	G.....	G.....	8028
VER.KE.x.TYO1	G.....	G.....	8024
COL.CM.x.CGU1	G.....	G.....	7310
GRV.ET.x.GRI_677	G.....	G.....	7907
MND_2.CM.98.CM16	G.....	G.....	8050
DRL.x.x.FAO	G.....	G.....	7741
RCLM.GA.x.GAB1	G.....	G.....	7587
RCLM.NG.x.NG411	G.....	G.....	7631
MND_1.GA.x.MNDGB1	G.....	G.....	7507
MND_2.GA.x.M14	G.....	G.....	7979
MND_2.x.x.5440	G.....	G.....	7683
MND.US.x.MNE027	G.....	G.....	7795
LST.CD.88.447	G.....	G.....	7080
LST.CD.88.485	G.....	G.....	7077
LST.CD.88.524	G.....	G.....	7083
LST.KE.x.lho7	G.....	G.....	8155
GSN.CM.99.CM166	G.....	G.....	7837
GSN.CM.99.CM71	G.....	G.....	8349
MON.CM.99.L1	G.....	G.....	7863
MON.NG.x.NG1	G.....	G.....	6488
MUS_1.CM.01.1085	G.....	G.....	7834
MUS_1.CM.01.CM1239	G.....	G.....	7803
MUS_2.CM.01.CM1246	G.....	G.....	7932
MUS_2.CM.01.CM2500	G.....	G.....	7855
DEN.CD.x.CD1	G.....	G.....	7957
DEB.CM.99.CM40	G.....	G.....	7640
DEB.CM.99.CM5	G.....	G.....	7561
TAL.CM.00.266	G.....	G.....	7730
TAL.CM.01.8023	G.....	G.....	7280
SUN.GA.98.L14	G.....	G.....	8223
SYK.KE.x.KE51	G.....	G.....	7560
SYK.KE.x.SYK173	G.....	G.....	7867

Table of PLV Complete Genomes alignments. Columns include sample names (e.g., H1B.FR.83.HXB2, MAC.US.x.239) and sequence coordinates (e.g., 8210, 8211, 8212). The table contains multiple rows of sequence data with various symbols like dashes and letters (A, C, G, T) representing nucleotide bases.

Accession	Region	Sequence	Position
H1B.FR.83.HXB2	Tat Rev intron end	GACCCACCTCCCAACC	8523
H1B.FR.83.HXB2	Tat Rev exon 2	N P P S Y F Q O Q T H I O Q	8524
H101.AE.TH.90.CM240	Tat Rev exon 2	T H L S P O P	8525
H102.AG.NG.x.IBNG	Tat Rev exon 2	CTTC ATCAT	8526
H103.AB.RU.97.KAL153.2	Tat Rev exon 2	TTAC ACCA	8527
H104.cpx.CY.94.CY032	Tat Rev exon 2	TTA AACC	8528
H1A.UG.85.LU45	Tat Rev exon 2	TGCG	8529
H1B.US.90.WEAU160	Tat Rev exon 2	TG	8530
H1C.ET.86.ETH2220	Tat Rev exon 2	TTA CA	8531
H1D.CD.84.84ZR085	Tat Rev exon 2	T G	8532
H1F1.BE.93.VI850	Tat Rev exon 2	TTA G	8533
H1G.SE.93.SE6165	Tat Rev exon 2	TTAC ACCA	8534
H1H.CF.90.056	Tat Rev exon 2	TTG G A	8535
H1J.SE.93.SE7887	Tat Rev exon 2	TTA A	8536
H1K.CM.96.MP535	Tat Rev exon 2	TTA A	8537
H1N.CM.95.YBF30	Tat Rev exon 2	TTA A	8538
H1O.BE.87.ANT70	Tat Rev exon 2	T CA AA ATCA	8539
H1P.CM.91.MVP5180	Tat Rev exon 2	T C G	8540
CPZ.CD.90.ANT	Tat Rev exon 2	A T CAAAAC	8541
CPZ.CM.01.SIVcpzCAM13	Tat Rev exon 2	TTA G A	8542
CPZ.CM.05.SIVcpzEK505	Tat Rev exon 2	TTA G	8543
CPZ.CM.05.SIVcpzL7	Tat Rev exon 2	TTTC C T	8544
CPZ.CM.05.SIVcpzMB66	Tat Rev exon 2	TTA G	8545
CPZ.CM.98.CAM3	Tat Rev exon 2	TTA G A	8546
CPZ.CM.98.CAM5	Tat Rev exon 2	TTA G A	8547
CPZ.GA.88.GAB1	Tat Rev exon 2	TTA GT	8548
CPZ.GA.88.GAB2	Tat Rev exon 2	TTA G A	8549
CPZ.TZ.01.TAN1	Tat Rev exon 2	TTAC A G G	8550
CPZ.US.85.CPZUS	Tat Rev exon 2	TTT A	8551
H2A.DE.x.BEN	Tat Rev exon 2	CC T G G T T A C T C A A C A G A T C C A T C C A C A A G	8552
H2A.GW.x.ALI	Tat Rev exon 2	CC T C G G T T A C T C A A C A G A T C C A T C C A C A A G	8553
H2A.SN.x.ST	Tat Rev exon 2	CC T C G G T T A C T C A A C A G A T C C A T C C A C A A G	8554
H2B.CH.86.D205	Tat Rev exon 2	CC T C T C T T A T C A A C A G A T C C T A T C C G C A A G	8555
H2G.CL.x.ABT96	Tat Rev exon 2	CC T C T C T T A T T T T C A G C A G A T C C A T C C G C A A G	8556
H2U.FR.96.L2034	Tat Rev exon 2	CCA C T C T T A T T T G A A C A G A T C C C T A T C C A A C G G	8557
MAC.US.x.239	Tat Rev intron end	CC A C C T C T T A T T T C A G C A G A C C C A T C C A A C A G	8558
MAC.US.x.239	Tat Rev exon 2	P P S Y F Q O Q T H I O Q	8559
MAC.US.x.251.1A11	Tat Rev exon 2	N P Y P Y T	8560
MAC.US.x.251.BK28	Tat Rev exon 2	CC A C T C T T A T T T C A G T A G A C C C A T C C A A C A G	8561
MAC.US.x.EMB1.3	Tat Rev exon 2	CC A C T C T T A T T T C A G T A G A C C C A T C C A A C A G	8562
SMM.SL.92.SL92B	Tat Rev exon 2	CC C C A T C T T A T C A T C A G C A G A T C C A T C C A C A G G	8563
SMM.US.x.H9	Tat Rev exon 2	CC T C G C G T T R T G T Y A G C A G A T C C C T A T C Y A R A C G	8564
SMM.US.x.PGM53	Tat Rev exon 2	CC T C G T T A T G T T C A G C A G A T C C C T A T C C A A C A G	8565
SMM.US.x.SIVsmH635F.L3	Tat Rev exon 2	CC T C G C T T A T T T C A G C A G A T C C C T A T C C A A C A G	8566
SAB.SN.x.SAB1C	Tat Rev exon 2	CC T C T C T T G T C G T C A G C A G A T C C C T A T C C A A C A G	8567
TAN.UG.x.TAN1	Tat Rev exon 2	CC T C T C T T A T T A T A C A G A T C C A T A T C C A C T G A A C	8568
VER.DE.x.AGM3	Tat Rev exon 2	CC T C T C T T A T T C A G A T C C A A T A T C C G C C T G	8569
VER.KE.x.AGM155	Tat Rev exon 2	CC T C T C T T A T T C A G A T C C A A T A T C C C A C G T G G	8570
VER.KE.x.TYO1	Tat Rev exon 2	CC T C T C T T A T T C A G A T C C A A T A T C C C A A C A G T G	8571
COL.M.C.GU1	Tat Rev exon 2	CC T C T C T T A T T A T A C A G A T C C A T A T C C A C T G A A C	8572
GRV.ET.x.GRI.677	Tat Rev exon 2	CC T C T C T T A T T C A G A T C C A A T A T C C C A G T T C A	8573
MND.2.CM.98.CM16	Tat Rev exon 2	CC T T T G T T T C A G G G A G A C T T C C C G T C C C C A C	8574
DRL.x.FAO	Tat Rev exon 2	CC T T T T T A A G G A C G A T A C C G C C C A G A G A G C	8575
RCM.GA.x.GAB1	Tat Rev exon 2	CC A C T C T T A T T G T C A G A T C C A T C C C A A C A A C A	8576
RCM.NG.x.NG411	Tat Rev exon 2	CC A C T T T G T A T C A T C A G C A G A T C C C T A T C C G A A T	8577
MND.1.GA.x.MNDGB1	Tat Rev exon 2	CC T T T A A A G G G G A T A T C T C C G A C T C G C C	8578
MND.2.GA.x.M14	Tat Rev exon 2	T G T T T A A G G G G A T A T C T C C G C C C C A C	8579
MND.2.x.x.5440	Tat Rev exon 2	T G T T T A A G G G G A T A T C T C C G C C C C A C	8580
MNE.US.x.MNE027	Tat Rev exon 2	CC A C T T C T T A T T C A G C A G A C C A T A T C C G A C A G	8581
LST.CD.88.447	Tat Rev exon 2	T G A A G G A T G C A G T G G G A G A A A A G A G A C C A A C A G A C C	8582
LST.CD.88.485	Tat Rev exon 2	T G A A G G A T G C A G T G G G A G A A A A G A G A C C A A C A G A C C	8583
LST.CD.88.524	Tat Rev exon 2	T G A A G G A T G C A G T G G G A G A A A A G A G A C C A A C A G A C C	8584
LST.KE.x.lho7	Tat Rev exon 2	T G A A G G A T G C A G T G G G A G A A A A G A G A C C A A C A G A C C	8585
GSM.CM.99.CN166	Tat Rev exon 2	TT T T A C A A	8586
GSM.CM.99.CN171	Tat Rev exon 2	TT T T A C A A	8587
MON.CM.99.LU1	Tat Rev exon 2	TT T T A C A A	8588
MON.NG.x.NG1	Tat Rev exon 2	TT T T A C A A	8589
MUS.1.CM.01.1085	Tat Rev exon 2	TT T T A C A A	8590
MUS.1.CM.01.CM1239	Tat Rev exon 2	TT T T A C A A	8591
MUS.2.CM.01.CM1246	Tat Rev exon 2	TT T T A C A A	8592
MUS.2.CM.01.CM1250	Tat Rev exon 2	TT T T A C A A	8593
DEN.CD.x.CD1	Tat Rev exon 2	TT T T A C A A	8594
DEB.CM.99.CM40	Tat Rev exon 2	TT T T A C A A	8595
DEB.CM.99.CM5	Tat Rev exon 2	TT T T A C A A	8596
TAL.CM.00.266	Tat Rev exon 2	TT T T A C A A	8597
TAL.CM.01.8023	Tat Rev exon 2	TT T T A C A A	8598
SUN.GA.98.1.14	Tat Rev exon 2	TT T T A T G T G A A A G G A T T T G C T C C A G A A A C T C C C C A A A C C A A C	8599
SYK.KE.x.KE51	Tat Rev exon 2	CC T G C G	8600
SYK.KE.x.SYK173	Tat Rev exon 2	CC T A C	8601

H1B.FR.83.HXB2
Rev exon 2
H101_AE.TH.90.CM240
H102_AG.NG.x.IBNG
H103_AB.RU.97.KAL153.2
H104_cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WEAU160
H1C.ET.86.ETH2220
H1D.CD.84.84ZR085
H1F1.BE.93.V1850
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.AN170
H1O.CM.91.MVP5180
CPZ_CD.90.ANT
CPZ_CM.01.SIVcpzCAM13
CPZ_CM.05.SIVcpzEK505
CPZ_CM.05.SIVcpzLBT
CPZ_CM.05.SIVcpzMB66
CPZ_CM.05.SIVcpzMT145
CPZ_CM.98.CAM5
CPZ_CM.98.CAM5
CPZ_GA.88.GAB1
CPZ_GA.88.GAB2
CPZ_TZ.01.TAN1
CPZ_US.85.CPZUS
H2A.DE.x.BEN
H2A.GW.x.ALI
H2A.SN.x.ST
H2B.CL.x.EHO
H2B.GH.86.DP05
H2G.CL.x.AB196
H2U.FR.96.L2034

GCTACCACCGCTTGAGAGACTTACTTGTATTGTAACGAGGATGTGGAACTTCTGGGA...
L P P P E R L T L D N E D C C G T S G G
S Y H R L R D L L I V T R I V E L L G
...
Rev end (TAA) in some lineages
...
TGGAACTCTCTACAGTATTGGAGTCAG 8639

MAC.US.x.239
Env
Nef
Rev exon 2
MAC.US.x.251.1A11
MAC.US.x.251.BK28
MAC.US.x.HM13
SMM.SI.92.SI.92B
SMM.US.x.EB9
SMM.US.x.PGM55
SMM.US.x.SIVsmH635F.L3
STM.US.x.STM
SAB.SN.x.SAB1C
PAN.US.x.PAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TYO1
COL.CM.x.CCU1
GRV.ET.x.GRI.677
MND.2.CM.98.M16
DRL.x.x.FAO
RCM.GA.x.GAB1
RCM.NG.x.NG411
MND.1.GA.x.MNDGB1
MND.2.GA.x.M14
MND.2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.52192B
LST.KE.x.lho7
GSN.CM.99.CN166
GSN.CM.99.CN71
MON.CM.99.LN1
MON.NG.x.NG1
MUS.1.CM.01.1085
MUS.1.CM.01.CM1239
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KK51
SYK.KE.x.SYK173

CT-GG-TATT-AGC-ACTG-AG-ACC--C-ATCG-GAGTATACCA-ATC--C-AACC-ATACTCCAGAGGCTCTCTGCGAC-CTACA-A--ATTCCGAGAAGTCCTCA--ACTGAA...
T W L F S N Q L R T L L S R V Y Q P L P Q P I L Q R L S A T L Q R I R E V L R T E
...
L A I O Q L O N L A I E S I P D P P T N T P E A L L C D P T E D S R S P O D *
CT-GG-TATT-AGC-ACTG-AG-ACC--C-ATCG-GAGTATACCA-ATC--C-AACC-ATACTCCAGAGGCTCTCTGCGAC-CTACA-A--ATTCCGAGAAGTCCTCA--ACTGAA...
...
Rev end (TAA) in some lineages
...
Nef start

Table with columns for accession numbers, sequence coordinates, and alignment scores. It includes sequences from HIV strains (e.g., H1B.FR.83.HXB2) and other sources (e.g., MAC.US.x.239), with a reference sequence for Env gp41, gp160 end.

H1B.FR.83.HXB2
 Nef
 H101_AE.TH.90.CM240
 H102_AG.NG.x.IBNG
 H103_AB.RU.97.KAL153_2
 H104_cpx.CY.94.CY032
 H1A1.UG.85.U455
 H1B.US.90.WEAT160
 H1C.ET.86.ETH2220
 H1D.CD.84.8ZR085
 H1F1.BE.93.VI850
 H1G.SE.93.SE6165
 H1H.CF.90.056
 H1I.SE.93.SE7887
 H1K.CM.96.MPS35
 H1N.CM.95.YBF30
 H1O.BE.87.ANT70
 H1O.CM.91.MVP5180
 CPZ.CD.90.ANT
 CPZ.CM.01.SIVcpz.CAM13
 CPZ.CM.05.SIVcpz.EK505
 CPZ.CM.05.SIVcpz.LB7
 CPZ.CM.05.SIVcpz.MB66
 CPZ.CM.05.SIVcpz.MT145
 CPZ.CM.98.CA173
 CPZ.CM.98.CA.M5
 CPZ.GA.88.GAB1
 CPZ.GA.88.GAB2
 CPZ.TZ.01.TANI
 CPZ.US.85.CPZUS
 H2A.DE.x.IB8
 H2A.GW.x.ALI
 H2A.SN.x.ST
 H2B.CL.x.EHO
 H2B.GH.86.D205
 H2G.CL.x.ABT96
 H2U.FR.96.12034

Nef start
 ATGGGTGGCAAGTGGTCAAAAAGTAGTGTGATTGGATGGCCTACTGTAAAGGAAAGAAATGAGACGA GCTGAGCCAGCAGCAG 8878
 M G G K W S K S S V I G W P T V R E R M R R A C T P A A .. 8452
 AA T C CA AG G AG T T A A AG A A CCTA A 8405
 G A C CA AG G AG T A A GCTC CT G A 8058
 A A CA AG G GAG T A G GCTCGAGTGAGCCAGAAAGATGAGGCGAGCTCAA A CCTG A 8287
 TA C CT G G A A A CT 8885
 G AT T C AG A G A A A ACTGATCCAAGGGAAAGAAAGAGAC CT 8423
 G A C A AG G G A A A CCTA 8155
 A A CA AG CGAG A AAC A CCCTA 8317
 A A GA GG G T CAG G GCTC C CT 8176
 GAA A A AG G A G GCACGACCAGCAGCAGACAGGGTGGGAA AC A 8065
 AAAG TT G C C AG A AGAAA C A A CAAACGCAAGAACCA A TAGAGC 8465
 AAA GCA T AG G AA T TGAG G AG A A T A ACTAGAACTTCCTGAGT A TGC C 8965
 GAATGCA AGC C AA T TGCA T AGAA A T A TCCTCCTCTGATCCTCAAC A TGT C 8988
 TAGTCTTGCTTTAAG T TC AGTG G GCA GAAA C AA A A GAG CCAAT 8511
 AAAAGTT G CC TG G G ACAA C TA T A CAGC AGATC 8385
 AA A CC G A GAA T AAG T A AAACGT 8410
 A C C G G C AGAG A C C T GA G A AC AA T 8523
 A C CC AG A AGAAA A G A A A CCAA 8411
 CAA G C G A G AGA C A C A CGA CTCA C TGCTT C 8315
 A C G G G AG T AC G A A CCA T 8621
 GCATTAAG AAC A T C G A GAG C AAG A GGA CCA 8951
 A CAA A T C CA TG A GCAG C A G C T G G ACT C GA 8289
 ATGGGAAACATATTTGGTAGA TGGCCT G C CG AAAG C TCGAAGA C T C TAACA CT A 8561
 AA T G CA AG A AGAA C AA C T A ACTCAGA AACAG 8911
 GGGACGC TATCAGCAG 9306
 AGGACAGCG TATCAGCAG 9302
 GGGACGCG TATCA CAG 8748
 9227
 AGGGCAAC ATACA GCAG 8691
 GCTCAGGGCCAA GCTAAGT A 8787

MAC.US.x.239
 Nef
 MAC.US.x.251_1A11
 MAC.US.x.251_BK28
 MAC.US.x.EMBL_3
 SMM.SL.92.SL92B
 SMM.US.x.H5
 SMM.US.x.RGM53
 SMM.US.x.SIVsmH635F_L3
 STM.US.x.STM
 SAB.SN.x.SAB1C
 TAN.UG.x.TANI
 VER.DE.x.AGM3
 VER.KE.x.9063
 VER.KE.x.AGM155
 VER.KE.x.TYO1
 COL.CM.x.CGU1
 GRV.ET.x.GRI_677
 MND_2.CM.98.CM16
 DRL.x.x.FAO
 RCM.GA.x.GAB1
 RCM.NG.x.NG411
 MND_1.GA.x.MNDGB1
 MND_2.GA.x.M14
 MND_2.x.x.5440
 MNE.US.x.MNE027
 LST.CD.88.447
 LST.CD.88.485
 LST.CD.88.524
 LST.KE.x.jho7
 GSN.CM.99.CN166
 GSN.CM.99.CN71
 MON.CM.99.L1
 MON.NG.x.NG1
 MUS_1.CM.01.1085
 MUS_1.CM.01.CM1239
 MUS_2.CM.01.CM1246
 MUS_2.CM.01.CM2500
 DEN.CD.x.CD1
 DEB.CM.99.CM40
 DEB.CM.99.CM5
 TAL.CM.00.266
 TAL.CM.01.8023
 SUN.GA.98.L14
 SYK.KE.x.KE51
 SYK.KE.x.SYK173

.....GGGACAGA-ATACAATCAG- 9263
 G Q K Y N Q
GGGACAGA-ATACAATCAG- 9259
GGGACAGA-ATACAATCAG- 9245
GGGACAGA-ATACAATCAG- 8394
 GGATGGATA-A-ATACAAT-T- 8714
 GCCTCAGA-ATATTGT-AG- 8752
 GGGACAG-GATACAGT-AG- 9200
 GCCTCAGA-ATATAGT-AG- 9285
 GCCTCAGAGATAAAT-A- 8914
 9087
 8906
 8411
 8903
 8890
 8874
 8151
 8766
 8909
 8606
 8452
 8535
 8378
 8838
 8545
 GGGACAGA-ATACAATCAG- 8740
 8035
 8032
 8035
 9110
 8772
 8772
 8788
 7413
 8780
 8749
 8845
 8801
GTTACTG-AGC-TGAT-AG-AC- 8878
 8559
 8704
 8254
 9181
 8422
 8723

TGGCAAGCAACAGGACGGACTAC ATGGGCAAGGCATTAACA
 TGGCAAGCATCCGACCAAATTGG ATGGGCAAGGCATCTCG

H1B.FR.83.HXB2
Nef
G1 AE.TH.90.CM240
H102.AG.NG.x.IBNG
H103.AB.RU.97.KAL153.2
H104.cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WE.AU160
H1C.ET.86.ETH2220
H1D.CD.84.84ZR085
H1F1.BE.93.VI850
H1G.SE.93.SE6165
H1H.CF.90.056
H1I.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.AN170
H1O.CM.91.MV5180
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpz.CAM13
CPZ.CM.05.SIVcpz.EK505
CPZ.CM.05.SIVcpz.LB7
CPZ.CM.05.SIVcpz.MB66
CPZ.CM.05.SIVcpz.MT145
CPZ.CM.98.CAM3
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.GAB2
CPZ.TZ.01.TANI
CPZ.US.85.CPZUS
H2A.DE.x.BE1
H2A.GW.x.ALI
H2A.SN.x.STI
H2B.CL.x.EHO
H2B.GH.86.D205
H2G.CL.x.ABT96
H2U.FR.96.12034

MAC.US.x.239
Nef
MAC.US.x.251.1A11
MAC.US.x.251.BK28
MAC.US.x.EMBL_3
SMM.SL.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F.L3
STM.US.x.STM
SAB.SN.x.SABIC
TAN.UG.x.TANI
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TYO1
COL.CM.x.CGU1
GRV.ET.x.GRI_677
MND_2.CM.98.CM16
DRL.x.FAO
RCM.GA.x.GAB1
RCM.NG.x.NG411
MND_1.GA.x.MNDGB1
MND_2.GA.x.M14
MND_2.x.x.5440
MNE.US.x.MNE027
LSTCD.88.447
LSTCD.88.485
LSTCD.88.524
LST.KE.x.Jho7
GSN.CM.99.CM166
GSN.CM.99.CM71
MON.CM.99.L1
MON.NG.x.NG1
MUS_1.CM.01.1085
MUS_1.CM.01.CM1239
MUS_2.CM.01.CM1246
MUS_2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173

H1B.FR.83.HXB2
H101.AE.TH.90.CM240
H102.AC.VI.x.IBNG
H103.AB.RU.97.KALI.153.2
H104.cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WEAU160
H1C.ET.86.ETH220
H1D.CD.84.R4Z.R085
H1E1.BE.93.VI850
H1G.SE.93.SE6165
H1H.CF.90.056
H1J.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.ANT70
H1O.CM.91.MVP5180
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.SIVcpzEK305
CPZ.CM.05.SIVcpzL87
CPZ.CM.05.SIVcpzMB66
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CAM3
CPZ.CM.98.CAM15
CPZ.GA.88.GAB1
CPZ.GA.88.GAB2
CPZ.TZ.01.TAN1
CPZ.US.85.CPZUS
H2A.DE.x.BEN
H2A.GW.x.ALI
H2A.SN.x.ALI
H2B.CL.x.EHO
H2B.GH.86.D205
H2G.CL.x.AB196
H2U.FR.96.L2034
MAC.US.x.239
Nef
MAC.US.x.251.LA11
MAC.US.x.251.BK28
MAC.US.x.EMBL_3
SMM.SL.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F.L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.UG.x.TAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.ACM155
VER.KE.x.PY01
COL.CM.x.CGU1
GRV.ET.x.GRI1.677
MND.2.CM.98.CM16
DRL.x.x.FAO
RCM.GA.x.GAB1
RCM.NG.x.NG411
MND.1.GA.x.MNDGB1
MND.2.GA.x.M14
MND.2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.lho7
GSN.CM.99.CN166
GSN.CM.99.CN71
MON.CM.99
MON.NG.x.NG1
MUS.1.CM.01.1085
MUS.1.CM.01.CM1239
MUS.2.CM.01.CM1246
MUS.2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
LAL.CM.01.8023
SYN.GA.98.1.14
SYK.KE.x.KE51
SYK.KE.x.SYK173

Table with columns for different HIV strains and their corresponding nucleotide sequences. Headers include NF-k-B-II, NF-k-B-I, Sp1-III, Sp1-II, and Sp1-I. The table lists sequences for various HIV strains such as H1B.FR.83.HXB2, H101.AE.TH.90.CM240, etc., and includes a MAC.US.x.239 section at the bottom.

	TAR element end		Poly-A signal	
H1B.FR.83.HXB2	..ACTGCTTAAGCCTCAATAAAG..CTTGCCTTGAG..TG.C 9632
H101.AE.TH.90.CM240A		 9199
H102.AC.NG			 9159
H103.AB.RU.97.KALI.53.2	CTAAGGAACCCACTGCTTAA	 8808
H104.cpx.CY.94.CY032			 9047
H1A1.UG.85.U455			GTG- 9092
H1B.US.90.WEAU160			 9632
H1C.ET.86.ETH2220			 9031
H1D.CD.84.847R085			 8975
H1F1.BE.93.VI850			 8903
H1G.SE.93.SE6165A		AA-CC 9074
H1H.CF.90.056			 8950
H1J.SE.93.SE7887			G-CG 8932
H1K.CM.96.MP535			 8604
H1N.CM.95.YBF30			 9182
H1O.BE.87.ANT70	..G-CGCG-	A 9753
H1O.CM.91.MVP5180	..GCG-	A 9749
CPZ.CD.90.ANT			T 9068
CPZ.CM.01.SIVcpzCAM13			AA- 9281
CPZ.CM.05.SIVcpzEK505			CT 9129
CPZ.CM.05.SIVcpzLB7A		CT 9167
CPZ.CM.05.SIVcpzMB66			A- 9277
CPZ.CM.05.SIVcpzMT145			CT 9205
CPZ.CM.98.CAM3			A- 9088
CPZ.CM.98.CAM5			TA 9261
CPZ.GA.88.GAB1			CT 9061
CPZ.GA.88.GAB2			CT 9061
CPZ.TZ.01.TANI	..GCG-TC	A-TG 9323
CPZ.US.85.CPZUS			CT 9697
H2A.DE.x.BEN	CC-G-A--GGCCGGTACTGGGCAGAGCGGCTCCACGCT.TGCTTGCTTAA..AGA--TTCAG-..AGA..A-CAI0233
H2A.GW.x.ALI	CC-G--GGCCGGCACTGGGCAGAGCGGCTCCACGCT.TGCTTGCTTAA..AGA--TTAA-..AGA..A-CAI0227
H2A.SN.x.ALI	CC-G--GGCCGGCACTGGGCAGAGCGGCTCCACGCT.TGCTTGCTTAA..AGA--TTAG-..AGA..A-CA 9672
H2B.CL.x.EHO	CC-GCA--GGCCGGTCTGGGCAGAGGGCTCCACGCT.TGCTTGCTTAA..AAC--TTA-..TA-..AGA..A-CAI0122
H2B.GH.86.D205	CC-GCA--AGGCCAGTGTGGGCAGAGCGGCTCCACGCT.TGCTTGCTTAA..AAC--TTAA-..AGA..A-CAI0150
H2G.CL.x.AB196	CC-GCA--GGCCGGTCTGGGCAGAGGGCTCCACGCT.TGCTTGCTTACAACGA--TTAAA-..AGA..A-TA 9584
H2U.FR.96.12034	CC-AAAG		 9607
	TAR element end		Poly-A signal	3' LTR R repeat end
MAC.US.x.239	CC-GCA--GGCCGGTCTGGGCAGAGTGACTCCAGCT.TGCTTGCTTAA..AG.C--TTCATT-..-AGA..A-TAI0155
MAC.US.x.251_1A11	CC-GCA--GGCCGGTCTGGGCAGAGTGACTCCAGCT.TGCTTGCTTAA..AG.C--TTCATT-..-AGA..A-TAI0150
MAC.US.x.251_BK28	CC-GCA--GGCCAGTCTGGGCAGAGTGGCTCCAGCT.TGCTTGCTTAA..AG.A--TTCATT-..-AGA..A-TAI0125
MAC.US.x.EMB1	CC-GCA--GGCCAGTCTGGGCAGAGTGGCTCCAGCT.TGCTTGCTTAA..AG.A--TTCATT-..-AGA..A-TA 9922
SMM.SL.92.SL92B	CC-GCA--GGCCAGTCTGGGCAGAGTGGCTCCAGCT.TGCTTGCTTAA..AG.A--TTCATT-..-AGA..A-CA 9609
SMM.US.x.H9	CC-GCA--GGCCGGTCTGGGCAGAGGGCTCCAGCT.TGCTTGCTTAA..AG.C--TTCG-..AAC-..AGA..A-CA 9609
SMM.US.x.PGM53	CC-GCA--GGCCGGTCTGGGCAGAGGGCTCCAGCT.TGCTTGCTTAA..AG.A--TTCG-..AAC-..AGA..A-TA 9638
SMM.US.x.SIVsmH635F_L3	CC-GCA--GGCCGGTCTGGGCAGAGGGCTCCAGCT.TGCTTGCTTAA..AG.A--TTCG-..AAC-..AGA..A-TA 9638
STM.US.x.STM	CC-GCA--GGCCGGTCTGGGCAGAGGGCTCCAGCT.TGCTTGCTTAA..AG.A--TTCG-..AAC-..AGA..A-TAI0091
SAB.SN.x.SAB1C	CC-GCA--GGCCAGTCTGGGCAGAGTGGCTCCAGCT.TGCTTGCTTAA..AG.A--TTCAT-..-AGA..A-TAI0189
TAN.UG.x.TANI	CT-AGAGCACTCTGGTAGGCTCTGGTAGACCTCGCTGGCAGTGGCCATTGCCAGTAGCAGAGACTCCGCTTGCTTGCTT.GATTC-CTC
VER.DE.x.AGM3	..TTG--GAAGGCTATTG--TG-..TG-..C-A-..A-T- 9698
VER.KE.x.9063	..TTG--GGAAGGCTATTG--TT-..CG-TG-..CA-AG-..CAG 9218
VER.KE.x.AGM155	..TTG--GAAGGCTATTG--TAC.T-..TG-..CA-AG-..CT 9719
VER.KE.x.TY01	..TTG--GAAGGCTATTG--TTG-..CA-AG-..CT 9698
COL.CM.x.CGU1	..TTG--GAAGGCTATTG--TTG-..CA-AG-..CT 9681
GRV.ET.x.GRI_677	..TTG--C-TATGGCTATTG--TC-..C-TCGC-..AG- 8728
MND_2.CM.98.CM16	..TTG--C-TATGGCTATTG--TC-..C-TCGC-..AG- 9524
DRL.x.x.FAO	..TTG--C-TATGGCTATTG--TC-..C-TCGC-..AG- 9518
RCM.GA.x.GAB1	TCTGG-TAGCTCGT.CAGGAGGCTTCTGGAGGTAGATCACTAGGGGCTGCTGGTCTTGGTCTGGTAGACTTCGCTACTTCTAGTGCCATGAAG-AGGGGAGTAGGCCCTGCTTATCCTAATTGCTCTTATAGCTTGAAGCTCC-A-CAT.TTAGATG-ATGAGT 9622
RCM.NG.x.NG411	CGGGTAGACCGCCAGTGGAGGCTCGGCCGGCTCAACGGGAGAGTACCGCTTGCTTATAGCTTGAAGCTCC-..ATGCCAG-..AGTTTAC.T 9355
MND_1.GA.x.MNDGB1	..TTG--GAAGGCTATTG--TATGCTCTTAGTT-AACCAG 9451
MND_2.GA.x.M14	..TTG--GAAGGCTATTG--TTAAC-TAGAA-..TTAGA-CA 9215
MND_2.x.x.5440	TCTGG-TAGCTCG.TCAGGCT.TTCTGGAAAGGCTTACGGGAGGACACTCTTGGTGGTCTGATAGACTCTAGCAATCCCTAGGCCCTGGAGATGTGGGATTGACTACCCTGCTTGTCTATTATTGGCCC-TA-CTGATTAGAATGAAGT 9538
MNE.US.x.MNE027	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
LST.CD.88.447	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
LST.CD.88.485	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
LST.CD.88.524	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
LST.KE.x.lho7	TAGGGAGC--GGGCTCTAGCTTGTACTTTAAA..GCCTTCAA-TTGCTCATTG-T-AAA-TA 9858
GSN.CM.99.CN166	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
GSN.CM.99.CN71	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
MON.CM.99.L1	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
MON.NG.x.NG1	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
MUS_1.CM.01.1085	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
MUS_1.CM.01.CM1239	AAG-TA--G-GCCCAATA		TAGGCTTTACCAGCC-..G-GCTTGAAGCAAAA-AAAAAAA 8232
MUS_2.CM.01.CM1246	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
MUS_2.CM.01.CM2500	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
DEN.CD.x.CD1	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
DEB.CM.99.CM40	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
DEB.CM.99.CM5	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
TAL.CM.00.266	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
TAL.CM.01.8023	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
UN.GA.98.L14	GAT-A--GGGGAGCTAGAGGCTCCTTGTCTTGCAGTACAAGC-TC-..-AG-..TTT 8845
SVK.KE.x.KE51	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
SVK.KE.x.SYK173	..TTG--GAAGGCTATTG--TAT-..-AGA..A-TA 9631
	GAGG-C-GGGTTGGCCGGCCGGTCTTCGGCAGCTCTCCGTTGCTTGCTTTGCTTACGCTGATCAGCA-..TGTGGTAAACGCAAG-A-..ACGT 9558

Extensive secondary structure in this region?
3' LTR R repeat end, 3' LTR U5 start
H1B.FR.83.HXB2
H101.AE.TH.90.CM240
H102.AC.NG.x.IB.NG
H103.AB.RU.97.KAL153_2
H104.cpx.CY.94.CY032
H1A1.UG.85.U455
H1B.US.90.WEAU160
H1C.ET.86.ETH220
H1D.CD.84.R4ZRO83
H1F1.BE.93.VI850
H1G.SE.93.SE6165
H1H.CF.90.056
H1J.SE.93.SE7887
H1K.CM.96.MP535
H1N.CM.95.YBF30
H1O.BE.87.ANT170
H1O.CM.91.MVP5180
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzLB7
CPZ.CM.05.SIVcpzMB66
CPZ.CM.05.SIVcpzMT145
CPZ.CM.98.CAM3
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.GAB2
CPZ.TZ.01.TANI1
CPZ.US.85.CPZUS
H2A.DE.x.BEN
H2A.CW.x.ALI
H2A.SN.x.ALI
H2B.CL.x.EHO
H2B.GH.86.D205
H2G.CL.x.AB196
H2U.FR.96.12034
3' LTR U5 start
MAC.US.x.239
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMB1_3
SMM.SL.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F_L3
STM.US.x.STM
SAB.SN.x.SAB1C
TAN.UG.x.TANI1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TYO1
COL.CM.x.CGU1
GRV.ET.x.GRI_677
MND_2.CM.98.CM16
DRL.x.x.FAO
RCM.GA.x.GAB1
RCM.NG.x.NG411
MND_1.GA.x.MNDGB1
MND_2.GA.x.M14
MND_2.x.x.5440
MNE.US.x.MNE027
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.lho7
GSN.CM.99.CM166
GSN.CM.99.CM71
MON.CM.99.L1
MON.NG.x.NG1
MUS_1.CM.01.1085
MUS_1.CM.01.CM1239
MUS_2.CM.01.CM1246
MUS_2.CM.01.CM2500
DEN.CD.x.CD1
DEB.CM.99.CM40
DEB.CM.99.CM5
TAL.CM.00.266
TAL.CM.01.8023
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173
3' LTR U5 end_
AG-T-.GT---TGT-CC-.ATC-.T-.CC-AGCC-CC-C-.GGT-.CAAC.TCGG-..-CTCAATA...ATAAGAAGA-CCTGGTCTGTTAGGACCTTTCTGCTT.TGGGAA-CCG-AGCA-----C-----
AG-T-.GT---TGT-CC-.ATC-.T-.CC-AGTC-CC-C-.GGT-.CAAC.TCGG-..-CTCGGTA...ATAAGAAGA-CCTGGTCTGTTAGGACCTTTCTGCTT.TGGGAA-CCG-AGCA-----C-----
AG-C-.GT---TGT-CC-.ATC-.T-.CC-AGTC-CC-C-.GGT-.CAAC.TCGG-..-CTCGGTA...ATAAGAAGA-CCTGGTCTGTTAGGACCTTTCTGCTT.TGAGAA-CCG-AGCA-----C-----
AGTC
AG---.GT---TGT-CC-.ATC-.T-.CC-AGTC-CC-C-.GGT-.CA-C.TCGG-..-CTCGAAT...CATAAGAAGA-CCTGGTCTGTTAGGACCTTTCTGC-----C-----
AG---.GT---TGT-CC-.ATC-.T-.CC-AGTC-CC-C-.GGT-.CA-C.TCGG-..-CTCGACA...CATAAGAAGA-CCTGGTCTGTTAGGACCTTTCTGCTT.TGGGAA-CCG-GCA-----C-----
AG---.GT-G-TGT-CC-.ATC-.T-.CC-AGTC-CC-C-.GGT-.CAAC.TCGG-..-CTCAA...GA-CCTGGTCTGTTAGGACCTTTCTGC-----C-----
G-A-.GCA---TGTGCC-.ATT-AT-CC-CAG...GACAA-CCTGGTCTACTAGGATCC...TGA...AGAA...CT...
-ATCTGAG.CAAGTGCC-.ATTG-GC-CCCTCTCCT...AGGGAA-CCT-GTTACTGGG...TTTCTTATC-AGGC-G-----C-----
AGACTT-.GAGAGA...ATTG-C-.GG-GCA-CCA-CACAAG...GG-.C-GTTCTTACTGGGTTCTCTCTCAGCCAG...GCG-G-A--C-----
-ATCTGAGTCAAGTGCCCT-ATTGAGC.CC-CAC-C-A-GCAAG...GG-.C-GTTCTTACTGGGTTCTCTCTCAGCCAG...GCG-G-A--C-----
-ATCTGAGTCAAGTG-CCT-ATTGAGC.CC-CAC-C-TTGAAG...GG-GAAGTTCTTACTGGGTTCTCTCTCAAACCCAG...GACCCAGGCCAG...GCG-G-A--C-----
-CCGTATAATTG-AGTCAA-TG-TCA-T-C-GCGCC-AGCCTCTAG.A...TG-A-CTCTTACTGGGTTCTCTG-A-CAGGT...G-G-.G-A--C-----
AAGTGTACTC-TGTC-T-TGTTT-TAGTGAG...ACTCTGG-TACT-GAG...TCAGACTAGGTGAGAGCTATCTCGAG-TT-AG-GAGAGTAAA-A-----C-----
G-A-GC.A---TGTGCTGTTT-ACCTC-CAGCA-T-A-GACTCT-GGT-GGGA...TCAGATTCTGTGGCAGAGAGCCTTGG...CT-AGAAAAATCCCTA-C-T-----
AATC-.AGT---TG-.CCCATTT-A...TCTCTCA-TGAACGACTCTGGGGT-G-GA...TCAGATATTGTGGCAGAGTACTGAA...CT-AGAAAA-TCC-AACCAG.
AAG-G.TACTC-TCCAT-TTGTC-TCAATAGAA...ACTCTGG-TACC-GAG...T---TCTGTGGCAGAGATTTGATAT-TAG-G-AGTAGAGAA-----C-----
AG-C-.GT---TGC-CCCAT-TCT.CCTAGTC-CC
CAAGCTAGT-CTCA-AGTAGTC-T-TC-CCTTC-CCC...TGGTTCA-GATC...T-CA-TAGAGAGATTGGACCTTGTGAT-GGGAC-GGCT-CCA-----
AAA-AAA-AAAAAAAAAAAAAAAAAAAA
ACTGCAAGCAA-TGT-CCCGTG-T-TCTC-TCTCC...TTCTAA-CCCTGTTCCAGTGTCTCAG-TAGGGAGC-A...TTACTGG-GCC-AGTATCC-G-CTG-CGG-----
C-GCT-C-CA-CTGT-C---TG-GT.CACC-CCCGA-A-GTG...

9719
9703
9201
8808
9050
9178
9031
8975
8903
9074
8953
8943
8604
9182
9754
9793
9068
9284
9132
9170
9281
9208
9170
9261
981
9146
9326
9781
10359
10353
9672
10242
10269
9599
9607
10279
10274
10249
9264
9613
9638
10184
10315
9892
10036
9784
9325
9815
9778
9794
9778
9623
9518
9731
9465
9557
9215
9529
9664
8588
8589
8594
9957
9414
9410
9448
8260
9420
9508
9484
9430
9679
9227
9158
9338
8845
10006
9068
9597