
HIV-1/HIV-2/SIV Complete Genomes

Introduction	355
Table of HIV-1/HIV-2/SIV Complete Genome Alignments	357
Nucleotide Alignments of HIV-1/HIV-2/SIV Complete Genomes	358

Alignment of HIV-1/HIV-2/SIV Complete Genomes

Complete genomes of HIV-1, HIV-2, and SIV sequences are represented in the alignment as follows:

HIV-1: One or two sequences per subtype from M, N, and O groups

HIV-2: Three A, two B, and one G sequence

SIV: Sequences from viruses infecting chimpanzees, the subspecies of African Green Monkeys—vervet, grivet, sabaues, tantalus—sooty mangabeys and macaques infected with sooty mangabey virus, L’Hoest monkey, and mandril.

Table 1 lists the sequence name, accession number, first author, and publication for each of the sequences in this alignment. See page 34 of this compendium for a review of the phylogeny of SIV, HIV and other lentiviruses. Together, these genomes represent the full breadth of diversity discovered to date in the primate lentivirus lineage. Non-primate lentiviruses such as equine infectious anemia virus (EIAV) and caprine arthritis/encephalitis virus (CAEV) are more distantly related to each other and to the primate lentiviruses, and are not included in this alignment.

The new alignment was based on the previous version, but entirely redone using manual editing with the program BioEdit,

<http://www.mbio.ncsu.edu/RNaseP/info/programs/BIOEDIT/bio-edit.html>

created by Tom Hall. It is now entirely 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 annotation is unchanged from last year’s compendium.

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 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. Three representative genomes have been translated in all three reading frames; HIV-1 subtype B strain HXB2, HIV-2 ALI, and SIV from a Sykes monkey. The translations are provided as a visual aid for finding landmarks in the genomes.

The HIV-2/SIV-SMM *vpx* gene is postulated to be a duplication of the *vpr* gene (Tristem *et al. Nature* 347:341–342 (1990)) 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.

HIV Database compendia from previous years, as well as electronic copies of these alignments in a variety of formats are available on our WWW site at

<http://hiv-web.lanl.gov/HTML/compendium.html>

and

http://hiv-web.lanl.gov/ALIGN_CURRENT/ALIGN-INDEX.html

respectively.

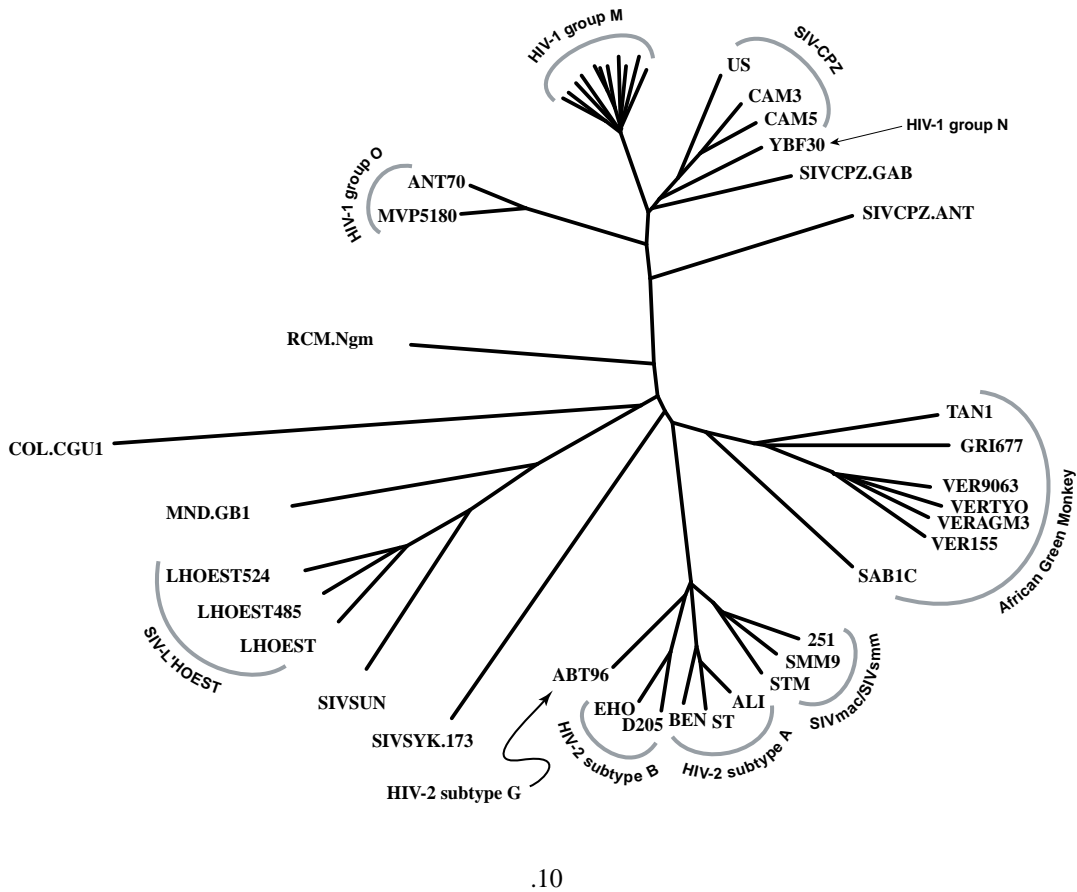


Figure 1. This phylogenetic tree was constructed from the complete lentivirus genome alignment. Columns in the alignment for which any sequence is represented by a gap were first removed. The tree was constructed using PHYLIP DNADIST with maximum likelihood distance estimation and a transition/transversion ratio of 1.6. The DNADIST output was then used to create a neighbor-joining tree using the WEIGHBOR program (The PHYLIP neighbor program gave similar results). WEIGHBOR is available for UNIX, Mac, and PC at

<http://t10.lanl.gov/billb/weighbor/index.html>.

A cut and paste interface to weighbor is available at

<http://bioweb.pasteur.fr/seqanal/interfaces/weighbor.html>

Some of the genomes are known to represent recombinants or mosaics between two or more different lineages, so this tree should not be assumed to represent the true phylogenetic history of these viruses. For example, the SABIC genome from the *sabaeus* subspecies of African green monkeys is recombinant between an African green monkey lentiviral lineage and a sooty mangabey viral lineage. When the AGM-like regions of its genome are used to build a phylogenetic tree, the SABIC lineage shares the same major branch with the other AGM lineages. Given these limitations, the tree is still useful for graphically representing the diversity and relationships between these viruses. It is readily apparent that the HIV-2 viruses are related to sooty mangabey viruses.

Table 1: Table of HIV-1/HIV-2/SIV Complete Genome Alignments

Name	Accession	Author	Reference
01_AE.TH.CM240	U54771	Carr, JK	<i>J Virol</i> 70 (9):5935–5943 (1996)
02_AG.NG.IBNG	L39106	Howard, TM	<i>ARHR</i> 10 (12):1755–1757 (1994)
03_AB.RU.KAL153-2	AF193276	Liitjola, K	<i>AIDS</i> 12 (14):1907–1919 (1998)
04_cpx.CY.94CY032-3	AF049337	Gao, F	<i>J Virol</i> 72 (12):10234–10241 (1998)
A.UG.U455	M62320	Oram, JD	<i>ARHR</i> 6 (9):1073–1078 (1990)
B.FR.HXB2	K03455	Wong-Staal, F	<i>Nature</i> 313 (6000):277–284 (1985)
C.ET.ETH2220	U46016	Salminen, MO	<i>ARHR</i> 12 (14):1329–1339 (1996)
D.CD.84ZR085	U88822	Gao, F	<i>J Virol</i> 72 (7):5680–98 (1998)
F1.BE.VI850	AF077336	Carr, JK	<i>Virology</i> 269 (1):95–104 (2000)
G.SE.SE6165	AF061642	Carr, JK	<i>Virology</i> 247 (1):22–31 (1998)
H.CF.90CF056	AF005496	Murphy, E	<i>ARHR</i> 9 (10):997–1006 (1993)
J.SE.SE7887	AF082394	Laukkanen, T	<i>ARHR</i> 15 (3):293–297 (1999)
K.CM.MP535	AJ249239	Peeters, M	Unpublished
N.CM.YBF30	AJ006022	Simon, F	<i>Nature Med</i> 4 (9):1032–1037 (1998)
O.CM.ANT70	L20587	Vanden Haesevelde, M	<i>J Virol</i> 68 (3):1586–1596 (1994)
O.CM.MVP5180	L20571	Gurtler, LG	<i>J Virol</i> 68 :1581–1585 (1994)
H2A.DE.BEN	M30502	Kirchhoff, F	<i>Virology</i> 177 (1):305–311 (1990)
H2A.GW.ALI	AF082339	Azevedo-Pereira, JM	Unpublished (1998)
H2A.SN.HIV2ST	M31113	Kumar, P	<i>J Virol</i> 64 (2):890–901 (1990)
H2B.CI.EHO	U27200	Rey-Cuille, MA	<i>Virology</i> 202 (1):471–476 (1994)
H2B.GH.D205	X61240	Kreutz, R	<i>ARHR</i> 8 (9):1619–1629 (1992)
H2G.CI.ABT96	AF208027	Brennan, CA	<i>ARHR</i> 13 (5):401–404 (1997)
COL.-.SIV-COL-CGU1	-	Peeters, M	Unpublished (2000)
CPZ.CD.CPZANT	U42720	Vanden Haesevelde, MM	<i>Virology</i> 221 (2): 346–350 (1996)
CPZ.CM.CAM3	AF115393	Corbet, S	<i>J Virol</i> 74 : 529–534 (2000)
CPZ.CM.CAM5	AJ271369	Souquiere, S	Unpublished
CPZ.GA.CPZGAB	X52154	Huet, T	<i>Nature</i> 345 (6273):356–359 (1990)
CPZ.US.CPZUS	AF103818	Gao, F	<i>Nature</i> 397 (6718):436–441 (1999)
GRIVET.ET.AGM_GRI-677	M66437	Fomsgaard, A	<i>Virology</i> 182 (1):397–402 (1991)
LHOEST.CD.SIVlhoest447	AF188114	Beer, BE	<i>J Virol</i> 74 (8):3892–3898 (2000)
LHOEST.CD.SIVlhoest485	AF188115	Beer, BE	<i>J Virol</i> 74 (8):3892–3898 (2000)
LHOEST.CD.SIVlhoest524	AF188116	Beer, BE	<i>J Virol</i> 74 (8):3892–3898 (2000)
LHOEST.KE.SIVlhoest	AF075269	Hirsch, VM	<i>J Virol</i> 73 (2):1036–1045 (1999)
MAC.US.MM251	M19499	Franchini, G	<i>Nature</i> 328 (6130):539–543 (1987)
MANDRIL.GA.MNDGB1	M27470	Tsujimoto, H	<i>Nature</i> 341 :539–541 (1989)
RCM.NG.SIV-RCM-NGM	-	Beer, B	Unpublished (2000)
SABAEUS.SN.SAB1C	U04005	Jin, MJ	<i>EMBO J</i> 13 (12):2935–2947 (1994)
SMM.US.SIVSMMH9	M80194	Courgnaud, V	<i>J Virol</i> 66 :414–419 (1992)
STM.US.STM	M83293	Novembre, FJ	<i>Virology</i> 186 (2):783–787 (1992)
SUN.GA.SIVSUN	AF131870	Beer, BE	<i>J Virol</i> 73 (9):7734–7744 (1999)
SYKES.KE.SYK173	L06042	Hirsch, VM	<i>J Virol</i> 67 (3):1517–1528 (1993)
TANTALUS.UG.TAN1	U58991	Stivahtis, GL	<i>Virology</i> 228 :394–399 (1997)
VERVET.DE.AGM3	M30931	Baier, M	<i>Virology</i> 176 (1):216–221 (1990)
VERVET.KE.AGM155	M29975	Johnson, PR	<i>J Virol</i> 64 (3):1086–1092 (1990)
VERVET.KE.AGMTYO	X07805	Fukasawa, M	<i>Nature</i> 333 :457–461 (1988)
VERVET.KE.AGM_VER-9063	L40990	Hirsch, VM	<i>J Virol</i> 69 (2):955–967 (1995)

B_FR_HXB2	GATGGTGCTCAAGCTAGTACCACTGAGCCAGAGAGTTAGAGAA...GCCAACAAAGGAGAGAACACACCGCTTTTACCCCTGTGAGC.....CTGCATGGAATGGATGACCCCGGAGAGAGAAGT	250
A_UG_U455	-G-----C-C-CTG-AG-----G.....CAG-G-----A-AT-----AA--C-----TGA--A-A-----	0
C_ET_ETH2220	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
D_CD_84ZR085	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
F1_BE_VI850	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
G_SE_SE6165	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
H_CF_90CF056	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
I_SE_SE7887	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
K_CM_MP535C	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
O1_AE_TH_CM240	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
O2_AG_NG_IBNG	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
O3_AB_RU_KAL153-2	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
O4_CPX_CV_94CY032-3	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
N_CM_YBF30	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
O_CM_MP5180	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	254
O_BE_ANT70	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	253
CPZ_US_CPZUS	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	250
CPZ_GA_CPZGAB	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	250
CPZ_CM_CAM3	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
CPZ_CM_CAM5	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
CPZ_CD_CPZANT	-GTCAG--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC--T-A-CTT-T.....AAT-----GCT--G--TG-AC-CG-G--GA-	0
H2A_GW_ALI_Nef	G_W_L_W_K_L_V_P_V_D.....V_P_O_E.....G_E_D_T_E_T_H_C_L_L_H_P_V_O.....T_S_R_H_D_D_T_H_G_E_T	Nef
H2A_GW_ALI	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	244
H2A_SN_HIV2ST	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	0
H2B_DE_BEN	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	244
H2B_GH_D205	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	247
H2B_CI_RHO	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	244
H2G_CI_ABT96	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	244
MAC_US_WM251	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	244
SMM_US_SIVSMMH9	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	0
STM_US_STM	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	0
TANTALUS_UG_TANI	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	241
GRIVET_ET_GRI-677	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	250
VERVET_KE_VER-9063	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	250
VERVET_KE_AGMTYO	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	0
VERVET_KE_AGMT155	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	250
VERVET_DE_AGM3	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	0
SABAEUS_SN_SABIC	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	244
SUN_GA_SIVSUN	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	247
MANDRIL_GA_MNDGB1	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	247
LHOEST_KE_loeest447	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	0
LHOEST_CD_loeest448	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	0
LHOEST_CD_loeest524	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	244
RCM_NG_SIV-NGM	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	0
COL_...SIV-CGU1	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	0
SYKES_KE_SYK173	-G-CTG-GG--A--T.....GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----GTCCC-C-----	138

B.FR.HXB2	GTTAGAGTGGAGGTTTGACACGCCCGCTAGCAATTCATCATGGCCCGAGAGCTGCATCCCGGAGTACTTCAAGAACTCGTGTACA.....	TCF-1 alpha	334
A.UG.U455	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
C.ET.ETH2220	AC--A-A--CA--ATCAT--T--ATCAT--GC--AACA--T--A-----TGC-AAA--C--A--CT-----CCC--G--..-A--TGACACTGC.....		0
D.CD.84ZR085	AC-GAT--A-----T--ATCT--GCAACACC--TG--T--TAIGATAACT--C--A--CT-----C-----GGA--A--A--ACTGCTG.....		0
F1.BE.VI850	C--ATC--CA-----CTGAA--G--TCGAGA--GA--TAA-----A-----CG-----A-----A-----ACTTTT--G.....		0
G.SE.SE6165	C--G--TC--C--C-----A--G--G--C--AAGA--T--T--A-----A--AA-----A-----A-----ACTTTT--G.....		0
H.CF.90CF056	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
J.SE.SE7887	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
K.CM.WP535C	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
01.AE.TH.CM240	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
02.AG.NG.IBNG	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
03.AB.RU.KAL153-2	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
04.CPX.CY.94CY032-3	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		325
N.CM.YBF30	AC--A-A--CA--ATCAT--T--ATCAT--GC--AACA--T--A-----TGC-AAA--C--A--CT-----CCC--G--..-A--TGACACTGC.....		0
O.BE.ANT70	AC-GAT--A-----T--ATCT--GCAACACC--TG--T--TAIGATAACT--C--A--CT-----C-----GGA--A--A--ACTGCTG.....		343
CPZ.US.CPZUS	C--ATC--CA-----CTGAA--G--TCGAGA--GA--TAA-----A-----CG-----A-----A-----ACTTTT--G.....		344
CPZ.GA.CPZGAB	C--G--TC--C--C-----A--G--G--C--AAGA--T--T--A-----A--AA-----A-----A-----ACTTTT--G.....		329
CPZ.CH.CAM3	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		335
CPZ.CD.CPZANT	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
H2A.GW.ALI Nef	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
H2A.GW.ALI	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
H2A.SN.HIV2ST	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
H2A.DE.BEN	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
H2B.GH.D205	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
H2B.CI.BHO	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
H2G.CI.BHT96	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
MAC.US.MM251I	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
SMW.US.SIVSMWH9	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
STM.US.STM	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
TANTALUS.UG.TANI	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
GRIVET.ET.GRI-677	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
VERVET.KE.VBR-9063	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
VERVET.KE.AGMTYO	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
VERVET.KE.AGM155	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
VERVET.DE.AGM3	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
SABAEUS.SN.SAB1C	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
SUN.GA.SIVSUN	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
MANDILL.GA.MNDGB1	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
LHOEST.KE.roest	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
LHOEST.CD.roest447	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
LHOEST.CD.roest485	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
LHOEST.CD.roest524	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
RCM.NG.SIV-NGM	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
COL.SIV-CGU1	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
SYKES.KE.SYK173	-----AT-----A-----T-----G-----TC-AACA-----GA-----T-----A-----AG-----CAGACT.....		0
L_V_W_R_F_D_P_K_L_A_H_D_Y_K_A_F_I_L_H_P_E_E_F_G_Y_K_S_G_L_P_E_D_E_W_K_A_R_L_K_A_R_G_I_A-----TT-----A-----CCTAAG--G--TCA-G--T--AA--TTTATT--A--C--A--G--A--TGG--T--AAG--C--GGCCTGCCAGAAGATGAGTGGAGGCAAGACTGAAAAGCAAGAGGGAT.		Nef	373
A-----T-----CA-----TC-ATG--G--C-A-A-CT--A-----TTCACT--T-C-A-A-AG-G--TGG-C--AAAG--C--GGATTGCCAGAAAAGAAATGGAAGGCAAAACTGAAAAGCAAGAGGGAT.			0
TC-TATC--CA-----TTC--T--G--A-G--T-TG--TTTCA-TAG--TT--A--AG-G--GG-T--CAG-C--GGCTACCAGAGGAGTGGAGGCTAGACTAAAAGCAAGAGGGAT.			373
CC-T-TC--CA-----TC--T-----A-G--CT-TG-----TTCAGCAG--TTC--A--G-G--TGG-T--TCAG-C--GGAATGCCAGAAAAGAGTGGAAAGGCTAAACTGNAGCAAGAGGAAT.			373
TC--C-----A-----TCCAACT-----C-ACAC-T-TGA---ATAT-CTAGAT-C-A--AG-G--GG-AGCAA--TCAGGCCCTGTACAGGAAGAGGTTAGAAGAAGGCTAACCCG.			363
-----AA-----TCCA-T-----G-----GT-----ACC-GGAC.....AGCTTTA--G--A-GC--TGGTCTTGTFTA--GAGAAATAGCCTAACCCAGGCTTCCGGT.....			0
A--G-T-----AA-----CCGAAGT--G--GG-GT--CGCC-GGAC.....ATGTTTA--G--A-GC--CG-ACATGC--A--GCC...TAGTTTCGCTGGTTAGCCTAACCCGTTG.....			37
C--G--TC--A-----CCTATGT-----G--G--GT--TGACC-AAACAGA-A-T--TTT--AG--A-GC--TGCAATAGTGA--GAGGAAGTACTAACCCGAGGCTTGTG.....			349
C--G--TC--A-----TCCAATGT--G--TG-----GT-----GACC--TC-AG-GA-T-CTTTACAG-----AT--TTCAAACAGTTGGTACAGGAAACTAGCCGACACACAGGCTTG.....			361
A--G--CA--A-----CCA-TA-----G-GG-CT-TG--T--TT-GAG-----C--A--C-GG-GCC--TCAGCACA--GGCTAACCAAAAACCCACATCCTATG.....			0
CC-G-T-----A-----TCC-ACT--G--C-A-G--TTT-G-----AT-C-AA-----C-T--AG-G-----C-TGTTACTAGTGTGAGTGGAGGTAGCTTAAACCCGAAACCG.			352
AC-TATC-----CA-----TCCAATGT--G--GG-CITTT--A-----TCA-AAA--G--T-C-A-GA-AC-A--TACTGCTTTCCGCTTTTACTGCAAGAGGAAGTACCCAAACCGC.			365
-C-TATC--CAC--CA--TCCAATGT--G--GG-CITTT--A-----TCA-AAA--G--T-C-A-GA-AC-A--TACTGCTTTCCGCTTTTACTGCAAGAGGAAGTACCCAAACCGC.			365
A-----CA-----A-----CA-TCC--T-T-----G-AG--T--GCA-GGTATAGA-----T-----TGGAG--GA-AAAGA--CAAGACCAAGTAACCCACTTCCCTCTGGAGTTGCCITTG.....			0
-----CA-----AC--CA--CC--A--G--C-A-ACA-CAGGCTGGGAGAT--GCC-GCTT--C---TAGAG--GAC-AAACAG--A--ACCACAAGAGCTGCAGTCAGCACTTTAGTAAGAACAATCTCCTTAGGA.			364
CC-CAT-----AC--CA--CC--A--G--C-A-ACA-CAGGCTGGGAGAT--GCC-GCTT--C---TAGAG--GAC-AAACAG--A--ACCACAAGAGCTGCAGTCAGCACTTTAGTAAGAACAATCTCCTTAGGA.			267

See Pohlmann, J Virol 72(7):5589-5598(1998) and similar publications for information on this enhancer region

B.FR.HXB2TCGAGCTTCT.ACA...AGGACTTTCCG.....CTGGGACTTTCCAGGG.	NF-kappa-B-I Sp1-III	377
A.UG.U455			0
C.ET.ETH2220			0
D.CD.84ZE085			0
F1.BE.VI850			0
G.SE.SEG165			0
H.CF.90CF056			0
J.SE.SE7887			0
K.CM.MP535C			0
O1.AE.TH.CM240			0
02.AG.NG.IBNG			0
03.AB.RU.KAL153-2GCTGACA-GA-.G-----G-----GC-----G		374
04.CPX.CY.94CY032-3			0
N.CM.YBF30			0
O.CM.MVP5180			385
O.BE.ANT70ACG-GA-A-----G-----CTGT--A-----A-----GACTGTGACACTGC-----C-T		413
CPZ.US.CPZUSTGCGTAAACCGCGAGGGCC.AATA-AC-----G-----TC...A-----TAA...GGGACTTTCCAA-----G-----AAG		404
CPZ.GA.CPZGARACTGGCCCAITGCC.ACARGAC-----G-----TCIGC-----AA-----G-----G-----A		396
CPZ.CM.CAM3			0
CPZ.CM.CAM5	CGCCCTGCGCAG-A--AAC-----G-----A-----TAA.....		54
CPZ.CD.CPZANT			0
H2A.GW.ALI Nef			Nef
H2A.GW.ALITTTGGTTCAGGCGGAGGAAGTAACACTG.....AAA-C.AGCTGAGACTGC-----A-----GAA--GGC-GTA-CCAG		474
H2A.SN.HIV2ST			0
H2A.DE.BENATAACAGGAACAACATCTTGGTCAAGGAGGAGTAGCTA.CTAAGAA-C.AGCTGAGGCTGC-----A-----GAA--GGC-GTA-CCAA		477
H2B.GH.D205TAAGCTTAACAGGACAGCAGCATAACAGGAAGACTAGCTACACTGCACAA.CTAAGAA-CTAGCAGACTGC-----A-----AAA--AGGC-GTA-CA-T		483
H2B.CI.BHOTAAGCTTAACAGGACAGCAGCATAACAGGAAGACTAGCTACACTGCACAA.CTAAGAA-CTAGCAGACTGC-----A-----AAA--AGGC-GTA-CA-T		481
H2G.CI.ABT96			0
MAC.US.MM251AAGAGCCCTTCTTAACATGGCTGACAAGAGGAA-CTCGCTGAGATAGC-----A-----CAA--GGA-GTT-T--G		441
SMM.US.SIVSMMH9			0
STM.US.STMAAGAGCCCTTCTTAAGATGGCTGACAAGAGGAA-CAAGCTGAGACAGC-----A-----CAA--GGC-GTA-.CAG		114
TANTALUS.UG.TANITAAGCTTAACAGGACAGCAGCATAACAGGAAGACTAGCTACACTGCACAA.CTAAGAA-CTAGCAGACTGC-----A-----AAA--AGGC-GTA-CA-T		417
GRIVET.ET.GRI-677			361
VERVET.KE.VBR-9063GTTAAGCCGTTGCCAGGGAGATGACAT-T--AAC-----G-----G-----A-----A-----C		431
VERVET.KE.AGMTYO			0
VERVET.KE.AGM155CGGTTTCCCTGGTTGCCCTAGGAGATGACAT-AAAGAAC-----G-----G-AC-----A--C		435
VERVET.DE.AGM3			0
SABAEUS.SN.SABLCGAGTTGTCATGTTGATGACAT-AAAGAAC-----G-----TCAG-----A-----GGG-GGAGACTG		415
SUN.GA.SIVSUNCATCCTCTTGCATCGCTAGGCAACGGGCTAGC.GCAT.....GC--GCTAGAT-----A-----GGG-GGAGACTG		427
MANDRILL.GA.MNDGB1			0
LHOEST.KE.roestAAACCGGACATCCTCTTTGCGAGCCCGGTTGCTAAGGCACCGGGCTAGCCGATGCGCAATG-----G-----GAGC-----GGACGGG-G-----GG--T.....		473
LHOEST.CD.roest447			0
LHOEST.CD.roest485			0
LHOEST.CD.roest524			0
RCM.NG.SIV-NGMGAAACAGGCATGCTACTAAGA--A-GCTGACACAGA-----ACTGGCGGCTGCGCGCTGGTGTAA-----A--ACT		452
COL.....SIV-CGU1			0
SYKES.KE.SYK173GGGG-CCG-----C-----GGGC.....GGGG-AA-G--CA--CT-CT		329

HIV-1/HIV-2/SIV complete genomes

	TAR element bulge loop	TAR element stem	
B.FR.HXB2CCAGATCTGAGCTGGAGCTCTCTGGCTA.ACTAGGAAACCACTGCTTAA.....AGCCT.....	525
A.UG.U455G-.....C-.....G-A.....A.....G-GA.....G-A.....	0
C.ET.ETH2220G-.....C-.....G-AG-A-C-..ACTG.....G-AG-A-C-..ACTG.....	0
D.CD.842E085G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	39
F1.BE.VI850G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	0
G.SE.SEG165AA-.....C-.....G-.....G-.....G-.....G-.....	0
H.CF.90CF056T-.....T-.....A-.....C-GTGA.....A-.....C-GTGA.....	0
J.SE.SE7887T-.....T-.....A-.....C-GTGA.....A-.....C-GTGA.....	0
K.CM.MP535CTA-.....TA-.....GCC-GC.....GCC-GC.....	0
01.AE.TH.CM240G-.....C-.....G-AG-A-C-..ACTG.....G-AG-A-C-..ACTG.....	85
02.AG.NG.IBNGG-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	533
03.AB.RU.KAL153-2G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	559
04.CPX.CV.94CX032-3AA-.....C-.....G-.....G-.....G-.....G-.....	550
N.CM.YBF30T-.....T-.....A-.....C-GTGA.....A-.....C-GTGA.....	544
O.CM.MVP5180G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	200
P.BE.ANT70G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	0
CPZ.US.CPZUSG-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	533
CPZ.GA.CPZGABG-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	559
CPZ.CM.CAM3G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	550
CPZ.CM.CAM5G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	544
CPZ.CD.CPZANTG-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	200
H2A.GW.ALIG-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	683
H2A.SN.HIV2STG-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	132
H2A.DE.BENG-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	687
H2B.GH.D205G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	693
H2B.CI.BHOG-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	690
H2G.CI.ABT96G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	55
MAC.US.WM251G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	639
SMH.US.SIVSMH9G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	134
STM.US.STMG-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	324
TANTALUS.UG.TANIG-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	574
GRIVET.ET.GRI-677G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	405
VERVET.KE.VER-9063G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	585
VERVET.KE.AGMTYOG-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	81
VERVET.KE.AGM155G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	589
VERVET.DE.AGM3G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	81
SABAEUS.SN.SABICG-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	644
SUN.GA.SIVSUNG-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	595
MANDRIL.GA.MNDGB1G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	150
LHOEST.KE.loeest447G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	629
LHOEST.CD.loeest448G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	0
LHOEST.CD.loeest485G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	0
LHOEST.CD.loeest524G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	685
RCM.NG.SIV-NGMG-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	98
COL.SIV-CGU1G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	475
SYKES.KE.SYK173G-.....C-.....CT.CTAGCT.....G.....CT.CTAGCT.....G.....	475

see Zhang, J. Virol 72(7):5664-5471(1998)
5'LTR U5 end <-| Lys tRNA primer binding site
AAAAATCTTAGCAG..TGCGCCCGAACACAGGG.ACCITGAA.....AGCGAAAGG.

B.FR.HXB2	AACTAGAGATCCCTCAGACCCCTTTTAGTCAGTGTGG.....	AC-A-A-T--A.....TC-	667
A.UG.U455	-----AGC-----C-G-----	-----G-----T-	136
C.ET.ETH2220	-----AC-C--A-T-A-A.....	-----G-----T-	46
D.CD.84ZB085	-----AGC-----C-G-----	-----G-----T-	181
F1.BE.VI850	-----AC-C--A-T-A-A.....	-----G-----T-	17
G.SE.SE6165	-----AC-C--A-T-A-A.....	-----G-----T-	50
H.CF.90CF056	-----AC-C--A-T-A-A.....	-----G-----T-	14
J.SE.SE7887	-----AC-C--A-T-A-A.....	-----G-----T-	0
K.CM.MP535C	-----AC-C--A-T-A-A.....	-----G-----T-	0
01.AE.TH.CM240	-----AC-C--A-T-A-A.....	-----G-----T-	214
02.AG.NG.IBNG	-----AC-C--A-T-A-A.....	-----G-----T-	195
03.AB.RU.KAL153-2	-----AC-C--A-T-A-A.....	-----G-----T-	672
04.CPX.CV.94CY032-3	-----AC-C--A-T-A-A.....	-----G-----T-	14
N.CM.YBF30	-----AC-C--A-T-A-A.....	-----G-----T-	225
O.CM.MVP5180	-----AC-C--A-T-A-A.....	-----G-----T-	669
O.BE.ANT70	-----AC-C--A-T-A-A.....	-----G-----T-	696
CPZ.US.CPZUS	-----AC-C--A-T-A-A.....	-----G-----T-	689
CPZ.GA.CPZGAB	-----AC-C--A-T-A-A.....	-----G-----T-	684
CPZ.CM.CAM3	-----AC-C--A-T-A-A.....	-----G-----T-	36
CPZ.CM.CAM5	-----AC-C--A-T-A-A.....	-----G-----T-	340
CPZ.CD.CPZANT	-----AC-C--A-T-A-A.....	-----G-----T-	13
H2A.GW.ALI	GGTGTTCACCTG.AGTA--AAGACCCCTGTC--TAGGACCCCTTCTGCTTTGGAAACGGAGCAGG--C	GT--GTT--GAA--CT-A.	885
H2A.SN.HIV2ST	GGTGTTCATCTAAAGTA--AAGACCCCTGTC--TAGGACCCCTTCTGCTTTGGAAACCAAGCCAGG--C	GT--GTT--GAA--CT-A.	336
H2A.DE.BEN	GGTGTTCATCTAGT-AC.AAGACCCCTGTC--TAGGACCCCTTCTGCTTTGGAAACCAAGCCAGG--C	GT--GTT--GAA--CT-A.	890
H2B.GH.D205	GGTGTCC-CTGA.....TAGACCCCTGTC--TAGGACCCCTTCTGCTTTGGAAACCAAGCCAGG--C	GT--GTT--GAA--CT-A.	886
H2B.CI.BHO	GGTGTCC-CTAG.....AAGACCCCTGTC--TAGGACCCCTTCTGCTTTGGAAACCAAGCCAGG--C	GT--GTT--GAA--CT-A.	886
H2B.CI.ABT96	GGTGTTC-CTAGC.....AAGACCCCTGTC--TAGGACCCCTTCTGCTTTGGAAACCAAGCCAGG--C	GT--GTT--GAA--CT-A.	258
MAC.US.WM251	GGTGTTC-G.AAVTAAG.....AAGACCCCTGTC--TAGGACCCCTTCTGCTTTGGAAACCAAGCCAGG--C	GT--GTT--GAA--CT-A.	840
SMI.US.SIVSMH9	GGTGTTC-G.AAVTAAG.....AAGACCCCTGTC--TAGGACCCCTTCTGCTTTGGAAACCAAGCCAGG--C	GT--GTT--GAA--CT-A.	840
STM.US.STM	GGTGTTC-G.AAVTAAG.....AAGACCCCTGTC--TAGGACCCCTTCTGCTTTGGAAACCAAGCCAGG--C	GT--GTT--GAA--CT-A.	336
TANTALUS.UG.TANI	GGTGTTC-G.AAVTAAG.....AAGACCCCTGTC--TAGGACCCCTTCTGCTTTGGAAACCAAGCCAGG--C	GT--GTT--GAA--CT-A.	517
GRIVET.ET.GRI-677	CTA--G--A-----GTTA-TGG.....GTTTCTTAT.....CCAGGCTGAG--T	GT--GTT--GAA--CT-A.	738
VERVET.KE.VER-9063	CTAG--GTGAA--TC.TTAC-G.....GGTT..CTCCTGTACCAGGTGG-G-A-C	GT--GTT--GAA--CT-A.	550
VERVET.KE.AGMTYO	-C-GTTTCT-A.TG.....GGTTCTCTCTGACCCAGCCGG-G-A-C	GT--GTT--GAA--CT-A.	765
VERVET.KE.AGMT155	-T..TTTCT-A.TG.....GGTTCTCTCTGACCCAGCCGG-G-A-C	GT--GTT--GAA--CT-A.	251
VERVET.DE.AGM3	G-AGTTTCT-A.TG.....GGTTCTCTCTGACCCAGCCGG-G-A-C	GT--GTT--GAA--CT-A.	751
SABAUS.SN.SABIC	---.TTTCT-A.TG.....GGTTCTCTCTGACCCAGCCGG-G-A-C	GT--GTT--GAA--CT-A.	250
SUN.GA.SIVSUNGGTTCTCTCTGACCCAGCCGG-G-A-C	GT--GTT--GAA--CT-A.	803
MANDRIL.GA.MNDGB1GGTTCTCTCTGACCCAGCCGG-G-A-C	GT--GTT--GAA--CT-A.	802
LHOEST.KE.loeest447GGTTCTCTCTGACCCAGCCGG-G-A-C	GT--GTT--GAA--CT-A.	305
LHOEST.KE.loeest485GGTTCTCTCTGACCCAGCCGG-G-A-C	GT--GTT--GAA--CT-A.	823
LHOEST.KE.loeest524GGTTCTCTCTGACCCAGCCGG-G-A-C	GT--GTT--GAA--CT-A.	0
LHOEST.CD.loeest485GGTTCTCTCTGACCCAGCCGG-G-A-C	GT--GTT--GAA--CT-A.	0
LHOEST.CD.loeest524GGTTCTCTCTGACCCAGCCGG-G-A-C	GT--GTT--GAA--CT-A.	887
RCM.NG.SIV-NGMGGTTCTCTCTGACCCAGCCGG-G-A-C	GT--GTT--GAA--CT-A.	295
COL.....SIV-CGU1GGTTCTCTCTGACCCAGCCGG-G-A-C	GT--GTT--GAA--CT-A.	634
SYKES.KE.SYK173GGTTCTCTCTGACCCAGCCGG-G-A-C	GT--GTT--GAA--CT-A.	

HIV-1/HIV-2/SIV complete genomes

HIV-1/HIV-2/SIV complete genomes

Gag p7 nucleocapsid binds these loops, see DeGuzman, Science 279:384-388(1988) packaging signal secondary structure, see Harrison J Virol 72:5886-5896(1998)

Table with columns for strain names (e.g., B.FR.HXB2, A.UG.U455, H2A.SN.HIV2ST), nucleotide sequences, and positions (e.g., 752, 217, 130, 267, 103, 155, 98, 72, 0, 319, 294, 774, 118, 307, 760, 785, 777, 768, 124, 425, 89, 998, 449, 1003, 996, 995, 367, 949, 444, 620, 852, 665, 875, 364, 862, 362, 918, 917, 419, 937, 0, 0, 1000, 410, 749). The table lists the complete genomes for various HIV-1, HIV-2, and SIV strains, showing their nucleotide sequences and corresponding positions.

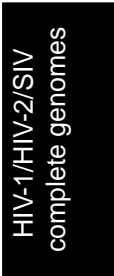
gag p7 nucleocapsid binds these loops, see DeGuzman, Science 279:1384-1388(1988) packaging signal secondary structure, see Harrison J Virol 72:5886-5896(1998)

B.FR.HXB2	...TAA--TT-----A.....	789
A.UG.U455	...AAATTTTA	256
D.CD.84ZB085	-T.GAA--TT-----	169
F1.BE.VI850	--GAATTTTT	306
G.SE.SEG165	...A---T-----G.....	144
H.CF.90CF056	...ATTTG-----	192
J.SE.SE7887	...AAAA--T-T-----	136
K.CM.MP535C	112
01.AE.TH.CM240	0
02.AG.NG.IBNG	354
03.AB.RU.KAL153-2	329
04.CPX.CV.94CY032-3	840
N.CM.YBF30	155
O.CM.MVP5180	347
O.BE.AN70	816
CPZ.US.CPZUS	840
CPZ.GA.CPZGAB	831
CPZ.CM.CM3	826
CPZ.CM.CM5	179
CPZ.CD.CPZANT	485
H2A.GW.ALI	153
H2A.SN.HIV2ST	1095
H2A.DE.BEN	547
H2B.GH.D205	1102
H2B.CI.BHO	1078
H2G.CI.ABT96	1079
MAC.US.WM251	465
SMI.US.SIVSMH9	1040
STM.US.STM	534
TANTALUS.UG.TANI	708
GRIVET.ET.GRI-677	901
VERVET.KE.VER-9063	726
VERVET.KE.AGMTYO	941
VERVET.KE.AGMTYO	431
VERVET.DE.AGM3	930
SIN.GA.SIVSUN	430
MANDRIL.GA.MNGB1	1016
LHOEST.KE.loeest447	1036
LHOEST.CD.loeest485	449
LHOEST.CD.loeest524	1056
RCM.NG.SIV-NGM	0
COL.SIV-CGU1	0
SYKES.KE.SYK173	1066
	534
	794

HIV-1/HIV-2/SIV complete genomes

B.FR.HXB2 Gag	M_G_A_R_A_S_V...L_S_G_G_E_L_D_R_W_E_K_I	p17
B.FR.HXB2	ATGGGTGGAGCGTCAGTA...TTAAGGGGGGAGATTAGATCGATGGGAAAAAT	845
A.UG.U455AA-A-.....TC-.....G	312
C.ET.ETH2220A-.....A-C-A-A-.....GCC-.....	225
D.CD.84ZB085A-.....A-.....G-.....GCT-.....	362
F1.BE.VI850A-.....A-.....GA-.....	200
G.SE.SEG165C-.....A-.....GCT-.....G	248
H.CF.90CF056T-.....C-A-.....GAT-.....G	192
J.SE.SE7887T-.....A-.....GC-.....G	168
K.CM.MP535CT-.....A-.....GC-.....G	56
O1.AE.TH.CM240T-.....A-.....GC-.....G	410
02.AG.NG.IBNGT-.....A-.....GC-.....G	885
03.AB.RU.KAL153-2T-.....A-.....GC-.....G	896
04.CPX.CV.94CY032-3T-.....A-.....GC-.....G	211
N.CM.YBP30G-.....CA-.....A-.....TC-.....	403
O.CM.MVP5180T-G-.....G-CA-.....A-.....GC-.....CG-.....	872
CPZ.US.CPZUST-G-.....G-CA-.....A-.....GC-.....CG-.....	896
CPZ.GA.CPZGABT-.....C-G-CA-.....CG-.....GCT-.....G	887
CPZ.CM.CAM3T-.....C-CA-.....A-.....GCT-.....G	882
CPZ.CM.CAM5CA-.....CG-.....GCT-.....G	235
CPZ.CD.CPZANTA-G-G-.....T-.....G-G-A-AGA-GC-.....AC-.....GT-.....	541
A-G-G-.....T-.....G-G-A-AGA-GC-.....AC-.....GT-.....	209
H2A.GW.ALI Gag	M_G_A_R_N_S_V...L_R_G_R_K_A_D_E_L_E_R_I	p17
H2A.GW.ALI	AAAC-C-C-...-G-A-A-A-A-GC-CGA-TA--G--	1151
H2A.SN.HIV2STAAC-C-C-...-G-A-AA-A-GC-CGA-TA--G--	603
H2A.DE.BENAAC-C-C-...-G-A-AA-A-GC-CGA-TA--G--	1158
H2B.GH.D205C-.....G-CA-.....A-.....GC-.....CGA-TA--G--	1134
H2B.CI.BHOC-.....G-CA-.....A-.....GC-.....CGA-TA--G--	1135
H2G.CI.BBT96C-.....G-CA-.....A-.....GC-.....CGA-TA--G--	521
MAC.US.MM251C-.....G-CA-.....A-.....GC-.....CGA-TA--G--	1096
SMM.US.SIVSMM49T-.....AAC-C-C-...-GTCA-.....AAGA-GC-.....GA-TA--G--	590
STM.US.STMC-.....AGC-C-C-...-GTCA-.....AA-A-GC-.....GA-TA--G--	764
TANTALUS.UG.TANIA-G-CA-C-G-...-GTCA-.....AA-A-GC-.....GA-TA--G--	957
GRIVET.ET.GRI-677C-G-G-TCAC-C-...-GTCA-.....A-AGCC-C-CACG-TG-G-G	782
VERVET.KE.VBR-9063G-GCTA-C-...-C-.....AA-A-C-.....C-A-TC-G	997
VERVET.KE.AGMTYOG-GCTA-C-...-C-.....ATA-A-C-.....C-A-TC-G	487
VERVET.DE.AGM3G-GCTA-C-...-C-.....ATA-A-C-.....C-A-TC-G	986
SABAEUS.SN.SAB1CG-GCTA-C-...-C-.....ATA-A-C-.....C-A-TC-G	486
SUN.GA.SIVSUNA-G-CA-C-...-G-GATA-.....A-TTG-GAGGA-C-TT--G-G-G	1072
MANDRILL.GA.MNDGB1AATG-GAAC-T-CC-...-GTTA-.....ACT-.....T-.....G-.....AA-IT-G-.....	1118
LHOEST.KE.roestAT-G-GAAC-T-C-...-C-C-.....TA-CA-ATTGA-A-AGAT-ITGT-GTG-.....	505
LHOEST.CD.roest447AT-G-GAAC-T-C-...-C-C-.....TA-CA-ATTGA-A-AGAT-ITGT-GTG-.....	1131
LHOEST.CD.roest485AT-G-GAAC-T-C-...-C-C-.....TA-CA-ATTGA-A-AGAT-ITGT-GTG-.....	56
LHOEST.CD.roest485AT-G-GAAC-T-C-...-C-C-.....TA-CA-ATTGA-A-AGAT-ITGT-GTG-.....	56
LHOEST.CD.roest524AT-G-GAAC-T-C-...-C-C-.....TA-CA-ATTGA-A-AGAT-ITGT-GTG-.....	56
RCM.NG.SIV-NGMG-.....A-T-.....GTCA-.....AAGA-G-.....GCT-.....C-GG-.....	1122
COL.SIV-CGU1CAACGAGCA-GG-C-G-...-G-GAA-AA-AC-.....AGACCTAC-G-G-.....	659
SYKES.KE.SYK173A-AGCG-G-.....G-C-ATA-.....CA-TA-.....GC-.....CA-AT-G-.....	853
SYKES.KE.SYK173 Gag	M_G_A_A_G_S_A_I_L_T_G_R_E_L_D_R_Y_E_K_I	p17

B.FR.HXB2 Gag	R_L_R_P_P_G_G_K_K_K_Y_K_L_K_H_I_V_W_A_S_R_E_L_E_R_F_A_V_N_P_G_L_L_E_L_E_T_S_E_G_C_R_O_I_L_G	p17
B.FR.HXB2	TCGGTTAAGCCAGGGGAAATAAATAAATAAATATAAGTATGGCAAGGAGCTAGACGATTCCAGTTATCCCTGGCCCTGTAGAAACATCAGAAGGCTGTAGACAATACTGGGA	975
<hr/>		
A.UG.U455	---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	442
C.ET.ETH2220	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	355
D.CD.84ZRO85	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	492
F1.BF.VI850	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	330
G.SE.SE6165	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	378
H.CF.J90CF056	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	322
J.SE.SE7887	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	298
K.CM.MP535C	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	186
O1.AE.TH.CM240	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	540
02.AG.NG.IBNG	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	515
03.AB.RU.KAL153-2	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1026
04.CPX.CY.94CY032-3	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	341
N.CM.YBF30	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	533
O.BE.ANT70	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1002
CPZ.US.CPZUS	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1026
CPZ.GA.CPZGAB	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1017
CPZ.CH.CM3	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1012
CPZ.CH.CM5	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	365
CPZ.CH.CM3	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	671
CPZ.CH.CPZANT	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	339
<hr/>		
H2A.GW.ALI Gag	R_L_R_P_P_G_G_K_K_K_Y_K_L_K_H_I_V_W_A_S_R_E_L_E_R_F_A_V_N_P_G_L_L_E_L_E_T_S_E_G_C_R_O_I_L_G	p17
H2A.GW.ALI	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1281
<hr/>		
H2A.SN.HIV2ST	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	733
H2A.DE.BEN	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1288
H2B.GH.D205	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1264
H2B.CI.BHO	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1265
H2G.CI.BBT96	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	651
MAC.US.MW251	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1226
SMW.US.SIVSMW9	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	720
STM.US.STM	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	894
TANTALUS.UG.TANI	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1087
GRIVET.ET.GRI-677	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	912
VERVET.KE.VBR-9063	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1127
VERVET.KE.AGMTYO	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	617
VERVET.KE.AGM155	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1116
VERVET.DE.AGM3	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	616
SABAEUS.SN.SABIC	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1202
SUN.GA.SIVSUN	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1248
MANDILL.GA.MNDGB1	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	635
LHOEST.KE.loeest	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1261
LHOEST.CD.loeest447	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	186
LHOEST.CD.loeest485	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	186
LHOEST.CD.loeest524	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	186
RCM.NG.SIV-NGM	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	1252
COL.....SIV-NGU1	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	789
<hr/>		
SYKES.KE.SYK173	---C---C---G---G---TC---A---T---C---C---C---T---G---A---A---C---C---C---T---T---G---G---A---CAG---A---CAG---A---CAG---	983
SYKES.KE.SYK173 Gag	R_L_R_P_P_G_G_K_K_K_Y_K_L_K_H_I_V_W_A_S_R_E_L_E_R_F_A_V_N_P_G_L_L_E_L_E_T_S_E_G_C_R_O_I_L_G	p17



B.FR.HXB2 Gag	E_E_Q.....N_K_S_K_K_K_A_O_O_A.....A_A_D.....T_G.....	p17
B.FR.HXB2	AAGAGCAA.....AACAAAAGTAAAGAAAAAGCAGCAAGCA.....GCAGCTGAC.....ACAGGA.....	1158
A.UG.U455	..AT..T..G..AC..C..GGA..G.....A.....	625
C.ET.ETH2220	..A.....G..C..C..A..G.....GACA.....	538
D.CD.84ZR085	..A.....C.....GG..G.....	678
F1.BF.VI850	..A.....C..C..C..GA.....A..G.....	513
G.SE.SE6165	..GATA.....G.....C..C..A..AT.....ATG..T.....	561
H.CF.90CF056	..ATA.....G.....C..C..A.....A.....	505
J.SE.SE7887	..ATA.....G.....C..C..A.....A.....	481
K.CM.MP535C	..A.....T.....G.....C..C..G.....A.....	360
O1.AE.TH.CM240	..ATA.....T.....G.....C..C..G.....G.....	723
02.AG.NG.IBNG	..ATA.....T.....G.....C..C..G.....G.....	698
03.AB.RU.KAL153-2	..ATA.....T.....G.....C..C..G.....G.....	1209
04.CP.XV.94CY032-3	..ATA.....T.....G.....C..C..G.....G.....	524
N.CM.MVP30	..GA..A..G.....GAACAGAC..GCCGAGCA..CC..G..AGO..GG.....G..CA.....	728
O.CM.MVP5180	..TAATG.....GCA..GC..G.....TCTGCG..AGCCGCTAAG.....A..AAACA.....GC.....	1182
O.BE.ANT70	..G..TA.....GTGC..ENTG.....C..GAA..A..G..A.....A..ACACA.....GC..C.....	1209
CPZ.US.CPZUSCGCA..C..TGG..G..CA.....AGCAA..ACTGAAAGTAACT..GAG..C..TGTGC.....GAG..C..AGT.....	1209
CPZ.GA.CPZGABCAGTT.....CGCCCTCCGAG..C..G..A..GT..GCCA.....C..GCAGC..G.....CAGCAGCA.....TCT.....	1225
CPZ.CM.CM3TGA..AGT.....TGC..G..CAC..AGC..G..A..GGAAGTAGCCAAACCCGCAAGCAG..C..G..G..A.....GAA.....	551
CPZ.CD.CPZANTTGA..AGT.....TGC..G..CAC..AGC..G..A..GGAAGTAGCCAAACCCGCAAGCAG..C..G..G..A.....TGGTGTACGCA.....	570
H2A.GW.ALI GagO..I..I..O..R..H..L..A..A..E.....I..E..T.....	p17
H2A.GW.ALIC..A..T..TAC..R..C..TCT..GC..GC..A.....AT..AAACA.....G..AGAAAATGCCCAAGCACAAGTAGCACAAGCA.....	1473
H2A.SN.HIV2STC..A..G..GTAC..G..C..TCT..GT..GC..A.....A..AAACT.....	925
H2B.DE.BENACT..GCAC..G..C..TCT..GT..GC..A.....A..T..GAACT.....	1480
H2B.GH.D205G..T..GCAC..G..C..TCT..GC..GC..AC.....A..AAA..A.....	1447
H2B.CI.BHOC..AGT..GTGC..G..C..TCT..GT..GT..A.....A..T..GAACT.....	1448
H2G.CI.ABT96C..G..T..GTGC..G..C..CCT..GT..GT..A.....A..GAACT.....	843
MAC.US.MM251C..A..T..GTGC..G..C..TCT..GT..GT..A.....A..GAACT.....	1418
SMM.US.SIVSMH9C..A..T..GTGC..G..C..TCT..GT..GT..A.....A..GAACT.....	912
STM.US.STMCGTGCATCTAGTGGAA..AG..AA..CTGC..G..C..TCT..GT..GT..A.....G..CAGCAGC..G.....	1086
TANTALUS.UG.TANIC..CT..CCATCTAGTGGAC..A..TGAG..AGC..GCTAA..A..A..AAT.....G..G.....	1297
GRIVET.ET.GRI-677TGTCGCATCTAGTGGAA..GG..AA..GAGC..C..AG..CACTAGT.....A..A..G..A.....	1104
VERVET.KE.VBR-9063C..CTGCATCTAGTGGAA..AG..AA..GTGC..C..AGACATCTAGT.....G..CAAA..G.....	1340
VERVET.KE.AGMTYOTGTCGCATCTAGTGGAC..AG..AA..CTGC..G..CA..CT..GT.....G..CAGC..G.....	830
VERVET.KE.AGM155G..AGTGCACGAGAAATG..CAG.....GCC..C..GGGA..TCTAGT..GC.....	1329
VERVET.DE.AGM3G..AGTGCACGAGAAATG..CAG.....GCC..C..GGGA..TCTAGT..GC.....	841
SABAEUS.SN.SABICT.....C..GCAAAATGAG..A.....	1427
SUN.GA.SIVSUNTCATAAC..AGNAAGGAGAR..GC..GGAGG..TG.....	1401
MANDRILL.GA.MNDGB1T.....A..C..G..C..C..C..G..G..G..G..G.....	800
LHOEST.KE.ROESTT.....A..C..ACTCAGGAAGAA..GGC..G.....	1417
LHOEST.CD.ROEST447T.....A..C..ACTCAGGAAGAC..GGC..G.....	342
LHOEST.CD.ROEST485TT.....GAACAAACCTGTG.....	336
LHOEST.CD.ROEST524GC..TGCCATCTAGTGGAC..AG..TGAG..TGC..GG..A..AG..AA.....	342
RCM.NG.SIV-NGMC..T..TAAAAAGCAGGCC..TG..T..GAA..T..GCC..GCAAGG..G..G..AG.....AA..AA..A.....	1465
COL.....SIV-CGU1C..T..TAAAAAGCAGGCC..TG..T..GAA..T..GCC..GCAAGG..G..G..AG.....AA..AA..A.....	1013
SYKES.KE.SYK1173C..ATGTAAATGGAAAAGATG..CC..CCAGC..C..TCT..GTGGA..AGT.....A..AA..AG.....	1210
SYKES.KE.SYK1173 Gag	E_A_C_N_W_K_D_D_P_P_A_T_S_G_G_Q_S.....E_N_S.....S_Q_N_M_A_S_E_T_S_S_G_Q_K.....	p17

B.FR.HXB2 Gag	H S N Q V S	Gag p17 \ / Gag p24	O N Y P I V Q N I O G Q M V H Q A I S P R T L N A	384
B.FR.HXB2	.CACAGCAAT...CAGGTACAGC...CAAAAATTACCTTATAGTCAGAACATCCAGGGCAAAATGGTACATCAGCCGATATACCTAGAACCTTTAAATGCA	p24
A.UG.U455	...AG--T...	...A--TGCA--A--CCA--C--T--G--C--G--C--G--	...T--T--G--C--C--C--G--G--	706
CT.EH2220	...GGA-AGG-C	...T--T--G--C--C--C--G--G--	...T--T--G--C--C--C--G--G--	625
D.CD.84ZB085	...A-----GC	...T-----A-----TC-T-A-----T-TC-----G-C	...T-----A-----TC-T-A-----T-TC-----G-C	768
F1.BE.VI850	...A-----C-----A	...T-----A-----TC-T-A-----T-TC-----G	...T-----A-----TC-T-A-----T-TC-----G	594
G.SE.SEG165	...A-AGA-C-A-A	...T-AGCT-A-A-G	...T-AGCT-A-A-G	574
H.CF.90CF056	...G-A-G	...T-TC-G-A-CC	...T-TC-G-A-CC	598
J.SE.SE7887	...GA-A	...T-TC-G-A-CC	...T-TC-G-A-CC	574
K.CM.MP535C	...AG--GC--A-A	...A--TGCA--A--C--C--TT--G--G--	...A--TGCA--A--C--C--TT--G--G--	450
01.AE.TH.CM240	...AG--GC--A-A	...A--TGCA--A--C--C--TT--G--G--	...A--TGCA--A--C--C--TT--G--G--	782
02.AG.NG.IBNG	...AG--GC--A-A	...A--TGCA--A--C--C--TT--G--G--	...A--TGCA--A--C--C--TT--G--G--	816
03.AB.RU.KAL153-2	...GGT--GC--A-TA--T...	...AGG--T--C--C--CTGT--A--A--C--C--G--G--	...AGG--T--C--C--CTGT--A--A--C--C--G--G--	1302
04.CPX.CV.94CY032-3	...CCT-GG	...AACA--TGCA--A--A--C--C--G--G--	...AACA--TGCA--A--A--C--C--G--G--	617
N.CM.YBF30	...T--C	...ATT-GT	...ATT-GT	1272
O.CM.MVP5180	...GCTGGC	...A-T-T	...GCTGGC	1296
O.BE.ANT70	...GA--GC	...GCT--T	...GA--GC	1299
CPZ.US.CPZUS	...GC--A-A-T	...AG--C	...GC--A-A-T	1312
CPZ.GA.CPZGAB	...GC--A-A-T	...AG--C	...GC--A-A-T	641
CPZ.CM.CAM3	...CATCT-AGTGGC-AAG-G	...AG--C	...CATCT-AGTGGC-AAG-G	944
CPZ.CM.CAM5	...P P S E Q G	...G N F P V O Q V A	...G N F P V O Q V A	663
CPZ.CD.CPZANT	...CACCT-G	...G-ACAGG-G	...CACCT-G	p24
H2A.GW.ALI Gag	...CACCT-G	...G-ACAGG-G	...CACCT-G	1563
H2A.GW.ALI	...CACCT-G	...G-ACAGG-G	...CACCT-G	1563
H2A.SM.HIV2ST	...CACCT-G	...GG-AA-G-A	...CACCT-G	1015
H2A.DE.BEN	...CACCT-G	...GG-AA-G-A	...CACCT-G	1570
H2B.GH.D205	...CACCT-G	...GG-AA-G-A	...CACCT-G	1531
H2B.CI.BHO	...CCA-C	...AGC	...CCA-C	1532
H2G.CI.BBT96	...CCACC	...AGT...G-AGAGGAGG-C	...CCACC	933
MAC.US.MM251	...CAICT-GC	...GGCAGAG-A	...CAICT-GC	1508
SMM.US.SIVSMMH9	...CACCT-G	...GGCAGAG-A	...CACCT-G	1002
STM.US.STM	...ACACT-G	...GGCAGAG-A	...ACACT-G	1176
TANTALUS.UG.TANI	...CACCT-G	...GGCAGAG-A	...CACCT-G	1387
GRIVET.ET.GRI-677	...CACCTGG	...GGC-AATCA	...CACCTGG	1194
VERVET.KE.VBR-9063	...CACCTGG	...GGC	...CACCTGG	1427
VERVET.KE.AGMTYO	...ACACCTGG	...GGC	...ACACCTGG	917
VERVET.DE.AGM3	...ACACCT-G	...GGT-G-TG	...ACACCT-G	928
SABAEUS.SN.SAB1C	...ACACCT-G	...GGT-G-TG	...ACACCT-G	1517
SUN.GA.SIVSUN	...GCA	...A-CTTCT	...GCA	1485
MANDRILL.GA.MNDGB1	...A-G-AAC-A	...ACAG-G	...A-G-AAC-A	878
LHOEST.KE.hoest	...GCA	...A-CTTCT	...GCA	1507
LHOEST.CD.hoest447	...GCA	...A-CTTCT	...GCA	423
LHOEST.CD.hoest485	...GCA	...A-CTTCT	...GCA	423
LHOEST.CD.hoest524	...GCA	...A-CTTCT	...GCA	417
RCM.NG.SIV-NGM	...AA1TACCAATAGTAACA-TT	...G-C-C-G-A-TCT-CT-GA-CATTAAATGCTTG	...AA1TACCAATAGTAACA-TT	1570
COL.---SIV-CGU1	...G	...CTC-AGGG-AG-GCAT-ACC-C-GAGCCCCAGG-C-T-GGAGCCTGGG-TAAGTG-GT-GAGGG-GGAAT	...G	1089
SYKES.KE.SYK173	GGTAGTAGCAGCAG...	GAAAAACAAGGAGCA-CA...-C-CCACCTAGA...GG--C-T-AC-GC-CAGA--TCCT--AAAT--GTG-A-T--ACA-GAG-CC--GT--G--C--AA--	GGTAGTAGCAGCAG...	1331
SYKES.KE.SYK173 Gag	V V Q Q E K Q K A A T P P P R G N Y P L L R N P Q N O W I H T G V P V R T L K T	GGTAGTAGCAGCAG...GAAAAACAAGGAGCA-CA...-C-CCACCTAGA...GG--C-T-AC-GC-CAGA--TCCT--AAAT--GTG-A-T--ACA-GAG-CC--GT--G--C--AA--	V V Q Q E K Q K A A T P P P R G N Y P L L R N P Q N O W I H T G V P V R T L K T	p24

B.FR.HXB2 Gag	W V K V V E E K A F S P E V I P M F S A L S E G A T P Q D L N T M L N T V G G H O A A	1381
B.FR.HXB2	TGGTAAAGTAGTAAAGGTTTCAGCCCAAGTAGTATCCATGTTTTCAGCAITTAACAGAAAGGAGCCACCCACAGAGATTAAACACCATGCTAAACACAGTGGGGGACATCAAGCAGCCA	
A.UG.U455	-----C-----A-----A-----T-----G-----T-----TG-----G-----TGT-----C-----G-----T-----	836
C.ET.ETH2220	-----G-----A-----A-----A-----T-----G-----A-----T-----	755
D.CD.84ZRO85	-----A-----A-----A-----A-----	898
F1.BF.VI850	-----G-----GA-----T-----T-----C-----CAC-----G-----T-----	724
G.SE.SE6165	-----A-----C-----T-----T-----A-----	784
H.CF.90CF056	-----A-----T-----T-----A-----	728
J.SE.SE7887	-----GA-----A-----A-----A-----T-----G-----	704
K.CM.WP535C	-----G-----G-----A-----T-----A-----C-----	580
O1.AE.TH.CM240	-----G-----A-----G-----T-----A-----C-----	946
02.AG.NG.IBNG	-----G-----A-----G-----A-----T-----TG-----G-----	912
03.AB.RU.KAL153-2	-----G-----A-----A-----A-----G-----G-----	1432
04.CPX.CY.94CY032-3	-----G-----A-----A-----A-----C-----G-----G-----	747
N.CM.WP5180	-----G-----GA-----G-----C-----T-----T-----A-----A-----A-----G-----CTC-----C-----T-----T-----A-----A-----A-----G-----A-----	945
O.BE.ANT70	-----G-----C-----C-----T-----A-----T-----T-----A-----ATG-----C-----T-----A-----C-----T-----A-----C-----T-----A-----C-----T-----A-----A-----A-----G-----CA-----A-----C-----GG-----TT	1402
CPZ.US.CPZUS	-----G-----C-----A-----C-----T-----A-----T-----A-----C-----T-----T-----CATG-----G-----G-----A-----TTT-----CT-----A-----T-----T-----TG-----CA-----A-----A-----G-----TT	1426
CPZ.GA.CPZGAB	-----CG-----A-----A-----C-----T-----A-----G-----C-----C-----G-----CATG-----C-----G-----G-----T-----C-----G-----T-----A-----G-----G-----A-----	1429
CPZ.CM.CM3	-----G-----G-----A-----A-----T-----G-----A-----T-----TC-----T-----G-----G-----T-----T-----TG-----A-----	1442
CPZ.CM.CM5	-----T-----G-----A-----A-----AA-----T-----A-----A-----CATG-----CC-----G-----A-----I-----	771
CPZ.CD.CPZANT	-----C-----G-----A-----A-----AAC-----T-----A-----A-----ATG-----CC-----G-----G-----T-----T-----C-----G-----G-----C-----T-----T-----TG-----C-----A-----A-----G-----C-----G-----G-----T	1074
	GTGT-----AAA-----AT-----C-----C-----T-----T-----T-----G-----A-----T-----T-----T-----G-----A-----T-----T-----TG-----T-----AC-----G-----	793
H2A.GW.ALI Gag	W V K L V E E K F G A E V V P G F O A L S E G C T P Y D I N O M L N C V G D H O A A	p24
H2A.GW.ALI	-----T-----A-----AAG-----G-----GG-----AG-----G-----AGGA-----C-----C-----C-----CTG-----G-----CT-----T-----A-----T-----TCAA-----T-----TTGT-----C-----AC-----	1693
H2A.SN.HIV25T	-----C-----G-----A-----AAG-----G-----GG-----AG-----G-----AGGA-----C-----C-----C-----CTG-----G-----CT-----T-----A-----T-----TCAA-----T-----TTGT-----C-----AC-----	1145
H2A.DE.BEN	-----T-----G-----A-----AAG-----G-----GG-----AG-----G-----AGGA-----C-----C-----C-----CTG-----G-----CT-----T-----A-----T-----TCAA-----T-----TTGT-----C-----AT-----C-----	1700
H2B.GH.D205	-----GT-----T-----A-----AAG-----G-----GG-----AG-----G-----AGGA-----C-----C-----C-----TG-----TT-----T-----A-----TCAG-----T-----TTGT-----A-----A-----A-----G-----	1661
H2B.CI.BHO	-----T-----T-----A-----AAG-----G-----GG-----AG-----G-----AGGA-----C-----C-----C-----GTG-----CTA-----T-----T-----A-----C-----TCAA-----T-----TTGT-----A-----A-----A-----G-----	1662
H2G.CI.BHT96	-----T-----G-----A-----AAG-----G-----GG-----AG-----G-----AGGA-----C-----G-----G-----C-----G-----C-----CT-----T-----CA-----T-----TCAG-----T-----TTGT-----A-----A-----A-----G-----T-----	1063
MAC.US.MM251	-----T-----GA-----G-----A-----AAA-----TG-----AG-----AG-----G-----AGGA-----C-----G-----G-----C-----G-----C-----CT-----T-----CA-----T-----TCAG-----T-----TTGT-----A-----A-----A-----G-----	1638
SWM.US.SIVSMMH9	-----T-----G-----G-----A-----AAA-----TG-----GG-----G-----AG-----G-----AGGA-----C-----G-----C-----CA-----G-----C-----CT-----Y-----C-----T-----Y-----C-----TCAA-----T-----TTGT-----A-----A-----A-----G-----	1132
STM.US.STM	-----GT-----G-----G-----A-----AAA-----TG-----TG-----G-----AG-----G-----AGGA-----C-----C-----C-----CA-----C-----CT-----T-----CA-----T-----TCAG-----T-----TTGT-----A-----A-----A-----G-----	1306
TANTALUS.UG.TANI	-----T-----AC-----T-----G-----A-----AAGG-----G-----GG-----A-----AG-----CCA-----C-----C-----G-----GTG-----TTAAGTT-----T-----CA-----T-----TCAG-----T-----TTGT-----A-----A-----A-----G-----G-----	1517
GRIVET.ET.GRI-677	-----C-----TGC-----G-----G-----A-----AAA-----TG-----AG-----GA-----AG-----G-----A-----CCA-----C-----C-----G-----GTCT-----T-----CT-----T-----G-----TCAG-----C-----TGT-----A-----A-----AC-----GG-----AT	1324
VERVET.KE.VER-9063	-----C-----G-----G-----A-----AAA-----TG-----AG-----GA-----AG-----G-----A-----CCA-----C-----C-----CA-----C-----CT-----T-----CA-----T-----TCAG-----T-----TTGT-----A-----A-----AC-----GG-----AT	1557
VERVET.KE.AGMTYO	-----C-----G-----G-----A-----AAA-----TG-----GG-----A-----AG-----CCA-----C-----C-----CA-----C-----CT-----T-----CA-----T-----TCAG-----T-----TTGT-----A-----A-----AT-----GG-----AT	1047
VERVET.KE.AGM155	-----C-----G-----G-----A-----AAA-----TG-----GG-----A-----AG-----CCA-----C-----C-----CA-----C-----CT-----T-----CA-----T-----TCAG-----T-----TTGT-----A-----A-----AC-----GG-----GC	1546
VERVET.KE.AGM3	-----C-----G-----GA-----T-----G-----A-----AAG-----TG-----G-----AG-----CA-----G-----AG-----CC-----G-----T-----TT-----T-----CA-----C-----TCAA-----T-----TTGT-----A-----A-----AC-----GG-----GC	1647
SUN.GA.SIVSUN	-----G-----GA-----T-----G-----A-----AAG-----CG-----AA-----G-----G-----C-----TC-----A-----C-----GA-----T-----ATGAC-----A-----TT-----T-----CA-----C-----TGCTG-----G-----TG-----CA-----A-----A-----AG-----C-----T-----	1615
MANDILL.GA.MNDGB1	-----G-----AC-----G-----A-----AAG-----CG-----AA-----G-----G-----C-----ACAG-----G-----A-----C-----A-----AA-----TTTG-----TG-----T-----G-----T-----TG-----A-----T-----A-----AT-----G-----C-----	1008
LHOEST.KE.hoest	-----G-----A-----A-----T-----GA-----GA-----GG-----AA-----ACAG-----GG-----A-----T-----CC-----CA-----AA-----TTTG-----G-----T-----C-----TGT-----T-----TG-----CA-----A-----A-----AC-----GG-----C-----	1637
LHOEST.CD.hoest447	-----G-----A-----A-----T-----GA-----GA-----GG-----AA-----ACAG-----GG-----A-----T-----CC-----CA-----AA-----TTTG-----G-----T-----C-----TGT-----T-----TG-----CA-----A-----A-----AC-----GG-----A-----	553
LHOEST.CD.hoest485	-----G-----A-----A-----T-----GA-----GA-----GG-----AA-----ACAG-----GG-----T-----TC-----GA-----AA-----TTG-----G-----C-----TGT-----T-----TG-----TA-----A-----A-----AC-----GG-----A-----	553
LHOEST.CD.hoest524	-----G-----A-----A-----T-----GA-----GG-----AG-----ACAG-----GG-----T-----G-----C-----G-----CTG-----TA-----CT-----T-----G-----TCAG-----T-----TG-----CA-----A-----A-----AC-----GT-----T	547
RCM.NG.SIV-NGM	GCT.....TTTCCT-----GCT-----C-----T-----AT-----TACG-----G-----A-----TAG-----TT-----T-----A-----G-----TTG-----C-----T-----TTT-----ATACT-----G-----TT-----C	1664
COL.....SIV-CGU1	-----GG-----G-----CT-----GA-----CTCC-----AAAA-----TGATG-----TCCA-----AG-----G-----AT-----A-----CAG-----TC-----A-----TTT-----TT-----TT-----T-----C-----C-----L-----N-----G-----M-----L-----N-----A-----V-----G-----D-----H-----O-----G-----	1461
SYKES.KE.SYK173	W V E A V N S K K F D A S I V P L F Q A L T E G F I P Y D L N G M L N A V G D H O A A	p24
SYKES.KE.SYK173 Gag	-----GG-----G-----CT-----GA-----CTCC-----AAAA-----TGATG-----TCCA-----AG-----G-----AT-----A-----CAG-----TC-----A-----TTT-----TT-----TT-----T-----C-----C-----L-----N-----G-----M-----L-----N-----A-----V-----G-----D-----H-----O-----G-----	

HIV-1/HIV-2/SIV complete genomes

B.FR.HXB2 Gag
B.FR.HXB2
A.UG.U455
C.ET.ETH2220
D.CD.842R085
F1.BE.VI850
G.SE.SE6165
H.CF.90CF056
J.SE.SE7887
K.CM.MP535C
O1.AE.TH.CM240
02.AG.NG.IBNG
03.AB.RU.KAL153-2
04.CPX.CY.94CY032-3
N.CM.MVP3180
O.BE.ANT70
CPZ.US.CPZUS
CPZ.GA.CPZGAB
CPZ.CM.CM3
CPZ.CM.CM5
CPZ.CD.CPZANT
H2A.GW.ALI Gag
H2A.GW.ALI
H2A.SN.HIV2ST
H2A.DE.BEN
H2B.GH.D205
H2B.CI.BHO
H2G.CI.ABT96
MAC.US.MM251I
SMW.US.SIVSMWH9
STM.US.STM
TANTALUS.UG.TANI
GRIVET.ET.GRI-677
VERVET.KE.VBR-9063
VERVET.KE.AGMTYO
VERVET.KE.AGM155
VERVET.DE.AGM3
SABAEUS.SN.SAB1C
SUN.GA.SIVSUN
MANDILL.GA.MNDGB1
LHOEST.KE.roest
LHOEST.CD.roest447
LHOEST.CD.roest485
LHOEST.CD.roest524
RCM.NG.SIV-NGM
COL.SIV-CGU1
SYKES.KE.SYK1173
SYKES.KE.SYK173 Gag

T L O F E O I G W M T N N P P I P V G E I Y K R W I I L G L N K I V R M Y S P T S
TACCTTTCAGGAAACAAATAGGATGCAAAAT.AATCCACCTATCCAGTAGGAGAAATTTATAAAGATGGATAATCCTGGGAATTAATAAATAGTAAGAAATGTATAGCCCTACCCAGC
C - - - - - A GGC G - C - C - G - - - - - G T T - - - - - G T T
C - - - - - A GGC G - C - C - G - - - - - G T - - - - - G T
C - - - - - A GC G - C - C - G - - - - - C - - - - - C T
T - - - - - G CGCC G - C - C - G - - - - - G - - - - - A
C - - - - - C GGC G - C - C - G - - - - - G - - - - - G
C - - - - - C GGC G - C - C - G - - - - - G - - - - - T
A - - - - - A C AGC - C - - - - - G - G - C - C - - - - - C
A - - - - - A G G - C - C - C - - - - - G - - - - - G
A - - - - - A GC C C - C - C - C - - - - - G
C - - - - - GGCA . GG - G - C T - - - - - GC T - G - T - - - - - G - T - - - - - C
C - A - G - A - G - T A T - - - - - T - G - G G G G C - A A C T - - - - - G - G - A - C - C - G - T - G - G - T - G - G - - - - - C
C - T - G - A - G - T C A C - - - - - C - A - - - - - C - C - G - A - - - - - G - G - T - A - C - C - G - A - C - T - A G T G - - - - -
C - A - G - - - - - G - G - C - - - - - G C A - C - C - T - - - - - G C A - - - - - G - G - A - G C - T - - - - - C T - C G T G - T
C - A - G - - - - - T - C T - - - - - G C A T - C - - - - - G - T G - - - - - G - G - T T - A - C - C - G - G - T - - - - - A G T A - T
C - T - A - A - - - - - G - T - C T - - - - - G C A - - - - - G - - - - - G - G - T - A - G C - T - C - G - - - - - G - - - - - A G T T - T
C - A G - G - - - - - G - G - G C A - - - - - C A C T C A C A C A G - G G - G G A G - - - - - G - C - C - G - - - - - G - G - G - C - C - G - - - - - G - - - - - C T
T V E F O I Q W M F P P V G N I Y R R W I O I G L Q K C V R M Y N P T N
C - A G - A G - A - - - - - G - C C A G - - - - - T T C - G G C C A C G G A A T - - - - - G - G - - - - - G A - C - C - C - G - - - - - C C A G A - A - G C - C - G - G T G T - C - G - - - - - C A - A - A - - - - - A
C - A G - A G - A - G - G - C C A G - - - - - T T T - G G C C A C - A A A T - - - - - G - A - - - - - A G C - C - G - - - - - A G C - C - G - G T G C - - - - - C - A - A - A - A
C - A G - A G - T - G - C C A G - - - - - T A T - G G C C A C - A A A T - - - - - C G - A - - - - - G - A - C - C - C - G - - - - - G A - C - C - G - G T G - - - - - C - A - C - A - A - T - A
C - A G - A G - G - C A G - - - - - T A C - G G G C C C - A A A T - - - - - G - A - C - G - - - - - G - A - C - G - - - - - T C A A T - A - G C - G - - - - - T G T - C C - A T - A
C - A G - A G - A - - - - - G - C A G - - - - - T A C - G A C C C C - A A A T - - - - - G - A - C - C - G - G - - - - - T C A T - A - G C - C - G - - - - - T G T - C C - G - - - - - T - A T
- R A - A G - A - - - - - T C A G - - - - - C A G C - G G C A G C - A A A C - - - - - A - T - Y - G - G - - - - - C C A R T - A - C - T C - A - - - - - T G T - C - G - - - - - Y - A T - A - A - A
T - A G - A G - I - - - - - C C A G - - - - - T A Y - G G C A A C - G A A C - - - - - C - A - - - - - C A - C - - - - - C G G - - - - - C C A A - - - - - A - - - - - A - A - A
C T - A - C A G - - - - - T C A - - - - - T A C - G G C A G C - A A A T - - - - - A - G - - - - - A - C - C - G - G - - - - - T C A A - A - G C - C - A - - - - - T G T - C - - - - - A - A - A
C T - T G - G C T - G - - - - - A G - - - - - C T T T T - G C T A G A G - A G A T - - - - - C G T - C - C - G - G - - - - - G - T - - - - - C - C - A - - - - - T G T - A - - - - - T - T C A
C T - A A - - - - - A G - - - - - A G - - - - - C C T T C - G C C A G A - A G A C - - - - - G - C - C A A - C - G - A - - - - - G - T - T T - - - - - C - C - A - - - - - G G - - - - - C A G - - - - - A
C T - A G - A - A - G T - A - - - - - T T A T - C A G C C C - C A G A G - A G A T - - - - - T - G - C T - C - G - C A G - G - C - G - - - - - T - A - - - - - G C - G C - G - - - - - A
C T - A G - A - G - G C - A G - - - - - C T A T - C - G C T C - G G G - A G A T - - - - - T - C C - C C G G - - - - - T - T - A - C - T C - A - G T G T - C - A - - - - - A - - - - - A G T A T C A
C - A G - G - A - G - G - A - - - - - A T A C - C A G C C C - A G G A G - A G A T - - - - - G - T - C C - C - G - G - - - - - G - C - C - G - - - - - T - A - - - - - C - A T - A G T G C T
C - A A - A - A - G - A - - - - - C C - C - G A C C A C - G A A T G - G - A A T - - - - - A - T - - - - - G - C - C - G - - - - - T C - C - C - G - - - - - T G T - C - A - - - - - A - A - A
A - - - - - A G - G - A - G - G - C A G - - - - - C T A T G C C T T A T G G A C A A - A G A G - T G - - - - - C A - - - - - C - C - G - C - A S A - C A - G G - A - G - G - - - - - T - - - - - C C - C A G A - - - - - T C A
A - - - - - A G - G - A - - - - - T - G - C - - - - - A T - T G C A A C - A A T G - A - - - - - A A T - - - - - A C - - - - - C - - - - - G - G - - - - - G - T - - - - - A - G - - - - - C A - G - - - - - T - A - T A
A - - - - - A G - A - - - - - C - C C - - - - - C T - T G C A G A C A - T - - - - - G - A G A T - - - - - A - G - - - - - G A - G - C - T - A - C T A - G G - A - - - - - G - - - - - C - C A A - A G T T - T
A - - - - - A G - A - - - - - G - C C - - - - - C C - T G C A G A C A - C - - - - - G - A G A T - - - - - A - G - - - - - G A - G - T - A A - C T A - G G - A - - - - - G - T - - - - - T C - C A A - A G T T - T
A - - - - - A G - A - - - - - G C A - C - - - - - C A - G A C A C A - C - - - - - G A T - - - - - G A - - - - - G A - - - - - G A - - - - - G A - C - C T A - G G - A - - - - - G - G - - - - - T C - C A A - A G T T - T
C - A T A - G C T - - - - - C C - T - G A G C C C A A C - A - T G - - - - - G - A - C - - - - - G G - A T - - - - - T G - G - T - G - C - A - - - - - T G T - G - A - - - - - A C - A G T A - T
C T - A G - A G C A - - - - - G - G - C T - - - - - G G A G A G C C A T A G C C - C A T T T A - G G G G C T G G A T - - - - - T - C A G A G - C T G G A - A A - G - - - - - C A G - T G C A - A - A T - T C T
C T - A A - G C T - - - - - A - - - - - T - T - G C C A G C A A C - A G - A - A - - - - - A - - - - - G - G - G - C - G - G - - - - - C - T - A - G C - G C - G - - - - - G C A G G - C - C - A - - - - - G T T T C T
S I A E Q I E W I T R Q N N P V Q V G E I Y R R W I I L G L Q K C V R M Y N P V S
C T - A A - G C T - - - - - A - - - - - T - T - G C C A G C A A C - A G - A - A - - - - - A - - - - - G - G - G - C - G - G - - - - - C - T - A - G C - G C - G - - - - - G C A G G - C - C - A - - - - - G T T T C T

Table with columns for sample names (e.g., B.FR.HXB2 Gag, A.UG.U455, H2A.GW.ALI Gag), sequence coordinates (1511-2367), and nucleotide sequences (e.g., ...AGAAG...ATTGTTAAGTGTTC...).

HIV-1/HIV-2/SIV complete genomes

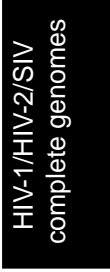
HIV-1/HIV-2/SIV complete genomes

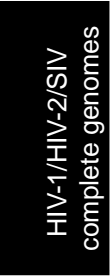
stem-loop after TTTTTT slip potentiates slippage
Gag-Pol 1 ribosomal slip site
Gag p7 end / Gag p1 start
D C T ... E R O A N F F L G K I W P S ... Y K G R P G N F L O S R P ... R A N
ATTGACT... GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC... AGAGCCAA

Accession	Region	Sequence	Position
B.FR.HXB2	Gag	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	p6
B.FR.HXB2	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2155
A.UG.U45E	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1604
C.ET.ETH2220	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1526
D.CD.84ZR085	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1675
F.BE.VI850	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1492
G.SE.SE6165	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1561
H.CF.90CF056	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1505
J.SE.SE7887	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1475
K.CM.MP535C	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1351
01.AE.TH.CM240	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1714
02.AG.NG.IBNG	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1680
03.AB.RU.KAL153-2	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2200
04.CPX.CY.94CY032-3	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1524
N.CM.YBP30	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1728
O.CM.MVF5180	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2197
O.BE.AN1770	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2221
CPZ.US.CPZUS	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2222
CPZ.GA.CPZGAB	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1560
CPZ.CM.CAM3	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1863
CPZ.CM.CAM5	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1863
CPZ.CD.CPZANT	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1603
H2A.GW.ALI	Gag	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	p6
H2A.GW.ALI	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2458
H2A.SN.HIV2ST	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1910
H2A.DE.BEN	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2465
H2B.GH.D205	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2438
H2B.CI.EHO	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2436
H2G.CI.ABT96	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1843
MAC.US.MM251	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2409
SNM.US.SIVSMH9	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1906
STM.US.STM	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2080
TANTALUS.UG.TANI	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2297
GRI.VET.ET.GRI-677	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2092
VERVET.KE.VER-9063	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2334
VERVET.KE.AGM150	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1815
VERVET.KE.AGM155	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2317
SABAEUS.SN.SABIC	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1829
SUN.GA.SIVSUN	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2457
MANDRIL.GA.MNDGB1	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2477
LHOEST.KE.hoest	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1792
LHOEST.CD.hoest447	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2459
LHOEST.CD.hoest448	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1375
LHOEST.CD.hoest524	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1375
LHOEST.CD.hoest524	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1369
RCM.NG.SIV-NGM	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2453
COL.---SIV-CGU1	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	1981
SYKES.KE.SYK173	Gag	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2248
SYKES.KE.SYK173	Gag	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	p6
SYKES.KE.SYK173	Pol	...GAGACAGGCTAAITTTT TAGGAAGAT... CTAGAGGAGGAGGAGGAGGAAITTTTCTCAGACGAGACC...	2458

B.FR.HXB2 GagT_P_P.....O_K_Q_E_P_I_D_K_E_L_Y_P_L_T_S_L_R_S_L_F_G_N_D_P_S_S_Q_-\$	Gag p6 end <-	p6
B.FR.HXB2 Pol1N_S_P.....S_E_A_G_A_D_R_O_G_T_V_S_F_N_F_O_P_O_V_T_L_W_O_R_P_L_V_T_I_K_I_G_Q	Pol TF / Pol p10 protease	p10
B.FR.HXB2AACTCCCC.....TCAGAAGCAGGAGCGATAGACAAGAACTGTATCTTTAATCTTCCCTAGGTCACCTTTGGCAACGAGCCCTCTGCAATAAAGATAGGGGGCAAC		2307
A.UG.U455G-CT.....-GC--A.....-T--A.....-G--A-AC.....-GT.....-T--G--A.....-A--A--G-		1753
C.ET.FTH2220A-CTTCT-C-A--A-GAGCTG-A--CAG-G-AG-C.....-AA.....-A-T-T.....-A.....-A--G-		1699
F.D.84ZR085CCCT-C-AGAA.....A-GAC-A.....AA.....-G-T.....-A.....-A--G-		1829
F.BE.VI850CC-CT-T.....GA-C-G-C-AAAGAC-GGG-ACTGTACC-C-G.....-AA.....-T-A.....-A--GA		1644
G.SE.SE6165C-CCTCC-C-A--A-GAGATG--AA-G-G-TA-ATC-C.....-AA.....-G-A.....-A--A--G-		1704
H.CF.90CF056C-C--T.....-C--A.....-A-C-GA-GG-CA-GGAACC-C-G.....-AA.....-T-T-G-A-A--A--GT		1654
J.SE.SE7887C-C-T.....C-C-A.....-A.....-G.....-G.....-G.....-GA.....-G--A--G-		1621
K.CM.MP535CCC-CT-T.....GAG-C-G-ACCAAGACA-GG-AC-GAGGCC.....-AA.....-A-T.....-AG--A--GT		1503
01.AE.TH.CM240ACCTC-CT-T.....GA-C-G-C-AAAGACA-GG-AC-T-CTCC.....-GT.....-T--G--A--A--A--G-		1881
02.AG.NG.IBNGGAAAGATAACCT-CTTA.....GA-C-G-C-AAAGACA-GG-AC-T-CTCC.....-GT.....-T--G--A--A--A--G-		1832
03.AB.RU.KAL153-2ACCTC-CT-T.....GA-AC-G-AC-AA-GACGGG-AC-G-ATCC.....-C-T.....-T--G--GA--A--A--G-		2352
04.CPX.CY.94CY032-3CT-T.....-T--A.....-GG.....-AA.....-G.....-T.....-AC.....-A--A--G-		1673
N.CM.YBP30ACAGGGAAGAG-TG-AGGA.....GA-CC-G-AG-A-AGGA--TCTCTGTACCCA.....-A.....-G-A.....-AAAG-G		1901
O.CM.MVP5180GGCAGTNAAGA-CAAGAGAA.....GT--A-GG-GATC-GG-A-G-C-A-TG-C.....-AA-C.....-GC--G-T.....-T-T-G		2327
O.BE.AN1770GGAGTGHAGGG-CAGAGAA.....G-A--A--A-AGG-GGCCCG--C-G.....-G-TG-C.....-AA-C.....-GC--GAG-T--C-C-G-		2361
CPZ.US.CPZUSGGAGAACAGGAAA.....GG-A--G.....-CCCT.....-AGC.....-AA.....-C--G.....-AG--CA--GA		2368
CPZ.GA.CPZGABGAAAGCAGGAGGAGGAGAGC-G.....TCT-TA.....ACC-C.....-A--C--G.....-AG--AA--G		1718
CPZ.CM.CAM3GAGAGGAGGAGGAGGAGAA.....AA-G-A-C.....CCTGTATCC-G.....-C--GG-G-A--AA--G		2009
CPZ.CM.CAM5GGAGAAAAG-AGGAGGA.....GA-G-A-C.....CCTGTATCC-AA.....-C.....C-TGG-G-AG--AA--GT		1746
CPZ.CD.CPZANTGAGGACTCAGAAAGG--T-AA.....GGG-G-G.....ACC-C-CGTA.....-AA-C.....AA-GA-GGA-G-TCTC--CA--A--A		1746
H2A.GW.ALI GagE_D_L_L_H_L_E.....O_G_E_T_P_H_K_E_V_T_E_D_L_L_H_L_N_S_L_F_G_K_D_Q_-\$		p6
H2A.GW.ALI PolG_G_L_T_A_P_R.....A_G_R_D_T_T_O_R_G_D_R_G_L_A_P_O_F_S_L_W_K_R_P_V_T_A_Y_I_E_G_Q		p10
H2A.GW.ALIGGAGACTTACTGCAC-T-GA.....G--GGAG-ACA-ACACA-AG--TGACAG-GGAC-GCTGCA.....-AT-T.....-A-A.....-AG-A.....-GCCT-C-T-A--C-G--		2685
H2A.SN.HIV25TAGAGACTTCTCAG-T-GA.....GA-AC-AG-ACA-ATCAGAG--ACGACAG-GGAC-GCTGCA.....-AT-T.....-A-A--AG-A.....-GC-C-TG-T-A--C-G--		2137
H2A.DE.BENGAGACTTCTCAG-T-GA.....G--GAG-ACA-TCAGAG--GACAG-GGAC-GCTGCA.....-AT-T.....-A-A--AG-A.....-GCCT-C-C-A--AT--G--		2692
H2B.GH.D205GATGGGACACA-CAGAGAGA.....GAGCCAG-AGA-CTACA-GG-TGACAG-GGA-GCTGCA.....-AT-T.....-AG-A.....-A-GC-TGT-C-A--T--GT		2668
H2B.CI.EHOACTAGGGAAGAAGCAGAAGGA.....GA-C-GAG-AGA-CTACA-GG-TGACAG-GGA-GCTGCA.....-AT-T.....-AG-A.....-A-GC-CT-T-A--T--T		2019
H2G.CI.ABT96GTTGGCAACGACGACAGAGA.....-GAA-CA-A-CTACA-GG-TGACAG-RGAH-GCTGCA.....-C-T.....-AGGA.....-TGC-T-C-AAA-T-G--		2585
MAC.US.NM251GTTGGCAACGACGACAGAGA.....AAGC-GAG-A--TTACA-GG-TGACAG-GGA-GCTGCA.....-AT-T.....-AGGA.....-TGCCT-T-T-AA-A-G--		2082
STM.US.STMGTTGGCAACGACGACAGAGA.....GAGC-G-A-ACA-CTACA-GG-TGACAG-GGA-GCTGCA.....-AT-T.....-AG-A.....-TGCCT-C-T-AA-AA-G--		2256
TANTALUS.UG.ITANIGGCCTGAGAGGAGGAGAGA.....AG-TC-ATT-AGAAA-CAGA-GG-A-GAAATGGAGACTA--T-G-T-C.....AGGA--AA-GAAGT-AG-CT-C-CA--C-G--		2497
GRIVET.KE.VER-6063ACAGTGGAGG-CAAGGAA.....A--CAAG-A-A-A-AG-G-A--G-G-GA-G-TT-C-TGAG-TC..C.....AGGA--AA-GAAGC-G-C-TC--AA-AAAGC		2292
VERVET.KE.AGM155ACAGTGGAGG-CAAGGAA.....GA-CA-CACCAA-A-ATCCAG-CTGGC-G-AGGA-ATT-TGAA-TC..C.....AG-A--AA-AAA--G-C-TT-T-AA-AGTC		2525
VERVET.KE.AGM155ACAGTGGAGG-CAAGGAA.....GA-TCCA-C-CAAT-ATCCAG-TTGGACCG-GGGA-ATT-TGAA-TC..C.....AG-A--AA-AAA--G-GT-T--AA--GTC		2517
SABAEUS.SN.SABLCACAGTGGAGG-CAAGGAA.....AA-CCCA-CA-CAGI-ATCC-G-TTGGACAG-GGA-ATT-TGAA-TC..C.....AG-A--AA-AAA--G-CT-TA--A-AAATG		2020
SUN.GA.SIVSUNACAGTGGAGG-CAAGGAA.....AA-CC-CTC-CCA-ACAGAG-CCCT-CGA-G-GG--AC-GC.....-AG-A--AA-AAA--G-TT-C--AA--GTC		2705
MANDRIL.GA.MNDG1TCAGAAATAGGG-GAGGG-T.....AGCAC--A-AG-AG-A-G-GGGA-G-AGGAGC-A-G-T-G-ATATG--CCAG-A--T-A-AG-AGG-CTTC-TAAT		2652
LHOEST.CD.hoest447TTTAGACACACAGCTGAAAG.....G-C-C-A-AT-AA-GAG--G-AG-GA-TCA.....G-TA-T.....CAAG-A--AA-AAGAGG-CTCAG-G-AT-TGTCA		1982
LHOEST.CD.hoest448CTCGGCAACAGCTGAGGAG.....G-C-C.....AGTCCC-C-G-ATGTG-G-CGAGCCCTGCTGAATTCCT...CCAG--AA-A-AA-GT--TT-G-CRAT		1563
LHOEST.CD.hoest485CTCGGCAACACA-CTGAGGAA.....G-AC.....GTTCC-C-G-ATGTG-G-CGAGCC-TGCTGAATTC-TCT..CCAG--AA-A-AA-AC-TCT-G-CRAT--G		1560
RCM.NG.SIV-NGMCTTAGACACA-CTGAGGAA.....G-AC-A.....AGTTC-C-G--TGTTG-G-CGAGCCCTGCTGAATTC-TCT..CCAG--AA-A-AA-AC-TCT-G-CRAT--G		1563
COL.--SIV-CGU1CAAACAGGAGCAG-GGAAAGA.....GA-C-G-A-AGT-GGAGAGCT-CGA-G-G-G-GT-C-AT-T-G-A--T-G-A--T-T-G-G-AG--AA--A--A		2076
SYKES.KE.SYK173 GagCAACAGGAGTGTCT-T-G.....GA-G-G-AGCAGTGG-AG--GG--ACCAA-AG--TC-ATC-C-T--C-T.....-AG-A--AA-GA-AGA-G-G-TG--AA--TG-C-		p6
SYKES.KE.SYK173 PolN_R_E_V_L_S_P.....K_E_S_S_G_K_E_E_T_K_S_L_Y_P_S_L_S_Q_L_S_L_W_R_R_P_M_I_E_V_D_V_E_G_D		p10
SYKES.KE.SYK173O_O_G_S_A_L_A.....E_G_E_Q_W_E_G_G_D_Q_E_S_L_S_L_S_Q_L_S_L_W_R_R_P_M_I_E_V_D_V_E_G_D		p10

B.FR.HXB2 Pol 1871
 B.FR.HXB2 2425
 A.UG.U455 1817
 C.ET.ETH2220 1947
 D.CD.842R085 1762
 F1.BF.V1850 1822
 G.SE.SE6165 1772
 H.CF.90CF056 1739
 J.SE.SE7887 1621
 K.CM.MP535C 1999
 O1.AE.TH.CM240 1950
 O2.AG.NG.IBNG 2470
 O3.AB.RU.KAL153-2 1791
 O4.CP.CY.94CY032-3 2019
 N.CM.MVP5180 2455
 O.BE.ANT70 2479
 CPZ.US.CPZUS 2485
 CPZ.GA.CPZGAB 2486
 CPZ.CM.CM5 1836
 CPZ.CM.CM5 2127
 CPZ.CD.CPZANT 1864
 H2A.GW.ALI Pol 2803
 H2A.GW.ALI 2255
 H2A.SN.HIV25T 2810
 H2A.DE.BEN 2786
 H2B.GH.D205 2781
 H2B.CI.BHO 2137
 H2G.CI.BHT96 2703
 MAC.US.MM251 2200
 SMM.US.SIVSMMH9 2374
 STM.US.STM 2621
 TANTALUS.UG.TANI 2422
 GRIVET.ET.GRI-677 2649
 VERVET.KE.VBR-9063 2139
 VERVET.KE.AGMTYO 2641
 VERVET.KE.AGM55 2144
 SABAUS.SN.SABIC 2823
 SUN.GA.SIVSUN 2776
 MANDRL.GA.MNDGB1 2106
 LHOEST.KE.hoest 1687
 LHOEST.CD.hoest447 1687
 LHOEST.CD.hoest485 1684
 LHOEST.CD.hoest524 1688
 RCM.NG.SIV-NGM 2774
 COL. SIV-CGU1 2200
 SYKES.KE.SYK173 2607
 SYKES.KE.SYK173 Pol 1
 L.V.Q.M.L.V.D.T.G.A.D.D.T.I.I.R.E.D.I.....Q.L.H.Q.P.W.S.P.K.I.V.G.L.G.N.I.T.V.R.Q
 P.V.E.V.L.L.D.T.G.A.D.D.S.I.V.E.G.I.E.L.....G.S.N.Y.T.P.K.I.V.G.G.I.G.G.F.I.N.V.K.E
 C-GT-T-T-C-C-T-A-G-G-G-AGAG-A.....GGGA-C-AC-AT-CC-AG-A-G-A-C-A-TACC-A-G-
 C-GTA-T-T-CC-T-A-G-C-GC-AGAG-A.....GGGA-C-AT-AT-GT-G-AG-A-G-A-C-A-TACC-A-G-
 CGGTA-TCT-C-G-T-C-T-A-G-G-C-GG-AGAA-A.....GGG-AC-AT-AC-CT-AG-A-G-A-A-CACC-A-G-
 C-GTA-TAT-C-C-TT-C-T-A-G-C-GG-AGAA-A.....GGTA-C-AT-AC-CC-AG-A-G-A-C-A-TACC-A-G-
 C-GTA-TA-T-TA-C-G-C-GG-AGAA-A.....GGCA-C-AT-AC-CC-AG-T-G-A-A-TACC-ATG-
 CTGTA-TAT-G-T-TA-TG-AC-G-AGAG-A.....GG-A-T-AT-AT-GT-AG-A-A--A-T-TACC-A-G-
 CC-TA-TAT-T-TA-TG-C-G-AGAA.....GGTCC-AT-AT-CC-T-AG-A-A--T-TACT-A-G-
 CTGTA-TA-G-C-T-T-A-TG-C-GG-AGAA.....GGGCT-CA-AC-CC-T-AG-A-A--T-TACC-A-G-
 C-GTAC-T-G-TA-A-G-G-AGAA-A.....GGTCA-AC-CC-T-AG-A-A--T-TACT-AGG-
 C-GT-C-T-G-TA-A-G-G-AGAA-A.....CAITTAGA-C-C-CGG-T-G-AG-A-T-A-GGGA--TC-A-
 CTGTAA-TC-G-TA-A-TC-A-GAC-CACTTCCC-ATAA-CC-CGTT-C-GG-AG-A-T-A-AGGC-TC-T-C-A-G
 CC-TTA-G-A-GC-C-CA-A-TA-ATGA-A.....CAACTAA--CCT-G-A-A-AGGAC-T--A-G
 CCGTTA-G-AT-G-TA-A-A-G-T-CAGA-A.....CAATTAT-TCC-G-C-A-AGGC-T--A-G
 CC-TCAG-AT-G-G-CA-TA-A-GCAGA-A.....CAATTATTC-C-G-AGG-T-A-A-G
 A-GTCACT-T-T-G-TCT-GA-AT.....GG-AC-AT-G-A-A-G-T-A-A-G-T-T-A-GC
 CTGT-ACT-T-G-C-G-A-G-GCT-GATTC-G.....TCCATAGA--TCAC--G-C-GG-AG-G-T-T-A-G-ACRA-T-G--A-G
 CT-TAAG-C-G-TA-T-GA-A-A.....AAATTAAT-AT-C-G-----A-T-AACT-A-G-
 C-GT-TC-A-G-G-C-T-G-TA-TT-C-ATCAG-A.....AGATAGA--CCT-AT-CT--G-G-AG--A-G-ACRA-T-G--A-G
 C-GT-TCI-TC-T-C-T-G-CA-C-TTC-ATTAG-T.....AGATAGA--CCT-AT-C-T-G-G-AG--A-G-ACRA-T-G--A-G
 A-GTAACT-C-G-TA-C-TTC-ATTAG-T.....AGATAGA--CCT-AT-CT--G-G-AG--A-G-ACRA-T-G--A-G
 AGGT-G-CT-TA-C-A-TT-A.....GATTTAGG-A-AT-G-----T-A-A-A-T-T-A-
 AGTGCC-G-G-TA-TA-GAAGA-G-T.....CAATTAA--CGCT--C-G--GT-CAA-TCA--A-G-AGCA--G--A-
 -GGT-C-ATGT--G-----TA-A-TAG-----GAAGA-A-T.....CAATTAA-ACCA-CC--TC-----AG--A-T-A-AAAC--T-C--T--G--
 L.V.Q.M.L.V.D.T.G.A.D.D.T.I.I.R.E.D.I.....Q.L.H.Q.P.W.S.P.K.I.V.G.L.G.N.I.T.V.R.Q





B.FR.HXB2 Pol	Y D Q I L I E I C G H K A I G T V L V G P T P V N I I G R N L L T T O I G C T	p10
B.FR.HXB2	GTATGATCACTATAGAAAATCTGTGACATAAAGCTATA. GGTACAGTATTAGGACCTACACCTGTCAACATAAATTTGGAAGAAATCTGTGACTCAGATTGGTTGCACT	2540
A.UG.U455	T T A A G A G G A A T G A T C	1986
C.ET.ETH2220	A A T A A G C C C C C C A	1932
CD.84ZR085	C C G G C C C C C C C	2062
F1.BF.V1850	A C T A A G G G G G G G T	1877
G.SE.SE6165	A G C T A A G G A G G G G G C	1937
H.CF.90CF056	G A G C G T G A G A A G G G G G C	1887
J.SE.SE7887	A G T A A G T A A G T A A G T A A G T A A G T A A G T	1854
K.CM.MP535C	A T A A G T A A G T A A G T A A G T A A G T A A G T	1736
O1.AE.TH.CM240	A T A A G T A A G T A A G T A A G T A A G T A A G T	2114
02.AG.NG.IBNG	A T A A G T A A G T A A G T A A G T A A G T A A G T	2065
03.AB.RU.KAL153-2	A T A A G T A A G T A A G T A A G T A A G T A A G T	2585
04.CPX.CY.94CY032-3	A T A A G T A A G T A A G T A A G T A A G T A A G T	1906
N.CM.MVP3180	A T A A G T A A G T A A G T A A G T A A G T A A G T	2134
O.BE.ANT70	A T A A G T A A G T A A G T A A G T A A G T A A G T	2570
CPZ.US.CPZUS	A T A A G T A A G T A A G T A A G T A A G T A A G T	2594
CPZ.GA.CPZGAB	A T A A G T A A G T A A G T A A G T A A G T A A G T	2600
CPZ.CM.CM3	A T A A G T A A G T A A G T A A G T A A G T A A G T	1951
CPZ.CM.CM5	A T A A G T A A G T A A G T A A G T A A G T A A G T	2242
CPZ.CD.CPZANT	A T A A G T A A G T A A G T A A G T A A G T A A G T	1979
H2A.GW.ALI Pol	Y E D V F I K V L N K R V K A T I M V G D T P I N I F G R N I L L T T A L G M S	p10
H2A.GW.ALI	AGTG-GAA--A-G-ACTAAATA-A-G-TA-A-. CC-CA-A-GAC--TGAC-C-AA--T-TT-C-CA-T--AGCCT-A-CATGT-A	2918
H2A.SN.HIV25T	A--A-AA-TG--GAA--AG-G-A-TAAATA-A-G-TA-G-. CC-CA-A-GAC--TGA-C-AA--TT-C--CA-TC--AGCCT-A-CATGT-A	2370
H2A.DE.BEN	A--CA-AA-TG--GAA--A-G-ACTAAATA-A-G-TA-G-. CC-CA-A-GAC--GA-C-AA--CT-C--A-TC--AGCCT-A-CATGT-A	2925
H2B.GH.D205	A--CA-AG-TG--GAA--G-AGTG-A-A-G-TA-GG-. CA-TA-A-GAC--GA-C-AA-A--TT-C--A-T--A-ACCT-G-CATG--	2901
H2B.CI.BHO	A--CA-BA-TG--GAA--G-AGTA-A-A-G-TA-G-. CA--A-GAC--GGAC-C-AA-A--TT-C--A-T--A-AGCT-A-CATG--	2896
H2G.CI.BRT96	A--CA-BA-TG--GAA--TA-G-GTA-A-A-G-ATACAG-. C-T-TA-GAC--GGAC-C-AA-A--T-CT-G-CA-T--AGT-A-C-CATGT--	2252
MAC.US.MM251	A--CA-BA-TG--AAA--G-T-TA-CA-A-GAAT--AG-. G--A-CA-GAC--GGAC-T-GA-T--TT-C--T-C-A-AGCTC-G-GAIGT--	2818
SMW.US.SIVSMW9	A--A-AG-TG--AAA--A-G-A-TA-CA-AGT-AT--AG-. A--TA-GAC--GAC-C-AA-T--TT-C--C-A-AGCTI-G-GAIGT-A	2315
STM.US.STM	A--T-A-AG-TG--AA--GG-A-TA-CA-G--ATA-A-. A--TA-TA-GACT--GAC-C-CA-T--TG-C-G-C-A--AGCTI-G-GAIGT-A	2489
TANTALUS.UG.TANI	A--C-GGG-G--AAG-C-T-GGAG-ATA-A-TCATC-CC-. CT-A-TC-A--AG-AA-T-T-A--A-TC-AG-A-CTT-A-AGCA-GA	2736
GRIVET.ET.GRI-677	AG-A-TG--GAAG--AGAA-ATA-A-TTTTG-G-. C-TC-TT-G-A-C-A-T--A-T--A--GCA-CATG-AA	2537
VERVET.KE.VBR-9063	A--A-CG-C-GGGAAG--A--AGAA-ATA-A-TTTTG-G-. A--A-GT--G-A-C-A-T--A-T--T-AC-AG-A-CAGCA-AGCT-GA	2764
VERVET.KE.AGMTYO	A--A-TG--GAAG--C-T-GGAA-ACA-G-TTTA-G-. A--CC-CA--G-A-T-CA-T--C-A--CT-T--G-C-GCA-TGCC-CGG	2254
VERVET.KE.AGM155	A--AG-G-T-GGGAAG--AG-T-GGAA-ACA-A-TTTTG-G-. G--CA-GA--AGC-T-CA-A--GCA-AGCA-AGC-AA	2756
SABAEUS.SN.SABIC	A--CC-CA-TCAGGAG--A--AGAA-ATA-A-CTTG-C-. C--CA-T-G--GAA--A-T-T-T-A--G-C-AG-A-AT-A-AGTT--	2259
SUN.GA.SIVSUN	A--A-TG--CT-TG--TA-G-ATA-A-C--C-. TA-G-ATA-A-C--C--C--A--AG-T-C-A-G--A-TC-TT-AGCCT-A-GG-A-GA	2938
MANDILL.GA.MNDGB1	A--AGAG-TG--CT-T--AGCA--A-A-CTA-AG-. AGCA--A-A-CTA-AG--G--C--I--C--A-AG-T-TG-A-G--A-T--A--GCA-G-GCC-AA	2221
LHOEST.KE.ROEST	A--CA-GG-TG--CT-T--AGCA--A-A-C-A-AG-. AGCA--A-A-C-A-AG--G--C--T--G--C--A--AG-T-TG-A--A-T--A--TG-G-AGCA-GA	2886
LHOEST.CD.ROEST447	A--CA-GG-TG--CT-T--AGCA--GA-G-C-A-AG-. AGCA--GA-G-C-A-AG--G--C--T--G--C--A--AG-T-TG-A--A-T--A--TG-G-AGCA-GA	1802
LHOEST.CD.ROEST485	A--TCAGGG-TG--TT-T--C--AGCA--GA-G-CCA-AG-. AGCA--GA-G-CCA-AG--G--C--T--G--C--A--AG-T-TG-A--A-T--A--TG-G-AGCA-GA	1799
LHOEST.CD.ROEST524	A--T-A-C-TTGTAAA--ACT-AGCA--GA-G-CTA-ACAT-. C-CAT--G--A--A-A-T--G-A--C-G--GCA-AT-A-AGC-T--	2889
RCM.NG.SIV-NGM	--AGAG-T-A-A-TC--A-G-GG-A--A-AG-TAC-G-. AGATA-C--A--AC-C-CA-A-T--G-A-T--A--AAG-AGGGA--TAGTA	2315
COL.SIV-CGU1		
SYKES.KE.SYK173	A--AGAA-T--TAGAT-CACTG-TGT-AA--CCTC--GA-A-AGAAAGCAAGTAGAG--G--TC-T--L-L-V-G-P-T-P-V-N-I-L-G-R-N-I-L-T-K-L-G-V-K	2737
SYKES.KE.SYK173 Pol	Y R N I R F T V V K P S G K R K Q V E G T L L V G P T P V N I L G R N I L T K L G V K	p10

B.FR.HXB2 Pol	T E M E K E G K I S K I G P E N P Y N T P V F A I K K K K D S T K W R K L V D F R E L N K	p51
B.FR.HXB2	CAGAGATGGAAGGAAAGGAAATTTCAAAAATTTGGGCTGAAATCCATACATACTCCAGTATTTGCCATAAAGAAAAGACAGTACTAAATGGAGAAAATTTAGTAGATTTTCAGAGAACCTTAATAA	2794
A.UG.U455	AT-----A-----A-----T-----G-----C-----G-----G-----C-----C-----	2240
C.ET.ETH2220	A-A-----GC-----A-----G-----G-----A-----G-----G-----G-----G-----G-----	2186
D.CD.84ZRO85	T-----T-----A-----A-----G-----A-----G-----G-----G-----G-----G-----	2316
F1.BE.VI850	T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	2131
G.SE.SE6165	A-----A-----G-----A-----A-----C-----A-----A-----G-----G-----G-----G-----G-----	2191
H.CF.90CF056	-----A-----A-----A-----C-----G-----A-----G-----G-----T-----A-----G-----G-----	2141
J.SE.SE7887	-----A-----A-----A-----C-----T-----C-----G-----G-----G-----G-----G-----G-----	2108
K.CM.MP535C	-----A-----A-----A-----C-----G-----G-----G-----G-----G-----G-----G-----G-----	1990
01.AE.TH.CM240	-----A-----A-----A-----C-----G-----G-----G-----G-----G-----G-----G-----G-----	2368
02.AG.NG.IBNG	-----A-----A-----A-----C-----G-----G-----G-----G-----G-----G-----G-----G-----	2319
03.AB.RU.KAL153-2	AG-----C-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	2839
04.CPX.CV.94CY032-3	-----A-----A-----A-----T-----G-----A-----A-----T-----A-----G-----T-----C-----C-----	2160
N.CM.MVP5180	A-----A-----C-----A-----C-----G-----A-----A-----A-----T-----A-----G-----T-----C-----C-----	2388
O.BE.ANT70	AG-----A-----C-----A-----C-----G-----A-----A-----A-----T-----A-----G-----T-----C-----C-----	2824
CPZ.US.CPZUS	AG-----A-----C-----A-----C-----G-----A-----A-----A-----T-----A-----G-----T-----C-----C-----	2848
CPZ.GA.CPZGAB	A-----A-----A-----G-----A-----G-----A-----G-----A-----A-----T-----A-----G-----T-----C-----C-----	2854
CPZ.CM.CM3	AG-----A-----A-----G-----A-----G-----A-----G-----A-----A-----T-----A-----G-----T-----C-----C-----	2855
CPZ.CM.CM5	AG-----A-----A-----G-----A-----G-----A-----G-----A-----A-----T-----A-----G-----T-----C-----C-----	2496
CPZ.CD.CPZANT	ATA-----T-----A-----GCA-----AAT-----G-----T-----G-----A-----A-----A-----T-----A-----G-----T-----C-----C-----	2233
H2A.GW.ALI Pol	E K T E R E G O L E E A P P T N P Y N T P T F A I K K K K D K K K W R M L I D F R E L N K	p51
H2A.GW.ALI	A-A-A-C-----GA-----CC-----T-----AGAGG-----GGCACCT-----AACT-----C-----T-----T-----C-----CAC-----A-----I-----K-----K-----K-----W-----R-----M-----L-----I-----D-----F-----R-----E-----L-----N-----K	3172
H2A.SN.HIV2ST	AGA-A-----G-----G-----CC-----GC-----AGAGG-----GGCACCT-----AACT-----T-----T-----C-----CAC-----A-----C-----G-----AA-----AC-----T-----GTC-----A-----A-----A-----C-----	2624
H2A.DE.BEN	A-A-A-----G-----G-----CC-----GC-----AGA-----GGCACCT-----AACT-----T-----T-----C-----CAC-----A-----T-----G-----AG-----AC-----T-----GTC-----GA-----T-----A-----A-----C-----	3179
H2B.GH.D205	A-A-A-----G-----AC-----T-----AGA-----G-----GCACCC-----ACT-----A-----A-----C-----A-----CACC-----C-----A-----G-----T-----AG-----AC-----T-----GTC-----A-----C-----T-----GT-----A-----C-----	3155
H2B.CI.BHO	A-A-A-----A-----G-----AC-----CT-----AGA-----G-----GGCGCT-----ACT-----A-----A-----G-----A-----G-----T-----G-----CAC-----C-----A-----G-----T-----GTC-----A-----A-----C-----	3150
H2G.CI.ABT96	ACA-A-----C-----T-----GGA-----G-----GCACCC-----CACC-----G-----T-----C-----A-----CAC-----A-----T-----C-----G-----G-----AA-----AC-----T-----GTC-----A-----C-----C-----G-----	2506
MAC.US.MM251	A-A-----T-----TC-----GT-----GGAGG-----GC-----CCC-----GACC-----T-----C-----C-----CAC-----T-----R-----GR-----T-----AG-----A-----G-----T-----GTC-----GA-----T-----G-----A-----A-----G-----	3072
SMW.US.SIVSMH9	A-A-A-----G-----A-----T-----CC-----GT-----AGAGG-----GC-----CCT-----AAC-----T-----T-----C-----C-----CAC-----T-----G-----A-----A-----G-----T-----GTC-----GA-----T-----A-----A-----C-----	2569
STM.US.STM	A-A-A-----G-----T-----C-----GC-----AGA-----G-----GC-----CCT-----AAC-----T-----T-----C-----C-----CAC-----T-----C-----CAC-----T-----GTC-----CGG-----A-----T-----T-----A-----C-----	2743
TANTALUS.UG.TANI	AG-----CTC-----A-----A-----T-----A-----T-----GG-----AGGA-----G-----T-----T-----A-----G-----TGT-----CCG-----T-----A-----A-----AC-----G-----G-----T-----GTC-----GA-----T-----CT-----C-----	2996
GRIVET.ET.GRI-677	A-C-A-----G-----A-----T-----A-----T-----GG-----A-----AGGA-----G-----T-----T-----A-----G-----TGT-----C-----GA-----G-----A-----A-----C-----T-----GTC-----T-----G-----A-----C-----	2797
VERVET.KE.VER-9063	A-C-A-----T-----A-----A-----CAGC-----G-----A-----AGGA-----G-----T-----T-----C-----A-----C-----TGT-----C-----GA-----G-----G-----A-----A-----C-----T-----GTC-----A-----A-----	3024
VERVET.KE.AGMTYO	CC-----AT-----A-----GC-----A-----CAGT-----G-----A-----AGGA-----G-----C-----A-----A-----TGT-----C-----A-----A-----AAGC-----G-----G-----A-----A-----T-----GTC-----G-----C-----T-----GT-----A-----	2514
VERVET.KE.AGMT55	A-ACAT-----A-----GG-----A-----A-----T-----AAGC-----GGG-----A-----AGGA-----C-----G-----A-----A-----CTGT-----G-----A-----AAT-----AC-----G-----T-----GTC-----G-----G-----C-----	3016
VERVET.DE.AGM3	ACC-----AT-----A-----G-----A-----AGC-----G-----A-----AGGA-----G-----G-----C-----G-----TG-----A-----AGT-----AC-----T-----GTC-----G-----T-----G-----A-----A-----	2519
SABAEUS.SN.SAB1C	A-----CT-----A-----AC-----AC-----TT-----GGA-----G-----A-----A-----A-----A-----T-----C-----C-----A-----A-----G-----G-----T-----CA-----G-----C-----GC-----GT-----A-----	3192
SUN.GA.SIVSUN	ATAGAT-----A-----GG-----A-----GCA-----A-----T-----GCA-----A-----A-----T-----T-----A-----A-----G-----G-----T-----AA-----A-----G-----T-----T-----A-----A-----GG	3151
MANDRILL.GA.MNDGB1	ATAGAT-----A-----GG-----A-----TCC-----G-----G-----AT-----A-----GG-----C-----T-----C-----A-----A-----A-----T-----C-----G-----G-----T-----AA-----A-----G-----T-----T-----A-----C-----	2481
LHOEST.KE.roest	ATAGAT-----A-----A-----A-----TCC-----G-----G-----AT-----A-----GG-----C-----T-----C-----A-----A-----A-----T-----C-----G-----G-----AA-----A-----G-----T-----A-----G-----C-----	3146
LHOEST.CD.roest447	ATAGAT-----A-----A-----A-----GCA-----ATG-----A-----GG-----C-----T-----T-----C-----A-----A-----A-----C-----T-----C-----A-----A-----G-----G-----C-----GA-----T-----T-----C-----A-----GC	2062
LHOEST.CD.roest485	ATAGAT-----A-----A-----A-----GCA-----ATG-----A-----GG-----C-----T-----T-----C-----A-----A-----A-----C-----T-----C-----A-----A-----G-----G-----C-----GA-----T-----T-----C-----A-----GC	2059
LHOEST.CD.roest524	AT-----CC-----G-----A-----T-----G-----A-----A-----A-----A-----A-----T-----T-----G-----T-----C-----A-----A-----A-----A-----T-----TCC-----A-----GC-----T-----G-----T-----G-----	3143
COL.SIV-CGU1	A-AGAT-----GCA-----G-----A-----AGA-----GCA-----AAITTA-----GG-----T-----A-----	2569
SYKES.KE.SYK173	A-----C-----CTC-----C-----AC-----T-----GGAG-----GC-----A-----T-----ACC-----T-----T-----T-----C-----TA-----T-----G-----AA-----AG-----G-----G-----T-----G-----TC-----CA-----T-----T-----A-----T-----A-----G-----	2991
SYKES.KE.SYK173 Pol	E D M L K A G O L E E K A S P T N P Y N T P V F V I R K K K D K K K W R M L I D F R E L N K	p51

B.FR.HXB2 Pol	K Y T A F T I P S I N N E T P G I R Y Q Y N V L P Q G W K G S P A I F Q S S M T K I L	P51
B.FR.HXB2	AAGTATACTGCATTACCATACCTAGTATAAACAATGAGACACACAGGATGATGATGACATGATGATCAGTACAAATGCTTCCACAGGATGGAAGGATCCACAGCAATATTCCAAAGTAGCATGACAAAATCTTTAG	3054
A.UG.U455	-----G-C-T-----T-----G-C-----A-T-----C-C-----C-----G-----T-----G-----	2500
C.ET.ETH2220	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	2446
D.CD.84ZR085	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	2576
F1.BF.VI850	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	2391
G.SE.SE6165	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	2451
H.CF.90CF056	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	2401
J.SE.SE7887	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	2368
K.CM.MP535C	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	2250
O1.AE.TH.CM240	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	2628
02.AG.NG.IBNG	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	2579
03.AB.RU.KAL153-2	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	3099
04.CP.CY.94CY032-3	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	2420
N.CM.MVP5180	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	2648
O.BE.ANT70	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	3084
CPZ.US.CPZUS	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	3108
CPZ.GA.CPZGAB	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	3114
CPZ.CH.CM3	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	3115
CPZ.CH.CM5	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	2465
CPZ.CD.CPZANT	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	2756
H2A.GW.ALI Pol	-----C-----C-----C-----A-----A-----A-----C-----C-----C-----C-----	2493
H2A.GW.ALI	Q Y T A F T L P S V N N A E P G K R Y I Y K V L P Q G W K G S P A I F Q H T M R O I L	P51
H2A.SN.HIV2ST	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	3432
H2A.DE.BEN	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	2884
H2B.GH.D205	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	3439
H2B.CI.BHO	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	3412
H2G.CI.ABT96	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	3410
MAC.US.MM251	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	2766
SMM.US.SIVSMMH9	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	2766
SFM.US.SFM	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	3332
TANTALUS.UG.TANI	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	2829
GRIVET.ET.GRI-677	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	3003
VERVET.KE.VER-9063	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	3256
VERVET.KE.AGM155	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	3057
VERVET.DE.AGM3	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	3284
SABAEUS.SN.SABIC	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	2774
SUN.GA.SIVSUN	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	3276
MANDILL.GA.MNDGB1	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	2779
LHOEST.KE.roest	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	3452
LHOEST.CD.roest447	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	3411
LHOEST.CD.roest485	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	2741
LHOEST.CD.roest524	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	3406
RCM.NG.SIV-NGM	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	2322
COL.SIV-CGU1	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	2322
SYKES.KE.SYK173	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	2319
SYKES.KE.SYK173 Pol	-----TC-----ATCAG-----ATCAG-----A-----ATA-----Y-----K-----V-----L-----P-----Q-----G-----W-----K-----G-----S-----P-----A-----I-----F-----Q-----H-----T-----M-----R-----O-----I-----L-----	3403

Pol p66 reverse transcriptase continued (RT is dimer of p51 and p66)

Pol p51 reverse transcriptase end / Pol p15 RNase H

Table of HIV-1/HIV-2/SIV complete genomes with columns for protein names (e.g., B.FR.HXB2 Pol, A.UG.U455) and nucleotide sequences. The sequences are aligned in rows, with protein names listed on the left and their corresponding positions (e.g., p15, 3949) on the right.

B.FR.HXB2 Pol	T L T T T N Q K T E L Q A I Y L A L Q D S G L E V N I V T D S Q Y A L G I I O A Q P	p15
B.FR.HXB2	CACCCCTAAGTACACAAATCAGAAGACTGAGTTTACAAAGCAATTTTCTAGCTTTTCAGGATTCGGGATTTAGAACTAGTAAACATAGTAAAGACTCACAAATAGCATTAGGAATCATTCAGCACAA	4079
A.UG.U455	TT-----G-----A-----A-----T-----CC-----C-----A-----C-----T-----G-----G-----G-----G-----	3525
D.CD.84ZR085	TT-----T-----A-----A-----G-----CC-----G-----A-----A-----C-----C-----G-----G-----C-----TG-----	3471
F1.BF.VI850	C-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3601
G.SE.SE6165	T-----G-----G-----C-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	3419
H.CF.90CF056	T-----G-----A-----A-----G-----A-----G-----A-----A-----G-----A-----A-----G-----A-----A-----G-----A-----A-----G-----	3476
J.SE.SE7887	T-----G-----A-----A-----G-----A-----G-----A-----A-----G-----A-----A-----G-----A-----A-----G-----A-----A-----G-----	3426
K.CM.MP535C	T-----A-----A-----AG-----A-----A-----C-----G-----T-----A-----A-----G-----C-----G-----G-----T-----A-----	3393
01.AE.TH.CM240	TT-----G-----G-----A-----A-----T-----G-----CC-----T-----C-----A-----C-----T-----T-----G-----G-----G-----G-----	3275
02.AG.NG.IBNG	TT-----G-----G-----A-----A-----T-----G-----CC-----T-----C-----A-----C-----T-----T-----G-----G-----G-----G-----	3653
03.AB.RU.KAL153-2	T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3604
04.CPX.CV.94CY032-3	T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	4124
N.CM.MBF30	T-----TA-----TG-----A-----C-----A-----G-----A-----A-----G-----A-----CGG-----T-----C-----T-----T-----G-----A-----TT-----G-----	3445
O.CM.VFP5180	A-----AGT-----GAA-----G-----C-----A-----G-----A-----ATG-----TG-----A-----TAA-----C-----C-----CAAGGAGC-----T-----G-----C-----ATCCT-----C-----	3673
O.BE.ANT70	A-----AAT-----GA-----G-----C-----C-----A-----AG-----A-----ATC-----G-----A-----TA-----C-----A-----CAA-----GA-----AC-----T-----TG-----C-----TG-----C-----TG-----C-----	4109
CPZ.US.CPZUS	A-----TT-----GAAA-----T-----T-----A-----G-----A-----C-----CAC-----GG-----A-----CT-----A-----AA-----A-----A-----AGATAC-----T-----TG-----C-----TG-----C-----TG-----C-----	4133
CPZ.GA.CPZGAB	T-----G-----T-----GAAA-----T-----C-----T-----C-----AG-----A-----A-----A-----G-----TT-----GCT-----G-----C-----C-----A-----ATCA-----C-----G-----T-----G-----T-----GT-----G-----	4139
CPZ.CM.CM3	A-----A-----T-----GAAA-----T-----C-----T-----AG-----A-----AC-----GAC-----TG-----G-----A-----A-----A-----G-----A-----A-----CAIAT-----T-----T-----C-----T-----T-----C-----	4140
CPZ.CM.CM5	A-----ATT-----GAAA-----T-----T-----T-----AG-----A-----AC-----AC-----AC-----TG-----G-----A-----A-----A-----CAIAT-----C-----T-----C-----T-----T-----T-----	3490
CPZ.CD.CPZANT	ACA-----CAAA-----G-----C-----C-----AC-----AG-----A-----A-----G-----G-----CT-----A-----G-----C-----AG-----AAC-----GGCCCC-----C-----T-----C-----T-----G-----GTACC-----	3781
H2A.GW.ALI Pol	V L E Q T T N Q O A E L E A F A L A V T D S G P K A N I I V D S Q Y V M G I V A G Q P	p15
H2A.GW.ALI	GGTA-----GAAC-----A-----T-----C-----C-----AG-----A-----A-----G-----CT-----CGCG-----G-----AG-----AAACA-----C-----A-----TCC-----A-----CC-----T-----TA-----GT-----T-----G-----T-----A-----G-----AG-----AGC-----GC-----G-----	4454
H2A.SN.HIV2ST	GCTAT-----GAGC-----A-----C-----C-----C-----AG-----A-----A-----G-----CT-----CGCA-----G-----AG-----AACA-----C-----A-----TCC-----A-----G-----CC-----TA-----GT-----T-----A-----G-----AG-----AGC-----GC-----	3906
H2A.DE.BEN	AGTAT-----GAAC-----A-----T-----C-----C-----G-----A-----A-----G-----TCT-----CGGA-----G-----AC-----GCA-----C-----A-----CCC-----A-----G-----T-----T-----CA-----GT-----T-----G-----T-----A-----G-----G-----AG-----AGC-----GC-----	4461
H2B.GH.D205	AGTGT-----GAAC-----G-----C-----AC-----AG-----A-----AC-----TG-----T-----GCAT-----A-----AACA-----C-----A-----A-----CC-----C-----T-----CA-----GT-----T-----G-----TCA-----G-----A-----AGCT-----G-----	4434
H2B.CI.BHO	AC-----RT-----GAAC-----A-----C-----AG-----A-----C-----TG-----T-----GCN-----AC-----A-----C-----A-----CC-----C-----G-----C-----T-----CA-----GT-----T-----G-----TCA-----G-----A-----AGCT-----	4432
H2G.CI.ABT96	PCT-----T-----GAGC-----G-----C-----C-----G-----R-----AC-----TG-----G-----T-----GCAA-----G-----C-----A-----TCC-----A-----C-----T-----CA-----GTR-----T-----C-----G-----TCA-----G-----R-----AGC-----GR-----	3788
MAC.US.MM251	AGTGT-----GAAC-----G-----T-----T-----AC-----AG-----A-----A-----G-----T-----CTCA-----G-----A-----ACA-----C-----A-----GCC-----A-----GAC-----T-----TA-----GT-----T-----C-----TTA-----G-----C-----AG-----AGC-----GT-----C-----	4354
SMW.US.SIVSMW9	ACTG-----GAAAC-----G-----T-----C-----AC-----AG-----A-----A-----GG-----CT-----T-----GCCA-----G-----C-----GC-----C-----A-----TCC-----A-----C-----T-----T-----GGT-----TCA-----G-----C-----AG-----AGC-----G-----	3851
STM.US.STM	AG-----TT-----GAAC-----G-----T-----T-----C-----C-----AC-----AG-----A-----A-----G-----T-----GCCA-----G-----C-----GC-----C-----A-----TCC-----A-----C-----T-----T-----GGT-----TCA-----G-----C-----AG-----AGC-----G-----	4025
TANTALUS.UG.TANI	AGAAAT-----GAG-----T-----C-----G-----C-----AC-----AG-----A-----A-----G-----G-----CCT-----A-----AA-----AC-----AGTAAAGCA-----T-----T-----G-----TCA-----G-----C-----AG-----AGC-----G-----	4278
GRIVET.ET.GRI-677	GGAAT-----GAG-----T-----C-----C-----AC-----AG-----A-----A-----AC-----G-----A-----GA-----G-----G-----AGT-----GA-----A-----T-----T-----C-----T-----T-----T-----A-----GAAC-----T-----GAC-----TGT-----	4082
VERVET.KE.VER-9063	G-----AAT-----GAGC-----G-----T-----C-----C-----AG-----A-----A-----G-----GT-----C-----AA-----AA-----G-----C-----G-----AGT-----GA-----A-----T-----CA-----G-----A-----G-----AT-----GAC-----G-----C-----	4306
VERVET.KE.AGMTYO	A-----AT-----GAAA-----T-----C-----AC-----AG-----AA-----A-----AC-----T-----AA-----AA-----G-----G-----A-----CAGT-----GCCTA-----T-----G-----T-----A-----G-----TT-----GAC-----C-----	3793
VERVET.KE.AGMT55	ACAG-----GAAA-----T-----C-----AC-----AG-----AC-----GAC-----C-----AA-----AA-----G-----C-----G-----AGC-----CCCTA-----C-----T-----G-----A-----G-----AT-----GAC-----G-----C-----	4298
VERVET.DE.AGM3	A-----AAT-----GAAA-----T-----C-----C-----C-----AG-----A-----A-----ATG-----C-----AA-----AA-----G-----AC-----AG-----A-----AGT-----GCCTA-----T-----A-----T-----A-----G-----AT-----AACT-----C-----C-----	3801
SABAEUS.SN.SAB1C	AG-----AT-----GAGA-----T-----T-----C-----C-----C-----AG-----A-----A-----G-----G-----G-----C-----CTGT-----C-----AAGA-----C-----T-----AGCA-----T-----A-----TA-----G-----T-----T-----AGC-----GGG-----C-----	4477
SUN.GA.SIVSUN	G-----AAT-----GGAA-----T-----C-----C-----C-----C-----AG-----A-----AC-----G-----G-----T-----AA-----AA-----G-----C-----G-----T-----AGTA-----GAGCTC-----T-----T-----A-----C-----A-----C-----CGTAAAGAG-----C-----	4433
MANDILL.GA.MNDGB1	AGAAATCTAA-----T-----C-----C-----C-----C-----AG-----A-----A-----A-----T-----G-----AA-----A-----C-----AG-----A-----AGTT-----T-----T-----T-----G-----C-----A-----C-----CGTAAAGAG-----C-----	4433
LHOEST.KE.roest	GGAT-----T-----C-----C-----C-----C-----AG-----A-----A-----AC-----G-----G-----G-----CTRA-----G-----A-----A-----A-----AGCAATAGTA-----T-----T-----T-----A-----GAA-----AT-----GTC-----CA-----AG-----	3766
LHOEST.CD.roest447	AGAAAT-----A-----T-----C-----C-----C-----C-----AG-----A-----A-----A-----G-----G-----TG-----C-----TAT-----G-----A-----A-----A-----RAGA-----CHGTAAAGCAGG-----C-----T-----TGA-----AA-----G-----TT-----ATCTCA-----GG-----T-----	4434
LHOEST.CD.roest485	AGAAAT-----A-----T-----C-----C-----C-----C-----AG-----A-----A-----A-----G-----G-----TG-----C-----TAT-----G-----A-----A-----A-----RAGA-----CHGTAAAGCAGG-----C-----T-----TGA-----AA-----G-----TT-----ATCTCA-----GG-----T-----	3350
LHOEST.CD.roest524	AGAAAT-----CAA-----T-----T-----T-----C-----G-----A-----A-----G-----G-----A-----TGT-----A-----A-----AGCAA-----C-----AC-----T-----T-----C-----T-----A-----GAAAG-----T-----AAGTCAGAG-----T-----	3347
RCM.NG.SIV-NGM	AGAAAT-----GAAA-----T-----T-----C-----C-----AG-----A-----A-----GG-----G-----A-----TA-----G-----C-----AA-----A-----C-----A-----AA-----AGG-----T-----T-----G-----C-----AT-----AGC-----G-----AGT-----	4428
COL.SIV-CGU1	--AT-----TCA-----AGGGT-----C-----AG-----A-----C-----GAC-----G-----G-----TG-----C-----AA-----AC-----GG-----CC-----AA-----AG-----A-----G-----TT-----G-----AGT-----TT-----G-----A-----ACT-----GG-----G-----	3848
SYKES.KE.SYK173	GTAT-----GGAGC-----A-----C-----AC-----G-----A-----A-----G-----G-----GCT-----GCTCA-----G-----A-----C-----AAA-----GACA-----T-----T-----GT-----T-----AT-----C-----TC-----ATGA-----TGC-----C-----	4279
SYKES.KE.SYK173 Pol	Y L E Q T T N Q O A E L E L E G L L M A L Q D S K D K V N I V D S Q Y S Y G I L M T C P	p15

HIV-1/HIV-2/SIV complete genomes

B.FR.HXB2 Pol B.FR.HXB2	A_V_H_V_A_S_G_Y_I_E_A_E_V_I_P_A_E_T_G_Q_F_T_A_Y_F_L_L_K_L_A_G_R_W_P_V_K_T_I_H_T_D_N_G	p31 4584
A.UG.U455	GCAGTTGATGAGCAGGATATAGAAAGCAGAAAGTATCCAGGAAACAGGAGGAGAAACAGGATGCCCAGTAAAACAATACATACTGACAATGGCA	
C.ET.ETH2220	-----C-C-C-C-----A-G-----C-A-AC-----GT-----C-A-----	4030
D.CD.84Z8085	-----C-C-C-G-----A-A-----C-AC-----G-----C-GGT-----A-T-----	3976
F1.BF.VI850	-----C-C-C-A-----G-----C-C-CA-A-----G-----GT-G-----A-----	4106
G.SE.SE6165	-----T-AG-C-----A-----C-C-CA-A-----G-----T-----A-----	3924
H.CF.90CF056	-----C-----C-----C-A-A-----C-A-A-----C-GT-----A-T-----	3981
J.SE.SE7887	-----C-C-C-C-----AA-----C-C-G-G-----C-A-C-----A-----	3931
K.CM.MP535C	-----T-C-C-----C-A-A-----G-----F-----A-A-----GT-----A-----	3780
02.AE.TH.CM240	-----C-C-G-----C-A-A-----C-C-A-AC-----C-----C-A-AC-----	3780
01.AG.NG.IBNG	-----C-C-G-----C-A-A-----C-C-GC-----C-----G-----C-A-C-----T-----	4158
03.AB.RU.KAL153-2	-----C-C-----A-A-----G-----C-A-A-----G-----C-----A-----	4109
04.CPX.CY.94CY032-3	-----G-T-----C-CT-----A-A-----C-C-CA-AC-----T-----A-----	4629
N.CM.MVP5180	-----G-C-----G-A-----T-----G-----A-----G-----T-----G-----A-----	3950
O.BE.ANT70	-----T-C-----A-AC-T-----A-A-----T-----C-----C-G-----T-----A-----	4178
CPZ.US.CPZUS	-----T-C-----G-A-----G-----A-----T-----C-----C-G-----T-----A-----	4614
CPZ.GA.CPZGAB	-----T-C-----G-A-----G-----A-----T-----C-----C-G-----T-----A-----	4638
CPZ.CM.CAM3	-----A-C-----G-C-----C-----T-----A-A-----G-----C-----A-----	4644
CPZ.CD.CPZANT	-----A-C-----G-C-----C-----T-----A-A-----G-----C-----A-----	4645
	-----A-C-----G-C-----C-----T-----A-A-----G-----C-----A-----	4645
	-----G-----G-----C-----T-----A-A-----G-----C-----A-----	3998
	-----A-----G-----C-----T-----A-A-----G-----C-----A-----	4286
	-----G-----T-----C-----T-----A-A-----G-----C-----A-----	4023
H2A.GW.ALI Pol H2A.GW.ALI	A_V_H_V_A_S_G_Y_I_E_A_E_V_I_P_A_E_T_G_Q_F_T_A_Y_F_L_L_K_L_A_G_R_W_P_V_K_T_I_H_T_D_N_G	p31 4959
H2A.SN.HIV25T	-----G-C-C-GCAG-T-----AAAGAC-----CTC-C-A-----C-G-FA-T-----A-C-CACT-G-C-A-T-G-----TG	
H2A.DE.BEN	-----T-A-----T-----C-C-CAG-T-----AAG-C-----G-----CTC-C-AC-----C-G-CA-T-----A-C-CATT-G-C-A-----TG	4411
H2B.GH.D205	-----A-T-A-----TC-----C-C-CAG-T-----AAG-C-----G-----CTC-C-AC-----C-G-CA-T-----A-C-CGACT-G-C-A-----C	4966
H2B.CI.BHO	-----C-----G-T-----G-A-A-CCA-----G-----AAGAC-G-----TCTC-C-AC-----G-G-CA-C-----TA-C-C-CACC-C-A-----C-TG	4939
H2G.CI.ABT96	-----G-----G-T-----G-A-A-CCA-----G-----AAGAC-G-----TCTC-C-AC-----G-G-CA-C-----TA-C-C-CACC-C-A-----C-TG	4937
MAC.US.MM251	-----G-----G-T-----G-A-A-CCA-----G-----AAGAC-G-----TCTC-C-AC-----G-G-CA-C-----TA-C-C-CACC-C-A-----C-TG	4293
SNM.US.SIVSMNH9	-----A-----G-T-----C-----G-----C-----CA-----G-----TA-----C-----G-----TA-----C-----CA-----C-----A-----T-----TG	4859
STM.US.STM	-----A-----G-T-----C-----G-----C-----CA-----G-----TA-----C-----G-----TA-----C-----CA-----C-----A-----T-----TG	4356
TANTALUS.UG.TANI	-----A-----G-T-----C-----G-----C-----CA-----G-----TA-----C-----G-----TA-----C-----CA-----C-----A-----T-----TG	4530
GRIVET.ET.GRI-677	-----C-----G-----C-----G-----C-----AG-TAGG-----T-----A-----G-----A-----A-----C-----A-----CATTA-T-----G-----T-----C-----C-----C-----C-----C-----G-----T-----TG	4783
VERVET.KE.VBR-9063	-----A-----G-----C-----G-----C-----AG-TAGG-----T-----A-----G-----A-----A-----C-----A-----CATTA-T-----G-----T-----C-----C-----C-----C-----C-----G-----T-----TG	4587
VERVET.KE.AGMTYO	-----G-----C-----C-----G-----C-----AG-TAGG-----T-----A-----G-----A-----A-----C-----A-----CATTA-T-----G-----T-----C-----C-----C-----C-----C-----G-----T-----TG	4811
VERVET.KE.AGM155	-----G-----C-----C-----G-----C-----AG-TAGG-----T-----A-----G-----A-----A-----C-----A-----CATTA-T-----G-----T-----C-----C-----C-----C-----C-----G-----T-----TG	4298
VERVET.DE.AGM3	-----G-----C-----C-----G-----C-----AG-TAGG-----T-----A-----G-----A-----A-----C-----A-----CATTA-T-----G-----T-----C-----C-----C-----C-----C-----G-----T-----TG	4803
SABAEUS.SN.SAB1C	-----T-----T-----TC-----G-----T-----A-----TAG-----A-----A-----G-----A-----GA-----C-----T-----A-----TA-----C-----CA-----C-----C-----A-----T-----AC	4306
SUN.GA.SIVSUN	-----T-----T-----TC-----G-----T-----A-----TAG-----A-----A-----G-----A-----GA-----C-----T-----A-----TA-----C-----CA-----C-----C-----A-----T-----AC	4982
MANDRILL.GA.MNDGB1	-----G-----G-----TTCA-----G-----G-----T-----AC-A-CAAT-----GCA-----A-----AG-----CAC-----GG-----GC-----CTGT-----CC-----TG-----C-----COA-----TA-----T-----GT-----A-----C-----C-----A-----T-----GC	4944
LHOEST.KE.roest	-----G-----G-----TTCA-----G-----G-----T-----AC-A-CAAT-----GCA-----A-----AG-----CAC-----GG-----GC-----CTGT-----CC-----TG-----C-----COA-----TA-----T-----GT-----A-----C-----C-----A-----T-----GC	4271
LHOEST.CD.roest447	-----G-----G-----TTCA-----G-----G-----T-----AC-A-CAAT-----GCA-----A-----AG-----CAC-----GG-----GC-----CTGT-----CC-----TG-----C-----COA-----TA-----T-----GT-----A-----C-----C-----A-----T-----GC	4945
LHOEST.CD.roest448	-----G-----G-----TTCA-----G-----G-----T-----AC-A-CAAT-----GCA-----A-----AG-----CAC-----GG-----GC-----CTGT-----CC-----TG-----C-----COA-----TA-----T-----GT-----A-----C-----C-----A-----T-----GC	4945
LHOEST.CD.roest485	-----G-----G-----TTCA-----G-----G-----T-----AC-A-CAAT-----GCA-----A-----AG-----CAC-----GG-----GC-----CTGT-----CC-----TG-----C-----COA-----TA-----T-----GT-----A-----C-----C-----A-----T-----GC	3861
LHOEST.CD.roest524	-----G-----G-----TTCA-----G-----G-----T-----AC-A-CAAT-----GCA-----A-----AG-----CAC-----GG-----GC-----CTGT-----CC-----TG-----C-----COA-----TA-----T-----GT-----A-----C-----C-----A-----T-----GC	3861
RCM.NG.SIV-NGM	-----A-C-----G-----ATCA-----C-----TC-----G-----TT-----A-----T-----A-----AAAC-----A-----C-----C-----G-----C-----GTGCA-----C-----A-----C-----G-----A-----A-----AC-----	3858
COL....SIV-CGU1	-----A-C-----G-----ATCA-----C-----TC-----G-----TT-----A-----T-----A-----AAAC-----A-----C-----C-----G-----C-----GTGCA-----C-----A-----C-----G-----A-----A-----AC-----	4933
	-----A-C-----CTGT-C-CTG-TCTGTCTGG-----ACCA-AT-AAAGAG-----G-----A-----G-----G-----A-----GTAGGCC-----CA-----G-----TA-----TCAG-----GAG-----T-----G-----CA-----G-----T-----C-----A-----T-----GC	4347
SYKES.KE.SYK173	-----AA-CAC-----T-----A-CA-GA-CT-AAA-AGG-----G-----G-----T-----TG-----CA-GC-A-----CA-----T-----A-----S-----R-----W-----P-----I-----K-----Q-----I-----H-----T-----D-----N-----G-----	4784
SYKES.KE.SYK173 Pol	A_V_N_T_A_S_G_Y_I_E_T_K_I_L_K_R_E_T_G_D_E_T_A_L_F_L_M_Q_I_A_S_R_W_P_I_K_Q_I_H_T_D_N_G	p31

B.FR.HXB2 Pol	S_N_F_T_G_A_T_V_R_A_A_C_W_W_A_G_I_K_Q_E_F_G_I_P_Y_N_P_Q_S_Q_G_V_V_E_S_M_N_K_E_L_K_K_I_I	P31
B.FR.HXB2	GCAATTTACCGGTGCTACGGTTAGGCGCCCTGTTGGTGGCGGAATCAAGCAGGAATTTGGAATTCCTACAATCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAAAGAAATTAAGAAAATTTAT	4714
A.UG.U455	A-C-G-A-AA-A-TAAAT-C-A-G-C-G-C-C-C-G	4160
C.ET.ETH2220	A-AA-G-A-AA-A-A-T-TC-A-A-T-C-C	4106
D.CD.84ZR085	A-A-A-A-A-C-A-T-A-A	4236
F1.BF.VI850	A-CG-A-A-AA-AT-AAAT-CA	4054
G.SE.SE6165	GA-G-G-T-G-A-A-A-AT-C-A	4111
H.CF.90CF056	A-C-G-T-G-A-A-A-TG	4061
J.SE.SE7887	A-C-G-A-AA-A-AAATG-C-A	4028
K.CM.MP535C	A-G-A-A-AA-AAATG-C-A	3910
01.AE.TH.CM240	A-G-A-A-AA-AAATG-CA-A	4288
02.AG.NG.IBNG	A-A-G-A-A-T-A-G-C	4239
03.AB.RU.KAL153-2	A-A-G-A-A-A-AT-A-T	4759
04.CP.CY.94CY032-3	TA-C-T-AA-AA-AA-AAAT-A-T	4080
N.CM.MVP5180	T-AA-AG-CA-G-AA-T-A-A-A-C-AC-A-T-G-G-A-T	4308
O.BE.ANT70	T-AA-A-A-TA-G-A-T-A-CAAC-AC-A-T-G-A-A-T	4744
CPZ.US.CPZUS	T-AA-AG-A-A-T-A-T-C-A-C-AC-A-A-G-A-C	4768
CPZ.GA.CPZGAB	T-AA-G-A-C-A-T-AAAC	4774
CPZ.CM.CM3	T-AA-AG-A-A-AA-AT	4775
CPZ.CM.CM5	T-AA-AG-A-A-AA-AT	4128
CPZ.CD.CPZANT	T-AA-AG-A-A-AA-AT	4153
H2A.GW.ALI Pol	V_N_F_T_S_Q_E_V_K_M_V_A_W_V_V_G_I_E_Q_S_F_G_V_P_Y_N_P_Q_S_Q_G_V_V_E_A_M_N_H_L_K_N_Q_I	P31
H2A.GW.ALI	TCACAGGAA-A-A-ATG-TAGCA-TT-T-AG-ATCC-G-A-T-G-A-T-A-C-C-R-G-G-A-G-A-M-N-H-L-K-N-Q-I	5089
H2A.SN.HIV25T	TTCACAGGAA-A-A-ATG-TGGCA-ATA-T-AG-A-ATCC-C-G-A-T	4541
H2A.DE.BEN	TTCACAGGAA-G-A-ATG-TGGCA-TA-T-AG-A-ATCC-G-A-T	5096
H2B.GH.D205	TCAC-A-GT-A-A-ATG-TAGCC-TA-AG-A-AACT-G-A-T	5069
H2B.CI.BHO	TTCACAGAT-G-AAATG-AGCC-ATA-G-AG-A-AAC-C-G-G-T	5067
H2G.CI.BHT96	TCCGAGA-G-A-NTG-TAGCC-TA-C-G-A-C-G-A-A	4423
MAC.US.MM251	TCCGAGA-A-A-ATG-TTGC-A-R-TG-A-ACC-R-GG-A-T-T	4989
SMW.US.SIVSMW9	ATCACAGAA-G-AAATG-TTGC-A-R-TG-A-ACC-R-GG-A-T-T	4486
STM.US.STM	TCCGAGA-A-A-ATG-TAGCC-GACA-G-AG-A-AAAC-G-A-T	4660
TANTALUS.UG.TANI	TGTTA-CAGGA-ACAA-TATT-GACA-G-AG-A-AAAC-G-A-T	4913
GRIVET.ET.GRI-677	TC-CAG-AT-GGCA-G-TG-C-GTAA-AG-CACCAC-A-T-T	4717
VERVET.KE.VBR-9063	T-TCCCAAGA-AGC-AAT-GAAG-AG-CACCAC-A-T-T	4941
VERVET.KE.AGMTYO	T-TCCCAAGA-GGCA-AATA-GAAA-TG-A-TAC-ACA-T-A-A-T	4428
VERVET.KE.AGMT55	T-TCCCAAGA-AGCT-TATG-C-GAAA-AG-A-CAC-ACG-TG-A-T	4933
VERVET.DE.AGM3	T-GTC-CAAGAA-AGCA-AATG-GAAA-AG-A-CAC-ACA-TG-A-T	4436
SABAEUS.SN.SABIC	T-TA-CCAGAA-AGCA-TATT-C-GAAA-AG-CAC-ACA-TG-A-T	5112
SUN.GA.SIVSUN	T-TA-AAAGAT-GGA-A-TA-TTA-C-AC-TAC-ACA-G-A-T	5074
MANDILL.GA.MNDGB1	T-TA-CAGAA-RGAAA-ATG-TTA-G-AG-A-CAC-ACA-G-A-T	4401
LHOEST.KE.roest	TGTAA-AAAGAA-ACAA-A-TAAT-ATA-G-AG-A-TAC-ACA-G-A-T	5075
LHOEST.CD.roest447	TGT-A-CAAGAA-ACAA-TAAT-ATA-TG-A-TAC-ACA-G-A-T	3991
LHOEST.CD.roest485	TGT-A-CAAGAA-ACAA-A-TAACC-ATA-TG-A-TAC-ACA-G-A-T	3988
LHOEST.CD.roest524	TGTAA-AAAGAT-ACA-A-TGACC-TTAA-T-AG-CAC-ACA-G-A-T	5063
RCM.NG.SIV-NGM	T-AA-AG-A-GCAA-A-A-TCA-AG-CACC-GA-A-A-T	4477
COL.SIV-CGU1	TGTAA-CAGCACT-A-A-AGTG-CTA-AGCA-CACCACA-CAC-T-TCA	4477
SYKES.KE.SYK173	TGT-A-AC-A-T-C-AA-A-A-TGT-C-C-AG-A-CACTACA-T	4914
SYKES.KE.SYK173 Pol	TGT-C-C-AG-A-CACTACA-T-A-A-C-C-A-C-C-G-A-A-GGT-T	P31

HIV-1/HIV-2/SIV complete genomes

B.FR.HXB2 Pol 4974
B.FR.HXB2 4974
A.UG.U455 4420
C.ET.ETH2220 4366
D.CD.84ZR085 4496
F1.BF.V1850 4314
G.SE.SE6165 4371
H.CF.90CF056 4321
J.SE.SE7887 4288
K.CM.MP535C 4170
O1.AE.TH.CM240 4170
02.AG.NG.IBNG 4548
03.AB.RU.KAL153-2 4499
04.CPX.CV.94CYO32-3 5019
N.CM.MVP5180 4340
O.BE.ANT70 4568
CPZ.US.CPZUS 5004
CPZ.GA.CPZGAB 5028
CPZ.CM.CM3 5034
CPZ.CM.CM5 5035
CPZ.CD.CPZANT 4388
H2A.GW.ALI Pol 4413
H2A.GW.ALI 5349
H2A.SN.HIV2ST 4801
H2A.DE.BEN 5356
H2B.GH.D205 5329
H2B.CI.BHO 5327
H2G.CI.ABT96 4683
MAC.US.MM251 5249
SMW.US.SIVSMW9 4920
STM.US.STM 4746
TANTALUS.UG.TANI 5173
GRIVET.ET.GRI-677 4977
VERVET.KE.VBR-9063 5201
VERVET.KE.AGMTYO 4688
VERVET.KE.AGM155 5193
VERVET.DE.AGM3 4696
SABAEUS.SN.SAB1C 5372
SUN.GA.SIVSUN 5334
MANDRILL.GA.MNDGB1 4661
LHOEST.KE.hoest 5335
LHOEST.CD.hoest447 4251
LHOEST.CD.hoest485 4251
LHOEST.CD.hoest524 4248
RCM.NG.SIV-NGM 5323
COL.....SIV-CGU1 4737
SYKES.KE.SYK1173 5168
SYKES.KE.SYK1173 Pol1 5168
SYKES.KE.SYK1173 Pol 5168

HIV-1/HIV-2/SIV complete genomes

HIV-1/HIV-2/SIV
complete genomes

B.FR.HXB2 Pol	V_V_I_Q_D_N_S...D_I_K_V_P_R_R_K_A_K_I_I_R_D_Y--K_Q_M_A--D_D_C_V_A_S_R_Q_D_E_D_-\$	Pol p31 integrase end <-	p31
B.FR.HXB2 Vif	TAGTAATACAGATAAATAGT...GACATAAAAGTAGTGCCAAAGAAAGAAAAGCAAAAGATCATTAGGGATTATGGGAGATGATGATGTTGGCAAGTAGACACAGGATGAGGATTAG....	M_E_N_R_W_Q_V_M_I_V_W_O_V_D_R_M_R_I_R_-\$ -> Vif start	Vif 5096
A.UG.U455	-----C-----T-----G-----A-----A-----G-----G-----A-----A-----	-----A-----G-----	4542
C.ET.FTH2220	-----C-----C-----G-----A-----A-----G-----G-----A-----A-----	-----A-----A-----	4488
F.I.BE.VI1850	-----C-----AC-----A-----G-----A-----A-----G-----G-----A-----A-----	-----T-----G-----	4618
G.SE.SE6165	-----C-----C-----A-----A-----A-----A-----G-----G-----A-----A-----	-----A-----A-----	4436
H.CF.90CF056	-----G-----C-----A-----A-----A-----A-----G-----G-----A-----A-----	-----A-----A-----	4443
J.SE.SE7887	-----C-----C-----T-----G-----A-----A-----A-----A-----G-----G-----A-----A-----	-----A-----A-----	4410
K.CM.MP535C	-----G-----C-----T-----G-----A-----A-----A-----A-----G-----G-----A-----A-----	-----A-----A-----	4292
01.AE.TH.CM240	-----G-----C-----AC-----T-----G-----A-----A-----A-----A-----G-----G-----A-----A-----	-----A-----A-----	4670
02.AB.RU.KAL153-2	-----G-----C-----C-----C-----T-----G-----A-----A-----A-----A-----G-----G-----A-----A-----	-----A-----A-----	4621
04.CPX.CY.94CY032-3	-----G-----T-----CG-----G-----T-----C-----C-----C-----T-----G-----A-----A-----A-----A-----G-----G-----A-----A-----	-----A-----A-----	5141
N.CM.YBP30	-----C-----AG-----A-----T-----G-----A-----A-----A-----A-----G-----G-----A-----A-----	-----G-----A-----C-----GAAAT	4695
O.CM.MVP5180	-----C-----GG-----A-----T-----G-----A-----A-----A-----A-----G-----G-----A-----A-----	-----AC-----A-----A-----A-----A-----ACA--AAG-G-AAGCAT	5131
I.BE.AN170	-----CA-----G-----A-----CGAG-----T-----G-----A-----A-----A-----A-----G-----G-----A-----A-----	-----AC-----A-----A-----A-----A-----ACA--AAG-G-AAGGT	5155
CPZ.US.CPZUS	-----G-----C-----G-----GAGAG-----T-----G-----A-----A-----A-----A-----G-----G-----A-----A-----	-----A-----A-----A-----A-----AG--AGCTT	5162
CPZ.CM.CAM3	-----G-----TA-----G-----CGAG-----T-----G-----A-----A-----A-----A-----G-----G-----A-----A-----	-----T-----A-----A-----A-----A-----AG-C-AGGCT	4515
CPZ.CM.CAM5	-----C-----CA-----CC-AGAG-----A-----T-----G-----C-----TC-----G-----T-----A-----A-----G-----G-----A-----A-----	-----A-----A-----A-----A-----C-AAACT	4803
CPZ.CD.CPZANT	-----G-----I-----V-----K-----V-----G-----T-----D-----I-----K-----V-----P-----R-----R-----K-----A-----K-----I-----I-----R-----D-----Y-----G-----R-----Q-----E-----L-----D-----S-----G-----P-----H-----L-----E-----G-----A-----R-----E-----D-----G-----	-----A-----A-----A-----A-----C-ATACT	4540
H2A.GW.ALI Pol	-----CA-----G-----TA-----G-----TAGGG-----CA-----A-----A-----A-----A-----G-----G-----C-----A-----C-----	-----G-----G-----C-----C-----A-----A-----GG-AG-CAA-AGCTG--AG-G--CCCCACTTGGAGGT-CCAG--GG-T...GG	p31 Vif 5473
H2A.GW.ALI Vif	-----CA-----G-----CA-----G-----TAGGG-----C-----A-----A-----A-----A-----G-----G-----C-----T-----C-----	-----G-----G-----C-----C-----A-----A-----GG-AG-CAA-AGATG--AGCG--TCCCACCTGGAGGT-CCAG--GG-T...GG	4925
H2A.SN.HIV25T	-----CA-----C-----A-----G-----TAGGG-----CA-----A-----A-----A-----A-----G-----G-----C-----T-----C-----	-----G-----G-----C-----C-----A-----A-----GG-AGCA--A-CTG--AG-A--CCCCACCTGGAGGT-CCAG--GG-T...GG	5480
H2A.DE.BEN	-----CA-----C-----A-----G-----TAGGG-----CA-----A-----C-----C-----A-----A-----G-----G-----C-----T-----C-----	-----G-----G-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	5453
H2B.GH.D205	-----CA-----C-----A-----G-----TAGGG-----CA-----A-----C-----C-----A-----A-----G-----G-----C-----T-----C-----	-----G-----G-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	5451
H2B.CI.EHO	-----CA-----C-----A-----G-----TAGGG-----CA-----A-----C-----C-----A-----A-----G-----G-----C-----T-----C-----	-----G-----G-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	4807
H2G.CI.ABT96	-----CA-----CT-----A-----G-----TAGGG-----CA-----A-----C-----T-----A-----A-----G-----G-----C-----T-----C-----	-----G-----G-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	5373
MAC.US.MW251	-----CA-----C-----A-----G-----TAGGG-----CA-----A-----C-----C-----A-----A-----G-----G-----C-----T-----C-----	-----G-----G-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	4870
SNM.US.SIVSMH9	-----CA-----CC-----A-----G-----TAGGG-----CA-----A-----C-----C-----A-----A-----G-----G-----C-----T-----C-----	-----G-----G-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	5044
STM.US.STM	-----C-----A-----A-----G-----TAGGG-----CA-----A-----G-----G-----A-----A-----G-----G-----C-----T-----C-----	-----T-----G-----A-----A-----G-----G-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	5286
TANTALUS.UG.ITANI	-----G-----T-----A-----AGG-----GAG-----G-----T-----G-----A-----A-----G-----G-----C-----T-----C-----	-----G-----G-----A-----A-----G-----G-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	5104
GRIVET.FT.GRI-677	-----G-----C-----CA-----AGGAGAA-----G-----C-----G-----A-----T-----C-----C-----	-----G-----G-----A-----A-----G-----G-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	5328
VERVET.KE.VER-9063	-----G-----CC-----CA-----G-----CGGA-----A-----C-----C-----A-----A-----G-----G-----C-----T-----C-----	-----A-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	4815
VERVET.KE.AGM150	-----G-----TC-----CA-----G-----AGG-----GAA-----A-----C-----C-----A-----A-----G-----G-----C-----T-----C-----	-----A-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	5320
VERVET.KE.AGM155	-----G-----C-----CA-----GAGG-----GTG-----A-----T-----A-----A-----A-----A-----G-----G-----C-----T-----C-----	-----A-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	4823
SABAEUS.SN.SABLC	-----G-----T-----AC-----GG-----A-----AT-----G-----AC-----A-----C-----T-----G-----A-----A-----G-----G-----C-----T-----C-----	-----A-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	5493
SUN.GA.SIVSUN	-----G-----C-----CA-----C-----GGG-----GAG-----A-----C-----CTG-----A-----T-----G-----A-----A-----G-----G-----C-----T-----C-----	-----A-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	4774
MANDRIL.GA.MNDGB1	-----G-----TT-----A-----T-----C-----AGAA-----G-----G-----G-----G-----A-----A-----A-----A-----G-----G-----C-----T-----C-----	-----A-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	5438
LHOEST.KE.hoest	-----G-----GG-----A-----AGGGGAG-----A-----C-----TTCTC-----C-----C-----C-----G-----G-----G-----G-----G-----G-----C-----T-----C-----	-----G-----G-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	5448
LHOEST.CD.hoest447	-----G-----GG-----A-----AGGAGAC-----A-----C-----TTCTC-----C-----C-----C-----G-----G-----G-----G-----G-----G-----C-----T-----C-----	-----G-----G-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	4364
LHOEST.CD.hoest448	-----G-----GG-----A-----AGGAGAG-----A-----AC-----TTCTC-----C-----C-----C-----G-----G-----G-----G-----G-----G-----C-----T-----C-----	-----G-----G-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	4364
LHOEST.CD.hoest524	-----G-----GG-----A-----G-----GGAGAG-----A-----AT-----TCTCTA-----C-----C-----G-----G-----G-----G-----G-----G-----C-----T-----C-----	-----G-----G-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	4361
RCM.NG.SIV-NGM	-----C-----G-----C-----G-----C-----AGAG-----TC-----G-----G-----C-----T-----G-----A-----A-----G-----G-----C-----T-----C-----	-----G-----G-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	5448
COL....SIV-CGU1	-----C-----C-----CAG-----A-----TC-CAAGGTAC-----T-----GTTT-----GAAA-----T-----CC-ACITCC-ATATGG-G-AGAT-TG--CAG--GAA-T--TT-TCC-ATGG-C-AA--AGCGAATAC	-----G-----G-----C-----C-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	4867
SYKES.KE.SYK173	-----GG-----G-----ACAG-GGAGGA-----TTT-----C-----A-----A-----CA-----CC-----G-----G-----C-----T-----C-----	-----A-----A-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	5292
SYKES.KE.SYK173 Pol	V_V_V_Q_I_T_E_F_G_D_I_F_A_V_P_R_R_K_A_K_I_I_T_D_H_G_E...R_M_D_S_G_S_H_V_E_N_D_P_K_I_D_-\$	-----A-----A-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	p31
SYKES.KE.SYK173 Vif	M_E_K...E_W_I_V_V_P_T_W_R_M_T_P_R_Q_I_-\$	-----A-----A-----A-----A-----GG-GGA-AA-G-TTG--TG-A--CC-AC-TGGAGG-TACCAG-C-GGCT...AG	Vif

Table with columns for isolate names (e.g., B.FR.HXB2 Vif, A.UG.U455, H2A.GW.ALI Pol), genomic coordinates (4570-5318), and nucleotide sequences. Includes a partial sequence at the bottom for H2A.GW.ALI Vif.

B.FR.HXB2 Vif	T T Y W G L H T G E R D W H	L G Q G V S I E W R K K R Y S T Q V D P E L A D	Vif
B.FR.HXB2	CAACATATTGGGTCTGCATACACAGAAAGACTGGCAT	TTGGFTCAGGAGTCTCCATAGAATGGAGAAAAGAGATATAGCACACAAAGTAGACCTGAACTAGCAGA	5351
A.UG.U455	G-----A-----A-----C-----T-----G-----	-----CTG-----A-----T-----C-----	4797
C.ET.ETH2220	T-----T-----A-----T-----A-----	-----ATTG-----G-----C-----A-----	4743
D.CD.84ZR085	A-----A-----A-----A-----A-----	-----G-----G-----G-----G-----	4873
F1.BF.VI850	G-----G-----CC-----A-----A-----	-----C-----CGG-----AG-----G-----A-----	4691
G.SE.SE6165	G-----G-----A-----A-----A-----A-----	-----T-----G-----C-----CGA-----A-----G-----	4748
H.CF.90CF056	C-----C-----A-----A-----T-----C-----	-----ACTG-----A-----G-----G-----GC-----G-----	4698
J.SE.SE7887	T-----T-----G-----AG-----C-----C-----	-----CG-----GA-----G-----C-----A-----GC-----G-----	4865
K.CM.MP535C	G-----G-----A-----AG-----C-----C-----	-----CTG-----A-----G-----C-----A-----A-----C-----	4547
O1.AE.TH.CM240	C-----C-----C-----C-----C-----	-----C-----G-----GA-----A-----A-----T-----C-----	4925
02.AG.NG.IBNG	A-----A-----A-----A-----A-----	-----A-----C-----G-----A-----A-----A-----CT-----	4876
03.AB.RU.KAL153-2	A-----A-----GC-----G-----CA-----	-----C-----CTC-----GA-----G-----T-----C-----G-----	5396
04.CPX.CV.94CY032-3	T-----T-----G-----AA-----C-----C-----	-----ACT-----GA-----A-----C-----AG-----T-----A-----G-----	4717
N.CM.MVP5180	C-----C-----AT-----AATGC-----G-----GGAA-----C-----	-----CAAT-----C-----GAG-----AA-----GA-----T-----T-----AC-----	4945
O.BE.ANT70	C-----C-----AT-----ATGC-----G-----TGAA-----	-----CT-----TAGT-----A-----AA-----AA-----GA-----T-----AC-----	5381
CPZ.US.CPZUS	T-----G-----AT-----A-----CA-----C-----	-----AC-----GGA-----CT-----C-----AG-----A-----T-----GA-----G-----	5405
CPZ.GA.CPZGAB	C-----C-----CCT-----AAGCC-----	-----CT-----GGATCT-----GT-----	5411
CPZ.CM.CM3	C-----C-----AT-----AATGC-----	-----C-----G-----T-----G-----	5415
CPZ.CM.CM5	A-----A-----ACA-----GC-----CAT-----C-----	-----A-----TG-----TC-----C-----AA-----T-----AC-----G-----	4765
CPZ.CD.CPZANT	G-----G-----TG-----TC-----AT-----A-----ATGT-----	-----A-----A-----T-----GAG-----AG-----T-----A-----AC-----T-----	4796
H2A.GW.ALI Vif	Q A Y W N L L T P E K G W L S S Y A V	R I T W Y T E K F W T D V T P D C A D	Vif
H2A.GW.ALI	AGG-----AAC-----A-----CC-----A-----GA-----	-----TCTCCTCTTATGCAGTA-----	5732
H2A.SN.HIV2ST	AGG-----C-----AAC-----A-----CC-----	-----A-----AGAC-----ACT-----TAT-----C-----G-----A-----A-----TCT-----G-----G-----T-----TAC-----A-----CTGT-----G-----	5184
H2A.DE.BEN	AAC-----A-----A-----A-----CC-----	-----A-----AGAT-----ACC-----TAT-----C-----G-----A-----AG-----TCT-----G-----G-----T-----TAC-----A-----CTGT-----	5739
H2B.GH.D205	A-----G-----G-----A-----CC-----	-----G-----AGAC-----AC-----TAT-----G-----G-----AG-----T-----TAT-----G-----T-----ACA-----TG-----G-----	5712
H2B.CI.BHO	A-----G-----C-----AAC-----	-----CC-----AGAC-----AC-----TAT-----G-----G-----C-----T-----TAT-----G-----T-----ACT-----TG-----G-----	5710
H2G.CI.BHT96	A-----GGG-----AACT-----	-----CCCT-----AG-----GA-----TTAAGCCAGTATGCAGTA-----	5066
MAC.US.MM251	A-----GGG-----CA-----T-----CC-----	-----TAC-----TAC-----G-----A-----T-----TAT-----G-----T-----GAG-----A-----GAC-----	5632
SMM.US.SIVSMMH9	A-----GGG-----AA-----T-----A-----CC-----	-----GRA-----GG-----TCAGTACTAATGCAAGT-----	5129
STM.US.STM	A-----GGG-----AA-----T-----A-----CC-----	-----G-----AGG-----ACC-----TACTCC-----GA-----AT-----TCT-----G-----G-----T-----C-----ACA-----A-----TTAT-----	5303
TANTALUS.UG.TANI	ATTGT-----C-----CAC-----A-----CCC-----	-----CAG-----GA-----TCTAGTAATGCTGTG-----	5559
GRIVET.ET.GRI-677	AT-----TT-----CACAA-----C-----CCC-----	-----G-----AGT-----TTGAG-----T-----CAA-----C-----CG-----C-----AA-----G-----ACT-----TG-----G-----G-----	5360
VERVET.KE.VBR-9063	ACTGC-----CA-----T-----A-----CC-----	-----CA-----A-----T-----G-----GCAGTAG-----	5587
VERVET.KE.AGMTYO	ACTCT-----CA-----T-----A-----G-----TCC-----	-----A-----CA-----A-----ATG-----AGCAATTTAG-----TCT-----A-----A-----G-----T-----CA-----TAC-----	5077
VERVET.KE.AGM155	ATCT-----CA-----A-----CA-----G-----AG-----	-----GA-----TATCAACTATGCTAGTAA-----	5582
VERVET.DE.AGM3	T-----AT-----CA-----T-----A-----CCC-----	-----G-----G-----A-----G-----T-----GG-----GA-----GGTT-----G-----GC-----T-----TGGAC-----G-----	5085
SABAEUS.SN.SAB1C	G-----GC-----G-----ACAC-----C-----CC-----	-----C-----A-----GG-----T-----AT-----RACCAGTGGGATCT-----	5752
SUN.GA.SIVSUN	G-----GC-----CA-----T-----G-----CA-----AT-----	-----GGT-----GA-----CTTAGTACTTAGCAGTAAGCA-----A-----AAGCA-----T-----TAGACC-----CCCTTTG-----G-----AGAG-----GGT-----T-----GA-----ACT-----A-----GG-----	5071
MANDRILL.GA.MNDGB1	-----I-----GKCA-----T-----G-----C-----G-----G-----	-----A-----A-----A-----A-----GTCI-----T-----GAGAC-----CAA-----G-----A-----T-----CAIT-----	5706
LHOEST.KE.roest	-----TGC-----GKCACT-----	-----G-----G-----GA-----ATCTCAGTGGGCAGTA-----	4622
LHOEST.CD.roest447	-----TGC-----GKCACT-----	-----G-----G-----GA-----ATCTCAGTGGGCAGTA-----	4622
LHOEST.CD.roest485	-----TT-----GCTCA-----	-----G-----C-----GG-----GA-----AT-----TCACAGTGGGCAGTA-----	4619
LHOEST.CD.roest524	-----A-----A-----C-----AT-----	-----AG-----GA-----TTAGAACCTATGGATAGGAA-----	5721
RCM.NG.SIV-NGM	-----TTATG-----A-----TACAG-----AGG-----AGT-----G-----	-----C-----T-----TAAGCTGGCAACAAGCAGTAGGG-----AC-----	5112
COL.SIV-CGU1	-----TTATG-----A-----TACAG-----AGG-----AGT-----G-----	-----C-----T-----TAAGCTGGCAACAAGCAGTAGGG-----AC-----	5112
SYKES.KE.SYK173	-----FTCC-----CACAA-----T-----G-----CCC-----	-----GA-----TTACATATGCAAGGAAAT-----	5548
SYKES.KE.SYK173 Vif	-----F-----Y-----H-----N-----L-----A-----P-----E-----R-----G-----W-----L-----H-----M-----Q-----G-----I-----	-----R-----I-----Q-----Y-----Q-----W-----N-----O-----W-----N-----T-----D-----L-----T-----P-----A-----V-----A-----D-----	Vif

HIV-1/HIV-2/SIV
complete genomes

B.FR.HXB2	Vif	5463	5463
B.FR.HXB2	Vif	5463	5463
A.UG.U455	Vif	4909	4909
C.ET.ETH2220	Vif	4855	4855
CD.84ZRO85	Vif	4985	4985
F1.BF.VI850	Vif	4803	4803
G.SE.SE6165	Vif	4860	4860
H.CF.90CF056	Vif	4777	4777
J.SE.SE7887	Vif	4659	4659
K.CM.MP535C	Vif	5037	5037
01.AE.TH.CM240	Vif	4988	4988
02.AG.NG.IBNG	Vif	5508	5508
03.AB.RU.KAL153-2	Vif	4829	4829
04.CX.CY.94CY032-3	Vif	5057	5057
N.CM.VBP5180	Vif	5493	5493
O.BE.ANT70	Vif	5517	5517
CPZ.US.CPZUS	Vif	5523	5523
CPZ.GA.CPZGAB	Vif	5527	5527
CPZ.CM.CM3	Vif	4877	4877
CPZ.CM.CM5	Vif	5165	5165
CPZ.CD.CPZANT	Vif	4908	4908
H2A.GW.ALI	Vif	5847	5847
H2A.GW.ALI	Vif	5847	5847
H2A.SN.HIV2ST	Vif	5299	5299
H2A.DE.BEN	Vif	5854	5854
H2B.GH.D205	Vif	5827	5827
H2B.CI.BHO	Vif	5825	5825
H2G.CI.ABT96	Vif	5181	5181
MAC.US.MM251	Vif	5747	5747
SNM.US.SIVSMH9	Vif	5244	5244
STM.US.STM	Vif	5418	5418
TANTALUS.UG.TANI	Vif	5686	5686
GRIVET.ET.GRI-677	Vif	5472	5472
VERVET.KE.VBR-9063	Vif	5705	5705
VERVET.KE.AGMTYO	Vif	5195	5195
VERVET.KE.AGM155	Vif	5700	5700
VERVET.DE.AGM3	Vif	5203	5203
SABAEUS.SN.SABIC	Vif	5870	5870
SUN.GA.SIVSUN	Vif	5871	5871
MANDILL.GA.MNDGB1	Vif	5189	5189
LHOEST.KE.ROEST	Vif	5818	5818
LHOEST.CD.ROEST447	Vif	4737	4737
LHOEST.CD.ROEST485	Vif	4737	4737
LHOEST.CD.ROEST524	Vif	4734	4734
RCM.NG.SIV-NGM	Vif	5833	5833
COL. SIV-CGU1	Vif	5217	5217
SYKES.KE.SYK173	Vif	5657	5657
SYKES.KE.SYK173	Vif	5657	5657

HIV-1/HIV-2/SIV complete genomes

O L I H L Y Y F D C F S D S A I R K A L L G H I V S P R C E Y O A G H NK
 CCAACTAATTCATCTGTTACTTTTTCAGACTCTGCTGTATAAGAAAGCCCTATTAGGACACATAGTTCAGCCCTAGGCTGTAATATCAAGCAGGACATAAC.....AAAG
 ---T---C---C---T---T---A---A---C---G---GA---A---A---A---A---T---T---G---A---C---T---T---G---A---T---A---A---A---A---A---A---T---
 ---C---A---C---T---G---A---A---C---G---A---A---A---T---T---G---T---T---G---A---C---T---T---G---A---T---T---G---A---T---A---A---A---A---A---A---T---
 ---A---A---C---A---A---A---A---C---G---A---A---A---T---T---G---T---T---G---A---C---T---T---G---A---T---T---G---A---T---T---G---A---T---A---A---A---A---A---A---T---
 ---G---C---A---A---A---A---C---C---G---A---A---A---T---T---G---A---C---T---T---G---A---T---T---G---A---T---T---G---A---T---T---G---A---T---A---A---A---A---A---A---T---
 ---T---G---C---A---A---A---A---C---G---A---A---A---T---T---G---A---C---T---T---G---A---T---T---G---A---T---T---G---A---T---T---G---A---T---A---A---A---A---A---A---T---
 ---A---A---G---A---A---A---A---C---G---A---A---A---T---T---G---A---C---T---T---G---A---T---T---G---A---T---T---G---A---T---T---G---A---T---A---A---A---A---A---A---T---
 ---A---A---G---A---A---A---A---C---G---A---A---A---T---T---G---A---C---T---T---G---A---T---T---G---A---T---T---G---A---T---T---G---A---T---A---A---A---A---A---A---T---
 ---A---G---A---A---A---A---A---A---C---G---A---A---A---T---T---G---A---C---T---T---G---A---T---T---G---A---T---T---G---A---T---T---G---A---T---A---A---A---A---A---A---T---
 ---AGGA---G---A---A---A---A---A---A---C---G---A---A---A---T---T---G---A---C---T---T---G---A---T---T---G---A---T---T---G---A---T---T---G---A---T---A---A---A---A---A---A---T---
 ---AGGA---G---A---A---A---A---A---A---C---G---A---A---A---T---T---G---A---C---T---T---G---A---T---T---G---A---T---T---G---A---T---T---G---A---T---A---A---A---A---A---A---T---
 ---A---G---A---A---A---A---A---A---C---G---A---A---A---T---T---G---A---C---T---T---G---A---T---T---G---A---T---T---G---A---T---T---G---A---T---A---A---A---A---A---A---T---
 TAG---T---G---C---CTCTC---A---T---T---G---T---T---A---A---C---G---A---A---A---T---T---G---A---C---T---T---G---A---T---T---G---A---T---T---G---A---T---A---A---A---A---A---A---T---
 TAG---G---A---A---A---A---A---A---C---G---A---A---A---T---T---G---A---C---T---T---G---A---T---T---G---A---T---T---G---A---T---T---G---A---T---A---A---A---A---A---A---T---
 TAG---G---A---A---A---A---A---A---C---G---A---A---A---T---T---G---A---C---T---T---G---A---T---T---G---A---T---T---G---A---T---T---G---A---T---A---A---A---A---A---A---T---
 T---A---G---A---A---C---AA---TCRAA---AG---C---C---A---AA---CC---G---A---G---A---ATTGA---ATAC---CC---C---CA---GAAG---GT.....GT.....C---CA
 ---S---L---I---H---G---T---Y---F---S---C---F---T---A---G---E---V---R---R---A---I---R---G---E---K---L---L---S---C---C---N---Y---P---Q---A---H---K---S
 ---TCC---A---GGCAC---T---CTCT---C---A---G---CAGG---AAG---GA---A---CAG---GG---A---AGT---ATTGT---CT---C---CA---T---T---CCCAG---CC---GTCA.....C---Q
 ---TCC---A---AGCAC---T---CTCT---C---A---G---CAGG---AAG---GA---A---CAG---GG---A---AGT---ATTAT---CT---C---CA---C---C---C---CA---C---GTAC.....GTAC
 ---ATC---A---AGCAC---T---CTCT---C---A---G---CAGG---AAG---GA---A---CAG---GG---A---AGT---ATTGT---CT---C---CA---C---C---C---CA---C---GTAC.....GTAC
 ---G---C---G---GG---C---T---CTCT---C---AA---AAG---G---GA---A---CAGG---G---A---AGA---ATTGT---CTAC---CA---C---C---T---CT---CG---AGGG.....CG-AGGG
 T---T---T---G---GG---C---T---CTCT---C---A---CTAA---AAG---G---GA---A---CAGG---G---A---AGA---ATTGT---CTAC---CA---C---C---T---CT---CG---AGGG.....CG-AGGG
 ---ATTI---C---G---AGCAC---T---CCCT---C---A---CGGA---AAG---G---A---CAGG---G---A---AGA---GCTGT---T---C---CA---G---TC---CGAG---CT---GGTAC.....GTGAC
 ---ACTT---C---G---GGCAC---T---CCCT---C---A---CGGA---AAG---G---A---CAGG---G---A---AGA---GCTGT---T---C---CA---G---TC---CGAG---CT---GGTAC.....GTGAC
 ---T---G---C---G---GGCAC---T---CCCT---C---A---CGGA---AAG---G---A---CAGG---G---A---AGA---GCTGT---T---C---CA---G---TC---CGAG---CT---GGTAC.....GTGAC
 ---TA---G---A---G---A---A---A---A---TAGA---C---CCA---C---TAGA---C---G---AGT---ACTGT---T---C---CAGG---TCAGAA---CT---GAAAC.....GAAAC
 ---AG---A---A---C---A---A---A---A---TAGA---C---CCA---C---TAGA---C---G---AGT---ACTGT---T---C---CAGG---TCAGAA---CT---GAAAC.....GAAAC
 TGG---A---G---A---C---AC---T---TCA---A---AAGG---C---TCA---A---TA---AGG---G---G---G---T---CGT---TTC.....CAC---TCA---G---A---A.....A
 TAGTA---A---AGTGC---T---A---TA---AA---AAGA---C---CCA---C---A---TC---GAGG---C---G---G---T---GTT---C.....C---G---T---C---A---G---AACCACA.....AACCACA
 ---TCCA---G---A---TGTTC---T---AC---TAGA---C---CCA---C---AC---AGG---A---GGT---C---T---TTC.....C---G---T---C---A---G---AACCACA.....AACCACA
 TAGC---G---A---ACCC---T---ACT---A---AAGG---C---C---G---A---C---G---GGT---C---T---TTC.....C---G---G---T---C---C---AG---C---GAAAAACA.....GAAAAACA
 ---TA---T---TG---OC---T---CC---A---TAGA---G---CCA---C---A---CAGG---G---A---AGTA---CTTTGG.....C---G---G---T---C---C---AG---C---GAAAAACA.....GAAAAACA
 T---GA---G---C---CT---TA---G---T---T---GGA---CTCTAGTG---C---GA---TA---G---GG---A---A---A---ACTTAAAGTT---AG---T---C---C---AGCC.....CAAGGT
 T---A---G---C---CTATA---A---T---A---C---CATG---AG---AG---T---TG---A---A---CAGG---G---AGA---GTTAAAGTT---AG---T---C---C---AGCC.....CAAGGT
 T---A---G---A---CTGTA---A---T---A---C---TAAAGG---AG---T---G---G---A---G---A---A---CTATGT---T---CC---G---...---C---TTT---T---C---AGGT.....AGGT
 T---A---G---A---CTGTA---A---T---A---C---TAAAGG---A---T---G---G---G---A---AGA---ATGG---ATTC---CTATG---CCA---T---C---CTATG---CCA---T---AGGT.....AGGT
 T---A---GC---G---A---G---T---GGTGC---TAGG---AC---GTTT---G---A---AGG---G---A---AGA---CTG---AC---C---A---ATG---CTA---T---AGGT.....AGGT
 TA---T---G---A---AAATC---G---T---CCT---T---A---TAA---A---T---C---A---A---C---G---GCA---GCTG---T---...---GG---T---G---GAA---GAGAT.....GAGAT
 ...A---GCACATCA---AAGGGGAAAGTGTGTACGCAT---GGAA---AAGC---A---TA---AGG---CTTACCCTGG---AT---...---T---TC---G---T---CCTA.....CCTA
 ---GC---T---A---AAAT---T---CCCT---C---A---CAAAGA---GG---C---ACC---AG---CAG---GG---AC---T---G---CAT---ACAC---CTGACC---C---...---CTGAC
 R---L---I---H---N---F---Y---F---P---C---F---T---A---R---A---V---N---Q---A---V---R---G---E---L---L---T---S---H---C---W---T---P---...---H---T---D
 ---R---L---I---H---N---F---Y---F---P---C---F---T---A---R---A---V---N---Q---A---V---R---G---E---L---L---T---S---H---C---W---T---P---...---H---T---D

B.FR.HXB2 Vif	Q K T K G H R G S H T M N G H S	Vif end in HIV-1 <-	Vif
B.FR.HXB2 Vpr	P E D Q G G P Q R E P H N E W T L E L		Vpr
B.FR.HXB2	CCAGAAGACCAGGGCCACAGAGGGACACACAAATGAATGGACACTAGAGCT		5623
A.UG.U455G-----G-----GT--GC-----A-G-T-----T-		5069
C.ET.ETH2220G-----G-----G-----G-----G-----G-----		5015
D.CD.84ZF085G-C-----G-----A-----T-----T-----T-----		5145
F1.BE.VI850G-----G-----A-----A-----CT-----G-----T-----A-----		4963
G.SE.SEG165G-----G-----A-----A-----CT-----G-----T-----A-----		5020
H.CF.90CF056G-----G-----A-----A-----CT-----G-----T-----A-----		4970
J.SE.SE7887G-----G-----A-----A-----CT-----G-----T-----A-----		4937
K.CM.MP535CG-----G-----A-----A-----CT-----G-----T-----A-----		4819
01.AE.TH.CM240G-----G-----A-----A-----CT-----G-----T-----A-----		5191
02.AG.NG.IBNGG-----G-----A-----A-----CT-----G-----T-----A-----		5148
03.AB.RU.KAL153-2G-----G-----A-----A-----CT-----G-----T-----A-----		5668
04.CPX.CY.94CY032-3G-----G-----A-----A-----CT-----G-----T-----A-----		4989
N.CM.YBR30G-----G-----A-----A-----CT-----G-----T-----A-----		5217
O.CM.MBP5180G-----G-----A-----A-----CT-----G-----T-----A-----		5653
O.BE.ANT70G-----G-----A-----A-----CT-----G-----T-----A-----		5677
CPZ.US.CPZUSG-----G-----A-----A-----CT-----G-----T-----A-----		5683
CPZ.GA.CPZGABG-----G-----A-----A-----CT-----G-----T-----A-----		5037
CPZ.CM.CAM3G-----G-----A-----A-----CT-----G-----T-----A-----		5325
CPZ.CM.CAM5G-----G-----A-----A-----CT-----G-----T-----A-----		5080
CPZ.CD.CPZANTG-----G-----A-----A-----CT-----G-----T-----A-----		5080
H2A.GW.ALI Vif# P Q R N S T T R #		Vif
H2A.GW.ALI VpxP R E T V P P G		Vpx
H2A.GW.ALIACCCAGAGAACAGTACCACAGGA# O W R R D Y R R G L R V A R Q D	5973
H2A.SN.HIV2STGGCCCCGAGAGACAAATACCACAGGAA-CAGTGGC-AAGA-ACT-TC-GAGAGGC-TTC-G-	5425
H2A.DE.BENGACCCGAGAGAAGGTACCCAGGAA-CAGTGG--AAGAGACC-TT-GAGAGGC-TTC-G-	5980
H2B.GH.D205CCGAGGAGAGAGTGCRCRAGGAA-CHG-G-C-AAGA-ACAGTA-GAGA-GCAITTC-CT--	5956
H2B.CI.EHOCCGAGGAGAGAGTGCRCRAGGAA-CHG-GGC-AAGA-ACAGTA-GAGAGGCATTC--	5954
H2G.CI.ABT96CCAGGGAGAGATMCCACRAGAAA-CHG-G-C-A-GAGACA-TC-GAGAGGCATTC--	5307
MAC.US.SIVSMH9GATCCAGGAGAGATCCACCTGGAA-CAGTGG--AAGAGACA-TC-GAGAGGC-TTC-	5870
STM.US.STMSRTCCAGGAGAGATCCACCTGGAA-CAGTGG--AAGA-ACRRTA-CCGAGGCYTT-C-	5367
TANTALUS.UG.ITANIGATCCAGAGAGAAATCCCCCGGAA-CAGTGG--AAGAGCA-TA-GA-GCAITTC-G-	5541
GRIVET.ET.GRI-677GATCCAGAGAAACCTTTACCAATGGTG--AT-TGG-AT-T-G-C--GTTGG-C	5833
VERVET.KE.VBR-9063GTTCCAGAGAAAACAGCCACAGGAAGT--GAT-TGG-AT-T-A-C-A--A-TGGG-C	5607
VERVET.KE.AGMTYOGATCCAGAGAGCAAGACAGGAGAAGT--ATTGG-AT-T-AGC--GTGGG-CA	5840
VERVET.KE.AGM155GATCCAGAGAGCAAGACAGGAGAAGT--ATTGG-AT-T-AGC--GTGGG-CA	5330
SABAEUS.SN.SABLGATCCAGAGAGCAAGACAGGAGAAGT--ATTGG-AT-T-AGC--GTGGG-C	5835
SUN.GA.SIVSUNGATCCAGAGAGCAAGACAGGAGAAGT--ATTGG-AT-T-AGC--GTGGG-C	5338
MANDRILL.GA.MNDGB1GGCTCCACCCAGTAGGAGGGATCCTCCNAGGATCCT--ATATGG-ATTGAGC--GTGGG-C	6050
LHOEST.KE.hoestGGCCAGAGCCAGTAGACACRACCA-G-G-----C-A-----TITGCAC--	5994
LHOEST.CD.hoest447GGCCAGAGAGATGACAAATATCAT-----A-T-CC-A-----TITGCAC--	5306
LHOEST.CD.hoest485CAGCAAAAGGTCCAGAAAACCTCCAG-TTT-A-----CA-----TGGGG-C-G--	5950
LHOEST.CD.hoest524CAGCAAAAGGTCCAGAAAACCTCCAG-TTT-A-----CA-----TGGGG-C	4869
RCM.NG.SIV-NGMCAGCAGCCGACCCAGAAAAGCCCTCAG-TTT-A-----CA-----TGGGG-C	4866
COL.---SIV-CGU1GAAAGATGCCAGAGCCCACTGGGGCT--AGATGTA.....GAGTTGGCCCC--	5968
SYKES.KE.SYK173TTGATCCTCCGAGATGCAAGAACACCATGGTCTTTCATCCCAAGAATGT-G-TT-BCC--A-GTGATCAATGTCACT-TAAA	5807
SYKES.KE.SYK173 VifLQSLPACARNTMVLHSHSKKCPVDPK-RDQCCHK	Vif
SYKES.KE.SYK173 VprFNP SQHGTPWFIPIRNVFLTPNVI_NV_T_V_K	Vpr

B.FR.HXB2 Vpr	L	E	E	L	K	N	E	A	V	R	H	F	P	R	I	W	L	H	G	L	G	O	H	I	Y	E	T	Y	G	D	T	W	A	G							
B.FR.HXB2	TTT	A	G	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G					
A.UG.U455	G	-	A	-	C	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-					
C.ET.ETH2220	C	-	C	-	G	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-					
D.CD.84Zr085	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
F1.BF.VI850	G	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-					
G.SE.SE6165	G	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-					
H.CF.90CF056	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
J.SE.SE7887	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
K.CM.MP535C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
O1.AE.TH.CM240	G	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-					
02.AG.NG.IBNG	G	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-					
03.AB.RU.KAL153-2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
04.CP.CY.94CY032-3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
N.CM.MVP5180	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-					
O.BE.ANT70	C	-	C	-	G	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-	C	-					
CPZ.US.CPZUS	C	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-			
CPZ.GA.CPZGAB	C	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-			
CPZ.CM.CM3	C	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-	A	-			
CPZ.CD.CPZANT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
H2A.GW.ALI Vif	S	R	G	L	K	Q	R	G	G	E	S	P	A	P	R	E	L	I	F	P	G	A	H	F	P	G	V	A	K	V	L	E	I	L	W	A	\$				
H2A.GW.ALI Vpx	V	E	A	L	N	R	E	A	V	N	H	L	P	R	E	L	I	F	Q	V	W	Q	R	S	W	R	Y	W	H	D	E	Q	G	M	S	O	S	Y	T		
H2A.GW.ALI	AG	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
H2A.SM.HIV2ST	GG	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
H2A.DE.BEN	CA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
H2B.GH.B205	AA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
H2B.CI.EHO	AA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
H2G.CI.ABT96	AG	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
MAC.US.MW251	AS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
STM.US.STM	GG	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
TANTALUS.UG.TAM1	GG	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
GRIVET.ET.GRI-677	GC	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
VERVET.KE.VBR-9063	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
VERVET.KE.AGM155	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
VERVET.KE.AGM155	GG	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
SABAUS.SN.SABLIC	GC	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
SUN.GA.SIVSUN	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
MANDRILL.GA.MNDGB1	CA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
LHOEST.KE.hoest	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
LHOEST.CD.hoest447	CA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
LHOEST.CD.hoest485	CA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
LHOEST.CD.hoest524	AA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
RCM.NG.SIV-NGM	GC	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
COL. SIV-CGU1	GG	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
SYKES.KE.SYK1173	GGC	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
SYKES.KE.SYK1173 Vif	G	R	T	G	S	D	R	S	I	O	A	F	Y	S	S	R	N	I	W	S	L	E	S	I	L	K	R	R	G	R	D	S	R	G	R	D	S	Y			
SYKES.KE.SYK1173 Vpr	A	E	L	V	V	T	E	A	S	K	H	F	T	P	Q	E	I	Y	W	S	L	E	S	I	L	K	R	E	A	G	T	D	S	P	T	M	A	W	E	R	T

HIV-1/HIV-2/SIV complete genomes

Strain	Genomic Region	Sequence	Position
B.FR.HXB2 Tat	-> Tat exon 1 start	M_E_P_V_D_P_R	
B.FR.HXB2 Vpr	HIV-1 Vpr normal end <-	R_R_A_R_N_G_A_S_R_S_	
B.FR.HXB2		GGAGGCAAGAAATGGAGCCAGTAGATCTCTAG	5850
A.UG.U455		-A--GC--G--	5295
C.ET.FTH2220		-A-----T-----	5241
D.CD.84ZR085		-A-----T-----	5371
F.I.BE.VI850		-A-G-T--GG--C--G--A	5189
G.SE.SE6165		-A-----T-----	5246
H.CF.90CF056		-A-----T-----	5196
J.SE.SE7887		-A--GG--G--	5163
K.CM.MP535C		-A--GG--G--	5045
O1.AE.TH.CM240		-A--GC--G--	5417
O2.AG.NG.IBNG		-A--GC--G--	5374
O3.AB.RU.KAL153-2		-A-----T-----	5894
O4.CPX.CY.94CY032-3		-A--G-C-C-GG--TGG--CCAG-AG	5212
N.CM.YBR30		-A--AG-ATGGRAACAGT--ATCC-AG	5440
O.CM.MVP5180		AC-C-AG-G--GA--AGA--A--ATGGA--CAAGTAGATCCTGA	5891
O.BE.AN770		-A--G--GAAAT--GATCC--TAGA--CCTGA	5906
CPZ.US.CPZUS		-A--AG-ATGGATCCAGT--ATCC-AA	5921
CPZ.GA.CPZGAB		-A--T-C-ATGGATCCAAT--ATCC-GA	5913
CPZ.CM.CAM3		-A--GAG--ATGGA-CCAGT--ATCC-GA	5260
CPZ.CM.CAM5		-A--GAG--ATGGAACCAAT--ATCY-AA	5548
CPZ.CD.CPZANT		-A--G--CATGGACCCCGT--AC-C-GA	5282
H2A.GW.ALI Vpx.Vpr	-< Vpr starts	M_T_E_A_P_T_E_F_P_P_A_G_M_G_P_H	
H2A.GW.ALI		TGACTGAGCACCAACAGAGTTTCCCGCGCGAATGGGACCCCA	6275
H2A.SN.HIV25T		TGACTGAAAGCACCAACAGAGTCTCCCGCGGAGGATGAGACCCAC	5727
H2A.DE.BEN		TGACTGAAAGCACCAACAGAGTCTCCCGCGGAGGATGAGACCCAC	6285
H2B.GH.D205		TGACTGAAAGCACCAACAGAGTCTCCCGCGGAGGATGAGACCCAC	6258
H2B.CI.EHO		TGGCAGAAGCCCGCCAGAGATTTCCCTCCAGAGAAGGAAACCCACA	6256
H2G.CI.ABT96		TGGCAGAAGCCCGCCAGAGATTTCCCTCCAGAGAAGGAAACCCACA	5601
MAC.US.MM251		GGCAGAAGAAATCCCGCCA	6161
SNM.US.SIVSMH9		ATGGAAGAAAGACCTCCAGAAAATGAGGCCCA	5659
STM.US.STM		A.ATGACAGAAAGACCTCCAGAAAATGAGGCCCA	5833
TANTALUS.UG.ITANI		A.ATGACACACACCTCCAGAGATGAGGCCCA	6074
GRIVET.ET.GRI-677		CCACCCCTCTCCAGGACTTGATA	5869
VERVET.KE.VER-9063		AGAG-ATGGACA--GAG-AG-G--	6105
VERVET.KE.AGM150		AGAG-ATGGACA--GGG-AG--AGAG-CAACCTCTCCAGGACTTGATA	5595
VERVET.KE.AGM155		A--GAG-ATGGACA--GGG-AG--AGAG-CAACCTCTCCAGGACTTGATA	6100
SABAUS.SN.SABLC		A--GAG-ATGGACA--GGG-AG--AGAG-CAACCTCTCCAGGACTTGATA	5603
SUN.GA.SIVSUN		CCCC-GTCTGGGA---TTCAG--GGAGCTGATCGCCCTCCAGGACTTGATA	6333
MANDRIL.GA.MNDGB1		ATGTCAACC--GGCC-C-A-CA--C-AAAGCCAGGCAAGGACATTAAGAGG	6259
LHOEST.KE.hoest447		T-CCCC-CTC--GGATTCAGCCA--GGAG--TAGATTTATA	5541
LHOEST.CD.hoest447		CCATC--C-G-A-C--ACCA--C-TA--TGAAGGACATCAACAGGCACTGA	6203
LHOEST.CD.hoest485		CTCAC--G-C-G-A-C--ACCA--C--ACTAAAGACATCAAAAGAGCCCTTA	5122
LHOEST.CD.hoest524		CTCAC--G-C-G-A-C--ACCA--C--ACTAAAGACATCAAAAGAGCCCTTA	5122
RCM.NG.SIV-NGM		CACACC--C-G-A-C--ACC--C-AACTGGAAGAGATACAAAGAGCCCTTA	5119
COL.---SIV-CGU1		-ACAG-G-CA-GCAA-G-AAGCAG--GAAAGAAATCCAGATTTCTCCGGGAATGTA	6271
SKYES.KE.SYK173 Tat		AAT--TCGCTTTCCTCTA--GCA--ACC-T--	5516
SKYES.KE.SYK173 Vpr		--TACC-CCATCC--TCTTAG	6042
SKYES.KE.SYK173 Vpr end in SIVSYKES <-		V_P_P_S_F_L_J	exon1

B.FR.HXB2 Rev	-> Rev start except in macaques and manglebeys	M A G R S G D S D E L I R T V R L I K L I Y O S S	Tat and Rev exon end / intron	exon1
B.FR.HXB2 Tat	G I S Y G R K K R A H O R R A H O N S O R T H O A S L S K O S	AGGATCTCTATGGCAGGAAGAAGCGGGA.....CAGCGACGAGACTCATCAG.....AACAGTCAGACTCATCAGGTTCTCTCAAAGCAG.TAA.....		exon1
B.FR.HXB2				6049
A.UG.U455	T.....A.....-CC.....G-C-C-A.....GG.....A-GA.....A-CT-A-C.....		5494
C.ET.ETH2220	T.....C.....-C.....C-A.....G.....A-GA.....AA-CT-A-C.....		5439
D.CD.842R085	T.....-C.....C.....C-C.....G.....A-GA.....A-C-T-A-C.....		5570
F1.BE.V1850	G.....-AC.....A-C-A.....G.....AGTG.....AA.....A-C-A-C.....		5388
G.SE.SE6165	AG.....-C.....G-A-C-A.....G.....A-AGC.....A-C-G-C.....G.....		5445
H.CF.90CF056	A.....T.....-C.....-AG.....A-C-GCA.....GTTTG.....AGA.....AA.....A-CT-A-C.....		5395
J.SE.SE7887	T.....-A.....-CC.....TC.....C-CT.....G.....A-AC.....A-CT-A-C.....		5362
K.OM.MP435C	T.....-A.....-CC.....G-A-C-C-T-C.....G.....G-AC.....AA-C-A-C.....		5244
O1.AE.TH.CM240	02.AG.NG.IBNG.....-GC.....G-A-C.....G.....G-AC.....AA-C-A-C.....		5572
02.AB.RU.KALI53-2	T.....-A.....-T.....G-T-T-A.....G.....G-CA.....AA-CT-A-C.....		6093
03.CPX.CY.94CY032-3	G.....-A.....T.....A.....AGA-C-GC-GCT.....GCAAGC.....TA-CAGA-A-A-C-G-C-G.....		5418
N.CM.YBF30	G.....-A.....AG.....G-GCCG.....GCA.....CAGA-A-A-C-G-C-G.....		5639
O.CM.MVP5180	G.....-A.....AG.....G-GCCG.....GCA.....CAGA-A-A-C-G-C-G.....		6087
O.BE.ANT70	G.....-A.....AG.....G-GCCG.....GCA.....CAGA-A-A-C-G-C-G.....		6099
CPZ.US.CPZUS	G.....-A.....AG.....G-GCCG.....GCA.....CAGA-A-A-C-G-C-G.....		6126
CPZ.GA.CPZGAB	G.....-A.....AG.....G-GCCG.....GCA.....CAGA-A-A-C-G-C-G.....		6118
CPZ.CM.CAM3	G.....-A.....AG.....G-GCCG.....GCA.....CAGA-A-A-C-G-C-G.....		5462
CPZ.CM.CAM5	G.....-A.....AG.....G-GCCG.....GCA.....CAGA-A-A-C-G-C-G.....		5750
CPZ.CD.CPZANT	G.....-A.....AG.....G-GCCG.....GCA.....CAGA-A-A-C-G-C-G.....		5490
H2A.GW.ALI Rev	-< Rev exon 1 start	M T E R A G E E D L O R K L R L L L L H O T N	Tat and Rev exon end / intron	exon1
H2A.GW.ALI Tat	G I W Y D R A K R R R R S P K K L R L L L L H O T N	C-G-A-GG---A-C-A---GGCA-C.....AGAA-TCTCC-AGAA-A.....T-A-G---TCGT---GC---G-CA---G.....		6691
H2A.GW.ALI				6143
H2A.SN.HIV2ST		C-G-A-GG---AAC-A---GGCA-A.....AGAA-ACTCCGAA-A.....C-A-G---TCGT---GC---G-CA---G.....		6701
H2A.DE.BEN		C-A-A-A---AGC-A---GGCA-AC.....AGAA-GACTCC-AGGA-A.....-C-A---C-TCG---GC---G-CA---G.....		6668
H2B.GH.D205		T-G-A-GT---A-C-CTC---GAA.....TCTGC-AA---A-G.....-AC-GC---GC---G-CA---G.....		6666
H2B.CI.EHO		G-G-A-GT---AAC-CTC---GAAA.....TCTTC-AA---G-A-G.....-AC-AG-T---GC---C-G---G.....		6014
H2G.CI.ABT96		C-AG-A-GT---AGCACT-GAA-A.....AGAAGT-CGAA---AGA-A-G.....G-A---TC-TT---GC---C-G---G.....		6577
MAC.US.MW251		G-G-A-GT---AGCA-TCACGAAA.....AGAA-ACTCCGAAA-G.....GC-A---G-A---AC-T---GC---C-A.....		6069
SMM.US.SIVSMMH9		T-A-A-AA---AGCA-C-ACGAA-A.....ACT-GAAG-AGA-A-G.....ACT-GAAG-AGA-A-G.....G-A---AC-T---GC---C-A.....		8243
STM.US.STM	-A---AGTCTCT-G-ATC.....GTTAAGAA-AG---A-G.....TCA-A---GAA-CT-G---C.....		8264
TANTALUS.UG.TANI	-A---CATGCTCT-G-AAA.....AGA-CTAAG-AGAA.....AGA-A---ACC-A---CGGC.....		6029
GRIVET.ET.GRI-677	-TG-TA---CATGCCCT-G-ACC.....AGAAAAC---T-T-G.....C-A---GCCTA---A-CAGG-C.....		6277
VERVET.KE.VER-9063	-TG-TA---CATGCCCT-G-ACC.....AGAAAAC---T-T-G.....C-A---GCCTA---A-CAGG-C.....		5767
VERVET.KE.AGMTYO	-TG-TA---CATGCCCT-G-ACC.....AGAAAAC---T-T-G.....C-A---GCCTA---A-CAGG-C.....		6272
VERVET.KE.AGM155	-TG-TA---CATGCCCT-G-ACC.....AGAAAAC---T-T-G.....C-A---GCCTA---A-CAGG-C.....		5775
VERVET.DE.AGM3	-TCG---TATGTCCT-G-CA.....AGAGTCTT-AGAAG.....ATTT-G---A-TCAAGT-TCTCT-C-CA-C---G.....		6475
SABAUS.SN.SABIC	-G-CA---TATGTCCT-G-G-AC.....AT-A---TC-ATATC.....-AGG-A---GCAAGGCTGT-CGAAG.....		6416
SUN.GA.SIVSUN	-A---AAGG---CGATGCT-C-G-RAAC-T.....GTA-C-G---CTAA-A-GA.....-TA-CTGG-AG-GGT-AGAAGCTAT-CGAAG.....		5734
MANDRIL.GA.MNDGB1	-GC-ACGA---TATGTC-C-G-RAAC-G.....AGA---GTT-CCT-GATA-TT.....GAGAC---GCNAGA---CTGTGG-CAG.....		5366
LHOEST.KE.hoest447	-G-ACGA---TATGTC-C-G-A.....A-AA---GTGCCAGGTATT.....GAAG---A-GNAGA---CTGTGG-CGG.....		5285
LHOEST.CD.hoest448	-G-ACGA---TATGTC-C-G-A.....A-AA---GTGCCAGGTATT.....GAAG---A-GNAGA---CTGTGG-CGG.....		5285
LHOEST.CD.hoest524	-G-ACGG---TATGTC-C-G-ACC.....AGAA---GTT-CCAGTA-TT.....GAAC-G-GCAGGA---CTGTGG-CAG.....		5282
RWM.NG.SIV-NGM	-AAA---CTTCT-G-GAA-A.....AGAG-CA-CC-GAA.....AAGAAG-G-A---T-A---C-TT-GA-AC---TG.....GTATGGAATGCTCTGTGTACCT.....		6721
COL.---SIV-CGU1		G-A---CG---AAG-A---T.....AG-GCTA-ACAAAC-CAGT-A.....AG-GCTA-ACAAAC-CAGT-A.....		5635
SYKES.KE.SYK173		G-A---A-T---C---ACCA-G-AA-C.....GCAGCT--C--TAT.....T-G-GG-T-A-AGCGC-C-T-CAGG.....		6218
SYKES.KE.SYK173 Tat	G I T Y A R P R K R A A R S I A R S I	Tat and Rev exon 1 end / intron		exon1
SYKES.KE.SYK173 Rev	M P D O G S E O L A V F L R M I A H L Q			exon1
	-< Rev start except in macaques and manglebeys			

HIV-1/HIV-2/SIV
complete genomes

(ACG start codon in HXB2, ATG in others)

B.FR.HXB2 Vpu
 B.FR.HXB2
 T O P I V A
 | -> Vpu start
 T O P I V A
 6086

.....GTAGTACATGTAACCAACCTATACCAATAGTACAA
GTAAT--CTT--T-AC--T-GGA--CTGG-- 5535
GTAC.CAAATAATAGATGTAAG--T-ATTAC--G-AA--GTAGATTAT-G-A--T-- 5498
GTAGGTTHAA--T--T--T--A--T-- 5607
GTAGAAAAAT--T--TC-TA-T-GTT-CC-A--GG- 5435
T--GT-AT--GT--T-- 5432
T-TCATA--T-T-TAVAT--GG-T--G-AT-G 5432
GTAATCAAGTA.-C-T--T-AT--T-G-A--C-- 5408
GTAGTCTAA--T-GTGT-CT-G--TTC-AT-G 5291
GTAATA--T-AC--T-GGA--TAGT-- 5657
GTAGTAATAAT--C-T--G-T--T-A--ACT-- 5619
T--T-CT-G--C-- 6130
GTATTA--T--TT--TTCTGGGA--CTGG-- 5459
GTAAC-C-TGT--TA--G-TGTC-T-G-G-T 5670
ACGC-G-T--T-RAGAAAC-CC--CT 6121
ACGC-T--T-A-GGGAC-CC-- 6133
CCC..T-TTG-CT-T--TAA-TGGTTGA-A--GTT 6144
GTAAGTTC--T-T-ACT-TGT--GTGGCT--TTC 6144
GTAAGTTC--CT-T-T-A-TGGGA-CA-A--GCT 5502
GTAAT-TT-G--CT-T-T-AT-TGGGAGCA-A--GTT 5790
CTC.....T--T-ACTAA--TTT..... 5511

H2A.GW.ALI Env
 H2A.GW.ALI
 M M S S R N O L
 | -> Env start
 M M S S R N O L
GT-T-ATGT--GTAG--ATCAGCTGC 9p120
GT-T-ATGT--GTAG--ATCAGCTGC 6718
GT-A-ATGTG--GTTAGG-ATCA-CT-T 6170
GT-T-G-C--GTTAGG-ATCAGCTGT 6728
GT-T-GC-TA-T-TAGC-GCGGCTGC 6695
GT-T-GC-A-G-TAAT-ATTACCT-C 6693
GT-T-GC-TA-C-TGGC-ATCAGCTGC 6041
GT-T-GG-TG-C-TGGG-ATCAGCTGC 6604
GT-T-GG-TG-C-TGGG-ATCAGCTGC 6096
GT-T-GCCTGCCCTGG--ATCA-CTGC 6270
GT-T-GG--AT--AGGG-AA--GGG 6291
GT-T-GGAGAT-G-TT--AA-AT-C 6056
GT-T-A--C-TTT-GGTA--TTC 6304
GT-T-AGTA-C-AT--ACCTT-G 5794
GT-T-AC-AGT-CTT-GG-A-TTTT 6289
GT-T-A-G-TG-C-TT-C-G..AT-G 5799
TAAH-T-A-G-T-C-TA--G--C-GCTTT 6505
GT-A-ATG-GGTGC--TGA-A-TTG 6443
GT..ATGAAGTGT--GGTC-TTTC 5758
GG-G-ATGG-ATGT--GG-T--G 6393
GAAATGAAGTGC--TGG-C--AT-C 5309
G-ATGACTGT-TT-AGC-CCT-C 5309
AAATGG-ATGT--GGTAAACAT-T 5306
GATGCTTACCTTAAAGTATGAGAGTAA--TCTTTA-AT-T-GCTAAGT--AT-GTA--G 6841
GATGCTTACCTTAAAGTATGAGAGTAA--TCTTTA-AT-T-GCTAAGT--ACT-GCT-G 5668
AAAATGG-AGCTTTT-G-ACITTAC- 6246
 M A F R T Y
 | -> Env start

A.UG.U455
 C.ET.ETH2220
 D.CD.84Z8085
 F.I.BE.VI850
 G.SE.SEG165
 H.CF.90CF056
 J.SE.SE7887
 K.CM.MP535C
 O1.AE.TH.CM240
 O2.AG.NG.IBNG
 O3.AB.RU.KAL153-2
 O4.CPX.CY.94CY032-3
 N.CM.YBF30
 O.CM.MVP5180
 O.BE.AN70
 CPZ.US.CPZUS
 CPZ.GA.CPZGAB
 CPZ.CM.CAM3
 CPZ.CM.CAM5
 CPZ.CD.CPZANT

H2A.SN.HIV2ST
 H2A.DE.BEN
 H2B.GH.D205
 H2B.CI.EHO
 H2G.CI.ABT96
 MAC.US.MM251
 SMM.US.SIVSMH9
 STM.US.STM
 TANTALUS.UG.TANI
 GRIVET.ET.GRI-677
 VERVET.KE.VBR-9063
 VERVET.KE.AGM150
 VERVET.KE.AGM155
 VERVET.DE.AGM3
 SABAEUS.SN.SABLC
 SUN.GA.SIVSUN
 MANDRIL.GA.MNDGB1
 LHOEST.KE.hoest
 LHOEST.CD.hoest447
 LHOEST.CD.hoest485
 LHOEST.CD.hoest485
 LHOEST.CD.hoest524
 RCM.NG.SIV-NGM
 COL.--.SIV-CGU1

SYKES.KE.SYK173
 SYKES.KE.SYK173 Env

B.FR.HXB2 Vpu 6186 I V A L V A I I A I I A I V V W S I V I I E Y R K I L R O R K I D R L TAGTAGCATTAGTAGTACCAATAAATAGCA....ATAGTGTGTGG...TCATAGTAAATAGAA....TAT...TAT...AGGAAAATATTAAAGACAAAGAAAATAGACAGGTT...R_K_I_L_R_O_R_K_I_D_R_L
5637 --AC--GGC--GA-----C--CT-----A--T-----GGT-----AAA-----AA--T--GC--AG--A-----T-----T-----
5598 C.ET.ETH2220 --CA-----CA-----C-----C-----A-----C--TAT-----T-----A-----T-----G-----T-----
5707 F1.BF.VI.850 --T-----T-----GC-----C-----C-----A-----T-----TAT-----T-----A-----A-----T-----AA-----
5535 G.SE.SE6165 --G-----G-----A-----A-----T-----T-----T-----T-----A-----G-----AG--A-----GA--A-----C-----
5592 G--CGCT--G-----AC--TTT--C--GCC-----G-----A-----G-----A-----G-----A-----A-----G-----AG-----
5529 --C--TT--G-----C--CT-----TT--C--I-----TT--C--I-----G-----A-----A-----T-----G-----TAT-----T-----A-----
5508 K.CM.MP535C --G--C--GA-----C--CT-----GC-----C-----A-----A-----A-----A-----T-----TAT-----T-----G-----GT-----
5757 02.AE.TH.CM240 --G--C-----T--C-----GC-----C-----A-----G-----AG-----TAT-----T-----A-----AGG--A-----G--AG-----
5719 J.SE.SE7887 --T--G-----G-----C-----T-----T-----G-----G-----T-----T-----T-----T-----A-----T-----
6230 04.AB.RU.KAL153-2 --G--C--G-----G-----C-----T-----T-----A-----TT-----T-----T-----A-----T-----GAG--G-----G--G-----C-----
5559 N.CM.YBF30 --G--G-----G--C-----T-----GC-----C-----A-----A-----C-----G-----AT-----C-----TAT--G-----G-----A-----ACACA-----
5770 O.CM.MVP5180 --T-----T-----AGT--CTTGTGTC--T-----AAT-----G-----C-----A-----A-----T-----TCT--TAAOC--T--G--ATT-----TTAGTGCA--AG--A-----GAT--G--GG--GCA--GA-----
6221 A--ATTA--TAGT--CTTTCG--GT-----T-----AAT-----G-----A-----T-----A-----G-----T-----A-----G-----G-----G-----A-----G-----G-----
6233 --A-----TT--G-----GCT-----G--A-----A-----G--T-----T-----T-----G-----GC-----AA--C-----GGAC--TGCT--G--TTG-----GGAGACAATA--A--T--A--GA--ATACTCA--C-----
6261 --CA--TT--G-----GCT-----T-----T-----T-----T-----T-----T-----G-----G-----A-----A-----A-----T-----A-----G-----ATAGCTT--G--C-----
5605 A--T--GTA-----G--A-----A-----I-----T-----T-----T-----T-----T-----T-----G-----G-----C-----T-----T-----A-----A-----G-----G-----C-----C-----A-----AAA-----
5893 --A-----T-----G--A-----A-----C-----T-----T-----T-----T-----T-----T-----T-----T-----A-----G-----G-----G-----GG-----AGAGAA--T--A--GA--GA--AG--G-----CCC--A--AA-----
5618 --A-----GAGTAT--CITTTTC--TGCCCT--TAGT-----T-----A-----T-----A-----C-----A-----A-----ATT--CTGT--ACCTAT--CTCTATA--GCTTTTATAA--TAT--T--AGCAGCAG-----T--G--T--ATA--G--AAA-----

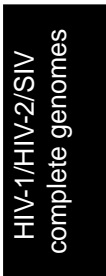
H2A.GW.ALI Env 9p120
H2A.GW.ALI L V T T I L L A S A C L V S A C K O Y V T V F Y Y C K O Q Y V T V F Y TATTGTAACAA--ATGTGA--TG--TT--TTAT.....TATTGTAACAA--ATGTGA--TG--TT--TTAT.....
6235 -T--T--CAGCT--GC-----T--GTGCTTGCTT--ATA.....TATTGC--CCAA--ATGTGA--TG--TT--CTAT.....TATTGTAACAA--ATGTGA--TG--TT--CTAT.....
6793 H2A.DE.BEN -T--T--TCA--TT--C--A--GTGCTTGCTT--GTA.....TATTGTAACAA--ATGTGA--TG--TT--CTAT.....TATTGTAACAA--ATGTGA--TG--TT--CTAT.....
6760 H2B.GH.D205 -CTA--T--GC--CC--GC--TAT--GCT--C--GT--GGTTT.....GTATCTAAGCA--ATGTTA--TG--CT--CTAT.....ATGCGAAGAC--TTGTC--TG--CT--CTAT.....
6758 H2B.CI.BHO -T--TA--C--CC--GC--TAT--GT--CTAT--GCTAT.....ATGCGAAGAC--TTGTC--TG--CT--CTAT.....ATGCGAAGAC--TTGTC--TG--CT--CTAT.....
6106 H2G.CI.ABT96 --TA--C--ACCT--RC--CTT--GT--CTTAT--GGTTC.....TATTGTACCAA--ATGTCA--TG--TT--TTAT.....TATTGTAC--CAA--ATGTCA--AG--CT--TTAT.....
8669 MAC.US.MM251 -TA--C--CA--CT--GC--TTT--GHG--CTAI--GGATC.....TATTGTAC--CAA--ATGTCA--AG--CT--TTAT.....TATTGTAC--CAA--ATGTCA--AG--CT--TTAT.....
6161 SMM.US.SIVSMH9 -TA--C--GC--CTGC--TY--GTGCTTC--GGAIT.....TATTGTAC--CAA--ATGTCA--AG--CT--TTAT.....TATTGTAC--CAA--ATGTCA--AG--CT--TTAT.....
6335 STM.US.STM -TA--C--CT--GC--TT--GTGCTTGCTT--GACC.....TATTGTAC--CAA--ATGTCA--AG--CT--TTAT.....TATTGTAC--CAA--ATGTCA--AG--CT--TTAT.....
6362 TANTALUS.UG.TANI -T--TT--G--A--TT--G--T--GCT--AT--GGACTGTATATGGACACA--ATATTA--AG--GT--TTAT.....TATTGCAC--CA--ATGTCA--GG--T--TTAT.....
6115 GRIVET.ET.GRI-677 -A--AT--GC--A--GGAT--GT--GG--AT--GGTAAAC.....C-----ATGTGA--AG--GT--TTAT.....C-----ATGTGA--AG--GT--TTAT.....
5853 VERVET.KE.VER-9063 -TA-----TA--CA-----G--AT--GTTCTT--TT--AGTGAAGA.....CAA-----ATGTGA--AG--GT--TTAT.....CAA-----ATGTGA--AG--GT--TTAT.....
6366 VERVET.KE.AGMTYO -G--A--TT--G--GA-----G--AT--GGG-----G--GT--AGTAAAG.....CAA-----ATGTGA--AG--GT--TTAT.....CAA-----ATGTGA--AG--GT--TTAT.....
6364 VERVET.KE.AGM155 -TT--GG--A-----GGAT--GG--G--GCTTAAATCAAAG.....CAGCA-----ATGTGA--AG--GT--TTAT.....CAGCA-----ATGTGA--AG--GT--TTAT.....
6561 SABAEUS.SN.SAB1C -GGA--CT-----A-----GGAT--GG--G--GCTTAAATCAAAG.....CAACAA-----ATGTGA--AG--GT--TTAT.....CAACAA-----ATGTGA--AG--GT--TTAT.....
6499 SUN.GA.SIVSUN -GGC--AGTGGTGTGGAGCT--G--GTGT--GTC.....GGTAGTATAGC--A--ACA--ATGTGA--AG--GT--TTAT.....GGTAGTATAGC--A--ACA--ATGTGA--AG--GT--TTAT.....
5819 MANDRILL.GA.MNDGB1 --GTTTC--CTT--GC--GTT-----GGTAGTATAGC--A--ACA--ATGTGA--AG--GT--TTAT.....GGTAGTATAGC--A--ACA--ATGTGA--AG--GT--TTAT.....
6449 LHOEST.KE.roest LHOEST.KE.roest -TC--GCTT--GC--TC--T-----GGTAGTATAGC--A--ACA--ATGTGA--AG--GT--TTAT.....GGTAGTATAGC--A--ACA--ATGTGA--AG--GT--TTAT.....
5374 LHOEST.CD.roest447 -TT--CTTGA--CA--TT--AG--G--CA-----GGGAGAAAACA--AGAGCA--ATGTAA--AG--GT--TTAT.....GGGAGAAAACA--AGAGCA--ATGTAA--AG--GT--TTAT.....
5371 LHOEST.CD.roest485 -TT--CTTGC--CT--TT--G--GC--CA-----GGGAG.....GCA--AGAGCA--ATGTAA--AG--GT--TTAT.....GGGAG.....GCA--AGAGCA--ATGTAA--AG--GT--TTAT.....
5374 LHOEST.CD.roest524 -GC--TTGC--GC--CTGCTGGG--C-----GGGAGAA.....AAGGAGAAAACA--ATGTAA--AG--GT--TTAT.....AAGGAGAAAACA--ATGTAA--AG--GT--TTAT.....
6903 RCM.NG.SIV-NGM -TA--AT--GG--T--GT-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----
COL.____.SIV--CGU1 --T--GTA-----GT--AGT--AT-----AAA-----GTAA--AG--TATCAA.....AAA-----GTAA--AG--TATCAA.....

SYKES.KE.SYK173 Env 6314
SYKES.KE.SYK173 I V C L F S L I S L G F M E K O Q Y V T V F Y Y T....T....GTC--TT--TAGCT-----GTTT--GGAITTTGGAA--A--CAACAA--ATGTAA--AG--T--TTAT.....

HIV-1/HIV-2/SIV
complete genomes

B.FR.HXB2 Env	-> Env gp160 start in HIV-1	M R V # K E K Y O H L L W	R A E D S G N E S E #	G E I S A L I V E	R W G W	gp120
B.FR.HXB2 Vpu						
B.FR.HXB2						6269
A.UG.U455						5717
C.ET.ETH2220						5678
D.CD.84ZR085						5787
F.J.BE.VI850						5615
G.SE.SE6165						5672
H.CF.90CF056						5609
J.SE.SE7887						5588
K.CM.MP535C						5462
O1.AE.TH.CM240						5837
O2.AG.NG.IBNG						5799
O3.AB.RU.KALI53-2						6307
O4.CPX.CY.94CY032-3						5639
N.CM.YBF30						5850
O.CM.MVP5180						6322
O.BE.ANT70						6334
CPZ.US.CPZUS						6358
CPZ.GA.CPZGAB						6348
CPZ.CM.CAM3						5697
CPZ.CM.CAM5						5985
CPZ.CD.CPZANT						5719
H2A.GW.ALI Env						gp120
H2A.GW.ALI						6783
H2A.SN.HIV2ST						6235
H2A.DE.BEN						6793
H2B.GH.D205						6760
H2B.CI.EHO						6758
H2G.CI.ABT96						6106
MAC.US.MM251						6669
SMM.US.SIVSMMH9						6161
STM.US.STM						6335
TANTALUS.UG.TANI						6362
GRIVET.FI.GRI-677						6115
VERVET.KE.VBR-9063						6386
VERVET.KE.AGMTYO						5853
VERVET.KE.AGM155						6364
VERVET.DE.AGM3						5864
SABAEUS.SN.SABIC						6561
SUN.GA.SIVSUN						6499
MANDRILL.GA.MNDGB1						5819
LHOEST.KE.loeest						6449
LHOEST.CD.loeest447						5374
LHOEST.CD.loeest485						5371
LHOEST.CD.loeest524						5374
RCW.NG.SIV-NGM						6903
COL.SIV-CGU1						5712
SYKES.KE.SYK173						6314
SYKES.KE.SYK173 Env						gp120

B.FR.HXB2 Vpu	E_M_G_H_H_A_P_W_D_V_D_D_V_D_D_V_D_D	Vpu end <-	Vpu
B.FR.HXB2 Env	R_W_G_T_M_E_L_G_M_L_M_I_C_S_A_T_E_K_L_W_V_T_V_Y_G_V_P_V_W_K_E_A_T		9pL20
B.FR.HXB2	AGATGGGCGCACATGCTCCCTTTGGGATGTGAAG.....ATCTGTAGTGTACAGAAAATTGTGGGTCACAGTCTATTATGGGGTACCTGTGGGAAGAACARACC.....T_T_L_F		6384
A.UG.U455	A-T-T-A-T-G-T-T-A-A-A.....AT-T-AA-G-C-C-C.....G.....C.....-A-T-GTT		5829
C.ED.ETH2220	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5793
D.CT.842R085	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5905
F.J.BE.VI1850	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5730
G.SE.SE6165	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5787
H.CF.90CF056	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5724
J.SE.SE7887	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5703
K.OM.MP435C	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5583
L.AE.TH.CM240	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5952
O2.AG.NG.IBNG	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5911
O3.AB.RU.KALI53-2	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6416
O4.CPX.CY.94CY032-3	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5754
N.CM.YBF30	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5971
O.CM.MVP5180	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6422
O.BE.ANT70	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6431
CPZ.US.CPZUS	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6455
CPZ.GA.CPZGAB	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6448
CPZ.CM.CAM3	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5803
CPZ.CM.CAM5	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6088
CPZ.CD.CPZANT	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5810
H2A.GW.ALI Env	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		9pL20
H2A.GW.ALI	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6823
H2A.SN.HIV2ST	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6275
H2A.DE.BEN	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6833
H2B.GH.D205	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6800
H2B.CI.BHO	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6798
H2G.CI.ABT96	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6146
MAC.US.MM251	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6709
SMM.US.SIVSMMH9	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6201
STM.US.STM	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6375
TANTALUS.UG.TANI	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6402
GRIVET.FT.GRI-677	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6155
VERVET.KE.VBR-9063	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6406
VERVET.KE.AGMTYO	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5893
VERVET.KE.AGM155	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6404
VERVET.DE.AGM3	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5904
SABAEUS.SN.SABIC	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6601
SUN.GA.SIVSUN	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6539
MANDRILL.GA.MNDGB1	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5859
LHOEST.KE.hoest	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6492
LHOEST.CD.hoest447	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5414
LHOEST.CD.hoest485	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5411
LHOEST.CD.hoest524	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5414
RCW.NG.SIV-NGM	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6943
COL. SIV-CGU1	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		5761
SYKES.KE.SYK173	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		6354
SYKES.KE.SYK173 Env	A-T-T-AGG-T-T-A.....T-A-GA-TG-GG-C.....-A-A-T-T-G.....C-T.....C-T		9pL20



HIV-1/HIV-2/SIV complete genomes

B.FR.HXB2 Env	C.A.S.D.A.K.A.Y.D.T.E.V.H.N.V.W.A.T.H.A.C.V.P.T.D.P.N.P.Q.E.V.V.L.V.N.V.T.E.N.F.N.M.W.K	9p120
B.FR.HXB2	GTGCATCAGATGCTAAAGCATATGATACAGAGTACATAATGTTTGGCCACACATGCTGTGTACCCACAGACCCCAAGTAGTATTGGTAAATGTGACAGAAAATTTTACCATGTGGAA	65114
A.UG.U455	-----G-A-G-----C-T-----A-AT-----C-----A-----G-----	5959
C.ET.ETH2220	-----A-----C-GT-----T-----G-----G-----T-G-GT-----A-----	5923
D.CD.84ZR085	-----T-----A-A-----C-C-----A-C-----T-----A-----C-----AA-----C-----	6035
F1.BE.VI850	-----A-G-----C-----C-----A-----T-----TC-----AA-----A-----G-----T-----	5860
G.SE.SE6165	-----T-----AG-G-T-----AAG-----C-----T-----A-----A-----TA-----A-----T-----G-----C-----T-----	5917
H.CF.90CF056	-----G-----T-----AAG-----C-----T-----A-----A-----GA-----G-----CA-----AG-----A-----G-----C-----T-----	5854
J.SE.SE7887	-----AG-----T-----AAA-----C-----T-----C-----G-----GA-----GAATC-----CC-----A-----C-----	5833
K.CM.MP535C	-----C-----G-----C-----G-----A-----G-----C-----C-----T-----A-----CA-----C-----A-----A-----G-----	5713
01.AE.TH.CM240	-----AG-----AG-----A-----A-----C-----T-----G-----G-----A-----C-----A-----A-----G-----	6041
02.AG.NG.IBNG	-----A-----CC-----AG-----A-----A-----C-----T-----A-----A-----C-----A-----A-----G-----	6546
03.AB.RU.KAL153-2	-----T-----CC-----AG-----A-----A-----C-----C-----A-----C-----A-----A-----G-----C-----	5884
04.CPX.CY.94CY032-3	-----T-----CCT-----ACAAGC-----T-----ACAG-----A-----A-----C-----T-----T-----T-----T-----T-----TC-----C-----A-----G-----C-----	6101
N.CM.YBF30	-----T-----CCT-----ACAAGC-----T-----AAAG-----A-----A-----A-----T-----T-----T-----T-----T-----TC-----C-----A-----G-----C-----	6101
O.BE.ANT70	-----C-----T-----A-----G-----CA-----GCAG-----CC-----C-----CA-----C-----T-----T-----T-----T-----T-----TC-----C-----A-----G-----C-----	6552
CPZ.US.CPZUS	-----C-----C-----G-----C-----AG-----CT-----A-----G-----A-----G-----A-----G-----A-----G-----A-----G-----C-----	6561
CPZ.GA.CPZGAB	-----C-----G-----GTC-----TCT-----A-----C-----A-----CC-----A-----A-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	6578
CPZ.CM.CM3	-----C-----G-----CTC-----CT-----A-----C-----A-----A-----CC-----C-----A-----A-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5933
CPZ.CM.CM5	-----C-----A-----A-----CTCCATGACAAG-----G-----C-----A-----A-----T-----AG-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	6218
CPZ.CD.CPZANT	-----C-----A-----A-----CTCCATGACAAG-----G-----C-----A-----A-----T-----AG-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5937
H2A.GW.ALI Env	C.A.....T.K.N.R.D.T.W.G.T.I.Q.C.L.P.D.D.Y.Q.E.....I.A.L.N.V.T.E.A.F.D.A.W.D	9p120
H2A.GW.ALI	-----CA-----AAATAGAG-----AC-----GA-----CATAACAG-----CT-----AGACA-----TGATG-----TTAT-----G-----A-----GCCT-----G-----GGC-----CG-----TGCA-----G-----	6929
H2A.SN.HIV25T	-----TA-----AAATAGAG-----AC-----GA-----CATAACAG-----CT-----G-----AGACA-----TGATG-----TTAT-----G-----A-----A-----GCCT-----G-----GGC-----CG-----GCA-----	6381
H2A.DE.BEN	-----TA-----AAATAGAG-----CAC-----GG-----CATAACAG-----CT-----G-----AGACA-----TGATG-----TTAT-----G-----A-----A-----TT-----GGC-----G-----TGCA-----	6939
H2B.GH.D205	-----CACAACACAGAG-----CAC-----GA-----TGTACAG-----CC-----C-----AGACA-----TGTGTC-----TACT-----G-----A-----CAG-----C-----CA-----A-----GGC-----G-----TGCA-----G-----	6906
H2B.CI.BHO	-----CAGAACACAGAG-----ACC-----GA-----TGTACAG-----CC-----C-----AGATA-----TGATG-----TATAC-----A-----C-----CAAT-----A-----A-----GGC-----G-----TGCC-----G-----	6904
H2G.CI.BHT96	-----CA-----AAATAGAG-----CACA-----G-----TACACAG-----CC-----T-----TGATA-----TCAATG-----TTACAGT-----A-----C-----GCAT-----G-----A-----GC-----G-----TGCC-----G-----	6252
MAC.US.MM251	-----CA-----AAATAGAG-----ACC-----GA-----AC-----CAG-----CC-----AGATA-----TGGTG-----TTATTC-----T-----I-----GGCC-----T-----I-----GC-----G-----TGCT-----G-----	8815
SMW.US.SIVSMW9	-----CAGAATAGGG-----AC-----GA-----CACARA-----CT-----G-----AGATA-----TGGTG-----TTATTC-----T-----T-----GGCAA-----T-----C-----GGC-----G-----TGCT-----G-----	6307
STM.US.STM	-----CA-----AAATAGGG-----AC-----GA-----AC-----CAA-----C-----AGACA-----TGGTG-----TTATTC-----T-----T-----GGCAA-----C-----GGC-----G-----TGCC-----G-----	6481
TANTALUS.UG.TANI	-----CCTAATACC-----CC-----A-----AC-----AAT-----A-----AGAT-----T-----ATG-----TTATAC-----T-----CA-----T-----CA-----T-----CT-----T-----G-----A-----G-----AGCA-----	6508
GRIVET.ET.GRI-677	-----GCCAATACC-----A-----G-----A-----CACCAA-----CA-----AGAT-----T-----ATG-----TTATAC-----G-----GCCTC-----CA-----T-----T-----GC-----CG-----GGCT-----	6259
VERVET.KE.VBR-9063	-----TCCCTACTACC-----GGC-----A-----T-----AC-----AAT-----CA-----AGAT-----T-----ATG-----TTATAC-----G-----GCCAT-----CA-----T-----T-----GCC-----G-----AGCC-----	6512
VERVET.KE.AGMTYO	-----TCCCACTACA-----GCT-----A-----T-----TAC-----AAT-----CA-----AGAT-----T-----ATG-----TTATAC-----T-----CCTC-----A-----C-----T-----CCA-----G-----GGCA-----GG-----	5999
VERVET.KE.AGM155	-----CCTACTACC-----GGT-----G-----A-----TAC-----AAT-----CA-----AGAT-----T-----ATG-----TTATAC-----T-----CCAC-----G-----A-----A-----GCCA-----G-----AGCA-----GC-----	6010
VERVET.DE.AGM3	-----CACCACACC-----GAC-----A-----A-----TAC-----AA-----CGA-----AGAT-----T-----ATG-----TACAC-----G-----GAT-----T-----TGAACCA-----GG-----AC-----A-----GCA-----A-----GTACC-----T-----CCT-----A-----C-----CAGAAAAT-----	6707
SABAEUS.SN.SAB1C	-----CCTAATA-----GCT-----G-----TG-----TAC-----T-----A-----CT-----G-----AGATTTA-----AA-----CCTATCC-----T-----G-----GCCA-----CA-----G-----TT-----G-----GAA-----TTT-----CAGA-----GG-----	6645
SUN.GA.SIVSUN	-----T-----CCTAATA-----GCT-----C-----T-----GACAT-----T-----CT-----G-----AGATTTG-----AA-----CCTAATCC-----G-----GCCA-----CA-----G-----TT-----G-----GAA-----TTT-----CAGA-----ACT-----	5965
MANDRILL.GA.MNDGB1	-----T-----CCTAATA-----GCT-----C-----T-----GACAT-----T-----CT-----G-----AGATTTG-----AA-----CCTAATCC-----G-----GCCA-----CA-----G-----TT-----G-----GAA-----TTT-----CAGA-----ACT-----	5965
LHOEST.KE.roest	-----T-----CCTAATA-----GCT-----A-----T-----GACAT-----A-----CA-----A-----AGACTTTG-----AGCTTTATCC-----C-----G-----A-----TCCATA-----TC-----CAAC-----T-----AGTTTAAA-----TT-----CTAAG-----	6598
LHOEST.CD.roest447	-----T-----CCTAATA-----GCT-----A-----T-----GACAT-----A-----CA-----A-----AGACTTTG-----AGCTTTATCC-----C-----G-----A-----TCCATA-----TC-----CAAC-----T-----AGTTTAAA-----TT-----CTAAG-----	5520
LHOEST.CD.roest485	-----T-----CCTAATA-----GCT-----A-----T-----TACAT-----A-----CC-----AGATCTG-----AACTTATCC-----C-----G-----A-----CCTTA-----TC-----CAAC-----T-----AACTTACTTAAA-----TT-----CTCAAG-----	5520
LHOEST.CD.roest524	-----T-----T-----TAAACGCAC-----ATCA-----GA-----GTTAAT-----A-----AGA-----GAGG-----TAT-----C-----CCA-----T-----G-----A-----G-----TGCT-----G-----	7049
RCM.NG.SIV-NGM	-----T-----TTCTCC-----CAG-----GA-----CAACAGGTTTG-----GG-----T-----AC-----TCCC-----TGGAAAG-----C-----T-----A-----CA-----AACA-----GCCAAATGT-----AC-----G-----AGC-----TTTGATCTTTT-----	5870
COL.....SIV-CCGU1	-----T-----TTCTCC-----CAG-----GA-----CAACAGGTTTG-----GG-----T-----AC-----TCCC-----TGGAAAG-----C-----T-----A-----CA-----AACA-----GCCAAATGT-----AC-----G-----AGC-----TTTGATCTTTT-----	5870
SYKES.KE.SYK173	-----A-----C-----CTCCCATTA-----AGGA-----GA-----T-----TA-----GAAT-----G-----GT-----CA-----CA-----CA-----AT-----A-----AG-----C-----CA-----T-----G-----G-----A-----ACTTTCACAGCATG-----	6457
SYKES.KE.SYK173 Env	C.T.....T.S.H.K.G.G.W.A.T.K.N.C.V.P.S.A.....D.Q.I.E.V.R.V.N.I.T.G.E.Y.F.P.A.W	9p120

B.FR.HXB2 EnvATAATGGAGAAAGGAGAGATATAAACTGCTCT.....TTCAATATCAGCAACGATTAAGAGGFNISSITSRG	gp120
B.FR.HXB2ATG.....	6724
A.UG.U455ATC-CA.....G-C-GAAC	6166
C.ET.ETH2220GATGGA-T-GG-A-A-G-T.....G-C-GAAC	6127
D.CD.84Z8085AT-GT-CC-T-AT-A-G-T-T.....A-C-GAAC	6242
F1.BF.VI850CGG-A-CC.....G-C-GTAG	6043
G.SE.SE6165G-C-GTAG	6151
H.CF.90CF056GGA-CAGG-G-AC-C-T.....G-A-C-TGAC	6055
J.SE.SE7887G-GTAGTCTT-ATATC-G-C-T.....A-CT-GAA	6037
K.CM.WP535CG-GG-AGGCC-C-A-T.....A-C-GAA	5920
O.AE.TH.CM240GG-ATATA-C-AT-AG-G-TA.....T-G-C-GAAC	6298
02.AG.NG.IBNGGACAT-T-G-A.....A-C-GAAG	6245
03.AB.RU.KAL153-2AA-A-TGATG-A-G.....C-C-GA-C-G	6747
04.CP.XV.94CY032-3G-ATT-A-GA.....G-A-C-GAA	6079
N.CM.WP5180GAC-A-GAT-CAA-C-A-G-T.....T-G-A-C-TGAGC	6317
O.BE.ANT70ATGA-ACA-T-AT-T-TAG.....T-G-A-C-TGTT	6747
CPZ.US.CPZUSC-C-CAA-TG-AACTT-G-G-G-TGAG.....T-G-A-C-TGAA	6738
CPZ.GA.CPZGABATAGCTGG.....T-A-CT-TGAA	6774
CPZ.CM.CM3AGTAGCTG-AAATACTACTGATATA.....T-G-G-CA-GAAT	6785
CPZ.CM.CM5AACAGACACTCTCCGCTCTGNA.....T-G-G-CT-TGAGG	6164
CPZ.CD.CPZANTGTAACACCA-A-ACAAACAACAACAATA.....GTAGATGGATGAAGC-C-G-A-TAAC.....T-G-G-CT-TGAGG	6428
H2A.GW.ALI EnvS_D_P_C_I_K_A_D_N_C_P.....R_G_L_G_D_E_M_V_NAGGGACTAGGGATGA-GAG--GGTCAA	gp120
H2A.GW.ALIGCGATCCATGCAT-A-GC-G-C.....	7166
H2A.SN.HIV2STGATTCCTCATGCAT-CGG-C-G-C-A-A.....GGTTGGGAGAAGA-GAG--GGTC-A	6615
H2A.DE.BENCTTCTA-CTGCAT--A-AC-C-CA--G-A.....GGATTAGGTTAGAGAG--G-TGCA	7197
H2B.GH.D205TGCAT-A-A-ATG-C-G-A-A.....GGACTAGAGAAGGAA--TGCA	7158
H2B.CI.BHOTGTATC-A-ATG-C-G-TG-A.....GGG-TAGACTAG-GAA--G-T-A	7150
H2G.CI.ABT96TGTAT-AT-TAGC-T-T-TA-A.....GGT-TAGACTAG-TG--GGTT--	6516
MAC.US.MM251TGTAT-CTCAG-T-T-TA-A.....GG-TTGAACAAGAGCAA--G-T-A	7085
SWM.US.SIVSMMH9TGTAT-AGA-GT-T-T-TA-A.....GG-TTGAACAAGAGCAA--GGT-A	6577
STM.US.STMTGT-T-AGT-AC-C-TA-A.....GG-TTGAAGAAGA-TCAC--GT--	6745
TANTALUS.UG.TANIGC---TT-T-AGA-T-GTTCT-TA-C.....TGC--GGCAGG-TAT-GG--A	6787
GRIVET.ET.GRI-677GA-CAG-G-A-CCAGC-TCC-A.....TGCA-TGCAGG-TATCA--A	6522
VERVET.KE.VBR-9063GA-C---G-A-CAGCTTC-T-TA-A.....TGC--GGCAGGTTA--T--A	6821
VERVET.KE.AGMTYOGAG---G-T-AGCCCTCT-T-TA-A.....TGC--GGCAGGTTA-T-G-A	6296
VERVET.KE.AGMT55GAG--A-TG-C-AG-CAGCTTC-T-TA-A.....TGC--GGCAGGTTA-T-G-A	6804
VERVET.DE.AGM3GAG-T-AA-TG--AGCCTGCTTC-T-TA-A.....TGC--GGCAGGTTA-T-G-A	6328
SABAEUS.SN.SABICGAG-T-AA-TG--AGCCTGCTTC-T-TA-A.....TGC--GGCAGGTTA-T-G-A	6989
SUN.GA.SIVSUNGAG-T-AA-TG--AGCCTGCTTC-T-TA-A.....TGC--GGCAGGTTA-T-G-A	6945
MANDILL.GA.MNDGB1GAG-T-AA-TG--AGCCTGCTTC-T-TA-A.....TGC--GGCAGGTTA-T-G-A	6217
LHOEST.KE.loeestTT-AG--AGATCTT-T---A-AGAGGGCTTAAIG-G-C--CAGAG--G-ATGC-A	6874
LHOEST.CD.loeest447TT-AT---GAC-T-T-TAT-AAATGGTTAAATG-G-CT-CTGAG-TT-AT---A	5814
LHOEST.CD.loeest485TT-AT---GAC-T-T-TA-AGATGGTTAAATG-G-CT-CTGAG-TT-AT---A	5811
LHOEST.CD.loeest524TT-AT---AGAT-C-T-TA-ACGATGGTTAAATG-G-CT-CAGA--TT-AT---G-A	5817
RCM.NG.SIV-NGMG-C-C---TG-C-TG--T-TCAA.....T-CGG-GGCAGGTTA-A--A	7334
COL.---SIV-CGU1ATTG-A-AG--TATCC-TGTT-T-T-AGA.....A-G---TCA-CAG-TTT-T-CTTAA	6119
SYKES.KE.SYK1173C-GT-CAGAGCCA-G-TTCAA	6673
SYKES.KE.SYK173 EnvN_S_T_E_P_R_F_N	gp120

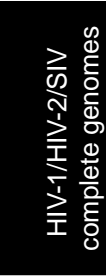
HIV-1/HIV-2/SIV complete genomes

B.FR.HXB2 Env	K V Q K E Y A F F Y K L D I I P I D N N	D T	T S Y	9p120
B.FR.HXB2	TAAGGTGCAGAAAGAAATATGCATTTTTTATAAACCCTTGATATAACCAATAGATAAT...	GATACT	...ACCAGCTAT	6797
A.UG.U455	AAA---T---C---G---A---CA---A...	AC-CA-AAC	...AT---T---	6242
C.ET.ETH2220	AAAAG---G---C---G---C---TA---C...	G-T	...TGAT---	6200
D.CD.84ZR085	AG-A-C---T---C---G---G---A...	GGG-AAAAATGAAATCAAT	...GATACITATGGT---CT---	6339
F1.BE.VI850	AAA---T---TT---C---GC---G---C---	A-C-A-AGC	...GTGAA---	6119
G.SE.SE6165	AAA---T---C---G---C---G---G---TA---	CA	...T---AT---	6221
H.CF.90CF056	CA---T---C---C---G---G---G---I---	A---G---	...TCAG---	6128
J.SE.SE7887	RAG-A-C---G---C---G---AA---G---	A-C-A	...AA-AT---	6110
K.CM.MP535C	AA-A-A---G---C---C---G---C---TA---GGC...	GGG-AATAACAGTAGT	...TGAA---C---	6005
O1.AE.TH.CM240	AA---G---TCC---C---G---G---A---T---AG	A-G-AGACTAGT	...GTGAG---	6377
02.AG.NG.1BNG	AAA---GATGC---C---G---G---A---TA---G-A...	A---A---GGT	...GTCAG---	6321
03.AB.RU.KAL153-2	A---A---A---C---C---G---G---G---A---TA---GC...	AGAGTGCCAAATTAATGGTAGTAAAGAAATAAT	...AGTACGAAGAG---	6814
04.CPX.CV.94CY032-3	AAA-A-C---TT---CT---TC---G---G---GT---GAA---G-G-C...	TCA-A-GCAGTAAAT	...TATAAT-AA-CA---	6185
N.CM.VBF30	AAA-G-C---A---G---TC---A---C---GT---TCA---C-G-GTAAGG-TA-G-C...	ACA-GCAGCACAAATAAGACAAAC	...GGA-A-CA---	6390
O.BE.ANT70	AAA-G---A---C---G---TC---A---C---GT---TCA---T-G-GGA-C-GA-G-G...	ACA-GCAGCACAAATAAGACAAAC	...AGC-AA-TG---	6832
CPZ.US.CPZUS	AAA-A-AC---TG---T---C---A---GT---GA---C---GCT---AT---CA---AAT...		...CA---	6841
CPZ.GA.CPZGAB	AA-A-AC-G-TC---T---C---A---GTGAG---G-GG---AAC---GG---	G-ACAAAC	...CA---	6858
CPZ.CM.CM3	AA-A-A---T---T---TC---A---GT---GA---C---G---TTG---T---	F-GAGG-AATAAGAG--A-GAT	...CA---	6237
CPZ.CD.CPZANT	AAA---A---ATGA---A---CA---A---GGA---C---T---GAAGTGTG-CG-CAACAAT---G---AAC---		...T---	6250
H2A.GW.ALI Env	C R F N M T G L Q R D K P K Q Y N E T T W Y S K D V V E P F N T T N	O T R C		9p120
H2A.GW.ALI	-TGTCG-TTC-CATGACA-G--ACAGAGAG-TAAGCCA-A-CAGTAT-AT-A-CA...	TGGTACTCAAAGATGTGGTTTGTGAACCAITTAACACCACCACAAAC...	CAG---T---R---C---	7284
H2A.SN.HIV25T	CTGTCA-TTC--TATGACA-G--AGAGAGG-TAAGA-A-A-C-TAT-AT-A-CA...	TGGTACTCAAAGATGTAGTCTGTGTAATCAAATGACACCAAGAAAGAG...	...AA-CA-G-	6730
H2A.DE.BEN	ATGT-A-TTC--TATGA-G-GG-AGAAC-AG-TAAGA-A-GG-GGTAT-AG-C-CA...	TGGTACTCAAAGATGTGGTTTGTGACCAACAACA...	...GCTGG--CA-G-	7303
H2B.GH.D205	ATGTAATTTT-GTATGAG-G-C-AAAAGAG-TGAGTCA-A-CA-TAT-A-C-CC...	TGGTACTCAAAGATTTAGAGTGTATAATACAGAGTATACC...	...G-A-GC	7270
H2B.CI.BHO	CTGCAATTC--ATGACA-G--AAAAGAG-TGAGTCA-A-CA-TAT-AG-C-CC...	TGGTAAACAAGATTTAGTGTGAAAGGGGCAAGAGTAAAT...	...GAA-G-AA-G-	7265
H2G.CI.BHT96	CTGCAATTC--CATGACA-GCC-AAAAGAG-T-AGAGA-AGGA-TAC-AT-G-C...	TGGTAA-TCAACAAGATTTGCAATGCAAGTCAATAGAGTGAAGT...	...GAG-GT--A-GC	6634
MAC.US.MM251	CTGTAATTC--CATGACA-GG-AAAAGAG-CAAGACA-AGAGTAC-AT-A-C...	TGGTACTTCAAGATTTGGTTGTGAACAGGAAAT...AGCAGTAAAGAA-G--A-GC	...GAG-GT--A-GC	7203
SMM.US.SIVSMH9	TGTAATTC--CATGACA-GG-AAAAGAG-CAAAA-G-G-GA-TAT-AT-A-CA...	TGGTAA-TCAAGAGATTTAGTGTGAAACAAAATAGCAATGAAATGAAACGAC-GT-AA-GC	...GTCACTGGCGAAGAG-G--A-GC	8698
STM.US.STM	TGTAATTC--TATGACA-G-C-AAAAGAG-TAAA-G-G-GAGTAC-AT-A-CG...	TGGTACTTACAGCATCTAATCTGTGAACAAAAT...	...GTCACTGGCGAAGAG-G--A-GC	8660
TANTALUS.UG.TANI	GTAAAA---T-TA--AGTAC--GG--G-T-AG-GT-GG-TGTGAGA-GG-A...	A-A-C-ACACACA	...GGCACG-GAG-T-G-	8672
GRIVET.ET.GRI-677	GTAAAA---TT-T...AGCA-GACC-GGT-TGA-C-GGAGT-GTCTGCA...	A-A-AGGAAGTAAAAG	...GGAAAGT-AGGAT-G-	6610
VERVET.KE.VBR-9063	C-AA-A-A---TT-T--T-G-AG-A-GG-TGA-C-CAGAG--CTATTGTA-G-C...	AGCT--AGCTCTAACTCT	...AGCAC-AGGAA-GC	6909
VERVET.KE.AGMTYO	C-AA-A-A---TT-T--T-TG-GG-G-GG-TGA-C-CAGA--CTATTGTA-A-	A-G--A...AATAGC	...ACTAGC-AAAGAG-GC	6381
VERVET.KE.AGM155	CC-AAA-A---TT-C...T-G-AG-A-GG-TGA-C-CAGA--CTTTGTA-GCG...	AG--A...TCGCATAAT	...GGGACA-AAAGAG-GC	6889
VERVET.DE.AGM3	C-AAAA---GT-T...T-G-GG-G-GG-TGA-C-CAGA--CATGTTGA-G-G...	G--G...AACAAAT	...TCTAAC-GAGAA-G-	6407
SABAEUS.SN.SAB1C	ATGTAAT-T-T-TGACA-GCC--G-AGAG-TTG-A-A-CGGAGTGA-C-A-C...	GA-AGGAATAACAAC	...CATACTGTGG-A-G-	7083
SUN.GA.SIVSUN	ATGTAAT-T-T-TGACA-GCC--G-AGAG-TTG-A-A-CGGAGTGA-C-A-C...	TT--GGGATACAGACTTACTTGCATGCTGAATGCTGAATGAC...	...CA-G-	7045
MANDRILL.GA.MNDGB1	ATGCA-T-T-T-T-ACA-G-C-A-GCAGGG-TTCGA-GGA-AT-G-ACA--...	TT--GAGGGATGATGTAAGTCTCAGAGAGGGGAGAACAGA...	...CT-GC	6314
LHOEST.KE.roest	ATGCA-T-T-T-T-ACA-G-C-A-GCAGGG-TTCGA-GGA-AT-G-ACA--...	TT--GATATGATGATGTAAGTCTCAGAGAGGGGAGAACAGA...	...CT-G-	6983
LHOEST.CD.roest447	GTGTAGAT-T-T-TTACA-G-C-A-GAGGG-CTG-AGG-C-GA-AT--AGC-G-G...	TT--GGTATGATGACAACTGTGATPAGAAATGGA...	...CT-G-	5914
LHOEST.CD.roest485	ATGTAAT-T-T-T-ACA-G-C-A-GAGGG-CTG-A-GA-AT--AGC-G-G...	TT--GGTATGATGACAACTGTGATPAGAAATGGA...	...CT-G-	5911
LHOEST.CD.roest524	ATGTTCAAT-T-T-TGACA-GCC--G-AGAG-TTCCA-GGA-GAGAT--AGC-A...	TT--GGTATGATGACAACTGTGATPAGAAATGGA...	...CT-GC	5917
RCM.NG.SIV-NGM	G-AAAAAT---T-A--ACACA...GGGT-CTCAAGAGATC-CTGGTGA-GGCA...	G-A-AAATAATACTAGGCAATGT...	...T---GGCA	7413
COL....SIV-CGU1	AG-TAGAA--G-GTT-G-ATT-GGA---CAGT-GAA---T---CTGTGC-T-GA---	A-A-A-GACAGTAAATAGT	...T---GGCA	6204
SYKES.KE.SYK173	-TGCTCTTTT--CTTGACA-GGGGG--CA-GG-TAAAA-ACAGCA-TAT-G-C-TTC...	TTCTA-AGGATGACTTGTGTAAGAAAGAGGCAATAGTAGC...	...TAT---	6779
SYKES.KE.SYK173 Env	C S F N L T G G F K D K Q Y R A F	F Y K D D L M K E G N S S	...Y Y	9p120

Accession	Strain	Genome Type	Position
B.FR.HXB2 Env	B.FR.HXB2	HIV-1	9p120 6927
A.UG.U455		HIV-1	6372
C.ED.ETH2220		HIV-1	6330
CD.84ZRO85		HIV-1	6469
F1.BF.VI850		HIV-1	6249
G.SE.SE6165		HIV-1	6251
H.CF.90CF056		HIV-1	6258
J.SE.SE7887		HIV-1	6240
K.CM.MP535C		HIV-1	6135
L.AE.TH.CM240		HIV-1	6507
02.AG.NG.IBNG		HIV-1	6451
03.AB.RU.KAL153-2		HIV-1	6944
04.CP.XY.94CY032-3		HIV-1	6315
N.CM.MVP5180		HIV-1	6520
O.BE.ANT70		HIV-1	6962
CPZ.US.CPZUS		HIV-1	6971
CPZ.GA.CPZGAB		HIV-1	6988
CPZ.CM.CM3		HIV-1	6367
CPZ.CM.CM5		HIV-1	6637
CPZ.CD.CPZANT		HIV-1	6380
H2A.GW.ALI Env	H2A.GW.ALI	HIV-2	9p120 7414
H2A.SN.HIV2ST		HIV-2	6860
H2A.DE.BEN		HIV-2	7433
H2B.GH.D205		HIV-2	7400
H2B.CI.BHO		HIV-2	7395
H2G.CI.ABT96		HIV-2	6764
MAC.US.MM251		HIV-2	7333
SMW.US.SIVSMW9		HIV-2	8628
STM.US.STM		HIV-2	6990
TANTALUS.UG.TANI		HIV-2	7002
GRIVET.ET.GRI-677		HIV-2	6740
VERVET.KE.VBR-9063		HIV-2	7039
VERVET.KE.AGMTYO		HIV-2	6511
VERVET.KE.AGMT55		HIV-2	7019
VERVET.DE.AGM3		HIV-2	6537
SABAEUS.SN.SAB1C		HIV-2	7213
SUN.GA.SIVSUN		HIV-2	7175
MANDRILL.GA.MNDGB1		HIV-2	6441
LHOEST.KE.roest		HIV-2	7113
LHOEST.CD.roest447		HIV-2	6044
LHOEST.CD.roest485		HIV-2	6041
LHOEST.CD.roest524		HIV-2	6047
RCM.NG.SIV-NGM		SIV	7543
COL.SIV-CGU1		SIV	6334
SYKES.KE.SYK173		SIV	6909
SYKES.KE.SYK173 Env		SIV	9p120

HIV-1/HIV-2/SIV complete genomes

loop tip
B.FR.HXB2 Env
B.FR.HXB2
A.UG.U455
C.ET.EH2220
D.CD.84Z8085
F1.BE.V1850
G.SE.SEG165
H.CF.90CF056
J.SE.SE7887
K.CM.MP535C
01.AE.TH.CM240
02.AG.NG.IBNG
03.AB.RU.KAL153-2
04.CPX.CV.94CY032-3
O.CM.MVP5180
O.BE.AN710
CPZ.US.CPZUS
CPZ.GA.CPZGB
CPZ.CM.CM3
CPZ.CM.CM5
CPZ.CD.CPZANT
H2A.GW.ALI Env
H2A.GW.ALI
H2A.SM.HIV2ST
H2A.DE.BEN
H2B.GH.D205
H2B.CI.BHO
H2G.CI.BBT96
MAC.US.MM251
SMM.US.SIVSMM49
STM.US.STM
TANTALUS.UG.TANI
GRIVET.ET.GRI-677
VERVET.KE.VBR-9063
VERVET.KE.AGMTYO
VERVET.KE.AGM155
VERVET.DE.AGM3
SABAEUS.SN.SAB1C
SUN.GA.SIVSUN
MANDRILL.GA.MNDGB1
LHOEST.KE.hoest
LHOEST.CD.hoest447
LHOEST.CD.hoest485
LHOEST.CD.hoest524
RCM.NG.SIV-NGM
COL.SIV-CGU1
SYKES.KE.SYK173
SYKES.KE.SYK173 Env
gp120
gp120

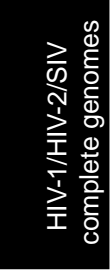


B.FR.HXB2 Env	I_A_S_K_L_R_E_O_F_G_N_N_K_T_I_I	9p120
B.FR.HXB2	TAGCTAGCAAAATTAAGAGAAACAATTTGGAAATAATAAACAATAATC.....	7304
A.UG.U455	GAAC--AGA--A-G--AAT--C-A-C-	6743
C.ET.ETH2220	AAAGAA--CA-A-G-C--CCCT--	6695
D.CD.84ZR085	A--G-A-C-TT-GAAC-C-CA-	6834
F1.BF.VI850	AAAGCAG--AGTCG-T--CCCT--	6617
G.SE.SE6165	C-GCAG--C-G-GA-ATC--AAT--G-GC--GCAGAG-AC	6728
H.CF.90CF056	-T--CAC--G--AT--C--GAC--	6623
J.SE.SE7887	-CA--C--	6602
K.CM.MP535C	AAAGAA--GC--A--T--AAG--GGA-C-T-	6503
01.AE.TH.CM240	A--GAA--G-C--AAT--G-CA-	6869
02.AG.NG.IBNG	T--CAC--AGACGT-C--AAG-C-CC-C-	6816
03.AB.RU.KAL153-2	-T--T--A--G--A--C--T-	7312
04.CP.XV.94CY032-3	AG-GAAG--G-A-AG-TC--CCCT--A-C-T-A-	6686
N.CM.VBP30	C-AGAGAG--A-AGA-ATCC-G-G-A-C-C-	6875
O.CM.VBP180	-TA-GG-ATTT-A-T-TTG-AAACC-A-CAG-G-AT	7344
O.BE.ANT70	GAA--GG-ATTT--T-G-AAAC--CAGGT-GT	7335
CPZ.US.CPZUS	CTTTTA--GAG-T-AT-A--AAG-A-GCA--CCTG--AGG-AGAGTTAATCCA..AAT	7345
CPZ.GA.CPZGAB	AAA--AGGCT--GC-ACCTC--CAAUCGA-CAGC-G-C-AC-AGACTCAACAGA..GCC	7377
CPZ.CM.CM3	CT-G--CAGCC--GC--G-TT--AC--CCT--C-CA-G-C-A-TTT..ACAGTA..AGT	6744
CPZ.CM.CM5	CC--TTAGCCA--...C-A-G-A-AT--A-GC-T-G--A-G--ACCTTTACACCA..AAC	7020
CPZ.CD.CPZANT	C--AGCA-GTTC-TGCG--G--T-GGAA--AGTAG-C-AC-A-C-AAT	6757
H2A.GW.ALI Env	V_K_E_T_L_V_K_H_P_R_Y_K_G_T_N_D_T	9p120
H2A.GW.ALI	GAAGAA-CCC-TGT-A---TCCCA-GT---AGG---C-ATGA-ACA.....N_Q	7790
H2A.SN.HIV2ST	GAAGCTA-CCC-TGC-A---TCCCA-GT---AGG---C-ACGA-ACA.....GAAAA	7236
H2A.DE.BEN	GAAGAA-CCC-TGT-C---TCCCA-T-C-AGG--TC-ATGATACA.....GGGA	7809
H2B.GH.D205	GAA--G-CCA-C-T-A---TCCCA-GT---AGG-GGTGC--AAAT..ATCACA	7782
H2B.CI.BHO	GAAGAG-CCA-T-A-A-T---TCCCA-T--TCAGG---C--ATATC	7768
H2G.CI.ABT96	GAAGRA--CG-GTCA-G---CCCA-CT---AGG---ATGA-AGA	7140
MAC.US.MM251	GAACAG-CCA-TGTCA---TCCCA-GT---C-GG--T-AC-ATACT	7706
SMW.US.SIVSMW9	GAAGGAMCC--GGTCA---TCCCA-GT---GGG--T-ATRAKAYT	7201
STM.US.STM	TAAGAA-CC--GGTCA---TCCCA-GT---CAGG--C-ATGA-AGC	7363
TANTALUS.UG.TANI	GAGA-AA-C-A-GTG--G-TTCCAA-G-G-AT-TCG-GGG-CAAT	7381
GRIVET.ET.GRI-677	CAGAGAG-G-AGA-GTGAATACT--CAG--GT...G-ATA	7116
VERVET.KE.VBR-9063	TAAGAG-A-TCA--TT-CCAA-G--GGT-TCGGGG--C-AAT	7421
VERVET.KE.AGMTYO	AAACA--A-GT--GTT-CCAA-G-C-GGT-T-A-GG--C-AAT	6893
VERVET.KE.AGM155	CAAGAGG-A-GT-A--TT-CCAA-G-ACGGT-CCA-GGC--C-AAT	7401
VERVET.DE.AGM3	AAA-ATG-A-GT-A--TT-CCAA-G--GAT-CCA-GG--C-AAT	6919
SABAEUS.SN.SAB1C	CAAGRA-C-A-GTGAG--T-CCAC--A--GT-T-GTGC-CAAT	7592
SUN.GA.SIVSUN	--GG-RACT--G--GC-GGTGACCTTAGC--C-TG-GT-ATGCTGT..ACATTT..GATAATATAAAGACATGCAGATTACC	7629
MANDRILL.GA.MNDGB1	--GC-RA-TGC--AG--GTTGAA--C-G-AGCC-TG-ATTAT-CAGAGGACGAGGACATGATCTTAAGAAACACTTGT	6913
LHOEST.KE.roest	--GA-A--TGC-T-A-AGG-ITAA--C--CAGCG-TG-ATTAT-CAAT..GGACAA..TGATAGATAACACAAAGCCTTGT	7555
LHOEST.CD.roest447	--GG-AA-TGC-TCA-A-G-ITAA--C--CAGCA-TG-ATTAT-CAAT..GGACAA..TGCTTAGGGAAACAGCCTTGT	6480
LHOEST.CD.roest485	--GA--G-TG--AGA-G-ITAA--AG-C-GC--TG-ATTAT-AGAA..GGAAAT..TGTTCCACACCAAGAAACCCATGT	6477
LHOEST.CD.roest524	AAAGAGGC-A-T-ACA-GAC-AAA--T--AGGT--CG-CAGCCT	6483
RCM.NG.SIV-NGM	CAAA-AG--T--C--CTTGG...CT-G-GCGC-TTT--CGC-AAAATA	7910
COL....SIV-CGU1		6737
SYKES.KE.SYK173	--CA-GAAC--GGCC--CGA--ACC...T-G-A--GT--C-AC	7286
SYKES.KE.SYK173 Env	V_H_E_Q_A_T_K_T...W_K_N_V_T_N	9p120

HIV-1/HIV-2/SIV complete genomes

B.FR.HXB2 EnvF K Q S S G G D P E I V T H S F N C G G E F F Y C N S T Q L F N S T W F N S T W	9p120
B.FR.HXB2TTAAAGCAATCCTCAGGAGGAGCCAGAAAATTGTAAACACACAGTATTTTCTACTGTAAATTCACACACAGTGTAAATAGTACTTTGGTAAATAGTACTTTGG	7424
A.UG.U455ATAATC---CCTAGC---TAT---AC---A---T---C---T---C---CA---T---GGC---C---T---AA---GGC---C---TG---CA	6869
C.ET.ETH2220ATAGAA---C---T---T---AC---A---T---C---T---C---A---A---C---T---A---C---C---AAAAC---AG---ACTGTT---AAT	6821
D.CD.84ZR085ATRAAT---A---C---T---AC---A---T---C---C---A---T---A---C---G---A---AA---TATCAGACAT	6960
F1.BF.V1850ATRAAA---C---T---AC---T---T---AC---T---A---A---A---G---A---G---G---C---AG---C---CC---A---GGCACT	6743
G.SE.SE6165ATRACG---CTC---TG---T---AC---A---T---C---T---A---A---G---A---G---G---C---TCA---G---GATCA---A---T	6854
H.CF.90CF056ATRACG---C---AA---AIG---G---A---A---T---A---T---C---A---A---T---GGG---T---C---GAA---TGCA---A---AAT	6749
J.SE.SE7887ACATTTAAACA---C---AA---C---A---T---A---A---A---A---A---A---A---G---AG---GAG---G---A---GCAT	6728
K.CM.MP535CATRAATC---C---A---C---T---AC---T---TCA---A---T---C---A---A---A---A---A---CC---AGGA---A---GAAAC	6995
02.AE.TH.CM240ATRAATC---GCTA---CC---T---TGT---T---AC---T---A---A---T---C---A---T---A---C---G---A---C---A---A---C	6942
03.AB.RU.KAL153-2GTC---T	7429
04.CP.XV.94CY032-3GCT---C---G---T---AC---A---T---C---AA---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A	6806
N.CM.MVP5180	AACCTTCAGGC---CGAG---GAGGAA---A---T---G---G---G---A---G---G---A---G---G---A---G---G---A---A---A---A---A---A	7005
O.BE.ANT70	TACCATATA---C---GCAG---A---TAGT---T---A---TG---G---A---G---C---A---T---A---C---C---A---T---A---A---A---A	7474
P.BE.ANT70	TAACATGACA---C---T---CAG---AGC---T---A---T---T---G---G---A---G---G---A---G---G---A---G---G---A---A	7462
CPZ.GA.CPZGABG---T---A---T---G---G---G---A---G---G---A---G---G---A---G---G---A---G---G---A---A	7441
CPZ.US.CPZUST---A---T---G---G---C---A---T---T---TG---A---A---A---A---A---A---A---A---A---A---A	7467
CPZ.CM.CM3C---C---A---A---T---G---G---A---G---G---A---G---G---A---G---G---A---A	6852
CPZ.CM.CM5C---C---A---A---T---G---G---A---G---G---A---G---G---A---G---G---A---A	7128
CPZ.CD.CPZANTGCGAACAACA---ATG---AC---T---CA---A---T---T---G---G---A---A---G---T---TT---G---AT---G---AT---G---AT---G---AT	6880
H2A.GW.ALI Env	I N F T K P G R G S D A E V V Y M W T N C R G E F L H C N M T W F L N W V E N K T G O	9p120
H2A.GW.ALI	AAITTAACCTTTTACA---A---C---GGAAAG---CTCA---TG---G---G---TATAATGT---GAC---C---CA---A---A---C---C---CATG---TTGGT---CC---C---T---GGTGGAAAC---A---CGGG---CA	7920
H2A.SN.HIV2ST	AAITTCGTTTTTA---AGC---C---GGAGA---C---CTCA---TG---G---C---TACATGT---GAC---C---CA---A---A---C---C---ATG---TTGGT---CC---C---T---GGTAGAAAAC---GA---CG---A---CA	7366
H2A.DE.BEN	AAITTAACCTTTTACG---A---C---GGGAG---CTCA---TG---G---G---C---TTTATGT---GAC---C---CA---A---A---C---C---CATG---TTGGT---CC---C---T---GGTAGAAAAC---G---ACCAACA	7939
H2B.GH.D205	AAAGTTAGTA---CAG---A---TGGAA---TTCA---TG---CAAC---TATATCT---GACC---A---A---C---A---C---CA---A---C---CTG---CTGGT---C---A---CT---GGTAGA---AAC---A---CA---A---ACA	7912
H2B.CI.BHO	GATRAAGCTTAGCAG---CG---GAG---A---CTCA---T---G---AAG---TATATCT---GAC---C---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A	7898
H2G.CI.ABT96	MATAACTTTTACA---CA---C---GGAGA---TTCA---T---G---AAGTTAATCT---GAC---CA---A---G---C---G---T---AATG---TTGGT---TC---A---T---G---TAGA---CAC	7258
MAC.US.MM251	TAAAATCAA1---A---C---GCTC---TGS---A---T---G---ACCTTCAATG---GACA---CA---A---A---G---CC---A---G---CC---A---A---G---CC---A---A---G---CC---A---A---G---CC	7836
SMM.US.SIVMMH9	ACRAATTAAGC---A---CAGCTC---AGG---A---T---G---ACCTTCAATG---GACA---CA---A---A---G---CC---A---A---G---CC---A---A---G---CC---A---A---G---CC	7331
STM.US.STM	AAAATAAAGA---AGT---GCTC---TGGG---A---T---T---GG---CACCTTCAATG---GACA---CC---T---A---C---A---T---C---AATG---AATGGT---TC---A---T---GGTAGAAGY---TA---AA---A---G---T	7493
TANTALUS.UG.TANI	AAGGCAGATT---GGCTAAGTAGACA---T---G---A---T---GCA---CT---ACATTT---GC---CCA---A---A---G---C---T---TGAT---TGCT---C---A---T---G---TAGAAAAC---GA---A---A---A---A---A	7511
GRIVET.ET.GRI-677	AAATATATACATC---G---GAAG---..AT---T---G---A---T---TCA---CG---AFTTTT---G---CCA---A---A---G---T---GATG---GACTGGT---TA---C---TA---CTAAACAA---CGA---CAGAAAGAT	7243
VERVET.KE.VBR-9063	AGAGAAATA---A---CT---GAGACAGTTTT---A---T---GCA---C---AFTTTT---G---CCAG---A---G---C---G---AATG---GACTGGT---TC---A---TA---TTGAATAA---C---ATCAGTAGAT	7551
VERVET.KE.AGMTYO	AGAACACATA---ACCT---AGACA---T---G---A---GCA---C---AFTTTT---G---CCAG---A---G---C---T---T---GATG---GACTGGT---TC---A---TA---CTGAATAA---A---CATGGGAT	7023
VERVET.KE.AGM155	AAAACAATACT---TT---AGACA---TTT---A---T---GCA---C---AFTTTAT---G---CAG---A---C---T---AATGGATGGT---TC---A---TACCTGAATAA---CG---CAGTAGAT	7531
VERVET.DE.AGM3	TGAAGAGATT---A---CT---GAGACT---TTT---A---T---TC---A---TTTCTT---TT---CA---T---C---T---AATGGATGGT---CC---CTA---TTGAATAA---GTCAGTAGAC	7049
SABAEUS.SN.SAB1C	AAACAAGATT---CTA---AGACA---T---G---A---T---CA---T---C---T---A---A---G---C---A---T---G---A---G---T---G---T---G---T---G---T---G---T---G---T---G---T---G---T---G---T	7722
SUN.GA.SIVSUN	TAAA---CCA---T---G---GAGCAT---CA---TG---C---CA---AG---T---TTG---TGA---G---CA---CA---T---A---G---T---A---G---T---A---G---T---A---G---T---A---G---T---A---G---T	7753
MANDRILL.GA.MNDGB1	TCTTTAAGATAGA---TT---TA---TA---TC---A---TAG---C---GCA---A---T---AUG---TGA---G---CA---CA---T---A---G---T---A---G---T---A---G---T---A---G---T---A---G---T	7043
LHOEST.KE.loeest447	ACCRATAGCTTAA---T---ACTAGGCAT---CA---TT---C---CA---AG---T---TTA---TGCA---CA---AA---A---A---G---T---C---CA---AG---T---TTG---TGCA---CA---AA---A---A---G---T---C	7685
LHOEST.CD.loeest448	ACCTATAGCAAAA---T---AC---AGAAG---G---AA---T---T---C---CA---AG---T---TTG---TGCA---CA---AA---A---A---G---T---C---CA---AG---T---TTG---TGCA---CA---AA---A---A---G---T---C	6610
LHOEST.CD.loeest485	ACCCTATAGCAAAA---T---ACGAGAAG---TA---T---TC---C---CA---AG---T---T---TGCAC---CA---AA---A---A---G---T---C---CA---AG---T---TTG---TGCA---CA---AA---A---A---G---T---C	6613
RCM.NG.SIV-NGM	GAATATAACTA---A---GATC---GT---AT---GAT---GGCAAA---TGTTTT---G---A---AT---A---A---G---T---GAT---GGCAAA---TGTTTT---G---A---AT---A---A---G---T---GAT---GGCAAA	8040
COL.---SIV-NGU1	TATCACTCCRAGG---TGG---AATAG---AC---A---T---T---G---A---G---G---C---ACTTTCCAC---TT---G---G---T---CATA---TT---C---T---G---T---TAGT---A---GG---ACA---GA---CC	6852
SYKES.KE.SYK173ACCACA---GG---G---TCCCAGC---G---A---T---T---GG---ACG---A---A---TT---G---CC---A---G---A---G---C---GT---T---TA---G---C---GC---AC---T---AC---AA---GGA---ACG---AA---T	7412
SYKES.KE.SYK173 EnvT T W R S Q P G G D L E V R T H W F Q C G G E F Y C N V S K L F A N I T N G N A S	9p120

B_FR.HXB2 Env 9p120
B_FR.HXB2 7527
A_UG.U455 6957
C_ET.ETH2220 6903
D_ED.84ZRO85 7048
F1_BE.VI850 6807
G_SE.SE6165 6933
H_CF.90CF056 6846
I_SE.SF7887 6822
K_CM.WP535C 6690
L_AE.TH.CM240 7080
M_AG.NG.IBNG 7036
N_AB.RU.KAL153-2 7523
O_CX.CV.94CY032-3 6903
P_CM.YBF30 7075
Q_BE.ANT70 7583
R_US.CPZUS 7565
S_GA.CPZGB 7523
T_CM.CM3 7546
U_CM.CM5 6934
V_CD.CPZANT 7246
W_CD.CPZANT 6962
X2A.GW.ALI Env 9p120
X2A.GW.ALI 7993
X2A.SN.HIV2ST 7439
Y2B.DE.BEN 8003
Z2B.GH.D205 7985
AA2B.CI.BHO 7968
AB2G.CI.ABT96 7364
AC2B.US.IVV251 7930
AD2B.US.SIVSMH9 7437
AE2B.US.STM 7602
AF2B.US.TAN1 7641
AG2B.US.GRI-677 7370
AH2B.US.VER-9063 7681
AI2B.US.VR-9063 7150
AJ2B.US.AGMTYO 7661
AK2B.US.AGM3 7179
AL2B.US.SAB1C 7837
AM2B.US.SUN 7826
AN2B.US.SUN 7116
AO2B.US.MNDGB1 7758
AP2B.US.LHOEST 6683
AQ2B.US.LHOEST 6680
AR2B.US.LHOEST 6686
AS2B.US.RCM.NG.SIV-NGM 8128
AT2B.US.COL. SIV-CGJ1 6916
AY2B.US.SYKES.KE.SYK173 7497
AZ2B.US.SYKES.KE.SYK173 Env 9p120



Accession	Strain	Genome Sequence
B.FR.HXB2 Env	B.FR.HXB2	Y_A_P_P_I_S_G.....Q_I_R_C_S_S_N_I_T_G_L_L_T_R_D_G_G_N.....S_N_N_E_S_E_I_F_R_P ATGCCCTCCATCAGTGGGA.....CAAAATTAGATGTTCAAAATATTACAGGCTGCTATTAAACAAGAGATGGTGGTAAT.....AGCAACAATGATCCGAGATCTTCAGACC
A.UG.U455		---CAA---A-G-GA---C---A-A---A-A---G---C-AAC.....T-CAA-AAAT---C-
C.ET.ETH2220		---TGAA---AT---A-TG---AG---C---A-A---C-G---C-T---A-CC-A.....GAGCCAC-T-GCACAAAA---A-
D.CD.842R085		---TGAA---G---AT---A-TG---AG---C---A-A---G---G---GC.....AT-T-C-C-AAT---C-C-T-
F1.BE.V1850		---TGCA---A-C---CA---AAC---GC---A-AA---G---TT---C-AG---C.....AACAAATACA-AT-CA-G---AGT-
G.SE.SE6165		---GCA---A-C---TG---GT---A-AA---G---TT---C-AG---C.....GGCTCTGCA---AAAT-T-C-
H.CF.90CF056		---TGCA---A-C---CG---CA---AA---A-A---G---C-T---A-A---CAGG.....GGGAATGGC-TG-G---GCA-T---C-
J.SE.SE7887		---TGCA---AGT---AC---AGC---AA---A-A---G---GT---AAT.....AT-T-C-C-TAAT---C-
K.CM.MP535C		---CAA---AT---A---GAT---C---A---G---C.....CAG-G---TTA---T---A
01.AE.TH.CM240		---GCA---AGC---AC---CAGC---G---AA---AA---G---C.....A-TA-T
02.AG.NG.IBNG		---A-A---A-T-C-G---GTTC---AC---A-C---C---T---A---A---G---TC---GA---TATAG---GCC---G---CCA-GAAACA---AG-ATAT-
03.AB.RU.KAL153-2		---A-A---A-T-C-G---GTTC---AC---A-C---C---T---A---A---A---TC---CACTT---CAACATGG.....ATTG---C---CA-GTGAAT---CAC-T-
04.CPX.CV.94CY032-3		---A-A---A-T-C-G---GTTC---AC---A-C---C---T---A---A---A---TC---CACTT---CAACATGG.....AAC---G---CA-CAAT-TA-CA-T-
N.CM.MVP5180		---A-A---A-T-C-G---GTTC---AC---A-C---C---T---A---A---A---TC---CACTT---CAACATGG.....AATAAC---G-T-AG-ATAATA-CAG-C-TG-
O.BE.ANT70		---A-A---A-T-C-G---GTTC---AC---A-C---C---T---A---A---A---TC---CACTT---CAACATGG.....AAC---G---G-TAA-CTL-CA-T-
CPZ.US.CPZUS		---A-A---A-T-C-G---GTTC---AC---A-C---C---T---A---A---A---TC---CACTT---CAACATGG.....AAC---G---G-TAA-CTL-CA-T-
CPZ.GA.CPZGAB		---A-A---A-T-C-G---GTTC---AC---A-C---C---T---A---A---A---TC---CACTT---CAACATGG.....AAC---G---G-TAA-CTL-CA-T-
CPZ.CM.CM3		---A-A---A-T-C-G---GTTC---AC---A-C---C---T---A---A---A---TC---CACTT---CAACATGG.....AAC---G---G-TAA-CTL-CA-T-
CPZ.CM.CM5		---A-A---A-T-C-G---GTTC---AC---A-C---C---T---A---A---A---TC---CACTT---CAACATGG.....AAC---G---G-TAA-CTL-CA-T-
CPZ.CD.CPZANT		---A-A---A-T-C-G---GTTC---AC---A-C---C---T---A---A---A---TC---CACTT---CAACATGG.....AAC---G---G-TAA-CTL-CA-T-
H2A.GW.ALI Env	H2A.GW.ALI	Y_L_P_P_R_E_G.....E_L_T_C_N_S_T_V_T_S_L_I_A_N_I_D_T_D_G.....N_Q_T_N_I_T TTG---T-GGAA---G-GT-G-CC-CAAC---CAG-A-CA-CT-A-TGCT-AC-TT-CACG-A-GGC.....AA-C---CAAAT-TTA-
H2A.SN.HIV25T		---GT-A-CC-CAAC---T-CAG-G-CA-CA-AA-TGCT-AC-TT-C-A-AG-G.....AA-C---CAAAT-TTA-
H2A.DE.BEN		---GT-A-CC-CAAC---T-CAG-G-CA-CA-AA-TGCT-AC-TT-C-A-AG-G.....CG-A-TC-T-C-AA-TTA-
H2B.GH.D205		---G-GC-TCC---AAT-C-G-G-TA-C-CA-GCC-AC-TCA-CTCA-AC.....A-CAGTACC-C-AAT-TTT-
H2B.CI.BHO		---G-C-CTCC---AAT-C-G-G---CA-C-CA-GCC-AC-TT-CT-GATAG---AAG.....AA-C-TT-CTAAT-TTA-
H2G.CI.BHT96		---G-TC-G-CG---AAT-C-CAG-C-CA-C-CA-GCC-AT-T---A-AGNA-GGCTAAC.....AA-C-T-GAAAT-TCAT
MAC.US.MM251		---G-CC-G-CG---AAC-C-CAG-G-CA-T-CA-GC-AC-T---T-GAC-G-GGA.....AA-C-A-CTAGT-TCA
SMW.US.SIVSMW9		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
STM.US.STM		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
TANTALUS.UG.TANI		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
GRIVET.ET.GRI-677		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
VERVET.KE.VBR-9063		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
VERVET.KE.AGMTYO		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
VERVET.KE.AGM155		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
VERVET.DE.AGM3		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
SABAEUS.SN.SAB1C		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
SUN.GA.SIVSUN		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
MANDRILL.GA.MNDGB1		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
LHOEST.KE.roest447		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
LHOEST.CD.roest448		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
LHOEST.CD.roest524		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
RCM.NG.SIV-NGM		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
COL.....SIV-CGU1		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
SYKES.KE.SYK173		---G-CC-G-CG---AAC-C-C-G-A-TA-C-CA-GC-CA-T---T-GATC---GGC.....AATA---C-AAT-TCA
SYKES.KE.SYK173 Env		Y_L_P_P_T_A_G.....H_I_K_C_T_S_N_V_T_A_V_L_T_D_I_E_Y_Y_P_G.....T_A_C-T-GAA-TTTA-

Rev responsive element secondary structure in RNA, see Charpentier, J Mol Evol 266:950-956(1997)
see HIV-1 complete genome annotation in Compendium p. 318-325 for stem-loop positions
B.FR.HXB2 Env
B.FR.HXB2
...
9p120 7748

Rev responsive element secondary structure in RNA, see Charpentier, J Mol Evol 266:950-956(1997) see HIV-1 complete genome annotation in Compendium p. 318-325 for stem-loop positions

Table with 2 columns: Accession IDs (e.g., B.FR.HXB2 Env, A.UG.U455, H2A.GW.ALI Env) and Nucleotide Sequences (e.g., AGCAGCAGAACAAATTGGCTG...AGGCGTATTGAGGGCCAGCAAGCATCTGTTGCCAACTCACAGTCTGGGGCATCCAGGCAAGATCCTGGCTGFGAAAGATACCTTAAGAGGATCAACAGCT)

HIV-1/HIV-2/SIV complete genomes

Accession	Strain	Genome
B.FR.HXB2 Env	B.FR.HXB2	L G I W G C S G K L I C T T A V P W N A S W S N K S L E O I W N H T CCTGGGATTTGGGGTTGCTCTGGAATACTCAATTCAGCACACTGCTGGATTCAGTGGAGTAATAAATCT.....CTGGAAACAGATTTGGAAATCACACG
A.UG.U455		---A-A---C---G---C---A---C---CT---A---GG-C-A---A---G-C-A---A---T
C.ET.ETH2220		---A---C---C---C---A---A---C---CT---AA---GG---G---A---T
D.CD.84ZB085		---A---A---A---A---A---C---CT---G---A---CT---GGG---A---T
F1.BE.VI850		---A---C---C---C---A---A---C---CT---A---G---G---GGG---A---T
G.SE.SEG165		---A---C---C---C---A---A---C---CT---TATA---TG---A---G---A---T
H.CF.90CF056		---A---C---C---C---A---A---C---CT---A---A---AGT---A---C---G---CA---T
J.SE.SE7887		---A---C---C---C---A---A---C---CT---TAT---GG---C---G---GA---T
K.CM.MP535C		---A---AC---C---C---A---A---C---CT---TCTTG---G---CA---T
01.AE.TH.CM240		---A---A---C---C---A---A---C---CT---T---G---CA---T
02.AG.NG.IBNG		---A---A---C---C---A---A---C---CT---T---TA---TG---C---A---G---A---T
03.AB.RU.KAL153-2		---A---A---C---C---A---A---C---CT---A---A---TG---A---A---T
04.CPX.CV.94CY032-3		---A---A---C---C---A---A---C---CT---TATA---TG---T---A---G---CA---T
N.CM.YBF30		---AA---TC---A---C---A---ACA---A---TAT---CA---AG---C---C---C---TA---CTCT.....TATA---TG---T---A---G---CA---T
O.CM.MVP5180		---AAAC---A---C---TAA---A---C---TTA---AT---A---AAAA---CA---ATCA---TCAGGA.....AGATATAATGAT---CAGT---G---CA---CTT
O.BE.AN70		---AA---CC---A---C---TAAA---G---AG---C---TA---AT---A---AAAA---AGA---CA---TAGGA.....AAC---AGC---G---CAC---TTA
CPZ.US.CPZUS		---A---TC---G---C---A---ACA---TTAT---CA---G.....A---CC---C.....CTCTCTAT---TGCT---GGCA---TGTA
CPZ.GA.CPZGAR		---A---CC---G---C---A---GGCTG---C---TTAT---CA---G.....CAACTC---CC---GGG---AGC.....AATCCACA---TG---C---GGG---A---TTTA
CPZ.CM.CM3		---A---CC---A---C---A---GG---C---TTAT---A---G.....AA---C---GC---T.....ACATCTAT---TG---C---G---A---TTTA
CPZ.CM.CM5		---AA---TC---A---C---A---GG---C---TTAT---CA---G.....GC---C---C---GC---C.....ACATCTAT---TG---C---G---A---TTTA
CPZ.CD.CPZANT		AT---A---CC---C---A---TG---AC---GG---G---CC---TCA---A---G.....AA---TCC---GTA---CTTCA---GCAAAACATGTGCAAAAGAACAGCATGATA---AC---TGT---G---AA---T---T
H2A.GW.ALI Env	H2A.GW.ALI	L N S W G C A F R O V C H T T V P W N N S L K P D W D N M G---AAAATTCAG---A---TG---ATTT---G---AAG---C---CA---A---A---A---GTAAA---AC---CC.....T---AA---CTGA---G---CA---T
H2A.SN.HIV2ST		A---AAAATTCAG---A---TG---GTTT---G---AAG---C---CA---A---A---A---GTAAA---GACACC.....T---AACG---CTGA---CA---T
H2A.DE.BEN		A---AAAATTCAG---A---TG---GTTT---G---AGG---C---CA---A---A---G---GTAAA---GAC---CT.....T---ATCG---CTGAC---AA---T---T
H2B.GH.D205		A---AAAATTCAG---A---TG---TTC---G---AGG---C---TCA---GA---A---A---CCAAA---GAAAC.....CAC---CA---A---A---T
H2B.CI.BHO		A---AAAATTCAG---A---TG---TTC---G---AGG---C---TCA---GA---A---A---GTAAA---GAA---CC.....TA---G---CAGAC---A---T
H2G.CI.BBT96		G---AAAATTCAG---A---TG---GTTT---G---AGG---C---CA---A---A---G---CCT---G---TGC.....T---G---CAGAA---A---T---T
MAC.US.MM251		RN---AAAATTCAG---A---TG---GTTT---G---AAG---C---CA---A---A---A---CCAAA---GCAA---T.....AAC---CAGAC---CA---TGAT
SMW.US.SIVSMH9		G---AAAATTCAG---A---TG---GTTT---G---AGG---C---CA---A---K---A---A---CCAAA---GACAA.....T---AC---CT---RC---CA---T---T
STM.US.STM		G---AAAATTCAG---A---TG---ATTC---AGG---G---TCA---A---G---C---TTC---ACAAT.....T---T---C---GA---CA---T---T
TANTALUS.UG.TANI		G---AAAATTCAG---A---TG---GT---G---AAG---A---TCA---CA---A---A---GTA---A---AAC.....AC---CAGAC---G---CA---T---T
GRIVET.ET.GRI-677		AT---AAAG---C---G---G---AT---G---AGG---C---TCAT---CA---A---A---C---ATGGCAAAAC.....A---ACG---CC---A---C---AA---T
VERVET.KE.VBR-9063		A---AAAATTCAG---G---G---GT---G---AAG---A---TCAT---AA---A---A---GAG---CCCTGG---CAAAT.....G---ACT---C---GA---C---AA---T---T
VERVET.KE.AGM155		GT---AAAAGC---G---G---AT---G---AAG---C---TCAT---AA---C---A---G---C---GTGG---A---AAT.....AC---CAGAG---A---T---T
VERVET.DE.AGM3		GT---AATGC---A---TG---TTC---GG---AGG---G---TCA---GA---G---C---T---GTA---A---AAC.....AG---ACC---CTGA---A---T---T
SABAEUS.SN.SABLC		TT---A---CTTC---G---TCAAT---G---AAG---A---CAT---AA---C---C---T---GTA---A---AAC.....A---ACT---CGAC---G---AA---T---T
SUN.GA.SIVSUN		G---TCTCAA---G---T---AT---GGC---AGG---G---TCAT---T---AGAG---C---AA---CAA---C.....A---CAGC---CC---A---C---A---CGAC
MANDRILL.GA.MNDG1		---T---CCTC---G---TCAAT---G---AGG---G---TCAT---AA---GAG---CAFA---A.....A---CAC---CC---A---C---A---CATCAGAA
LHOEST.KE.hoest		G---T---CCTC---G---TCAAT---G---AAG---A---TCAT---AA---AGAA---CTA---A.....A---AACT---CT---A---CAAGGGAT
LHOEST.CD.hoest447		G---T---CCTC---G---TCAAT---G---AA---A---TCAT---AA---GAA---TA---A.....A---AACT---CT---A---CAAGGGAT
LHOEST.CD.hoest485		A---T---CCTC---G---TCAGT---G---AAG---G---TCAT---AA---A---A---TA---A.....G---TACT---CT---A---CAAGGGAT
RGM.NG.SIV-NGM		AT---AAAT---CA---TT---C---TG---AT---G---AAG---G---TCA---AA---G---GAG---TATA.....AAAC---CAGAG---CA---ACA
COL.---SIV-CGU1		G---T---CC---GCATT---G---TG---AAAC---TG---AA---T---GA---AT---C---CAAG---CC.....GGAGA---GG---TCCA---C---GA---T---T
SYKES.KE.SYK173		---T---TCA---A---A---G---ATTCC---G---AG---C---CAT---A---A---A---G---G---GAAAAGCA---TG---A---C---C---CAAC.....TTCTGT---CCA---CACAG---AA---T---T
SYKES.KE.SYK173 Env		L S N W G C A F R O V C H T T A V P W E K A C G N N S N F C P K P Q W K N M G---AAAATTCAG---A---TG---ATTT---G---AAG---C---CA---A---A---A---GTAAA---AC---CC.....T---AA---CTGA---G---CA---T

Accession	Strain	Genomic Coordinates
B.FR.HXB2 Env	B.FR.HXB2	172-188
A.UG.U455		195-211
C.ET.ETH2220		218-234
D.CD.84ZR085		241-257
F1.BE.VI850		264-280
G.SE.SE6165		287-303
H.CF.90CF056		310-326
J.SE.SE7887		333-349
K.CM.MP535C		356-372
01.AE.TH.CM240		379-395
02.AG.NG.IBNG		402-418
03.AB.RU.KAL153-2		425-441
04.CX.CY.94CY032-3		448-464
N.CM.MVP5180		471-487
O.BE.ANT70		494-510
CPZ.US.CPZUS		517-533
CPZ.GA.CPZGB		540-556
CPZ.CH.CM3		563-579
CPZ.CH.CM5		586-602
CPZ.CD.CPZANT		609-625
H2A.GW.ALI Env	H2A.GW.ALI	632-648
H2A.SN.HIV25T		655-671
H2A.DE.BEN		678-694
H2B.GH.D205		701-717
H2B.CI.BHO		724-740
H2G.CI.ABT96		747-763
MAC.US.MM251		770-786
SMW.US.SIVSMW9		793-809
STM.US.STM		816-832
TANTALUS.UG.TANI		839-855
GRIVET.ET.GRI-677		862-878
VERVET.KE.VBR-9063		885-901
VERVET.KE.AGMTYO		908-924
VERVET.KE.AGM3		931-947
SABAEUS.SN.SAB1C		954-970
SUN.GA.SIVSUN		977-993
MANDRILL.GA.MNDGB1		1000-1016
LHOEST.KE.hoest		1023-1039
LHOEST.CD.hoest447		1046-1062
LHOEST.CD.hoest485		1069-1085
LHOEST.CD.hoest524		1092-1108
RCM.NG.SIV-NGM		1115-1131
COL.....SIV-CGU1		1138-1154
SYKES.KE.SYK173		1161-1177
SYKES.KE.SYK173 Env		1184-1200
9p41		1207-1223
8232		1230-1246
7662		1253-1269
7608		1276-1292
7753		1299-1315
7509		1322-1338
7650		1345-1361
7551		1368-1384
7539		1391-1407
7392		1414-1430
7785		1437-1453
8228		1460-1476
7602		1483-1499
7783		1506-1522
8306		1529-1545
8279		1552-1568
8240		1575-1591
8272		1598-1614
7648		1621-1637
7954		1644-1660
7691		1667-1683
9p41		1690-1706
8677		1713-1729
8123		1736-1752
8693		1759-1775
8669		1782-1798
8655		1805-1821
8066		1828-1844
8620		1851-1867
8127		1874-1890
8289		1897-1913
8322		1920-1936
8054		1943-1959
8368		1966-1982
7837		1989-2005
8345		2012-2028
7866		2035-2051
8518		2058-2074
8537		2081-2097
7827		2104-2120
9469		2127-2143
7394		2150-2166
7391		2173-2189
7397		2196-2212
8803		2219-2235
7633		2242-2258
8208		2265-2281
9p41		2288-2304

HIV-1/HIV-2/SIV complete genomes

Table with columns for strain names (e.g., B.FR.HXB2 Rev, H2A.GW.ALI Tat), genomic coordinates (exon2, 9p41, 8036), and nucleotide sequences. The sequences are aligned across multiple lines for each strain, showing conserved regions and variations. The table ends with 'end Rev exon 2 <->'.

B.FR.HXB2 Rev	L_G_T	Q_G_V_G	S_P_Q_I_L	exon2
B.FR.HXB2 Env	R_R_G_W	E_A_L_K_Y_W	F_A_L_K_Y_W	gp41
B.FR.HXB2	CGCAGGGGTGG	GAAGCCCTCAAAATTTGG	GAAGCCCTCAAAATTTGG	8612
A.UG.U455	CGCAGCAGCCTCAAGGACTGA-ACT	G-G	G-G	8063
C.ET.ETH2220	CGCAGCAGTCTCAAGGACTGA-AG	A-T	A-T	8009
D.CD.84ZB085		A-T	A-T	8133
F1.BE.VI850	A-G	T	T	7889
G.SE.SE6165	CGCAGCAGCCTCAAGGACTGA-ACT	-GT-T	-GT-T	8051
H.CF.90CF056	A-A	A	A	7931
J.SE.SE7887	A-GG	T	T	7919
K.CM.MP535C	CACAGCAGTCTCAAGGACTGA-AC	G	G	7772
01.AE.TH.CM240	CACAACTGCTCAAGGACTGA-ACT	G	G	8186
02.AG.NG.IBNG	ATA	G	G	8139
03.AB.RU.KAL153-2	CAGAGTCTCAGCAGGACTG-AACTACT- AAT	CT-AGA-C-C-T	CT-AGA-C-C-T	8608
04.CPX.CY.94CY032-3	CTGGACTGTGGATCCTGGGA-AA-A-ACAATT	-TTGT-G-CT-T	-TTGT-G-CT-T	7985
O.CM.MVP5180	CTGGACTGTGGATCCTGGGA-AA-A-ACAATT	A-T-TTGT-G-CT-T	A-T-TTGT-G-CT-T	8184
O.BE.AN70	CAGAACATCAACAGGACTG-CTACT-T	A-T-AA-G-GCAC-C	A-T-AA-G-GCAC-C	8670
CPZ.US.CPZUS	CAGAGCCTCAACAGGACTG-AACA-CT-A	AG-CTG-ACGGG-GA- ACTGTGCTT	AG-CTG-ACGGG-GA- ACTGTGCTT	8641
CPZ.GA.CPZGAB	CAGAGCCTCAACAGGACTG-AACA-CT-A	A-CTT-AG-G-CCT	A-CTT-AG-G-CCT	8669
CPZ.CM.CM3	AGAATTACTTTCAGACTGTGCAA-CATCT-GA	A-CAAT	A-CAAT	8049
CPZ.CM.CAM5		A-CAAT	A-CAAT	8355
CPZ.CD.CPZANT		A-CAAT	A-CAAT	8092
H2A.GW.ALI Rev	P_T_D_L_P_E_S_S_E_S_A_D_N_N_O_G	L_A_E_T_K	L_A_E_T_K	exon2
H2A.GW.ALI Env	I_L_O_P_I_F_R_S_L_O_R_A_L_T_I_R	D_W_L_R_L_K	D_W_L_R_L_K	gp41
H2A.GW.ALI	ATCCTCCAACCGGATCTTCGGAGTCTTCAGAGGCGCTGACA-CAATCA	CTGG-G-G-CT-AAA	CTGG-G-G-CT-AAA	9099
H2A.SN.HIV2ST	ACCCCTCAACTAATCCAGACTCTCCGAGACTTCGACAGCA-TCA-A	CTGG-G-G-G-T-AAC	CTGG-G-G-G-T-AAC	8545
H2A.DE.BEN	ACCCGGCAGTGAATCCAGACTTCAGAGACTTAACAGCAATCA	CTGG-G-G-G-CT-AA	CTGG-G-G-G-CT-AA	9103
H2B.GH.D205	ATCCTCCACAAATCTCA-CAAATCTC	C-C-A	C-C-A	9067
H2B.CI.EHO	ACCCCTC-AACCA-TACTC	C-C-G	C-C-G	9044
H2G.CI.ABT96	ACTCTC-AACC...CT-TTCAAAGAATCTCCAGGACCTCCAAGCAATTAGA	CAT	CGGTAGAA	8488
MAC.US.MM251	ATCCTCAACCAATCTCCAGAGCTCTCCGACCTTA-A	TTC-A	GGAC-GAA	9042
SMM.US.SIVSMMH9	ATCCTCAGCAGTGTCCAGAGCCTTCAAGAGCTG-AG-A-CCC-T	T-A	G-GT-GA	8549
STM.US.STM	ACCCCTCAACCAATCTCCAGAGTCTCCAGACTG-AG-AAITTA-A	T-G	G-GT-GA	8711
TANTALUS.UG.ITANI	AGCCTCAGC-CTCTTCTT	C-CTG-T-G-C-GGA	C-CTG-T-G-C-GGA	8723
GRIVET.ET.GRI-677		A	GGACAGCA	8413
VERVET.KE.VBR-9063		T-G-GCGCT	T-G-GCGCT	8730
VERVET.KE.AGMTYO		T-GGA-AGCT	T-GGA-AGCT	8184
VERVET.KE.AGM155		T-G-AGGCT	T-G-AGGCT	8707
VERVET.DE.AGM3		T-G-AGGCT	T-G-AGGCT	8228
SABAEUS.SN.SABLC		T-G-AGGCT	T-G-AGGCT	8904
SUN.GA.SIVSUN	IGCCCTC-AGGACAUCCAA	C-GAGAAG-G-C-GCTA	C-GAGAAG-G-C-GCTA	8998
MANDRILL.GA.MNDGB1	ATITGAATAT	ACCCT-AGAGGT-GGCT	ACCCT-AGAGGT-GGCT	8195
LHOEST.KE.hoest	ATITGAATAT	ACCAG-AG-G-AGC	ACCAG-AG-G-AGC	8936
LHOEST.CD.hoest447	ATITGAATAT	ACCAG-AG-G-AGC	ACCAG-AG-G-AGC	7861
LHOEST.CD.hoest485	ATITGAATAT	ACCAG-AG-G-AGC	ACCAG-AG-G-AGC	7858
LHOEST.CD.hoest524	ATITGAATAT	ACCAG-AG-G-AGC	ACCAG-AG-G-AGC	7864
RCM.NG.SIV-NGM	ATITGGGTGGATAGACTAAGCTAAGATGGGTGGCAA	ACCCT-AG-G-TCCCT	ACCCT-AG-G-TCCCT	9234
COL.---SIV-CGU1	AGCTTTGTTGGTCAATTCAGCAATGGGAT-CTTACTGTC	CTGC-T-AAC-TCCGCCA	CTGC-T-AAC-TCCGCCA	7968
SYKES.KE.SYK173				8543
SYKES.KE.SYK173 Env				Env

HIV-1/HIV-2/SIV complete genomes

B.FR.HXB2 Env	R_A_I_R_H_I_P_R_R_I_R_O_G_L_E_R_I_L_L_S	gp160 and gp41 Env end <-	-> Nef starts	M_G_K_W_S_K_S_S_V_I_G_W_P_T_V_R_E	Nef
B.FR.HXB2	AGAGCTATTCGCCACATACCTAGAAAGAAATAGACAGGCTTGGAAAGGATTTTGGTATAAA.G	ATCGGTGGCAAGTGGTCAAAAAGTAGTGTGATGGATGGCCCTACTGTAAAGGAAA	8851
A.UG.U455	---T-A---A---C---A---A---GC---G---C	---C---A---A---GC---G---C	8302
C.ET.ETH2220	T-CT-A---A---GCAGC-A	T-CT-A---A---GCAGC-A	8248
D.CD.84ZB085	-A---G-G-TT---C-C---T-A---GC---G	-A---G-G-TT---C-C---T-A---GC---G	8372
F1.BE.VI850	-G-T-A---C-C---GCA---G	-G-T-A---C-C---GCA---G	9128
G.SE.SEG165	-C-T-A---C-C---A---AGC---A	-C-T-A---C-C---A---AGC---A	8290
H.CF.90CF056	-T---TT---A---AGC---A	-T---TT---A---AGC---A	8171
J.SE.SE7887	-T-A-TT---A---AGC---A	-T-A-TT---A---AGC---A	8146
K.CM.MP535C	-C---T---A---C---A---C---A---AGC---A	-C---T---A---C---A---C---A---AGC---A	8011
L.AE.TH.CM240	-G-TA---A---C---GCA---A---GC---A	-G-TA---A---C---GCA---A---GC---A	8425
O2.AG.NG.IBNG	-G-TA---A---C---GCA---A---GC---A	-G-TA---A---C---GCA---A---GC---A	8378
03.AB.RU.KAL153-2	-G-TA---A---C---GCA---A---GC---A	-G-TA---A---C---GCA---A---GC---A	8847
04.CPX.CV.94CY032-3	-G-GA-ATTA---C-A---A---AGCAC---A	-G-GA-ATTA---C-A---A---AGCAC---A	8224
N.CM.YBF30	CA-GAT-C-TT---C-A---A---TGCA---A---C---AG-G	CA-GAT-C-TT---C-A---A---TGCA---A---C---AG-G	8423
O.CM.MVP5180	-C-GA---TA---C-A---G---T---A---A---G---AT-G	-C-GA---TA---C-A---G---T---A---A---G---AT-G	8919
O.BE.AN70	-AT---ATA---A---GC-C---T---G---A---C---A---G---A---GC---AT	-AT---ATA---A---GC-C---T---G---A---C---A---G---A---GC---AT	8946
CPZ.US.CPZUS	-GC---TG---T---A---A---AGC---A	-GC---TG---T---A---A---AGC---A	8978
CPZ.GA.CPZGAB	-GC---TA---T---A---A---AGC---A	-GC---TA---T---A---A---AGC---A	8924
CPZ.CM.CM3	-GC---AA---GG---AG---GC-C---C	-GC---AA---GG---AG---GC-C---C	8288
CPZ.CM.CM5	8594
CPZ.CD.CPZANT	8311
H2A.GW.ALI Env	R_R_I_L_A_V_P_R_R_I_R_O_G_A_E_I_A_L_L_S	G_G_Y_S_Q_S_Q_G_G_S_G_R_G_Q_K_L_P_S_C_E	gp41
H2A.GW.ALI Nef	-GAGG-A-T-GCAG-C-A-G-G-C-G	-GAGG-A-T-GCAG-C-A-G-G-C-G	Nef
H2A.GW.ALI	9282
H2A.SM.HIV2ST	-G-GA-A-TTGCGAG-C-A-G-C-G	-G-GA-A-TTGCGAG-C-A-G-C-G	8728
H2A.DE.BEN	-G-GA-A-T-GCAG-T-A-C-G	-G-GA-A-T-GCAG-T-A-C-G	9286
H2B.GH.B205	-A-CAT-GCG-C-C-G	-A-CAT-GCG-C-C-G	9250
H2B.CI.EHO	-GG-A-T-GCA-G	-GG-A-T-GCA-G	9227
H2G.CI.ABT96	-TGG-C-T-GCA-C	-TGG-C-T-GCA-C	8671
MAC.US.MM251	-G-GA-A-T-GCA-C-MR-C-C-C	-G-GA-A-T-GCA-C-MR-C-C-C	9225
STM.US.SIVSNMH9	-AGG-CG-GGC-C-A-GC-C	-AGG-CG-GGC-C-A-GC-C	8732
STM.US.STM	-G-GA-A-T-G-ACAT-A-GC-C	-G-GA-A-T-G-ACAT-A-GC-C	8894
TANTALUS.UG.TAM1	-G-GA-A-T-G-ACAT-A-GC-C	-G-GA-A-T-G-ACAT-A-GC-C	8906
GRIVET.ET.GRI-677	-GCAC-CAT-AG-GT-A-G-C	-GCAC-CAT-AG-GT-A-G-C	8596
VERVET.KE.VBR-9063	-GAAC-CGT-A-GT-A-G-C	-GAAC-CGT-A-GT-A-G-C	8913
VERVET.KE.AGMTYO	-G-A-CAT-A-TCT-A-G-GC	-G-A-CAT-A-TCT-A-G-GC	8367
VERVET.KE.AGMTYO	-G-A-CAT-A-TCT-A-G-GC	-G-A-CAT-A-TCT-A-G-GC	8890
SABAEUS.SN.SABLC	-G-AG-CAT-GCAAT-A-G-C	-G-AG-CAT-GCAAT-A-G-C	8411
SUN.GA.SIVSUN	-A-TGG-A-A-AACA-G-GGA-T	-A-TGG-A-A-AACA-G-GGA-T	9087
MANDRILL.GA.MNDGB1	-CGTC-AT-G-----TC-----TC	-CGTC-AT-G-----TC-----TC	9181
LHOEST.KE.hoest	-TTC-GATT-GGAGG-AGA-TG-G	-TTC-GATT-GGAGG-AGA-TG-G	8378
LHOEST.CD.hoest447	-TTC-GATT-GGAGG-AGA-TG-G	-TTC-GATT-GGAGG-AGA-TG-G	9110
LHOEST.CD.hoest485	-TTC-GATT-GGAGG-AGA-TG-G	-TTC-GATT-GGAGG-AGA-TG-G	8035
LHOEST.CD.hoest485	-TTC-GATT-GGAGG-AGA-TG-G	-TTC-GATT-GGAGG-AGA-TG-G	8032
RCM.NG.SIV-NGM	-G-TGGCAGG-ACGT-GTGC	-G-TGGCAGG-ACGT-GTGC	8035
COL.____.SIV-CGU1	9390
SYKES.KE.SYK173 Env	GAG-CG-GCAGC-T-G-GC-G	GAG-CG-GCAGC-T-G-GC-G	8723
SYKES.KE.SYK173 Nef	E_A_V_A_A_L_P_R_R_L_R_O_G_L_E_I_Y_L_N_S	E_A_V_A_A_L_P_R_R_L_R_O_G_L_E_I_Y_L_N_S	gp41
SYKES.KE.SYK173 Nef	R_P_L_O_P_C_R_G_G_F_D_K_A_W_R_S_T_L_T_E	R_P_L_O_P_C_R_G_G_F_D_K_A_W_R_S_T_L_T_E	Nef

HIV-1/HIV-2/SIV complete genomes

HIV-1/HIV-2/SIV complete genomes

Accession	Strain	Genome Type	Sequence	Position
B.FR.HXB2	Nef	HIV-1	GAATGACGGA.....GCT.....E P A A D R V G A A S R D L E K H G A I T.....S S	8933
B.FR.HXB2	Nef	HIV-1GCT.....GAGCCAGCAGATAGGGTGGGAGCAGCATCTCGAGACCTGGAAAAACATGGCAATCACA.....AGTAG	8933
A.UG.U455	Nef	HIV-1GAAA.....CCTG.....A-AG-A-A.....T-A-TT-A-T-T.....G.....C-.....C-	8384
C.ET.ETH2220	Nef	HIV-1ACTGATCCAGGGAAGAAGAGC.....G.....T-A-C-T-T.....G-C-T.....C-	8330
D.CD.84ZR085	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8478
F1.BE.VI850	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8210
G.SE.SE6165	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8372
H.CF.90CF056	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8253
J.SE.SE7887	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8231
K.CM.MP535C	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8120
01.AE.TH.CM240	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8507
02.AG.NG.IN8G	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8460
03.AB.RU.KAL153-2	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8932
04.CP.CY.94CY032-3	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8342
N.CM.YBF30	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8520
O.BE.ANT70	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	9043
CPZ.US.CPZUS	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	9019
CPZ.GA.CPZGAB	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8966
CPZ.CM.CAM3	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	9006
CPZ.CM.CAM5	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8370
CPZ.CD.CPZANT	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8676
H2A.GW.ALI	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	8387
H2A.GW.ALI	Nef	HIV-1A-.....CCTA.....AG-A-A.....T-A-C-G-GG.....T.....C-.....C-	9345
H2A.SN.HIV25T	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8791
H2A.DE.BEN	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	9349
H2B.GH.D205	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	9265
H2B.CI.BHO	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	9242
H2G.CI.ABT96	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8734
MAC.US.MM251	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	9288
SWM.US.SIVSMMH9	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8795
STM.US.STM	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8957
TANTALUS.UG.TANI	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8927
GRIVET.ET.GRI-677	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8614
VERVET.KE.VBR-9063	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8940
VERVET.KE.AGMTYO	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8394
VERVET.KE.AGM155	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8917
VERVET.DE.AGM3	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8438
SABAEUS.SN.SAB1C	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	9117
SUN.GA.SIVSUN	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	9181
MANDRILL.GA.MNDGB1	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8378
LHOEST.KE.roest	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	9110
LHOEST.CD.roest447	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8035
LHOEST.CD.roest485	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8032
LHOEST.CD.roest524	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8035
RCM.NG.SIV-NGM	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	9390
COL.____.SIV-CGU1	Nef	HIV-1G-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8151
SYKES.KE.SYK173	Nef	SIVG-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8738
SYKES.KE.SYK173	Nef	SIVG-G-CAGCGGTATCA-CAG-G-AT-T-ATGA-TACCCCTGGAGAACCC-GCA-.....GAA-	8738

B.FR.HXB2 Nef	N T A A T N A A	C A W L E A O E E E E F	V G F	P V T P Q V P L R P	Nef
B.FR.HXB2	CAATACAGCAGCTACCAATGCTGCT	TGTGCCTGGTAGAAGCACAAGAGGAGGAG	GTGGGTTTT	CCAGTCACACCTCAGGTACTCTTTAAGACCA	9030
A.UG.U455	T-T-T-T-T-CAG-	G-G-G-G-GA-C	A-C-C	T-GG-A	8481
D.ET.ETH2220	C-C-C-C-AT-C-	G-G-G-G-GA-C	A-C-C	T-GG-A	8430
F1.BF.VI850	A-T-T-T-C-	G-G-G-G-GA-C	A-C-C	T-GG-A	8578
G.SE.SE6165	TAG-A-T-T-C-AC	G-G-G-G-GA-C	A-C-C	T-GG-A	8307
H.CF.90CF056	C-AT-C-A-	G-G-G-G-GA-C	A-C-C	T-GG-A	8472
J.SE.SE7887	T-T-T-T-CG-A-	GCC-C-G-GHA	A-C	GG-A	8353
K.CM.MP535C	TT-CAC-AT-C-A	G-G-G-G-GA-C	A-C	T-GG-A	8328
O1.AE.TH.CM240	ATAGAT-A	T-GAG-C	A-C	TG-G	8598
O2.AG.NG.IBNG	CAA-T-CC-A	T-T-A-T	A-C	GG-A	8577
O3.AB.RU.KAL153-2	AT-AT-A	G-G-G-G	A-C	GG-G	9029
O4.CPX.CY.94CY032-3	T-C-AC	AAAA-T	A-C	GG-A	8442
N.CM.MVP5180	TAG-AC-AT-AAAG	A-A-A-A-AGAG	A-C	ACCG-A	8620
O.BE.ANT70	TC-TCAA-A	CT-A-TC-CAGC-CA-A	A-C-C	A-G-C	9116
CPZ.US.CPZUS	C-CAG-A-CCAAA-	CT-A-TC-AGT-CC-A	A-C	AG-C	9103
CPZ.GA.CPZGAB	CC-AG-T-CAAA	CTG-T-T-GA-ATG-CA-T	A-A-C	T-TCCG-A	9103
CPZ.CM.CM3	C-C-CAG-AT-CAGA-	CTG-T-T-A-ATG-CA-TC-CA-T	AGAA-A-C	T-G-C	8473
CPZ.CM.CM5	A-A-AT-CAGA-C	CTG-T-T-A-ATG-CA-TC-T	AGAA-A-C	T-G-C	8776
CPZ.CD.CPZANT	-G-AG-AAAA-A	GTAAT-ACCTATTC-A-G-TC-TACA	A-GGGACA-C	T-G-G	8484
H2A.GW.ALI Nef	E K E L Y K O Q	N M D D D D L D S L V G V	S V T P R V Q L R T	A G G A	9448
H2A.GW.ALI	AG-G-A-A-TTGTA--GCAACAG	AA-ATGGATGAT-T-AITTT-T-T-T-TAGCCTTA-A-GG-C	T-T-T--AAGA-AA		
H2A.SN.HIV2ST	G...A-GCT-GTA--GCAACAA	AA-ATGGATGAT-T-AITTC-T-T-T-T-C	CTA-A-GG-C	T-T-T--AAGA-AA	8888
H2B.DE.BEN	AC-G-A-ATTTGAT-GGCAGCA	AA-ATGGATGAT-T-AITTC-T-T-T-T-C	CTA-A-AG	T-T-T--AAGA-AA	9449
H2B.GH.D205	GGGG--A-GGGAGGCAACAA	GATCC-T-A-AGT-T-T-C	TGAA-AGCCATATGTAAG-C-AATA	ACG-G-GA	9362
H2B.CI.BHO	GGGAG--A-GGGAGGAAACAA	GATTC-ATG-G-T-T	CAATGAA-A-GG-C	GT-A-G-CGG-C-AC-G-G	9336
H2G.CI.BHT96	AGCA-AGTT--ATA--GGYAGCAA	AAATAATCAT--T-G-TAGT-T	TA-TGACTTA-A-AG-A	TAC-CA-A-C-CC	9834
MAC.US.MM251	AGCA-A-IT--ATA--GAAACAA	AA-ATGGATGAT-ATG-G-A-T-T	T-C	ITG-A-GG-A	9388
SMW.US.SIVSMW9	AGCA-A-IT--G-TAT-GACACAA	ACATGGATGAT-TG-ATA-T-A-T	T-T-C	ITTA-A--G	8889
STM.US.STM	AGCA-A-IT--G-TAT-GACACAA	ACATGGATGAT-T-ATG-G-A-T-T	TA-T	CTA-A-AG-A	9057
TANTALUS.UG.TANI	AGCA-A-IT--G-TAT-GACACAA	ACATGGATGAT-T-ATG-G-A-T-T	TA-T	CTA-A-AG-A	9024
GRIVET.ET.GRI-677	GGCCGGCTGCAAG-A-GA-AC	ACATGGATGATGAG-C	A-A	A-A	8071
VERVET.KE.VBR-9063	TGGA-G-CGCTGCAAG-A-GA-AC	GAATGGATGAGTGG-ATG-T-T	A-A	A-A	9037
VERVET.KE.AGMTYO	GG-A-GGAGGCT-CAAG-A-GA-AC	ACCTGGGATGAGTGG-ATG-T-T	A-A	A-A	8491
VERVET.KE.AGMT55	AGGA-G-CGCTACAGG-C-GG-AC	ACCTGGGATGAGTGG-ATG-T-T	A-A	A-A	9014
SABAEUS.SN.SAB1C	AGGA-G-AA-TTACAAG-A-GA-A	AAATGGGATGAGTGG-ATG-T-T	A-A	A-A	8535
SUN.GA.SIVSUN	TGTA-A-ATGAAACATC-A-GAAG	CAGC--C-TGGT-T-ATG-G-T-A-A	A-A	A-C	9214
MANDRILL.GA.MNDGB1	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	T-A	T-A	8466
LHOEST.KE.roest	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	T-A	T-A	9110
LHOEST.CD.roest447	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	T-A	T-A	8035
LHOEST.CD.roest485	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	T-A	T-A	8035
LHOEST.CD.roest524	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	T-A	T-A	9502
RCM.NG.SIV-NGM	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	T-A	T-A	8269
COL.....SIV-CGU1	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	T-A	T-A	
SYKES.KE.SYK173	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	T-A	T-A	8862
SYKES.KE.SYK173 Nef	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	AGATGAGAA...ACTTGAC-TAA-ACA-AG-TC-G	T-A	T-A	Nef

HIV-1/HIV-2/SIV complete genomes

Accession	Strain	Genome	Position
B.FR.HXB2	Nef	CAATAAAGGAGAGAACACCAGC...TTGTTTACACCCCTGTGAGCGTGCATGGATGGATGAC...PFE R E V L E W R F D S R L A F H H...V A	9381
B.FR.HXB2	Nef	CAATAAAGGAGAGAACACCAGC...TTGTTTACACCCCTGTGAGCGTGCATGGATGGATGAC...PFE R E V L E W R F D S R L A F H H...V A	9381
A.UG.U455	Nef	T-C-GG...A-T...C-AC...A-AT-AA...AG-A-T...GAA...A...TAC...G...AA-A...AGA-T	8832
C.ET.ETH2220	Nef	G-G...A-T...C-G...A-A...G-T...GAA...A...TAC...G...AA-A...AGA-T	8781
D.CD.842R085	Nef	C-G...A-T...C...A-CT-A...A-A...G...GAA...A...TAC...G...AA-A...AGA-T	8929
F1.BF.VI850	Nef	G...A-T...C-A...CA-CT-A...A-A...G...GAA...A...TAC...G...AA-A...AGA-T	8658
G.SE.SE6165	Nef	G...A-T...C...CA-CT-A...A-A...G...GAA...A...TAC...G...AA-A...AGA-T	8623
H.JF.90CF056	Nef	A-A-T...C...CA-CT-A...A-A...G...GAA...A...TAC...G...AA-A...AGA-T	8704
J.SE.SE7887	Nef	G-A-T...C...CA-CT-A...A-A...G...GAA...A...TAC...G...AA-A...AGA-T	8679
K.CM.MP353C	Nef	C-AE...TH...CM240...G...GAA...A...TAC...G...AA-A...AGA-T	8568
02.AG.NG.IBNG	Nef	G...A-T...C...CA-CT-A...A-A...G...GAA...A...TAC...G...AA-A...AGA-T	8949
03.AB.RU.KAL153-2	Nef	C-A...RU...KAL153-2...G...GAA...A...TAC...G...AA-A...AGA-T	8908
04.CP.XV.94CY032-3	Nef	C-A...XV...94CY032-3...G...GAA...A...TAC...G...AA-A...AGA-T	9364
N.CM.MVP5180	Nef	C-G...A-T...C-G...A-A...G-T...GAA...A...TAC...G...AA-A...AGA-T	8793
O.BE.ANT70	Nef	G...TAC-A-TG-AGATGCTAGTC-TC-T-A-CTT-TAAT-AGCT-G-T...G-AC-CG-G-AC-A-A-CA-T-ATCAT-GC-AACA-T...A-A	8971
P.UM.UM70	Nef	AGGA-TAC-TGTG-G-GGCTAATC-CC-G-T-A-CAT-TGCC-AT-T-A-T...A-AC-T-A-A-A-T-ATCT-GGAACACC-T...T-T-T	9494
CPZ.US.CPZUS	Nef	C-A-A-TA...C-C-G-T...CA-AT-T-A...A-A-T...GAAC-TG-C-C-ATC-CA...CTGAA-G-TGCGAGA...AGA-T	9420
CPZ.GA.CPZGAB	Nef	T-A-T...A-T...C-G-T...CA-AT-T-A...A-A-T...GAAC-TG-C-C-ATC-CA...CTGAA-G-TGCGAGA...AGA-T	9454
CPZ.CM.CM3	Nef	C-C...AT-TA...C-C-T...CA-AT-AA...A-A-T...GAAC-TG-C-C-ATC-CA...CTGAA-G-TGCGAGA...AGA-T	8824
CPZ.CD.CPZANT	Nef	AT...GG-A-TA...C-CC-C...CTT-TACA...AGAT-GA...AC-T-A-GA-C-AGA-GAA...GATGAGAAGA...A-T-T-A	9127
CPZ.CD.CPZANT	Nef	AT...GG-A-TA...C-CC-C...CTT-TACA...AGAT-GA...AC-T-A-GA-C-AGA-GAA...GATGAGAAGA...A-T-T-A	8820
H2A.GW.ALI	Nef	T E T...H C...L L H P V Q T S R H D D...T H G E T L V W R F D P K L A H D Y...K A	9793
H2A.GW.ALI	Nef	T E T...H C...L L H P V Q T S R H D D...T H G E T L V W R F D P K L A H D Y...K A	9793
H2A.SN.HIV2ST	Nef	G-G-GACT...CA-T...AG-G-T-A-CACAAACAAGCA-T-T...C-ITG-G-ACA-TT...CC-ACG-T-T-AGCT...A	9233
H2A.DE.BEN	Nef	G-G-GACT...CA-T...AG-G-T-A-CACAAACAAGCA-T-T...C-ITG-G-ACA-TT...CC-ACG-T-T-AGCT...A	9234
H2B.GH.D205	Nef	GG-GG-A-ACC...CAIT-T...C-AA-G...G-CACAGA-CTCTCATG...T...GA-C-TG-G-GACA-T...CA...TC-ATG-G-C-A-A-CT...A	9710
H2B.CI.BHO	Nef	GG-GG-A-ACC...CAIT-T...C-AA-G...G-CACAGA-CTCTCATG...T...GA-C-TG-G-GACA-T...CA...TC-ATG-G-C-A-A-CT...A	9681
H2G.CI.BBT96	Nef	GG-GG-A-ACC...CAIT-T...C-AA-G...G-CACAGA-CTCTCATG...T...GA-C-TG-G-GACA-T...CA...TC-ATG-G-C-A-A-CT...A	9179
MAC.US.MM251	Nef	GG-GG-A-ACC...CAIT-T...C-AA-G...G-CACAGA-CTCTCATG...T...GA-C-TG-G-GACA-T...CA...TC-ATG-G-C-A-A-CT...A	9733
SMW.US.SIVSMM49	Nef	GG-GG-A-ACC...CAIT-T...C-AA-G...G-CACAGA-CTCTCATG...T...GA-C-TG-G-GACA-T...CA...TC-ATG-G-C-A-A-CT...A	9234
STM.US.STM	Nef	GG-GG-A-ACC...CAIT-T...C-AA-G...G-CACAGA-CTCTCATG...T...GA-C-TG-G-GACA-T...CA...TC-ATG-G-C-A-A-CT...A	9405
TANTALUS.UG.TANI	Nef	GG-GG-A-ACC...CAIT-T...C-AA-G...G-CACAGA-CTCTCATG...T...GA-C-TG-G-GACA-T...CA...TC-ATG-G-C-A-A-CT...A	9366
GRIVET.ET.GRI-677	Nef	TG-G-GA...CA-T...C-A-G-T-A-CACAGA-GGG-AGATCC-TGGATAGATC-TG...C-G-TC-A...CCGAATG-G-GG-GT...AAACC-G	9062
VERVET.KE.VBR-9063	Nef	TG-G-GA...CA-T...C-A-G-T-A-CACAGA-GGG-AGATCC-TGGATAGATC-TG...C-G-TC-A...CCGAATG-G-GG-GT...AAACC-G	9388
VERVET.KE.AGMTYO	Nef	ATG-G-A...CAIT-T...C-AG-C-T-A-C-CAGA-GGA-AGATCCA-TGGATCAGCC-TG...GA-C-G-T-A...TCCTATGT-G-A-GT...ACC-A	8842
VERVET.KE.AGMTYO	Nef	ATG-G-A...CAIT-T...C-AG-C-T-A-C-CAGA-GGA-AGATCCA-TGGATCAGCC-TG...GA-C-G-T-A...TCCTATGT-G-A-GT...ACC-A	9365
VERVET.DE.AGM3	Nef	TTG-G-GA...CAIT-T...C-G-G-T-A-CACRAG-AGG-RAAGACC...TGGAATCAACT-TG...C-G-TC-A...TCC-ATCT-G-C-AT-T...ACC-T	8886
SABAEUS.SN.SABLC	Nef	TTG-G-GA...CAIT-T...C-G-G-T-A-CACRAG-AGG-RAAGACC...TGGAATCAACT-TG...C-G-TC-A...TCC-ATCT-G-C-AT-T...ACC-T	9559
SUN.GA.SIVSUN	Nef	T-G-G-GNAC...CA-T...CCTAAT-T...TCTCAGC-YACT-CATTCAG...CAG-AA-G-GTTA-T-NAGACCCT...G-AGAGACTTAT-TGGA...C-C-ACTC...	9515
MANDRILL.GA.MNGB1	Nef	T-G-G-GNAC...CA-T...CCTAAT-T...TCTCAGC-YACT-CATTCAG...CAG-AA-G-GTTA-T-NAGACCCT...G-AGAGACTTAT-TGGA...C-C-ACTC...	8811
LHOEST.KE.hoest	Nef	TCAG-T-AGACA...CAAT...C-T-G-TT-AGCCAG-GGG-TCC-A...T...CTCCAGC-ATG-CCCA-A-T...A-AG...GAGCC-CAI...TCGGAC-CI-AT-T-GC-TA-GAC-ITG...CA-CAT	9459
LHOEST.CD.hoest447	Nef	TCAG-T-AGACA...CAAT...C-T-G-TT-AGCCAG-GGG-TCC-A...T...CTCCAGC-ATG-CCCA-A-T...A-AG...GAGCC-CAI...TCGGAC-CI-AT-T-GC-TA-GAC-ITG...CA-CAT	8384
LHOEST.CD.hoest485	Nef	TCAG-T-CCAGC...CAAT...C-C-G-T-AGCCAG-GGG-TCCA-G...T...CTCCAGC-ATG-CCCA-A-T...A-AG...GAGCC-CAI...TCGGAC-CI-AT-T-GC-TA-GAC-ITG...CA-CAT	8381
LHOEST.CD.hoest524	Nef	TCAG-T-AGAGA...CAAT-T...C-C-G-T-AGCCAG-GGG-TCCA-A...T...CTCCAGC-ATG-CCCA-A-T...A-AG...GAGCC-CAI...TCGGAC-CI-AT-T-GC-TA-GAC-ITG...CA-CAT	8384
RCM.NG.SIV-NGM	Nef	G-G-GA-G...CAIT-T...C-CC-G-T-A-CAGAACTAG-GGG-TCC-A...A...CTG-G-G-A-ACC-T-T...T...GAG-T...CA-GG	9847
COL...SIV-CGU1	Nef	G-G-GA-G...CAIT-T...C-CC-G-T-A-CAGAACTAG-GGG-TCC-A...A...CTG-G-G-A-ACC-T-T...T...GAG-T...CA-GG	8575
SYKES.KE.SYK173	Nef	TG-GG-G...CAT-CACTCC-CC-C-T-T-A...GCCTC--ACA-GGAGTTCCAGCAT-GAGA-CCTCATG-G-ACTTCAACCCACCCTG--TATA--CCAGG-TGG...-A-ATG	9204
SYKES.KE.SYK173	Nef	TG-GG-G...CAT-CACTCC-CC-C-T-T-A...GCCTC--ACA-GGAGTTCCAGCAT-GAGA-CCTCATG-G-ACTTCAACCCACCCTG--TATA--CCAGG-TGG...-A-ATG	9204

B.FR.HXB2 NefR_E_L_H_P_E_Y_F_K_N_C.....\$	Nef end <-	
B.FR.HXB2CGAGAGCTGCATCCGGAGTACTCAAGAAGCTGC...TGA.....	TCF-1 alpha	
A.UG.U455TAT.....	T-AT-AG-A	
C.ET.ETH2220C-.....	T-A-AG-	
D.CD.84ZR085	AA-AAA-T.....	T-A-C-AG-A	---CATGAGCTC.
F.I.BE.VI850AGA.....	A-AG-	
G.SE.SEG165TAAA.....	A-AG-	
H.CF.90CF056A-A-.....	T-A-AG-	
J.SE.SE7887C-A-.....	A-AG-	
K.CM.MP535CA-A-.....	AT-AG-	
01.AE.TH.CM240A-.....	T-A-AG-	
02.AG.NG.IBNGT-A-AG-	T-A-AG-	
03.AB.RU.KAL153-2TT-A-AG-	TT-A-AG-	
04.CPX.CY.94CY032-3TT-A-AG-	TT-A-AG-	
N.CM.YBR30TGC-AAA-C-A-CT-CCC-G-AA	CG-GGA-A	
O.CM.MVP5180	ATCATAACT-C-A-CT-C-GGA-A	CG-GGA-A	
O.BE.AN70AA-.....	CG-	---ATGCC.
CPZ.US.CPZUSA-AA-.....	A-AG-	---ACTI-AGACTGGCGATCGGCACCAAGACTGC.
CPZ.GA.CPZGABA-AA-.....	T-A-C-	
CPZ.CM.CAM3A-AA-.....	T-ATGGAG-CA-AAACC-CCGCCCTGGCAGTAGAARCTGC.	
CPZ.CM.CAM5A-AA-.....	T-ATGGAG-CA-AAACC-CCGCCCTGGCAGTAGAARCTGC.	
CPZ.CD.CPZANTA-AGA-C-T-A-T-AG-	GAG-AGAC-C.	
H2A.GW.ALI NefF_I_L_H_P_E_E_F_G_Y_K.....S_G_L_P_E_D_E_W_K_A_R_L_K_A_R_G_I_P_F_S_K_N_R_N_S_-\$	Nef end <-	
H2A.GW.ALIITTTAITT--A-C-A--G-A--TGG-T--AAG...-C-GGCCTGCCAGAAGATGAGTGGAAAGCAAGACTGAAAGCAAGAGGGGATACCATTTAGTAAGAAACAGGAACAGCCTGAT		
H2A.SM.HIV2STTTTATT-GAT-C-A--G-G--TGG-T--AAG...-C-GGCCTGCCAGAAGATGAGTGGAAAGCAAGACTGAAAGCAAGAGGGGATACCATTTAGTAAGAAACAGGAACAGCCTGAT		
H2A.DE.BENTTCACT--T-C-A-AG-G--TGG-C--AAG...-C-GGATTTGCAGAGAAAGATGGAAGGCAAGACTGAAAGCAAGAGGGGATACCATTTAGTAAGAAACAGGAACAGCCTGAT		
H2B.GH.D205TTCA-TAG-TT--A-AG-G--GG-T--CAG...-C-GGGTACAGAGGAGGAGTGGAAAGCTAGACTTAAAGCAAGAGGGGATACCATTTAGTAAGAAACAGGAACAGCCTGAT		
H2B.CI.EHOTTCAGCAG-TTC--A--G-G--TGG-T--TCAG...-C-GGAATGCCAGAAAGAGTGGAAAGCTTAACTGAGCAAGAGGGGATACCATTTAGTAAGAAACAGGAACAGCCTGAT		
H2G.CI.ABT96TTTAT--T--T--G-G--G-TGGAAGTCAG-GTACCAAGACAGTATGGAAGAAACAACTAAACAGAGAGGACTGCCTATAGAAATAAGCAAGAAAGG...-C-GGCCTGCCAGAAGATGAGTGGAAAGCAAGACTGAAAGCAAGAGGGGATACCATTTAGTAAGAAACAGGAACAGCCTGAT		
NAC.US.SIVSMH9TAT-CTAGAT-C-A-AG-G--GG-AGCAA-T...CAGCCCTGCAGAGAGAGGTTAGAAAGGCTAA		
STM.US.STMTTT-TTGA--C-A-AG-G--TGG-T--CAG...-C-GGCCTGCCAGAAGATGAGTGGAAAGCAAGACTGAAAGCAAGAGGGGATACCATTTAGTAAGAAACAGGAACAGCCTGAT		
TANTALUS.UG.ITANITTT-TTGA--C-A-AG-G--TGG-T--CAG...-C-GGCCTGCCAGAAGATGAGTGGAAAGCAAGACTGAAAGCAAGAGGGGATACCATTTAGTAAGAAACAGGAACAGCCTGAT		
GRIVET.ET.GRI-677GACAGT-TA-GGACAT-C-TGGTCTTGTAAAG...A-GAAGTAG		
VERVET.KE.VBR-9063GACAT-T-TA-GGACAT-C-GAAC-TGCAAAAG...C-CTAG		
VERVET.KE.AGM150AACAGA-A-T-TTT-AG-A-GC-TGCAATA...GTGAAGAGGAAGTAG		
VERVET.KE.AGM155AATCG-GA-T-CTTTACTG-A-GC-TGGGCTG...GTGAAGAGGAAGTAG		
SABAEUS.SN.SABL1TC-AG-GA-T-CTTTACTG--AT-TTCRACA...GTTGGTACGGAARCTAG		
SUN.GA.SIVSUNAACAGAAAT--TCACTG-A-GC-TG-TCWT...C-ORAGGGAAGTAG		
MANDRILL.GA.MNDGB1I-GAG--C-A--C-GG-GCC-TCAGCA...CA-GGCTAA		
LHOEST.CD.hoest447	TATGATTTAGGCCAT-C-AAA-C-T--AG-G--C-TGTTACTA-TCTGAGTGGAGGTTAGCTTAAACCCATCCCTTCCTTCGCTAGGCAAGC...-C-GGCCTGCCAGAAGATGAGTGGAAAGCAAGACTGAAAGCAAGAGGGGATACCATTTAGTAAGAAACAGGAACAGCCTGAT		
LHOEST.CD.hoest485TA-C-TA-CACC-T-AGTGT-CA-TA-GGTAAGT-TTGAACATCTGCCATTTGAAAGGAAGATGAGTGGAAAGCAAGACTGAAAGCAAGAGGGGATACCATTTAGTAAGAAACAGGAACAGCCTGAT		
LHOEST.CD.hoest524TCR-AAA-G-T-C-A-GA-AC-A-TACTGCT-TCGCTTT...GACTCGAGAGAAAGTACCAAAACCCGCAACCCGAGCCATCCCTTCCTTCGAGCCCG		
RCM.NG.SIV-NGMTCAGAAA-A-TA-AAG-A-AC-AGCA-CAGCTATTGGCTTT...CAGAGAAAGTACCAAAACCCGCAACCCGAGCCATCCCTTCCTTCGAGCCCG		
COL.--SIV-CGU1TCAGAAA-A-TA-AAG-A-AC-AGCA-CATCTCTTAGCTTG...AACTGTAAGAGAAAGTACCAAAACCCGCAACCCGAGCCATCCCTTCCTTCGAGCCCG		
SYKES.KE.SYK173TCAGAAA-A-TA-AAG-T-C-AGCAACTATGATTTAATCTCGAGAGAAAGTACCAAAACCCGCAACCCGAGCCATCCCTTCCTTCGAGCCCG		
SYKES.KE.SYK173 NefTATAGA--T--T--TGGA...GAGAAAAGAAACAAGCAAGCAAGTACCA...CTTCTCTGGAGTTGCTTGGAAAACAGCCGATGC		
-A-C-G-TGCCATA--ATGG-GG-CTGA.....		
GCCCGCT--GTTA--AGACAA-CAGAAACACCA-CAGCTGGAGTACACACTTAGTAAGAAACACTCTCTAGGAG...ACTCCATGGTGTGACAAGCTCGGCCAGGGGA		
A_R_L_Q_L_E_R_Q_T_G_K_P_Q_E_I_Q_S_A_L_S_K_N_I_S_-\$	Nef end <-	

HIV-1/HIV-2/SIV complete genomes

HIV-1/HIV-2/SIV complete genomes

B.FR.HXB2CATCGAGCTTGTACAAA.....	NF-kappa-B-IIGGGACTTTCGG.CTG.....	NF-kappa-B-I <Sp1-III>	Sp1-IIIGGGACTTTCAGGGGCGG.TG.GCCT	9473
A.UG.U455AAG.ACTG.CTGA--CA--AG--GAC.....A--GA--GA--G--T--GT--	8932				
C.ET.ETH2220CACAGA--G--AC--TTCCGGCC.....A--G--AG--AG--G--T--	8874				
D.ED.84ZR085GACTGTGCACACAGAGATT.GCTG.ACACAGAA--ATC--A--A--A--A--G--CG--TCC..AGAG	8975				
F1.BF.V1850CACAGAA--TTGC--AC--AG.....AC--G--G--A--T--	8972				
G.SE.SE6165CACAGAA--T.C--G--A--ACAGAA--T.TC--G--A--G--G--A--T--	8917				
H.CF.90CF056CAAGAA--TTTC--AG--G.....CAAGAA--TTTC--AG--G.....GA--GAG--	8791				
J.SE.SE7887CAAGAA--TTTC--AG--G.....CAAGAA--TTTC--AG--G.....GA--GAG--	8772				
K.CM.MP535CCAAGAA--TTTC--ACT.....A--A--GA--GTG--	8598				
01.AE.TH.CM240CACAGAA--TTGC--AC.....CACAGAA--TTGC--AC.....GA--GTG--	9040				
02.AG.NG.IBNGCACAGAC.TGCT.GACATGAA--TTGC--AC.....A--GA--GTG--	8999				
03.AB.RU.KAL153-2CACAGAA--TTGC--AC--A--C--C--GA--GTG--	9467				
04.CP.CY.94CY032-3CAAGGACTTACTGCTG--CAA.....ATACT.....GA--GTG--	8888				
N.CM.MVP5180CTGACACTGC.....CTGACACTGC.....A--GTG--	9082				
O.BE.ANT70TGTAACGGC.GCAGGGCAATA.A--GCTG--TG.....A--C--T--GAGG--ATAAGG	9595				
CPZ.US.CPZUSTGACTCTG.....A--T--GAGG--ACAGGG	9538				
CPZ.GA.CPZGABTGCCAGCG.CCTG.CCAGTAGA.A--GCTG--AA.....G--A--G--G--GTC	9563				
CPZ.CM.CM5TGACAA.....TAA.....G--A--G--G--TTC	8927				
CPZ.CD.CPZANTTAACAGCG.GATATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....TAA.....A--G--A--G--TC	9228				
H2A.GW.ALIAACAGC--GAG--AC--A--AAC--GCTGA..ATAGCA.....G--GA--T--	9929				
H2A.SN.HIV2STACAGC--AAC--GC--AG--TGCA.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9990				
H2A.DE.BENCTAAG--AAC--GC--AGG--TGCA.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9430				
H2B.GH.D205GAAAG--AACTAGCA--AC--TGCA.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9991				
H2B.CI.EHOGAAAG--AACTAGC--A--TGCA.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9910				
H2G.CI.ABT96AC--AGC--AC--GCA.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9881				
MAC.US.MM251ACA--A.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9337				
SMM.US.SIVSMMH9ACA--A.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9884				
TANTALUS.UG.TANIGCCTTAGGA--TGACATA--AG--AGT.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9399				
GRIVET.ET.GRI-677TGTCAGC.ACTT.TGCCGT.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9299				
VERVET.KE.VER-9063CTAACCCAGCTTTGGTTAAGCGTTGCCAG.GAGATGAC--TT--GAAACTGTCACAG.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9506				
VERVET.KE.AGMTYOCCAGCCAGCCTCGGTTTACACAT.CACC.ATGGAGAT--CA--AAA--ACTGTTGAC.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9160				
VERVET.KE.AGM155CCGACACAGCTTCCGTTTCTGTTGCC.TAGAGAT--CA--AAG--ACTGTTGAC.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9523				
VERVET.DE.AGM3CTAACCCAGCCTTTGGTTAAGCACATCACCATTGTTGAT--CA--AAG--ACTGTTGAC.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9506				
SABAEUS.SN.SABICCCACAAAACACA.TCCT.ATGGAG--TCA--G--TG--TG--CATTA.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9035				
SUN.GA.SIVSUNGGCT--G--CATGC--GCT--GAGTA.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9678				
MANDRILL.GA.MNDGB1GCTAGC--T--CTGA.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9684				
LHOEST.KE.loeestAGGA--T--CTGA.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	8981				
LHOEST.CD.loeest447ACAAGC--T--CTGA.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9640				
LHOEST.CD.loeest485GAGAAGAT--CTGA..CAC--GAA.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	8559				
LHOEST.CD.loeest524G--T--CTGA.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	8562				
RWM.NG.SIV-NGMACTG--CGCTGCGCTGGTGTAA.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	8565				
COL.____SIV-CGU1TTC.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	10017				
SYKES.KE.SYK173AG--C--GGCGGGG--G.....GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	8655				
G--G--A--T--GATAAAA--CG--TGA--AGAGC--G--CTTGTGTCT.....G--G--A--T--	9328				

B.FR.HXB2	TAA.CTAGAGATCC.....CTCAGAC..CCTTTTAGTCAGTGTGAAAAATCTCTAGCA.	3' LTR end <-	9719
A.UG.U455-AC-A--A-T.GTGTA-----		9178
C.ET.ETH2220-AC-A--A-T.GTGTA-----		9031
D.CD.84ZR085-AC-A--A-T.GTGTA-----		8975
F1.BF.VI850-AC-A--A-T.GTGTA-----		8903
G.SE.SE6165-AC-A--A-T.GTGTA-----		9074
H.CF.90CF056-AC-A--A-T.GTGTA-----		8953
J.SE.SE7887-AC-A--A-T.GTGTA-----		8943
K.CM.WP535C-AC-A--A-T.GTGTA-----		8598
O1.AE.TH.CM240-AC-A--A-T.GTGTA-----		9203
02.AG.NG.IBNG-AC-A--A-T.GTGTA-----		9201
03.AB.RU.KAL153-2-AC-A--A-T.GTGTA-----		9715
04.CPX.CY.94CY032-3-AC-A--A-T.GTGTA-----		9050
N.CM.YBF30-AC-A--A-T.GTGTA-----		9182
O.CM.WVP5180-AC-A--A-T.GTGTA-----		9793
O.BE.ANT70-AC-A--A-T.GTGTA-----		9802
CPZ.US.CPZUS-AC-A--A-T.GTGTA-----		9781
CPZ.GA.CPZGAB-AC-A--A-T.GTGTA-----		9811
CPZ.CH.CAM3-AC-A--A-T.GTGTA-----		9170
CPZ.CH.CAM5-AC-A--A-T.GTGTA-----		9261
CPZ.CD.CPZANT-AC-A--A-T.GTGTA-----		9068
H2A.GW.ALI-AC-A--A-T.GTGTA-----		10353
H2A.SN.HIV2ST-AC-A--A-T.GTGTA-----		9672
H2A.DE.BEN-AC-A--A-T.GTGTA-----		10359
H2B.GH.D205-AC-A--A-T.GTGTA-----		9978
H2B.CI.EHO-AC-A--A-T.GTGTA-----		10242
H2G.CI.ABT96-AC-A--A-T.GTGTA-----		9599
MAC.US.MM251-AC-A--A-T.GTGTA-----		10249
SMM.US.SIVSMMH9-AC-A--A-T.GTGTA-----		9638
STM.US.STM-AC-A--A-T.GTGTA-----		8892
TANTALUS.UG.TANI-AC-A--A-T.GTGTA-----		9752
GRIVET.ET.GRI-677-AC-A--A-T.GTGTA-----		9418
VERVET.KE.VER-9063-AC-A--A-T.GTGTA-----		9779
VERVET.KE.AGMTYO-AC-A--A-T.GTGTA-----		9170
VERVET.KE.AGML55-AC-A--A-T.GTGTA-----		9758
VERVET.DE.AGM3-AC-A--A-T.GTGTA-----		9294
SABAEUS.SN.SAB1C-AC-A--A-T.GTGTA-----		10009
SUN.GA.SIVSUN-AC-A--A-T.GTGTA-----		10006
MANDRIL.GA.MNDGB1-AC-A--A-T.GTGTA-----		9215
LHOEST.KE.loeest-AC-A--A-T.GTGTA-----		9957
LHOEST.CD.loeest447-AC-A--A-T.GTGTA-----		8588
LHOEST.CD.loeest485-AC-A--A-T.GTGTA-----		8589
LHOEST.CD.loeest524-AC-A--A-T.GTGTA-----		8594
RCM.NG.SIV-NGM-AC-A--A-T.GTGTA-----		10412
COL.____.SIV-CGU1-AC-A--A-T.GTGTA-----		8728
SYKES.KE.SYK173-AC-A--A-T.GTGTA-----		9597
H2A.GW.ALI-AC-A--A-T.GTGTA-----		10353
H2A.SN.HIV2ST-AC-A--A-T.GTGTA-----		9672
H2A.DE.BEN-AC-A--A-T.GTGTA-----		10359
H2B.GH.D205-AC-A--A-T.GTGTA-----		9978
H2B.CI.EHO-AC-A--A-T.GTGTA-----		10242
H2G.CI.ABT96-AC-A--A-T.GTGTA-----		9599
MAC.US.MM251-AC-A--A-T.GTGTA-----		10249
SMM.US.SIVSMMH9-AC-A--A-T.GTGTA-----		9638
STM.US.STM-AC-A--A-T.GTGTA-----		8892
TANTALUS.UG.TANI-AC-A--A-T.GTGTA-----		9752
GRIVET.ET.GRI-677-AC-A--A-T.GTGTA-----		9418
VERVET.KE.VER-9063-AC-A--A-T.GTGTA-----		9779
VERVET.KE.AGMTYO-AC-A--A-T.GTGTA-----		9170
VERVET.KE.AGML55-AC-A--A-T.GTGTA-----		9758
VERVET.DE.AGM3-AC-A--A-T.GTGTA-----		9294
SABAEUS.SN.SAB1C-AC-A--A-T.GTGTA-----		10009
SUN.GA.SIVSUN-AC-A--A-T.GTGTA-----		10006
MANDRIL.GA.MNDGB1-AC-A--A-T.GTGTA-----		9215
LHOEST.KE.loeest-AC-A--A-T.GTGTA-----		9957
LHOEST.CD.loeest447-AC-A--A-T.GTGTA-----		8588
LHOEST.CD.loeest485-AC-A--A-T.GTGTA-----		8589
LHOEST.CD.loeest524-AC-A--A-T.GTGTA-----		8594
RCM.NG.SIV-NGM-AC-A--A-T.GTGTA-----		10412
COL.____.SIV-CGU1-AC-A--A-T.GTGTA-----		8728
SYKES.KE.SYK173-AC-A--A-T.GTGTA-----		9597