

II

HIV-1/SIVcpz Complete Genomes

Introduction	164
Table of features annotated in the HIV-1/SIVcpz Complete Genome Alignments	165
Table of HIV-1/SIVcpz Complete Genome Alignments	167
Notes on full-length HIV-1/SIVcpz Complete Genome Alignments	170
Nucleotide Alignments of HIV-1/SIVcpz Complete Genomes	188

Alignment of HIV-1/SIVcpz Genomes

This year again many new full-length viral sequences have become available, originating from diverse geographic origins and representing the spectrum of known HIV variation. We have decided to publish only full length HIV-1/CPZ sequences in our printed nucleotide alignment section, as this set is now becoming an adequate representation of the overall diversity of the virus.

As of December 2000 there were 202 complete or nearly complete (defined as greater than 8,000 consecutive basepairs of sequence) HIV-1 genomes in the database. Of these, some were not included in the printed alignment, as they are very closely related to a sequence already included in the alignment, and our intent is to print a hardcopy alignment representative of global diversity. The complete alignment including all sequences is available at our web site.

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

One hundred and twenty two HIV-1 sequences plus viral strains isolated from chimpanzees comprise the printed alignment. In phylogenetic analyses, the CPZ sequences are the simian-derived viruses most similar to HIV-1; in fact HIV-1 M, N and O group sequences are roughly as distant from one another as they are from the CPZ sequences (see Figure 1 in Section III page 356 of this compendium).

The sequences in this section are identified by their common name preceded by the HIV subtype designations and country of origin. The primary sequence reference, country of origin, database accession number, and brief notes describing the isolate and sequence, with some additional relevant references, can be found in Table 2, page 170. The sequences that have been found to be recombinants with portions of the genetic sequences associated with different subtypes are indicated by listing all of the subtypes in the prefix to the name. For example, the prefix AG simply indicates that some regions of the sequence are subtype A-like, others G-like. The subtypes are organized alphabetically and not meant to reflect the proportion of either subtype in the mosaic genome.

This alignment was generated by using the HMMER Hidden Markov Model sequence alignment software developed by Sean Eddy.

<http://genome.wustl.edu/eddy/hmmer.html>

An iterative process was used involving alignment of the genomes using HMMER, followed by hand-editing (using an in-house revised version of the MASE alignment editing program (Faulkner, D., and Jurka, J., *Trends in Biochem. Science* **13**:321–322 (1988)), and BioEdit.

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

The resulting final alignment is not suggested to be an “optimal alignment” with the absolute minimum number of gaps and mismatches. It is a compromise between optimal alignment, readability, and an attempt to keep insertions and deletions from altering the protein reading frame presentation. Most gaps have been introduced in multiples of 3 bases to maintain open reading frames when translated directly from the alignment. Frameshifting gaps were added at the gag-pol slip site, at the end of pol, and at the end of vif.

After the final alignment was generated, a HMMER model was built with the hmmb program, using this alignment as the input or training set. The final HMMER model based on the full length genomes has been tested here with partial genomes as well. Using the HMMER -R option for ragged ends (gaps inserted at the ends of sequences are given very low weight) the HMMER program did a reasonable job of aligning the complete and partial env genes to each other. The model was used again to align the complete genomes plus the env gene sequences, and in this case all sequences were reasonably aligned to each other.

The annotation

The annotation for the precursor peptide cleavage sites in Gag and Gag-Pol is based on the information published in Tozser91, LeGrice89. The annotation of the Gag-Pol ribosomal slip site is based on information published in Reil93, Kollmus94, Le89. The annotation for the cis-acting transcriptional activation domains in the LTR section is based on information published in Zhang97, Estable96, Montano97, Gao96. There are a varying number of NF- κ B binding sites in C subtype sequences, with

some sequences carrying an additional site (Gao96, Carr96, Montano97). The annotation for the Rev responsive element (the RRE) is based on Charpentier97.

The HXB2 nucleotide reference sequence is translated into all three reading frames at the top of the alignment using the single character amino acid designation. At the bottom of the alignment, protein sequences, based on the HXB2 sequence are indicated; the HIV genome has many overlapping coding regions, and all are shown. For more complete annotation of functional domains see the protein sequence alignments in Part IV.

Features of HIV-1 annotated in the alignment that follows				
DNA	Page	Protein		
		Frame 1	Frame 2	Frame 3
5'LTR U3 start	188			
TCF-1 α	192			
NF- κ B-I, II	194			
SpI, II, III	194			
TATA Box	196			
5' LTR U3 end \ / R repeat begins	196			
mRNA start site	196			
Poly-A signal	196			
5' LTR repeat end \ / U5 start	198			
5' LTR U5 end	198			
Lys tRNA primer binding site	198			
Gag binding loops	200			
	202	Gag-Pol start		
	210	p17 \ / p24		
	220	p24 \ / p2		
	220	p2 \ / p7		
	224	p7 \ / p1 \ / p6		
	228	Gag end		Pol protease start
	232			protease \ / p51 RT
	254			p51 RT \ / p15 RNase H
	258			p15 RNaseH \ / p31 integrase
	272	Vif start		Pol end
	280	Vif end		Vpr start
	284		Tat ex1 start	Vpr end
	286			Rev ex1 start
	288		Tat ex1 end	Rev ex1 end
	288		Vpu start	
	292			
	294		Vpu end	gp120 Env start
V3 loop start	308			
Rev res. element start	318			
Stem I, IIA, IIIB	320			gp120 \ / gp41
Stem IIB', IIC, IIC', IIA'	322			
Stem III-IV, III-IV', V, V', I'	322			
Stem I', Rev res. element end	324			
	330	Tat ex2 start	Rev ex2 start	
	332	Tat ex2 end		
	334		Rev ex2 end	
	336	Nef start		Env end
TCF-1 α binding	348	Nef end		
NF- κ -BII	348			
NF- κ -BI, SPI-III, II, I	350			
TATA box	350			
3' LTR U3 end \ / repeat start	352			
3' LTR repeat end ' LTR U5 start	352			
3' LTR U5 end	354			

- [Carr96] J.K. Carr, M.O. Salminen, C.Koch, D.Gotte, A.W. Artenstein, P.A. Hegerich, L.S. D., D.S. Burke, F.E. McCutchan. Full-length sequence and mosaic structure of a human immunodeficiency virus type 1 isolate from Thailand. *J Virol* 70:593543, 1996.
- [Charpentier97] B.Charpentier, F.Schultz, M.Rosbash. A dynamic in vivo view of the HIV-1 Rev-RRE interaction. *J Mol Biol* 266:950962, 1997.
- [Estable96] M.C. Estable, B.Bell, A.Merzouki, J.S. Montaner, M.V. O'Shaughnessy, I.J. Sadowski. Human immunodeficiency virus type 1 long terminal repeat variants from 42 patients representing all stages of infection display a wide range of sequence polymorphism and transcription activity. *J Virol* 70:405362, 1996.
- [Gao96] F.Gao, D.L. Robertson, S.G. Morrison, H.Hui, S.Craig, J.Decker, P.N. Fultz, M.Gerard, G.M. Shaw, B.H. Hahn, P.M. Sharp. The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype {(A/E)} recombinant of African origin. *J Virol* 70:701329, 1996.
- [Kollmus94] H.Kollmus, A.Honigman, A.Panet, H.Hauser. The sequences of and distance between two cis-acting signals determine the efficiency of ribosomal frameshifting in human immunodeficiency virus type 1 and human t-cell leukemia virus type ii in vivo *J Virol* 68:608791, 1994.
- [Le89] S.Y. Le, J.H. Chen, J.V. Maizel. Thermodynamic stability and statistical significance of potential stem-loop structures situated at the frameshift sites of retroviruses. *Nucleic Acids Res* 17:614352, 1989.
- [LeGrice89] S.F. Le Grice, R.Ette, J.Mills, J.Mous. Comparison of the human immuno-deficiency virus type 1 and 2 proteases by hybrid gene construction and trans-complementation. *J Biol Chem* 264}:149028, 1989.
- [Montano97] M.A. Montano, V.A. Novitsky, J.T. Blackard, N.L. Cho, D.A. Katzenstein, M.Essex. Divergent transcriptional regulation among expanding human immunodeficiency virus type 1 subtypes. *J Virol* 71:865765, 1997.
- [Reil93] H.Reil, H.Kollmus, U.H. Weidle, H.Hauser. A heptanucleotide sequence mediates ribosomal frameshifting in mammalian cells. *J Virol* 67:557984, 1993.
- [Tozser91] J.Tozser, I.Blaha, T.D. Copeland, E.M. Wondrak, S.Oroszlan Comparison of the HIV-1 and HIV-2 proteinases using oligopeptide substrates representing cleavage sites in gag and gag-pol polyproteins. *FEBS Lett* 281:7780, 1991.
- [Zhang97] L.Zhang, Y.Huang, H.Yuan, B.K. Chen, J.Ip, D.D. Ho. Genotypic and phenotypic characterization of long terminal repeat sequences from long-term survivors of human immunodeficiency virus type 1 infection. *J Virol* 71:560813, 1997.

Table 1. Table of HIV-1/SIVcpz complete genome alignments. Only one of each related set of sequences is included in the alignment displayed here, but all are available at the HIV database web site:<http://hiv-web.lanl.gov>

Name	Accession	Country	Author	Reference
A.BY.97BL006	AF193275	BYELORUSSIA	Liitsola, K	<i>ARHR</i> 16 (11):1047–1053 (2000)
A.KE.Q23	AF004885	KENYA	Poss, M	<i>J Virol</i> 73 (7):5255–5264 (1999)
A.SE.SE6594	AF069672	SWEDEN	Carr, JK	<i>AIDS</i> 13 (14):1819–1826 (1999)
A.SE.SE7253	AF069670	SWEDEN	Carr, JK	<i>AIDS</i> 13 (14):1819–1826 (1999)
A.SE.SE7535	AF069671	SWEDEN	Carr, JK	<i>AIDS</i> 13 (14):1819–1826 (1999)
A.SE.SE8538	AF069669	SWEDEN	Carr, JK	<i>AIDS</i> 13 (14):1819–1826 (1999)
A.SE.SE8891	AF069673	SWEDEN	Carr, JK	<i>AIDS</i> 13 (14):1819–1826 (1999)
A.SE.UGSE8131	AF107771	SWEDEN	Laukkanen, T	-
A.UG.92UG037	U51190	UGANDA	Gao, F	<i>J Virol</i> 70 (3):1651–1657 (1996)
A.UG.U455	M62320	UGANDA	Oram, JD	<i>ARHR</i> 6 (9):1073–1078 (1990)
A2.CD.97CDKTB48	AF286238	CONGO	Feng, G	Unpublished (2000)
A2.CY.94CY017-41	AF286237	CYPRUS	Gao, F	Unpublished (2000)
A2C.ZM.ZAM184	U86780	ZAMBIA	Salminen, MO	<i>Virology</i> 213 (1):80–86 (1995)
A2D.-.97KR004	AF286239	-	Gao, F	Unpublished (2000)
A2G.CD.97CDKP58	AF316544	CONGO	Gao, F	Unpublished (2000)
AC.IN.21301	AF067156	INDIA	Lole, KS	<i>J Virol</i> 73 (1):152–160 (1999)
AC.RW.92RW009	U88823	RWANDA	Gao, F	<i>J Virol</i> 72 (7):5680–5698 (1998)
AC.SE.SE9488	AF071474	SWEDEN	Carr, JK	Unpublished
ACD.SE.SE8603	AF075702	SWEDEN	Carr, JK	Unpublished
ACG.BE.VII035	AJ276595	BELGIUM	Janssens, W	<i>ARHR</i> 16 (12):1183–1189 (2000)
AD.SE.SE6954	AF075701	SWEDEN	Carr, JK	Unpublished
AD.SE.SE7108	AF071473	SWEDEN	Carr, JK	Unpublished
ADHK.NO.97NOGIL3	AJ237565	NORWAY	Jonassen, TO	<i>ARHR</i> 16 (1):49–58 (2000)
ADK.CD.MAL	X04415	CONGO	Alizon, M	<i>Cell</i> 46 (1):63–74 (1986)
AG.BE.VII197	AJ276596	BELGIUM	Janssens, W	<i>ARHR</i> 16 (12):1183–1189 (2000)
AG.NG.92NG003	U88825	NIGERIA	Gao, F	<i>J Virol</i> 72 (7):5680–98 (1998)
AGHU.GA.VI354	AF076474	GABON	Carr, JK	Unpublished
AGU.CD.Z321	U76035	CONGO	Choi, DJ	<i>ARHR</i> 13 (4):357–361 (1997)
AJ.BW.BW2117	AF192135	BOTSWANA	Novitsky, VA	<i>ARHR</i> 16 (10):1015–1020 (2000)
B.AU.MBC200	AF042100	AUSTRALIA	Oelrichs, RB	<i>ARHR</i> 14 (9):811–814 (1998)
B.AU.MBC925	AF042101	AUSTRALIA	Oelrichs, RB	<i>ARHR</i> 14 (9):811–814 (1998)
B.CN.RL42	U71182	CHINA	Graf, M	<i>ARHR</i> 14 (3):285–288 (1998)
B.DE.D31	U43096	GERMANY	Kreutz, R	<i>ARHR</i> 8 (9):1619–1629 (1992)
B.DE.HAN	U43141	GERMANY	Sauermann, U	<i>ARHR</i> 6 (6):813–823 (1990)
B.ES.89SP061	AJ006287	SPAIN	Olivares, I	<i>ARHR</i> 14 (18):1649–1651 (1998)
B.FR.HXB2	K03455	FRANCE	Wong-Staal, F	<i>Nature</i> 313 (6000):277–284 (1985)
B.GA.OYI	M26727	GABON	Huet, T	<i>AIDS</i> 3 (11):707–715 (1989)
B.GB.CAM1	D10112	U.K.	McIntosh, AA	Unpublished (1991)
B.GB.MANC	U23487	U.K.	Zhu, T	<i>Nature</i> 374 (6522):503–504 (1995)
B.NL.3202A21	U34604	NETHERLANDS	Guillon, C	<i>ARHR</i> 11 (12):1537–1541 (1995)
B.TW.TWCYS	AF086817	TAIWAN	Huang, LM	Unpublished
B.US.AD8	AF004394	U.S.	Theodore, TS	<i>ARHR</i> 12 (3):191–194 (1996)
B.US.BC	L02317	U.S.	Ghosh, SK	<i>Virology</i> 194 :858–864 (1993)
B.US.DH123	AF069140	U.S.	Shibata, R	<i>J Virol</i> 69 (7):4453–4462 (1995)
B.US.JRCSF	M38429	U.S.	O'Brien, WA	<i>Nature</i> 348 :69–73 (1990)
B.US.JRFL	U63632	U.S.	O'Brien, WA	<i>Nature</i> 348 :69–73 (1990)
B.US.MNCG	M17449	U.S.	Gurgo, C	<i>Virology</i> 164 (2):531–536 (1988)
B.US.NY5CG	M38431	U.S.	Willey, RL	<i>PNAS USA</i> 83 (14):5038–5042 (1986)

B.US.P896	U39362	U.S.	Kim, FM	<i>J Virol</i> 69 (3):1755–1761 (1995)
B.US.RF	M17451	U.S.	Starcich, BR	<i>Cell</i> 45 (5):637–648 (1986)
B.US.SF2	K02007	U.S.	van Beveren, C	<i>RNA Tumor Viruses</i> , 2nd Edition 2:1124–1141; Cold Spring Harbor Laboratory, (1985)
B.US.WCIPR9018	U69591	U.S.	Fang, G	<i>J AIDS</i> 12 (4):352–357 (1996)
B.US.WEAU160	U21135	U.S.	Tozser, J	<i>FEBS Letters</i> 281 :77–80 (1991)
B.US.WR27	U26546	U.S.	Salminen, MO	<i>Virology</i> 213 (1):80–86 (1995)
B.US.YU2	M93258	U.S.	Li, Y	<i>J Virol</i> 65 :3973–3985 (1991)
BF1.BR.93BR029-4	AF005495	BRAZIL	Gao, F	<i>J Virol</i> 72 (7):5680–5698 (1998)
C.BR.92BR025	U52953	BRAZIL	Gao, F	<i>J Virol</i> 70 (3):1651–1667 (1996)
C.BW.96BW01B03	AF110959	BOTSWANA	Novitsky, VA	<i>J Virol</i> 73 (5):4427–4432 (1999)
C.BW.96BW0402	AF110962	BOTSWANA	Novitsky, VA	<i>J Virol</i> 73 (5):4427–4432 (1999)
C.BW.96BW0502	AF110967	BOTSWANA	Novitsky, VA	<i>J Virol</i> 73 (5):4427–4432 (1999)
C.BW.96BW1104	AF110969	BOTSWANA	Novitsky, VA	<i>J Virol</i> 73 (5):4427–4432 (1999)
C.BW.96BW1210	AF110972	BOTSWANA	Novitsky, VA	<i>J Virol</i> 73 (5):4427–4432 (1999)
C.BW.96BW15C02	AF110974	BOTSWANA	Novitsky, VA	<i>J Virol</i> 73 (5):4427–4432 (1999)
C.BW.96BW16B01	AF110976	BOTSWANA	Novitsky, VA	<i>J Virol</i> 73 (5):4427–4432 (1999)
C.BW.96BW17B03	AF110980	BOTSWANA	Novitsky, VA	<i>J Virol</i> 73 (5):4427–4432 (1999)
C.ET.ETH2220	U46016	ETHIOPIA	Salminen, MO	<i>ARHR</i> 12 (14):1329–1339 (1996)
C.IN.93IN101	AB023804	INDIA	Mochizuki, N	<i>ARHR</i> 15 (14):1321–1324 (1999)
C.IN.93IN904	AF067157	INDIA	Lole, KS	<i>J Virol</i> 73 (1):152–160 (1999)
C.IN.93IN905	AF067158	INDIA	Lole, KS	<i>J Virol</i> 73 (1):152–160 (1999)
C.IN.93IN999	AF067154	INDIA	Lole, KS	<i>J Virol</i> 73 (1):152–160 (1999)
C.IN.94IN11246	AF067159	INDIA	Lole, KS	<i>J Virol</i> 73 (1):152–160 (1999)
C.IN.95IN21068	AF067155	INDIA	Lole, KS	<i>J Virol</i> 73 (1):152–160 (1999)
D.CD.84ZR085	U88822	CONGO	Gao, F	<i>J Virol</i> 72 (7):5680–98 (1998)
D.CD.ELI	K03454	CONGO	Alizon, M	<i>Cell</i> 46 (1):63–74 (1986)
D.CD.NDK	M27323	CONGO	Spire, B	<i>Gene</i> 81 :275–284 (1989)
D.CD.Z2Z6	M22639	CONGO	Srinivasan, A	<i>Gene</i> 52 :71–82 (1987)
D.KE.MB2059	AF133821	KENYA	Neilson, JR	<i>J Virol</i> 73 (5):4393–4403 (1999)
D.UG.94UG1141	U88824	UGANDA	Gao, F	<i>J Virol</i> 72 (7):5680–5698 (1998)
F1.BE.VI850	AF077336	BELGIUM	Carr, JK	<i>Virology</i> 269 (1):95–104 (2000)
F1.BR.93BR020-1	AF005494	BRAZIL	Gao, F	<i>J Virol</i> 72 (7):5680–5698 (1998)
F1.FI.FIN9363	AF075703	FINLAND	Laukkanen, T	Unpublished
F1.FR.MP411	AJ249238	FRANCE	Triques, K	<i>ARHR</i> 16 (2):139–151 (2000)
F2.CM.MP255	AJ249236	CAMEROON	Triques, K	<i>ARHR</i> 16 (2):139–151 (2000)
F2.CM.MP257	AJ249237	CAMEROON	Triques, K	<i>ARHR</i> 16 (2):139–151 (2000)
F2KU.BE.VII1126	AF076475	BELGIUM	Carr, JK	Unpublished
G.BE.DRCBL	AF084936	BELGIUM	Debyser, Z	<i>ARHR</i> 14 (5):453–459 (1998)
G.FI.HH8793-12-1	AF061641	FINLAND	Salminen, MO	<i>ARHR</i> 8 (9):1733–1742 (1992)
G.NG.92NG083	U88826	NIGERIA	Gao, F	<i>J Virol</i> 72 (7):5680–5698 (1998)
G.SE.SE6165	AF061642	SWEDEN	Carr, JK	<i>Virology</i> 247 (1):22–31 (1998)
H.BE.VI991	AF190127	BELGIUM	Janssens, W	<i>AIDS</i> 14 (11):1533–1543 (2000)
H.BE.VI997	AF190128	BELGIUM	Janssens, W	<i>AIDS</i> 14 (11):1533–1543 (2000)
H.CF.90CF056	AF005496	C.A.R.	Gao, F	<i>J Virol</i> 72 (7):5680–5698 (1998)
J.SE.SE7022	AF082395	SWEDEN	Laukkanen, T	<i>ARHR</i> 15 (3):293–297 (1999)
J.SE.SE7887	AF082394	SWEDEN	Laukkanen, T	<i>ARHR</i> 15 (3):293–297 (1999)
K.CD.EQTBI1C	AJ249235	CONGO	Triques, K	<i>ARHR</i> 16 (2):139–151 (2000)
K.CM.MP535	AJ249239	CAMEROON	Triques, K	<i>ARHR</i> 16 (2):139–51 (2000)
MO.CM.97- CAMP645MO	AJ239083	CAMEROON	Peeters, M	<i>J Virol</i> 73 (9):7368–7375 (1999)
N.CM.YBF106	AJ271370	CAMEROON	Souquiere, S	Unpublished
N.CM.YBF30	AJ006022	CAMEROON	Simon, F	<i>Nature Med</i> 4 (9):1032–1037 (1998)

O.CM.ANT70	L20587	CAMEROON	Vanden-Haesevelde, M	<i>J Virol</i> 68 (3):1586–1596 (1994)
O.CM.MVP5180	L20571	CAMEROON	Gurtler, LG	<i>J Virol</i> 68 :1581–1585 (1994)
01_AE.CF.90CF11697	AF197340	C.A.R.	Anderson, JP	<i>J Virol</i> 74 (22):10752–10765 (2000)
01_AE.CF.90CF402	U51188	C.A.R.	Gao, F	<i>J Virol</i> 70 (10):7013–7029 (1996)
01_AE.CF.90CF4071	AF197341	C.A.R.	Anderson, JP	<i>J Virol</i> 74 (22):10752–10765 (2000)
01_AE.TH.93TH057	AF197338	THAILAND	Anderson, JP	<i>J Virol</i> 74 (22):10752–10765 (2000)
01_AE.TH.93TH065	AF197339	THAILAND	Anderson, JP	<i>J Virol</i> 74 (22):10752–10765 (2000)
01_AE.TH.93TH253	U51189	THAILAND	Gao, F	<i>J Virol</i> 70 (10):7013–7029 (1996)
01_AE.TH.95TNIH047	AB032741	THAILAND	Auwanit, W	<i>ARHR</i> 16 (3):283–289 (2000)
01_AE.TH.CM240	U54771	THAILAND	Carr, JK	<i>J Virol</i> 70 (9):5935–5943 (1996)
01_AE.TH.TH022	AB032740	THAILAND	Auwanit, W	<i>ARHR</i> 16 (3):283–289 (2000)
02_AG.CM.MP807	AJ286133	CAMEROON	Montavon, C	<i>J AIDS</i> 23 (5):363–74 (2000)
02_AG.FR.DJ263	AF063223	FRANCE	Carr, JK	<i>Virology</i> 247 (1):22–31 (1998)
02_AG.FR.DJ264	AF063224	FRANCE	Carr, JK	<i>Virology</i> 247 (1):22–31 (1998)
02_AG.GH.G829	AF184155	GHANA	Candotti, D	<i>J Med Virol</i> 62 (1):1–8 (2000)
02_AG.NG.IBNG	L39106	NIGERIA	Howard, TM	<i>ARHR</i> 10 (12):1755–1757 (1994)
02_AG.SE.SE7812	AF107770	SWEDEN	Laukkanen, T	Unpublished
02_AG.SN.MP1211	AJ251056	SENEGAL	Toure-Kane, C	<i>ARHR</i> 16 (6):603–609 (2000)
02_AG.SN.MP1213	AJ251057	SENEGAL	Toure-Kane, C	<i>ARHR</i> 16 (6):603–609 (2000)
03_AB.RU.KAL153-2	AF193276	RUSSIA	Liitsola, K	<i>ARHR</i> 16 (11):1047–53 (2000)
03_AB.RU.RU98001	AF193277	RUSSIA	Liitsola, K	<i>ARHR</i> 16 (11):1047–53 (2000)
04_cpx.CY.94CY032-3	AF049337	CYPRUS	Gao, F	<i>J Virol</i> 72 (12):10234–10241 (1998)
04_cpx.GR.97PVCH	AF119820	GREECE	Nasioulas, G	<i>ARHR</i> 15 (8):745–758 (1999)
04_cpx.GR.97PVMY	AF119819	GREECE	Nasioulas, G	<i>ARHR</i> 15 (8):745–758 (1999)
05_DF.BE.VI1310	AF193253	BELGIUM	Laukkanen, T	<i>Virology</i> 269 (1):95–104 (2000)
05_DF.BE.VI961	AF076998	BELGIUM	Laukkanen, T	<i>Virology</i> 269 (1):95–104 (2000)
06_cpx.AU.BFP90	AF064699	AUSTRALIA	Oelrichs, RB	<i>ARHR</i> 14 (16):1495–1500 (1998)
06_cpx.ML.95ML127	AJ288982	MALI	Montavon, C	<i>ARHR</i> 15 (18):1707–1712 (1999)
06_cpx.ML.95ML84	AJ245481	MALI	Montavon, C	<i>ARHR</i> 15 (18):1707–1712 (1999)
06_cpx.SN.97SE1078	AJ288981	SENEGAL	Montavon, C	<i>ARHR</i> 15 (18):1707–1712 (1999)
10_CD.TZ.BFL061	AF289548	TANZANIA	Koulinska, I	<i>ARHR</i> 17 (5):423–431 (2001)
10_CD.TZ.BFL071	AF289549	TANZANIA	Koulinska, I	<i>ARHR</i> 17 (5):423–431 (2001)
10_CD.TZ.BFL110	AF289550	TANZANIA	Koulinska, I	<i>ARHR</i> 17 (5):423–431 (2001)
11_cpx.CM.MP818	AJ291718	CAMEROON	Peeters, M	Unpublished (2000)
11_cpx.FR.MP1298	AJ291719	FRANCE	Peeters, M	Unpublished (2000)
11_cpx.FR.MP1307	AJ291720	FRANCE	Peeters, M	Unpublished (2000)
11_cpx.GR.GR17	AF179368	GREECE	Paraskevis, D	<i>ARHR</i> 16 (9):845–855 (2000)
CPZ.CD.CPZANT	U42720	CONGO	Vanden-Haesevelde, MM	<i>Virology</i> 221 (2):346–350 (1996)
CPZ.CM.CAM3	AF115393	CAMEROON	Corbet, S	<i>J Virol</i> 74 :529–534 (2000)
CPZ.CM.CAM5	AJ271369	CAMEROON	Muller-Trutwin, M	<i>J Med Primatol</i> 29 (3–4):166–172 (2000)
CPZ.GA.CPZGAB	X52154	GABON	Huet, T	<i>Nature</i> 345 (6273):356–359 (1990)
CPZ.US.CPZUS	AF103818	U.S.	Gao, F	<i>Nature</i> 397 (6718):436–441 (1999)

Table 2. Notes on selected full-length HIV-1/SIVcpz complete genomes

Name	Accession	Country	Author	Reference
A.KE.Q23	AF004885	KENYA	Poss, M	<i>J Virol</i> 73 (7):5255–5264 (1999)
<p>This subtype A sequence was derived from a woman from Mombasa, Kenya, who had been recently infected with HIV-1. The blood sample was drawn in June 13, 1994. An env gene fragment from a PCR amplification from an earlier blood sample (July 1993) was published in Poss, M., <i>et al.</i> (ARHR 13(6):493–499 (1997)). The full length sequence was kindly released prior to publication by M. Poss and colleagues, U. Washington. Many env sequences from this same patient are available with accession numbers AF004893 and AF047979-AF048346.</p>				
A.SE.SE6594	AF069672	SWEDEN	Carr, JK	<i>AIDS</i> 13 (14):1819–1826 (1999)
<p>This sequence was from a 29 year old male sampled in Sweden in 1993, who is believed to have been infected in Uganda via heterosexual contact, before moving to Sweden. The patient was CDC-B3 at the time of sampling. This isolate has a syncytium-inducing phenotype and uses the CXCR4 coreceptor for entry. Small sections of the 5' and 3' LTRs are not included in this sequence.</p>				
A.SE.SE7253	AF069670	SWEDEN	Carr, JK	<i>AIDS</i> 13 (14):1819–1826 (1999)
<p>This subtype A sequence is from a 27 year old male living in Sweden, who is thought to have been infected in Somalia via heterosexual contact. The patient was CDC stage C3 when sampled in 1994. The virus is NSI and uses the CCR5 coreceptor. The patient's CD4 count was zero. Virus was cocultured with donor PBMC before PCR amplification and direct sequencing. Small sections of the 5' and 3' LTRs are not included in this sequence.</p>				
A.SE.SE7535	AF069671	SWEDEN	Carr, JK	<i>AIDS</i> 13 (14):1819–1826 (1999)
<p>This sequence is from a 45 year old male sampled in Sweden in 1994, who is believed to have been infected in Uganda via heterosexual contact, before moving to Sweden. The patient was CDC-A1 at the time of sampling. This isolate has a syncytium-inducing phenotype and uses the CXCR4 coreceptor for entry. Small sections of the 5' and 3' LTRs are not included in this sequence.</p>				
A.SE.SE8538	AF069669	SWEDEN	Carr, JK	<i>AIDS</i> 13 (14):1819–1826 (1999)
<p>This sequence was generated in Sweden from a cultured blood sample from a 24-year old Tanzanian woman. Probable country of infection is Tanzania. The virus is NSI, CCR5+ and CXCR4-. Small sections of the 5' and 3' LTRs are not included in this sequence.</p>				
A.SE.SE8891	AF069673	SWEDEN	Carr, JK	<i>AIDS</i> 13 (14):1819–1826 (1999)
<p>This sequence was generated in Sweden from a PBMC co-culture of a sample taken from a 26-year-old Ugandan woman. Probable infection country is Uganda, risk factor heterosexual. The virus is CCR5+, CXCR4-. Small sections of the 5' and 3' LTRs are not included in this sequence.</p>				
A.SE.UGSE8131	AF107771	SWEDEN	Laukkanen, T	-
<p>This subtype A sequence is from a 32 year old female living in Sweden, who is thought to have been infected in Uganda via heterosexual contact. The patient was CDC stage C3 when sampled in 1995. The virus is SI and can use both the CCR5 and CXCR4 coreceptor. Virus was cocultured with donor PBMC before PCR amplification and direct sequencing. Small sections of the 5' and 3' LTRs are not included in this sequence.</p>				
A.UG.U455	M62320	UGANDA	Oram, JD	<i>ARHR</i> 6 (9):1073–1078 (1990)
<p>This sequence is from the 1985 Ugandan isolate U455. It was cloned in phage, and has defective env, vpr, and vpu. The env ORF in this sequence is interrupted by an in-frame stop codon beyond the COOH end of the V5 region. This sequence clusters with subtype A HIV-1.</p>				
A.UG.92UG037	U51190	UGANDA	Gao, F	<i>J Virol</i> 70 (3):1651–1657 (1996)
<p>This sequence is from a complete genome PCR amplified from proviral DNA. The patient was a 31 year old asymptomatic female from Entebe, Uganda. See also U09124, U09127. 92UGO37 is one of a set of three complete genomes from a study linking the HIV-1 epidemic in the heterosexual population in Thailand to an A/E recombinant. It is obtained through the WHO Global Programme on AIDS (WHO</p>				

Network, (ARHR 10(11):1327–1343 (1994))) and comes from an asymptomatic 31-year old female from Entebe, Uganda; she had not taken any anti-retroviral therapy prior to sampling. The risk factor for infection was heterosexual contact. The isolate 92UGO37 was established and propagated by short term cocultivation with normal donor lymphocytes and then the near full length genome was PCR amplified and sequenced. 92UGO37 is subtype A. An LTR sequence is available under accession number U51287 and an additional env/nef sequence with accession number U09127. There is an inframe stop codon in pol at position 3144 in this clone. The isolate from which this sequence was derived is NSI and uses CCR5 or CCR8 (Bjorndal, A., *et al.*, J Virol 71(10):7478–87 (1997) and Rucker, J., *et al.*, J Virol 71(12):8999–9007 (1997). See also Gao, F., *et al.*, J Virol 70(10):7013–7029 (1996). This sequence was kindly made available prior to publication, and is now published (Gao F., *et al.*, J Virol 72(7):5680–5698 (1997)). Biotypes were determined by MT-2 syncytium assay; however, both syncytium-inducing (SI) and non-syncytium-inducing (NSI) variants may be present in the viral “swarm” for each isolate. Recent studies indicate that NSI isolates contain predominantly CCR5-using variants while most SI isolates contain both CXCR4 (SI) and CCR5 (NSI) variants. Some SI isolates may contain dual-tropic variants that use both CXCR4 and CCR5 co-receptors. The isolate 93UG037 is available from the NIH AIDS Reagent program, and is NSI R5.

- | | | | | |
|---|----------|-----------|--------------|----------------------------|
| A2.CD.97CDKTB48 | AF286238 | CONGO | Feng, G | Unpublished (2000) |
| A2.CY.94CY017-41 | AF286237 | CYPRUS | Gao, F | Unpublished (2000) |
| B.AU.MBC200 | AF042100 | AUSTRALIA | Oelrichs, RB | ARHR 14(9):811–814 (1998) |
| Isolate MBC200. Date of sample and source of isolation:18/3/1986, Melbourne, Australia. Biological source:Peripheral blood co-cultured with donor PBMC by the Victorian Infectious Diseases Reference Laboratory. The patient was a Caucasian homosexual male, diagnosed with AIDS in December 1985 at which time the T4:T8 ratio was 0.2. Virus isolation:Biological cloning by three rounds of limiting dilution in donor PBMC. Sequencing: Derived directly from the biologically cloned isolate. Hirt supernatant DNA was obtained from low-passage number donor PBMC culture and sequence derived from overlapping PCR products. All open reading frames are intact and the nucleic acid sequence clusters within subtype B in p17, pol, env and nef. Viral phenotype:Produces high levels of syncytia in PBMC and MT-2 cells. Grows well in Jurkat cells and primary macrophages (Kiernan,R. ARHR 6(6):743–752) | | | | |
| B.AU.MBC925 | AF042101 | AUSTRALIA | Oelrichs, RB | ARHR 14(9):811–814 (1998) |
| Isolate MBC925. Date of isolation:15/5/1987, Melbourne, Australia. Biological source:Cerebrospinal fluid cultured with donor PBMC by the Victorian Infectious Diseases Reference Laboratory. The patient was a Caucasian homosexual male, diagnosed with AIDS in October 1986. Neurological symptoms were present. In December 1986 the T4:T8 ratio was 0.14. The patient died in May 1987. Virus isolation:Biological cloning by three rounds of limiting dilution in donor PBMC. Sequencing: Derived directly from the biologically cloned isolate. Hirt supernatant DNA was obtained from low-passage number donor PBMC culture and sequence derived from overlapping PCR products. All open reading frames are intact and the nucleic acid sequence clusters within subtype B in p17, pol, env and nef. Viral phenotype:Non-syncytium inducing in PBMC. The virus does not replicate in MT-2 cells or Jurkat cells. Good growth in primary macrophages. | | | | |
| B.CN.RL42 | U71182 | CHINA | Graf, M | ARHR 14(3):285–288 (1998) |
| RL42 was isolated from an asymptomatic IVDU, infected by needle sharing, in Dehong prefecture of Yunnan province South of China. This is near the Laos and Thailand golden drug triangle. The isolate was generated by Prof. Dr. Shao Yiming from the Chinese Academy of Preventive Medicine, Beijing, China. This sequence is of the Thai B' subtype (a subset of subtype B), which is the most prevalent subtype of HIV-1 found in the Yunnan province of Southwest China. | | | | |
| B.DE.D31 | U43096 | GERMANY | Kreutz, R | ARHR 8(9):1619–1629 (1992) |
| The patient from which this virus is derived, has never been well described. It is only shown as HIV1-D31 in figure 3 of the paper. The complete genome has been sequenced. | | | | |
| B.DE.HAN | U43141 | GERMANY | Sauermann, U | ARHR 6(6):813–823 (1990) |
| Isolate HAN was isolated from a 39 year old homosexual German patient with AIDS related complex, in 1986. This patient died from complications of AIDS in 1987. HAN was highly cytopathic in MT-2 T cell | | | | |

line, it was able to productively infect MT-4, H9 or Jurkat cell lines. Genomic DNA from infected MT-2 cells was used to prepare a lambda phage genomic library. Two full-length clones, HAN2/2 and HAN2/3 were purified. HAN2/3 was used for DNA sequencing, and has a defective env gene.

B.ES.89SP061 AJ006287 SPAIN Olivares, I *ARHR* 14(18):1649–1651 (1998)

This sequence was generated from a molecular clone. This clone in turn was derived from a biological clone out of a co-cultured sample from a 4-year-old boy who was presumably infected via his mother, who was a Spanish intravenous drug user. The sample was taken in 1989; the infection date is unknown. The virus is subtype B, the clone has the SI phenotype but does not display the “SI-indicating” mutations in the V3 loop.

B.FR.HXB2 K03455 FRANCE Wong-Staal, F *Nature* 313(6000):277–284 (1985)

This sequence was from provirus cloned in lambda phage and is derived from the IIIB isolate related to LAI. This clone has been extensively studied. sites; tat mRNA and other transcript boundaries. Sequence for [25] kindly provided in computer-readable form by L.Ratner, 19-AUG-1986. The HXB2R sequence is being used as a reference genome for all the HIV entries because it has been derived from a demonstrably infectious clone. Hence not all of the “sites” references above were concerned with this isolate. Since the earliest appearance of this sequence in the HIV database and in GenBank (prior to the 1987 publication of [3]), the 5’LTR appears to have been derived from BH10 rather than HXB2. With this corrected version, the 3’LTR only is shown with annotation of differences from [1] so as to facilitate coordination with [3] and with the latest information on splice sites. Revisions were made by [5], [6], and [7] with approval of the principal author of [3]. These affect restriction site analyses, in particular upstream of the gag cds start and in the envelope cds; the latter affect the RRE sequence at 7266. Many of the revisions bring the HXB2 sequence closer to the BH10 sequence, yet these were the differences receiving the greatest attention in [3]. Possibly mutation has occurred in the sub-cloning and outgrowth, or possibly the more recent HXB2 subclones are BH10 contaminants. Be that as it may, HXB2 clones currently in use have a sequence most closely approximated by the sequence below. For a full comparison of the IIIB/LAV sibling sequences, see Part III. 2 additional changes to the sequence of HIVHXB2R were made per MarvinReitz (personal correspondence), 6/91: the “t” at site 8383 was changed to “c”, and the “g” at site 8427 was changed to “a”. The vpU cds not annotated below do not possess a start codon in the normal position (bases 5608 to 5610; “ACG”). Schwartz *et al.*, *J. Virol.* 64: 2519–2529, state that HXB2 does not produce vpU protein (it remains an infectious clone). The minimal continuous RRE (CAR) of 204 nt is defined by [7] to start at coordinate 7327. Dr. Seth Pincus *et al.* [8] report a single base deletion after codon 686 of the envelope gene in an “E variant” set of clones, which results in premature termination of translation and the production of a truncated gp160, causing a marked decrease in the expression of envelope on the surface of the infected cell. This sequence is from the French isolate LAI (formerly BRU) which is also referred to as IIIB. (Wain-Hobson85). Also see:(Alizon86), (Lukashov95b) and (Wain-Hobson91). GenBank accession numbers K02013, L23090-L23103, X01762, L48380-L48399, M64178-M64223, M64406-M64415 and M64768-M64775, AF033819. Other sequences which are of this type include:PV22, K02083; MFA, M33943 (Stevenson90); un-named, Z11530; BH8, K02011; BH10, M15654; TH4, L31963; MCK1, D86068; PM213, D86069; F12CG, Z11530; and HXB, K03455, M38432, M64775 and M14100. The variation of the IIIB isolate in culture was studied by (Lockey96), GenBank accession numbers U54647, U54649, U54651, U54653, U54655, U54657, U54659, U54665, U54667, U54681, U54683, U54685 and U54689. The variation of IIIB/LAI in 9 years of infection in a chimpanzee has been studied by Fultz *et al.* unpublished, GenBank accession numbers U56866-U56883 and U56888-U56899. The IIIB/LAI isolate of HIV-1 has also been extensively studied in cases such as the infected lab worker. See for example (Reitz94), (Pincus94) U12030-U12055. The tropism of isolates from the lab worker for primary PBMCs and failure to grow in T-cell lines was localized to the V3-loop by Lishan Su *et al.*(Su97). Recombinant virus pNL4–3, with envelope from LAI(BRU) and gag-pol from NY5 has also been studied:(Adachi86) GenBank accession number M19921, (Duensing95) GenBank accession number L42371 and (Salminen95) GenBank accession number U26942. Other GenBank entries with IIIB-LAI sequences can be found in the patented sequences section and in the cloning vector section (for example U19867 and A00647)

B.GA.OYI M26727 GABON Huet, T *AIDS* 3(11):707–715 (1989)

This sequence is derived from the Gabonese isolate OYI, designated elsewhere as isolate 397, was obtained from a healthy HIV-1 infected individual presenting an atypical Western Blot. This sequence is from a lambda phage clone, the cloned provirus being functionally defective. The vpu gene does not have a start

codon. Phylogenetic analysis reveals that the sequence is closely related to the North American isolate SF2 and the European virus HAN (across the genome). This is the first report of a virus from Africa that clusters with North American rather than Zairean viruses: OYI and SF2 differ by approximately 7% in envelope. The single C → S substitution at residue 22 of the OYI tat protein renders it inactive, but may not account for the avirulence of the virus. Sibling sequences for OYI(397) are available (see *Human Retroviruses and AIDS 1989* page I-A-181).

- B.GB.CAM1** D10112 U.K. McIntosh, AA Unpublished (1991)
This sequence is from the British isolate CAM1. It has a defective vpu gene. McIntosh A, and Karpas A, Thesis (1991), Cambridge University, England. GenBank accession numbers D10112, D00917
- B.GB.MANC** U23487 U.K. Zhu, T *Nature* **374**(6522):503–504 (1995)
Kindly provided in electronic form by Dr. David Ho, Aaron Diamond AIDS Research Center, New York City, (212)-725-0018. This sequence ostensibly represents HIV-1 captured by PCR amplified from the 1959 sample “Manchester sailor” kidney tissue (see Corbitt. G., *et al.*, *Lancet* 336:51 (1990)). The sequence of the complete genome is available, and it is indistinguishable from contemporary subtype B HIV-1 sequences in phylogenetic analysis. This information together with the observation that additional tissue samples were HIV PCR negative, suggests that the HIV clone that came from MANC kidney sample was very likely to be a contemporary clinical contaminant. The sequence was assembled from multiple PCR amplified fragments. All reading frames in this sequence are intact.
- B.NL.3202A21** U34604 NETHERLANDS Guillon, C *ARHR* **11**(12):1537–1541 (1995)
This sequence is from a complete genome of a macrophage tropic, NSI clone, from an isolate taken from the PBMC of a patient who was in transition from NSI to SI phenotype bulk phenotype. An SI isolate from the same patient has also been completely sequenced, ACH320.2A.1.2 The patient, isolates and phenotype of the molecular clones are described in Groenink, M, *J Virol* 65:1968–1975 (1991).
- B.TW.TWCYS** AF086817 TAIWAN Huang, LM Unpublished
This as-yet unpublished sequence contains a frameshifting single-base insertion in the pol gene, at 3203–3209 where a run of 7 “A”s should have only 6.
- B.US.DH123** AF069140 U.S. Shibata, R *J Virol* **69**(7):4453–4462 (1995)
The DH12 isolate has been extensively characterized. It is dual-tropic, using CCR3, CCR5 and CXCR4 coreceptors. Chimeric molecular clones with the macrophage-tropic AD8 isolate have been made, showing that either V1-V2 or V3 regions of env from DH12 can confer the ability to use CXCR4 onto AD8. The DH12 isolate was passaged in human and chimpanzee PBMCs prior to cloning. A complete genome sequence is available. See also GenBank accession number AF069139.
- B.US.JRCSF** M38429 U.S. O’Brien, WA *Nature* **348**:69–73 (1990)
JRCSF and JRFL (see U63632) were isolated from cerebral spinal fluid and brain tissue respectively of the patient JR, who died with Kaposi’s sarcoma and severe AIDS encephalopathy (*Science* 236:819–822, 1987). JRCSF is from an infectious lambda phage clone of the 1986 isolate JRCSF. Both clones are infectious, but JRFL productively infects macrophages while JRCSF does not. (Peripheral blood was not available from the patient). Second receptor usage for this isolate was defined by Simmons *et al.*, *J Virol* 70:8355 (1996) and Zhang *et al.*, *Nature* 383:768 (1996). The JRCSF and JRFL env nucleotide sequences differ by at least 3%. Both manifest insertions in nef previously reported for HIVBRVA. Patient JR was a 29 year old homosexual male with a history of multiple sexual partners, including one who died of AIDS with AIDS related dementia. Patient JR died in June 1986 and had AIDS related dementia. Complete autopsy findings are reported in [3]. The entry with accession number U45960 is from a mutated clone of JR-CSF which shows a wider range of tropism for cultured cells and increased syncytium-inducing capability. The sequence was kindly provided in computer-readable form by Irvin Chen, UCLA School of Medicine, Los Angeles. On July 16, 1997, Christopher Buck wrote:... [there are] some variations between the published sequence of proviral clone JR-CSF and the actual sequence of plasmid pYK-JRCSF (AIDS Reagent Program catalog #2708). The substitutions are as follows: C7808 → G G7809 → C G8232 → A resulting in these amino acid substitutions in envelope reading frame: R525 → A G666 → D The isolate JR-CSF is available from the NIH AIDS Reagent program, and is NSI R5. References: Simmons *et al.*, *J Virol* 70:8355 (1996) MEDLINE:97126031 Zhang *et al.*, *Nature* 383:768 (1996) MEDLINE:97048157 *Science* 236:819–822 (1987) MEDLINE:87206194.

- B.US.JRFL U63632 U.S. O'Brien, WA *Nature* **348**:69–73 (1990)
- Patient JR was a 29 year old homosexual male with a history of multiple sexual partners, including one who died of AIDS with AIDS related dementia. Patient JR died in June 1996 and had AIDS related dementia. Both clones are infectious, but JRFL productively infects macrophages while JRCSF does not (peripheral blood was not available from the patient). Both JRFL and JRCSF isolates manifest insertions in nef previously reported for HIVBRVA. The PCR clones sequenced were constructed from DNA either amplified directly from brain tissue (JRBR) or from in vitro cultured brain isolates (JRBC). Based on a threshold of at least 95% similarity, the sequences from DNA directly amplified from brain tissue were grouped into two subtypes. The sequences derived from cultured brain isolates showed less than 5% divergence from one another and were grouped into one subtype. This sequence is from an infectious lambda phage clone of the 1986 isolate JRFL, derived from the brain of a patient who died from Kaposi's sarcoma and severe AIDS encephalopathy. The infectious clone JRCSF was isolated from the CSF of the same patient. Both are NSI. The other receptor defined for this isolate is: CCR3. Second receptor usage for this isolate was defined by Deng *et al.*, *Nature* 388:296 (1997) and Farzan *et al.*, *J Exp Med* 186:405 (1997). Kindly provided in computer-readable form by IS Chen, UCLA School of Medicine, Los Angeles CA 90024. JRFL and JRCSF (See M38429 and U45960) were isolated from frontal lobe and cerebral spinal fluid respectively, of patient JR, who died with Kaposi's sarcoma and severe AIDS encephalopathy (*Science* 236, 819–822, 1987). See also: Pang, J., *et al.*, *JAIDS*, 4:1082–92 (1991) and Klasse, P. J., *et al.*, *ARHR* 12:347–350 (1996). The isolate JR-FL is available from the NIH AIDS Reagent program, and is NSI R5. References: O'Brien, W. A., *et al.*, *Nature* 348:69–73 (1990) MEDLINE 91043044 Pang, S., Koyanagi, Y., Miles, S., Wiley, C., Vinters, H. V. and Chen, I. S. High levels of unintegrated HIV-1 DNA in brain tissue of AIDS dementia patients. *Nature* 343, 85–89 (1990) MEDLINE 90114416 Clark, G. L. and Vinters, H. V. Dementia and ataxia in a patient with AIDS *Western J. of Med.* 146 (1), 68–72 (1987) MEDLINE 87151465 Simmons *et al.*, *J Virol* 70:8355 (1996) MEDLINE 97126031 Zhang *et al.*, *Nature* 383:768 (1996) MEDLINE 97048157
- B.US.MNCG M17449 U.S. Gurgo, C *Virology* **164**(2):531–536 (1988)
- MN is from one of the earliest available isolates, and is a commonly used reference and vaccine strain. The MN isolate was taken from a 6 year old male pediatric AIDS patient from the area of Newark, New Jersey, USA in 1984. His mother was an IV drug user who died of pneumonia in 1982. His father was also HIV seropositive. Other sequences from this patient from the 1984 blood sample and from a 1987 sample taken shortly before death (U72495) are available. See also L48364-L48379. The MN sequence was cloned from the isolate in lambda phage. The coding sequences for pol, nef and vpu are prematurely truncated; pol shows an in-frame stop codon at 3783, nef and vpu are prematurely truncated at position 9357 and position 6142 respectively. A set of V3 sequences from this isolate are available (Accession #s L48364-L48379, Lukashov, V. and Goudsmit, J., *AIDS* 9:1307–1311 (1995). In Lori *et al* [2], the pol gene is not noted to be defective in the ST.1 clone of HIV-1 isolate MN. Another complete genome of the MN isolate is available with accession number AF075719 and it too has defective genes, although not pol nor vpu. This sequence was kindly provided in computer readable format by M. Reitz, N.C.I., Bethesda, MD. 20892 U.S.A. The isolate MN is available from the NIH AIDS Reagent program, and is X4.
- B.US.NY5CG M38431 U.S. Willey, RL *PNAS USA* **83**(14):5038–5042 (1986)
- This sequence is from the 1984 New York T-cell tropic isolate NY5. It was cloned in lambda phage and is not replication competent. It has a defective vpu gene due to the loss of the start codon. Supernatant DNA extracted from A3.01 cells infected with the NY5 HIV isolate stock was digested with EcoRI and cloned into lambda WESB. The insert is an EcoRI permuted single LTR clone and was then transferred into pBR322. See also GenBank accession number K03346, for an env sequence of this isolate. A computer readable copy of this sequence was kindly provided by Chuck Buckler, 01-NOV-1988.
- B.US.RF M17451 U.S. Starcich, BR *Cell* **45**(5):637–648 (1986)
- RF (also designated HAT because the virus was isolated from a patient of Haitian descent) is among the first isolates, and is among the commonly used reference and vaccine study strain. The sequence is from the full-length lambda phage clone HAT-3, from isolate RF, cultured in HUT-78 cells. RF is from a 28 year old symptomatic Haitian male, who moved to the U.S. at age 25, in 1980 and was sampled in 1983, shortly before his death in Dec. 1983. He had no history of IV drug use, homosexuality or blood transfusions. In October 1983, he had 20 lb weight loss, giardia with diarrhea, thrush, and diffuse lymphadenopathy. His CD4/CD8 ratio was 0.08. Primary culture from a November 1983 blood sample was co-cultured on HUT-

78 cells. GenBank accession numbers are M17451 and M1250. RF has defective gag and vpu genes. Several env genes are available from this isolate. GenBank accession numbers are U30778-U30781. The sequenced clone did not have the base “a” at position 640 required for gag translation. Two differences in the restriction map arise in comparison to an earlier published map for lambda-HAT (Hahn, B.H. *et al.*, PNAS USA 82:4813 (1985)): i) a Bgl-II site is found at position 193 of the sequence and ii) the HindIII site reported by Hahn *et al.* at position 2000 is not present in this sequence. See also Reitz M., *et al.*, ARHR 8:1950 (1992)

B.US.SF2 K02007 U.S. van Beveren, CP *RNA Tumor Viruses*, 2nd Edition
2:1124–1141; Cold Spring Harbor Laboratory, (1985)

This sequence is from an infectious phage clone from the US isolate ARV-2. ARV-2/SF2 was isolated from the PBMC of a patient with oral candidiasis after co-culture with mitogen stimulated PBMCs, (Levy, J., *et al.*, Science 225:840–842, (1984)). It is a standard reference strain, and has been used for vaccine studies. The isolate SF2 is available from the NIH AIDS Reagent program, and is SI R5X4.

B.US.WCIPR9018 U69591 U.S. Fang, G *J AIDS* 12(4): 352–357 (1996)

This subtype B sequence is from a sample taken in 1990. No information is available about the patient or properties of the virus.

B.US.WEAU160 U21135 U.S. Tozser, J *FEBS Letters* 281:77–80 (1991)

Sequence kindly provided by Sajal K. Ghosh, UAB, Birmingham. A cytopathic HIV-1 virus was cloned from an acutely infected patient in 1990. The clone WEAU 1.60 is replication competent and upon transfection produces highly cytopathic T-cell tropic virus. The clone and the viral isolate from which it was derived are syncytium-inducing (SI). GenBank accession number U21135. The WEAU 1.60 clone was obtained from a coculture of the patient’s PBMCs, first with normal donor PHA-stimulated lymphocytes for 14 days, then with the H9 T-cell line for another 14 days. The patients’ blood specimen was obtained 15 days after the onset of clinical symptoms of acute (primary) infection, and 35 days after a single sexual encounter with a partner whose virus was proven phylogenetically to be responsible for the transmission event. The patient is identified as “Patient #1” in N. Engl. J. Med. 324:954–960 (1991) and as “WEAU 0575” in Science 259:1749–1754 (1993). The patient is also discussed in Borrow *et al.*, Nat Med, 3:205–11 (1997). The WEAU 1.60 clone has been completely sequenced from a plasmid. It is subtype B. There is a deletion of a single T at position 9069, resulting in a frameshift mutation and premature termination of nef. The frameshifting deletion in nef was NOT present in the patients’ uncultured PBMCs where instead there is a “T”. The nef gene was not interrupted in 10 of 10 clones analyzed by PCR sequencing from the uncultured PBMCs. It has been sequenced in its entirety by two different labs (G. Shaw and L. Hood) with 100% concordance.

B.US.WR27 U26546 U.S. Salminen, MO *Virology* 213(1):80–86 (1995)

The clinical isolate WR27 was from a USA patient with a first seropositive sample in 1987 and with clinical progression to Walter Reed stage 5 (Disease stages described in Redfield *et al.* New Engl. J. Med. 314:131–32 (1986)). Blood was drawn for viral isolation in Sept. 1988. The virus was cultured on seronegative donor PBMCs prior to PCR amplification of nearly complete genome from PBMC proviral DNA. The viral sequence had a V3 loop predictive of an SI phenotype. All reading frames were intact except for rev, which had an inframe stop codon in both exons.

B.US.YU2 M93258 U.S. Li, Y *J Virol* 65:3973–3985 (1991)

Other receptors defined for this isolate are:CCR3, GPR15. Second receptor usage for this isolate was defined by Choe *et al.*, Cell 85:1135 (1996) and Farzan *et al.*, J Exp Med 186:405 (1997). Observed phenotype may not correspond with this exact sequence, but to the same isolate. A complete genome from another isolate from this same patient (clone YU-10 accession number M93259) has also been sequenced. YU-2 is a lambda phage clone that is replication competent. It was from the uncultured brain tissue of a patient with AIDS dementia complex. YU2 and YU10 differ by 0.26% in the nucleotide sequence. YU2 was fully replication competent after reconstruction in both primary T lymphocytes and monocyte-macrophages. YU-2 has a defective vpu gene due to the loss of the start codon. See also Li, Y., *et al.*, J Virol 65:3973–3985 (1991). Kindly provided in computer-readable format by Beatrice Hahn.

- C.BR.92BR025 U52953 BRAZIL Gao, F *J Virol* **70**(3):1651–1667 (1996)
 This sequence is from a PCR clone from a primary isolate that is part of a set obtained through WHO Global Programme on AIDS (WHO Network, ARHR 10:1327–1344 (1994)). It is from a 23 year old male hemophilia patient from Porto Alegre, Brazil. He had seroconverted more than 1.2 months prior to the date this blood sample was collected in 1992. He was asymptomatic, and had not taken any anti-retroviral therapy prior to sampling. 92BR025 was established and propagated by short-term co-cultivation with normal donor lymphocytes, and then the near full length genome was PCR amplified and sequenced. The HIV isolate exhibited an NSI phenotype, when assayed by the WHO. See also entries with accession numbers U09126, U15121 and U52953. The full length genome is clone 8, 92BR025.8. This clone has two inframe stop codons in pol at positions 2141, and 3115, and a frame shift mutation at position 4131. This sequence was kindly made available prior to publication. Additional env, nef and ltr region sequences are available from this isolate: U09126, U09132, U51282, and U15121. Biotypes were determined by MT-2 syncytium assay; however, both syncytium-inducing (SI) and non-syncytium-inducing (NSI) variants may be present in the viral swarm” for each isolate. Recent studies indicate that NSI isolates contain predominantly CCR5-using variants while most SI isolates contain both CXCR4 (SI) and CCR5 (NSI) variants. Some SI isolates may contain dual-tropic variants that use both CXCR4 and CCR5 co-receptors. The isolate 92BR025 is available from the NIH AIDS Reagent program, and is NSI R5. A small region of Gag, beginning very near the junction of p1/p6 (position 2131 in the HXB2R K03455 genome; 1471 in this genome) and ending before the end of p6 (position 2237 in HXB2R; position 1579 in this genome) was discovered to be subtype B at the HIV Database. This region includes 3 in/dels that are highly indicative of either subtype B or C, that contribute to the B-like nature of this region in 92BR025, but which would be overlooked if gapstripping were used prior to bootscanning.
- C.BW.96BW01B03 AF110959 BOTSWANA Novitsky, VA *J Virol* **73**(5): 4427–4432 (1999)
 This sequence comes from a study of 23 full length C subtype sequences from Gaborone, Botswana, from eight individuals, sampled in 1996. Three full length sequences from this patient are available: AF110959 - AF110961.
- C.BW.96BW0402 AF110962 BOTSWANA Novitsky, VA *J Virol* **73**(5): 4427–4432 (1999)
 This sequence comes from a study of 23 full length C subtype sequences from Gaborone, Botswana, from eight individuals, sampled in 1996. Five full length sequences from this patient are available: AF110962 - AF110966.
- C.BW.96BW0502 AF110967 BOTSWANA Novitsky, VA *J Virol* **73**(5): 4427–4432 (1999)
 This subtype C sequence is from Botswana. It has 4 NF-Kappa B binding sites where most subtype C have 3, and most other subtypes have just 2. It was kindly provided as a reference strain prior to publication by Dr. Vlad Novitsky, and is part of a study of multiple 23 full length C subtype sequences from Gaborone, Botswana, from eight individuals, sampled in 1996. Two full length sequences from this patient are available: AF110967-AF110968
- C.BW.96BW1104 AF110969 BOTSWANA Novitsky, VA *J Virol* **73**(5): 4427–4432 (1999)
 This sequence comes from a study of 23 full length C subtype sequences from Gaborone, Botswana, from eight individuals, sampled in 1996. Three full length sequences from this patient are available: AF110969-AF110971.
- C.BW.96BW1210 AF110972 BOTSWANA Novitsky, VA *J Virol* **73**(5): 4427–4432 (1999)
 This sequence comes from a study of 23 full length C subtype sequences from Gaborone, Botswana, from eight individuals, sampled in 1996. There is only one full length sequence from this patient.
- C.BW.96BW15C02 AF110974 BOTSWANA Novitsky, VA *J Virol* **73**(5): 4427–4432 (1999)
 This sequence comes from a study of 23 full length C subtype sequences from Gaborone, Botswana, from eight individuals, sampled in 1996. Three full length sequences from this patient are available: AF110973-AF110975.
- C.BW.96BW16B01 AF110976 BOTSWANA Novitsky, VA *J Virol* **73**(5): 4427–4432 (1999)
 This sequence comes from a study of 23 full length C subtype sequences from Gaborone, Botswana, from eight individuals, sampled in 1996. Three full length sequences from this patient are available: AF110976-AF110978.

- C.ET.ETH2220 U46016 ETHIOPIA Salminen, MO *ARHR* **12**(14):1329–1339 (1996)
ETH2220 is the first reported (almost full length) subtype C sequence from Ethiopia. The patient from which this clone was obtained was taken in 1986. In its genomic organization, this clone closely resembles subtype A, B, and D isolates except that the core promoter contains three potential binding sites for the transcription factor NF- κ B instead of containing two. This is a feature that was preserved in other Ethiopian C subtype samples, as well as C viruses from Zambia. This sequence was cloned as a PCR amplified near full length genome, and has a defective tat gene.
- C.IN.93IN101 AB023804 INDIA Mochizuki, N *ARHR* **15**(14):1321–1324 (1999)
The isolate 93IN101 is available from the NIH AIDS Reagent Program, and is NSI R5. Expressed Env protein reacted efficiently with the sera of HIV-1-infected patients of India, but not with sera from HIV-1 + people in Japan.
- C.IN.93IN904 AF067157 INDIA Lole, KS *J Virol* **73**(1):152–160 (1999)
This subtype C sequence is one of several complete genomes from India. It is derived from a cocultured PBMC sample taken March 27, 1993 from a 28 year old woman from Pune in Maharashtra State, India who seroconverted in 1992 following a blood transfusion; NSI phenotype. A small section of the 5' LTR sequence present in the 301904 virus is not included in this sequence. This sample is available in plasmid form through the NIH AIDS Research and Reference Reagent Program and is named p93IN301904. The isolate 93IN904 is available from the NIH AIDS Reagent program, and is R5. Reference:Lole, K.S. *et al.*, *J Virol* **73**(1):152–60 (1999)
- C.IN.93IN905 AF067158 INDIA Lole, KS *J Virol* **73**(1):152–160 (1999)
This subtype C sequence is one of several complete genomes from India. It is derived from an primary culture PBMC sample taken March 27, 1993 from a woman from Pune in Maharashtra State, India who seroconverted in 1992 following a blood transfusion; NSI phenotype. A small section of the 5' LTR sequence present in the 301905 virus is not included in this sequence. This sample is available in plasmid form through the NIH AIDS Research and Reference Reagent Program and is named 93IN905, and is R5.
- C.IN.93IN999 AF067154 INDIA Lole, KS *J Virol* **73**(1):152–160 (1999)
This subtype C sequence is one of several complete genomes from India. This subtype C sequence is one of several complete genomes from India. It is derived from primary PBMC cocultures taken April 24, 1993 from a 52 year old man from Pune in Maharashtra State, India who seroconverted in 1992; risk factors included sex with men and commercial sex workers; NSI phenotype. This sample is available in plasmid form through the NIH AIDS Research and Reference Reagent Program and is named 93IN999, and is R5. A small section of the 5' LTR sequence present in the 301999 virus is not included in this sequence.
- C.IN.94IN11246 AF067159 INDIA Lole, KS *J Virol* **73**(1):152–160 (1999)
This subtype C sequence is one of several complete genomes from India. It is derived from primary PBMC cocultures taken Oct. 25, 1994 from a 26 year old man from Pune in Maharashtra State, India who seroconverted in 1994. His major risk factor for HIV infection was sexual contact with commercial sex workers. A small section of the 5' LTR sequence present in the 11246 virus is not included in this sequence. This sample is available in plasmid form through the NIH AIDS Research and Reference Reagent Program and is named p94IN11246.
- C.IN.95IN21068 AF067155 INDIA Lole, KS *J Virol* **73**(1):152–160 (1999)
A small section of the 5' LTR sequence present in the 21068 virus is not included in this sequence. It is derived from primary PBMC coculture taken Feb. 18, 1995 from a 21 year old man from Pune in Maharashtra State, India who seroconverted in 1994. His only identified risk factor for HIV infection was genital ulcer disease. This sample is available in plasmid form through the NIH AIDS Research and Reference Reagent Program and is named p95IN21068.
- D.CD.84ZR085 U88822 CONGO Gao, F *J Virol* **72**(7):5680–98 (1998)
Sample 84ZR085 was obtained from an AIDS patient from Zaire. The near full length genome was cloned in phage and sequenced. D-84ZR085 is subtype D. There was a frame shift mutation in gag/pol, position 1692. This isolate was obtained from Thomas Jefferson University, and isolate phenotyping information was not available. This sequence was kindly made available prior to publication. Phylogenetic analysis of

the pol protease region done at the HIV-DB indicates that this sequence region is an outlier of the subtype D clade.

D.CD.ELI K03454 CONGO Alizon, M *Cell* **46**(1):63–74 (1986)

This sequence is of a phage clone derived from the Zairean isolate ELI. ELI was recovered in 1983 from a 24 year old woman with AIDS. All reading frames in this sequence are intact. This sequence is from the Zairean isolate ELI. The complete genomic sequence and an infectious clone are available. In *Human Retroviruses and AIDS 1995* pages III-45 and III-47, ELI was listed as an unlikely D/A mosaic, with only gp41 being weakly A-like. Entry with accession number M27949 is from this same isolate.

D.CD.NDK M27323 CONGO Spire, B *Gene* **81**:275–284 (1989)

Kindly provided prior to publication by J.-C. Chermann, Pasteur Institute, Marseille. This is an infectious molecular clone of an isolate estimated to be 10,000 times more cytopathic in vitro than a prototypical HIV-1 (IIIB/LAI). Only minor sequence differences appear to be responsible for the “acute biological effect”. This sequence clusters with HIV-1 subtype D in phylogenetic analysis. It was cloned in phage and is replication competent. All reading frames in this sequence are intact. The NDK virus was isolated from a Zairean man residing in France. He had AIDS and infected his wife; she infected her lover. The cytopathicity of this virus was found not to be localized to env, and might be regulated by regions of gag, vpr and env.

D.CD.Z2Z6 M22639 CONGO Srinivasan, A *Gene* **52**:71–82 (1987)

An infectious molecular clone of this virus was reconstructed. It was cloned in phage and sequenced from the isolate Z2, also called CDC-Z34. All reading frames in this sequence are intact. Total cell DNA extracted from A3.01 cells, infected with the Z2 HIV isolate stock, was digested with XhoI and cloned into lambda J1. The insert is an XhoI permuted single LTR clone, which was then transferred into pIBI. In the sequence below, position one is the first base of the single LTR of the clone, while the last base (9081) is the one just before the LTR of the intact circle. HIVZ2Z6 is an infectious clone and is from the same isolate (CDC Z34; P.Feorino) as HIVZ6. The entry with accession numbers K03458 and M16322, is also from the same isolate. It is defective in Vpr, whereas this entry has an intact Vpr CDS.

D.KE.MB2059 AF133821 KENYA Neilson, JR *J Virol* **73**(5):4393–4403 (1999)

The entry submitted to GenBank, with 10,035 bp of sequence, included non-HIV DNA at both the 5' and 3' ends. The HIV sequence begins at position 363 with ATCCTTGATCTTTGGGT which is not quite at the very 5' end of the normal HIV-1 5' LTR, about 52 bp are missing. Similarly, the 3' end of the genome is incomplete, ending at 9422 with CCAGAAAAGACAAGAGATC, which does not complete the 3' LTR of HIV-1. Another sequence (env gene only) from this patient is found with accession number AF004887.

D.UG.94UG1141 U88824 UGANDA Gao, F *J Virol* **72**(7):5680–5698 (1998)

Sample 94UG114 was obtained from an asymptomatic 31-year-old man from Butuku, Uganda as part of the WHO/UNAIDS study. He had not taken any anti- retroviral therapy prior to sampling. His risk factor for infection was heterosexual contact. The near full length genome was PCR amplified from a short term culture of a PBMC sample and sequenced. The isolate from which this sequence was derived is NSI by an MT-2 assay. This sequence was kindly made available prior to publication by Feng Gao. Biotypes were determined by MT-2 syncytium assay; however, both syncytium-inducing (SI) and non-syncytium-inducing (NSI) variants may be present in the viral “swarm” for each isolate. Recent studies indicate that NSI isolates contain predominantly CCR5-using variants while most SI isolates contain both CXCR4 (SI) and CCR5 (NSI) variants. Some SI isolates may contain dual-tropic variants that use both CXCR4 and CCR5 co- receptors. The isolate 94UG114 is available from the NIH AIDS Reagent program, and is NSI R5.

F1.BE.VI850 AF077336 BELGIUM Carr, JK *Virology* **269**(1):95–104 (2000)

Small sections of the 5' and 3' LTRs are not included in this sequence. This sequence was isolated from a Belgian man in 1993 whose wife was infected in Zaire (now called the Dem Rep of the Congo). This sequence was kindly provided prior to publication by J. Carr *et al.* This sequence was originally classified as subtype F, but because F is readily subdivided into two distinct categories, it has been reclassified as subtype F1. See Triques *et al.*, *Virology*. 259(1):99–109 (1999).

- F1.BR.93BR020-1 AF005494 BRAZIL Gao, F *J Virol* **72**(7):5680–5698 (1998)
 This sequence was originally classified as subtype F, but because F is readily subdivided into two distinct categories, it has been reclassified as subtype F1. See Triques *et al.* *Virology*. 259(1):99–109 (1999). This sample is part of a set of sequences generated through the WHO Global Programme on AIDS (WHO Network, ARHR 10:1327–1344 (1994) and came from an asymptomatic HIV seropositive bisexual contact. The isolate 92BR020 was established and propagated by short term co-cultivation with normal donor lymphocytes, and then the near full length genome was PCR amplified and sequenced. The isolate 92BR020 was described as syncytium inducing (SI) using an MT-2 assay. An envelope gene sequence from this isolate is described in Gao, F., *et al.*, *J Virol* 70:1651–1657 (1996). This sequence was kindly made available prior to publication, and was eventually published in Gao, F. *et al.*, *J Virol* 72(7):5680–98 (1998). There were no defective genes. A summary of isolates with known co-receptor usage can be found in the HIV database reviews.
- F1.FI.FIN9363 AF075703 FINLAND Laukkanen, T Unpublished
 The virus was isolated in 1993 from a Finnish male who claims that he was most likely infected in Finland in 1985 by a Kenyan woman, however, he had multiple partners. This sequence was originally classified as subtype F, but because F is readily subdivided into two distinct categories, it has been reclassified as subtype F1.
- F1.FR.MP411 AJ249238 FRANCE Triques, K *ARHR* **16**(2):139–151 (2000)
 Prior to 1999, the F subtype was made up of 3 distinctive clusters and very diverse viruses. Triques and colleagues have defined a new subtype K, and broken F into two subclusters, F1 and F2. This sequence is a complete genome reference for F1. This isolate was from a French patient who believes he was infected when deployed in Chad or Yugoslavia.
- F2.CM.MP255 AJ249236 CAMEROON Triques, K *ARHR* **16**(2):139–151 (2000)
 Prior to 1999, the F subtype was made up of 3 distinctive clusters and very diverse viruses. Triques and colleagues have defined a new subtype K, and broken F into two subclusters, F1 and F2. This sequences is a complete genome reference for F2.
- F2.CM.MP257 AJ249237 CAMEROON Triques, K *ARHR* **16**(2):139–151 (2000)
 Prior to 1999, the F subtype was made up of 3 distinctive clusters and very diverse viruses. Triques and colleagues have defined a new subtype K, and broken F into two subclusters, F1 and F2. This sequences is a complete genome reference for F2.
- G.BE.DRCBL AF084936 BELGIUM Debyser, Z *ARHR* **14**(5):453–459 (1998)
 Clinical details are discussed in *ARHR* 14(5):453–9 (1998) and the analysis of the complete genome is in Oelrichs *et al.*, *ARHR* 15(6):585–9 (1999). A pregnant 26 year old women was sampled who had lived in Zaire (now called the Dem Rep of the Congo) until 1993, then moved to Belgium. She was diagnosed with AIDS and had a low CD4 when sampled in 1996. Her G subtype virus was not detected by Amplicor Monitor or Nasba RNA kits, although she was found to have a high viral load by branched DNA. The sequence was kindly provided prior to publication by R. Oelrichs *et al.* This complete genome sequence shows the same pattern of phylogenetic associations as 92NG083 (U88826), HH8793 (AF061640) and SE6165 (AF061642). These four (and other subtype G sequences) form their own clade (subtype G) when the complete gag, pol or env genes are included in the analysis. However, all G subtype genomes have some ambiguous A/G regions in the central part of the genome in a phylogenetically indistinct region in the the accessory gene region from the beginning of vif to the beginning of vpu. The coordinates of this region are 5055 to 6297 on HXB2, and there are several subtypes which become difficult to resolve unambiguously in this region: A, G, and the circulating recombinant forms which resemble the prototypes AE(CM240) and AG(IbNG) (per. comm., Jean Carr). Of particular note is that a region of gp41 from these viruses clusters with the AE and IbNG circulating recombinant forms. See J. Carr *et al* *Virology* 247:22–31 (1998) and F. Gao *et al* *J Virol* 72(7):5680–5698 (1998) for analyses of the other genomes with this pattern. At this time (Jan. 1999) it is not clear whether the AE(CM240) circulating recombinant form is AEG triple recombinant, or if the above four genomes are AEG triple recombinant, or if an evolutionary anomaly rather than recombination is the basis for this pattern.

- G.FI.HH8793-12-1 AF061641 FINLAND Salminen, MO *ARHR* 8(9):1733–1742 (1992)
Sections of the 5' and 3' LTRs are not included in this sequence. This sample was taken in June 1993 (Jean Carr, Pers. Communication)
- G.NG.92NG083 U88826 NIGERIA Gao, F *J Virol* 72(7):5680–5698 (1998)
This sequence is from a PCR clone from a primary culture from the NSI isolate 92NG083; the sample was taken in 1992 from an AIDS patient from Jos, Nigeria. The isolate was originally called JV1083, but was renamed 92NG083 to be consistent with WHO nomenclature. The full length clone has an altered initiation codon at position 157, an inframe stop codon at position 360 in gag, and a vpu frameshift mutation at position 5462.
- G.SE.SE6165 AF061642 SWEDEN Carr, JK *Virology* 247(1):22–31 (1998)
Siblings sequences from the same blood sample:L40743, L40761, L40752 Set: Two female sex partners of this individual
Set_IDs:L40744, L40745, L40753, L40754, L40762, L40763.
Sections of the 5' and 3' LTRs are not included in this sequence. This patient (6165) was infected in Congo and moved to Sweden. He had tested HIV positive (ELISA and Western Blot) approximately 18 months prior to infecting patient 6167 via heterosexual intercourse and 19 months prior to infecting patient 6168, also via heterosexual intercourse. He had low (35 per ul) CD4 count and dermatological problems but no AIDS defining illness at the time of sampling in 1993.
- H.BE.VI991 AF190127 BELGIUM Janssens, W *AIDS* 14(11):1533–1543 (2000)
- H.BE.VI997 AF190128 BELGIUM Janssens, W *AIDS* 14(11):1533–1543 (2000)
- H.CF.90CF056 AF005496 C.A.R. Gao, F *J Virol* 72(7):5680–5698 (1998)
This sequence clusters with available HIV-1 subtype H sequences in phylogenetic analysis, and is the first available full length H subtype sequence. The isolate comes from Bangui, in the Central African Republic, and was sampled in 1990, from an asymptomatic individual, who had no anti-retroviral therapy. The isolate had an NSI phenotype by an MT-2 assay, and the sample was obtained from the Pasteur Institute, Bangui. The isolate 90CF056 was established and propagated by short term cocultivation with normal donor lymphocytes, and then the near full length genome was PCR amplified, cloned and sequenced. The isolate was at one point designated 90CR056, but was changed to 90CF056 as CR stands for Costa Rica, and CF for Central African Republic. The first genetic characterization of this virus isolate (an env V3 sequence designated 4056, GB accession number L11497, Murphy *et al.*, *ARHR* 9:997–1006 (1993)) left the subtype designation as unclassified, but a second study of this env region sequence classified it as subtype H (W. Janssens, *ARHR* 10:877–879 (1994)). This sequence was kindly made available prior to publication, and was the first subtype H full length genome available (Gao, F. *et al.*, *J Virol* 72(7):5680–98 (1998)) The patient was heterosexual, asymptomatic, and the biological phenotype of the isolate was NSI. There were no defective genes in the sequence.
- J.SE.SE7022 AF082395 SWEDEN Laukkanen, T *ARHR* 15(3):293–297 (1999)
This sequence is from a woman who was infected in Zaire (now called the Democratic Republic of the Congo) between 1981 and 1986. Blood for sequencing was drawn in 1993. She was asymptomatic with a CD4 count of 184. The sequence was kindly provided prior to publication by M. Salminen. This sequence is from the same individual as SE7022 described by T. Leitner *et al* *ARHR* 11(8):995–997 (1985), see accession numbers L41177 and L41179 for env and gag genes from this individual. Other examples of subtype J have been found in Gambia, see accession numbers U33099, U33100 and U33102.
- J.SE.SE7887 AF082394 SWEDEN Laukkanen, T *ARHR* 15(3):293–297 (1999)
This sequence is from a male who was infected in Sweden between 1993 and 1994. Blood for sequencing was drawn in 1994. He was asymptomatic with a CD4 count of 567. The sequence was kindly provided prior to publication by M. Salminen. This sequence is from the same individual as SE7887 described by T. Leitner *et al* *ARHR* 11(8):995–997 (1985), see accession numbers L41176 and L41178 for env and gag genes from this individual.

- K.CD.EQTB11C AJ249235 CONGO Triques, K *ARHR* **16**(2):139–151 (2000)
Prior to 1999, the F subtype was made up of 3 distinctive clusters and very diverse viruses. Triques and colleagues have defined a new subtype K, and broken F into two subclusters, F1 and F2. This sequence is a complete genome reference for subtype K.
- K.CM.MP535 AJ249239 CAMEROON Triques, K *ARHR* **16**(2):139–51 (2000)
Prior to 1999, the F subtype was made up of 3 distinctive clusters and very diverse viruses. Triques and colleagues have defined a new subtype K, and broken F into two subclusters, F1 and F2. This sequence is a complete genome reference for subtype K.
- N.CM.YBF106 AJ271370 CAMEROON Souquiere, S Unpublished
YBF106 is from a 51 year old male patient with AIDS, from the central part of Cameroon. He was found to be HIV-seropositive in October, 1997 and died of AIDS in December, 1998.
- N.CM.YBF30 AJ006022 CAMEROON Simon, F *Nature Med* **4**(9):1032–1037 (1998)
YBF30 uses CCR5 as a co-receptor, and was also tested on cell lines expressing CCR2b and CCR3 and did not utilize these coreceptors. YBF30 grew on chimpanzee PBMCs, but did not replicate in T-cell lines (MT-2 and HUT78). YBF30 infection of cells was entirely blocked by RANTES alone or in combination with MIP-1alpha and MIP-1beta. YBF30 was isolated from a 40 year old woman who had never traveled outside Cameroon. She presented with *Histoplasma capulatum* infection of the colon in May, 1995 and died of AIDS (cachexia, neurological involvement and suspected disseminated histoplasmosis) in December, 1995. YBF30 was isolated from a May, 1995 blood sample and YBF31 from a December 1995 blood sample. YBF30 and YBF31 are greater than 98% identical to each other and less than 85% identical to HIV-1 M group, HIV-1 O group and SIV-CPZ sequences. The authors propose labeling this and similar viruses as N (between M and O, and also non-O non-M) group HIV-1. The N group designation is based upon sero-epidemiological surveys in Cameroon, and another sequence (YBF105 not yet submitted to the databases) which indicate that more than one patient is infected with this clade of HIV-1. 700 stored sera collected between 1988 and 1997 were serologically tested, with a peptide-based EIA. 611 (87%) were reactive with M-group. 65 (9%) were reactive with O group. 8 were indeterminate. 16 (2%) were reactive with SIV-CPZ and not M or O group, 3 of these were strongly reactive with YBF30 peptides. A partial Pol gene was sequenced for one (YBF105).
- O.CM.ANT70 L20587 CAMEROON Vanden-Haesevelde, M *J Virol* **68**(3):1586–1596 (1994)
ANT70 was isolated from the first O group infection discovered, and the very divergent LTR sequence was published in 1990 (de Leys, R., *et al.*, *J Virol* 64:1207- 1216 (1990)). The isolate came from CDC stage II infected 19 year old female with unusual serological reactivity, who progressed to CDC stage III before publication of [1]. Both a man and his wife were infected with an O group HIV-1. The wife seroconverted in March, 1987. Both the husband and wife were originally from Cameroon, and living in Belgium. The husband was CDC stage III at the time virus was isolated from the wife for sequencing. The wife had a CD4:CD8 ratio of 0.25. Supernatant from the original coculture of wife PBMCs plus PHA-stimulated donor PBMCs, was used to infect MOLT4 clone 8 cells and MT-4 cells. Syncytia were formed in both these cell lines. After several weeks culture, chronically infected cell lines were obtained that shed virus, and supernatant from these stable lines were used for viral RNA isolation. HIV-1 O group viruses have the same genetic organization as HIV-1 M group viruses. For a review see Korber, B., *et al.*, *Human Retroviruses and AIDS Database, Part III*, 41–56 1996.
- O.CM.MVP5180 L20571 CAMEROON Gurtler, LG *J Virol* **68**:1581–1585 (1994)
This isolate was derived from a Cameroonian woman, sampled in 1991, who died of AIDS in 1992. The viral isolate MVP-5180 was grown in several human T-cell lines and the monocytic U937 line. The isolate MVP5180 is available from the NIH AIDS Reagent program, and is SI R5X4. A summary of isolates with known co-receptor usage can be found in the HIV database reviews.
- O_SN_MP1299 AJ302646 SENEGAL Toure-Kane, C Unpublished (2001)
- O_SN_MP1300 AJ302647 SENEGAL Montavon, C Direct Submission (2001)

- 01_AE.CF.90CF11697 AF197340 C.A.R. Anderson, JP *J Virol* **74**(22):10752–10765 (2000)
 In a study of the origins of subtype E, or CRF01, four full length genomes were obtained, two from Thailand, two from the Central African Republic. While several previous studies discussed the recombinant nature of this clade of viruses, Anderson *et al.* questioned this conclusion, noting the regions designated A and E form distinct subclusters, but are monophyletic. The nature of the origins of CRF01 remains controversial. This sequence was derived from end point diluted proviral DNA, amplified using nested PCR, from a sample from an AIDS patient in Bangui, the Central African Republic taken in 1990. V3 sequences from the same isolate are also available: L11519 116977X (clone 7) and L11520 HIV116976X clone 6.8 and L11521 11697–1 clone 1
- 01_AE.CF.90CF402 U51188 C.A.R. Gao, F *J Virol* **70**(10):7013–7029 (1996)
 One of a set of three complete genomes from a study linking the HIV-1 epidemic in the heterosexual population in Thailand to an A/E recombinant. 90CF402, previously named CAR-E 4002 or 90CR402, was obtained from a man from Bangui, Central African Republic, who had lymphadenopathy, diarrhea, severe weight loss and recurrent respiratory infections. He was infected through heterosexual contact, but the year of infection is unknown. 90CR402, GenBank accession number U51188, was first adapted to growth in chimpanzee cells, expanded in chimpanzee cells, and then re-expanded in human PBMCs before lambda cloning and sequencing. The sequence has a defective vpu due to the loss of a start codon, and a defective vif gene. The pattern of subtype A-E recombination of breakpoints is shared between A-E subtype sequences from Thailand and from the Central African Republic, suggesting a shared ancestral recombinant virus that arose prior to the subsequent epidemics in the two areas.
- 01_AE.CF.90CF4071 AF197341 C.A.R. Anderson, JP *J Virol* **74**(22):10752–10765 (2000)
 In a study of the origins of subtype E, or CRF01, four full length genomes were obtained, two from Thailand, two from the Central African Republic. While several previous studies discussed the recombinant nature of this clade of viruses, Anderson *et al.* questioned this conclusion, noting the regions designated A and E form distinct subclusters, but are monophyletic. The nature of the origins of CRF01 remains controversial. This sequence was derived from end point diluted proviral DNA, amplified using nested PCR, from a sample from an AIDS patient in Bangui, the Central African Republic taken in 1990.
- 01_AE.TH.93TH057 AF197338 THAILAND Anderson, JP *J Virol* **74**(22):10752–10765 (2000)
 In a study of the origins of subtype E, or CRF01, four full length genomes were obtained, two from Thailand, two from the Central African Republic. While several previous studies discussed the recombinant nature of this clade of viruses, Anderson *et al.* questioned this conclusion, noting the regions designated A and E form distinct subclusters, but are monophyletic. The nature of the origins of CRF01 remains controversial. This sequence was derived from end point diluted proviral DNA, amplified using nested PCR, from a sample from an IVDU in Bangkok, Thailand taken in 1993. This sample was first characterized in Wasi *et al.*, *AIDS* 1995 9(8):843–9 PMID: 7576317
- 01_AE.TH.93TH065 AF197339 THAILAND Anderson, JP *J Virol* **74**(22):10752–10765 (2000)
 In a study of the origins of subtype E, or CRF01, four full length genomes were obtained, two from Thailand, two from the Central African Republic. While several previous studies discussed the recombinant nature of this clade of viruses, Anderson *et al.* questioned this conclusion, noting the regions designated A and E form distinct subclusters, but are monophyletic. The nature of the origins of CRF01 remains controversial. This sequence was derived from end point diluted proviral DNA, amplified using nested PCR, from a sample from an IVDU in Bangkok, Thailand taken in 1993. This sample was first characterized in Wasi *et al.*, *AIDS* 1995 9(8):843–9 PMID: 7576317
- 01_AE.TH.93TH253 U51189 THAILAND Gao, F *J Virol* **70**(10):7013–7029 (1996)
 This virus was isolated in 1993 from a 21 year old man from Chiang Mai, Thailand, who had end-stage AIDS. The isolate was previously designated CMU010, or 302053. The isolate was expanded in donor PBMCs, then in H9 cells, then a lambda phage clone was generated and sequenced. The sequence has a defective vpu due to the lack of a start codon and a defective env gene. Like other “E” subtype viruses from both Asia and Africa, large stretches of the genome are associated with the A subtype, and all share a common mosaic pattern of A/E breakpoints, suggesting that the currently identified A/E recombinants all share a common ancestor. The breakpoints are mapped in pages III-25–30, of *Human Retroviruses and AIDS 1997* (<http://hiv-web.lanl.gov/COMPENDIUM/1997/hahn.pdf>).

- 01_AE.TH.95TNIH047 AB032741 THAILAND Auwanit, W *ARHR* **16**(3):283–289 (2000)
 Whole blood from an asymptomatic heterosexual 46-year-old Thai man in Ubonrathchathanee province was transported to the National Institute of Health in Nonthaburi, Thailand. PBMCs were separated and co-cultivated with PHA-stimulated donor PBMCs. DNA from p24 antigen-positive culture was used to amplify three fraction of proviral DNA. The complete genomic sequence of the provirus was determined by the compilation of three clones containing different parts of the viral genome.
- 01_AE.TH.CM240 U54771 THAILAND Carr, JK *J Virol* **70**(9):5935–5943 (1996)
 Blood from an asymptomatic heterosexual 21-year-old Thai man was transported from Thailand to the USA where PBMCs were separated and co-cultivated with PHA-stimulated donor PBMCs. DNA from p24 antigen-positive culture was used to amplify the proviral DNA. The complete genomic sequence of the provirus was determined by the compilation of three clones containing different parts of the viral genome. CM240 is an example of a Thai subtype E virus, which is a mosaic of a clade A virus and clade E virus, with the gag gene (and other regions) of subtype E viral genome falling within clade A in phylogenetic analysis. This is the pattern of A-E sequences found throughout Asia and Africa, and no full length E subtype reference strain has been identified (as for 93TH253). Carr *et al.*, provide detailed analysis of the breakpoints, and point out that the A/E mosaic genomes have a natural pseudotype structure where the external envelope protein spikes on the virion essentially are contributed by the E subtype, and the rest of the viral proteins have a subtype A origin. See also the env sequence from the same isolate (L14572), Mascola J., *et al.*, (JID 169:48–54 (1993)).
- 01_AE.TH.92TH022 AB032740 THAILAND Auwamt, W *ARHR* **16**(3):283–289 (2000)
 Whole blood from an asymptomatic heterosexual 25-year-old Thai man in Ubonrathchathanee province was transported to the National Institute of Health in Nonthaburi, Thailand. PBMCs were separated and co-cultivated with PHA-stimulated donor PBMCs. DNA from p24 antigen-positive culture was used to amplify three fractions of proviral DNA. The complete genomic sequence of the provirus was determined by the compilation of three clones containing different parts of the viral genome.
- 02_AG.CM.MP807 AJ286133 CAMEROON Montavon, C *J AIDS* **23**(5):363–74 (2000)
 This sequence was obtained in a study documenting the high prevalence of CRF02_AG (IbNG-like) viruses in West and West Central Africa, between 60–84% of viruses from this region were IbNG-like. 219 viruses were sequenced in env and gag, and most were associated with CRF02_AG phylogenetically and shared a common breakpoint structure. This was one of three full length sequences obtained.
- 02_AG.FR.DJ263 AF063223 FRANCE Carr, JK *Virology* **247**(1):22–31 (1998)
 Carr *et al.* states that this virus was from a French foreign legion soldier assigned to peace-keeping duties in Djibouti, referencing Louwagie *et al* (*J Virol* 69(1):263- 271 (1995)). The Louwagie paper does not mention the French soldiers and only states that the blood sample was from Djibouti. The sample was isolated in 1991. There are several sequences which share AG recombination breakpoints with IbNG, and are essentially the same recombinant recirculating form; IbNG is the prototype, sharing a similar structure with DJ264 and DJ263.
- 02_AG.FR.DJ264 AF063224 FRANCE Carr, JK *Virology* **247**(1):22–31 (1998)
 A small section of LTR sequence present in the DJ263 virus is not included in this file. Carr *et al.* states that this virus was from a French foreign legion soldier assigned to peace-keeping duties in Djibouti, referencing Louwagie *et al* (*J Virol* 69(1):263–271 (1995)). However the Louwagie paper does not mention the French soldiers and only states that the blood sample was from Djibouti. The sample was isolated in 1991. There are several sequences which share AG recombination breakpoints with IbNG, and are essentially the same recombinant recirculating form; IbNG is the prototype, sharing a similar structure with DJ264 and DJ263.
- 02_AG.GH.G829 AF184155 GHANA Candotti, D *J Med Virol* **62**(1):1–8 (2000)
 This sequence is an A/G recombinant of the CRF02_AG form. This sequence was derived from a 29-year-old Ghanaian woman who developed AIDS but was HIV-antibody seronegative. The lack of detectable antibody response might be due to either a highly pathogenic HIV-1 that disabled development of the typical humoral immune response or to a host-related dysfunction of the immune system.

- 02_AG.NG.IBNG L39106 NIGERIA Howard, TM *ARHR* 10(12):1755–1757 (1994)
 HIV-1 IbNg was isolated from the PBMCs of an apparently healthy 23 year old man from Nigeria. The patient's PBMCs were cocultured with PHA-stimulated donor PBMCs from an HIV sero-negative donor. After confirming HIV infection in the culture, a mixture of cells and culture supernatant were used to infect a second culture of donor PBMCs, with fresh PHA-stimulated PBMCs added on days 4 and 6. The cultured cells were harvested on day 8 and cytoplasmic RNA was harvested. RT-PCR was used to amplify the complete HIV-1 genome in 5 overlapping segments. The partial env gene sequence (U48628) was originally designated subtype A (Howard, T., *et al.*, *ARHR* 10:1755–1757 (1994)); as was the full length genome. The full length sequence was eventually shown to be an A-G recombinant with multiple cross-over points (Gao F, *et al.*, *J. Virol* 70:7013 (1996)). The breakpoints are mapped in (Robertson, D., *et al.*, part III pages 25–30, of the 1997 compendium). The IbNg sequence has a 16 bp insertion within the Lys-tRNA primer binding site, just 3' of the 5' LTR. It also has a single nucleotide deletion in tat cds at position 5449. See also the entry with accession number U48628, which is from another isolate taken from this same individual. There are many sequences which share recombination breakpoints with IbNG, and are essentially the same recombinant circulating form; IbNG is the prototype for CRF02_AG, and shares a similar structure with DJ264 and DJ263 (Carr *et al.*, *Virology* 247:22–31 (1998)).
- 02_AG.SE.SE7812 AF107770 SWEDEN Laukkanen, T Unpublished
- 02_AG.SN.MP1211 AJ251056 SENEGAL Toure-Kane, C *ARHR* 16(6):603–609 (2000)
 This full length genome came from a study of subtype prevalence in Senegal, and CRF02_AG IbNg like strains were found to be most common.
- 02_AG.SN.MP1213 AJ251057 SENEGAL Toure-Kane, C *ARHR* 16(6):603–609 (2000)
 This full length genome came from a study of subtype prevalence in Senegal, and CRF02_AG IbNg like strains were found to be most common.
- 03_AB.RU.KAL153-2 AF193276 RUSSIA Liitsola, K *ARHR* 16(11):1047–53 (2000)
 This is the first complete genome sequence of the gag-A/env-B circulating recombinant form which is common among IV drug users in the Kaliningrad region of Russia, and so is considered the prototype of CRF03_AB. The nonrecombinant subtype A and subtype B parents of this recombinant are common in southern Ukraine and in Russia. A gag gene sequence from this same patient is available with accession number AF082414. This sequence was kindly provided by Mika Salminen in 1998 as a CRF03 reference strain, prior to the published description.
- 03_AB.RU.RU98001 AF193277 RUSSIA Liitsola, K *ARHR* 16(11):1047–53 (2000)
 This sequence is subtype A/B recombinant of the CRF03_AB form for which the Kaliningrad sequence KAL153 is a prototype (accession AF193276). The sample was taken in Kaliningrad from a male IDU in 1998.
- 04_cpx.CY.94CY032-3 AF049337 CYPRUS Gao, F *J Virol* 72(12):10234–10241 (1998)
 This sample, like others in this study (see also subtypes A, B, C, and F) were collected in February 1994 from the AIDS clinic in Nicosia, Cyprus. Patient HO31 was a 24 year old asymptomatic female known to have been HIV seropositive for at least 5 years. Patient HO32 was a 35 year old asymptomatic male, also seropositive for at least 5 years. Both were IDUs who had lived in Greece and used IV drugs there, before moving to Cyprus. DNA was extracted from patient PBMCs and PCR amplified. Products were cloned and sequenced. Two env gene clones from HO32 and one from HO31 were sequenced (accession numbers U28672, U28673 and U28685). Complete genome is only available for HO32 (CY032). For patient information see Kostrikis, L.G. *et al.* (*J Virol* 69:6122–6130 (1995)). This sequence has the same genetic recombination breakpoints as PVCH and PVMY, and 94CY032 is the prototype of the circulating recombinant form. The Gao *et al.* 1998 paper characterizes this subtype in greater detail, presenting the first published account of this full length genome. The analysis of C2-V3 env gene sequences confirmed that 94CY032.3 was closely related to sequences previously classified as subtype I. However, the remainder of its genome various regions in which 94CY032.3 was significantly clustered with either subtype A or subtype G. Only regions in vpr, nef, and the middle portions of pol and env, formed independent lineages roughly equidistant from all other known subtypes. Since these latter regions most likely have a common origin, Gao *et al.* classified them all as subtype I, and report that 94CY032 represents a triple recombinant

(A/G/I) with at least 11 points of recombination crossover. Since subtype I is now obsolete and has been found to consist of segments of subtype K as well as unknown regions, the new designation for this CRF should be AGKU.

- 04_cpx.GR.97PVCH AF119820 GREECE Nasioulas, G *ARHR* **15**(8):745–758 (1999)
 The sequence isolated from patient GR11 (accession AF049292) is from the same patient as 97PVCH AF049292. The patient was a 32 year old male IVDU with symptoms (CDC stage B3) in 1991, when sampled. He is no longer living. For patient information see information on patient GR11 in Nasioulas, G. *et al.* (*ARHR* 14(8):685–90 (1998)). This sequence has the same genetic recombination breakpoints as CY032 and PVMY, and is one of the circulating recombinant forms of which CY032 is the prototype.
- 04_cpx.GR.97PVMY AF119819 GREECE Nasioulas, G *ARHR* **15**(8):745–758 (1999)
 The sequence was isolated from a 13 year old whose mother and father were IVDUs. The isolate is also called GR84. This sequence has the same genetic recombination breakpoints as CY032 and PVCH, and is one of the circulating recombinant forms of which CY032 is the prototype. See the previous entry for more information on the mosaic pattern.
- 05_DF.BE.VII310 AF193253 BELGIUM Laukkanen, T *Virology* **269**(1):95–104 (2000)
 This sequence is subtype D/F recombinant and is one of the CRF05_DF circulating recombinant form. Another genome of this form is available with accession number AF076998. CRF05_DF sequences have been found in individuals with links to the Democratic Republic of the Congo; VII310 was isolated from a woman who was infected by a man who traveled frequently in the DRC.
- 05_DF.BE.VI961 AF076998 BELGIUM Laukkanen, T *Virology* **269**(1):95–104 (2000)
 Small sections of the 5' and 3' LTRs are not included in this sequence. This sequence seems to be recombinant/mosaic between subtypes D and F. Another sequence of the env gene from this same patient is available with accession number X96530. CRF05_DF sequences have been found in individuals with links to the Democratic Republic of the Congo; VI961 was isolated from the partner of a seropositive man from the DRC.
- 06_cpx.AU.BFP90 AF064699 AUSTRALIA Oelrichs, RB *ARHR* **14**(16):1495–1500 (1998)
 HIV-1 from Burkina Faso, identified in Australia. The patient is a 32 year old African male who acquired the infection heterosexually in 1991. The patient was diagnosed in August 1996 at which time he had a CD4 count of 125. This sequence was derived by PCR directly from patient PBMCs when therapy-naive. The genome of this virus is a mosaic between subtypes A, G and J. The LTR is subtype J, a segment of about 950 bp at the beginning of gag is subtype A, the remainder of gag and part of the protease sequence are subtype G. Most of the pol gene cannot be assigned to a subtype. The mid-genome accessory region is mostly subtype J. gp120 is subtype G. The 3' region of gp41, the third exons of tat and rev, and the nef gene are subtype J. The subtype G protease region clusters tightly in phylogenetic analysis, with three sequences from the Ivory Coast (accession numbers AF000482, AF000491 and AF000492), indicating that this is possibly a circulating recombinant form with 4 isolates from 2 different countries obtained to date (Aug 1999); BFP90 is now considered the prototype isolate of CRF06_cpx.
- 06_cpx.ML.95ML127 AJ288982 MALI Montavon, C *ARHR* **15**(18):1707–1712 (1999)
 This sequence was submitted with 95ML84, (AJ245481), and classified as the same CRF, however no details are provided about the source or phylogeny in the Montavon *et al.* paper.
- 06_cpx.ML.95ML84 AJ245481 MALI Montavon, C *ARHR* **15**(18):1707–1712 (1999)
 This sequence is AGJK recombinant with an identical mosaic pattern to BFP90, the prototype for CRF06, with accession number AF064699. This sequence is from a sample collected as part of a survey of female commercial sex workers in Bamako, Mali in 1995. The sample was taken from a 35 year old woman with no signs of illness, whose blood was seropositive for both HIV-1 and HIV-2. Montavon *et al.* note that HIV-2 sequences of the HIV-2 subtype A form were sequenced from this patient, but no database entry has yet been created and also that sequences from 95ML84 were analyzed in Peeters *et al* *ARHR* 14(1):51–58 (1998) and classified as subtype G. But the closest sequence among those was labeled 95ML74 and is only 89% identical to this sequence, and there is no ML84 in that publication. The authors have been contacted for further clarification of this issue (04/14/00). This CRF is found in several West African nations, Senegal, Mali, and Burkina Faso.

- 06_cpx.SN.97SE1078 AJ288981 SENEGAL Montavon, C *ARHR* **15**(18):1707–1712 (1999)
This sequence is noted in Montavon *et al.* to be most closely associated with the CRF06_cpx virus AJ245481 from Mali, ML84. CRF06 is found in several West African nations, Senegal, Mali, and Burkina Faso. They present in depth characterization of the sample from Mali, and note that SE1078 is similar.
- 10_CD.TZ.BFL061 AF289548 TANZANIA Koulinska, I *ARHR* **17**(5):423–431 (2001)
This sequence was isolated from a perinatally infected infant from Dar es Salaam, Tanzania in 1996, and was part of a set of full length HIV-1 genomes characterizing CRF10_CD.
- 10_CD.TZ.BFL071 AF289549 TANZANIA Koulinska, I *ARHR* **17**(5):423–431 (2001)
This sequence was isolated from a perinatally infected infant from Dar es Salaam, Tanzania in 1996, and was part of a set of full length HIV-1 genomes characterizing CRF10_CD.
- 10_CD.TZ.BFL110 AF289550 TANZANIA Koulinska, I *ARHR* **17**(5):423–431 (2001)
This sequence was isolated from a perinatally infected infant from Dar es Salaam, Tanzania in 1996, and was part of a set of full length HIV-1 genomes characterizing CRF10_CD.
- 11_cpx.CM.MP818 AJ291718 CAMEROON Peeters, M Unpublished (2000)
- 11_cpx.FR.MP1298 AJ291719 FRANCE Peeters, M Unpublished (2000)
- 11_cpx.FR.MP1307 AJ291720 FRANCE Peeters, M Unpublished (2000)
- 11_cpx.GR.GR17 AF179368 GREECE Paraskevis, D *ARHR* **16**(9):845–855 (2000)
This subject that this sequence was isolated from was born in 1971 in Zaire, probably infected heterosexually in the DRC in 1991, and moved to in Greece in 1992. The genome was thought of to be a mosaic of A/G/E/J.
- CPZ.CD.SIVCPZANT U42720 CONGO Vanden-Haesevelde, MM *Virology* **221**(2): 346–350 (1996)
CPZANT is a simian immunodeficiency virus, phylogenetically linked to HIV-1. It was isolated from a captured wild chimpanzee from Zaire. This is the third SIV strain linked to HIV-1, after SIVCPZ-GAB (X52154) and SIVCPZ-GAB2 (U11495) were isolated from chimps in Gabon. Another chimpanzee virus was sequenced in 1998, and published in 1999 with accession number AF103818 (CPZ-US). The chimpanzee viral sequences are genetically more closely related to the HIV-1 sequences derived from infected humans than are HIV-2 strains or other SIVs. SIVCPZ-ANT is considered to be an outgroup of HIV-1 and is used to suggest the possibility of various introductions of HIV-1 into the human population.
- CPZ.CM.SIV-CPZCAM3 AF115393 CAMEROON Corbet, S *J Virol* **74**:529–534 (2000)
This complete genome is from a wild-caught infant chimpanzee of the *Pan troglodytes troglodytes* subspecies. Cam3 was captured at age approximately one year, in 1992 near a village on the southern border of Cameroon with Gabon. Cam3 died suddenly in 1998 with subacute pneumonia. Cam3 was housed for a time with Cam4, who apparently became infected by Cam3 while in captivity. Cam4 remains healthy in 1999, apart from episodes of diarrhea and fever. Cam4 is a *Pan troglodytes vellerosus* subspecies, and the infection of Cam4 by Cam3 in captivity is believed to have occurred through fighting. Cam3 and Cam4 were housed together from time of capture in 1992 through death of Cam3 in 1998.
- CPZ.CM.CAM5-SIVCPZ AJ271369 CAMEROON Muller-Trutwin, M *J Med Primatol* **29**(3–4):166–172(2000)
This env gene fragment is from a wild-caught infant chimpanzee of the *Pan troglodytes troglodytes* subspecies. Cam5 was captured at age less than one year, in 1998 in the central province of Cameroon, and was housed for just a short while at the Yaounde Zoo because he died with diarrhea soon after arrival. He is believed to have been infected in the wild. His mother was killed by hunters. An env gene sequence from the same isolate is available with accession number AF135498. CAM5 grows on human PBMC and uses CCR5 as a co-receptor.

CPZ.GA.SIVCPZGAB X52154 GABON Huet, T *Nature* **345**(6273):356–359 (1990)

The CPZ genome is more closely related to HIV-1s than to any other HIV or SIV viral sequences, but it is more divergent from prototypical HIV-1 than any other isolate, with possible exception of the partially characterized ANT70. CPZ is especially different with respect to the vpU gene product. Also see CPZGAB2, U11495 for a sequence fragment from an additional chimpanzee caught in Gabon

CPZ.US.SIVCPZUS AF103818 U.S. Gao, F *Nature* **397**(6718):436–441 (1999)

This full length molecular clone of a simian immunodeficiency virus, which infected an African wild-caught chimpanzee (Marilyn) who was the only chimpanzee identified as virus infected during a serosurvey of 98 chimpanzees in 1985. Marilyn had never been used in AIDS research and had not received human blood products after 1969. She died in captivity in 1985 after giving birth to stillborn twins. The complete genome was sequenced from 4 overlapped PCR fragments, amplified in 1998 from spleen tissues frozen at autopsy in 1985. Recovery of infectious virus from the frozen tissue was attempted but unsuccessful.

|→ 5' LTR, U3 start
 ...TGGAGGGCTAATTCACTCCCAACGAGACAAAGATATCCTTTGATCTGTGGATCTACACACACACAGGCTACTTCCTCGATTAGCAGAACTACACACCAGGG...CCAGGGATC...AGATATCCACT

B.FR.HXB2R 121
 CONSENSUS A 0
 A.UG.U455 0
 A.SE.UGSE8891 0
 A.SE.TZSE8538 0
 A.SE.UGSE6594 0
 A.KE.Q2317 0
 A.SE.UGSE7535 0
 A.SE.SOSE7253 0
 A.UG.92UG037 0
 A.SE.SE8131 0
 A2.CD.97CDKTB48 0
 A2.CY.94CY017 0
 CONSENSUS B 97
 B.AU.MBC925 121
 B.AU.MBC200 118
 B.GA.OYI 0
 B.GB.MANC 66
 B.CN.RL42 122
 B.DE.D31 0
 B.DE.HAN2 0
 B.ES.89SEP061 0
 B.NL.ACH320A 121
 B.TW.LM49 121
 B.US.DHL23 121
 B.US.SF2 122
 B.US.NY5CG 121
 B.US.WCIPR18 0
 B.US.YU2 0
 B.US.JRC5F 118
 B.US.JRFL 122
 B.US.MN 0
 B.US.BCSG3 121
 B.US.WEAU160 0
 B.US.RF 0
 B.US.WR27 0
 CONSENSUS C 0
 C.BR.92BR025 0
 C.BW.96BW0402 0
 C.BW.96BW1104 0
 C.BW.96BW15C02 0
 C.BW.96BW01B03 0
 C.BW.96BW0502 0
 C.BW.96BW16B01 0
 C.BW.96BW1210 0
 C.ET.ETH2220 0
 C.IN.301999 0
 C.IN.21068 0
 C.IN.301905 0
 C.IN.301904 0
 C.IN.93IN101 0
 C.IN.11246 0
 D.UG.94UG114 121
 D.CD.NDK 0
 D.CD.ELI 0
 D.KE.MB2059 0
 D.CD.2226 0
 D.CD.84ZR085 122
 F1.BR.93BR020 121
 F1.FI.FIN9363 0
 F1.BE.VI850 0
 F1.FR.MF411 0
 F2.CM.MF255C 0
 F2.CM.MF257C 0

|-> 5' LTR U3 start
 ...TGGARGGGCTAATTCACCCCAACGAGACAGATATATCTTGTGATCTGTGGTCTACACACACAGGCTACTTCCTGATTAGCAGAACTACACACAGGG...CCAGGGATC...AGATATCCACT
 A-----C-C--A--G-----TA-T--CA---A-T-----A-----G-----CT-----TC-----TC-----TC-----

B.FR.HXB2R 121
 G.BE.DRCBL 83
 G.NG.92NG083 0
 G.SE.SE6165 0
 G.FI.HH87932 0
 H.BE.VI991 0
 H.BE.VI997 0
 H.CF.90CF056 0
 J.SE.SE91733 0
 J.SE.SE92809 0
 K.CM.MP535C 0
 K.CD.FOR11C 0
 O.CM.ANT70 0
 O.CM.MP5180 122
 O.SN.MP1299 122
 O.SN.MP1300 122
 N.CM.YBF30 0
 N.CM.YBF106 0
 CONSENSUS 01 64
 01.AE.TH.93TH253 121
 01.AE.TH.CM240 121
 01.AE.CF.90CF402 121
 01.AE.TH.92TH022 73
 01.AE.TH.95THH047 73
 01.AE.CF.11697 58
 01.AE.TH.93TH057 64
 01.AE.TH.93TH065 64
 01.AE.CF.4071 46
 CONSENSUS 02 0
 02.AG.SE.SE7812 0
 02.AG.GH.G829 0
 02.AG.DJ.DJ263 0
 02.AG.DJ.DJ264 0
 02.AG.NG.IBNG 0
 02.AG.SN.MP1211 0
 02.AG.SN.MP1213 0
 02.AG.CM.MP807 0
 02.AG.RU.RU98001 0
 03.AB.RU.KAL153-2 106
 04.CPX.GR.97PVMY 121
 04.CPX.GR.97PVCH 122
 04.CPX.CY.94CY032-3 0
 05.DF.BE.VI1310 0
 06.CPX.ML.95ML84 0
 06.CPX.AU.BFP90 121
 06.CPX.WI.95WI127 121
 06.CPX.SN.97SEL078 121
 10.CD.BEL061 0
 10.CD.BEL110 0
 10.CD.BEL1071 0
 11.CPX.CM.MP818 121
 11.CPX.FR.MP1298 121
 11.CPX.FR.MP1307 121
 11.CPX.GR.GR17 0
 CPZ.GA.SIVCPZGAB 121
 CPZ.US.SIVCPZUS 121
 CPZ.CM.SIVCPZCAM3 0
 CPZ.CM.SIVCPZCAM5 0
 CPZ.CD.SIVCPZANT 0

B. FR. HXB2R	GACCTTTGGATGGTCTACAGCTAGTACCAGTTGACCCAGAACTTAGAAGAA.....GCCAACAAAGGAGAGAAACACACGCTTGTACCCCTGTGAGCCCTGCATGGAATGGATGACCCGGAG	242
G. BE. DRCEI	-----T---A-G---T---TCAG-G---G-----T-----A-----CA-CT---A-----G-----GAA--C.....	204
G. NG. 92MG083		0
G. SE. SE6165		0
G. FI. HH87932		0
H. BE. VI991		0
H. BE. VI997		0
H. CF. 90CF056		0
J. SE. SE91733		0
J. SE. SE92809		0
K. CM. MP535C		0
K. CD. EOTB11C		0
O. CM. ANI70	-----TG-TT--A-----GTCAGA--AG-GC---AG-CTAGGA...G-T-C-T-T---GGG-T-ATC-CC-G--T--A-CAT-TGCC-----T-T--A--TA-AC-T...	246
O. CM. MP5180	-----TG-TT--A-G-----GTCAG--AG-GC---GAG-CTGGGT...-T-C-AT-AG-TG-T--TC-TC-T-A-CTT-TAAT-----GCT--G--TG-AC-C...	246
O. SN. MP1299	-----G---TA-TT--T-----GTCAGA-CTG--GC---G---CTAGGA...-GT-T---GGG-T--C-CC-G--T--A-CTT--AAC-CT-T-A--AAAC-C...	246
O. SN. MP1300	-----G---CTA-TT-----GTCAGA-CTG--GC---CTAGGA...-T-GT-T--C-GGG-T-AMC-CC-G--T--A--TT--AAC-CT-T-A--T-AC-C...	246
N. CM. YBF30		0
N. CM. YBF106		0
CONSENSUS 01	gtgttttgatggtcttcaagctagtaccagttgaccccaagagaagtagagag.....gacaacaagagagaaacaactgctgtttacaccccatgagccagcatggaatagaggacgaaga...	185
01_AE.TH.93TH253	-TGT-----T---A-----C---AGAG-AG--G-G.....-A---A-T-C-----CA-----A---A-G---GAA--A...	242
01_AE.TH.CM240	-TGT-----T---AGAG-AG--G-G.....-A---A-T-C-----CA-----A---A-G---GAA--A...	242
01_AE.CF.90CF402	ATGT-----T---C---AGAG-AG--G-G.....-A---A-T-C-----CA-----A---A-G---GAA--A...	194
01_AE.TH.92TH022	-TGT-----T---G---C---AGAG-AG--G-G.....-A---A-T-C-----CA-----A---A-G---GAA--A...	194
01_AE.CF.11697	-TGT-----T---A---A---AG-AG--G-G.....-A---A-T-C-----CA-----AA-----G-A---TGAA--A...	179
01_AE.TH.93TH057	-TGT-----T---T-----C---AGAG-AG--G-G.....-A---A-T-C-----CC-----A---G---GAA--A...	185
01_AE.TH.93TH065	-TGT-----T---C---AGAG-AG--G-G.....-A---A-T-C-----CA-----A---A-G---GAA--A...	185
01_AE.CF.4071	-TGT-----T---AGAG-AG--G-G.....-A---A-T-TC-----CA-----A---A-G---TGAA--A...	167
CONSENSUS 02		0
02_AG.SE.7812		0
02_AG.GH.G829		0
02_AG.DJ.DJ263		0
02_AG.DJ.DJ264		0
02_AG.NG.IBNG		0
02_AG.SN.MP1211		0
02_AG.SN.MP1213		0
02_AG.CM.M807		0
03_AB.RU.98001		0
03_AB.RU.KALI53-2		0
04_CFX.GR.97FVNY	A--A-----G-----C-C---CTG-AG--G-----G-----CAG-G-----A-C-A-----A-AT---AA--C-----TGA--A...	227
04_CFX.GR.97FVCH	-TG-----T-----A-T---C-G-G-G-----G-----AT-CTGC-----A-T-T--C-G-----A-A---A-----G-TGAA-----	242
04_CFX.CY.94CY0323	-TG-----T-----A-T---C-G-G-G-----G-----TG-----A-T-T--C-G-----A-A---A-----G-TGAA-----	243
05_DF.BE.VI1310		0
05_DF.BE.VI961		0
06_CFX.ML.95ML84		0
06_CFX.AU.BPP90	-T-----T-----T-----G-----CUTT-CT-----A-T-C-----CA-AT--A-----GCA--G--TGAA--A...	242
06_CFX.ML.95ML127	-T-----T-----C---T---AG-G-AG--G-----G-----AA-CT-----A---C---CA-T--A-----GC--G--TGAA--A...	242
06_CFX.SN.97SE1078	-T-----T-----T-G-----C-T---A-G-AG--G-----G-----AG-CT-----CA-T--A-----G-----T-A--A...	242
10_CD.BEL061		0
10_CD.BEL110		0
10_CD.BEL071		0
11_CFX.CM.M818	-TG-----T-----C---AGAG-AG--G-----G-----TG-----T-A-T--C-----CA-----A-----TGAA--A...	242
11_CFX.FR.MP1298	-TG-----T-----T---AG-G-AG--G-----G-----TG-----A--TC-C-----T--CA-AT---AA-----TGAA--A...	242
11_CFX.FR.MP1307	-TG-----T-----T---C---A-AG-AG--G-----G-----TG-----A-T-C-C-----T--CA-A--A-----TGAA--A...	242
11_CFX.GR.GR17		0
CPZ.GA.SIVCPZGAB	CTGT-----T---A-----G-CC-GACAGAG--C-G-----C-----TG-----T--A-T-C-----G-T--CA-AT-T-A-----G--A--TGA--C...	242
CPZ.US.SIVCPZUS	A--A-A---GC-----TT---T---C---CACAGA--G-AG---GCGG.....-T-----T-----C-A-A--TAC--C-G-----CA-T--A-----A--TGAAC-T...	242
CPZ.CM.SIVCPZCAM3		0
CPZ.CM.SIVCPZCAM5		0
CPZ.CD.SIVCPZANT		0

HIV-1/SIVcpz complete genomes

B.FR.HXB2R
AGAGAAGTGTAGAGGTTTGACAGCCGCTAGCATTTTCATCACATGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGA.....TCF-1alpha- |-----CATC
CONSENSUS A 336
A.UG.U455 0
A.SE.UGSE8891 0
A.SE.TZSE8538 0
A.SE.UGSE6594 0
A.KE.Q2317 0
A.SE.UGSE7535 0
A.SE.SOSE7253 0
A.UG.92UG037 0
A.SE.SE8131 0
A2.CD.97CDKTB48 0
A2.CY.94CY017 0
CONSENSUS B 291
B.AU.MBC925 336
B.AU.MBC200 333
B.GA.OYI 0
B.GB.MANC 281
B.GB.CAMI 337
B.CN.RL42 0
B.DE.D31 0
B.DE.HAN2 0
B.ES.89SEP061 0
B.NL.ACH320A 336
B.TW.LM49 336
B.US.DHL23 336
B.US.SF2 337
B.US.NY5CG 336
B.US.WCIPR18 0
B.US.YU2 333
B.US.JRCSF 333
B.US.JRFL 0
B.US.MN 335
B.US.BCSG3 0
B.US.WEAU160 336
B.US.RF 0
B.US.WR27 0
CONSENSUS C 0
C.BR.92BR025 0
C.BW.96BW0402 0
C.BW.96BW1104 0
C.BW.96BW15C02 0
C.BW.96BW01B03 0
C.BW.96BW0502 0
C.BW.96BW16B01 0
C.BW.96BW1210 0
C.ET.ETH2220 0
C.IN.301999 0
C.IN.21068 0
C.IN.301905 0
C.IN.301904 0
C.IN.93IN101 0
C.IN.11246 0
D.UG.94UG114 0
D.CD.NDK 0
D.CD.ELI 0
D.KE.MB2059 0
D.CD.Z226 0
D.CD.84ZR085 0
F1.BR.93BR020 0
F1.FI.FIN9363 0
F1.BE.VI850 0
F1.FR.MF411 0
F2.CM.MF255C 0
F2.CM.MF257C 0

B. FR. HXB2R	AGAGAAGTTTAGAGTGGAGGTTTGCACAGCCCGCTAGCAATTCATCACATGGCCCGAGAGCTGCATCCCGAGTCTCAAGAACTGCTGAC-TCF-1alpha-.....CATC	336
G. BE. DRCLB1	-----C-G-T-----AA-----TA-----CGGAGA-----A-----AG-----	-----A-----AG-----	298
G. NG. 92NG083			0
G. SE. SE6165			0
G. FI. HH87932			0
H. BE. VI991			0
H. BE. VI997			0
H. CF. 90CF056			0
J. SE. SE91733			0
J. SE. SE92809			0
K. CM. MP535C			0
K. CD. FOR11C			0
O. CM. ANT70			0
O. CM. MPF5180	AACTGCTGA-C-G	350
O. SN. MP1299	CCC-G-AACTGA.....	339
O. SN. MP1300	CT-----G-AACTGA.....	349
N. CM. YBF30	TCTC-GGA-A-AACTGTTGA.....	349
N. CM. YBF106	G-----A-----	0
CONSENSUS 01	agagaagtgtgatgtggaagtgttgcagagttccctagcacgaa?acacatagccccgagaaactgcacccagagtaactataaagactgctga.....T-AT-AG-----caaaa	278
01_AE.TH.93TH253	-----C-GATA-----A-----TGC-----A-----G-----A-----T-----AT-----AG-----	-----A-----AG-----	336
01_AE.TH.CM240	-----C-AT-----A-----TTC-----CGAAGA-----A-----T-----C-----A-----AG-----A-----	-----A-----AG-----	359
01_AE.TH.90CF402	-----C-GATA-----A-----TTC-----CGAA-A-C-A-----C-A-----AT-----AG-----A-----	-----A-----AG-----	288
01_AE.TH.92TH022	-----C-GAT-----A-----TGC-----CGAA-A-----A-----C-----A-----AT-----AG-----A-----	-----A-----AG-----	288
01_AE.TH.95THH047	-----C-GAT-----A-----TTC-----CGAAGA-----A-----G-----A-----AG-----A-----	-----A-----AG-----	273
01_AE.CF.11697	-----C-GAT-----A-----TTC-----CGAAGA-----A-----CA-----A-----AT-----AG-----A-----	-----A-----AG-----	279
01_AE.TH.93TH057	-----C-GAT-----A-----TTC-----CGAA-AT-----A-----T-----AT-----AG-----A-----	-----A-----AG-----	279
01_AE.TH.93TH065	-----C-GAT-----A-----TTC-----CGAAGA-----A-----A-----A-----AG-----A-----	-----A-----AG-----	261
01_AE.CF.4071	-----C-GAT-----A-----TTC-----CGAAGA-----A-----A-----A-----AG-----A-----	-----A-----AG-----	0
CONSENSUS 02			0
02_AG.SE.7812			0
02_AG.GH.G829			0
02_AG.DJ.DJ263			0
02_AG.DJ.DJ264			0
02_AG.NG.IBNG			0
02_AG.SN.MP1211			0
02_AG.SN.MP1213			0
02_AG.CM.M5807			0
03_AB.RU.98001			0
03_AB.RU.KAL153-2	T-A-AG-----CACAGACTGCTGACAT	333
04_CPX.GR.97PVMY	-----C-A-----A-----T-----TG-----ACAGA-----T-----A-----AG-----	-----T-----A-----AG-----	336
04_CPX.GR.97PVCH	-----A-----A-----A-----T-----T-----CA-A-----A-----AG-----	-----T-----A-----AG-----	337
05_DF.BE.VI1310			0
06_CPX.ML.95ML84	AAA-----TTT-A-AG-----CAAGGCTGCTGA-AA	348
06_CPX.AU.BFP90	TC-----CGAAGA-----A-----C-----TTT-A-AG-----AA	336
06_CPX.MI.95ML127	TC-----CGAAGA-----A-----TTT-AT-AG-----AA	357
06_CPX.SN.97SEL078	TC-----G-----CGGAGA-----A-----A-----CAGCCTCACAAGACTGCTGA-AA	0
10_CD.BEL061			0
10_CD.BEL110			0
10_CD.BEL1071			0
11_CPX.CM.MP818	-----A-----GT-----TC-----CGAA-A-----A-----C-T-A-AG-AA.....GACTGCTGA--CA	342
11_CPX.FR.MP1298	-----AT-----A-----TC-----CGAA-A-----A-----C-T-A-AG-----	-----A-----AG-----	336
11_CPX.FR.MP1307	-----AT-----A-----TTC-----AGAAGA-----A-----GG-AT-G-----A.....GAACTGCTGA--CA	343
11_CPX.GR.GR17			0
CPZ.GA.SIVCPZGAB	A-G-G-C-AAGA-T-T-A-----A-AA-----ACTTTAG.....ACTGGCGCATGGG...--CA
CPZ.US.SIVCPZUS	-----G-C-----C-----ATC-----CA-----CTGAA-G-TCGGAGA-----GA-TAA-----A.....TGGCTGTAACCGCGCAGCGC--AT	352
CPZ.CM.SIVCPZCAM3			353
CPZ.CM.SIVCPZCAM5		GGGCGCTGGC.AG-A	14
CPZ.CD.SIVCPZANT			0

B.FR.HXB2R	NF-kappa-B-II	NF-kappa-B-I	Sp1-III	Sp1-II	Sp1-I	417
CONSENSUS A	GGACTTTCGCC	TGGGACTTTCAG	GGAGCGGTGCCTG	GGCGGGA	CTGGGAGTGGCGA	0
A.UG.0455	GGACTTTCGCC	TGGGACTTTCAG	GGAGCGGTGCCTG	GGCGGGA	CTGGGAGTGGCGA	0
A.SE.UGSE8891	GGACTTTCGCC	TGGGACTTTCAG	GGAGCGGTGCCTG	GGCGGGA	CTGGGAGTGGCGA	0
A.SE.TZSE8538	GGACTTTCGCC	TGGGACTTTCAG	GGAGCGGTGCCTG	GGCGGGA	CTGGGAGTGGCGA	0
A.SE.UGSE6594	GGACTTTCGCC	TGGGACTTTCAG	GGAGCGGTGCCTG	GGCGGGA	CTGGGAGTGGCGA	0
A.KE.Q2317	GGACTTTCGCC	TGGGACTTTCAG	GGAGCGGTGCCTG	GGCGGGA	CTGGGAGTGGCGA	0
A.SE.UGSE7535	GGACTTTCGCC	TGGGACTTTCAG	GGAGCGGTGCCTG	GGCGGGA	CTGGGAGTGGCGA	0
A.SE.SOSE7253	GGACTTTCGCC	TGGGACTTTCAG	GGAGCGGTGCCTG	GGCGGGA	CTGGGAGTGGCGA	0
A.UG.92UG037	GGACTTTCGCC	TGGGACTTTCAG	GGAGCGGTGCCTG	GGCGGGA	CTGGGAGTGGCGA	0
A.SE.SE8131	GGACTTTCGCC	TGGGACTTTCAG	GGAGCGGTGCCTG	GGCGGGA	CTGGGAGTGGCGA	0
A2.CD.97CDKTB48	GGACTTTCGCC	TGGGACTTTCAG	GGAGCGGTGCCTG	GGCGGGA	CTGGGAGTGGCGA	0
A2.CD.94CY017	GGACTTTCGCC	TGGGACTTTCAG	GGAGCGGTGCCTG	GGCGGGA	CTGGGAGTGGCGA	0
CONSENSUS B	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	370
B.AU.MBC925	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	418
B.AU.MBC200	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	415
B.GA.OYI	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	0
B.GB.MANC	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	363
B.GB.CAMI	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	419
B.CN.RL42	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	0
B.DE.D31	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	0
B.DE.HAN2	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	0
B.ES.89SP061	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	0
B.NL.ACH320A	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	418
B.IW.LM49	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	416
B.US.DH123	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	417
B.US.SF2	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	417
B.US.NY5CG	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	0
B.US.WCIPR18	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	0
B.US.YU2	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	414
B.US.JRGSF	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	418
B.US.JRFL	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	0
B.US.MN	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	416
B.US.RCSG3	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	0
B.US.WEAU160	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	417
B.US.RF	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	0
B.US.WR27	gggactttccgc	tggggaattccag	gtgagcgtg	ggcggga	c?ggggagtgccga	0
CONSENSUS C	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.BR.92BR025	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.BW.96BW0402	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.BW.96BW1104	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.BW.96BW15C02	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.BW.96BW01B03	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.BW.96BW0502	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.BW.96BW16B01	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.BW.96BW1210	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.ET.ETH2220	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.IN.201999	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.IN.210688	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.IN.301905	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.IN.301904	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.IN.93IN101	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
C.IN.11246	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
D.UG.94UG114	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	417
D.CD.NDK	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
D.CD.ELI	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
D.KE.MB2059	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
D.CD.2226	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
D.CD.84ZE085	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	417
FI.BR.93BR020	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	417
FI.FI.FIN9363	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
FI.BE.VI850	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
FI.FR.MF411	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
F2.CM.MF255C	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0
F2.CM.MF257C	gagcttctacaa	tgggactttccag	ggagcgtg	ggcggga	ctgggagtgccga	0

B. FR. HXB2R	GAAGTTGCTACAA	GGGACTTTCGC	NF-kappa-B-II	CT	TGGGACTTTCAG	GGAGCGCTGGCTG	Sp1-III	CTGGGAGTGGCGA	Sp1-I	CTGGGAGTGGCGA	GCCTCA	417
G. BE. DRCEBL	--AG--GAC-AG											382
G. NG. 92NG083												0
G. SE. SE6165												0
G. FI. HH87932												0
H. BE. VI1991												0
H. BE. VI1997												0
H. CF. 90CF056												0
J. SE. SE91733												0
J. SE. SE92809												0
K. CM. MP535C												0
K. CD. EOPB11C												0
O. CM. AN770												0
O. CM. MPF5180	A-A	---A---	AGCAAGACTGCTGACACTGC									452
O. SN. MP1299	A-A	---GAC-CTGC	AGACTGCTGACACTG									425
O. SN. MP1300	A-A	---GAC-CTGC	AGCAGACTGCTGACACGGC									452
N. CM. YBF30												0
N. CM. YBF106												0
CONSENSUS 01	gaagtttctaact	aggactt.cccg	tggggactttccagg	gaggtgtggccgg	ggcggag	ttggggagtggtgcta	accctca					359
01 AE. TH. 93TH253	--AG-T--ACT	A	---G- GA-GTGT-G-C-G	---AG	---T	---TT	---A					417
01 AE. TH. CM240												0
01 AE. CF. 90CF402	-GAG-T--ACT	A	---G- GA-GTGT-G-CTG	---G	---T	---T	---A					440
01 AE. TH. 92TH022	-AG-T--ACT	A	---G- GA-GTGT-G-C-G	---AG	---T	---T	---A					369
01 AE. TH. 95TH1047	-AG-T--A-C	A	---G- GA-GTGT-G-C-G	---AA-A	---T	---T	---A					369
01 AE. CF. 11697	-AG-T--ACC	A	---G- GA-GTGT-G-CTG	---AG	---T	---T	---A					354
01 AE. TH. 93TH057	-AG-T--ACT	A	---G- GA-GTGT-G-C-G	---AG	---T	---T	---A					360
01 AE. TH. 93TH065	-AG-T--ACT	A	---G- GA-GTGT-G-C-G	---AG	---T	---T	---A					360
01 AE. CF. 4071	-AG-TA--ACT	A	---G- A-GTGT-G-CTG	---AG	---T	---T	---A					342
CONSENSUS 02												0
02 AG. SE. 7812												0
02 AG. GH. G829												0
02 AG. DJ. DJ263												0
02 AG. DJ. DJ264												0
02 AG. NG. I8NG												0
02 AG. SN. MP1211												0
02 AG. SN. MP1213												0
02 AG. CM. ME607												0
03 AB. RU. 98001												0
03 AB. RU. KALI53-2	--AG--GAC	AA	---G- T--TT	---AG	---T	---T	---A					414
04 CPX. GR. 97PVMY	--AG--GAC	AA	---G- T--C	---A-AG	---T	---T	---A					420
04 CPX. GR. 97PVGH	--AG--GAC	AA	---G- T--C	---A-G	---T	---T	---A					421
04 CPX. CV. 94CY0323												0
05 DF. BE. VI1310												0
05 DF. BE. VI1961												0
06 CPX. ML. 95ML84												0
06 CPX. AU. BFP90	--AG-T--ATG											429
06 CPX. MI. 95MI127	--AG-T--AGG											417
06 CPX. SN. 97SEL078	--AG-T--GAC											439
10 CD. BFL061												0
10 CD. BFL110												0
10 CD. BFL071												0
11 CPX. CM. MP818	--AG--GAC	AC	---G- T--T	---G	---T	---T	---A					424
11 CPX. FR. MP1298	--AG--A-GAC											418
11 CPX. FR. MP1307	--AG--GAC											425
11 CPX. GR. GR17												0
CPZ. GA. SIVCPZGAB	AGAAC	GA.CTCTGC	---G-G-AGA-GTG-T-GG	---A	---G	---T	---T					438
CPZ. US. SIVCPZUS	A-AAC	GA.CTG	---G-AA-G-T-G-T-A	---G	---A	---C	---TTT					444
CPZ. CM. SIVCPZCM3												0
CPZ. CM. SIVCPZCM5	--AAC	---GA	---A--G-A-GCT-A	---G	---ACA	---C	---TTT					93
CPZ. CD. SIVCPZANT												0

HIV-1/SIVcpz complete genomes

	TATA Box	5' ITR U3 end	5' ITR R repeat begin	5' ITR U5 end	5' ITR R repeat begin	Poly-A signal	
B.FR.HXB2R	GA.TCCTGCATATAAGCAGCTGGCTTTTGGCTGTACTGGGTCTCTCTGGTTAGACCAAGACTGAGCCTGGGAGCTCTCTGG...	CTAACTAGGGAACCCACTGCTT...	AAGCCTCAATA...	AAGCTTTG	???????	-----	537
CONSENSUS A							0
A.UG.U455							7
A.SE.UGSE8891							0
A.SE.TZSE8538							0
A.SE.UGSE6594							0
A.KE.Q2317							7
A.SE.UGSE7535							0
A.SE.SOSE7253							0
A.UG.92UG037							0
A.SE.SE8131							0
A2.CD.97CDKTB48							0
A2.CD.94CY017							0
CONSENSUS B	ga.tgctgcataataagcagctgttttgccctgactgggtctctctgtgttagaccagatctgagcctgggagctctctctgg...	ctagtaggggaacccactgctt...	aagcctcaata...	aagcttbg	???????		489
B.AU.MBC925							538
B.AU.MBC200							534
B.GA.OYI							83
B.GB.MANC							483
B.GB.CAMI							539
B.CN.RL42							0
B.DE.D31							83
B.DE.HAN2							7
B.ES.89SP061							84
B.NL.ACH320A							538
B.IW.LM49							536
B.US.DHI23							535
B.US.SF2							537
B.US.NY5CG							537
B.US.WCIPR18							83
B.US.YU2							533
B.US.JRGSF							538
B.US.JRFL							0
B.US.MN							536
B.US.RCSG3							84
B.US.WEAU160							537
B.US.RF							51
B.US.WR27							0
CONSENSUS C	?? ???...	??...	??...	??...	??...	????????	0
C.BR.92BR025							0
C.BW.96BW0402							26
C.BW.96BW1104							0
C.BW.96BW15C02							27
C.BW.96BW01B03							26
C.BW.96BW0502							26
C.BW.96BW16B01							26
C.BW.96BW1210							0
C.ET.ETH2220							0
C.IN.301999							0
C.IN.210688							0
C.IN.301905							0
C.IN.301904							0
C.IN.93IN101							0
C.IN.11246							0
D.UG.94UG114							537
D.CD.NDK							0
D.CD.ELI							0
D.KE.MB2059							81
D.CD.Z226							83
D.CD.84ZE085							537
F1.BR.93BR020							537
F1.FI.FIN9363							51
F1.BE.VI850							0
F1.FR.MF411							0
F2.CM.MF255C							0
F2.CM.MF257C							0

Table with columns for TATA Box, 5' ITR U3 end, ITR R repeat, and Poly-A signal. It lists 100 HIV-1/SIVcpz complete genome sequences with their corresponding nucleotide positions (572-537) and Poly-A signal positions (549-537). The sequences are grouped by accession number and include strain identifiers like B.FR.HXB2R and various SIVcpz strains (01-02, 03-05, 06-10).

HIV-1/SIVcpz complete genomes

Accession	Genomic Region	Sequence	Position
B.FR.HXB2R	5' LTR R	repeat end \ / U5 start	
CONSENSUS A	Extensive secondary structure in this region. See Rizavi J Virol 67:2681-8 (1993)	Lys tRNA primer binding site	
A.UG.U455	5' LTR	5' ITR U5 end <	650
A.SE.UGSE8891			0
A.SE.TZSE8538			0
A.SE.UGSF6594			0
A.KE.Q2317			119
A.SE.UGSE7535			0
A.SE.SOSE7253			0
A.UG.92UG037			0
A.SE.SE8131			28
A2.CD.97CDKTB48			0
A2.CI.94CIY017			0
CONSENSUS B			502
B.AU.MBC925			651
B.AU.MBC200			647
B.GA.OYI			196
B.GB.MANC			596
B.GB.CAMI			652
B.CN.RL42			27
B.DE.D31			196
B.DE.HAN2			120
B.ES.89SP061			199
B.NL.ACH320A			651
B.IW.LMH9			649
B.US.DHL23			648
B.US.SF2			651
B.US.NYSGG			650
B.US.WCIPR18			196
B.US.YU2			646
B.US.JRGSF			651
B.US.JRFL			0
B.US.MN			649
B.US.RCSG3			197
B.US.WEAU160			650
B.US.RF			164
B.US.WR27			30
CONSENSUS C			35
C.BR.92BR025			0
C.BW.96BW0402			139
C.BW.96BW1104			39
C.BW.96BW15C02			140
C.BW.96BW01B03			138
C.BW.96BW0502			137
C.BW.96BW16B01			138
C.BW.96BW1210			39
C.ET.FTH2220			28
C.IN.301999			28
C.IN.210688			28
C.IN.301905			28
C.IN.301904			28
C.IN.93IN101			649
C.IN.11246			28
D.UG.94UG114			28
D.CD.NDK			0
D.CD.ELI			194
D.KE.MB2059			196
D.CD.Z226			650
F1.BR.93BR020			650
F1.FI.FIN9363			164
F1.BE.VI850			0
F1.FR.MF411			1
F2.CM.MF255C			0
F2.CM.MF257C			0

5' LTR R	5' LTR	Extensive secondary structure in this region.	Lys tRNA primer binding site	5' ITR U5 end <-	
repeat end \ / U5 start	See R1avi J Virol 67:2681-8 (1993)				
B.FR.HXB2R	CCTTGAGTG...CTTCAGTAGTGTGCCCCGTC	...TTGTGTGACTCTGGTAACTAGAGATCCC...TCAGAC...CCCTTTAGTCAGTGTGG...AANAATCTTA...GCAGTGGCCCGCCGAAACAGG	...AC-C--AT-	AT...-A-AGC-AAAGTT--T--	650
G.BE.DRCBL	-GCT...T		615
G.NG.92NG083	-GCT...T		21
G.SE.SE6165			33
G.FI.HH87932			33
H.BE.VI191			28
H.BE.VI197			0
H.CF.90CF056			0
J.SE.SE91733			0
J.SE.SE92809			0
K.CM.MP535C			0
K.CD.EQPE11C			0
O.CM.AMT70			680
O.CM.MPF5180			653
O.SN.MP1299			680
O.SN.MP1300			680
N.CM.YBF30			208
N.CM.YBF106			208
N.CM.YBF106			208
CONSENSUS 01			649
01.AE.TH.93TH253			196
01.AE.TH.93TH253			672
01.AE.TH.93TH253			672
01.AE.TH.93TH253			601
01.AE.TH.93TH253			601
01.AE.TH.93TH057			586
01.AE.TH.93TH065			591
01.AE.TH.93TH065			591
01.AE.CF.4071			573
CONSENSUS 02			0
02.AG.SE.7812			25
02.AG.GH.G829			0
02.AG.DJ.DJ263			18
02.AG.DJ.DJ264			19
02.AG.NG.IBNG			179
02.AG.SN.MP1211			0
02.AG.SN.MP1213			0
02.AG.CM.ME807			0
03.AB.RU.98001			28
03.AB.RU.KALI53-2			543
04.CPX.GR.97PVMY			651
04.CPX.GR.97PVGH			647
05.CPX.CY.94CYO323			0
05.DF.BE.VI1310			26
05.DF.BE.VI1961			1
06.CPX.ML.95ML84			0
06.CPX.AU.BFP90			661
06.CPX.WI.95ML127			648
06.CPX.SN.97SEL078			671
10.CD.BFL061			27
10.CD.BFL110			27
10.CD.BFL071			27
11.CPX.CM.MP818			656
11.CPX.FR.MP1298			650
11.CPX.FR.MP1307			650
11.CPX.GR.GR17			660
CPZ.GA.SIVCPZGAB			0
CPZ.US.SIVCPZUS			669
CPZ.CM.SIVCPZCAM3			673
CPZ.CM.SIVCPZCAM5			19
CPZ.CD.SIVCPZANT			324

HIV-1/SIVcpz complete genomes

Gag p7 nucleocapsid binds these loops, see DeGuzman, Science 279:384-8
Packaging signal secondary structure, see Harrison, J Virol 72:5886-96

B.FR.HXB2RACCTGAAGCG.AAAGGANA.CCAGAGG...AGCT...CTCTGACG...CA..GGACTCGGCTTGCTGA.ACGCCGACCGCAAGAGCCGAGGG.CGGCGA	740
CONSENSUS A	?.....TT.....TC.....G-T-A-A-..	0
A.UG.U455TC.....TT.....T.....G-T-A-..	205
A.SE.UGSE8891TT.....A...-T.....G-T-A-..	0
A.SE.TZSE8538TC.....TT.....T.....G-T-A-..	205
A.SE.UGSE6594TT.....A...-T.....G-T-A-..	0
A.KE.Q2317TT.....A...-T.....G-T-A-..	0
A.SE.UGSE7535TT.....A...-T.....G-T-A-..	0
A.SE.SOSE7253TT.....A...-T.....G-T-A-..	0
A.UG.92UG037TT.....A...-T.....G-T-A-..	112
A.SE.SE8131TT.....A...-T.....G-T-A-..	138
A2.CD.97CDKTB48TT.....A...-T.....G-T-A-..	86
A2.CY.94CY017TT.....A...-T.....G-T-A-..	108
CONSENSUS BTT.....A...-T.....G-T-A-..	691
B.AU.MBC925TT.....A...-T.....G-T-A-..	741
B.AU.MBC200TT.....A...-T.....G-T-A-..	737
B.GA.OYITT.....A...-T.....G-T-A-..	285
B.GB.MANCTT.....A...-T.....G-T-A-..	686
B.GB.CAM1TT.....A...-T.....G-T-A-..	742
B.CN.RL42TT.....A...-T.....G-T-A-..	116
B.DE.D31TT.....A...-T.....G-T-A-..	286
B.DE.HAW2TT.....A...-T.....G-T-A-..	210
B.ES.89SEP061TT.....A...-T.....G-T-A-..	290
B.NL.ACH320ATT.....A...-T.....G-T-A-..	741
B.TW.LM49TT.....A...-T.....G-T-A-..	739
B.US.DH123TT.....A...-T.....G-T-A-..	738
B.US.SF2TT.....A...-T.....G-T-A-..	741
B.US.NY5CGTT.....A...-T.....G-T-A-..	740
B.US.WC1PR18TT.....A...-T.....G-T-A-..	286
B.US.YU2TT.....A...-T.....G-T-A-..	736
B.US.JRC5FTT.....A...-T.....G-T-A-..	741
B.US.JRFLTT.....A...-T.....G-T-A-..	63
B.US.MNTT.....A...-T.....G-T-A-..	736
B.US.BCSG3TT.....A...-T.....G-T-A-..	287
B.US.WEAU160TT.....A...-T.....G-T-A-..	740
B.US.RFTT.....A...-T.....G-T-A-..	254
B.US.WR27TT.....A...-T.....G-T-A-..	120
CONSENSUS CTT.....A...-T.....G-T-A-..	124
C.BR.92BR025TT.....A...-T.....G-T-A-..	89
C.BW.96BW0402TT.....A...-T.....G-T-A-..	228
C.BW.96BW1104TT.....A...-T.....G-T-A-..	128
C.BW.96BW15C02TT.....A...-T.....G-T-A-..	229
C.BW.96BW01B03TT.....A...-T.....G-T-A-..	227
C.BW.96BW0502TT.....A...-T.....G-T-A-..	226
C.BW.96BW16B01TT.....A...-T.....G-T-A-..	228
C.BW.96BW1210TT.....A...-T.....G-T-A-..	118
C.ET.FTH2220TT.....A...-T.....G-T-A-..	117
C.IN.30I999TT.....A...-T.....G-T-A-..	117
C.IN.21068TT.....A...-T.....G-T-A-..	117
C.IN.30I905TT.....A...-T.....G-T-A-..	117
C.IN.30I904TT.....A...-T.....G-T-A-..	117
C.IN.93IN101TT.....A...-T.....G-T-A-..	739
C.IN.11246TT.....A...-T.....G-T-A-..	117
D.UG.94UG114TT.....A...-T.....G-T-A-..	87
D.CD.NDKTT.....A...-T.....G-T-A-..	284
D.CD.ELITT.....A...-T.....G-T-A-..	286
D.KE.MB2059TT.....A...-T.....G-T-A-..	740
D.CD.7226TT.....A...-T.....G-T-A-..	255
D.CD.84ZR085TT.....A...-T.....G-T-A-..	85
F1.BR.93BR020TT.....A...-T.....G-T-A-..	83
F1.FI.FIN9363TT.....A...-T.....G-T-A-..	91
F1.BE.VI850TT.....A...-T.....G-T-A-..	0
F1.FR.MF411TT.....A...-T.....G-T-A-..	0
F2.CM.MF255CTT.....A...-T.....G-T-A-..	0
F2.CM.MF257CTT.....A...-T.....G-T-A-..	0

Gag p7 nucleocapsid binds these loops, see DeGuzman, Science 279:384-8
 Packaging signal secondary structure, see Harrison, J Virol 72:5886-96

B.FR.HXB2R	G.....ACCTGAAGCG.AAAGGAAA.CCAGAGG...AGCT...CTCTCGAGC...CA..GGACTCGGCTTGCTGA.AGCGCGCAGCGCAAGAGGGGAGGG.CGGCGA	740
G.BE.DRCBL	-ACTTTGAAAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-A-A-.....A-.....	725
G.NG.92NG083	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-A-A-.....A-.....	107
G.SE.SEG165	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-A-A-.....A-.....	143
G.FI.HH87932	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-A-A-.....A-.....	142
H.BE.VI991T-.....CA-.....A-.....T-.....G-T-A-A-.....A-.....	118
H.BE.VI997T-.....A-.....T-.....G-T-A-A-.....A-.....	59
H.CF.90CF056T-.....TA-.....A-.....T-.....G-T-A-A-.....A-.....	86
J.SE.SE91733T-.....A-.....T-.....G-T-A-A-.....A-.....	60
J.SE.SE92809T-.....A-.....T-.....G-T-A-A-.....A-.....	60
K.CM.MP535CT-.....A-.....T-.....G-T-A-A-.....A-.....	0
K.CD.EQB1LICT-.....A-.....T-.....G-T-A-A-.....A-.....	0
O.CM.ANT170T-.....T-.....G-ARAG-AAC.....AC-G-.....AC-G-.....T-A-C-CT-.....A-.....AACTC-	773
O.CM.MVP5180T-.....T-.....G-ARAG-AAC.....AC-G-.....AC-G-.....T-A-C-CT-.....A-.....AACTC-	747
O.SN.MP1299T-.....T-.....G-ARAG-AAA.....C-CGAC.....G-AC-G-.....T-A-C-CT-.....A-.....AACTC-	774
O.SN.MP1300T-.....T-.....G-ARAG-AAA.....C-CGAC.....G-AC-G-.....T-A-C-CT-.....A-.....AACTC-	774
N.CM.YBF30T-.....TAG-.....G-CT-GAA.....A-.....C-T-G.....T-A-A-G.....G-.....	296
N.CM.YBF106T-.....TAG-.....G-CT-GAA.....A-.....C-T-G.....T-A-A-G.....G-.....	297
CONSENSUS 01	-?ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-a-a-A-.....g-gag?-.....	701
01.AE.TH.93TH253	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-a-a-A-.....A-.....	759
01.AE.TH.CM240	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-a-a-A-.....A-.....	307
01.AE.CF.90CF402	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-a-a-A-.....A-.....	782
01.AE.TH.92TH022	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-a-a-A-.....A-.....	711
01.AE.TH.95THH047	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-a-a-A-.....A-.....	711
01.AE.CF.11697	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-a-a-A-.....A-.....	696
01.AE.TH.93TH057	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-a-a-A-.....A-.....	701
01.AE.TH.93TH065	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-a-a-A-.....A-.....	701
01.AE.CF.4071	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....C-.....G-.....T-A-G-A-.....A-.....	684
CONSENSUS 02T-.....A-.....T-.....G-.....T-A-G-A-.....A-.....	0
02.AG.SE.7812	-ACTTTG.....AAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-A-A-.....A-.....	129
02.AG.GH.6829T-.....A-.....T-.....G-.....T-A-G-A-.....A-.....	0
02.AG.DJ.DJ263T-.....A-.....T-.....G-.....T-A-G-A-.....A-.....	105
02.AG.DJ.DJ264T-.....A-.....T-.....G-.....T-A-G-A-.....A-.....	106
02.AG.NG.18NGACTTGAAGCGGT.AAATAGGG-TC-.....TT-.....A-.....T-.....G-T-.....A-.....	282
02.AG.SN.MP1211T-.....A-.....T-.....G-.....T-A-G-A-.....A-.....	0
02.AG.SN.MP1213T-.....A-.....T-.....G-.....T-A-G-A-.....A-.....	0
02.AG.CM.MP807T-.....A-.....T-.....G-.....T-A-G-A-.....A-.....	0
03.AB.RU.98001T-.....A-.....T-.....G-.....T-A-G-A-.....A-.....	114
03.AB.RU.KAL153-2T-.....A-.....T-.....G-.....T-A-G-A-.....A-.....	543
04.CPX.GR.97PVMY	-ACCCGAAAGTGAAGTTTAATAGGG-TC-.....TT-.....G-.....T-.....G-T-A-A-.....A-.....	761
04.CPX.GR.97PVGH	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....G-.....T-.....G-T-A-A-.....A-.....	757
04.CPX.CY.94CY0323TTGAAAGTGAAGTTTAATAGG-TC-.....TT-.....A-.....T-.....G-T-A-A-.....A-.....	106
05.DF.BE.V11310T-.....TAG-.....A-.....AA.....A-.....TA-.....	116
05.DF.BE.V1961T-.....TAG-.....A-.....AA.....A-.....TA-.....	92
06.CPX.ML.95ML84	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-A-A-.....A-.....	0
06.CPX.AU.BFP90	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-A-A-.....A-.....	771
06.CPX.ML.95ML127	-ACTTGAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-A-A-.....A-.....	758
06.CPX.SN.97SEL078	-ACCCGAAAGCGAAAGTTTAATAGGG-TC-.....TT-.....A-.....T-.....G-T-A-A-.....A-.....	782
10.CD.BFL061T-.....TAG-.....A-.....	118
10.CD.BFL110T-.....TAG-.....A-.....	117
10.CD.BFL071T-.....TAG-.....A-.....	117
11.CPX.CM.MP818T-.....C-.....G-A-.....G-T-A-A-.....A-.....	747
11.CPX.FR.MP1298T-.....TAG-.....A-.....T-.....G-T-A-A-.....A-.....	740
11.CPX.FR.MP1307T-.....TAG-.....A-.....T-.....G-T-A-A-.....A-.....	751
11.CPX.GR.GR17T-.....TAG-.....A-.....T-.....G-T-A-A-.....A-.....	55
CPZ.GA.SIVCPZGABT-.....TA-.....G-CT-GAA.....C-.....T-.....A-.....A-.....	756
CPZ.US.SIVCPZUST-.....TA-.....CTG-AAC.....T-.....CG-.....C-.....T-.....A-.....A-.....AACT	762
CPZ.CM.SIVCPZCM3T-.....TTTG-.....T-.....G-.....TA-.....A-.....G-.....T-.....A-.....A-.....AACT	111
CPZ.CM.SIVCPZCM5T-.....T-.....AAG-.....T-.....G-.....TA-.....A-.....G-.....T-.....A-.....A-.....AACT	414
CPZ.CD.SIVCPZANTT-.....GGAG.....T-.....C-.....A-.....AAC.....G-.....T-.....A-.....GTA-.....GA	74

B.FR.HXB2R	CTTGGTGAGTACGCAAAA	TTTTTGACTAGCGGA	GGCTAGAAG	End packaging loops	- -> Gag and Gag-Pol CDS start	833
CONSENSUS A	????????????????	?????????????g	ggctagaag	-----A	-----gt-----c-----ggAaaLta--9C	64
A.UG.U455	-----T-----T	-----T-----T	-----T-----T	-----A	-----AA-----AA-----AA-----AA	300
A.SE.UGSE8891	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	38
A.SE.TZSE8538	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	38
A.SE.UGSE6594	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	38
A.KE.Q2317	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	297
A.SE.UGSE7535	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	65
A.SE.SOSE7253	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	40
A.UG.92UG037	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	203
A.SE.SE8131	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	229
A2.CD.97CDKTB48	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	180
A2.CY.94CY017	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	202
CONSENSUS B	g-aaaaa?..?tt	T-----T	a-??	ga	ag-a--aa-c-gg-aga-a-taa	784
B.AU.MBC925	-----A-----T	-----T-----T	G-----G	AG	-----A-----A-----A-----A	835
B.GA.OYI	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	379
B.GB.MANC	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	779
B.GB.CAM1	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	835
B.CN.RL42	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	205
B.DE.D31	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	379
B.DE.HAN2	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	302
B.ES.89SP061	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	382
B.NL.ACH320A	-----C-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	835
B.ITW.LM49	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	832
B.US.DHL23	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	830
B.US.SF2	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	834
B.US.NY5CG	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	833
B.US.WCIPR18	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	379
B.US.YU2	-----A-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	831
B.US.JRC5F	-----G-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	833
B.US.JRFL	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	156
B.US.MN	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	830
B.US.BC5G3	-----G-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	830
B.US.WEAU160	-----G-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	832
B.US.RF	-----G-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	347
B.US.WR27	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	212
CONSENSUS C	acttT? .ta	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	219
C.BR.92BR025	TTTT .TA	TTTT .TA	-----T-----T	ag	-----At-----gA-g-g-A-a-aa	183
C.BW.96BW0402	TTTT .TA	TTTT .TA	-----T-----T	-----T-----T	-----T-----T-----T-----T	323
C.BW.96BW1104	TTTT .TA	TTTT .TA	-----T-----T	-----T-----T	-----T-----T-----T-----T	223
C.BW.96BW15C02	TTTT .TA	TTTT .TA	-----T-----T	-----T-----T	-----T-----T-----T-----T	324
C.BW.96BW01B03	TTTT .TA	TTTT .TA	-----T-----T	-----T-----T	-----T-----T-----T-----T	323
C.BW.96BW0502	TTTT .TA	TTTT .TA	-----T-----T	-----T-----T	-----T-----T-----T-----T	320
C.BW.96BW16B01	TTTT .TA	TTTT .TA	-----T-----T	-----T-----T	-----T-----T-----T-----T	323
C.BW.96BW1210	TTTT .TA	TTTT .TA	-----T-----T	-----T-----T	-----T-----T-----T-----T	213
C.ET.FTH2220	TTTT .TA	TTTT .TA	-----T-----T	-----T-----T	-----T-----T-----T-----T	211
C.IN.301999	TTTT .TA	TTTT .TA	-----T-----T	-----T-----T	-----T-----T-----T-----T	211
C.IN.21068	TTTT .TA	TTTT .TA	-----T-----T	-----T-----T	-----T-----T-----T-----T	211
C.IN.301905	TTTT .TA	TTTT .TA	-----T-----T	-----T-----T	-----T-----T-----T-----T	211
C.IN.301904	TTTT .TA	TTTT .TA	-----T-----T	-----T-----T	-----T-----T-----T-----T	211
C.IN.93IN101	TTTTA .TA	TTTTA .TA	-----T-----T	-----T-----T	-----T-----T-----T-----T	833
C.IN.11246	TTTT .TA	TTTT .TA	-----T-----T	-----T-----T	-----T-----T-----T-----T	211
D.UG.94UG114	T-----T	T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	182
D.CD.NDK	T-----T	T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	377
D.CD.ELI	T-----T	T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	379
D.KE.MB2059	T-----T	T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	835
D.CD.2226	T-----T	T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	350
D.CD.84Z8085	T-----T	T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	350
F1.BR.93BR020	T-----T	T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	176
F1.FI.FIN9363	T-----T	T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	174
F1.BE.VI850	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	168
F1.FR.MF411	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	44
F2.CM.MF255C	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	44
F2.CM.MF257C	-----T-----T	-----T-----T	-----T-----T	-----T-----T	-----T-----T-----T-----T	44

B. FR. HXB2R	ATGGGAAAATTCGGTTAAGCCAGGGGAAAGAAAATAATATAATTAACATAGTATGGCAACGAGGAGCTAGAACGATTGCGAGTTAACTCCCTGGCCCTTTGAAACAATCAGAGGCTGTAGA	963
CONSENSUS A	a-----a-g-aa-a-g-c-g-a-c-ag-----c-----c-----g-aag-ttg-c-c-a-----t-----a-----a-a-tcCa	194
A.UG.U455	-----c-----c-g-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----cAG	430
A.SE.UGSE8891	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----cCA	168
A.SE.TZSE8538	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----cCA	168
A.SE.UGSE6594	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----cCA	427
A.KE.Q2317	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----cCA	195
A.SE.UGSE7535	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----cCA	170
A.SE.SOSE7253	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----cCA	333
A.UG.92UG037	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----cCA	359
A.SE.SE8131	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----cCA	310
A2.CD.97CDKTB48	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----cCA	332
A2.CY.94CY017	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----cCA	914
CONSENSUS B	a-----a-aaa-c-gt-a-g-g-----g-----g-----g-aag-ttg-c-c-a-----t-----g-----g-----a-----a-----c-t	965
B.AU.MBC925	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	962
B.AU.MBC200	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	962
B.GA.OYI	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	509
B.GB.MANC	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	909
B.GB.CAMI	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	965
B.CN.RL42	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	335
B.DE.D31	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	432
B.DE.HAN2	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	509
B.ES.89SP061	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	512
B.NL.ACH320A	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	965
B.TM.LM49	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	962
B.US.DH123	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	960
B.US.SF2	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	964
B.US.NY5CG	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	963
B.US.WCIPR18	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	509
B.US.YU2	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	961
B.US.JRCSF	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	963
B.US.JRFL	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	286
B.US.MN	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	960
B.US.BCSG3	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	510
B.US.WEAU160	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	962
B.US.RF	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	477
B.US.WR27	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	342
CONSENSUS C	a-----a-ag-g-g-g-----g-acac-tgc-a-a-cc-ag-a-----g-----g-aaga-t-ac-ta-c-----g-t-agaga-t-a-a-----a	349
C.BR.92BR025	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	313
C.BW.96BW0402	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	453
C.BW.96BW1104	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	454
C.BW.96BW15C02	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	454
C.BW.96BW01B03	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	453
C.BW.96BW0502	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	450
C.BW.96BW16B01	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	453
C.BW.96BW1210	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	353
C.ET.ETH2220	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	343
C.IN.301999	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	342
C.IN.21068	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	341
C.IN.301905	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	341
C.IN.301904	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	341
C.IN.93JN101	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	963
C.IN.11246	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	341
D.UG.94UG114	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	312
D.CD.NDK	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	507
D.CD.ELI	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	509
D.KE.MB2059	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	509
D.CD.42R085	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	965
F1.BR.93BR020	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	480
F1.FI.FIN9363	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	963
F1.BE.VI850	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	306
F1.FR.MF411	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	304
F2.CM.MF255C	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	174
F2.CM.MF257C	-----g-----g-a-g-----g-a-g-----t-----g-aa-c-c-----t-----g-----g-----a-----a-----c-t	174

HIV-1/SIVcpz complete genomes

B. FR. HXB2R	ACAAGATAGAGGACGACCAAAACAAAGTAAG	GCAGCAGCT	GACACAGGA	..	CACAGCAAT	..	CA	1169
CONSENSUS A	-L--Ra--aATac--T- g-9ca--C--gAc-caa- g-	---gc----	g--gca--a	1170
A.UG.UG55	-T--A--ATA--T-G-AC--C--GA--A--	---	A--	1171
A.SE.UGSE8891	-T-AT--ATA--T-G-AC--C--GA--A--	---	A--	1172
A.SE.TZSE8538	A.SE.UGSE6594	-T-AT--ATA--T-G-AC--C--GA--A--	---	A--	1173
A.KE.Q2317	-T-AT--ATA--T-G-AC--C--GA--A--	---	A--	1174
A.SE.UGSE7535	-T-AT--ATA--T-G-AC--C--GA--A--	---	A--	1175
A.SE.SOSE7253	-T-AT--ATA--T-G-AC--C--GA--A--	---	A--	1176
A.UG.92UG037	-T-AT--ATA--T-G-AC--C--GA--A--	---	A--	1177
A.SE.SE131	-T-AT--ATA--T-G-AC--C--GA--A--	---	A--	1178
A2.CD.97CDKTB48	-T-AT--ATA--T-G-AC--C--GA--A--	---	A--	1179
A2.CY.94CY017	-T-AT--ATA--T-G-AC--C--GA--A--	---	A--	1180
CONSENSUS B	-g-aga--aa-a-t-a-?a--aagca--aa	..gcagcact????	-a-a-g-	1181
B.AU.MBC925	-G-G-G-	---	A--	1182
B.GA.OYI	-T--	---	A--	1183
B.GB.MANC	-G-A	---	A--	1184
B.GB.CAMI	-G-A	---	A--	1185
B.CN.RL42	-T--	---	A--	1186
B.DE.D31	-T--	---	A--	1187
B.DE.HAN2	-T--	---	A--	1188
B.ES.89SP061	-T--	---	A--	1189
B.NL.ACH320A	-T--	---	A--	1190
B.TM.LM49	-T--	---	A--	1191
B.US.DH123	-G--	---	A--	1192
B.US.NY5CG	-T--	---	A--	1193
B.US.WCIPR18	-G-A	---	A--	1194
B.US.YU2	-G-A	---	A--	1195
B.US.JRCSF	-A--	---	A--	1196
B.US.JRFL	-A--	---	A--	1197
B.US.MN	-G-A	---	A--	1198
B.US.BCSG3	-G-A	---	A--	1199
B.US.WEAU160	-T--	---	A--	1200
B.US.RF	-T--	---	A--	1201
B.US.WR27	-G--	---	A--	1202
CONSENSUS C	-C-g-g-g-A-a-a- -sg-C--a-aac-g-g	..-ga-gagct	-acgacgg?	1203
C.BR.92BR025	-A--	---	A--	1204
C.BW.96BW0402	-A--	---	A--	1205
C.BW.96BW104	-A--	---	A--	1206
C.BW.96BW15C02	-A--	---	A--	1207
C.BW.96BW01B03	-A--	---	A--	1208
C.BW.96BW0502	-A--	---	A--	1209
C.BW.96BW16B01	-T--	---	A--	1210
C.BW.96BW1210	-T--	---	A--	1211
C.ET.ETH2220	-A--	---	A--	1212
C.IN.301999	-A--	---	A--	1213
C.IN.21068	-A--	---	A--	1214
C.IN.301905	-A--	---	A--	1215
C.IN.301904	-A--	---	A--	1216
C.IN.93JN101	-A--	---	A--	1217
C.IN.11246	-A--	---	A--	1218
D.UG.94UG114	-A--	---	A--	1219
D.CD.NDK	-A--	---	A--	1220
D.CD.ELI	-A--	---	A--	1221
D.KE.MB2059	-A--	---	A--	1222
D.CD.Z226	-A--	---	A--	1223
D.CD.84ZR085	-G-C	---	A--	1224
F1.BR.93BR020	-T-C	---	A--	1225
F1.FI.FIN9363	-C	---	A--	1226
F1.BE.VI850	-G-T	---	A--	1227
F1.FR.MF411	-T-C-C	---	A--	1228
F2.CM.ME255C	-T-C-C	---	A--	1229
F2.CM.ME257C	-T-C-C	---	A--	1230

B. FR. HXB2R
G. BE. DRBEL
G. NG. 92NG083
G. SE. SE6165
G. FI. HH87932
H. BE. VI991
H. BE. VI997
H. CF. 90CF056
H. SE. SE91733
J. SE. SE92809
K. CM. MP535C
K. CD. E07B11C
O. CM. AN170
O. CM. MP5180
O. SN. MP1299
O. SN. MP1300
N. CM. YBF300
N. CM. YBF106
CONSENSUS 01
01 AE. TH. 93TH253
01 AE. TH. CM240
01 AE. CF. 90CF402
01 AE. TH. 92TH022
01 AE. TH. 95TNLH047
01 AE. CF. 11697
01 AE. TH. 93TH057
01 AE. TH. 93TH065
01 AE. CF. 4071
CONSENSUS 02
02 AG. SE. 7812
02 AG. GH. G829
02 AG. DJ. DJ263
02 AG. DJ. DJ264
02 AG. NG. INBG
02 AG. SN. MP1211
02 AG. SN. MP1213
02 AG. CM. MP807
03 AB. RU. 98001
03 AB. RU. KALI53-2
04 CPX. GR. 97PVMY
04 CPX. GR. 97PVCH
04 CPX. CY. 94CY0323
05 DF. BE. V11310
05 DF. BE. V1961
06 CPX. ML. 95ML84
06 CPX. AU. BPP90
06 CPX. ML. 95ML127
06 CPX. SN. 97SE1078
10 CD. BEL061
10 CD. BEL110
10 CD. BEL171
11 CPX. CM. MP818
11 CPX. FR. MP1298
11 CPX. FR. MP1307
11 CPX. GR. GR17
CPZ. GA. STVCPZAB
CPZ. US. STVCPZUS
CPZ. CM. STVCPZAM3
CPZ. CM. STVCPZAM5
CPZ. CD. STVCPZANT

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz complete genomes

B. FR. HXB2R	ATGTTTTCAGCATATCAGAAAGAGCCACCCCAAGTTTAAACCCCATGCTAAACACAGCTGGGGGGACATCAAGCAGCCATGCAAAATGTTAAAGACCCATCAATGAGGAAGCTGCAGATGGGATA	1429
CONSENSUS A	a---c---a---c---c---acaa--t---t---tg---g---caT---	855
A.UG.U455	-----G-----GGGT-C---T---TG---G---IGT---	884
A.SE.UGSE8891	-----C-----T---A---G---T---T---C---C---T---	934
A.SE.TZSE8538	-----C-----T---TG---G---T---T---C---C---T---	634
A.SE.UGSE6594	-----C-----T---TG---G---T---T---C---C---T---	634
A.KE.Q2317	-----C-----T---TG---G---T---T---C---C---T---	893
A.SE.UGSE7535	G---C---T---T---TG---G---T---T---C---C---T---	658
A.SE.SOSE7253	-----C-----T---TG---G---T---T---C---C---T---	636
A.UG.92UG037	-----C-----T---TG---G---T---T---C---C---T---	799
A.SE.SE8131	-----C-----T---TG---G---T---T---C---C---T---	825
A2.CD.97CDKTB48	-----A-----T---C---T---T---T---T---T---T---C---	794
A2.CY.94CY017	-----A-----T---C---T---T---T---T---T---T---C---	786
CONSENSUS B	-----a---t---c---t---a---c---c---t---a---c---t---a---t---g---c---	1380
B.AU.MBC925	-----G-----T---	1434
B.AU.MBC200	-----T---	1428
B.GA.OYI	-----C-----T---	975
B.GB.MANC	-----T---	1375
B.GB.CAMI	-----A---	1431
B.CN.RL42	-----A---	801
B.DE.D31	-----A---	975
B.DE.HAN2	-----A---	898
B.ES.89SP061	-----A---	978
B.NL.ACH320A	-----C---G---G---	1431
B.TM.LM49	-----T---	1428
B.US.DH123	-----T---	1426
B.US.SF2	-----T---	1436
B.US.NY5CG	-----T---	1429
B.US.WCIPR18	-----C---	975
B.US.YU2	-----C---	1427
B.US.JRC5F	-----T---C---A---	1429
B.US.JRFL	-----T---C---A---	746
B.US.MN	-----T---	1435
B.US.BCSG3	-----T---	976
B.US.WEAU160	-----T---	1428
B.US.RF	-----T---	943
B.US.WR27	-----A---at---a---a---tt-a-c---t---aT---g---t---a-a-g---a---t---c---g---g---aa---t---t---	808
CONSENSUS C	-----A---at---a---a---tt-a-c---t---aT---g---t---a-a-g---a---t---c---g---g---aa---t---t---	808
C.BR.92BR025	-----A---C---C---T---	773
C.BW.96BW0402	-----A---C---C---T---	913
C.BW.96BW1104	-----A---C---C---T---	810
C.BW.96BW15C02	-----A---C---C---T---	911
C.BW.96BW01B03	-----A---C---C---T---	910
C.BW.96BW0502	-----A---C---C---T---	907
C.BW.96BW16B01	-----A---C---C---T---	910
C.BW.96BW1210	-----A---C---C---T---	813
C.ET.ETH2220	-----A---C---C---T---	803
C.IN.301999	-----A---C---C---T---	799
C.IN.21068	-----A---C---C---T---	798
C.IN.301905	-----A---C---C---T---	798
C.IN.301904	-----A---C---C---T---	798
C.IN.93IN101	-----A---C---C---T---	1420
C.IN.11246	-----A---C---C---T---	789
D.UG.94UG114	-----A---C---C---T---	778
D.CD.NDK	-----A---C---C---T---	964
D.CD.ELI	-----A---C---C---T---	975
D.KE.MB2059	-----A---C---C---T---	1431
D.CD.Z226	-----A---C---C---T---	1432
D.CD.84ZR085	-----A---C---C---T---	1432
F1.BR.93BR020	-----G---T---T---T---T---T---T---C---G---	946
F1.FI.FIN9363	-----G---T---T---T---T---T---T---C---G---	760
F1.BE.VI850	-----G---T---T---T---T---T---T---C---G---	764
F1.FR.MF411	-----G---T---T---T---T---T---T---C---G---	772
F1.FR.MF411	-----G---T---T---T---T---T---T---C---G---	628
F2.CM.ME255C	-----G---T---T---T---T---T---T---C---G---	628
F2.CM.ME257C	-----G---T---T---T---T---T---T---C---G---	640

B. FR. HXB2R	ATGTTTTCAGCATTTACAGAGGAGCCACCACCAAGATTAAACACCATGCTAAACACAGTGGGGGACATCAAGCAGCCATGCAAAATGTTAAAGAGACCACTCAATGAGGAAGTGCAGAAATGGGATA	1429
G. BE. DRGB1	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----C-----C-----	1393
G. NG. 92NG083	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	935
G. SE. SE6165	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	832
G. FI. HH87932	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	831
H. BE. VI991	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	821
H. BE. VI997	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	776
H. CF. 90CF056	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	756
J. SE. SE91733	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	753
J. SE. SE92809	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	752
K. CM. MP535C	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	628
K. CD. E0TB11C	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	628
O. CM. ANT70	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1475
O. CM. MP6180	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1450
O. SN. MP1299	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1477
O. SN. MP1300	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1477
N. CM. YBF30	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	993
N. CM. YBF106	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	994
N. CM. YBF106	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	994
CONSENSUS 01	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1387
01. AE. TH. 93TH253	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1446
01. AE. TH. CM240	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	994
01. AE. CF. 90CF402	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1471
01. AE. TH. 92TH022	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1398
01. AE. TH. 93TH057	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1384
01. AE. TH. 93TH065	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1387
01. AE. CF. 4071	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1372
CONSENSUS 02	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	652
02. AG. SE. 7812	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	807
02. AG. GH. G829	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	651
02. AG. DJ. DJ263	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	783
02. AG. DJ. DJ264	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	784
02. AG. NG. I8NG	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	960
02. AG. SN. MP1211	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	632
02. AG. SN. MP1213	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	629
02. AG. CM. M807	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	591
03. AB. RU. 98001	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	805
03. AB. RU. KALI53-2	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1204
04. CPX. GR. 97PVMY	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1443
04. CPX. GR. 97PVCB	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1446
04. CPX. CY. 94CY0323	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	795
05. DF. BE. VI1310	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	808
06. DF. BE. VI961	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	784
06. CPX. ML. 95ML84	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	666
06. CPX. AU. BPP90	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1454
06. CPX. ML. 95ML127	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1446
06. CPX. SN. 97SE1078	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1473
10. CD. BFL061	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	808
10. CD. BFL110	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	806
10. CD. BFL1071	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	806
11. CPX. CM. M818	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1435
11. CPX. FR. MP1298	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1432
11. CPX. FR. MP1307	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1441
11. CPX. GR. GR17	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	740
CPZ. GA. STVCPZGAB	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1490
CPZ. US. STVCPZUS	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1477
CPZ. CM. STVCPZCAM3	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	819
CPZ. CM. STVCPZCAM5	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	1122
CPZ. CD. STVCPZANT	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	841
Gag	-----G-----G-----T-----T-----G-----G-----T-----T-----G-----T-----T-----T-----G-----G-----	Gag

HIV-1/SIVcpz complete genomes

Genomic alignment data showing source identifiers (e.g., B.FR.HXB2R), nucleotide sequences, and positions (1510-971). Includes a sequence logo at the bottom: R_V_H_P_V_H_A_G_P_I_A_P_G_O_M_R_E_P_R_G_S_D_I_A_G_T_T_S_T_L_Q_E_O_I_G_W_M_T_..._N_N_P

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz Complete Genomes

Table with columns: Strain ID, Gag p24 Capsid end, Gag p2 start, Gag p2 end, Gag p7 start, Gag-Pol fusion TF protein start, and HIV-1/SIVcpz complete genomes. The table contains multiple rows of genomic data for various HIV-1 and SIVcpz strains, including B.FR.HXB2R, G.BE.DRCBL, G.NG.92NG083, G.SE.SE6165, G.FI.HH87932, H.BE.VI991, H.BE.VI997, H.CF.90CF056, J.SE.SE91733, J.SE.SE92809, K.CM.MP535C, K.CD.EQIB11C, O.CM.AN170, O.CM.MVP5180, O.SN.MP1299, O.SN.MP1300, N.CM.YBF30, N.CM.YBF106, CONSENSUS 01, 01.AE.TH.93TH253, 01.AE.TH.CM240, 01.AE.CF.90CF402, 01.AE.TH.92TH022, 01.AE.TH.95THH047, 01.AE.CF.11697, 01.AE.TH.93TH057, 01.AE.TH.93TH065, 01.AE.CF.4071, CONSENSUS 02, 02.AG.SE.7812, 02.AG.GH.G829, 02.AG.DJ.DJ263, 02.AG.DJ.DJ264, 02.AG.NG.IBNC, 02.AG.SN.MP1211, 02.AG.SN.MP1213, 02.AG.CM.MP807, 03.AB.RU.98001, 03.AB.RU.KAL153-2, 04.CPX.GR.97PVMY, 04.CPX.GR.97PVGH, 04.CPX.CY.94CX0323, 05.DF.BE.VI1310, 05.DF.BE.VI961, 06.CPX.MI.95MI84, 06.CPX.AU.BFP90, 06.CPX.ML.95ML127, 06.CPX.SN.97SML1078, 10.CD.BFL061, 10.CD.BFL110, 10.CD.BFL1071, 11.CPX.CM.MP818, 11.CPX.FR.MP1298, 11.CPX.FR.MP1307, 11.CPX.GR.GR17, CPZ.GA.SIVCPZGAB, CPZ.US.SIVCPZUS, CPZ.CM.SIVCPZCM3, CPZ.CM.SIVCPZCM5, CPZ.CD.SIVCPZANT, B.FR.HXB2R, Gag p24 Capsid end, Gag p2 start, Gag p2 end, Gag p7 start, Gag-Pol fusion TF protein start, HIV-1/SIVcpz complete genomes.

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz complete genomes

B. FR. HXB2R	GAGAGCAATTTT...AGGAACCAAGAAAGATTGTAAGTGTTC	1252
CONSENSUS A	--g-g--t-t-t...99GGc-ag-aa...9GA	1269
A.UG.U455	...GG-CG...GAA	1501
A.SE.UGSE8891	...A-GG-G-A...GAA	1251
A.SE.TZSE8538	...AGG...GAA	1251
A.SE.UGSE6594	...GG-G-A...GGA	1251
A.KE.Q2317	...A-GG-G-A...GAA	1510
A.SE.UGSE7535	...GG-G-AG...GGA	1275
A.SE.SOSE7253	...GGA-G-A...GGA	1256
A.UG.92UG037	...A-GG-G-A...GGA	1416
A.SE.SE8131	...GG-G-A...AGA	1442
A2.CD.97CDKTB48	...A-GGT...A...GAA	1411
A2.CY.94CY017	...GGT...A...GAA	1409
CONSENSUS B	-ag--c-t-tt...ggaac-aa-g-aeg-cg-t-g-t-tc-at	2003
B.AU.MBC925	...-A-C-C-A	2057
B.AU.MBC200	...-A-C-C-A	2051
B.GA.OYI	...-A-C-C-A	1595
B.GB.MANC	...-A-C-C-A	1998
B.GB.CAMI	...-A-C-C-A	2054
B.CN.RL42	...-A-C-C-A	1439
B.DE.D31	...-A-C-C-A	1598
B.DE.HAN2	...-A-C-C-A	1521
B.ES.89SP061	...-A-C-C-A	1601
B.NL.ACH320A	...-A-C-C-A	2054
B.TM.LM49	...-A-C-C-A	2051
B.US.DH123	...-A-C-C-A	2049
B.US.SF2	...-A-C-C-A	2059
B.US.NY5CG	...-A-C-C-A	2052
B.US.WCIPR18	...-A-C-C-A	1598
B.US.YU2	...-A-C-C-A	2050
B.US.JRCFS	...-A-C-C-A	2052
B.US.JRFL	...-A-C-C-A	1369
B.US.MN	...-A-C-C-A	2058
B.US.BCSG3	...-A-C-C-A	1599
B.US.WEAU160	...-A-C-C-A	2051
B.US.RF	...-A-C-C-A	1566
B.US.WR27	...-A-C-C-A	1431
CONSENSUS C	g-g-agg--t-tt...Aagg-cct-aa-9Aattgt-A-t	1428
C.BR.92BR025	A--A--G...-AAGG-CT-A-GA-C-A--A-C	1393
C.BW.96BW0402	...-AAGG-CT-A-GA	1430
C.BW.96BW1104	...-AAGG-CT-A-GA	1531
C.BW.96BW15C02	...-AAGG-CT-A-GA	1530
C.BW.96BW01B03	...-AAGG-CT-A-GA	1527
C.BW.96BW0502	...-AAGG-CT-A-GA	1530
C.BW.96BW16B01	...-AAGG-CT-A-GA	1433
C.BW.96BW1210	...-AAGG-CT-A-GA	1423
C.ET.ETH2220	...-AAGG-CT-A-GA	1416
C.IN.301999	...-AAGG-CT-A-GA	1415
C.IN.21068	...-AAGG-CT-A-GA	1415
C.IN.301905	...-AAGG-CT-A-GA	1415
C.IN.301904	...-AAGG-CT-A-GA	2037
C.IN.93JN101	...-AAGG-CT-A-GA	1406
C.IN.11246	...-AAGG-CT-A-GA	1590
D.UG.94UG114	...-AAGG-CT-A-GA	1601
D.CD.NDK	...-AAGG-CT-A-GA	2057
D.CD.ELI	...-AAGG-CT-A-GA	2058
D.KE.MB2059	...-AAGG-CT-A-GA	1572
D.CD.84ZR085	...-AAGG-CT-A-GA	1380
F1.BR.93BR020	...-AAGG-CT-A-GA	1381
F1.FI.FIN9363	...-AAGG-CT-A-GA	1389
F1.BE.VI850	...-AAGG-CT-A-GA	1348
F1.FR.MF411	...-AAGG-CT-A-GA	1348
F2.CM.ME255C	...-AAGG-CT-A-GA	1260
F2.CM.ME257C	...-AAGG-CT-A-GA	

	Gag-p7	nucleocapsid end	Gag p1 start	Gag p1 end	Gag p6 start	
B.FR.HXB2R	CAAATGAAGATTGACT	...GAGAGAG...GCTAATTT	TTTAGGGAGATCTGGCCCTTC	...CTACAA...GGGAAG	...GCCAGGGAATTTCTTCAGACGACC	2148
CONSENSUS A	---a-c-c---a	---a-t---aa-t---	---a-t---aa-t---	---agc-a...g-g---	---a-a---c---g---c---	1365
A.UG.U455	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---A...G---	---C---C---C---	1397
A.SE.UGSE8891	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---A...G---	---A...C---C---	1347
A.SE.TZSE8538	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---AG...A-G---	---A...C---C---	1347
A.SE.UGSE6594	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---AG...A-G---	---A...C---C---	1347
A.KE.Q2317	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---AG...A-G---	---A...C---C---	1606
A.SE.UGSE7535	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---AGA...G---	---G...A...C---A---T---	1371
A.SE.SOSE7253	---g-c-c-c---	---a-g---a-g---	---a-g---a-g---	---AG...G---	---C---C---C---	1352
A.UG.92UG037	---g-c-c-c---	---a-g---a-g---	---a-g---a-g---	---AG...G---	---C---C---C---	1512
A.SE.SE8131	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---G...A...C---C---	1538
A2.CD.97CDKTB48	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---A...A...G---	---C---C---C---GA---	1507
A2.CI.94CIY017	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---A...A...G---	---C---C---C---GA---	1505
CONSENSUS B	---aa-t-tacc????-aga	---t-a---	---g-t-t-c-t---	---cac-a...g-ga---	---csg-ga-t-t-t---cg---	2099
B.AU.MBC925	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---G...A---	2147
B.AU.MBC200	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---G...A---	1691
B.GA.OYI	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---G...A---	2094
B.GB.MANC	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---G...A---	2150
B.GB.CAMI	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---G...A---	1535
B.CN.RL42	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---G...A---	1694
B.DE.D31	---g-c-c-a---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---C---C---A---T---	1617
B.DE.HAN2	---g-c-c-a---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---C---C---A---T---	1697
B.ES.89SP061	---g-c-c-a---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---C---C---A---T---	2150
B.NL.ACH320A	---g-c-c-a---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---C---C---A---T---	2147
B.IW.LM49	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---C---C---A---T---	2145
B.US.DH123	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---C---C---A---T---	2155
B.US.SF2	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---C---C---A---T---	2148
B.US.NY5CG	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---C---C---A---T---	1694
B.US.WCIPRL18	---c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---C---C---A---T---	2146
B.US.YU2	---g-c-c---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---C---C---A---T---	2148
B.US.JRGSF	---g-c-c---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---C---C---A---T---	1465
B.US.JRFL	---g-c-c---	---a-t---aa-t---	---a-t---aa-t---	---G...G---	---C---C---A---T---	2151
B.US.MN	---g-c-c---	---a-t---aa-t---	---a-t---aa-t---	---G...G---	---C---C---A---T---	1695
B.US.RCSG3	---g-c-c---	---a-t---aa-t---	---a-t---aa-t---	---G...G---	---C---C---A---T---	1668
B.US.WEAU160	---g-c-c---	---a-t---aa-t---	---a-t---aa-t---	---G...G---	---C---C---A---T---	2147
B.US.RF	---c-c-aatgag-gac---	---a-t---aa-t---	---a-t---aa-t---	---C...G---	---C---C---A---T---	1527
B.US.WR27	---c-c-yn---	---c-m---	---c-m---	---C...G---	---C---C---A---T---	1524
CONSENSUS C	---a-c-a-t---	---ga-g-a---	---a-aaa-t---	---c-a...g---	---g-ttc-t---gc---cc---	1489
C.BR.92BR025	---g-c-c---	---g-c-c---	---a-t---aa-t---	---C...G---	---A...C---A---A---	1626
C.BW.96BW0402	---c-c-c---	---g-g-g---	---a-t---aa-t---	---C...G---	---C...C---C---G---	1526
C.BW.96BW1104	---c-c-c---	---g-g-g---	---a-t---aa-t---	---C...G---	---C...C---C---G---	1627
C.BW.96BW15C02	---c-c-c---	---g-g-g---	---a-t---aa-t---	---C...G---	---C...C---A---A---T---	1626
C.BW.96BW01B03	---c-c-c---	---g-g-g---	---a-t---aa-t---	---C...G---	---C...C---A---A---T---	1623
C.BW.96BW0502	---c-c-c---	---g-g-g---	---a-t---aa-t---	---C...G---	---C...C---A---A---T---	1626
C.BW.96BW16B01	---c-c-t---	---g-g-g---	---g-a-t-t---	---C...G---	---CC---C---C---C---	1529
C.BW.96BW1210	---c-c-t---	---g-g-g---	---g-a-t-t---	---C...G---	---CC---C---C---C---	1519
C.ET.FTH2220	---c-c-c---	---g-g-g---	---g-a-t-t---	---A...G---	---C---C---C---C---	1512
C.IN.301999	---c-c-c---	---g-g-g---	---a-t---aa-t---	---C...G---	---C---C---C---C---	1511
C.IN.210688	---c-c-c---	---g-g-g---	---a-t---aa-t---	---C...G---	---C---C---C---C---	1511
C.IN.301905	---c-c-c---	---g-g-g---	---a-t---aa-t---	---C...G---	---C---C---C---C---	2133
C.IN.301904	---c-c-c---	---g-g-g---	---a-t---aa-t---	---C...G---	---C---C---C---C---	1502
C.IN.93IN101	---c-c-c---	---g-g-g---	---a-t---aa-t---	---C...G---	---C---C---C---C---CCC---AGC---	1686
C.IN.11246	---c-c-c---	---g-g-g---	---a-t---aa-t---	---C...G---	---G...G...C---C---A---	1697
D.UG.94UG114	---c-c-c---	---g-g-g---	---a-t---aa-t---	---C...G---	---G...G...C---C---A---	2154
D.CD.NDK	---c-c-c---	---g-g-g---	---g-a-t-t---	---C...G---	---G...G...C---C---A---	1668
D.CD.ELI	---c-c-c---	---g-g-g---	---g-a-t-t---	---C...G---	---G...G...C---C---A---	1477
D.KE.MB2059	---c-c-g---	---g-g-g---	---a-t---aa-t---	---C...G---	---C---A---C---CA---A---G---	1485
D.CD.2226	---c-c-g---	---g-g-g---	---a-t---aa-t---	---C...G---	---C---A---C---CA---A---G---	1344
D.CD.842R085	---g-c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---A...G---	---C---A---C---CA---A---G---	1344
F1.BR.93BR020	---c-c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---A...G---	---C---A---C---CA---A---G---	1344
F1.FI.FIN9363	---c-c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---A...G---	---C---A---C---CA---A---G---	1344
F1.BE.VI850	---c-c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---A...G---	---C---A---C---CA---A---G---	1344
F1.FR.MF411	---c-c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---A...G---	---C---A---C---CA---A---G---	1344
F2.CM.MF255C	---c-c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---A...G---	---C---A---C---CA---A---G---	1356
F2.CM.MF257C	---c-c-c-c---	---a-t---aa-t---	---a-t---aa-t---	---A...G---	---C---A---C---CA---A---G---	1356

	Gag-p7 ribosomal slip site	Gag p7 nucleocapsid end // Gag p1 start	CTACAA...GGGAAG...CCAGGGAATTTCTTCAGACGAGACC...A	Gag p1 end // Gag p6 start	
B.FR.HXB2R	CAAATGAAGAATGTACT...	GAAGACAG...GCTAAATT.TTTAGGAGATCTGGCCTTC...	CTACAA...GGGAAG...CCAGGGAATTTCTTCAGACGAGACC...A		2148
G.BE.DRCBL	A-C-A.....A	A-T.....A	A.....G	A-C-C-C-A-G	2115
G.NG.92NG083	A-C-G.....A	A-T.....A	A.....G	A-C-C-C-A-G	1520
G.SE.SE6165	C-C-G.....A	A-T.....A	A.....G	A-C-C-C-A-G	1554
G.FI.HH87932	C-C-G.....A	A-T.....A	A.....G	A-C-C-C-A-G	1552
H.BE.VI991	C-C-A.....A	A-T.....A	A.....G	A-C-C-C-A-G	1543
H.BE.VI997	G-C-C-A.....A	A-T.....A	A.....G	A-C-C-C-A-G	1478
H.CF.90CF056	G-C-C-A.....A	A-T.....A	A.....G	A-C-C-C-A-G	1498
J.SE.SE91733	C-C-C.....A	A-T.....A	A.....G	A-C-C-C-A-G	1469
K.CM.MP8309	C-C-C.....A	A-T.....A	A.....G	A-C-C-C-A-G	1468
K.CD.EQPBL1C	C-C-C.....A	A-T.....A	A.....G	A-C-C-C-A-G	1344
K.CM.MP835C	C-C-C.....A	A-T.....A	A.....G	A-C-C-C-A-G	1344
O.CM.AN770	C-C-GAAT...GA-A	C-ATA.....CT	A-AG	A-C-C-A-G	2115
O.SN.MVFS180	C-AAAAAT...GA	-ATA.....C	GGGGGG...CACG	A-C-C-A-G	2115
O.SN.MP1299	C-AA...AAT-GA	-ATA.....C	GGGGGG...CACG	A-C-C-A-G	2190
O.SN.MP1300	C-GA...AAT-GA	-ATA.....C	GGGGGG...CACG	A-C-C-A-G	2114
N.CM.YBF30	AAAAATGAA-GA	-G-T-C.....T	A.....G	A-C-C-A-G	2214
N.CM.YBF106	AAAAATGAG-GA-M	-A-GG-T-C.....T	A.....G	A-C-C-A-G	1725
CONSENSUS 01	C-C-C.....g	aaa-t-tc.....tc	Aa-g-a	g-g-t-c-g-c	2100
01.AE.TH.93TH253	C-C-C.....G	A-T.....A	A.....G	A-C-C-A-G	1707
01.AE.TH.CM240	C-C-C.....A	A-T.....A	A.....G	A-C-C-A-G	1707
01.AE.CF.90CF402	C-C-C.....-A	A-T.....A	A.....G	A-C-C-A-G	2184
01.AE.TH.92TH022	C-C-C.....C	A-T.....A	A.....G	A-C-C-A-G	2111
01.AE.TH.95TH047	C-C-C.....C	A-T.....A	A.....G	A-C-C-A-G	2111
01.AE.CF.11697	C-C-C.....C	A-T.....A	A.....G	A-C-C-A-G	2097
01.AE.TH.93TH057	C-C-C.....C	A-T.....A	A.....G	A-C-C-A-G	2101
01.AE.TH.93TH065	C-C-C.....C	A-T.....A	A.....G	A-C-C-A-G	2100
01.AE.CF.4071	C-C-C.....ga	t.....t	AG	A-C-C-A-G	2085
CONSENSUS 02	-a-c-t.....t	t.....t	ag-g-g	a-t-c-g	1361
02.AG.SE.7812	C-C-C.....G	A-T.....A	A.....G	A-C-C-A-G	1520
02.AG.GH.G829	C-C-C.....G	A-T.....A	A.....G	A-C-C-A-G	1364
02.AG.DJ.DJ263	C-C-C.....G	A-T.....A	A.....G	A-C-C-A-G	1496
02.AG.DJ.DJ264	C-C-C.....G	A-T.....A	A.....G	A-C-C-A-G	1497
02.AG.NG.IBNG	C-C-C.....G	A-T.....A	A.....G	A-C-C-A-G	1673
02.AG.SN.MP1211	C-C-C.....-A	A-T.....A	A.....G	A-C-C-A-G	1345
02.AG.SN.MP1213	C-C-C.....G	A-T.....A	A.....G	A-C-C-A-G	1342
02.AG.CM.ME807	-G-C-C-C.....G	A-T.....A	A.....G	A-C-C-A-G	1304
03.AB.RU.98001	C-C-A.....C	A-T.....A	A.....G	A-C-C-A-G	1518
03.AB.RU.KALI53-2	C-C-C.....A	A-T.....A	A.....G	A-C-C-A-G	1917
04.CPX.GR.97PVMY	C-C-C.....C	A-T.....A	A.....G	A-C-C-A-G	2165
04.CPX.GR.97PVGH	C-C-C.....C	A-T.....A	A.....G	A-C-C-A-G	1517
04.CPX.CY.94CY0323	C-C-C.....G	A-T.....A	A.....G	A-C-C-A-G	1530
05.DF.BE.VII1310	C-C-C.....G	A-T.....A	A.....G	A-C-C-A-G	1503
05.DF.BE.VI961	C-C-C.....G	A-T.....A	A.....G	A-C-C-A-G	1388
06.CP.ML.95ML84	C-C-A.....A	A-T.....A	A.....G	A-C-C-A-G	2170
06.CP.AU.BFP90	C-C-A.....A	A-T.....A	A.....G	A-C-C-A-G	2168
06.CP.ML.95ML127	C-C-C.....A	A-T.....A	A.....G	A-C-C-A-G	2195
06.CP.SN.97SEL078	C-C-C.....A	A-T.....A	A.....G	A-C-C-A-G	1530
10.CD.BFL061	C-C-A.....GAAAGA	A-T.....A	A.....A	A-C-C-A-G	1525
10.CD.BFL110	C-C.....C	A-T.....A	A.....A	A-C-C-A-G	1525
10.CD.BFL071	C-C.....A	A-T.....A	A.....A	A-C-C-A-G	2148
11.CP.CM.MP818	C-C-C.....C	A-T.....A	A.....A	A-C-C-A-G	2145
11.CP.FR.MP1298	C-C-C.....C	A-T.....A	A.....A	A-C-C-A-G	2154
11.CP.FR.MP1307	C-C-C.....C	A-T.....A	A.....A	A-C-C-A-G	1453
11.CP.GR.GR17	C-C-C.....C	A-T.....A	A.....A	A-C-C-A-G	2215
CPZ.GA.SIVCPZGAB	C-C-A.....GA	AGGT.....TG	CG-G...C-G	C-C-G-A-A	214
CPZ.US.SIVCPZUS	C-C-AGCAGGAA-C	AGGT-T-CC	A-CGTGAGC-GG-ATCAAAGAGA	C-C-C-G-A-A	214
CPZ.CM.SIVCPZCM3	C-C-AGCTGGA-A	AGGT-T-C	G-GG-C-C-GG-ATCAAAGAGA	C-C-C-G-G-GAA	1553
CPZ.CM.SIVCPZCM5	A-C-NACTGGA-GC	GAT-T-C	G-GG-C-C-GG-ATCAAAGAGA	C-C-C-G-G-GAA	1856
CPZ.CD.SIVCPZANT	--GC-T--A-C--C-AGCAACAA-T-C-GGAATA-TA	-----ACCGACC-CA	G-GGTC...GT-CAG	-----C-G-G-G-AGGAGNAGTAGT...G	1996
B.FR.HXB2R	CAAATGAAGAATGTACT...	GAAGACAG...GCTAAATT.TTTAGGAGATCTGGCCTTC...	CTACAA...GGGAAG...CCAGGGAATTTCTTCAGACGAGACC...A		2148
B.FR.HXB2R	---Q_M_K_D_C_T_E_R_Q---A_N_F_L_L_G_K_I_W_P_S_Y_K_G_R---P_G_N_F_L_O_S_R_P---	---Gag p7 nucleocapsid end // Gag p1 start		---Gag p1 end // Gag p6 start	---Gag

B. FR. HXB2R	AGAGAGACTTCAGTCTGGGTAGACACACAC	AGAGAGACTTCAGTCTGGGTAGACACAC	TCGCCCTCAGAAGCAGGA	GCCTGAT	2225
G. BE. DRCEI	C-C-A-G-TC-AG-T-G	C-C-A-G-TC-AG-T-G	C-T-C-C-	-A-A-	2189
G. NG. 92NG083	T-C-A-C-G-TC-A-AG-T-G	T-C-A-C-G-TC-A-AG-T-G	C-T-C-C-	-A-A-	1594
G. SE. SE6165	C-C-A-C-G-TC-A-AG-T-G	C-C-A-C-G-TC-A-AG-T-G	C-T-C-C-	-A-A-	1628
G. FI. HH87932	C-C-A-C-G-TC-A-AG-T-G	C-C-A-C-G-TC-A-AG-T-G	C-T-C-C-	-A-A-	1626
H. BE. VI991	C-C-A-C-G-TC-A-AG-T-G	C-C-A-C-G-TC-A-AG-T-G	C-T-C-C-	-A-A-	1617
H. BE. VI997	C-C-A-C-G-TC-A-AG-T-G	C-C-A-C-G-TC-A-AG-T-G	C-T-C-C-	-A-A-	1552
H. CF. 90CF056	C-C-A-C-G-TC-A-AG-T-G	C-C-A-C-G-TC-A-AG-T-G	C-T-C-C-	-A-A-	1575
J. SE. SE91733	C-C-A-C-G-TC-A-AG-T-G	C-C-A-C-G-TC-A-AG-T-G	C-T-C-C-	-A-A-	1540
J. SE. SE92809	C-C-A-C-G-TC-A-AG-T-G	C-C-A-C-G-TC-A-AG-T-G	C-T-C-C-	-A-A-	1539
K. CM. MP535C	C-C-A-C-G-TC-A-AG-T-G	C-C-A-C-G-TC-A-AG-T-G	C-T-C-C-	-A-A-	1539
K. CD. E0TB11C	C-C-A-C-G-TC-A-AG-T-G	C-C-A-C-G-TC-A-AG-T-G	C-T-C-C-	-A-A-	1418
O. CM. AN170	GATG-GAGGAAGT-AAG-ACAG-A	GATG-GAGGAAGT-AAG-ACAG-A	C-T-T-GA	AAT-A	2280
O. CM. MP5180	ATG-GAGGAAGT-AAG-ACAG-A	ATG-GAGGAAGT-AAG-ACAG-A	---GT-A	AGG-GG	2280
O. SN. MP1299	GATGAC-GAGGAA-T-AAG-ACA	GATGAC-GAGGAA-T-AAG-ACA	---GT-A	GG-GA	2295
O. SN. MP1300	GATGAC-GAGGAA-T-AAG-ACA	GATGAC-GAGGAA-T-AAG-ACA	---GT-A	GG-GA	2279
N. CM. YBF30	CT-T-ATG-T-CA-AG-AG-GC-ACAGGGAAGGATG-AGG-C	CT-T-ATG-T-CA-AG-AG-GC-ACAGGGAAGGATG-AGG-C	GG-AC-A-G	GA	2279
N. CM. YBF106	CT-T-ATG-T-CA-AG-AG-GC-ACAGGGAAGGATG-AGG-C	CT-T-ATG-T-CA-AG-AG-GC-ACAGGGAAGGATG-AGG-C	GG-AC-A-G	AG-C	1816
CONSENSUS 01	gca-?aac-9gg-ATG-Aagaataaac????????ctctccgaa	gca-?aac-9gg-ATG-Aagaataaac????????ctctccgaa	---A-A-	AA-C	1814
01 AE. TH. CM240	C-A-A-GGG-ATG-A	C-A-A-GGG-ATG-A	---A-A-	-A-A	2173
01 AE. CF. 90CF402	C-A-A-GGG-ATG-A	C-A-A-GGG-ATG-A	---A-A-	-A-A	2209
01 AE. TH. CM240	ATG-A-GGG-ATG-A	ATG-A-GGG-ATG-A	CT-TTC-C	-A-A	1796
01 AE. TH. 92TH022	C-A-A-GGG-ATG-A	C-A-A-GGG-ATG-A	CT-TTC-C	-A-A	2258
01 AE. TH. 95TNIH047	C-A-A-GGG-ATG-A	C-A-A-GGG-ATG-A	CT-TTC-T	-A-A	2185
01 AE. CF. 11697	C-A-A-GGG-ATG-A	C-A-A-GGG-ATG-A	CT-TTC-T	-A-A	2185
01 AE. TH. 93TH057	C-A-A-GGG-ATG-A	C-A-A-GGG-ATG-A	CT-TTC-T	-A-A	2171
01 AE. TH. 93TH065	C-A-A-GGG-ATG-A	C-A-A-GGG-ATG-A	CT-TTC-C	-A-A	2175
01 AE. CF. 4071	C-A-A-GGG-ATG-A	C-A-A-GGG-ATG-A	CT-TTC-T	-A-A	2171
CONSENSUS 02	C-A-A-GGG-ATG-A	C-A-A-GGG-ATG-A	CT-TTC-C	-A-A	2159
02 AG. SE. 7812	a-c-gct-tg-Atgg-A-g-t-a	a-c-gct-tg-Atgg-A-g-t-a	ct-ct-t-cga	Ac-ag	1435
02 AG. GH. G829	G-C-TG-ATA-A	G-C-TG-ATA-A	CT-T-C	-A-GG	1594
02 AG. DJ. DJ263	G-C-TG-ATA-A	G-C-TG-ATA-A	CT-T-C	-A-GG	1438
02 AG. DJ. DJ264	G-C-TG-ATA-A	G-C-TG-ATA-A	CT-T-C	-A-GG	1570
02 AG. NG. IENG	G-C-TG-ATA-A	G-C-TG-ATA-A	CT-T-C	-A-GG	1571
02 AG. SN. MP1211	G-C-TG-ATA-A	G-C-TG-ATA-A	CT-T-C	-A-GG	1747
02 AG. SN. MP1213	G-C-TG-ATA-A	G-C-TG-ATA-A	CT-T-C	-A-GG	1419
02 AG. CM. M807	G-C-TG-ATA-A	G-C-TG-ATA-A	CT-T-C	-A-GG	1416
03 AB. RU. 98001	C-A-A-TG-ATG-A	C-A-A-TG-ATG-A	CT-TG-C	-A-A	1378
03 AB. RU. KALI53-2	C-A-A-TG-ATG-A	C-A-A-TG-ATG-A	CT-TG-C	-A-A	1378
04 CPX. GR. 97PVMY	C-C-A-A-TG-ATG-A	C-C-A-A-TG-ATG-A	C-T-C-T-A	-A-A-A	1992
04 CPX. GR. 97PVCH	C-C-A-A-TG-ATG-A	C-C-A-A-TG-ATG-A	C-T-C-T-A	-A-A-A	1991
04 CPX. CY. 94CY0323	C-C-A-A-TG-ATG-A	C-C-A-A-TG-ATG-A	C-T-C-T-A	-A-A-A	2239
05 DF. BE. VI1310	C-C-A-A-TG-ATG-A	C-C-A-A-TG-ATG-A	C-T-C-T-A	-A-A-A	2342
05 DF. BE. VI961	C-C-A-A-TG-ATG-A	C-C-A-A-TG-ATG-A	C-T-C-T-A	-A-A-A	1591
06 CPX. ML. 95ML84	C-C-A-A-TG-ATG-A	C-C-A-A-TG-ATG-A	C-T-C-T-A	-A-A-A	1604
06 CPX. AU. BPP90	C-C-A-A-TG-ATG-A	C-C-A-A-TG-ATG-A	C-T-C-T-A	-A-A-A	1577
06 CPX. ML. 95ML127	C-C-A-A-TG-ATG-A	C-C-A-A-TG-ATG-A	C-T-C-T-A	-A-A-A	1462
06 CPX. SN. 97SE1078	C-C-A-A-TG-ATG-A	C-C-A-A-TG-ATG-A	C-T-C-T-A	-A-A-A	2244
10 CD. BEL061	ACCTCAGCCCACT-C	ACCTCAGCCCACT-C	C-T-C-C-G-A	-A-A	2242
10 CD. BEL110	C-T-A-A	C-T-A-A	C-T-C-C	-A-A	2284
10 CD. BEL1071	C-T-A-A	C-T-A-A	C-T-C-C	-A-A	1601
11 CPX. CM. MP818	C-T-A-A	C-T-A-A	C-T-C-C	-A-A	1990
11 CPX. FR. MP1298	C-T-A-A	C-T-A-A	C-T-C-C	-A-A	1599
11 CPX. FR. MP1307	C-T-A-A	C-T-A-A	C-T-C-C	-A-A	2227
11 CPX. GR. GR17	C-T-A-A	C-T-A-A	C-T-C-C	-A-A	2219
CPZ. GA. STVCPZGAB	C-T-A-A	C-T-A-A	C-T-C-C	-A-A	2233
CPZ. US. STVCPZUS	C-T-A-A	C-T-A-A	C-T-C-C	-A-A	1527
CPZ. CM. STVCPZCAM3	C-T-A-A	C-T-A-A	C-T-C-C	-A-A	2285
CPZ. CM. STVCPZCAM5	C-T-A-A	C-T-A-A	C-T-C-C	-A-A	2285
CPZ. CD. STVCPZANT	C-T-A-A	C-T-A-A	C-T-C-C	-A-A	2285
Gag-Pol TF	R-A-N-S-P-T	R-R-E-I-O-V-W-G-R-D-N-N	S-P-S-E-A-G	A-D	
E. FR. HXB2R	GAGCAACAGCCACC	AGAGAGACTTCAGTCTGGGTAGACACAC	TCCCTCAGAAGCAGGA	GCCTGAT	2225
Gag P6	E-P-T-A-P-P	E-E-S-F-R-S-G-V-E-T-T-T	P-P-O-K-O-E	P-I	

HIV-1/SIVcpz complete genomes

G_BE.DRCBL 2536
G_NG.92NG083 1941
G_SE.SE6165 1869
G_FI.HH87932 1973
H_BE.VI991 1961
H_BE.VI997 1896
H_CF.90CF056 1919
J_SE.SE91733 1887
J_SE.SE92809 1886
K_CM.MP535C 1768
K_CD.EOTB11C 1768
O_CD.AN170 2627
O_CD.MP5180 2602
O_SN.MP1299 2626
O_SN.MP1300 2626
N_CM.YBF30 2166
N_CM.YBF106 2164
CONSENSUS 01 2523
01_AE.TH.93TH253 2559
01_AE.TH.CM240 2146
01_AE.CF.90CF402 2608
01_AE.TH.92TH022 2535
01_AE.TH.95TNIH047 2535
01_AE.CF.11697 2521
01_AE.TH.93TH057 2525
01_AE.TH.93TH065 2517
01_AE.CF.4071 2509
CONSENSUS 02 1784
02_AG.SE.7812 1944
02_AG.GH.G829 1788
02_AG.DJ.DJ263 1920
02_AG.DJ.DJ264 1921
02_AG.NG.IBNG 2097
02_AG.SN.MP1211 1769
02_AG.SN.MP1213 1766
02_AG.CM.MP807 1728
03_AB.RU.98001 1941
03_AB.RU.K81153-2 2341
04_CFX.GR.97FVNY 2386
04_CFX.GR.97FVCH 2389
04_CFX.CY.94CY0323 1938
05_DF.BE.VII1310 1954
06_CFX.ML.95ML84 1927
06_CFX.AU.BFP90 1815
06_CFX.ML.95ML127 2600
06_CFX.SN.97SE1078 2589
10_CD.BEL061 2631
10_CD.BEL110 1951
10_CD.BEL1071 1949
11_CFX.CM.ME818 2575
11_CFX.FR.MP1298 2566
11_CFX.FR.MP1307 2623
11_CFX.GR.GR17 1874
CPZ_GA.SIVCPZGAB 2633
CPZ_US.SIVCPZUS 2632
CPZ_CM.SIVCPZCAM3 1983
CPZ_CD.SIVCPZCAM5 2274
CPZ_CD.SIVCPZANT 2011

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_Q_I_G_C_T_L_N_F_P_I_S_P_I_E_T_V_Pol
Protease end / Pol p66 and p51 RT start

HIV-1/SIVcpz complete genomes

B. FR. HXB2R	ACCAGTAAAATTAAGCCAGGATGGATGCCCGCCAAAGATTAAACAATGGCCATTGACAGAGAAAGAAAATAAAAGCATTAGTAGAAATTTTGTACA.GGGGATATCCCCAAAG
G. BE. DRGB1
G. NG. 92NG083
G. SE. SE6165
G. FI. HH87932
H. BE. VI991
H. BE. VI997
H. CF. 90CF056
J. SE. SE91733
J. SE. SE92809
K. CM. MP535C
K. CD. F0TB11C
O. CM. ANT70
O. CM. MP5180
O. SN. MP1299
O. SN. MP1300
N. CM. YBF30
N. CM. YBF106
CONSENSUS 01
01 AE. TH. 93TH253
01 AE. TH. CM240
01 AE. CF. 90CF402
01 AE. TH. 92TH022
01 AE. TH. 95TNH047
01 AE. CF. 11697
01 AE. TH. 93TH057
01 AE. TH. 93TH065
01 AE. CF. 4071
CONSENSUS 02
02 AG. SE. 7812
02 AG. GH. G829
02 AG. DJ. DJ263
02 AG. DJ. DJ264
02 AG. NG. IBNG
02 AG. SN. MP1211
02 AG. SN. MP1213
02 AG. CM. M807
03 AB. RU. 98001
03 AB. RU. KALI53-2
04 CPX. GR. 97FVWY
04 CPX. GR. 97FVCH
04 CPX. CY. 94CY0323
05 DF. BE. VI1310
05 DF. BE. VI961
06 CPX. ML. 95ML84
06 CPX. AU. BPP90
06 CPX. ML. 95ML127
06 CPX. SN. 97SE1078
10 CD. BEL061
10 CD. BEL110
10 CD. BEL071
11 CPX. CM. M818
11 CPX. FR. MP1298
11 CPX. FR. MP1307
11 CPX. GR. GR17
CPZ. GA. STVCPZGAB
CPZ. US. STVCPZUS
CPZ. CM. STVCPZCAM3
CPZ. CM. STVCPZCAM5
CPZ. CD. STVCPZANT

P_V_K_L_K_P_G_M_D_G_P_K_V_K_Q_W_P_L_T_E_E_K_I_K_A_L_V_E_I_C_T_T...E_M_E_K_E_G_K_I_S_K

HIV-1/SIVcpz complete genomes

B. FR. HXB2R	ATTGGCCCTGAAATCCATACTCCAGTATTTGCCATAAAGAAAAGACAGACTACTAATAAAGGAGAACTTAAATCAAGAACTTCAAGACTTCCTGGGAAAGTTCAATTAG	2826
CONSENSUS A	-----t-a-c-----t-----t-----t-----g-a-ag-c-cac--aa-----a-t-----t-----t-----c-----g-c-----a-----t-a-c-t-----gttcaattag	2842
A.UG.0455	-----t-----t-----t-----c-----c-----c-----g-----c-----a-----a-----a-----g-----g-----g-----c-----a-----a-----a-----	2872
A.SE.UGSE8891	-----c-----c-----c-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2018
A.SE.UGSE8891	-----c-----c-----c-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2025
A.SE.TZSE8538	-----g-----t-----t-----t-----a-----g-----a-----c-----g-----c-----c-----g-----c-----a-----a-----a-----g-----c-----a-----g-----c-----	2010
A.SE.UGSE6594	-----a-----t-----t-----t-----a-----g-----a-----c-----g-----c-----c-----g-----c-----a-----a-----a-----g-----c-----a-----g-----c-----	2284
A.KE.Q2317	-----c-----c-----c-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2043
A.SE.UGSE7535	-----g-----t-----t-----t-----a-----g-----a-----c-----g-----c-----c-----g-----c-----a-----a-----a-----g-----c-----a-----g-----c-----	2030
A.SE.SOSE7253	-----a-----t-----t-----t-----a-----g-----a-----c-----g-----c-----c-----g-----c-----a-----a-----a-----g-----c-----a-----g-----c-----	2193
A.UG.92UG037	-----a-----t-----t-----t-----a-----g-----a-----c-----g-----c-----c-----g-----c-----a-----a-----a-----g-----c-----a-----g-----c-----	2198
A.SE.SE8131	-----c-----c-----c-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2164
A2.CD.97CDKTB48	-----a-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2183
A2.CY.94CY017	-----a-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2177
CONSENSUS B	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2837
B.AU.MBC925	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2825
B.AU.MBC200	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2369
B.GA.OYI	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2772
B.GB.MANC	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2828
B.GB.CAMI	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2213
B.CN.RL42	-----c-----c-----c-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2372
B.DE.D31	-----c-----c-----c-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2292
B.DE.HAN2	-----c-----c-----c-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2384
B.ES.89SP061	-----c-----c-----c-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2828
B.NL.ACH320A	-----c-----c-----c-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2825
B.TM.LM49	-----c-----c-----c-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2817
B.US.DH123	-----c-----c-----c-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2833
B.US.SF2	-----c-----c-----c-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2826
B.US.NY5CG	-----c-----c-----c-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2372
B.US.WCIPR18	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2824
B.US.YU2	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2143
B.US.JRC5F	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2841
B.US.JRFL	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2373
B.US.MN	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2825
B.US.BCSG3	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2343
B.US.WEAU160	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2905
B.US.RF	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2190
B.US.WR27	-----g-----t-----t-----t-----g-----t-----t-----t-----c-----a-----a-----ag-ga-----a-----c-----c-----tt-at-----	2164
CONSENSUS C	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2292
C.BR.92BR025	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2186
C.BW.96BW0402	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2293
C.BW.96BW1104	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2292
C.BW.96BW15C02	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2319
C.BW.96BW01B03	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2292
C.BW.96BW0502	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2198
C.BW.96BW16B01	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2198
C.BW.96BW1210	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2205
C.ET.ETH2220	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2177
C.IN.301999	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2177
C.IN.21068	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2177
C.IN.301905	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2177
C.IN.301904	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2177
C.IN.93JN101	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2168
C.IN.11246	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2181
D.UG.94UG114	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2361
D.CD.NDK	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2372
D.CD.ELI	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2828
D.KE.MB2059	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2829
D.CD.Z226	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2829
D.CD.84ZR085	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2348
F1.BR.93BR020	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2154
F1.FI.FIN9363	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2155
F1.BE.VI850	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2163
F1.FR.MF411	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2028
F2.CM.ME255C	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2022
F2.CM.ME257C	-----t-----t-----t-----c-----g-----g-----c-----a-----g-----g-----c-----a-----a-----g-----c-----a-----g-----c-----	2034

B. FR. HXB2R	ATTGGGCTGAAATCCATCACTCCAGTATTTGCCATAAGAAAAAGACGACTACTTAATGGAGAAATTAGTAGATTTTCAGAGAACTTAATAGAGAACTCAAGACTTCTGGGAGTTCAATTAG	2826
G. BE. DRGB1	-----T-C-----A-----C-----G-----G-----G-----C-----A-----G-----G-----C-----C-----	2190
G. NG. 92NG083	-----C-----T-----C-----A-----C-----G-----G-----G-----C-----A-----A-----G-----G-----C-----C-----	2195
G. SE. SE6165	-----C-----A-----A-----C-----G-----G-----G-----C-----A-----A-----G-----G-----C-----C-----	2203
G. FI. HH87932	-----A-----G-----G-----G-----G-----G-----C-----A-----A-----G-----G-----C-----C-----	2207
H. BE. VI991	-----A-----C-----G-----C-----A-----A-----G-----GA-----G-----G-----C-----A-----A-----A-----G-----	2215
H. BE. VI997	-----A-----G-----C-----A-----A-----G-----G-----G-----C-----A-----A-----G-----G-----C-----C-----	2150
H. CF. 90CF056	-----A-----G-----GC-----A-----A-----G-----G-----T-----G-----C-----A-----A-----G-----G-----	2173
J. SE. SE91733	-----G-----T-----C-----C-----C-----A-----A-----G-----G-----T-----T-----C-----A-----A-----G-----G-----	2141
J. SE. SE92809	-----G-----T-----C-----T-----C-----G-----G-----T-----T-----C-----A-----A-----G-----G-----	2140
K. CM. MP535C	-----G-----T-----C-----T-----C-----G-----G-----T-----T-----C-----A-----A-----G-----G-----	2022
K. CD. F0TB11C	-----A-----A-----A-----A-----TA-----C-----T-----A-----G-----TG-----G-----G-----T-----A-----A-----G-----G-----	2022
O. CM. AN170	-----A-----A-----T-----T-----A-----TA-----C-----T-----A-----G-----T-----G-----C-----T-----A-----A-----A-----G-----	2881
O. CM. MP5180	-----A-----A-----T-----T-----A-----TA-----C-----T-----A-----G-----T-----C-----G-----C-----T-----A-----A-----T-----G-----	2856
O. SN. MP1299	-----A-----A-----T-----T-----A-----TA-----C-----T-----A-----G-----T-----C-----G-----C-----T-----G-----T-----A-----C-----	2880
O. SN. MP1300	-----A-----A-----T-----T-----A-----TA-----C-----T-----A-----G-----T-----C-----G-----C-----T-----G-----T-----A-----C-----	2880
N. CM. YBF30	-----A-----G-----T-----C-----A-----T-----T-----A-----G-----T-----C-----T-----C-----T-----G-----T-----A-----A-----G-----C-----	2418
N. CM. YBF106	-----A-----G-----T-----C-----A-----T-----T-----A-----G-----T-----C-----T-----C-----T-----G-----T-----A-----A-----G-----C-----	2418
CONSENSUS 01	g-----a-----g-----t-----a-----a-----gg-----c-----c-----a-----g-----a-----at-----g-----c-----Aa-----g-----T-----	2777
01 AE. TH. 93TH253	-----A-----A-----T-----T-----A-----A-----g-----g-----c-----c-----c-----c-----A-----A-----G-----T-----	2813
01 AE. TH. CM240	-----A-----A-----T-----T-----A-----A-----g-----g-----c-----c-----c-----c-----A-----A-----G-----T-----	2400
01 AE. CF. 90CF402	-----A-----A-----T-----T-----A-----A-----g-----g-----c-----c-----c-----c-----A-----A-----G-----T-----	2862
01 AE. TH. 92TH022	-----A-----A-----T-----T-----A-----A-----g-----g-----c-----c-----c-----c-----A-----A-----G-----T-----	2789
01 AE. TH. 95TNIH047	-----A-----A-----T-----T-----A-----A-----g-----g-----c-----c-----c-----c-----A-----A-----G-----T-----	2789
01 AE. CF. 11697	-----A-----A-----T-----T-----A-----A-----g-----g-----c-----c-----c-----c-----A-----A-----G-----T-----	2785
01 AE. TH. 93TH057	-----A-----A-----T-----T-----A-----A-----g-----g-----c-----c-----c-----c-----A-----A-----G-----T-----	2779
01 AE. TH. 93TH065	-----A-----A-----T-----T-----A-----A-----g-----g-----c-----c-----c-----c-----A-----A-----G-----T-----	2779
01 AE. CF. 4071	-----A-----A-----T-----T-----A-----A-----g-----g-----c-----c-----c-----c-----A-----A-----G-----T-----	2771
CONSENSUS 02	-----g-----g-----a-----cc-----aa-----aag-----T-----t-----a-----a-----t-----t-----a-----c-----g-----ac-----t-----	2038
02 AG. SE. 7812	-----A-----A-----T-----T-----A-----A-----g-----g-----a-----cc-----aa-----aag-----T-----t-----a-----a-----t-----t-----	2198
02 AG. GH. G829	-----A-----A-----T-----T-----A-----A-----g-----g-----a-----cc-----aa-----aag-----T-----t-----a-----a-----t-----t-----	2042
02 AG. DJ. DJ263	-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----C-----C-----C-----C-----	2174
02 AG. DJ. DJ264	-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----C-----C-----C-----C-----	2175
02 AG. NG. I8NG	-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----C-----C-----C-----C-----	2351
02 AG. SN. MP1211	-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----C-----C-----C-----C-----	2023
02 AG. SN. MP1213	-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----C-----C-----C-----C-----	2020
02 AG. CM. M5807	-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----C-----C-----C-----C-----	1982
03 AB. RU. 98001	-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----C-----C-----C-----C-----	1986
03 AB. RU. KALI53-2	-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----C-----C-----C-----C-----	2595
04 CPX. GR. 97FVWY	-----A-----A-----T-----T-----A-----A-----GA-----C-----A-----G-----G-----C-----A-----A-----G-----G-----	2840
04 CPX. GR. 97FVCH	-----A-----A-----T-----T-----A-----A-----GA-----C-----A-----G-----G-----C-----A-----A-----G-----G-----	2843
04 CPX. CY. 94CY0323	-----A-----A-----T-----T-----A-----A-----GA-----C-----A-----G-----G-----C-----A-----A-----G-----G-----	2192
05 DF. BE. VI1310	-----C-----C-----A-----A-----T-----T-----A-----A-----GA-----C-----A-----G-----G-----C-----A-----A-----G-----G-----	2208
05 DF. BE. VI961	-----C-----C-----A-----A-----T-----T-----A-----A-----GA-----C-----A-----G-----G-----C-----A-----A-----G-----G-----	2181
06 CPX. ML. 95ML84	-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	2069
06 CPX. AU. BFP90	-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	2854
06 CPX. ML. 95ML127	-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	2843
06 CPX. SN. 97SE1078	-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	2885
10 CD. BEL101	-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	2205
10 CD. BEL110	-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	2194
10 CD. BEL1071	-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	2203
11 CPX. CM. M5818	-----T-----C-----C-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----C-----A-----G-----	2829
11 CPX. FR. MP1298	-----T-----C-----C-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----C-----A-----G-----	2820
11 CPX. FR. MP1307	-----C-----C-----T-----C-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----C-----A-----G-----	2877
11 CPX. GR. GR17	-----C-----C-----T-----C-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----C-----A-----G-----	2128
CPZ. GA. STVCPZGAB	-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	2887
CPZ. US. STVCPZUS	-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	2886
CPZ. CM. STVCPZCAM3	-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	2237
CPZ. CM. STVCPZCAM5	-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	2528
CPZ. CD. STVCPZANT	-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	2565
Pol p51 RT	-----I-----G-----P-----E-----N-----P-----Y-----N-----T-----P-----V-----F-----A-----I-----K-----K-----K-----D-----S-----T-----K-----K-----K-----D-----S-----T-----K-----K-----K-----D-----S-----T-----Q-----D-----F-----W-----E-----V-----Q-----L-----	Pol

B. FR. HXB2R	GAATACCACATCCC...GCAGGGTTAAAGAAAGAAAATAATCAGTAAACAGTACTGATGATGGGTCATGATCCATATTTTTCAGTTCCCTTAGATGAGACATTTAGAGGAACTACTGCACTTTACCACTACCTAGTAT	2853
G. BE. DRCEI	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2917
G. NG. 92NG083	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2917
G. SE. SE6165	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2922
G. FI. HH87932	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2930
H. BE. VI991	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2350
H. BE. VI997	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2354
H. CF. 90CF056	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2342
J. SE. SE91733	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2277
J. SE. SE92809	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2300
K. CM. MP535C	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2268
K. CD. E0TB11C	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2267
O. CM. AN170	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2149
O. CM. MP5180	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	3008
O. SN. MP1299	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2983
O. SN. MP1300	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	3007
N. CM. YBF30	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	3007
N. CM. YBF106	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2547
CONSENSUS 01	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2545
01 AE. TH. 93TH253	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2940
01 AE. TH. CM240	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2940
01 AE. CF. 90CF402	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2527
01 AE. TH. 92TH022	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2989
01 AE. TH. 95TNIH047	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2916
01 AE. CF. 11697	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2916
01 AE. TH. 93TH057	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2802
01 AE. TH. 93TH065	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2906
01 AE. CF. 4071	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2898
CONSENSUS 02	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2898
02 AG. SE. 7812	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2164
02 AG. GH. G829	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2325
02 AG. DJ. DJ263	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2169
02 AG. DJ. DJ264	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2301
02 AG. NG. I8NG	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2302
02 AG. SN. MP1211	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2478
02 AG. SN. MP1213	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2150
02 AG. CM. ME807	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2147
03 AB. RU. 98001	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2109
03 AB. RU. KALI53-2	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2323
04 CPX. GR. 97FVNY	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2722
04 CPX. GR. 97FVCH	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2867
04 CPX. CY. 94CY0323	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2970
05 DF. BE. V11310	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2319
05 DF. BE. V1961	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2335
06 CPX. ML. 95ML84	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2308
06 CPX. AU. BPP90	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2196
06 CPX. ML. 95ML127	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2981
06 CPX. SN. 97SE1078	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2970
10 CD. BEL061	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	3012
10 CD. BEL110	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2332
10 CD. BEL1071	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2321
11 CPX. CM. ME818	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2350
11 CPX. FR. MP1298	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2956
11 CPX. FR. MP1307	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2947
11 CPX. GR. GR17	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	3004
CPZ. GA. SIVCPZGAB	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2255
CPZ. US. SIVCPZUS	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	3014
CPZ. CM. SIVCPZCAM3	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	3013
CPZ. CM. SIVCPZCAM5	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2364
CPZ. CD. SIVCPZANT	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2655
Pol p51 RT	-----G-----G-----A-----G-----G-----A-----T-----A-----A-----C-----T-----C-----T-----	2392

HIV-1/SIVcpz complete genomes

B. FR. HXB2R
 CONSENSUS A
 A.UG.0455
 A.SE.UGSE8891
 A.SE.TZSE8538
 A.SE.UGSE6594
 A.KE.Q2317
 A.SE.UGSE7535
 A.SE.SOSE7253
 A.UG.92UG037
 A.SE.SE8131
 A2.CD.97CDKTB48
 A2.CY.94CY017
 CONSENSUS B
 B.AU.MBC925
 B.AU.MBC200
 B.GA.OYI
 B.GB.MANC
 B.GB.CAMI
 B.CN.RL42
 B.DE.D31
 B.DE.HAN2
 B.ES.89SP061
 B.NL.ACH320A
 B.TM.LM49
 B.US.DHL23
 B.US.SF2
 B.US.NY5CG
 B.US.WCIPR18
 B.US.YU2
 B.US.JRCSF
 B.US.JRFL
 B.US.MN
 B.US.BCSG3
 B.US.WEAU160
 B.US.RF
 B.US.WR27
 CONSENSUS C
 C.BR.92BR025
 C.BW.96BW0402
 C.BW.96BW1104
 C.BW.96BW15C02
 C.BW.96BW01B03
 C.BW.96BW0502
 C.BW.96BW16B01
 C.BW.96BW1210
 C.ET.ETH2220
 C.IN.301999
 C.IN.21068
 C.IN.301905
 C.IN.301904
 C.IN.93JIN101
 C.IN.11246
 D.UG.94UG114
 D.CD.NDK
 D.CD.ELI
 D.KE.MB2059
 D.CD.2226
 F1.BR.93BR020
 F1.FI.FIN9363
 F1.BE.VI850
 F1.FR.MF411
 F2.CM.MF255C
 F2.CM.MF257C

...CATCAGAAAGAACCTCCATTTGGATGGGTTATGAACCTCCATCGATAAATGGACAGTACAGCTGACTGTCATGATGCATACAGAACTTAGTGGGAA
 T-----T-----T-----C-g-----C-g-----C-g-----gag-----g-a-a-a-gc-----T-----A-----g-a
 A-----A-----A-----C-----C-----C-----CA-----A-----A-----A-----A-----A-----A-----A
 A.SE.UGSE8891 A-----A-----A-----C-----C-----C-----AAC-----A-----A-----A-----A-----A-----A
 A.SE.TZSE8538 A-----A-----A-----G-----G-----G-----C-----C-----C-----A-----T-----A-----A-----A
 A.SE.UGSE6594 A-----A-----A-----A-----G-----C-----C-----C-----C-----C-----C-----A-----T-----A-----A
 A.KE.Q2317 A-----A-----A-----A-----G-----G-----C-----C-----C-----C-----C-----C-----A-----T-----A
 A.SE.UGSE7535 A-----A-----A-----A-----G-----G-----C-----C-----C-----C-----C-----C-----A-----T-----A
 A.SE.SOSE7253 A-----A-----A-----A-----G-----G-----C-----C-----C-----C-----C-----C-----A-----T-----A
 A.UG.92UG037 A-----A-----A-----A-----G-----G-----C-----C-----C-----C-----C-----C-----A-----T-----A
 A.SE.SE8131 A-----A-----A-----A-----G-----G-----C-----C-----C-----C-----C-----C-----A-----T-----A
 A2.CD.97CDKTB48 A-----A-----A-----A-----G-----G-----C-----C-----C-----C-----C-----C-----A-----T-----A
 A2.CY.94CY017 A-----A-----A-----A-----G-----G-----C-----C-----C-----C-----C-----C-----A-----T-----A
 ?-----g-a-a-t-----C-----t-----a-c-t-c-t-a-a-a-g-----gtg-g-a-----a-c-c-----g-gt-----A
 T-----T-----T-----G-----G-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----A
 A-----G-----C-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G
 B.US.WR27 B-----U-----S-----R-----2-----7-----A-----A-----A-----A-----A-----A-----A-----A-----A
 B.US.WR27 B-----U-----S-----R-----2-----7-----A-----A-----A-----A-----A-----A-----A-----A-----A
 CONSENSUS C C-----C-----C-----T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.BR.92BR025 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.BW.96BW0402 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.BW.96BW1104 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.BW.96BW15C02 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.BW.96BW01B03 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.BW.96BW0502 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.BW.96BW16B01 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.BW.96BW1210 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.ET.ETH2220 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.IN.301999 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.IN.21068 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.IN.301905 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.IN.301904 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.IN.93JIN101 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 C.IN.11246 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 D.UG.94UG114 T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 D.CD.NDK C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 D.CD.ELI C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 D.KE.MB2059 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 D.CD.2226 C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 F1.BR.93BR020 F-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 F1.FI.FIN9363 F-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 F1.BE.VI850 F-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 F1.FR.MF411 F-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 F2.CM.MF255C F-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C
 F2.CM.MF257C F-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C

3337

2853

2783

2526

2521

2795

2554

2541

2704

2709

2675

2694

3388

3348

3336

2880

3283

3339

2724

2883

2803

2895

3339

3337

3328

3347

2883

3335

3349

2654

3352

2884

3336

2854

2716

2701

2675

2803

2697

2804

2803

2830

2803

2709

2729

2716

2688

2888

2888

3310

2679

2692

2872

2883

3341

3340

2859

2665

2666

2874

2539

2533

2545

```

B.FR.HXB2R      ...CATCAGAAAGAACCTCCATTTGGATGGGTATGAACTCCATCCCTGATAAATGGACAGCTACAGCCCTATAGTGTCCGACAAAAGACAGCTGGACTGTGCAATGACATACAGAACTTGTAGTGGGAA
G.BE.DRCBL      ...T-----A-----G-----G-----C-----A-----CA-----A-----A-----C-----G-----A-----A-----T-----T-----A-----A-----A-----
G.ING.92NG083  ...A-----G-----A-----G-----C-----G-----G-----A-----C-----G-----AGAT-----T-----A-----A-----A-----A-----A-----
G.SE.SE6165     ...A-----G-----A-----G-----C-----G-----C-----A-----C-----C-----G-----G-----A-----T-----A-----A-----A-----A-----A-----
G.FI.HH87932   ...A-----G-----T-----C-----C-----G-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
H.BE.VI991     ...C-----T-----A-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----
H.BE.VI997     ...C-----T-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
H.CF.90CF056   ...C-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
J.SE.SE91733   ...T-----C-----A-----G-----G-----A-----A-----G-----A-----G-----AGAT-----T-----A-----A-----A-----A-----A-----
J.SE.SE92809   ...T-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----A-----AGA-----T-----A-----A-----A-----A-----A-----
K.CM.MP535C    ...C-----T-----C-----C-----G-----G-----C-----G-----T-----C-----G-----T-----C-----G-----T-----A-----A-----A-----A-----
K.CD.E0TB11C  ...C-----T-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----
O.CM.AN70      ...-A-----G-----C-----T-----A-----A-----A-----A-----C-----A-----C-----G-----T-----C-----G-----T-----G-----T-----A-----A-----
O.CM.MP5180    ...C-----T-----T-----A-----A-----G-----G-----G-----G-----G-----C-----A-----C-----G-----T-----C-----G-----T-----A-----A-----
O.SN.MP1299    ...-A-----G-----C-----T-----T-----G-----G-----G-----G-----G-----A-----C-----A-----C-----G-----T-----C-----G-----T-----A-----A-----
O.SN.MP1300    ...-A-----G-----C-----T-----G-----G-----G-----G-----G-----A-----C-----A-----C-----G-----T-----C-----G-----T-----A-----A-----
N.CM.YBF30     ...C-----G-----C-----G-----C-----A-----C-----A-----C-----A-----A-----A-----T-----A-----A-----A-----A-----A-----
N.CM.YBF106    ...-A-----G-----G-----A-----G-----C-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
CONSENSUS 01  ...g-----c-----t-----a-----t-----C-----G-----C-----a-----t-----C-----G-----a-----a-----c-----a-----T-----g-----a-----
01.AE.TH.93TH253  ...G-----G-----G-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----
01.AE.TH.CM240   ...G-----G-----G-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----
01.AE.TH.92TH022 ...G-----G-----G-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----
01.AE.TH.95TNLH047 ...G-----G-----G-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----
01.AE.CF.11697   ...G-----G-----G-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----
01.AE.TH.93TH057 ...G-----G-----G-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----
01.AE.TH.93TH065 ...G-----G-----G-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----G-----C-----
01.AE.CF.4071   ...-T-----T-----a-----a-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
CONSENSUS 02  ...a-----a-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
02.AG.SE.7812    ...T-----T-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
02.AG.GH.G829    ...T-----T-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
02.AG.DJ.DJ263   ...T-----T-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
02.AG.DJ.DJ264   ...T-----T-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
02.AG.NG.IBNG    ...T-----T-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
02.AG.SN.MP1211  ...-G-----G-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
02.AG.SN.MP1213  ...T-----T-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
02.AG.CM.M8807   ...T-----T-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
03.AB.RU.98001   ...T-----T-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
03.AB.RU.KALI53-2 ...T-----T-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
04.CPX.GR.97FVMY ...C-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
04.CPX.GR.97FVCH ...C-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
04.CPX.CY.94CY0323 ...C-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
05.DF.BE.VI1310 ...C-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
05.DF.BE.VI961   ...G-----C-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
06.CPX.ML.95ML84 ...C-----C-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
06.CPX.AU.BPP90  ...C-----C-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
06.CPX.ML.95ML127 ...C-----C-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
06.CPX.SN.97SE1078 ...C-----C-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
10.CD.BET061     ...-G-----G-----T-----T-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----
10.CD.BEL10     ...T-----G-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----
10.CD.BEL1071   ...T-----G-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----
11.CPX.CM.MP818  ...C-----C-----C-----C-----A-----G-----A-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
11.CPX.FR.MP1298 ...C-----C-----C-----C-----A-----G-----A-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
11.CPX.FR.MP1307 ...C-----C-----C-----C-----A-----G-----A-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
11.CPX.GR.GR17   ...C-----C-----C-----C-----A-----G-----A-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
CPZ.GA.SIVCPZGAB ...A-----G-----A-----C-----T-----T-----A-----G-----G-----C-----A-----C-----T-----C-----A-----A-----G-----G-----T-----A-----
CPZ.US.SIVCPZUS  ...G-----G-----A-----A-----T-----T-----A-----C-----G-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
CPZ.US.SIVCPZCAM3 ...G-----G-----A-----A-----T-----T-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----
CPZ.CM.SIVCPZCAM5 ...A-----G-----A-----C-----T-----T-----G-----A-----G-----A-----G-----A-----G-----A-----G-----A-----G-----A-----G-----A-----
CPZ.CD.SIVCPZANT ...T-----AA-----A-----A-----G-----T-----A-----A-----C-----C-----T-----T-----AAA-----AA-----A-----G-----C-----TGAT-----A-----T-----

```

Pol p51 RT

B_FR_HXB2R
G_BE_DRCBL
G_ING_92NG083
G_SE_SE6165
G_FI_HH87932
H_BE_VI991
H_BE_VI997
H_CF_90CF056
J_SE_SE91733
J_SE_SE92809
K_CM_MP535C
K_CD_F07B11C
O_CM_ANT70
O_CM_MP5180
O_SN_MP1299
O_SN_MP1300
N_CM_YBF30
N_CM_YBF106
CONSENSUS 01
01_AE_TH_93TH253
01_AE_TH_CM240
01_AE_CF_90CF402
01_AE_TH_92TH022
01_AE_TH_95TNIH047
01_AE_CF_11697
01_AE_TH_93TH057
01_AE_TH_93TH065
01_AE_CF_4071
CONSENSUS 02
02_AG_SE_7812
02_AG_GH_G829
02_AG_DJ_DJ263
02_AG_DJ_DJ264
02_AG_NG_IBNG
02_AG_SN_MP1211
02_AG_SN_MP1213
02_AG_CM_M807
03_AB_RU_98001
03_AB_RU_KALI53-2
04_CFX_GR_97PVMY
04_CFX_GR_97PVCH
04_CFX_CY_94CY0323
05_DF_BE_VI1310
05_DF_BE_VI961
06_CFX_ML_95ML84
06_CFX_AU_BPP90
06_CFX_ML_95ML127
06_CFX_SN_97SE1078
10_CD_BEL101
10_CD_BEL110
10_CD_BEL1071
11_CFX_CM_MP818
11_CFX_FR_MP1298
11_CFX_FR_MP1307
11_CFX_GR_GRI17
CPZ_GA_SIVCPZGAB
CPZ_US_SIVCPZUS
CPZ_CM_SIVCPZCAM3
CPZ_CM_SIVCPZCAM5
CPZ_CD_SIVCPZANT
L_N_W_A_S_O_I_Y_P_G_I_K_V_R_Q_L_C_K_L_L_R_G_T_K_A_L_T_E_V_I_P_L_I_T_E_A_E_L_E_L_A_E_N
Pol p51 RT

B. FR. HXB2R
G. BE. DRCEI
G. NG. 92NG083
G. SE. SE6165
G. FI. HH87932
H. BE. VI991
H. BE. VI997
H. CF. 90CF056
J. SE. SE91733
J. SE. SE92809
K. CM. MP535C
K. CD. E07B11C
O. CM. AN70
O. CM. MP5180
O. SN. MP1299
O. SN. MP1300
N. CM. YBF30
N. CM. YBF106
N. CM. YBF106
CONSENSUS 01
01 AE. TH. 93TH253
01 AE. TH. CM240
01 AE. CF. 90CF402
01 AE. TH. 92TH022
01 AE. TH. 95TNIH047
01 AE. CF. 11697
01 AE. TH. 93TH057
01 AE. TH. 93TH065
01 AE. CF. 4071
CONSENSUS 02
02 AG. SE. 7812
02 AG. GH. G829
02 AG. DJ. DJ263
02 AG. DJ. DJ264
02 AG. NG. I8NG
02 AG. SN. MP1211
02 AG. SN. MP1213
02 AG. CM. M807
03 AB. RU. 98001
03 AB. RU. KALI53-2
04 CPX. GR. 97PVMY
04 CPX. GR. 97PVCH
04 CPX. CY. 94CY0323
05 DF. BE. VI1310
05 DF. BE. VI961
06 CPX. ML. 95ML84
06 CPX. AU. BPP90
06 CPX. ML. 95ML127
06 CPX. SN. 97SE1078
10 CD. BFL061
10 CD. BFL110
10 CD. BFL1071
11 CPX. CM. M818
11 CPX. FR. MP1298
11 CPX. FR. MP1307
11 CPX. GR. GR17
CPZ. GA. STVCPZGAB
CPZ. US. STVCPZUS
CPZ. CM. STVCPZCAM3
CPZ. CM. STVCPZCAM5
CPZ. CD. STVCPZANT

HIV-1/SIVcpz complete genomes

B. FR. HXB2R
 TACAAGCAATTTATAGCTTTGCGGATTTCAGTAAGTAAACATAGTAGTAAACAGACTCACATATGCATTTAGGAATCATTCAAGCACAACCCAGATCAAAAGTGAATCAGAGTTAGTCAATCAAAATAAT
 -a aT-aa cca ctag ct-----A-a C-aa-a ac-----ca-----g rat--tta-a c---g-a-a---CAGG-----a-----gt-----c-----
 -G-T-----C-----C-----T-----T-----G-----G-----G-----T-----G-----G-----AA-----
 A. SE. UGSE8891
 A. SE. UGSE8891
 A. SE. TZSE8538
 A. SE. UGSE6594
 A. KE. Q2317
 A. SE. UGSE7535
 A. SE. SOSE7253
 A. SE. U2UG037
 A. SE. SE8131
 A2. CD. 97CDKTB48
 A2. CY. 94CY017
 CONSENSUS B
 B. AU. MBC925
 B. AU. MBC200
 B. GA. OYI
 B. GB. MANC
 B. GB. CAMI
 B. CN. RL42
 B. DE. D31
 B. DE. HAN2
 B. ES. 89SP061
 B. NL. ACH320A
 B. TW. LM49
 B. US. DHL23
 B. US. SF2
 B. US. NY5CG
 B. US. WCIPL18
 B. US. YU2
 B. US. JRCSF
 B. US. JRFL
 B. US. MN
 B. US. BCSG3
 B. US. WEAU160
 B. US. RF
 B. US. WR27
 CONSENSUS C
 C. BR. 92BR025
 C. BW. 96BW0402
 C. BW. 96BW1104
 C. BW. 96BW15C02
 C. BW. 96BW01B03
 C. BW. 96BW0502
 C. BW. 96BW16B01
 C. BW. 96BW1210
 C. ET. ETH2220
 C. IN. 301999
 C. IN. 21068
 C. IN. 301905
 C. IN. 301904
 C. IN. 93IN101
 C. IN. 11246
 D. UG. 94UG114
 D. CD. NDK
 D. CD. ELI
 D. KE. MB2059
 D. CD. Z226
 D. CD. 84ZR085
 F1. BR. 93BR020
 F1. FI. FIN9363
 F1. BE. V1850
 F1. FR. MF411
 F2. CM. MF255C
 F2. CM. MF257C

4114
 3230
 3560
 3303
 3313
 3298
 3572
 3331
 3318
 3481
 3486
 3452
 3471
 4064
 4125
 4113
 3657
 4060
 4116
 3501
 3660
 3580
 3672
 4116
 4114
 4105
 4121
 4114
 3660
 4126
 3431
 4129
 3661
 4113
 3631
 3493
 3478
 3452
 3580
 3474
 3581
 3579
 3607
 3580
 3486
 3506
 3493
 3465
 3465
 3465
 4087
 3456
 3469
 3649
 3660
 4118
 4117
 3636
 3442
 3443
 3454
 3316
 3322

B. FR. HXB2R	TACAAAGCAATTTATCTAGCTTTCGAGGATTCGGGATTTAGAAAGTAAACATAGTAAACAGACTCACAAATATGCAATTTAGGAATCATTCAAGCACAAACAGATCAAAAGTTGAATCAGAGTTAGTCAATCAAAATAAT	4114
G. BE. DRCEI	-----T-----C-----C-----ACA-----C-----AGG-----G-----AA-----	4071
G. NG. 92NG083	-----T-----C-----A-----C-----AA-----C-----AGG-----G-----A-----	3483
G. SE. SE6165	-----G-----G-----A-----C-----AA-----C-----AGG-----G-----A-----	3511
G. FI. HH87932	-----C-----G-----G-----A-----G-----A-----GCC-----T-----A-----G-----AC-----	3514
H. BE. VI991	-----A-----G-----A-----G-----A-----GCC-----T-----A-----G-----AC-----	3503
H. CF. 90CF056	-----A-----A-----G-----G-----T-----G-----G-----T-----A-----G-----T-----	3438
J. SE. SE91733	-----C-----T-----T-----A-----G-----A-----GA-----T-----G-----T-----C-----A-----G-----T-----	3461
J. SE. SE92809	-----C-----C-----T-----A-----G-----A-----GC-----T-----G-----G-----T-----A-----A-----G-----A-----	3429
K. CM. MP535C	-----C-----C-----G-----T-----A-----A-----C-----A-----C-----G-----A-----A-----G-----A-----	3310
K. CD. E0TB11C	-----CC-----T-----A-----A-----C-----A-----C-----A-----A-----T-----A-----A-----T-----	3310
O. CM. AN170	-----ATG-----G-----A-----TA-----C-----A-----CA-----GA-----AC-----T-----G-----CG-----CTCCT-----C-----TACA-----G-----CCCTA-----TC-----G-----	4169
O. CM. MP5180	-----ATG-----TG-----A-----TAA-----C-----A-----CAAGGAC-----T-----T-----G-----C-----ATCCT-----C-----ACA-----G-----C-----CCCTA-----TC-----G-----	4144
O. SN. MP1299	-----ATG-----G-----A-----TAT-----C-----A-----CAAGGAC-----T-----T-----G-----C-----ATCCT-----C-----ACA-----G-----C-----CCCTC-----TC-----G-----	4168
O. SN. MP1300	-----ATG-----G-----A-----TAT-----C-----A-----CAAGGAC-----T-----T-----G-----C-----ATCCT-----C-----ACA-----G-----G-----CCCTC-----TC-----G-----	4168
N. CM. YBF106	-----T-----CCT-----A-----G-----C-----A-----A-----G-----A-----CGG-----T-----C-----T-----T-----G-----TA-----G-----A-----CT-----A-----G-----G-----GC-----	3708
N. CM. YBF30	-----T-----G-----CCa-----T-----C-----a-----a-----t-----a-----t-----c-----g-----CAGG-----a-----ag-----ac-----a-----a-----	4065
CONSENSUS 01	-----T-----G-----CC-----T-----C-----A-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	4101
01 AE. TH. CM240	-----T-----G-----CC-----T-----C-----A-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	3688
01 AE. CF. 90CF402	-----T-----G-----CC-----T-----C-----A-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	4150
01 AE. TH. 92TH022	-----T-----G-----CC-----T-----C-----A-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	4077
01 AE. TH. 95TH1047	-----T-----G-----CC-----T-----C-----A-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	4077
01 AE. CF. 11697	-----T-----G-----CC-----T-----C-----A-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	4067
01 AE. TH. 93TH057	-----T-----G-----CC-----T-----C-----A-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	4067
01 AE. TH. 93TH065	-----T-----G-----CC-----T-----C-----A-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	4067
01 AE. CF. 4071	-----T-----G-----CC-----T-----C-----A-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	4059
CONSENSUS 02	-----T-----T-----G-----CG-----T-----C-----G-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	4050
02 AG. SE. 7812	-----T-----T-----ca-----C-----g-----t-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	3222
02 AG. GH. G829	-----T-----T-----C-----C-----T-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	3486
02 AG. DJ. DJ263	-----T-----T-----C-----C-----T-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	3330
02 AG. DJ. DJ264	-----T-----T-----C-----C-----T-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	3462
02 AG. NG. I8NG	-----T-----T-----C-----C-----T-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	3463
02 AG. SN. MP1211	-----T-----T-----C-----C-----T-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	3639
02 AG. SN. MP1213	-----T-----T-----C-----C-----T-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	3311
02 AG. CM. M807	-----T-----T-----C-----C-----T-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	3308
03 AB. RU. 98001	-----T-----T-----C-----C-----T-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	3270
03 AB. RU. KAL153-2	-----T-----T-----C-----C-----T-----C-----A-----C-----C-----T-----T-----G-----C-----CAGG-----G-----AG-----GC-----	3484
04 CPX. GR. 97PVMY	-----CT-----C-----CT-----A-----A-----C-----C-----A-----G-----A-----G-----A-----T-----T-----	3883
04 CPX. GR. 97PVCB	-----CT-----C-----CT-----A-----A-----C-----C-----A-----G-----A-----G-----A-----T-----T-----	4128
04 CPX. CY. 94CY0323	-----CT-----C-----CT-----A-----A-----C-----C-----A-----G-----A-----G-----A-----T-----T-----	4131
05 DF. BE. V11310	-----C-----C-----A-----A-----A-----A-----C-----C-----A-----G-----A-----G-----A-----T-----T-----	3496
05 DF. BE. V1961	-----T-----T-----A-----A-----A-----A-----C-----C-----G-----G-----T-----A-----G-----A-----T-----T-----	3469
06 CPX. ML. 95ML84	-----A-----A-----C-----A-----A-----C-----C-----A-----G-----A-----G-----A-----T-----T-----	3357
06 CPX. AU. BPP90	-----A-----A-----C-----A-----A-----C-----C-----A-----G-----A-----G-----A-----T-----T-----	4142
06 CPX. ML. 95ML127	-----A-----A-----C-----A-----A-----C-----C-----A-----G-----A-----G-----A-----T-----T-----	4131
06 CPX. SN. 97SE1078	-----C-----A-----A-----A-----A-----C-----C-----T-----T-----G-----C-----T-----A-----G-----A-----T-----T-----	4173
10 CD. BEL061	-----C-----A-----A-----A-----A-----C-----C-----T-----T-----G-----C-----T-----A-----G-----A-----T-----T-----	3493
10 CD. BEL110	-----C-----A-----A-----A-----A-----C-----C-----T-----T-----G-----C-----T-----A-----G-----A-----T-----T-----	3482
10 CD. BEL1071	-----C-----A-----A-----A-----A-----C-----C-----T-----T-----G-----C-----T-----A-----G-----A-----T-----T-----	3491
11 CPX. CM. M818	-----G-----C-----C-----C-----A-----A-----A-----A-----C-----C-----T-----T-----G-----C-----T-----A-----G-----A-----T-----T-----	4117
11 CPX. FR. MP1298	-----G-----C-----C-----C-----A-----A-----A-----A-----C-----C-----T-----T-----G-----C-----T-----A-----G-----A-----T-----T-----	4108
11 CPX. FR. MP1307	-----G-----C-----C-----C-----A-----A-----A-----A-----C-----C-----T-----T-----G-----C-----T-----A-----G-----A-----T-----T-----	4165
11 CPX. GR. GR17	-----G-----C-----C-----C-----A-----A-----A-----A-----C-----C-----T-----T-----G-----C-----T-----A-----G-----A-----T-----T-----	3416
CPZ. GA. STVCPZGAB	-----A-----G-----TT-----GCT-----G-----C-----A-----ATCA-----C-----G-----T-----G-----T-----GT-----G-----G-----A-----G-----	4175
CPZ. US. STVCPZUS	-----GAC-----GG-----A-----CT-----A-----AA-----A-----A-----AGAATAC-----T-----TG-----C-----T-----G-----C-----CT-----T-----G-----G-----	4174
CPZ. CM. STVCPZCAM3	-----GAC-----TG-----A-----A-----A-----A-----G-----A-----A-----CATAT-----T-----T-----C-----T-----T-----C-----G-----A-----G-----T-----	3525
CPZ. CM. STVCPZCAM5	-----AC-----TG-----G-----A-----A-----G-----A-----A-----CATAT-----C-----T-----T-----T-----T-----T-----T-----G-----A-----AGGGTCT-----	3816
CPZ. CD. STVCPZANT	-----G-----G-----G-----CT-----A-----G-----C-----AG-----AAC-----GGCCC-----C-----T-----C-----T-----G-----T-----G-----GTACC-----CCCC-----GG-----AG-----	3553

p15 RNase

HIV-1/SIVcpz complete genomes

	Pol_p31 Integrase
B.FR.HXB2R	4241
CONSENSUS A	4241
A.UG.U455	3457
A.SE.UGSE8891	3687
A.SE.TZSE8538	3430
A.SE.UGSE6594	3440
A.KE.Q2317	3425
A.SE.UGSE7535	3699
A.SE.SOSE7253	3458
A.UG.92UG037	3445
A.SE.SE8131	3608
A2.CD.97CDKTB48	3613
A2.CT.94CY017	3579
CONSENSUS B	3598
B.AU.MBC925	4191
B.AU.MBC200	4240
B.GA.OYI	3784
B.GB.MANC	4187
B.GB.CAMI	4243
B.CN.RL42	3628
B.DE.D31	3787
B.DE.HAN2	3707
B.ES.89SP061	3799
B.NL.ACH320A	4243
B.IW.LM49	4241
B.US.DH123	4232
B.US.SF2	4248
B.US.NY5CG	4241
B.US.WC1PRL8	3787
B.US.YU2	4239
B.US.JRGSF	4253
B.US.JRFL	3558
B.US.MN	4256
B.US.RCSG3	3788
B.US.WEAU160	4240
B.US.RF	3758
B.US.WR27	3620
CONSENSUS C	3605
C.BR.92BR025	3579
C.BW.96BW0402	3707
C.BW.96BW1104	3601
C.BW.96BW15C02	3708
C.BW.96BW01B03	3706
C.BW.96BW0502	3734
C.BW.96BW16B01	3707
C.BW.96BW1210	3613
C.ET.FTH2220	3633
C.IN.301999	3620
C.IN.21068	3592
C.IN.301905	3592
C.IN.301904	3592
C.IN.93IN101	4214
C.IN.11246	3583
D.UG.94UG114	3596
D.CD.NDK	3776
D.CD.ELI	3787
D.KE.MB2059	4245
D.CD.Z226	4244
D.CD.842R085	3763
F1.BR.93BR020	3659
F1.FI.FIN9363	3570
F1.BE.VI850	3561
F1.FR.MF411	3443
F2.CM.MF255C	3437
F2.CM.MF257C	3449

Table listing HIV-1/SIVcpz complete genomes with columns for Accession ID (e.g., B_FR_HXB2R), Genomic Region (e.g., ATAGATAGGCCCAAGAT), and Position (e.g., 4359). The table includes a long list of accession numbers and their corresponding genomic coordinates.

HIV-1/SIVcpz complete genomes

B. FR. HXB2R	AGCTAAAGGAGGAGCCATGCACTAGTCCAGGAATATGCAACTAGTACCAITTTAGAGGAAAGTTTCCCTGGTAGCAGTTCATGTAGCCAGTGGATATATAGAACGAGA	4489
G. BE. DRCEI	-T-----G-----T-----T-----G-----A-----A-----A-----A-----C-----C-----C-----C-----	4446
G. NG. 92NG083	-A-----G-----T-----T-----T-----A-----A-----A-----A-----C-----C-----C-----C-----	3558
G. SE. SE6165	-G-----G-----T-----T-----T-----A-----A-----A-----A-----C-----C-----C-----C-----	3886
G. FI. HH87932	-T-----T-----T-----T-----T-----A-----A-----A-----A-----C-----C-----C-----C-----	3889
H. BE. VI991	-G-----G-----T-----T-----T-----C-----C-----C-----C-----A-----A-----A-----A-----	3878
H. BE. VI997	-G-----G-----T-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----	3813
H. CF. 90CF056	-G-----G-----T-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----	3836
J. SE. SE91733	-G-----G-----T-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----	3804
J. SE. SE92809	-G-----G-----T-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----	3803
K. CM. MP535C	-G-----G-----T-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----	3685
K. CD. EOTB11C	-G-----G-----T-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----	3686
O. CM. ANT70	-TA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	4544
O. CM. MP5180	-TA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	4519
O. SN. MP1299	-TA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	4543
O. SN. MP1300	-TA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	4543
N. CM. YBF30	-G-----G-----T-----T-----T-----G-----G-----G-----G-----G-----G-----G-----G-----	4083
N. CM. YBF106	-G-----G-----T-----T-----T-----G-----G-----G-----G-----G-----G-----G-----G-----	4081
CONSENSUS 01	-a-----g-----t-----a-----t-----c-----a-----c-----c-----c-----c-----c-----c-----	4440
01 AE. TH. 93TH253	-A-----T-----T-----T-----T-----T-----A-----T-----A-----T-----T-----T-----T-----	4476
01 AE. TH. CM240	-A-----T-----T-----T-----T-----T-----A-----T-----A-----T-----T-----T-----T-----	4063
01 AE. CF. 90CF402	-G-----G-----T-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----	4525
01 AE. TH. 92TH022	-G-----G-----T-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----	4452
01 AE. TH. 95TNLH047	-A-----A-----T-----T-----T-----T-----T-----A-----T-----A-----T-----T-----T-----T-----	4452
01 AE. CF. 11697	-A-----A-----T-----T-----T-----T-----T-----A-----T-----A-----T-----T-----T-----T-----	4438
01 AE. TH. 93TH057	-A-----A-----T-----T-----T-----T-----T-----A-----T-----A-----T-----T-----T-----T-----	4438
01 AE. TH. 93TH065	-A-----A-----T-----T-----T-----T-----T-----A-----T-----A-----T-----T-----T-----T-----	4442
01 AE. CF. 4071	-A-----A-----T-----T-----T-----T-----T-----A-----T-----A-----T-----T-----T-----T-----	4434
CONSENSUS 02	-c-----a-----g-----g-----g-----g-----g-----g-----g-----g-----g-----g-----	4425
02 AG. SE. 7812	-G-----G-----T-----T-----T-----T-----T-----A-----A-----A-----A-----C-----G-----	3697
02 AG. GH. G829	-G-----G-----T-----T-----T-----T-----T-----A-----A-----A-----A-----C-----G-----	3861
02 AG. DJ. DJ263	-G-----G-----T-----T-----T-----T-----T-----A-----A-----A-----A-----C-----G-----	3705
02 AG. DJ. DJ264	-G-----G-----T-----T-----T-----T-----T-----A-----A-----A-----A-----C-----G-----	3837
02 AG. NG. IBNG	-G-----G-----T-----T-----T-----T-----T-----A-----A-----A-----A-----C-----G-----	3838
02 AG. SN. MP1211	-G-----G-----T-----T-----T-----T-----T-----A-----A-----A-----A-----C-----G-----	4014
02 AG. SN. MP1213	-G-----G-----T-----T-----T-----T-----T-----A-----A-----A-----A-----C-----G-----	3686
02 AG. CM. M5807	-G-----G-----T-----T-----T-----T-----T-----A-----A-----A-----A-----C-----G-----	3683
03 AB. RU. 98001	-AT-----AT-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	3845
03 AB. RU. KALI53-2	-AT-----AT-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	3859
04 CPX. GR. 97FVNY	-G-----G-----T-----T-----T-----T-----T-----T-----C-----C-----C-----C-----	4258
04 CPX. GR. 97FVCH	-G-----G-----T-----T-----T-----T-----T-----T-----C-----C-----C-----C-----	4503
04 CPX. CY. 94CY0323	-G-----G-----T-----T-----T-----T-----T-----T-----C-----C-----C-----C-----	4506
05 DF. BE. VII1310	-T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	3855
05 DF. BE. VI961	-G-----G-----T-----T-----T-----T-----T-----T-----C-----C-----C-----C-----	3871
06 CPX. ML. 95ML84	-G-----G-----T-----T-----T-----T-----T-----T-----C-----C-----C-----C-----	3845
06 CPX. AU. BPP90	-G-----G-----T-----T-----T-----T-----T-----T-----C-----C-----C-----C-----	3732
06 CPX. ML. 95ML127	-G-----G-----T-----T-----T-----T-----T-----T-----C-----C-----C-----C-----	4517
06 CPX. SN. 97SE1078	-A-----A-----T-----T-----T-----T-----T-----T-----C-----C-----C-----C-----	4506
10 CD. BET061	-G-----G-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	4548
10 CD. BEL110	-G-----G-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	3868
10 CD. BEL1071	-G-----G-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	3857
11 CPX. CM. M5818	-A-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	3866
11 CPX. FR. MP1298	-A-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	4492
11 CPX. FR. MP1307	-A-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	4483
11 CPX. GR. GR17	-A-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	4540
CPZ. GA. STVCPZGAB	-G-----G-----T-----T-----T-----T-----T-----T-----C-----C-----C-----C-----	3791
CPZ. US. STVCPZUS	-C-----C-----T-----T-----T-----T-----T-----T-----C-----C-----C-----C-----	4550
CPZ. CM. STVCPZCAM3	-A-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	4549
CPZ. CM. STVCPZCAM5	-T-----T-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	3903
CPZ. CD. STVCPZANT	-CG-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	4191
Integrase	Q_L_K_G_E_A_M_H_G_Q_V_D_C_S_P_G_I_W_Q_L_D_C_T_H_L_E_G_K_V_I_L_V_A_V_H_V_A_S_G_Y_I_E_A_E	3928

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz
complete genomes

```

B. FR. HXB2R          AAGTTATCCAGCAGAAACAGGGCAGGAACACATACATGAGCAATGGCAGCAATTTACCCGGTCTAGCGTTAGGGCCGCCCTGT
CONSENSUS A          ---t-c-----a-g-ga-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
A.UG.U455            A---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
A.SE.UGSE8891       A---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
A.SE.TZSE8538       A---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
A.SE.UGSE6594       A---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
A.KE.Q2317          A---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
A.SE.UGSE7535       A---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
A.SE.SOSE7253       A---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
A.UG.92UG037       A---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
A.SE.SE8131         A---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
A2.CD.97CDKTB48    A---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
A2.CY.94CY017      A---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
CONSENSUS B          a-t-----c-g---g-a-----c-fc-ct-----aaaaa-c-a-t-a-----cagc-t---ccA-tacta 9---ag-----gt
B.AU.MBC925        G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.GA.MBC200        G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.GA.OYI           G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.GB.MANC          G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.GB.CAMI          G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.CN.RL42          G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.DE.D31           G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.DE.HAN2          G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.ES.89SP061       G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.NL.ACH320A       G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.TM.LM49          G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.US.DH123         G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.US.SF2           G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.US.NY5CG         G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.US.WCIPR18       G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.US.YU2           G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.US.JRCSF        G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.US.JRFL          G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.US.MN            G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.US.BCSG3        G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.US.WEAU160       G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.US.RF            G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
B.US.WR27          G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
CONSENSUS C          g---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---k-
C.BR.92BR025       G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
C.BW.96BW0402      G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
C.BW.96BW1104      G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
C.BW.96BW15C02     G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
C.BW.96BW01B03     G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
C.BW.96BW0502      G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
C.BW.96BW16B01     G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
C.BW.96BW1210     G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
C.ET.ETH2220       G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
C.IN.301999         G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
C.IN.21068          G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
C.IN.301905         G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
C.IN.301904         G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
C.IN.93JIN101      G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
C.IN.11246          G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
D.UG.94UG114       G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
D.CD.NDK           G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
D.CD.ELI           G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
D.KE.MB2059        G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
D.CD.42R085        G---g-----g-----c-c-ac-----gt-g---c-a-c---gag--aa-a-c---
F1.BR.93BR020      G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
F1.FI.FIN9363      G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
F1.BE.V1850        G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
F1.FR.MF411        G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
F2.CM.ME255C       G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
F2.CM.ME257C       G---c---a-a-a---g---g-----c-tta-ac-a-----R-----W-W-----t-gT-t-c-cCA-tgc-g-a-tNg-A---
    
```


B. FR. HXB2R	TGGTGGGGGGAATCAAGCAGGAAATTTGGAAATCCCTCAATCCCAAGTCAAGGAGTAGTAGAATCTATGAATAAAGAATTAAGAAAATATAGACAGGTAAGAGATCAGGCTGACATCTTAAAGA	4749
G. BE. DRCEI	AAACCA	4706
G. NG. 92NG083	AAATCA	4118
G. SE. SE6165	AAATCA	4146
G. FI. HH87932	AAATCA	4149
H. BE. VI991	AATTC	4138
H. BE. VI997	AATTC	4073
H. CF. 90CF056	AATTC	4096
J. SE. SE91733	AATTC	4064
J. SE. SE92809	AATTC	4063
K. CM. MP535C	AATTC	3946
K. CD. E0TB11C	AATTC	3946
O. CM. ANI70	AATTC	4804
O. CM. MP5180	AATTC	4779
O. SN. MP1299	AATTC	4803
O. SN. MP1300	AATTC	4803
N. CM. YBF30	AATTC	4803
N. CM. YBF106	AATTC	4343
CONSENSUS 01	AATTC	4341
01 AE. TH. 93TH253	AATTC	4700
01 AE. TH. CM240	AATTC	4736
01 AE. CF. 90CF402	AATTC	4323
01 AE. TH. 92TH022	AATTC	4785
01 AE. TH. 95THH047	AATTC	4712
01 AE. CF. 11697	AATTC	4698
01 AE. TH. 93TH057	AATTC	4702
01 AE. TH. 93TH065	AATTC	4702
01 AE. CF. 4071	AATTC	4885
CONSENSUS 02	AATTC	3957
02 AG. SE. 7812	AATTC	4121
02 AG. GH. G829	AATTC	3965
02 AG. DJ. DJ263	AATTC	4097
02 AG. DJ. DJ264	AATTC	4098
02 AG. NG. I8NG	AATTC	4274
02 AG. SN. MP1211	AATTC	3946
02 AG. SN. MP1213	AATTC	3946
02 AG. CM. MB807	AATTC	3905
03 AB. RU. 98001	AATTC	4119
03 AB. RU. KALI53-2	AATTC	4518
04 CPX. GR. 97PVMY	AATTC	4763
04 CPX. GR. 97PVCB	AATTC	4766
04 CPX. CY. 94CY0323	AATTC	4115
05 DF. BE. VI1310	AATTC	4131
05 DF. BE. VI961	AATTC	4105
06 CPX. ML. 95ML84	AATTC	3992
06 CPX. AU. BPP90	AATTC	4777
06 CPX. ML. 95ML127	AATTC	4766
06 CPX. SN. 97SE1078	AATTC	4808
10 CD. BFL061	AATTC	4128
10 CD. BFL110	AATTC	4117
10 CD. BFL071	AATTC	4126
11 CPX. CM. MB818	AATTC	4752
11 CPX. FR. MP1298	AATTC	4743
11 CPX. FR. MP1307	AATTC	4800
11 CPX. GR. GR17	AATTC	4051
CPZ. GA. STVCPZGAB	AATTC	4810
CPZ. US. STVCPZUS	AATTC	4809
CPZ. CM. STVCPZCAM3	AATTC	4163
CPZ. CM. STVCPZCAM5	AATTC	4451
CPZ. CD. STVCPZANT	AATTC	4188

Integrase

HIV-1/SIVcpz complete genomes

B_FR_HXB2R
 G_BE_DRCBI
 G_ING_92NG083
 G_SE_SE6165
 G_FI_HH87932
 H_BE_VI991
 H_BE_VI997
 H_CF_90CF056
 J_SE_SE91733
 J_SE_SE92809
 K_CM_MP535C
 K_CD_E0TB11C
 O_CM_ANT70
 O_CM_MPF5180
 O_SN_MP1299
 O_SN_MP1300
 N_CM_YBF30
 N_CM_YBF106
 N_CM_YBF106
 CONSENSUS_01
 01_AE_TH_93TH253
 01_AE_TH_CM240
 01_AE_CF_90CF402
 01_AE_TH_92TH022
 01_AE_TH_95TNLH047
 01_AE_TH_93TH057
 01_AE_TH_93TH065
 01_AE_CF_4071
 CONSENSUS_02
 02_AG_SE_7812
 02_AG_GH_G829
 02_AG_DJ_DJ263
 02_AG_DJ_DJ264
 02_AG_NG_IBNG
 02_AG_SN_MP1211
 02_AG_SN_MP1213
 02_AG_CM_M807
 03_AB_RU_98001
 03_AB_RU_KALI53-2
 04_CPX_GR_97FVMY
 04_CPX_GR_97FVCH
 04_CPX_CY_94CY0323
 05_DF_BE_VI1310
 05_DF_BE_VI961
 06_CPX_ML_95ML84
 06_CPX_AU_BFP90
 06_CPX_ML_95ML127
 06_CPX_SN_97SE1078
 10_CD_BET061
 10_CD_BEL110
 11_CPX_CM_MP818
 11_CPX_FR_MP1298
 11_CPX_FR_MP1307
 11_CPX_GR_GR17
 CPZ_GA_SIVCPZGAB
 CPZ_US_SIVCPZUS
 CPZ_CM_SIVCPZCAM3
 CPZ_CM_SIVCPZCAM5
 CPZ_CD_SIVCPZANT

CAGCAGTACAATGGCAGTATTTCACCAATTTTAAAGAAAGGGGGG...ATTGGGGGTACAGTCCAGGGGAAAGAAATAGTACAGTACATAGCAAGACAGATACAACTTAAGAAATTACAAAAACA

4876
 4873
 4845
 4273
 4276
 4265
 4200
 4223
 4191
 4190
 4072
 4073
 4931
 4906
 4930
 4950
 4470
 4468
 4827
 4863
 4450
 4912
 4839
 4839
 4825
 4829
 4819
 4812
 4084
 4248
 4092
 4224
 4225
 4401
 4073
 4070
 4032
 4246
 4845
 4890
 4893
 4242
 4258
 4232
 4119
 4904
 4893
 4935
 4354
 4344
 4879
 4870
 4927
 4178
 4937
 4936
 4290
 4578
 4315

Pol
 T_A_V_Q_M_A_V_F_I_H_N_F_K_R_K_G_G...I_G_G_Y_S_A_G_E_R_I_V_D_I_I_A_T_D_I_O_T_K_E_L_O_K_Q

HIV-1/SIVcpz
 complete genomes

B. FR. HXB2R	5003	AATTACAAA...	ATTCAAATTTTCGGGTTTATTACAGGACAGCAGCAATCCACTTTTGAAAGAGGACAGCAGCAAGCTCCTCTGAAAGAGGTGAAGGGCAGTAGTCAAGTAATAGTAGCATATAAAA	5003
CONSENSUS A	4219	a--ac-----	g--t--aa-----	4219
A.UG.0455	4449-t-----	g--c--ca-----	4449
A.SE.UGSE8891	4192-t-----	g--t--aa-----	4192
A.SE.TZSE8538	4202-t-----	g--t--aa-----	4202
A.SE.UGSE6594	4187-t-----	g--t--aa-----	4187
A.KE.Q2317	4461-t-----	g--t--aa-----	4461
A.SE.UGSE7535	4220-t-----	g--t--aa-----	4220
A.SE.SOSE7253	4207-t-----	g--t--aa-----	4207
A.UG.92UG037	4370-t-----	g--t--aa-----	4370
A.SE.SE8131	4374-t-----	g--t--aa-----	4374
A2.CD.97CDKTB48	4343-t-----	g--t--aa-----	4343
A2.CY.94CY017	4360-t-----	g--t--aa-----	4360
CONSENSUS B	4953	-a-----	aa--a-t-----	4953
B.AU.MBC925	5014-t-----	g--t--aa-----	5014
B.AU.MBC200	5002-t-----	g--t--aa-----	5002
B.GA.OYI	4546-t-----	g--t--aa-----	4546
B.GB.MANC	4949-t-----	g--t--aa-----	4949
B.GB.CAMI	5005-t-----	g--t--aa-----	5005
B.CN.RL42	4390-t-----	g--t--aa-----	4390
B.DE.D31	4549-t-----	g--t--aa-----	4549
B.DE.HAN2	4469-t-----	g--t--aa-----	4469
B.ES.89SP061	4561-t-----	g--t--aa-----	4561
B.NL.ACH320A	5005-t-----	g--t--aa-----	5005
B.TW.LM49	5004-t-----	g--t--aa-----	5004
B.US.DH123	4994-t-----	g--t--aa-----	4994
B.US.SF2	5010-t-----	g--t--aa-----	5010
B.US.NY5CG	5003-t-----	g--t--aa-----	5003
B.US.WCIPR18	4549-t-----	g--t--aa-----	4549
B.US.YU2	5001-t-----	g--t--aa-----	5001
B.US.JRCSF	5015-t-----	g--t--aa-----	5015
B.US.JRFL	4320-t-----	g--t--aa-----	4320
B.US.MN	5018-t-----	g--t--aa-----	5018
B.US.BCSG3	4550-t-----	g--t--aa-----	4550
B.US.WEAU160	5002-t-----	g--t--aa-----	5002
B.US.RF	4520-t-----	g--t--aa-----	4520
B.US.WR27	4382-t-----	g--t--aa-----	4382
CONSENSUS C	4367	-W-----	TT--w-c-----	4367
C.BR.92BR025	4340-t-----	g--t--aa-----	4340
C.BW.96BW0402	4469-t-----	g--t--aa-----	4469
C.BW.96BW1104	4363-t-----	g--t--aa-----	4363
C.BW.96BW15C02	4470-t-----	g--t--aa-----	4470
C.BW.96BW01B03	4468-t-----	g--t--aa-----	4468
C.BW.96BW0502	4496-t-----	g--t--aa-----	4496
C.BW.96BW16B01	4469-t-----	g--t--aa-----	4469
C.BW.96BW1210	4375-t-----	g--t--aa-----	4375
C.ET.ETH2220	4395-t-----	g--t--aa-----	4395
C.IN.301999	4382-t-----	g--t--aa-----	4382
C.IN.21068	4354-t-----	g--t--aa-----	4354
C.IN.301905	4354-t-----	g--t--aa-----	4354
C.IN.301904	4354-t-----	g--t--aa-----	4354
C.IN.93JIN101	4976-t-----	g--t--aa-----	4976
C.IN.11246	4345-t-----	g--t--aa-----	4345
D.UG.94UG114	4358-t-----	g--t--aa-----	4358
D.CD.NDK	4549-t-----	g--t--aa-----	4549
D.CD.ELI	5007-t-----	g--t--aa-----	5007
D.KE.MB2059	4525-t-----	g--t--aa-----	4525
D.CD.Z226	5006-t-----	g--t--aa-----	5006
D.CD.84ZR085	4331-t-----	g--t--aa-----	4331
F1.BR.93BR020	4332-t-----	g--t--aa-----	4332
F1.FI.FIN9363	4343-t-----	g--t--aa-----	4343
F1.BE.VI850	4205-t-----	g--t--aa-----	4205
F1.FR.MF411	4199-t-----	g--t--aa-----	4199
F2.CM.MF255C	4211-t-----	g--t--aa-----	4211
F2.CM.MF257C	-t-----	g--t--aa-----	

HIV-1/SIVcpz Complete Genomes

B. FR. HXB2R	AATTACAAA...ATTCAAAATTTTCGGGTTTTATTCAGGAGCAGCAGAGATTCACCTTTGGAAAGAGGACAGCAGCAGCAATAGTACATATAAA	5003
G. BE. DRGB1	-----G-C-A-A-----A-A-----C-----C-AC-A--	4960
G. NG. 92NG083	-----G-C-G-G-----C-A-A-----G-C-AC-A--	4972
G. SE. SE6165	-----G-C-G-G-----C-A-A-----C-AC-A--	4400
G. FI. HH87932	T-----T-----A-----C-AC-A--	4403
H. BE. VI991	MC-----A-----T-----C-AC-A--	4327
H. BE. VI997	T-----A-----A-----C-AC-A--	4350
H. CF. 90CF056	T-----C-----A-----C-AC-A--	4318
J. SE. SE91733	C-----G-C-A-A-----G-C-A-A--	4317
J. SE. SE92809	-----G-C-A-A-----A-A-----C-AC-A--	4199
K. CM. MP535C	-----G-A-A-----A-A-----C-AC-A--	4195
K. CD. EOTB11C	-----G-C-ITA-C-----G-C-A-----C-----G-A-----T-T-G	5058
O. CM. ANT70	-----G-C-ITA-C-----G-C-A-----C-----G-A-----T-T-G	5033
O. CM. MP5180	-----G-C-ITA-----G-C-A-----C-----AG-A-----T-G	5057
O. SN. MP1299	-----G-ITA-----G-C-A-----C-----AG-A-----T-G	5057
O. SN. MP1300	-----G-ITA-----G-C-A-----C-----AG-A-----T-G	4957
N. CM. YBF30	-----G-C-A-A-----C-A-T-G-----A-----CG-G-T--	4954
N. CM. YBF106	-----G-C-A-A-----C-A-T-G-----A-----CG-G-T--	4954
CONSENSUS 01	-----G-C-A-A-----a-A-----a-C-----T--	4990
01 AE. TH. 93TH253	-----G-C-A-A-----A-A-----C-----T--	4990
01 AE. TH. CM240	-----G-C-A-A-----A-A-----C-----T--	4577
01 AE. CF. 90CF402	-----G-C-A-A-----A-A-----C-----T--	5039
01 AE. TH. 92TH022	-----G-C-A-A-----A-A-----C-----T--	4966
01 AE. TH. 95THLH047	-----G-C-A-A-----A-A-----C-----T--	4966
01 AE. CF. 11697	-----G-C-A-A-----A-A-----C-----T--	4962
01 AE. TH. 93TH057	-----G-C-A-A-----A-A-----C-----T--	4966
01 AE. TH. 93TH065	-----G-C-A-A-----A-A-----C-----T--	4966
01 AE. CF. 4071	-----G-C-A-A-----A-A-----C-----T--	4946
CONSENSUS 02	-----G-C-A-----A-----C-----T--	4939
02 AG. SE. 7812	g-c-a-t-----t-----g-c-a-t-g-a-g-a-c-----g-c-t-----g	4211
02 AG. GH. G829	G-----G-C-CA-----A-A-----G-C-G-----T-G	4375
02 AG. DJ. DJ263	G-----G-C-CA-----A-A-----G-C-G-----T-G	4219
02 AG. DJ. DJ264	G-----G-C-CA-----A-A-----G-C-G-----T-G	4351
02 AG. NG. I8NG	G-----G-C-CA-----A-A-----G-C-G-----T-G	4352
02 AG. SN. MP1211	G-----G-C-CA-----A-A-----G-C-G-----T-G	4528
02 AG. SN. MP1213	G-----G-C-CA-----A-A-----G-C-G-----T-G	4200
03 AB. CM. M807	-----G-C-A-----A-A-----G-C-G-----T-G	4194
03 AB. RU. 98001	-----G-C-A-----A-A-----G-C-G-----T-G	4159
03 AB. RU. KALI53-2	-----G-C-A-----A-A-----G-C-G-----T-G	4373
04 CPX. GR. 97FVMY	-----G-C-A-----A-A-----G-C-G-----T-G	4772
04 CPX. GR. 97FVCH	-----G-C-A-----A-A-----G-C-G-----T-G	4772
04 CPX. CY. 94CY0323	A-----G-C-A-----G-----G-C-C-----C-----T--	5017
05 DF. BE. V11310	-----G-C-A-----G-----G-C-C-----C-----T--	5020
05 DF. BE. V1961	-----G-C-G-----G-----G-C-C-----C-----T--	4369
06 CPX. ML. 95ML84	-----G-C-G-----G-----G-C-C-----C-----T--	4385
06 CPX. AU. BPP90	-----G-C-A-----A-A-----G-C-C-----C-----T--	4360
06 CPX. ML. 95ML127	-----G-C-A-----A-A-----G-C-C-----C-----T--	4246
06 CPX. SN. 97SE1078	-----G-C-A-----A-A-----G-C-C-----C-----T--	5031
10 CD. BEL061	-----G-C-A-----A-A-----G-C-C-----C-----T--	5020
10 CD. BEL110	-----G-C-ITA-----C-A-A-----C-----A-----T--	5062
10 CD. BEL1071	-----G-C-ITA-----C-A-A-----C-----A-----T--	4371
11 CPX. CM. M818	-----G-C-GA-----T-----G-C-C-----C-----T--	4380
11 CPX. FR. MP1298	-----G-C-A-----A-A-----G-C-C-----C-----T--	5006
11 CPX. FR. MP1307	-----G-C-A-----A-A-----G-C-C-----C-----T--	4997
11 CPX. GR. GR17	-----G-C-A-----A-A-----G-C-C-----C-----T--	5054
CPZ. GA. SIVCPZGAB	A-----A-----A-----A-----G-C-C-----C-----T--	4305
CPZ. US. SIVCPZUS	-----G-C-A-----C-----G-CCT-A-----G-----C-----AG-G-AC-G	5064
CPZ. CM. SIVCPZCAM3	-----G-C-A-----T-----C-----GAG-GG-T--	5063
CPZ. CM. SIVCPZCAM5	-----G-C-A-----T-----C-----GAG-GG-T--	4417
Integrase	-----G-C-A-----T-----C-----GAG-GG-T--	4705
	-----G-C-TG-G-----C-A-T-G-----C-CA-----CC-AGAG-A-----T--G	4442
	-----I_T_K_.I_Q_N_F_R_V_Y_R_D_S_R_N_P_L_W_K_G_P_A_K_L_L_W_K_G_E_G_A_V_V_I_Q_D_N_S_D_I_K_	

HIV-1/SIVcpz
complete genomes

Table with columns for sample ID (e.g., B.FR.HXB2R), coordinates (e.g., 5131, 5088), and sequence (e.g., GTAGTCCAGAAAAGCAAGATGATAGGAGTTATGAAACACGATGGCCAGGTGATGATTTCTGCCAAGTAGACAGAGTAGGAT...TAGAACATGGAAGAGTTTACTTAAACACCATATGTATG). Includes a legend for Vif protein start and Pol p31 Integrase end.

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz complete genomes

B. FR. HXB2R	GCATACGAGAGAAAGACTGGCATTTGGGTCAGGAGTCTCCATAGAAATGGAGGAAAAGAGATATAGCACACAAGTAGACCCTGGAGCCCTGACATCATCTGATTACTTTGACTGTCTTTT	5385
CONSENSUS A	---ta-----a-----c-----t-g-t-g---t-t-ca-a-a-g---g-a-aagaga-t-gc-c---ag-gac-----a-----tctgtt-tt-t-t-----t-----	4801
A.UG.U455	---aa-c-----c-----t-g-----t-g-----c-----cta-aa-----aa-9at-----c-tg-----a-t-----tc-c-----ttt-9-c-----	4831
A.SE.UGSE8891	---a-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----a-a-----	4834
A.SE.TZSE8538	---AG-A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	4569
A.KE.Q2317	---A-----c-----t-g-----t-g-----c-----ctg-t-----a-----t-----c-ag-----c-g-----a-act-----a-c-----t-----t-----	4843
A.SE.UGSE7535	---A-----c-----t-g-----t-g-----c-----ct-ga-----a-----t-----c-g-----c-g-----g-a-----a-c-----c-----t-----c-----	4602
A.SE.SOSE7253	---A-----c-----t-g-----t-g-----c-----ct-ga-----a-----t-----c-g-----c-g-----g-a-----a-c-----c-----t-----c-----	4589
A.UG.92UG037	---A-----c-----t-g-----t-g-----c-----ct-ga-----a-----t-----c-g-----c-g-----g-a-----a-c-----c-----t-----c-----	4752
A.SE.SE8131	---A-----c-----t-g-----t-g-----c-----ct-ga-----a-----t-----c-g-----c-g-----g-a-----a-c-----c-----t-----c-----	4756
A2.CD.97CDKTB48	---A-----c-----t-g-----t-g-----c-----ct-ga-----a-----t-----c-g-----c-g-----g-a-----a-c-----c-----t-----c-----	4725
A2.CY.94CY017	---A-----c-----t-g-----t-g-----c-----ct-ga-----a-----t-----c-g-----c-g-----g-a-----a-c-----c-----t-----c-----	4742
B.AU.MBC925	g-t-----a-----c-----t-g-t-g---t-t-ca-a-a-g---g-a-aagaga-t-gc-c---ag-gac-----a-----tctgtt-tt-t-t-----t-----	5334
B.AU.MBC200	---a-----c-----t-g-----t-g-----c-----cta-aa-----aa-9at-----c-tg-----a-t-----tc-c-----ttt-9-c-----	5384
B.GA.OYI	---a-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	4928
B.GB.MANC	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	5331
B.GB.CAMI	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	5387
B.CN.RL42	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	4772
B.DE.D31	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	4931
B.DE.HAN2	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	4851
B.ES.89SP061	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	4943
B.NL.ACH320A	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	5387
B.TM.LM49	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	5386
B.US.DH123	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	5376
B.US.SF2	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	5392
B.US.NY5CG	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	5385
B.US.WCIPR18	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	4931
B.US.YU2	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	5383
B.US.JRCSF	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	5397
B.US.JRFL	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	4702
B.US.MN	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	5400
B.US.BCSG3	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	4932
B.US.WEAL160	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	5384
B.US.RF	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	4902
B.US.WR27	---A-----c-----t-g-----t-g-----c-----ctg-ga-----a-----t-----c-g-----c-g-----a-t-----c-c-----t-----y-a-----	5384
CONSENSUS C	g-a-----t-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4748
C.BR.92BR025	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4722
C.BW.96BW0402	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4851
C.BW.96BW1104	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4745
C.BW.96BW15C02	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4852
C.BW.96BW01B03	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4850
C.BW.96BW0502	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4878
C.BW.96BW16B01	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4851
C.BW.96BW1210	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4757
C.ET.ETH2220	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4777
C.IN.301999	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4764
C.IN.21068	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4736
C.IN.301905	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4736
C.IN.301904	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4736
C.IN.93JIN101	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	5358
C.IN.11246	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4727
D.UG.94UG114	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4740
D.CD.NDK	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4920
D.CD.ELI	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4931
D.KE.MB2059	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	5389
D.CD.Z226	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	5388
D.CD.84ZR085	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4907
F1.BR.93BR020	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4713
F1.FI.FIN9363	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4714
F1.BE.VI850	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4725
F1.FR.MF411	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4867
F2.CM.ME255C	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4884
F2.CM.ME257C	---T-----g-----c-----c-----a-----ga-t-g-sa-ga-ga-t-gc-----ca-g-----g-----g-----g-c-c-g-a-----g-----t-----a-----c-----t-----t-----	4596

B. FR. HXB2R	GCATACAGGAGAAAGAGACTGGCCATTTGGCTCAGGGAGTCTCCATAGAAATGGAGGAAAAGAGATATAGACACACAACTAGACCAACTAATTCATCTGTTACTTTGACTGTTTT	5385
G. BE. DRGB1	-----C-A-A-A-C-G-T-G-T-----T-C-C-GGA-----T-G-C-----C-T-T-A-T-----C-T-T-A-T-----	5342
G. NG. 92NG083	-----A-A-----G-----C-G-G-A-----T-G-----A-CAC-----T-G-----C-T-----T-----	4754
G. SE. SE6165	-----A-----T-----G-----C-G-G-A-----A-----T-----C-----G-----C-----T-----T-----	4782
G. FI. HH87932	-----A-----T-----G-----C-G-G-A-----A-----T-----C-----G-----C-----T-----T-----	4785
H. BE. VI991	-----T-----T-----AC-GG-A-----G-----T-----C-----G-----T-----T-----C-----T-----T-----	4774
H. BE. VI997	-----A-----C-----ACTG-A-----G-----G-----G-----G-----T-----C-----T-----T-----	4709
H. CF. 90CF056	-----A-----T-----C-----CGG-A-----G-----C-----A-----G-----G-----G-----A-----C-----T-----	4732
J. SE. SE91733	-----T-----C-----C-----G-G-A-----G-----C-----A-----G-----G-----G-----A-----C-----T-----	4700
J. SE. SE92809	-----T-----G-----A-----C-----CTG-A-----G-----C-----A-----A-----T-----G-----G-----A-----G-----T-----	4699
K. CM. MP535C	-----G-----A-----C-----CTG-A-----G-----C-----A-----A-----T-----G-----G-----G-----A-----A-----T-----	4581
K. CD. E07B11C	-----ATGC-----G-----TGAA-----A-----T-----G-----TAGT-----CAAT-----C-----GAG-----NA-----GA-----T-----AC-----AGGA-----G-----A-----C-----T-----AC-----	4577
O. CM. AN170	-----ATGC-----G-----TGAA-----A-----T-----G-----TAGT-----CAAT-----C-----GAG-----NA-----GA-----T-----AC-----AGGA-----G-----A-----C-----T-----AC-----	5440
O. CM. MP5180	-----ATGC-----G-----TGAA-----A-----T-----G-----TAGT-----CAAT-----C-----GAG-----NA-----GA-----T-----AC-----AGGA-----G-----A-----C-----T-----AC-----	5415
O. SN. MP1299	-----ATGC-----G-----TGAA-----A-----T-----G-----TAGT-----CAAT-----C-----GAG-----NA-----GA-----T-----AC-----AGGA-----G-----A-----C-----T-----AC-----	5439
O. SN. MP1300	-----ATGC-----G-----TGAA-----A-----T-----G-----TAGT-----CAAT-----C-----GAG-----NA-----GA-----T-----AC-----AGGA-----G-----A-----C-----T-----AC-----	5439
N. CM. YBF30	-----AACA-----CAGTCT-----C-A-A-T-R-----A-----ACT-----GA-A-----C-G-A-----C-AG-----T-----T-----A-----G-----A-----G-----A-----A-----T-----C-----	4979
N. CM. YBF106	-----AACA-----CAGTCT-----C-A-A-T-R-----A-----ACT-----GA-A-----C-G-A-----C-AG-----T-----T-----A-----G-----A-----G-----A-----A-----T-----C-----	4977
CONSENSUS 01	-----g-a-----AG-c-----a-----T-----g-----C-----GAGA-----a-----t-----a-----T-----C-----ca-----t-----g-----t-----a-----C-----a-----	5336
01 AE. TH. 93TH253	-----A-----AG-----C-----AG-----T-----T-----G-----G-----G-----G-----A-----T-----C-----C-----A-----A-----T-----G-----	5372
01 AE. TH. CM240	-----A-----AG-----C-----AG-----T-----T-----G-----G-----G-----G-----A-----T-----C-----C-----A-----A-----T-----G-----	4959
01 AE. CF. 90CF402	-----A-----AG-----C-----AG-----T-----T-----G-----G-----G-----G-----A-----T-----C-----C-----A-----A-----T-----G-----	5421
01 AE. TH. 92TH022	-----G-----AG-----A-----A-----T-----T-----G-----G-----G-----G-----A-----T-----C-----C-----A-----A-----T-----G-----	5348
01 AE. TH. 95TH1047	-----G-----AG-----A-----A-----T-----T-----G-----G-----G-----G-----A-----T-----C-----C-----A-----A-----T-----G-----	5348
01 AE. CF. 11697	-----G-----AG-----A-----A-----T-----T-----G-----G-----G-----G-----A-----T-----C-----C-----A-----A-----T-----G-----	5348
01 AE. TH. 93TH057	-----G-----AG-----A-----A-----T-----T-----G-----G-----G-----G-----A-----T-----C-----C-----A-----A-----T-----G-----	5348
01 AE. TH. 93TH065	-----G-----AG-----A-----A-----T-----T-----G-----G-----G-----G-----A-----T-----C-----C-----A-----A-----T-----G-----	5328
01 AE. CF. 4071	-----ta-----aa-----c-----c-----c-----a-----aagc-----g-----ga-----g-----t-----aa-----T-----C-----C-----G-----G-----C-----C-----A-----T-----	5321
CONSENSUS 02	-----T-----G-----T-----G-----T-----A-----G-----G-----G-----G-----A-----T-----C-----C-----G-----G-----C-----C-----T-----gac-----	4591
02 AG. SE. 7812	-----T-----G-----T-----G-----T-----A-----G-----G-----G-----G-----A-----T-----C-----C-----G-----G-----C-----C-----T-----gac-----	4757
02 AG. GH. G829	-----G-----G-----T-----G-----T-----A-----G-----G-----G-----G-----A-----T-----C-----C-----G-----G-----C-----C-----T-----ACT-----	4601
02 AG. DJ. DJ263	-----G-----G-----T-----G-----T-----A-----G-----G-----G-----G-----A-----T-----C-----C-----G-----G-----C-----C-----T-----ACT-----	4733
02 AG. DJ. DJ264	-----G-----G-----T-----G-----T-----A-----G-----G-----G-----G-----A-----T-----C-----C-----G-----G-----C-----C-----T-----ACT-----	4734
02 AG. NG. I8NG	-----G-----G-----T-----G-----T-----A-----G-----G-----G-----G-----A-----T-----C-----C-----G-----G-----C-----C-----T-----ACT-----	4910
02 AG. SN. MP1211	-----G-----G-----T-----G-----T-----A-----G-----G-----G-----G-----A-----T-----C-----C-----G-----G-----C-----C-----T-----ACT-----	4910
02 AG. SN. MP1213	-----G-----G-----T-----G-----T-----A-----G-----G-----G-----G-----A-----T-----C-----C-----G-----G-----C-----C-----T-----ACT-----	4582
02 AG. CM. M5807	-----G-----G-----T-----G-----T-----A-----G-----G-----G-----G-----A-----T-----C-----C-----G-----G-----C-----C-----T-----ACT-----	4576
03 AB. RU. 98001	-----G-----G-----T-----G-----T-----A-----G-----G-----G-----G-----A-----T-----C-----C-----G-----G-----C-----C-----T-----ACT-----	4541
03 AB. RU. KALI53-2	-----G-----G-----T-----G-----T-----A-----G-----G-----G-----G-----A-----T-----C-----C-----G-----G-----C-----C-----T-----ACT-----	4755
04 CPX. GR. 97FVWY	-----GC-----A-----AG-----A-----T-----G-----G-----R-----CT-----GA-----C-----T-----A-----C-----G-----G-----C-----C-----T-----T-----	5154
04 CPX. GR. 97FVCH	-----GC-----A-----AG-----A-----T-----G-----G-----R-----CT-----GA-----C-----T-----A-----C-----G-----G-----C-----C-----T-----T-----	5399
04 CPX. CY. 94CY0323	-----GC-----G-----CA-----C-----T-----G-----G-----G-----CT-----GA-----G-----T-----C-----G-----G-----T-----C-----G-----C-----T-----T-----	5402
05 DF. BE. VII1310	-----GC-----G-----CA-----C-----T-----G-----G-----G-----CT-----GA-----G-----T-----C-----G-----G-----T-----C-----G-----C-----T-----T-----	4751
05 DF. BE. VI961	-----GC-----G-----CA-----C-----T-----G-----G-----G-----CT-----GA-----G-----T-----C-----G-----G-----T-----C-----G-----C-----T-----T-----	4767
06 CPX. ML. 95ML84	-----AA-----G-----A-----C-----A-----C-----A-----ATTG-----A-----A-----G-----AG-----G-----G-----AGC-----G-----G-----T-----T-----	4742
06 CPX. AU. BPP90	-----AA-----G-----A-----C-----A-----C-----A-----ATTG-----A-----A-----G-----AG-----G-----G-----AGT-----G-----A-----A-----C-----T-----	4628
06 CPX. ML. 95ML127	-----AA-----G-----A-----C-----A-----C-----A-----ATTG-----A-----A-----G-----AG-----G-----G-----AGT-----G-----A-----A-----C-----T-----	5413
06 CPX. SN. 97SE1078	-----AA-----G-----A-----C-----A-----C-----A-----ATTG-----A-----A-----G-----AG-----G-----G-----AGT-----G-----A-----A-----C-----T-----	5402
10 CD. BEL061	-----AA-----G-----A-----C-----A-----C-----A-----ATTG-----A-----A-----G-----AG-----G-----G-----AGT-----G-----A-----A-----C-----T-----	5444
10 CD. BEL110	-----AA-----G-----A-----C-----A-----C-----A-----ATTG-----A-----A-----G-----AG-----G-----G-----AGT-----G-----A-----A-----C-----T-----	4764
10 CD. BEL107	-----AA-----G-----A-----C-----A-----C-----A-----ATTG-----A-----A-----G-----AG-----G-----G-----AGT-----G-----A-----A-----C-----T-----	4753
11 CPX. CM. MP818	-----AA-----G-----A-----C-----A-----C-----A-----ATTG-----A-----A-----G-----AG-----G-----G-----AGT-----G-----A-----A-----C-----T-----	4762
11 CPX. FR. MP1298	-----AA-----G-----A-----C-----A-----C-----A-----ATTG-----A-----A-----G-----AG-----G-----G-----AGT-----G-----A-----A-----C-----T-----	5388
11 CPX. FR. MP1307	-----AA-----G-----A-----C-----A-----C-----A-----ATTG-----A-----A-----G-----AG-----G-----G-----AGT-----G-----A-----A-----C-----T-----	5379
11 CPX. GR. GR17	-----AA-----G-----A-----C-----A-----C-----A-----ATTG-----A-----A-----G-----AG-----G-----G-----AGT-----G-----A-----A-----C-----T-----	5436
CPZ. GA. STVCPZGAB	-----AAAGCC-----G-----C-----A-----C-----T-----TC-----CT-----GGATCT-----GT-----G-----G-----T-----CTTCAC-----TAG-----T-----G-----C-----CTCTC-----A-----T-----T-----C-----	5449
CPZ. US. STVCPZUS	-----AAAGCC-----G-----C-----A-----C-----T-----TC-----CT-----GGATCT-----GT-----G-----G-----T-----CTTCAC-----TAG-----T-----G-----C-----CTCTC-----A-----T-----T-----C-----	5445
CPZ. CM. STVCPZCAM3	-----AAAGCC-----G-----C-----A-----C-----T-----TC-----CT-----GGATCT-----GT-----G-----G-----T-----CTTCAC-----TAG-----T-----G-----C-----CTCTC-----A-----T-----T-----C-----	4799
CPZ. CM. STVCPZCAM5	-----AAAGCC-----G-----C-----A-----C-----T-----TC-----CT-----GGATCT-----GT-----G-----G-----T-----CTTCAC-----TAG-----T-----G-----C-----CTCTC-----A-----T-----T-----C-----	5087
CPZ. CD. STVCPZANT	-----AAAGCC-----G-----C-----A-----C-----T-----TC-----CT-----GGATCT-----GT-----G-----G-----T-----CTTCAC-----TAG-----T-----G-----C-----CTCTC-----A-----T-----T-----C-----	4830
Vif	-----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_I_A_D_Q_L_I_H_L_Y_Y_F_D_C_F-----	

B. FR. HBX2R	TCAGACTCGTATAAAGGCCCTTATTAGGACACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5506
G. BE. DRGEL	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5463
G. NG. 92NG083	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5463
G. SE. SE6165	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4875
G. FI. HH87932	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4905
H. BE. VI991	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4906
H. BE. VI997	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4895
H. CF. 90CF056	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4830
J. SE. SE91733	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4853
J. SE. SE92809	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4821
K. CM. MP535C	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4820
K. CD. E0TB11C	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4702
O. CM. ANT70	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4698
O. CM. MP5180	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5561
O. SN. MP1299	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5526
O. SN. MP1300	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5560
N. CM. YBF30	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5500
N. CM. YBF106	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5098
CONSENSUS 01	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5457
01. AE. TH. 93TH253	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5493
01. AE. TH. CM240	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5080
01. AE. CF. 90CF402	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5542
01. AE. TH. 95TNIH047	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5469
01. AE. CF. 11697	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5455
01. AE. TH. 93TH057	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5459
01. AE. TH. 93TH065	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5449
01. AE. CF. 4071	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5442
CONSENSUS 02	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4711
02. AG. SE. 7812	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4878
02. AG. GH. G829	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4722
02. AG. DJ. DJ263	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4854
02. AG. DJ. DJ264	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4855
02. AG. NG. INBNG	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5031
02. AG. SN. MP1211	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4703
02. AG. SN. MP1213	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4697
02. AG. CM. MP807	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4862
03. AB. RU. 98001	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4876
03. AB. RU. KALI53-2	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5275
04. CPX. GR. 97FVBY	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5203
04. CPX. GR. 97FVCH	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5275
04. CPX. CY. 94CY0323	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4872
05. DF. BE. VI1310	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4888
05. DF. BE. VI961	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4863
06. CPX. ML. 95ML84	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4749
06. CPX. AU. BPP90	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5534
06. CPX. ML. 95ML127	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5523
06. CPX. SN. 97SE1078	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5565
10. CD. BEL061	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4885
10. CD. BEL110	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4874
10. CD. BEL1071	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4883
11. CPX. CM. MP818	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5509
11. CPX. FR. MP1298	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5500
11. CPX. FR. MP1307	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5557
11. CPX. GR. GR17	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4808
CPZ. GA. STVCPZGAB	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5570
CPZ. US. STVCPZUS	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5566
CPZ. CM. STVCPZCM3	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4920
CPZ. CM. STVCPZCM5	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	5208
CPZ. CD. STVCPZANT	-----A-C-G-A-A-A-A-G-A-A-C-G-G-G-T-T-TAGGACATAGTTACCCCTAGGTGTTAACTCAAGGAGGACATATAAAGGAGGACGTTAACACAGGACGACGTTAATA	4960
Vif	-----S-D-S-A-I-R-K-A-L-L-G-H-I-V-S-P-R-C-E-Y-Q-A-G-H-N-K-V-G-S-L-I-Q-Y-L-A-L-A-L-I-I-----T	

HIV-1/SIVcpz complete genomes

B.FR.HXB2R CAAAAGANTA...AAGCCACCTTTGCTAGTGTACAAACTCACAGAGGATAGATGGAAACACAGCCCGAGAGACCACAGGGCCACAGAGGACCCACAGGACCCACAGGACCTTTAGTGGAGTTTACAGC...
G.NE.DRCBL 5632
G.BE.92NG083 5889
G.SE.SE6165 5001
G.FI.HH87932 5029
H.BE.VI991 5032
H.BE.VI991 5021
H.CF.90CF056 4956
H.CF.90CF056 4979
J.SE.SE91733 4947
J.SE.SE92809 4946
K.CM.MP535C 4828
K.CD.F0PB11C 4824
O.CM.ANF70 5877
O.CM.MPF5180 5862
O.SN.MP1299 5886
O.SN.MP1300 5826
N.CM.YBF30 5224
N.CM.YBF106 5583
CONSENSUS 01 5619
01.AE.TH.93TH253 5200
01.AE.TH.93TH253 5595
01.AE.TH.92TH022 5591
01.AE.TH.95THU047 5851
01.AE.CF.11697 5855
01.AE.TH.93TH057 5575
01.AE.TH.93TH065 5368
01.AE.CF.4071 4837
CONSENSUS 02 5004
02.AG.SE.7812 4848
02.AG.GH.G829 4980
02.AG.DJ.DJ263 4981
02.AG.NG.IBNG 5157
02.AG.SN.MP1211 4829
02.AG.SN.MP1213 4823
02.AG.CM.MB807 4788
03.AB.RU.98001 5002
04.CPX.GR.97FVMY 5401
04.CPX.GR.97FVCH 5846
04.CPX.CY.94CY0323 5649
05.DF.BE.VI1310 4998
05.DF.BE.VI961 5014
06.CPX.ML.95ML84 4989
06.CPX.AU.BFP90 4875
06.CPX.WI.95WI127 5660
06.CPX.SN.97SEL078 5649
10.CD.BE1061 5691
10.CD.BE110 5011
10.CD.BE1071 5000
11.CPX.CM.MP818 5009
11.CPX.FR.MP1298 5635
11.CPX.FR.MP1307 5626
11.CPX.GR.GR17 5677
CPZ.GA.SIVCPZGAB 4934
CPZ.US.SIVCPZUS 5696
CPZ.CM.SIVCPZCAM3 5692
CPZ.CM.SIVCPZCAM5 5046
CPZ.CD.SIVCPZANT 5334
CPZ.CD.SIVCPZANT 5089

HIV-1/SIVcpz complete genomes

B. FR. HXB2R	GCTTAAGATGAGCTTTAGACATTTTCCTAGGATTTGGCTCCATGGCTTAGGGCAACATATCTATGAACTTATGGGACTTGGGAGGAGTGAAGCCATAATAAGAAATCTGCAACAACTGCTG	5762
G. BE. DRGB1	-----CC-----T-----A-----G-----T-----A-----C-----T-----A-----G-----T-----A-----A-----A-----A-----A-----A-----	5719
G. NG. 92NG083	-----C-----T-----A-----G-----A-----G-----A-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5131
G. SE. SE6165	-----C-----T-----A-----G-----A-----G-----A-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5139
G. FI. HH87932	-----C-----T-----A-----G-----A-----G-----A-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5162
H. BE. VI991	A-----G-----G-----A-----T-----A-----A-----G-----A-----G-----A-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5151
H. BE. VI997	-----G-----C-----G-----C-----A-----A-----G-----A-----G-----A-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5086
H. CF. 90CF056	-----A-----G-----G-----G-----A-----G-----A-----T-----C-----A-----A-----G-----A-----G-----A-----T-----T-----T-----T-----	5109
J. SE. SE91733	-----A-----G-----G-----G-----A-----G-----A-----T-----C-----A-----A-----G-----A-----G-----A-----T-----T-----T-----T-----	5077
J. SE. SE92809	-----A-----G-----G-----G-----A-----G-----A-----T-----C-----A-----A-----G-----A-----G-----A-----T-----T-----T-----T-----	5076
K. CM. MP535C	-----GA-----T-----A-----T-----A-----G-----A-----G-----A-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	4958
K. CD. F07B11C	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	4954
O. CM. ANI70	-----A-----AGCA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5792
O. CM. MP5180	-----G-----AGCA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5787
O. SN. MP1299	A-----A-----AGCA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5816
O. SN. MP1300	A-----A-----AGCA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5816
N. CM. YBF30	AT-----A-----GC-----C-----A-----A-----G-----A-----A-----C-----A-----A-----C-----A-----A-----C-----A-----A-----T-----A-----	5356
N. CM. YBF106	AT-----A-----GC-----C-----A-----A-----G-----A-----A-----C-----A-----A-----C-----A-----A-----C-----A-----A-----T-----A-----	5354
CONSENSUS 01	t-----A-----a-----c-----c-----c-----c-----t-----A-----G-----A-----G-----A-----G-----A-----G-----A-----G-----A-----G-----A-----	5713
01 AE. TH. 93TH253	-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	5749
01 AE. TH. CM240	-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	5330
01 AE. CF. 90CF402	-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	5798
01 AE. TH. 92TH022	-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	5725
01 AE. TH. 95TNLH047	-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	5722
01 AE. CF. 11697	-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	5711
01 AE. TH. 93TH057	-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	5715
01 AE. TH. 93TH065	-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	5705
01 AE. CF. 4071	-----A-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	5898
CONSENSUS 02	g-----t-----AC-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----	4967
02 AG. SE. 7812	-----A-----AC-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----	5134
02 AG. GH. G829	-----A-----AC-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----	4978
02 AG. DJ. DJ263	-----A-----AC-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----	5110
02 AG. DJ. DJ264	-----A-----AC-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----	5111
02 AG. NG. IENG	-----A-----AC-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----	5287
02 AG. SN. MP1211	-----A-----AC-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----	4959
02 AG. SN. MP1213	-----A-----AC-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----	4953
02 AG. CM. M5807	-----A-----AC-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----t-----	4818
03 AB. RU. 98001	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	5132
03 AB. RU. KALI53-2	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	5531
04 CPX. GR. 97FVWY	-----A-----G-----T-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----	5776
04 CPX. GR. 97FVCH	-----A-----G-----T-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----	5779
04 CPX. CY. 94CY0323	-----A-----G-----T-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----	5128
05 DF. BE. VI1310	-----A-----G-----T-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----	5144
05 DF. BE. VI961	-----A-----G-----T-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----	5119
06 CPX. ML. 95ML84	-----A-----G-----T-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----	5005
06 CPX. AU. BPP90	-----A-----G-----T-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----	5790
06 CPX. ML. 95ML127	-----A-----G-----T-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----	5779
06 CPX. SN. 97SE1078	-----A-----G-----T-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----A-----TC-----	5821
10 CD. BEL061	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	5141
10 CD. BEL110	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	5130
10 CD. BEL1071	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	5139
11 CPX. CM. MP818	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	5765
11 CPX. FR. MP1298	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	5807
11 CPX. FR. MP1307	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	5806
11 CPX. GR. GR17	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	5064
CPZ. GA. STVCPZGAB	AT-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5826
CPZ. US. STVCPZUS	A-----A-----AG-----A-----G-----T-----GC-----C-----C-----A-----A-----A-----A-----A-----A-----A-----A-----	5822
CPZ. CM. STVCPZCAM3	A-----A-----AG-----A-----G-----T-----GC-----C-----C-----A-----A-----A-----A-----A-----A-----A-----A-----	5176
CPZ. CM. STVCPZCAM5	A-----A-----AG-----A-----G-----T-----GC-----C-----C-----A-----A-----A-----A-----A-----A-----A-----A-----	5464
CPZ. CD. STVCPZANT	AA-----A-----A-----G-----A-----A-----C-----ACA-----C-----ACAT-----A-----ACA-----C-----T-----AA-----TTGGG-----A-----C-----A-----T-----A-----AG-----	5219
	-----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-----V-----E-----A-----I-----I-----R-----I-----L-----O-----Q-----L-----L-----	Vpr

HIV-1/SIVcpz complete genomes

Strain	Genomic Region	Sequence	Position
B.FR.HXB2R	TTTTATCCATT...TTTCAGAAATTGGGTTCGACATAGCAGAAATAGCGCTTT	ACTCGA.....CAG...AGGAGCAAGA...AATGGAGCCAGTAGATCCTAGACTAG	Tat cds start --> Vpr cds end <--
G.BE.DRCBL1AGG-A-A-T-G-G-G...T-G...T-G...C-G-G-A	5855
G.NG.92NG083A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5814
G.SE.SE6165A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5223
G.FI.HH87932A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5251
H.BE.VI991A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5254
H.BE.VI997A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5243
H.CF.90CF056A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5178
J.SE.SE91733A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5201
J.SE.SE92809A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5169
K.CM.MP535CA-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5168
K.CD.FORB11CA-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5050
O.CM.ANF70A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5046
O.CM.MPF5180A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5912
O.SN.MP1299A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5896
O.SN.MP1300A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5920
N.CM.YBF30A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5445
N.CM.YBF106A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5443
CONSENSUS 01A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5805
01.AE.TH.93TH253A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5842
01.AE.TH.CM240A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5422
01.AE.CF.90CF402A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5890
01.AE.TH.92TH022A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5817
01.AE.TH.95THH047A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5814
01.AE.CF.11697A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5803
01.AE.TH.93TH057A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5807
01.AE.TH.93TH065A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5797
01.AE.CF.4071A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5788
CONSENSUS 02A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5059
02.AG.SE.7812A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5226
02.AG.GH.G829A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5070
02.AG.DJ.DJ263A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5202
02.AG.DJ.DJ264A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5203
02.AG.NG.IBNGA-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5379
02.AG.SN.MP1211A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5051
02.AG.SN.MP1213A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5045
02.AG.CM.MB807A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5010
03.AB.RU.98001A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5224
03.AB.RU.KALI53-2A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5623
04.CPX.GR.97PVMYA-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5878
04.CPX.GR.97PVCHA-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5884
04.CPX.CV.94CV0323A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5224
05.DF.BE.VI1310A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5236
05.DF.BE.VI961A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5211
06.CPX.ML.95ML84A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5097
06.CPX.AU.BFP90A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5882
06.CPX.ML.95ML127A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5871
06.CPX.SN.97SEL078A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5913
10.CD.BEL061A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5233
10.CD.BEL110A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5222
10.CD.BEL1071A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5231
11.CPX.CM.MP818A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5857
11.CPX.FR.MP1298A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5848
11.CPX.FR.MP1307A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5899
11.CPX.GR.GR17A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5156
CPZ.GA.SIVCPZGABA-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5918
CPZ.US.SIVCPZUSA-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5926
CPZ.CM.SIVCPZCAM3A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5265
CPZ.CM.SIVCPZCAM5A-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5553
CPZ.CD.SIVCPZANTA-A-T-G-G-G...T-G...T-G...T-G...C-G-G-A	5287

F_I_H_#_F_R_I_G_C_R_H_S_R_I_G_V_...T_R_...O...R_R_A_R_...N_G_A_S_R_S_\$__Vpr protein end
 frameshift in HXB2
 Tat protein start M_E_P_V_D_P_R_L_ Tat



Accession	Sequence	Rev cds start	Rev protein start
B.FR.HXB2R	AGCCCTGGAGCATCCAGGAGTCCAGCTAAACTGCTTGTACCAATTGCTATTGTAATAAAGTGTTCCTTCCATTGCGCAAGTTTGTTCATCAAAAGACCTTAGGCAATCTCTATGGCAGGAGGAGCGG	5985	Tat
G.NE.DRCBL	---	5941	Rev
G.NG.92NG083	---	5353	
G.SE.SE6165	---	5381	
G.FI.HH87932	---	5384	
H.BE.VI991	---	5373	
H.BE.VI997	---	5308	
H.CF.90CF056	---	5331	
J.SE.SE91733	---	5299	
J.SE.SE92809	---	5298	
K.CM.MP535C	---	5180	
K.CD.F0PBL1C	---	5176	
O.CM.AN70	---	6042	
O.CM.MP5180	---	6026	
O.SN.MP1299	---	6050	
O.SN.MP1300	---	6050	
N.CM.YBF30	---	5575	
N.CM.YBF106	---	5573	
CONSENSUS 01	---	5935	
01.AE.TH.93TH253	---	5972	
01.AE.TH.CM240	---	5952	
01.AE.CF.90CF402	---	6020	
01.AE.TH.92TH022	---	5947	
01.AE.TH.95THH047	---	5944	
01.AE.CF.11697	---	5933	
01.AE.TH.93TH057	---	5937	
01.AE.TH.93TH065	---	5927	
01.AE.CF.4071	---	5918	
CONSENSUS 02	---	5189	
02.AG.SE.7812	---	5356	
02.AG.GH.G829	---	5200	
02.AG.DJ.DJ263	---	5332	
02.AG.DJ.DJ264	---	5333	
02.AG.NG.IBNG	---	5508	
02.AG.SN.MP1211	---	5181	
02.AG.SN.MP1213	---	5175	
02.AG.CM.MB807	---	5140	
03.AB.RU.98001	---	5554	
03.AB.RU.KAL153-2	---	5753	
04.CPX.GR.97PVMY	---	6008	
04.CPX.GR.97PVCH	---	6014	
04.CPX.CY.94CY0323	---	5354	
05.DF.BE.VI1310	---	5366	
05.DF.BE.VI961	---	5341	
06.CPX.ML.95ML84	---	5227	
06.CPX.AU.BFP90	---	6012	
06.CPX.ML.95ML127	---	6001	
06.CPX.SN.97SEL078	---	6043	
10.CD.BEL061	---	5362	
10.CD.BEL110	---	5352	
10.CD.BEL1071	---	5361	
11.CPX.CM.MP818	---	5987	
11.CPX.FR.MP1298	---	5978	
11.CPX.FR.MP1307	---	6029	
11.CPX.GR.GR17	---	5286	
CPZ.GA.SIVCPZGAB	---	6048	
CPZ.US.SIVCPZUS	---	6056	
CPZ.CM.SIVCPZCAM3	---	5395	
CPZ.CM.SIVCPZCAM5	---	5683	
CPZ.CD.SIVCPZANT	---	5417	

HIV-1/SIVcpz complete genomes

Table with columns for sequence identifiers (e.g., B.FR.HXB2R, G.BE.DRCBL), sequence alignments (GAGA...), and amino acid translations (Tat Rev exon end // intron start). The table lists 58 different HIV-1/SIVcpz strains and their corresponding genomic sequences and protein products.

HIV-1/SIVcpz complete genomes

Table with columns for accession numbers (e.g., B.FR.HXB2R, G.BE.DRCBL, G.NG.92NG083), sequence identifiers (e.g., ATAGTAGCAATAGTAGTACGATATAGCAATATAGCAATAGTGTGTCGATAGTAACTATAGAA), and amino acid sequences (e.g., A-T-TGC-A--T-GGCT-A--T-GC-TCC--GCCA-A-TGT-GG-GTA-G-AT-TATAGAA). The table lists 62 HIV-1/SIVcpz complete genomes.

HIV-1/SIVcpz complete genomes

	Env	cds	start	->
B.FR.HXB2R	CTAATAGAAAGCAGACAGACAGTGGCA	ATGAGAGTGAAGGAGAAA-ATCAGCACTTGTGGAGATGGGGGTG
CONSENSUS A	A-aga	-t-g--cacagagA-t-gt-aaca-ta-t
A.UG.U455	A-G	-t-g--TACAGAGGA-C
A.SE.UGSE8891	A-GT	-t-g--TACAGATGA-GG-AA
A.SE.TZSE8538	A-G	-t-g--TACAGAGGA-C-C
A.KE.Q2317	A-G	-t-g--TACAGAGGA-T
A.SE.UGSE6594	A-G	-t-g--TACAGAGGA-T
A.KE.Q2317	A-G	-t-g--TACAGAGGA-T
A.SE.UGSE7535	A-GC	-t-g--TACAGAGGA-T
A.SE.SOSE7253	A-G	-t-g--TACAGAGGA-T
A.UG.92UG037	A-G	-t-g--TACAGAGGA-T
A.SE.SE8131	A-G	-t-g--TACAGAGGA-T
A2.CD.97CDKTB48	A-GT	-t-g--TACAGAGGA-G-AA-A-G
A3.CY.94CY017	A-GT	-t-g--TACAGAGGA-C
CONSENSUS B	a-gdag-a-g	-g-t-a-g-atacagagAa-a--cac-tg-??
B.AU.MBC925	A-GT	-t-g--TACAGAGGA
B.AU.MBC200	A-G	-t-g--TACAGAGGA-G
B.GA.OYI	A-G	-t-g--TACAGAGGA-G
B.GB.MANC	A-GG	-t-g--TACAGAGGA-G
B.GB.CAMI	A-G	-t-g--TACAGAGGA-G
B.CN.RL42	A-G	-t-g--TACAGAGGA-G
B.DE.D31	A-G	-t-g--TACAGAGGA-G
B.DE.HAN2	A-G	-t-g--TACAGAGGA-G
B.ES.89SEP061	A-GC	-t-g--TACAGAGGA-G
B.NL.ACH320A	A-G	-t-g--TACAGAGGA-G
B.ITW.LM49	A-G	-t-g--TACAGAGGA-G
B.US.DHL23	A-G	-t-g--TACAGAGGA-G
B.US.SF2	A-G	-t-g--TACAGAGGA-G
B.US.NY5CG	A-G	-t-g--TACAGAGGA-G
B.US.WC1PR18	A-G	-t-g--TACAGAGGA-G
B.US.YU2	A-G	-t-g--TACAGAGGA-G
B.US.JRCSF	A-G	-t-g--TACAGAGGA-G
B.US.JRFL	A-G	-t-g--TACAGAGGA-G
B.US.MN	A-GC	-t-g--TACAGAGGA-G
B.US.BCSG3	A-G	-t-g--TACAGAGGA-G
B.US.WEAU160	A-GTA	-t-g--TACAGAGGA-G
B.US.RF	A-G	-t-g--TACAGAGGA-G
B.US.WR27	A-G	-t-g--TACAGAGGA-G
CONSENSUS C	Att-5g-a--gc	-g-t-ag-g--cactgagAat-atcaa-A-G
C.BR.92BR025	A-T	-g-t-ag-g--cactgagA-GA-A-A-G
C.BW.96BW0402	A-T	-g-t-ag-g--cactgagA-GA-A-A-G
C.BW.96BW1104	A-T	-g-t-ag-g--cactgagA-GA-A-A-G
C.BW.96BW15C02	A-T	-g-t-ag-g--cactgagA-GA-A-A-G
C.BW.96BW01B03	A-T	-g-t-ag-g--cactgagA-GA-A-A-G
C.BW.96BW0502	A-T	-g-t-ag-g--cactgagA-GA-A-A-G
C.BW.96BW16B01	A-T	-g-t-ag-g--cactgagA-GA-A-A-G
C.BW.96BW1210	A-T	-g-t-ag-g--cactgagA-GA-A-A-G
C.ET.FTH2220	ACT	-g-t-ag-g--cactgagA-GA-A-A-G
C.IN.301999	A-T	-g-t-ag-g--cactgagA-GA-A-A-G
C.IN.21068	A-T	-g-t-ag-g--cactgagA-GA-A-A-G
C.IN.301905	A-T	-g-t-ag-g--cactgagA-GA-A-A-G
C.IN.301904	A-T	-g-t-ag-g--cactgagA-GA-A-A-G
C.IN.93IN101	A-T	-g-t-ag-g--cactgagA-GA-A-A-G
C.IN.11246	A-T	-g-t-ag-g--cactgagA-GA-A-A-G
D.UG.94UG114	A-G	-t-g--TACAGAGGA-GA-A-A-G
D.CD.NDK	A-G	-t-g--TACAGAGGA-GA-A-A-G
D.CD.ELI	A-G	-t-g--TACAGAGGA-GA-A-A-G
D.KE.MB2059	A-G	-t-g--TACAGAGGA-GA-A-A-G
D.CD.2226	A-G	-t-g--TACAGAGGA-GA-A-A-G
D.CD.84ZR085	A-G	-t-g--TACAGAGGA-GA-A-A-G
F1.BR.93BR020	A-GC	-t-g--TACAGAGGA-GA-A-A-G
F1.FI.FIN9363	A-G	-t-g--TACAGAGGA-GA-A-A-G
F1.BE.VI850	A-G	-t-g--TACAGAGGA-GA-A-A-G
F1.FR.MF411	A-G	-t-g--TACAGAGGA-GA-A-A-G
F2.CM.MF255C	A-G	-t-g--TACAGAGGA-GA-A-A-G
F2.CM.MF257C	A-G	-t-g--TACAGAGGA-GA-A-A-G

HIV-1/SIVcpz Complete Genomes

B. FR. HXB2R	CTAATAGAAAGCAGACAGTGGCA	CTAATAGAAAGCAGACAGTGGCA	ATGAGAGTGAAGGAGA	AAATATCAGCACTTTGGAGATGGGGGTG	6268
G. BE. DRCEBL	A-G	A-G	-G-TACAGAGG	A-GG-A-A	6232
G. NG. 92NG083	A-G	A-G	-G-TACAGAGG	A-GG-A-A	5834
G. SE. SE6165	A-G	A-G	-C-G-TACAGAGG	A-GGT-A-A	5672
G. FI. HH87932	A-G	A-G	-G-TACAGAGG	A-GG-A-A	5675
H. BE. VI991	A-G	A-G	-C-TACAGAGG	A-CAAG	5647
H. BE. VI997	A-C	A-C	-T-AG	CA-A-G	5574
H. SE. 90CF056	A-G	A-G	-C-TACAGAGG	A-CAAG-A	5609
J. CE. SE91733	A-G	A-G	-C-TACAGAGG	A-CAAG-A	5589
J. SE. SE92809	A-G	A-G	-C-TACAGAGG	A-CAAG-A	5588
K. CM. MP535C	A-G	A-G	-G-TGACAGAG	A-GG-ACA-G	5468
K. CD. FOPE11C	A-G	A-G	-C-G-TACAGAGG	A-GG-ACA-G	5464
O. CM. AN70	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6320
O. CM. MP5180	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6307
O. SN. MP1299	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6328
O. SN. MP1300	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6328
N. CM. YBF30	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5871
N. CM. YBF106	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5844
CONSENSUS 01	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6220
01 AE. TH. 93TH253	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6256
01 AE. TH. CM240	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5837
01 AE. CF. 90CF402	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6305
01 AE. TH. 92TH022	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6232
01 AE. TH. 95THH047	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6229
01 AE. CF. 11697	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6318
01 AE. TH. 93TH057	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6222
01 AE. TH. 93TH065	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6212
01 AE. CF. 4071	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6201
CONSENSUS 02	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5477
02 AG. SE. 7812	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5477
02 AG. GH. G829	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5491
02 AG. DJ. DJ263	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5623
02 AG. DJ. DJ264	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5624
02 AG. NG. I8NG	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5799
02 AG. SN. MP1211	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5477
02 AG. SN. MP1213	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5466
02 AG. CM. ME807	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5440
03 AB. RU. 98001	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5826
04 CPX. GR. 97PVMY	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6025
04 CPX. GR. 97PVCH	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6302
04 CPX. CY. 94CY0323	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6308
05 DF. BE. V11310	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5639
05 DF. BE. V1961	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5657
06 CPX. ML. 95ML84	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5632
06 CPX. AU. BFP90	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5518
06 CPX. ML. 95ML127	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6305
06 CPX. SN. 97SEL078	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6292
10 CD. BEL061	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6334
10 CD. BEL110	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5855
10 CD. BEL1071	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5855
11 CPX. CM. MP818	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6277
11 CPX. FR. MP1298	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6269
11 CPX. FR. MP1307	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6322
11 CPX. GR. GR17	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5575
CPZ. GA. SIVCPZGAB	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6336
CPZ. US. SIVCPZUS	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	6352
CPZ. CM. SIVCPZCAM3	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5685
CPZ. CM. SIVCPZCAM5	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5973
CPZ. CD. SIVCPZANT	A-G	A-G	-A-CA-T-A	A-CA-A-G-A	5707
B. FR. HXB2R	CTAATAGAAAGCAGACAGTGGCA	CTAATAGAAAGCAGACAGTGGCA	ATGAGAGTGAAGGAGA	AAATATCAGCACTTTGGAGATGGGGGTG	6268
Vpu	L_I_D_R	L_I_E_R_A_E_D_S_G	N_E_S_E_E	#_I_S_A_L_V_E_M_G_V	Env
Env	Env protein start	M_R_V_K_E	Y_Q_H_L_W_R_W_G_W		

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz complete genomes

B. FR. HXB2R	GATGCTAAAGCATATGATACAGAGGTACATATAATTGTTGGGCGCACACATGCTGTACCCACAGACCCCAACCCACAGAGAAATTTTAAACATCTGGAAAAATGACA	6522
CONSENSUS A	----ta-----g-tac-----sgtg-----c-----l-t-a-----c-----c-----c-----a-----a-----c-----a-----at-----g-----t-----aac-----	5743
A.UG.U455	-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5967
A.SE.UGSE8891	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5717
A.SE.TZSE8538	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5742
A.SE.UGSE6594	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5701
A.KE.Q2317	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5986
A.SE.UGSE7535	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5745
A.SE.SOSE7253	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5735
A.UG.92UG037	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5898
A.SE.SE131	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5873
A2.CD.97CDKTB48	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5865
A2.CY.94CY017	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5882
CONSENSUS B	----t-t-a-g-ata-gat-c-g-a-----t-----t-----ca-t-c-g-----c-----ca-----ag-gtat-ggaa-t-g-----aa-t-----a-c-----t-----a-c-----	6457
B.AU.MBC925	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6529
B.AU.MBC200	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6517
B.GA.OYI	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6061
B.GB.MANC	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6464
B.GB.CAMI	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6520
B.CN.RL42	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5905
B.DE.D31	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6064
B.DE.HAN2	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5984
B.ES.89SP061	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6079
B.NL.ACH320A	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6523
B.TW.LM49	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6519
B.US.DH123	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6509
B.US.SF2	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6527
B.US.NY5CG	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6474
B.US.WCIPR18	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6067
B.US.YU2	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6519
B.US.JRC5F	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6530
B.US.JRFL	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5835
B.US.MN	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6533
B.US.BCSG3	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6042
B.US.WEAU160	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6517
B.US.RF	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6037
B.US.WR27	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5984
CONSENSUS C	----t-t-a--t--agaaa-a-g-----c--ct-aca-----a-c-----ac-----a-----a-----a-----a-----a-----a-----g-----t-----g-----g-----a-----a-----c-----	5890
C.BR.92BR025	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5877
C.BW.96BW0402	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5998
C.BW.96BW1104	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5897
C.BW.96BW15C02	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5991
C.BW.96BW01B03	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6023
C.BW.96BW0502	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5996
C.BW.96BW16B01	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5909
C.ET.ETH2220	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5931
C.IN.301999	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5897
C.IN.21068	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5869
C.IN.301905	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5899
C.IN.301904	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5889
C.IN.93JN101	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6491
C.IN.11246	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5854
D.UG.94UG114	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5873
D.CD.NDK	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6053
D.CD.ELI	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6064
D.KE.MB2059	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6522
D.CD.2226	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6521
D.CD.84ZR085	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	6043
F1.BR.93BR020	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5854
F1.FT.FIN9363	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5854
F1.BE.VI850	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5968
F1.FR.MF411	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5125
F2.CM.MF255C	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5723
F2.CM.MF257C	-----g-----g-----g-----a-----g-----t-----c-----t-----g-----c-----t-----g-----c-----t-----g-----g-----g-----g-----a-----a-----	5729

B. FR. HXB2R	... AAGRAATGATACTAATACCAATAGTACTAGC	... GGGAAATGATTAATGAGAAAGAGAGATAAAAAATGCTCTTTCAATATCAGACACAAGCATAAGAGCTAAG	61728
CONSENSUS A	... ??????atlaa?aaacc-acagaca??	... aatattccccacac??a?g?gaa-aaa-g-a-c-c---a-g-c---gnaac-agg-a--g	5920
A.UG_U0455	... CCAI--AC--TC--GTC--CC--AT--GCTC--GTC--CCTAAT	... AAT-TC-CAGATGGA-T--GG-A--A-G--T--	6170
A.SE.UGSE8891	... T--A-GTC-----CAVAT	... ATC-CC-GTGAAC...GC...A--G--	5911
A.SE.UGSE8538	... GT--C---CG--GGTC	... TAT-ATG-C-CC-AC...-TGCA--A-G--	5945
A.SE.UGSE6594	... CAAG-GTC--C--G-G-A--	... TC-AC-C-C-GAC...-G-A--	5889
A.SE.UGSE7535	... ATC--GTC--G--AC--A--ATAAT	... A-T-TC-GC-AC-GC...-CG-AG-C--G-G--	6168
A.SE.SOSE7253	... A-GCC...CG-A-CCG	... AATGTC-AC-C-C...-T-T-GA--G-G--	5960
A.UG.92UG037	... TC-CCA--A--TC-----C--TC--	... AAT-CTCAG-T-ACAT--G--A--	5920
A.SE.SEB131	... GTCA-G-C-C--C-C-C-TGCC	... AACCTC-C-GAT-CC-T...-A-G-G--	6079
A2_CD_97CDKTB48	... CACC-C-C-C-GAA	... CCGC-GA-C--A--	6044
A2_CV_94CY017	... TC-CGAG--CC--G--C	... CCCA-GCCCATTT--T-A--	6085
CONSENSUS B	... aagaatgctaataac-tagtagtagt????????????????	... aatattccccacac??a?g?gaa-aaa-g-a-c-c---a-g-c---gnaac-agg-a--g	6069
B.AU.MBC925	... GA--G--C-AT-CC-C--TAGCGACACAGTAGGAGAG	... AGA--AAG-G---C---A---G---	6156
B.AU.MBC200	... CGA--G--C-AT-CC-A--TAGC	... A-G--G---TGA--AG---T-C---A--	6738
B.GA.OYI	... CCGA--G--C-AT-CC-CA--TAGT	... T-GA-C---AT--C-C---A--	6285
B.GB.MANC	... C--A-C-AT-CC-C--ATAATCCGCTAGTGGAGT	... T-G-GC---AT--G-G--A--	6685
B.GB.CAMI	... A--A-C--GG-CC-A--TGATTTGGACAGG	... A-GA-GGAGA--A-AT--G--C---CT--A--	6732
B.CN.RL42	... GA--G--C-AT-CC-C--TAGT	... AC-TG-CAGG-GAAAT--A--C--	6096
B.DE.D31	... GA--G--C-AT-CC-A--TAGT	... T-G-G---G---G--G--	6273
B.DE.HAN2	... GGGG--GA--C--A--CC-A--TAGTCTACTAACCCACTAGTAGTA	... CG-GGA--G-G--A--A--C--	6178
B.ES.89SP061	... GCGA--G--C-AT-CC-C--TAGC	... ACA-TGGA---GTT--A--	6266
B.NL.ACH320A	... CCAATG--A--G-ATG--C-A-C-C-TT	... A-TCA---AT--G---G---G-C---A--G-A--	6716
B.US.DH123	... GA-GG--C-AT-TTG-AG-A-G-TACTATAAATCAAT	... T--A-GAA...GAAATA---A---T--A--	6718
B.US.SF2	... CGA--G--C-AT-CC-A--TAGT	... A-G---T---G---C---T---AA--	6730
B.US.NY5CG	... CGA--G--C-TATGCC-A-G--TAGT	... AGCTGG-AA-AG---T---A--	6668
B.US.WCIPR18	... TACTGTAC--AA-G--C-ATGCC-A--T	... T--G---T---A---G---C---T---A--	6288
B.US.YU2	... CGA--G--C-AT-CC-C--TAGC	... T-GA-C---A---G---C---A--	6722
B.US.JRCSF	... A-G--C-AT-CC-C--TAGT	... A-G---C---G---C---A--	6730
B.US.JRFL	... A-G--C-AT-CC-C--TAGT	... A-G---C---G---C---A--	6035
B.US.MN	... GGA--C--AT-CC-A--TACTGTAAACAATAGTAATAGC-A-G-CA	... AA-GG---A---G---C---A--	6757
B.US.BCSG3	... GA--G--C-AT-CC-C--CTAATCCCTAGTGTAGC	... T--A--A--G--A--G--	6254
B.US.WEAU160	... GA--GAG-C-AT-CC-C--TAGTGG	... G-G-C---CA-G---G---C---A--	6738
B.US.FR	... TTGA--GG--C-ATGCC-C--TAGC	... G-C-A-G---C---A---G---C---T---A--	6246
B.US.WR27	... TGGG-A-GC--T-G--GC--AGAATACTACTACCAATAGTGTAGT	... GA-C---A---G---C---A---G---C---T---A--	6118
CONSENSUS C	... ??????aataacaa?aaataata?e?????????aataag??????????	... gta-c-a-gaaa-a-g-a-eg	6074
C.BW.92BR025	... AC--T-G-CIA--AC--A-CI	... AT-AC-----GTTTCC....	6216
C.BW.96BW0402	... GT--C-C-A-GG--G-A--T	... GTTCCC....	6082
C.BW.96BW1104	... TGA--G--C-G-T--AC-A-ATACC	... CC-----CT--A--	6200
C.BW.96BW15C02	... A-GG-C-TA-C-GA-ACCTA-CAG	... GA--A-----	6177
C.BW.96BW0502	... C-C-C-AT-T-A-C-TG	... ATTGAT-AC---GTA-T-G---A-G	6223
C.BW.96BW16B01	... G-A-GGG--G-AT-CC-C--TGG-TAAT	... AAGACCTATGAT-AC...AGTGGG--A-G	6211
C.BW.96BW1210	... GT--CC-GA---A-GCTACC	... TAC-AT-AT...AAA-T-T---A	6106
C.ET.ETH2220	... C-AC--AGGT--A--A--T	... AAT-TGCC...T-AT-A-G--	6131
C.IN.301999	... GCT-CC-A-GT--G-T-CC--A-G-TAAT	... CT-CT-GT...A-TGGG-A--A	6124
C.IN.21068	... GT-C-G-A-GG--C-C-A-G-C	... TAC-ATGA...GCAT--G---A	6072
C.IN.301905	... A-A-A-G--GT-CCCAC-A-GAGC	... TAC-CA...GC-T-...A	6042
C.IN.301904	... ACCCAA--GGT-CCTAC-A-GATGAG	... GTG-AGCGTG-AG--A--G---A	6066
D.UG.94UG114	... TC--G--CC-A-CC-CA	... A-C-AT-A...ATGG-AGCGTG-AG--A--G---A	6888
D.CD.ELI	... TT--GGA--C-A-GGC--T-TGG-G-AC-AT	... TC-CT-CA...GA--G---G---A--	6049
D.KE.MB2059	... GTGATGAG--GT--CA-TG-AG-ATAAT	... AATGTC-CAGAGGAATA-G...G	6264
D.CD.84ZR085	... GA-CAGC--G-AT--TC-A-GAC-CT	... CCG-C--G-A-CC--A--G---G---A--	6121
F1.FI.FIN9363	... CC--C--GAC--GAC-C-T-TCTGACCAAGC	... C-AGGA--TCC-AG-CA-C	6065
F1.BE.VI850	... CCA--AC...CCGGAAGTCA	... A-T-CCC--AGGA--ACC-G-CA-C	6063
F1.FR.MF411	... ACC-CT-C-C-GC--CCGGAAGTCA	... A-T-CCC--AGGA--ACC-G-CA-C	5934
F2.CM.MF255C	... AATGTC-CCAGCTTGG--GGAAG--CRAAG-GGA	... GT-C---A---G---C---A--	5926
F2.CM.MF257C	... AAT--T-CCACCTTGG--C-T--GT--CC-T	... TCG-AG--A--G---T---A--C---GAG--A---A---A---A---	5941

HIV-1/SIVcpz Complete Genomes

B. FR. HXB2R
G. BE. DRGB1
G. NG. 92NG083
G. SE. SE6165
G. FI. HH87932
H. BE. VI991
H. BE. VI997
H. CF. 90CF056
J. SE. SE91733
J. SE. SE92809
K. CM. MP535C
K. CD. E07B11C
O. CM. AN170
O. CM. MP5180
O. SN. MP1299
O. SN. MP1300
N. CM. YBF30
N. CM. YBF106
CONSENSUS 01
01 AE. TH. 93TH253
01 AE. TH. CM240
01 AE. TH. 92CF402
01 AE. TH. 92TH022
01 AE. TH. 95TNH047
01 AE. TH. 93TH057
01 AE. TH. 93TH065
01 AE. CF. 4071
CONSENSUS 02
02 AG. SE. 7812
02 AG. GH. G829
02 AG. DJ. DJ264
02 AG. NG. INBNG
02 AG. SN. MP1211
02 AG. SN. MP1213
02 AG. CM. M807
03 AB. RU. 98001
03 AB. RU. KALI53-2
04 CPX. GR. 97PVMY
04 CPX. GR. 97PVCH
04 CPX. CY. 94CY0323
05 DF. BE. V11310
06 DF. BE. V1961
06 CPX. ML. 95ML84
06 CPX. AU. BPP90
06 CPX. ML. 95ML127
06 CPX. SN. 97SE1078
10 CD. BEL061
10 CD. BEL110
10 CD. BEL107
11 CPX. CM. M818
11 CPX. FR. MP1298
11 CPX. FR. MP1307
11 CPX. GR. GR17
CPZ. GA. STVCPZGAB
CPZ. US. STVCPZUS
CPZ. CM. STVCPZCAM3
CPZ. CM. STVCPZCAM5
CPZ. CD. STVCPZANT

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz Complete Genomes

305

<p>B. FR. HXB2R G. BE. DRCEI G. NG. 92NG083 G. SE. SE6165 H. FI. HH87932 H. BE. VI991 H. BE. VI997 H. CF. 90CF056 J. SE. SE91733 J. SE. SE92809 K. CM. MP535C K. CD. E0TB11C O. CM. AN170 O. CM. MP5180 O. SN. MP1299 O. SN. MP1300 N. CM. YBF30 N. CM. YBF106 CONSENSUS 01 01 AE. TH. 93TH253 01 AE. TH. CM240 01 AE. CF. 90CF402 01 AE. TH. 92TH022 01 AE. TH. 95TH047 01 AE. CF. 11697 01 AE. TH. 93TH057 01 AE. TH. 93TH065 01 AE. CF. 4071 CONSENSUS 02 02 AG. SE. 7812 02 AG. GH. G829 02 AG. DJ. DJ263 02 AG. DJ. DJ264 02 AG. NG. I8NG 02 AG. SN. MP1211 02 AG. SN. MP1213 03 AB. CM. M807 03 AB. RU. 98001 03 AB. RU. KALI53-2 04 CPX. GR. 97PVMY 04 CPX. GR. 97PVCH 04 CPX. CY. 94CY0323 05 DF. BE. VI1310 05 DF. BE. VI961 06 CPX. ML. 95ML84 06 CPX. AU. BPP90 06 CPX. ML. 95ML127 06 CPX. SN. 97SE1078 10 CD. BEL061 10 CD. BEL101 10 CD. BEL110 11 CPX. CM. M818 11 CPX. FR. MP1298 11 CPX. FR. MP1307 11 CPX. GR. GR17 CPZ. GA. SIVCPZGAB CPZ. US. SIVCPZUS CPZ. CM. SIVCPZCAM3 CPZ. CM. SIVCPZCAM5 CPZ. CD. SIVCPZANT</p>	<pre> CAAAGTTGTAACACCTCAGTCATT...ACACAGGCGCTGCCAAAGGTATCCCTTTGAG...CCAAATCCCATACATTATTGTGCCCCGGCTGGTTTTCGGATTCATAAAAATAAATAAAGACCGTTCAATGG T-A-----TGT--AC-----T-----T-----A-T-----AAT-----C-----T-A-----T-A-----G-----GTGG-----A-----A----- --CA-----TGT--AC-----A-----T-----T-----A-----A-----T-----T-----G-----GGG-----GA-----A----- T-A-----TGT--AC-----C-----A-----T-----A-----T-----A-----T-----C-----C-----C-----C-----GG-----GA----- T-A-----TGT--T-----G-----A-----A-----T-----C-----C-----C-----C-----C-----C-----C-----C-----GGG-----A-----AA-----T----- T-A-----C-----T-----T-----C-----G-----T-----T-----A-----T-----C-----C-----C-----C-----C-----C-----C-----C----- T-AC-----T-----T-----T-----T-----C-----G-----T-----C-----A-----T-----T-----A-----C-----A-----A----- T-AC-----T-----A-----A-----T-----G-----T-----C-----A-----T-----T-----A-----A-----C-----C-----A-----A-----AT----- T-A-----T-----AC-----G-----A-----A-----A-----A-----A-----C-----C-----G-----G-----A-----G-----C-----A-----G----- --T-A-----T-A-----AC-----C-----G-----A-----G-----A-----T-----A-----A-----A-----T-----CT-----T-----G-----C-----GC-----CAGAA-----T----- T-A-----T-A-----AT-----C-----AG-----A-----G-----G-----AGT-----C-----C-----T-----AA-----A-----A-----A-----C-----CT-----T-----G-----C-----CAGAC-----T----- TT-A-----TT-A-----AC-----AG-----AA-----G-----AG-----G-----G-----G-----G-----G-----G-----C-----C-----CAGG-----T----- TT-A-----TT-A-----AC-----G-----G-----T-----T-----ACT-----A-----T-----T-----A-----C-----C-----CT-----T-----G-----C-----CAGG-----T----- T-A-----T-A-----CTG-----G-----A-----T-----T-----ACT-----A-----C-----A-----AC-----A-----C-----C-----A-----G-----G-----AGGA-----AC-----T-----G----- T-AAA-----TGT-----ATA-----CACAGC-----GTG-----A-----T-----T-----ACTA-----A-----C-----A-----AC-----A-----C-----C-----A-----G-----G-----AGGA-----AT-----T-----G----- T-aA-----ataCT-----t-----AG-----g-----T-----t-----gg-----t-----t-----t-----ta-----T-----A-----A-----C-----C-----A-----G-----G-----t-----AT-----C----- T-A-----T-----T-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C----- T-A-----T-----T-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A----- T-CA-----T-----T-----T-----T-----A-----A-----T-----A-----A-----T-----A-----T-----A-----T-----G-----G-----A-----AT----- T-A-----T-----TGAT-----T-----AG-----A-----T-----C-----A-----A-----T-----A-----T-----A-----T-----G-----G-----A-----AT----- T-CA-----T-----T-----T-----T-----AG-----A-----T-----C-----A-----A-----T-----A-----T-----A-----T-----G-----G-----A-----AT----- T-A-----GT-----T-----AG-----T-----C-----A-----A-----T-----T-----C-----A-----T-----A-----A-----T-----T-----G-----G-----A-----AT----- T-A-----T-----T-----T-----T-----AG-----T-----C-----A-----A-----T-----T-----C-----A-----T-----A-----A-----T-----T-----G-----G-----A-----AT----- T-CA-----T-----C-----AG-----T-----A-----A-----T-----T-----C-----A-----T-----A-----A-----T-----T-----G-----G-----A-----AT----- Ta-A-----a-----T-----gc-----c-----c-----g-----t-----a-----at-----t-----c-----at-----c-----a-----t-----t-----a-----a-----g-----t-----a?-----g-----taag?-----a----- T-A-----T-----T-----C-----c-----c-----g-----t-----a-----at-----t-----c-----at-----c-----a-----a-----c-----a-----g-----g-----G-----G-----A----- TC-A-----C-----T-----AG-----D-----J-----D-----J-----2-----6-----3-----A-----A-----A-----A-----G-----A-----G-----A-----G-----G-----A----- TC-A-----C-----T-----AG-----D-----J-----D-----J-----2-----6-----4-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A----- T-A-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C----- T-A-----T-----T-----AC-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C----- T-A-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C----- T-----T-----G-----G-----A-----A-----T-----C-----A-----A-----A-----A-----T-----T-----C-----A-----A-----A-----A-----C-----C-----G-----A-----A----- T-----T-----G-----G-----A-----A-----T-----C-----A-----A-----A-----A-----T-----T-----C-----A-----A-----A-----A-----C-----C-----G-----A-----A----- T-----T-----TC-----CAA-----CA-----A-----T-----C-----C-----A-----A-----A-----A-----G-----A-----A-----T-----G-----G-----A-----AT-----C-----CC----- T-A-----G-----AC-----A-----T-----C-----C-----G-----T-----A-----T-----C-----C-----G-----T-----A-----T-----C-----C-----G-----G-----A-----AT-----C----- T-A-----T-----T-----AC-----A-----T-----T-----A-----A-----T-----A-----A-----T-----A-----T-----A-----T-----G-----G-----A-----AT-----C----- T-A-----T-----T-----AC-----A-----T-----T-----A-----A-----T-----A-----A-----T-----A-----T-----A-----T-----G-----G-----A-----AT-----C----- T-----T-----CTG-----A-----A-----T-----T-----A-----A-----T-----A-----A-----T-----A-----T-----A-----T-----G-----G-----A-----AT-----C----- TT-AA-----T-----T-----GT-----ACA-----AAT-----AG-----GT-----C-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A----- TT-A-----C-----T-----TA-----C-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A----- T-A-----C-----T-----A-----C-----CT-----C-----A-----A-----C-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C----- T-A-----C-----T-----A-----C-----CT-----C-----A-----A-----C-----A-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C----- T-A-----T-----TA-----ACT-----A-----A-----T-----A-----A-----T-----A-----A-----T-----A-----A-----T-----A-----A-----T-----A-----A-----T----- GGCAC-----C-----A-----ACT-----C-----A-----AT-----GA-----T-----CTA-----T-----A-----A-----A-----A-----A-----A-----T-----AT-----G-----G-----G-----AG-----G-----GAT-----T-----CA----- T-----S-----C-----N-----T-----S-----V-----I-----T-----O-----A-----C-----P-----K-----V-----S-----F-----E-----P-----I-----P-----I-----H-----Y-----C-----A-----P-----A-----G-----F-----A-----I-----L-----K-----C-----N-----K-----T-----F-----N-----G </pre>
---	---

HIV-1/SIVcpz complete genomes

Table with 4 columns: Accession ID, Nucleotide Sequence, Reverse Complement, and Source. The table lists 100 different HIV-1/SIVcpz complete genomes, including various strains like B_FR_HXB2R, G_BE_DRCBL, and others. Each entry shows the full genomic sequence and its reverse complement.



Accession	Sequence	Accession	Sequence	Accession	Sequence
B.FR.HXB2R	TTCTTACATAGGAAAAATAGAAATATG	..AGACAAGCA	..CATTTGAACATTAGTACAGCAAAATGGAATACACACTTTTAAACACAGATAGCTTAGCAAAATTAAGAGA
CONSENSUS A	-CTA-gca-ca--tg-c	..ata-c--gga--a	..c-----aaUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
A.UG.U455	-CTA-gt-c-t	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
A.SE.UGSE8891	-CTA-g-c-g-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
A.SE.TZSE8538	-CTA-g-c-g-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
A.SE.UGSE6594	-CTA-g-c-g-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
A.KE.Q2317	-CTA-g-c-g-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
A.SE.UGSE7535	-CTA-g-c-g-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
A.SE.SOSE7253	-CTA-g-c-g-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
A.UG.92UG037	-CTA-g-c-g-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
A.SE.SE8131	-CTA-g-c-g-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
A2.CD.97CDKTB48	-CTA-g-c-g-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
A2.CY.94CY017	-CTA-g-c-g-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
CONSENSUS B	tttAtacacaggggaaaacaata-aga-t	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.AU.MBC925	-TA-g-c-g-t	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.AU.MBC200	-TA-g-c-g-t	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.GA.OYI	-CA-g-c-aa-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.GB.MANC	-CA-gt-c-a-gcc	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.GB.CAML	-TA-g-c-c-ac-g	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.CN.RL42	GGTA-c-c-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.DE.D31	-TA-g-c-g	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.DE.HAN2	-TA-g-c-g	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.ES.89SEP061	-TA-g-c-c-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.NL.ACH320A	-TA-g-c-c-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.TW.LM49	-TAC-gc-a	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.US.DHL23	-TA-c-c-g	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.US.SF2	-TA-c-c-g	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.US.NY5CG	-CTA-g-g-a	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.US.WC1PR18	-TA-g-c-g-t	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.US.YU2	-GTA-c-g	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.US.JRCSF	-TA-c-c-g	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.US.JRFL	-TA-c-c-g	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.US.MN	-TA-c-c-g	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.US.BCSG3	A-ta-c-c-g	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.US.WEAL160	-TA-c-c-g	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.US.RF	-TA-c-c-g	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
B.US.WR27	-TA-g-c-c-c	..ATA--gg--A	..c-----aUG-c-g-agaTcaaaa	---aaaa-t-ac-a?-gg-a-ctaaac-----gaa?
CONSENSUS C	-TA-gata-gtT	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.BR.92BR025	-CTA-g-ac-g-ga?	..ATAa-a-g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.BW.96BW0402	-CTA-g-c-g-g	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.BW.96BW1104	-CTA-g-c-a-gc	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.BW.96BW15C02	-CTA-g-c-a-c	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.BW.96BW01B03	-CTA-g-g-c-g	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.BW.96BW0502	-CTA-g-c-g-c	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.BW.96BW16B01	-CTA-g-c-g-c	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.BW.96BW1210	-CTA-g-c-g-c	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.ET.FTH2220	-CTA-g-c-g-c	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.IN.301999	-CTA-g-c-g-c	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.IN.21068	-CTA-g-c-g-c	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.IN.301905	-CTA-g-c-g-c	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.IN.301904	-CTA-g-c-g-c	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.IN.93IN101	-CTA-g-c-g-c	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
C.IN.11246	-CTA-g-c-g-c	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
D.UG.94UG114	-CTA-g-c-g-c	..ATA--g-c-a	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
D.CD.NDK	-CTA-t-c-g	..TCA-G-TC-TA-A	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
D.KE.ELI	-CTA-c-c-g	..ACA-G-A-T-TA-A	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
D.KE.MB2059	-CTA-c-c-g	..ACA-G-A-T-TA-A	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
D.CD.2226	-CTA-c-c-g	..ACA-G-A-T-TA-A	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
D.CD.84ZR085	-TA-c-c-g-t	..ATA--g-c-c	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
F1.BR.93BR020	-TA-g-c-g	..ATA--g-c-c	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
F1.FI.FIN9363	-CTA-g-c-gc	..ATA--g-c-c	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
F1.BE.VI850	-CTA-g-c-gc	..ATA--g-c-c	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
F1.FR.MF411	-CTA-g-c-gc	..ATA--g-c-c	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
F2.CM.MF255C	-CTA-g-c-gc	..ATA--g-c-c	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA
F2.CM.MF257C	-CTA-g-c-gc	..ATA--g-c-c	..c-----ac-g-gaagTgaaa	---atgaaa-tt-c-aaggg-agtGCCAAA-TA-G-GA-TTA

Main body of the document containing a large block of DNA sequence with various annotations and labels on the left side. The sequence is organized into lines, with labels such as 'B. FR. HXB2R', 'G. BE. DRCEL8', etc., on the left. The sequence itself is a mix of uppercase letters (A, C, G, T) and dashes, representing nucleotide bases and gaps. There are also some lowercase letters and symbols interspersed. The right side of the sequence has some labels like 'Env gp120' and 'Env gp160'.

HIV-1/SIVcpz complete genomes

B.FR.HXB2R AACAATTAATCTTT AAGCAATCC TCAGGAGGGGACCCAGAAATTTGAAACGCCACAGTCTTTTACTTCTAGTAATTCACAA
T-C-A-TTT
-ATC-A-AT
G-NG.92NG083 G-GC-GAA CTC
G.SE.56165 G-GC-C CTC
G.FI.HH87932 G-GC-C CTC
H.BE.VI991 T-A C-A-C-C G
H.BE.VI997 T-C T G
H.CF.90CF056 G-GAAC G
J.SE.SE91733 C-AAT AAT
J.SE.SE92809 C-AAT AAT
K.CM.MP535C C-T C-T-TAATA CCA-C-AA C
K.CD.E0TB11C C-A-C-A C
O.CM.AN170 GTT-AC-A C C
O.CM.MP5180 GTT-AC-A C C
O.SN.MP1299 GTT-AC-C CA-C G
O.SN.MP1300 GTT-AC-C CA-C G
N.CM.YBF106 G-CT-Aateaat??? C
CONSENSUS 01 C-CT-CAGAAGCT-GAGAGGAAAT-A T
01.AE.TH.93TH253 aag-ca-atic caa-C-c??? A
01.AE.TH.CM240 G-C-AAT G
01.AE.TH.92CF402 G-C-AAT G
01.AE.TH.92TH022 G-C-AAT G
01.AE.TH.95TNIH047 G-C-AAT G
01.AE.CF.11697 G-C-AAT G
01.AE.TH.93TH057 G-C-AAT G
01.AE.TH.93TH065 G-C-AAT G
01.AE.CF.4071 G-C-AAT G
CONSENSUS 02 G-ct-ttaa?aat ??paca-ctc gcta-ccc c
02.AG.SE.7812 GT-T-AT G
02.AG.GH.G829 G-T-A-AG G
02.AG.DJ.DJ264 G-T-A-AG G
02.AG.NG.IBNG G-T-A-AG G
02.AG.SN.MP1211 T-T-C-AAT G
02.AG.SN.MP1213 T-T-C-AAT G
03.AB.RU.98001 AAC
03.AB.RU.KALI53-2 AAC
04.CPX.GR.97PMV AACAA-C-A A
04.CPX.GR.97PMVCH AACAA-C-A A
05.DF.BE.VI1310 AACAA-C-A A
05.DF.BE.VI961 AACAA-C-A A
06.CPX.ML.95ML84 AACAA-C-A A
06.CPX.AU.BPP90 AACAA-C-A A
06.CPX.ML.95ML127 AACAA-C-A A
06.CPX.SN.97SE1078 AACAA-C-A A
10.CD.BE1061 AACAA-C-A A
10.CD.BE110 AACAA-C-A A
10.CD.BE1071 AACAA-C-A A
11.CPX.CM.MP818 AACAA-C-A A
11.CPX.FR.MP1298 AACAA-C-A A
11.CPX.FR.MP1307 AACAA-C-A A
11.CPX.GR.GR17 AACAA-C-A A
CPZ.GA.SIVCPZGAB AACAA-C-A A
CPZ.US.SIVCPZUS AACAA-C-A A
CPZ.CM.SIVCPZCAM3 AACAA-C-A A
CPZ.CM.SIVCPZCAM5 AACAA-C-A A
CPZ.CD.SIVCPZANT AACAA-C-A A
..... K_T_I_I_F_..K_Q_S_..S_G_G_D_P_E_I_V_T_H_S_F_N_C_G_G_E_F_Y_C_N_S_T_
_Q_F_G_N_N_.....K_T_I_I_F_..K_Q_S_..S_G_G_D_P_E_I_V_T_H_S_F_N_C_G_G_E_F_Y_C_N_S_T_
Env gp120

HIV-1/SIVcpz complete genomes

B. FR. HXB2R	CAACTGTTTAACTAGTACTGGTTTAATAGTACTTGGACTACTGAAAGGCTGAAATAAACAACCTGAAAGAACTGACACAACTCACTCCCTCCATCGAGATAAACAACA	7491
CONSENSUS A	gccc-g-----t-c-cttysaat?????aaat??cactgacatgtaaaatgagccggagtcgcaaat-a-cl-a-cl-c-c-ca-c-----ag	6679
A.UG.0455	AGC-----C-T-----AAGGCAGCAI-----IG-GGCC--A-G--T-A-T--A--T-----G	6921
A.SE.UGSE8891	AGC-----C-C-G-AGC-A-GCC--CGAGCCGAT-----G-A-GTC--A--G--T-A--T-----GG	6556
A.SE.TZSE8538	AGC-----C-C-G-AGACGA-AAC-C-C--G-A-GAC--AA-G--T-A--T-----G	6664
A.SE.UGSE6594	AGC-----C-C-G-AGGT-A--GCACCTT-G-A--CACTGAC-GC-CTCAG-GTC-A--T-A--T-----G	6928
A.KE.Q2317	AGC-----C-ATT-GC-AA--TG-A--CAATG-C-GC-GGAGTCTGATGA--T-A--T-----AG	6729
A.SE.UGSE7535	AGC-----C-C-AGC-TACAAATGAC--GG-TA--G-G-GTC--A--T-A--T-----G	6686
A.SE.SOSE7253	AGC-----G-A--GGAAATGAC--CA--TAAATT-C--GTC--A--T-A--T-----G	6852
A.UG.92UG037	AGC-----G-A--GGAAATGAC--CA--TAAATT-C--GTC--A--T-A--T-----G	6866
A2.CD.97CDKTB48	AGC-----C-A--GAAAATG-C-A--CAAAACAG--T--G-GTC--A--T-A--G--A--	6804
A2.CY.94CY017	AGC-----G-G--GG-C-A-GG-AC-T-G-AC-GGCC-AC-CACTT-A-A-CACC-A--GA-GT--A--T-----	6866
CONSENSUS B	caac-----aatagt-t-----aat?????????????aaatg?aaatg?taaaataataactaaagcaaatg?cta-c-ca-c-ca-c-g	7410
B.AU.MBC925	AGC-----G-----AAGAA--GAA--GT-ATGTC-AI--C-CTG-A--AAI-A--T-----G	7513
B.AU.MBC200	AGC-----G-----AAATTAAAT-G-C-TGGATGA-----G-A-----T-----C-A--T-----C	7513
B.GA.OYI	AGC-----G-----AA-TACTA-----GGCA-A-AGCACTGAA--TA--T-A-----T	7024
B.GB.MANC	AGC-----G-----AA-C-GGGATGA--AG-A-----G--AA-TAAT-C-GA-AT--A-----T	7439
B.GB.CAMI	AGC-----G-----C-GTT--A-GG-ACCTT-G-A--TACTGA-GGGTTA-A-A-CRGTGAAAGA-AT--T-A--T-----	7489
B.CN.RL42	AGC-----G-----AAGA-----GTACTT-GG--G-T--AC--A--AG--T-A--T-----	6838
B.DE.D31	AGC-----G-----AATACTAA-G-GTCA-A-A-CAC-A--GA--T-A--T-----	7012
B.DE.HAN2	AGC-----G-----AAA-ACCTA-TA-TTGG--TGA-A-T--A--T--TA-G--T-----	6947
B.ES.89SP061	AGC-----G-----AAGG-ACCTGATGTC-AATAAC-C-G-AGGAA--T-A--A-----T	7063
B.NL.ACH320A	AGC-----G-----AAGAT-C-GGGA-T-TTA-TG-A-GGT-AA-TAAC-A--A-AT--A-----T	7489
B.TW.LM49	AGC-----G-----AAGCT-CC-G-ACITTT-AAATCC-CT-A-----A--A--A--T-----T	7462
B.US.DH123	AGC-----G-----AAG-----T-----T-----A-----T-----	7463
B.US.SF2	AGC-----A--A--AGGTTAA-TCAGA-TG-AGGA--A-----T-----TA-----T	7490
B.US.MB5CG	AGC-----G-----C-GTT--A-A-T-C-TGGA-T-ATA-TG-A-GGT-A--TAAAC-A--A--T--TA-----	7422
B.US.WC1PR18	AGC-----G-----AA-A--TACTAG-----TGAAGCACTG-A-GA-AT--A-----T	7045
B.US.YU2	AGC-----G-----AA-TACTAG--GTTA-A-A-CRGTG-AAAGA-AT--A-----T	7452
B.US.JRCSF	AGC-----G-----AAGA-----AA-----G-GG-----C--TA-----	7478
B.US.JRFL	AGC-----G-----AAA-A-ACCT-AAAGGTCA--T-ACACT-A-G-AA-T--T-----	6780
B.US.MN	AGC-----G-----AAG-T-A-A-ACITTT-AAAT--CT-A-GGTC--A-A--AT--A--T-A--A--	7505
B.US.BCSG3	AGC-----G-----AAGGG--A-----A-G-AGT-CTGA--GGTCAGA--CACTG-A-GA-AT--A-----	7020
B.US.WEAU160	AGC-----G-----AAGCT-A-GG-ACITTT-AAAG-----CTGAA-GG-C-GA-A--AT--A-----	7492
B.US.RF	AGC-----G-----AAG-----G-----A-----A-----	7030
B.US.WR27	AGC-----G-----AA-G-ACTT-BAAG-----GT-----TGCTC-GA-A-A-TT--GA-----	6866
CONSENSUS C	agcc-----a-agtaacataca?t?????????????pa?????aaat??actaaagtaatacaaatcaacc-c-caa-ccc-----a--a	6906
C.BR.92BR025	AGC-----A-ACACACCT-A--G-AC--AAAC-T-CTGGAAAC--A-A-TCA-TC--AA-----	6828
C.BW.96BW0402	AGC-----G-GC-ACAGCTT--A--G--GA-CCACTGGTCT-A--TAC-A--CA-C--A-----	6967
C.BW.96BW1104	AGC-----C-G-ATAAGAGCC-----G-AC--C-A-TCA--A-----	6824
C.BW.96BW15C02	AGC-----A-ACA-GG--CAGTTATAGGG--CAGAAAGTAAAT--TCA-TC--A-----	6948
C.BW.96BW0502	AGC-----A-ACAGTC-A-GTACAGA-GG--T-TAAGCCC--ACA-TC--TG-----	6931
C.BW.96BW16B01	AGC-----A-ACAACCC-G-TTTAATGGT-CAG-A-A-A--TT-AC-GG--C--AA--A--T-----	6971
C.ET.ETH2220	AGC-----C-AAAAAC-A-G-ACAGAC--A-TA-TA--G-C-GGT--A-TAAT-CA--AC--AA--A--T-----	6968
C.IN.301999	AGC-----G-----C-CTG-TTAA-G-ACATACTG-----CC-TCT-A-TAC-G-GACA-AC--A--T-A--G-----	6848
C.IN.21068	AGC-----G-----A-ACAACCTAAGTT--T-GCAGTAC--A-TTA-AC--A--A-----	6867
C.IN.301905	AGC-----G-----A-ACAACCTAAGTT--CA-GA-A--TAC-ACITTA--C--T--AA-----	6887
C.IN.301904	AGC-----G-----A-ACAAC--AAUGA--AA-GA-TAATCA-G-G-A-TCC-AC-CA-AC--AA-----	6981
C.IN.11246	AGC-----G-----A-ACA-GGAGCC-----GG-CAG-A-G-A-TTC--CTCA-TC--AA-----	6817
D.UG.94UG114	AGC-----G-----A-AAAC-ACATAC-TGCCTAATGGT-CAG-A-G-A-TTC--ACTCA--T--AA-----	6817
D.CD.NDK	AGC-----G-----A-AAAGG-ACATAC-TG--AAAT--G-GCAGCATGAATTCAG-TAC--AA-----	7454
D.KE.MB2059	AGC-----G-----A-AAAGGAAT-A-G--T--AGG-G-T-A-AC-CC-GA--G--A--A-----	6806
D.CD.2226	AGC-----G-----A-AAAG--G-A--A--C-GACTAAT-G-CAGGTTCAAT-A-G--G--A-----	6995
D.CD.42R085	AGC-----G-----A-AAAT--G-A--A--A-ATTACAGAGTCA--T-A-AGCAC--ACACA-AC--A--A-----	7030
F1.BR.93BR020	AGC-----G-----A-AAAAAATGA--AT-TCAAAT--GGG--AATGAGACT--A-----	7461
F1.FI.FIN9363	AGC-----G-----A-AAAT--G-AAATCAGAATGG-CTAAT--G-CAGAGTC--ATGA-A-AC--A--A-----	7484
F1.BE.VI850	AGC-----G-----A-AAA-TAT-A-G-CAT-----G-CAGGGTT-AAUGA-ACA-TC--A--A-----	7012
F1.FR.MF411	AGC-----G-----A-AAAGAC-AAAA-C-----A-G-TTC-A-G--T--T-----TC-----	6798
F2.CM.ME255C	AGC-----G-----A-AAA-A-AGTACC-----AAT-A-G--T--T-----TC-G-----	6771
F2.CM.ME255C	AGC-----G-----A-AAAGAA-G-GAAAG-----TAC-A-G--T--TT-----T-A-----	6852
F2.CM.ME257C	AGC-----G-----A-AAAGAA--AGA-G-G-----GCT-A-A-ACAC--AC-A-AT--T-----G-----	6850
F2.CM.ME257C	AGC-----G-----A-AAACAC--AA-AG-A-----AATGAC-T-CTGTCTAAT--C--GAG--A--A--T-----T-----	6886

Rev Responsive Element
|- (RRE) Stem I

B. FR. HXB2RAATGAGTCCGAGATCTTCAGACCTGGAGGAGGATATGAGGCAAAATGGAGAGTGAATTAATATAAATGTAAGTAAATTAACCACTTAGGAGTAGCACCACCCAGGCAAGAGAGA	7733
G. BE. DRCEBLGCACAGT	7672
G. NG. 92NG083GAC-G-CACAGAG	7059
G. SE. SEG165	AAATACA-G-AGT	7151
G. FI. HH87932GCA-G-A-T	7103
H. BE. VI991CAA-TAAT-TT-C	7132
H. BE. VI997TAAT-TT-C	7047
H. BE. VI997GCA-AAAT-T-C	7052
H. CE. 90CF056	T...GAG-GCA-T-AC	7040
J. SE. SE91733	T...GAG-GCA-T-AC	6992
J. SE. SE92809GCA-T-C	6916
K. CM. MP535C	..AATACT-GGACTGAG	7775
K. CD. EQIB11CAGC--CA-CAAT-TA-CA-T	7789
O. CM. ANI70CA-GTGAAT-T-CAC-T	7825
O. CM. MPF5180CA-CAAT-T-CA-T	7798
O. SN. MP1299	CACCAAC-CAGCAA-CC-CA-T	7278
O. SN. MP1300CCA-GAAACA-AG-ATAT-CTC	7233
N. CM. YBF30CCA-GAAACA-AG-ATAT-CTC	7645
N. CM. YBF106GCA-AAAT-T-C	7286
CONSENSUS 01GCA-AAAT-T-C	7796
O1 AE. TH. 93TH253GCA-AAAT-T-C	7681
O1 AE. TH. CM240GCA-AAAT-T-C	7690
O1 AE. TH. 90CF402GCA-AAAT-T-C	7885
O1 AE. TH. 92TH022GCA-AAAT-T-C	7855
O1 AE. TH. 95THH047GCA-AAAT-T-C	7659
O1 AE. CF. 11697GCA-AAAT-T-C	6895
O1 AE. TH. 93TH057GCA-AAAT-T-C	7120
O1 AE. TH. 93TH065GCA-AAAT-T-C	7108
O1 AE. CF. 4071GCA-AAAT-T-C	7239
CONSENSUS 02GCA-AAAT-T-C	6928
O2 AG. SE. 7812GCA-AAAT-T-C	7067
O2 AG. GH. G829GCA-AAAT-T-C	7929
O2 AG. DJ. DJ263GCA-AAAT-T-C	6935
O2 AG. NG. I8NGGCA-AAAT-T-C	6916
O2 AG. SN. MP1211GCA-AAAT-T-C	7048
O2 AG. SN. MP1213GCA-AAAT-T-C	7453
O2 AG. CM. MP807GCA-AAAT-T-C	7812
O3 AB. RU. 98001GCA-AAAT-T-C	7139
O3 AB. RU. KAL153-2GCA-AAAT-T-C	7090
O4 CPX. GR. 97PVMYGCA-AAAT-T-C	6937
O4 CPX. GR. 97PVGHGCA-AAAT-T-C	7723
O4 CPX. CV. 94CY0323GCA-AAAT-T-C	7768
O5 DF. BE. V11310GCA-AAAT-T-C	7098
O5 DF. BE. V1961GCA-AAAT-T-C	7117
O6 CPX. ML. 95ML84GCA-AAAT-T-C	7727
O6 CPX. AU. BFP90GCA-AAAT-T-C	7765
O6 CPX. ML. 95ML127GCA-AAAT-T-C	7000
O6 CPX. SN. 97SE1078GCA-AAAT-T-C	7735
O6 CPX. SN. 97SE1078GCA-AAAT-T-C	7755
10 CD. BFL061GCA-AAAT-T-C	7749
10 CD. BFL110GCA-AAAT-T-C	7156
10 CD. BFL071GCA-AAAT-T-C	
11 CPX. CM. MP818GCA-AAAT-T-C	
11 CPX. FR. MP1298GCA-AAAT-T-C	
11 CPX. FR. MP1307GCA-AAAT-T-C	
11 CPX. GR. GR17GCA-AAAT-T-C	
CPZ. GA. SIVCPZGABGCA-AAAT-T-C	
CPZ. US. SIVCPZUSGCA-AAAT-T-C	
CPZ. CM. SIVCPZCM3GCA-AAAT-T-C	
CPZ. CM. SIVCPZCM5GCA-AAAT-T-C	
CPZ. CD. SIVCPZAMTGCA-AAAT-T-C	
Env gp120N_E_S_E_I_F_R_P_G_G_D_M_R_D_N_W_R_S_E_L_Y_K_Y_K_V_V_K_I_E_P_L_G_V_A_P_T_K_A_K_R	Env

HIV-1/SIVcpz complete genomes

Strain	RRE Stem I cont.	Env gp120 end \ Env gp41 start	Stem I	Stem IIA	Stem IIIB	Position
B. FR. HXB2R	AGAGTGGTG.....	CAGAGAGAAAAGACAGTG...GGA...ATAGAGCTTTG...TTCCCTTGGG...TTCTTTGGACAGCAGGAGCACTATGGCGGACGCTCAATGACCCCTGA	A-A-A	G-G-G	G-G-G	7836
G. BE. DRCEBL	G.....	T.....	A-C	A-C	A-C	7775
G. NG. 92NG083	G.....	T.....	C-G	C-G	C-G	7162
G. SE. SE6165	G.....	T.....	G-C	G-C	G-C	7254
G. FI. HH87932	GG.....	T.....	G-G	G-G	G-G	7206
H. BE. VI991	G.....	T.....	G-G	G-G	G-G	7235
H. BE. VI997	G.....	T.....	G-G	G-G	G-G	7150
H. CF. 90CF056	G.....	T.....	G-G	G-G	G-G	7155
J. SE. SE91733	G.....	T.....	G-G	G-G	G-G	7150
J. SE. SE92809	G.....	T.....	G-G	G-G	G-G	7143
K. CM. MP535C	G.....	T.....	G-G	G-G	G-G	6996
K. CD. EQPBL1C	G.....	T.....	G-G	G-G	G-G	7019
O. CM. AN70	CC--CA-AAGCCTAGAACT...T	T-G	ATGC-A	T-G	G-TC-AA-T	7890
O. CM. MPF5180	CC-A-AA-AACAITCACCCCT--C-G	T-G	ATGC-A	T-G	G-TC-AA-T	7890
O. SN. MP1299	CC--CA-AGCAGTGGCACT...A	T-G	ATGC-A	T-G	G-TC-AA-T	7940
O. SN. MP1300	CC--CA-AGCAGTGGCACT...GA	T-G	ATGC-A	T-G	G-TC-AA-T	7913
N. CM. YBF30	C-CACA--AGT	CCCTTT	C-T-GC	T-G	TC-T-G	7384
N. CM. YBF106	C-CACA--AGT	CCCTTT	C-T-GC	T-G	TC-T-G	7339
CONSENSUS 01	G.....	T.....	a-a	A-T	a-a	7748
01 AE. TH. 93TH253	G.....	T.....	A-A	A-T	A-A	7820
01 AE. TH. CM240	G.....	T.....	A-A	A-T	A-A	7389
01 AE. CF. 90CF402	G.....	T.....	A-A	A-T	A-A	7899
01 AE. TH. 92TH022	G.....	T.....	A-A	A-T	A-A	7784
01 AE. TH. 95THH047	G.....	T.....	A-A	A-T	A-A	7783
01 AE. CF. 11697	G.....	T.....	A-A	A-T	A-A	7788
01 AE. TH. 93TH057	G.....	T.....	A-A	A-T	A-A	7773
01 AE. TH. 93TH065	G.....	T.....	A-A	A-T	A-A	7758
01 AE. CF. 4071	G.....	T.....	A-A	A-T	A-A	7762
CONSENSUS 02	G-aa	ag-T???	g-C	l-C	g-C	6998
02 AG. SE. 7812	G.....	T.....	C-T	C-T	C-T	7226
02 AG. GH. G829	G.....	T.....	C-G	C-G	C-G	7031
02 AG. DJ. DJ263	G.....	T.....	C-G	C-G	C-G	7211
02 AG. DJ. DJ264	G.....	T.....	C-G	C-G	C-G	7170
02 AG. NG. I8NG	G.....	T.....	C-G	C-G	C-G	7342
02 AG. SN. MP1211	G.....	T.....	C-G	C-G	C-G	7038
02 AG. SN. MP1213	G.....	T.....	C-G	C-G	C-G	7018
02 AG. CM. ME607	G.....	T.....	C-G	C-G	C-G	7016
03 AB. RU. 98001	G.....	T.....	G-G	G-G	G-G	7151
03 AB. RU. KALI53-2	G.....	T.....	G-G	G-G	G-G	7556
04 CPX. GR. 97PVMY	C-G	C	C	C-CTTG	C	7887
04 CPX. GR. 97PVGH	C.....	T.....	A-C	A-C	A-C	7935
04 CPX. CV. 94CY0323	C.....	T.....	G-CA	G-CA	G-CA	7206
05 DF. BE. VI1310	CA	A	A	A	A	7242
05 DF. BE. VI961	CA	A	A	A	A	7193
06 CPX. ML. 95ML84	A	GG	G-T	G-T	G-T	7040
06 CPX. AU. BFP90	A	GG	G-T	G-T	G-T	7884
06 CPX. ML. 95ML127	A	GG	G-T	G-T	G-T	7829
06 CPX. SN. 97SEL078	A	GG	G-T	G-T	G-T	7874
10 CD. BFL061	G-A	A-A	A-A	A-A	A-A	7187
10 CD. BFL110	G-A	A-A	A-A	A-A	A-A	7201
10 CD. BFL071	G-A	A-A	A-A	A-A	A-A	7220
11 CPX. CM. MP818	G	AGTG	G	G	G	7814
11 CPX. FR. MP1298	G	AGTG	G	G	G	7833
11 CPX. FR. MP1307	G	AGTG	G	G	G	7871
11 CPX. GR. GR17	G	AGTG	G	G	G	7103
CPZ. GA. SIVCPZGAB	CATACA--GCAAGACAG..AAG-C-C-G	CCITTC-T..C-G-AC	C	T	C	7873
CPZ. US. SIVCPZUS	CATACA--A	CCITTC	C-T-AC	-G	TC-T	7841
CPZ. CM. SIVCPZCM3	CATACA--A	CCITTC	C-G-T-GC	T	A-G	7249
CPZ. CM. SIVCPZCM5	CACAC--	TCA	C-G-T-GC	T	A-G	7555
CPZ. CD. SIVCPZANT	CC--AAA-AAACAACAC.....TCC--C	GCA-T	C-AC	C	TC-A	7268
Env gp120/gp41	_R_V_V.....	Q_R_E_K_R_A_V...G...I_G_A_L_L...F_L_G...F_L_G_A_A_G_S_T_M_G_A_A_S_M_T_L				Env

HIV-1/SIVcpz complete genomes

Stem IIB'	III'	Stem IIC'	Stem IIC'	Stem IIA'	Stem III-IV'	Stem III-IV'	Stem III-IV'	Stem V	Stem V'	Stem I'	
B.FR.HX82R	CGGTACAGCC	ACG	CGG	TTGGTATTAG	TTCAGTCCAGCAG	CAACAATTTGCTGAGG	GCTATTCGAGCGCG	AACAGCAGTCTGGG	GCATCAACGAGCTCCAG	GCAGAGAATCCTT	7966
CONSENSUS A	---	---	---	---	---	---	---	---	---	---	7150
A.UG.U455	---	---	---	---	---	---	---	---	---	---	7396
A.SE.UGSE8891	---	---	---	---	---	---	---	---	---	---	7131
A.SE.TZSE8538	---	---	---	---	---	---	---	---	---	---	7183
A.SE.UGSE6594	---	---	---	---	---	---	---	---	---	---	7136
A.KE.Q2317	---	---	---	---	---	---	---	---	---	---	7403
A.SE.UGSE7535	---	---	---	---	---	---	---	---	---	---	7207
A.SE.SOSE7253	---	---	---	---	---	---	---	---	---	---	7161
A.UG.92UG037	---	---	---	---	---	---	---	---	---	---	7327
A.SE.SE8131	---	---	---	---	---	---	---	---	---	---	7344
A2.CD.97CDKTB48	---	---	---	---	---	---	---	---	---	---	7276
AG.CT.94CY017	---	---	---	---	---	---	---	---	---	---	7338
CONSENSUS B	---	---	---	---	---	---	---	---	---	---	7884
B.AU.MBC925	---	---	---	---	---	---	---	---	---	---	7991
B.AU.MBC200	---	---	---	---	---	---	---	---	---	---	7505
B.GA.OYI	---	---	---	---	---	---	---	---	---	---	7926
B.GB.MANC	---	---	---	---	---	---	---	---	---	---	7967
B.GB.CAMI	---	---	---	---	---	---	---	---	---	---	7319
B.CN.RL42	---	---	---	---	---	---	---	---	---	---	7493
B.DE.D31	---	---	---	---	---	---	---	---	---	---	7428
B.DE.HAN2	---	---	---	---	---	---	---	---	---	---	7544
B.ES.89SP061	---	---	---	---	---	---	---	---	---	---	7970
B.NL.ACH320A	---	---	---	---	---	---	---	---	---	---	7937
B.IW.LM49	---	---	---	---	---	---	---	---	---	---	7977
B.US.DHA23	---	---	---	---	---	---	---	---	---	---	7935
B.US.SF2	---	---	---	---	---	---	---	---	---	---	7971
B.US.NYSG	---	---	---	---	---	---	---	---	---	---	7903
B.US.WCIPR18	---	---	---	---	---	---	---	---	---	---	7535
B.US.YU2	---	---	---	---	---	---	---	---	---	---	7927
B.US.JRGSF	---	---	---	---	---	---	---	---	---	---	7953
B.US.JRFL	---	---	---	---	---	---	---	---	---	---	7955
B.US.MN	---	---	---	---	---	---	---	---	---	---	7983
B.US.RCSG3	---	---	---	---	---	---	---	---	---	---	7498
B.US.WEAU160	---	---	---	---	---	---	---	---	---	---	7973
B.US.RF	---	---	---	---	---	---	---	---	---	---	7511
B.US.WR27	---	---	---	---	---	---	---	---	---	---	7344
CONSENSUS C	---	---	---	---	---	---	---	---	---	---	7278
C.BR.92BR025	---	---	---	---	---	---	---	---	---	---	7303
C.BW.96BW0402	---	---	---	---	---	---	---	---	---	---	7445
C.BW.96BW1104	---	---	---	---	---	---	---	---	---	---	7293
C.BW.96BW15C02	---	---	---	---	---	---	---	---	---	---	7423
C.BW.96BW01B03	---	---	---	---	---	---	---	---	---	---	7412
C.BW.96BW0502	---	---	---	---	---	---	---	---	---	---	7449
C.BW.96BW16B01	---	---	---	---	---	---	---	---	---	---	7449
C.BW.96BW1210	---	---	---	---	---	---	---	---	---	---	7323
C.ET.ETH2220	---	---	---	---	---	---	---	---	---	---	7342
C.IN.301999	---	---	---	---	---	---	---	---	---	---	7374
C.IN.21068	---	---	---	---	---	---	---	---	---	---	7377
C.IN.301905	---	---	---	---	---	---	---	---	---	---	7356
C.IN.301904	---	---	---	---	---	---	---	---	---	---	7298
C.IN.93IN101	---	---	---	---	---	---	---	---	---	---	7941
C.IN.11246	---	---	---	---	---	---	---	---	---	---	7268
D.UG.94UG114	---	---	---	---	---	---	---	---	---	---	7284
D.CD.NDK	---	---	---	---	---	---	---	---	---	---	7470
D.CD.ELI	---	---	---	---	---	---	---	---	---	---	7502
D.KE.MB2059	---	---	---	---	---	---	---	---	---	---	7936
D.CD.2226	---	---	---	---	---	---	---	---	---	---	7959
D.CD.842R085	---	---	---	---	---	---	---	---	---	---	7487
F1.BR.93BR020	---	---	---	---	---	---	---	---	---	---	7970
F1.FI.FIN9363	---	---	---	---	---	---	---	---	---	---	7562
F1.BE.VI850	---	---	---	---	---	---	---	---	---	---	7443
F1.FR.MF411	---	---	---	---	---	---	---	---	---	---	7133
F2.CM.MF255C	---	---	---	---	---	---	---	---	---	---	7125
F2.CM.MF257C	---	---	---	---	---	---	---	---	---	---	7158

Stem IIB'	Stem IIC	Stem IIA'	Stem III-IV	Stem III-IV	Stem V	Stem V'	Stem I'
B. FR. HX22R	CGGTACAGCCAGCAAAATTTCTGCTAGTGTAGTGTGCTAGGAGCAACAATTTCTGAGGGCTATTGAGCCGCAACAGCATCTGTTTGCACCTCAGTCTGGGGCACTCAGCAGCTCCAGGCAAGAAATCTCT	C	A	A	G	A	7966
G. BE. DRCEBL	T	C	A	A	A	T	7905
G. NG. 92NG083	C	C	A	A	A	T	7292
G. SE. SE6165	T	C	A	AGG	C	A	7384
G. FI. HH87932	T	C	A	A	A	T	7336
H. BE. VI991	T	C	A	A	A	T	7365
H. BE. VI997	T	C	A	A	AC	A	7280
H. CF. 90CF056	G	G	A	A	AC	A	7285
J. SE. SE91733	G	G	A	A	AC	A	7280
J. SE. SE92809	T	C	A	A	AM	A	7273
K. CM. MP535C	T	C	A	A	A	T	7126
K. CD. EOPB1C	T	C	A	A	A	T	7149
O. CM. ANP70	---A-CACACT-GC--AAG	---A	A	A	A	A	8020
O. SN. MPF5180	---A-CACAGTG-C--AAG	---A	G	CC	---A	A	8037
O. SN. MP1299	---A-CA-TCT-GA--AAG	---A	G	CC	---A	A	8070
O. SN. MP1300	---A-CACACT-GA--AAG	---A	G	CC	---A	A	8043
N. CM. YBF30	---C-GAC	---A	G	T	C	T	7514
N. CM. YBF106	---A-C-GAC	---A	G	T	C	T	7469
CONSENSUS 01	---A-C-GAC	---A	G	T	C	T	7878
01 AE. TH. 93TH253	---A-C-GAC	---A	G	T	C	T	7950
01 AE. TH. CM240	---A-C-GAC	---A	G	T	C	T	7519
01 AE. CF. 90CF402	---A-C-GAC	---A	G	T	C	T	8029
01 AE. TH. 92TH022	---A-C-GAC	---A	G	T	C	T	7914
01 AE. TH. 95THH047	---A-C-GAC	---A	G	T	C	T	7923
01 AE. CF. 11697	---A-C-GAC	---A	G	T	C	T	7918
01 AE. TH. 93TH057	---A-C-GAC	---A	G	T	C	T	7903
01 AE. TH. 93TH065	---A-C-GAC	---A	G	T	C	T	7888
01 AE. CF. 4071	---A-C-GAC	---A	G	T	C	T	7892
CONSENSUS 02	---A-C-GAC	---A	G	T	C	T	7128
02 AG. SE. 7812	---A-C-GAC	---A	G	T	C	T	7356
02 AG. GH. G829	---A-C-GAC	---A	G	T	C	T	7161
02 AG. DJ. DJ263	---A-C-GAC	---A	G	T	C	T	7341
02 AG. DJ. DJ264	---A-C-GAC	---A	G	T	C	T	7300
02 AG. NG. I8NG	---A-C-GAC	---A	G	T	C	T	7472
02 AG. SN. MP1211	---A-C-GAC	---A	G	T	C	T	7168
02 AG. SN. MB1213	---A-C-GAC	---A	G	T	C	T	7148
02 AG. CM. ME607	---A-C-GAC	---A	G	T	C	T	7146
03 AB. RU. 98001	---A-C-GAC	---A	G	T	C	T	7281
03 AB. RU. KAL153-2	---A-C-GAC	---A	G	T	C	T	7886
04 CPX. GR. 97PVMY	---A-C-GAC	---A	G	T	C	T	8017
04 CPX. GR. 97PVGH	---A-C-GAC	---A	G	T	C	T	8065
04 CPX. CV. 94CY0323	---A-C-GAC	---A	G	T	C	T	7372
05 DF. BE. V11310	---A-C-GAC	---A	G	T	C	T	7323
05 DF. BE. V1961	---A-C-GAC	---A	G	T	C	T	7323
06 CPX. ML. 95ML84	---A-C-GAC	---A	G	T	C	T	7170
06 CPX. AU. BFP90	---A-C-GAC	---A	G	T	C	T	8014
06 CPX. ML. 95ML127	---A-C-GAC	---A	G	T	C	T	7959
06 CPX. SN. 97SEL078	---A-C-GAC	---A	G	T	C	T	8004
10 CD. BFL061	---A-C-GAC	---A	G	T	C	T	7317
10 CD. BFL110	---A-C-GAC	---A	G	T	C	T	7331
10 CD. BFL071	---A-C-GAC	---A	G	T	C	T	7350
11 CPX. CM. MP818	---A-C-GAC	---A	G	T	C	T	7944
11 CPX. FR. MP1298	---A-C-GAC	---A	G	T	C	T	7963
11 CPX. FR. MP1307	---A-C-GAC	---A	G	T	C	T	8001
11 CPX. GR. GR17	---A-C-GAC	---A	G	T	C	T	7233
CPZ. GA. STVCPZGAB	---A-C-GAC	---A	G	T	C	T	8003
CPZ. US. STVCPZUS	---A-C-GAC	---A	G	T	C	T	7971
CPZ. CM. STVCPZCM3	---A-C-GAC	---A	G	T	C	T	7379
CPZ. CM. STVCPZCM5	---A-C-GAC	---A	G	T	C	T	7685
CPZ. CD. STVCPZANT	---A-C-GAC	---A	G	T	C	T	7398

Env gp41

HIV-1/SIVcpz complete genomes

	Stem I'	RRE end <-	
B.FR.HXB2R	GCCTGTGGAAAGATACCTAAAGGATCAACGCTCCTGGGATTTGGGGTTGCTCTCGAAAACCTCATTTGCCACACTGCTGCTTGGAAATGGATTTGGAGTAAAT	8072
CONSENSUS A	g-----aa-----c-----a-----A-----A-----c-----ac-----c-----a-----aa-----c-----ct-----gt-----a-----	7256
A.UG.U455	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7502
A.SE.UGSE8891	-----T-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	7237
A.SE.TZSE8538	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7289
A.SE.UGSE6594	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7242
A.KE.Q2317	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	7509
A.SE.UGSE7535	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7313
A.SE.SOSE7253	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7267
A.UG.92UG037	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	7433
A.SE.SE8131	-----T-----R-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7450
A2.CD.97CDKTB48	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	7382
AG.CI.94CI017	-----g-----g-----g-----g-----g-----g-----g-----g-----g-----g-----g-----g-----g-----g-----g-----	7444
CONSENSUS B	-----g-----g-----g-----g-----g-----g-----g-----g-----g-----g-----g-----g-----g-----g-----g-----	7990
B.AU.MBC925	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	8091
B.AU.MBC200	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	8097
B.GA.OYI	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7611
B.GB.MANC	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	8032
B.GB.CAMI	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	8073
B.CN.RL42	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	7425
B.DE.D31	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	7599
B.DE.HAN2	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	7534
B.ES.89SP061	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	7850
B.NL.ACH320A	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	8076
B.IW.LM49	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	8043
B.US.DH123	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	8041
B.US.SF2	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	8077
B.US.NY5CG	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	8009
B.US.WC1PRL8	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	7641
B.US.YU2	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	8033
B.US.JRGSF	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	8059
B.US.JRFL	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	7361
B.US.MN	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	8089
B.US.RCSG3	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	7604
B.US.WEAU160	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	8079
B.US.RF	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	7817
B.US.WR27	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	7450
CONSENSUS C	-----ta-----A-----a-----at-----a-----a-----g-----c-----ga-----a-----c-----a-----c-----gct-----g-----c-----ctcc-----t-----	7384
C.BR.92BR025	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7409
C.BW.96BW0402	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7551
C.BW.96BW1104	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7399
C.BW.96BW15C02	-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7529
C.BW.96BW01B03	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7518
C.BW.96BW0502	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7555
C.BW.96BW16B01	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7429
C.BW.96BW1210	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7448
C.ET.FTH2220	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7480
C.IN.301999	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7443
C.IN.21068	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7362
C.IN.301905	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7404
C.IN.301904	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	8047
C.IN.93IN101	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7374
C.IN.11246	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7390
D.UG.94UG114	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7576
D.CD.NDK	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7608
D.CD.ELI	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	8042
D.KE.MB2059	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	8065
D.CD.2226	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7593
D.CD.842E085	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7376
F1.BR.93BR020	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7368
F1.FI.FIN9363	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7349
F1.BE.VI850	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7239
F1.FR.MF411	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7231
F2.CM.MF255C	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7364
F2.CM.MF257C	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	

Strain	Genomic Region	Sequence	Position
B. FR. HXB2R	Stem I'	GCGTGTGGAAAGATACCTAAAGGATCAACACCTCTGGGATTTGGGGTCTCTCGAATACTCCTGTGCGCCTTTGGAAATCGTAGTTGGAGTAAAT	8072
G. BE. DRGBL	Stem I'	---C---A---C---C---A---C---CA---AA---C---CA---C---CA---AA---C---CA---C---CA---	8011
G. NG. 92NG083	Stem I'	---A---A---C---A---C---C---A---A---C---C---C---A---A---C---C---C---C---CA---C---CA---	7398
G. SE. SE6165	Stem I'	---A---A---C---A---C---C---A---A---C---C---C---A---A---C---C---C---C---CA---C---CA---	7490
G. FI. HH87932	Stem I'	---C---A---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---	7442
H. BE. VI991	Stem I'	---A---A---C---A---C---C---A---A---C---C---C---A---A---C---C---C---C---CA---C---CA---	7471
H. BE. VI997	Stem I'	---A---A---C---A---C---C---A---A---C---C---C---A---A---C---C---C---C---CA---C---CA---	7386
H. CF. 90CF056	Stem I'	---A---A---C---A---C---C---A---A---C---C---C---A---A---C---C---C---C---CA---C---CA---	7391
J. SE. SE91733	Stem I'	---A---A---C---A---C---C---A---A---C---C---C---A---A---C---C---C---C---CA---C---CA---	7386
J. SE. SE92809	Stem I'	---A---A---C---A---C---C---A---A---C---C---C---A---A---C---C---C---C---CA---C---CA---	7379
K. CM. MP835C	Stem I'	---A---A---C---A---C---C---A---A---C---C---C---A---A---C---C---C---C---CA---C---CA---	7332
K. CD. EQP811C	Stem I'	---A---A---C---A---C---C---A---A---C---C---C---A---A---C---C---C---C---CA---C---CA---	7955
O. CM. ANP70	Stem I'	A---CT---A---CC---TA---C---A---G---A---G---A---G---A---G---G---A---G---A---G---A---G---A---G---A---	8126
O. CM. MPF5180	Stem I'	A---CT---A---CC---TTA---C---A---G---G---A---G---A---G---G---A---G---A---G---A---G---A---G---A---	8143
O. SN. MP1299	Stem I'	A---CT---A---CC---TAA---C---A---G---A---G---A---G---G---A---G---A---G---A---G---A---G---A---	8185
O. SN. MP1300	Stem I'	A---CT---A---CC---TAA---C---A---G---A---G---A---G---G---A---G---A---G---A---G---A---G---A---	8152
N. CM. YBF30	Stem I'	T---A---A---T---G---G---A---A---AA---CC---A---C---A---A---A---A---A---A---A---A---A---A---	7620
N. CM. YBF106	Stem I'	T---GA---A---T---G---A---G---A---A---AA---TC---A---C---A---A---A---A---A---A---A---A---	7575
CONSENSUS 01	Stem I'	g a t a t g a g a	7984
01 AE. TH. 93TH253	Stem I'	A---T---A---AC---C---A---A---C---C---A---C---C---A---C---C---A---C---C---C---CA---C---CA---	8056
01 AE. TH. CM240	Stem I'	A---T---A---AC---C---A---A---C---C---A---C---C---A---C---C---A---C---C---C---CA---C---CA---	7625
01 AE. CF. 90CF402	Stem I'	A---T---A---AC---C---A---A---C---C---A---C---C---A---C---C---A---C---C---C---CA---C---CA---	8135
01 AE. TH. 92TH022	Stem I'	A---T---A---AC---C---A---A---C---C---A---C---C---A---C---C---A---C---C---C---CA---C---CA---	8020
01 AE. TH. 95THH047	Stem I'	A---T---A---AC---C---A---A---C---C---A---C---C---A---C---C---A---C---C---C---CA---C---CA---	8029
01 AE. CF. 11697	Stem I'	A---T---A---AC---C---A---A---C---C---A---C---C---A---C---C---A---C---C---C---CA---C---CA---	8024
01 AE. TH. 93TH057	Stem I'	A---T---A---AC---C---A---A---C---C---A---C---C---A---C---C---A---C---C---C---CA---C---CA---	8009
01 AE. TH. 93TH065	Stem I'	A---T---A---AC---C---A---A---C---C---A---C---C---A---C---C---A---C---C---C---CA---C---CA---	7994
01 AE. CF. 4071	Stem I'	A---T---A---AC---C---A---A---C---C---A---C---C---A---C---C---A---C---C---C---CA---C---CA---	7998
CONSENSUS 02	Stem I'	---C---g---aga---?g---c---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	7233
02 AG. SE. 7812	Stem I'	---C---A---GC---GT---T---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	7462
02 AG. GH. G829	Stem I'	---C---A---GC---GT---T---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	7267
02 AG. DJ. DJ263	Stem I'	---C---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---	7447
02 AG. DJ. DJ264	Stem I'	---C---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---	7406
02 AG. NG. I8NG	Stem I'	---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---	7578
02 AG. SN. MP1211	Stem I'	---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---	7374
02 AG. SN. MP1213	Stem I'	---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---	7354
02 AG. CM. ME607	Stem I'	---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---	7552
02 AB. RU. 98001	Stem I'	---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---G---	7387
03 AB. RU. KALI53-2	Stem I'	---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---	7792
04 CPX. GR. 97PVMY	Stem I'	---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---	8123
04 CPX. GR. 97PVGH	Stem I'	---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---	8171
04 CPX. CV. 94CY0323	Stem I'	---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---C---	7442
05 DF. BE. VII1310	Stem I'	---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---	7478
05 DF. BE. VI961	Stem I'	---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---	7429
06 CPX. ML. 95ML84	Stem I'	---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---	7276
06 CPX. AU. BFP90	Stem I'	---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---	8120
06 CPX. ML. 95ML127	Stem I'	---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---G---A---	8065
06 CPX. SN. 97SE1078	Stem I'	---GT---G---C---G---C---G---C---G---C---G---C---G---C---G---C---G---C---G---C---	8110
10 CD. BFL061	Stem I'	---G---C---G---C---G---C---G---C---G---C---G---C---G---C---G---C---G---C---	7423
10 CD. BFL110	Stem I'	---G---C---G---C---G---C---G---C---G---C---G---C---G---C---G---C---G---C---	7437
10 CD. BFL071	Stem I'	---G---C---G---C---G---C---G---C---G---C---G---C---G---C---G---C---G---C---	7456
11 CPX. CM. MP818	Stem I'	---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---	8050
11 CPX. FR. MP1298	Stem I'	---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---	8069
11 CPX. FR. MP1307	Stem I'	---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---	8107
11 CPX. GR. GR17	Stem I'	---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---	7339
CPZ. GA. STVCPZGAB	Stem I'	---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---	8112
CPZ. US. STVCPZUS	Stem I'	---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---	8077
CPZ. CM. STVCPZCM3	Stem I'	---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---	7485
CPZ. CM. STVCPZCM5	Stem I'	---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---	7791
CPZ. CD. STVCPZANT	Stem I'	---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---T---G---G---A---	7528

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz complete genomes

B. FR. HXB2R CONSENSUS AAAATCTCTGGAACAGATTGGAATCAC.....ACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTCACACAGCTTAATACACACTCTTAACTGAAGATCGCAAAAACCACGAGAAAAGAAATG	8193
A.UG.0455aa-t-tcagaatg-g--A--gata.....t-a-----gc-at-----T-TaacC-a---ag---g-gaa-----g-----a	7276
A.SE.UGSE8891GT--C-A-AGTG--A--G-GA--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7323
A.SE.TZSE8538A--A-AGTG--A--G-GA--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7358
A.SE.UGSE6594A--A-AGTG--A--G-GA--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7410
A.KE.Q2317TAC-TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7630
A.SE.UGSE7535A-A-TG-A--A--G-A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7434
A.SE.SOSE7253TACACT--A--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7388
A.UG.92UG037GA--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7554
A.SE.SEB131-a-t-tcggataag--aaAa-c.....t-c-----t-g-g-----a-Ga-aa-gac-----agetta--T-caact-aa-tg-g---g-g-ac-a-a-----g	7503
A2.CD.97CDKTB48A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7565
A2.CY.94CY017A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7503
CONSENSUS BA--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	8111
B.AU.MBC925A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	8212
B.AU.MBC200A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	8218
B.GA.OYIA--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7732
B.GB.MANCA--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	8153
B.GB.CAMIA--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	8194
B.CN.RL42A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7546
B.DE.D31A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7720
B.DE.HAN2A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7655
B.ES.89SP061A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7771
B.NL.ACH320AA--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	8197
B.TW.LM49A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	8164
B.US.DH123A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	8162
B.US.SF2A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	8198
B.US.NB5CGA--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	8130
B.US.WC192A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7762
B.US.YU2A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	8154
B.US.JRCSFA--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	8180
B.US.JRFLA--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7482
B.US.MNA--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	8210
B.US.WEAL160A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7725
B.US.RFA--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	8200
B.US.WR27A--A--TG--A--G--A--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7738
CONSENSUS Cgat-tcaagaag-g--gaAa.....t-c-----t-g-a-agt-t-ca-aaAC--T-cagg-gc--a-ac-g-a-a-----g-aaaa-----	7571
C.BR.92BR025-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7505
C.BW.96BW0402-C-AA--GG-A--G-GA--T-A-----C-C-A-----G-A-----T-TCAAC--G-----G-----G-----G	7672
C.BW.96BW1104-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7520
C.BW.96BW15C02-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7650
C.BW.96BW01B03-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7639
C.BW.96BW0502-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7676
C.BW.96BW16B01-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7676
C.BW.96BW1210-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7550
C.ET.ETH2220-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7601
C.IN.301999-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7569
C.IN.21068-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7864
C.IN.301905-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7833
C.IN.301904-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7825
C.IN.93IN101-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7825
C.IN.11246-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	8168
D.UG.94UG114-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7495
D.CD.NDK-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7511
D.CD.ELI-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7697
D.KE.MB2059-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7729
D.CD.2226-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	8163
D.CD.84ZR085-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7714
F1.BR.93BR020-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7497
F1.FI.FIN9363-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7479
F1.BE.VI850-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7470
F1.FR.MF411-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7560
F2.CM.ME255C-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7352
F2.CM.ME257C-G--AA--GG-T--A-----T-T-----G-T-----G-T-----T-AGG-GC--C-----G-----G-----G-----G	7385

Tat Rev intron end // Tat Rev exon 2 start Premature stop in Tat <-|

B.FR.HXB2R TGTACTTTTATAGTCAATAGACTTAGCCAGGGATATTCACCATTTTCAGACCCACTCCCA...ACCCGAGG...GGA...CCCGACAGCCCGAGGAATAGAAAGAGGATGGAGAGAGA 8444

A.UG.U455 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7627

A.SE.UGSE8891 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7674

A.SE.TZSE8538 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7609

A.SE.UGSE6594 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7661

A.KE.Q2317 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7614

A.SE.UGSE7535 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7881

A.SE.UGSE7253 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7685

A.UG.92UG037 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7639

A.SE.SE8131 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7805

A2.CD.97CDT848 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7754

A3.CY.94CY017 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7816

CONSENSUS B -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8362

B.AU.MBC925 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8463

B.GA.OYI -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7983

B.GB.MANC -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8404

B.GB.CAM1 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8445

B.CN.RL42 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7797

B.DE.D31 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7971

B.DE.HAN2 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7906

B.ES.89SP061 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8022

B.NL.ACH320A -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8448

B.ITW.LM49 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8415

B.US.DHL23 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8413

B.US.SF2 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8449

B.US.NY5CG -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8381

B.US.WC1PR18 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8013

B.US.YU2 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8405

B.US.JRCSF -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8431

B.US.JRFL -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7733

B.US.MN -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8461

B.US.BCSG3 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7976

B.US.WEAU160 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8451

B.US.RF -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7989

B.US.WR27 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7922

CONSENSUS C -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7556

C.BR.92BR025 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7781

C.BW.96BW0402 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7923

C.BW.96BW1104 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7771

C.BW.96BW15C02 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7901

C.BW.96BW01B03 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7890

C.BW.96BW0502 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7927

C.BW.96BW16B01 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7801

C.BW.96BW1210 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7820

C.ET.FTH2220 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7852

C.IN.301999 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7915

C.IN.21068 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7734

C.IN.301905 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7776

C.IN.301904 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8419

C.IN.93IN101 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7746

C.IN.11246 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7762

D.UG.94UG114 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7948

D.CD.NDK -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8414

D.CD.ELI -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 8437

D.CD.Z226 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7965

D.KE.MB2059 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7748

D.CD.84ZR085 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7740

F1.BR.93BR020 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7721

F1.FI.FIN9363 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7611

F1.BE.VI850 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7611

F1.FR.MF411 -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7603

F2.FR.MF255C -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7636

F2.CM.MF255C -a-c-t-t-g-g-a-a-t-t-t-g-g-g-t-t-c-atac-a-aa-Aag...gt...t-a-----ggg-A-----g-----CA 7636

B. FR. HXB2R	GACAGACAGATCCATTGATGACCGGATCTTGGCACTTATCTGGGACGATCTCCGGAGCCTCTGCTCTTACAGCTTACCCAGGATTTACTCTTGTATTAACGAGGATTTGTGGAAC	8574
G. BE. DRCEBL	-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	8513
G. NG. 92NG083	-----C-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7900
G. SE. SE6165	-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7992
G. FI. HH87932	-----A-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7944
H. BE. VI991	-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7973
H. BE. VI997	-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7888
H. CF. 90CF056	-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7893
J. SE. SE91733	-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7888
J. SE. SE92809	-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7881
K. CM. MP535C	-----A-----A-----AG-----G-----C-----T-----A-----T-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----	7733
K. CD. FOE11C	-----A-----A-----AG-----G-----C-----T-----A-----T-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----	7756
O. CM. AN70	-----G-----GCC-----G-----GG-----ACC-----CGCC-----C-----A-----G-----C-----GT-----G-----ACAG-----C-----CA-----A-----CAA-----NAAT-----G-----GG-----CT-----	8622
O. CM. MP5180	-----G-----GCC-----G-----GG-----ACC-----CGCC-----C-----A-----G-----C-----GT-----G-----ACAG-----C-----CA-----A-----CAA-----NAAT-----G-----GG-----CT-----	8622
O. SN. MP1299	-----G-----GCC-----G-----GG-----ACC-----CGCC-----C-----A-----G-----C-----GT-----G-----ACAG-----C-----CA-----A-----CAA-----NAAT-----G-----GG-----CT-----	8648
N. CM. YBF300	-----G-----GCT-----G-----A-----CC-----CGCC-----C-----A-----T-----C-----GT-----G-----ACAG-----C-----CA-----A-----CAA-----NAAT-----T-----G-----GG-----	8690
N. CM. YBF106	-----G-----G-----G-----G-----G-----T-----CA-----T-----G-----G-----G-----C-----CA-----A-----TGA-----CT-----C-----CA-----G-----C-----G-----	8125
CONSENSUS 01	-----G-----g-----g-----cg-----gc-----a-----a-----g-----c-----ag-----c-----ag-----c-----ag-----c-----ag-----c-----ag-----c-----ag-----c-----ag-----	8083
01 AE. TH. 93TH253	-----G-----G-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	8484
01 AE. TH. CM240	-----G-----G-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	8561
01 AE. CF. 90CF402	-----G-----G-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	8127
01 AE. TH. 92TH022	-----G-----G-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	8637
01 AE. TH. 95THH047	-----G-----G-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	8522
01 AE. CF. 11697	-----G-----G-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	8531
01 AE. TH. 93TH057	-----G-----G-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	8526
01 AE. TH. 93TH065	-----G-----G-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	8511
01 AE. CF. 4071	-----A-----A-----AG-----G-----C-----T-----A-----T-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	8496
CONSENSUS 02	-----a-----g-----a-----cg-----tc-----a-----g-----c-----ag-----t-----gc-----g-----c-----c-----ag-----t-----gc-----g-----c-----c-----ag-----	8500
02 AG. SE. 7812	-----G-----G-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7732
02 AG. GH. G829	-----G-----G-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7964
02 AG. DJ. DJ263	-----G-----G-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7769
02 AG. DJ. DJ264	-----G-----G-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7949
02 AG. NG. IBNG	-----A-----A-----AGGG-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7908
02 AG. SN. MP1211	-----A-----A-----AGGG-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	8080
02 AG. SN. MP1213	-----G-----G-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7779
02 AG. CM. MP807	-----G-----G-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7756
03 AB. RU. 98001	-----C-----C-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7754
03 AB. RU. KALI53-2	-----C-----C-----G-----G-----G-----T-----A-----G-----GC-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	7889
04 CPX. GR. 97FVNY	-----AT-----GA-----G-----T-----C-----G-----G-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	8294
04 CPX. GR. 97PVCH	-----AG-----G-----T-----C-----C-----G-----G-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	8628
04 CPX. CV. 94CV0323	-----AG-----G-----T-----C-----C-----G-----G-----C-----A-----G-----C-----G-----G-----C-----G-----C-----G-----	8673
05 DF. BE. V11310	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7947
05 DF. BE. V1961	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7980
06 CPX. ML. 95ML84	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7931
06 CPX. AU. BFP90	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7778
06 CPX. ML. 95ML127	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	8622
06 CPX. SN. 97SEL078	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	8567
10 CD. BEL101	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	8612
10 CD. BEL110	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7925
10 CD. BEL1071	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7959
11 CPX. CM. MP818	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7955
11 CPX. FR. MP1298	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	8552
11 CPX. FR. MP1307	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	8574
11 CPX. GR. GR17	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	8609
CPZ. GA. SIVCPZGAB	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7847
CPZ. US. SIVCPZUS	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	8614
CPZ. CM. SIVCPZCAM3	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	8582
CPZ. CM. SIVCPZCAM5	-----G-----G-----A-----A-----G-----C-----T-----CA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	7990
CPZ. CD. SIVCPZANT	-----G-----GAT-----G-----GG-----GGCC-----GCA-----C-----G-----T-----G-----G-----G-----C-----A-----G-----G-----T-----GAT-----GC-----AC-----GT-----AC-----GT-----	8296
Env gp41	D_R_D_R_S_I_R_L_V_N_G_S_L_A_L_I_W_D_D_L_R_S_L_C_L_F_S_Y_H_R_L_R_D_L_L_L_I_V_T_R_I_V_E	Env
Rev	R_Q_R_Q_I_H_S_I_S_E_R_I_L_G_T_Y_L_G_R_S_A_E_P_V_P_L_Q_L_P_P_L_E_R_L_T_L_L_D_C_N_E_D_C_G_T	Rev
Tat	T_E_T_D_P_F_D_S_Tat end most strains	Tat

HIV-1/SIVcpz complete genomes

B. FR. HXB2R ... Rev cds end <-| ... TTTCTGGAGCC ... AAGAGCTCAAGGAGCTGAGACT ... AGGGGTGGAGCCCTCAATATTTGGTGGAAATCTCTACAGATTTGGAGTTCAGGAACTAAGAATAGTCTTACCTTGCATTAATCCACAGCCAT ...

HIV-1/SIVcpz complete genomes


```

B.FR_HXB2R      GGTGGCAAGTGCTCAAAAAGTACTGTGATTGGGCTACTCTTAAGGGAAAGAANTGAGACGAGCTT.....GAGCCAGCAGCAGATAGGTTGGGAGCAG
CONSENSUS A    --tggc-a-----aagta-catagtggatggcc-gagg-l-yy---gaa-aagacaaa-t.....????????????????????????????????????????
A.UG_U455      -G-----CAGAG-G-A-----GAG-T-A-----GA-A-----GCTG-----A-AG-A-A-----
A.SE_UGSE8891  -A-CAAG-G-G-----AG-T-G-----GC-A-----AG-T-----A-AG-A-A-----
A.SE_TZSE8538  -A-CAAAGAG-----T-GAG-T-----GAGA-T-----A-G-A-A-----A-AG-A-A-----
A.SE_UGSE6594  -CA-AG-G-----GAGA-T-----GAGA-T-----T-C-----CCTTCAG-----A-AG-A-A-----
A.KE_Q2317     -AA-AG-G-----GAG-T-----AA-AG-T-----G-----CCT-----CCAG-A-A-----
A.SE_UGSE7535  -CA-AG-G-----GAG-T-----GC-A-----AGA-----CCT-ACAG-A-A-----
A.SE_SOSE7253  -G-----T-CA-AG-G-----GAG-T-----A-A-----CCTACAGCAGCAGGAAAGAAAAGAACAGCCCTTA-----A-AG-A-A-----
A.UG_92UG037   -AA-----G-T-G-----GAG-T-----A-A-----CGAGCTCCAGCTCATACTCCAGCTCTTA-----ACAG-A-A-----
A.SE_SE8131    -G-----G-CA-ACCA-----G-A-T-----G-----CCTCCACAGCAGCAAGAACAGCAGTGTCTCCAG-----C-G-A-A-----
A2_CD_97CDKTB48
A2_CV_94CY017  -G-----c-ag-g-----acgtagtgtagtggg-----cc-ac-g-----g-----gagacg-9ct-????????????????????????????-atg-g
CONSENSUS B    -A-AT-----TGCC-----GCCCAGGG.....GAAAGAGGAGAGCAGCTT-----G-----
B.AU_MBC925    -T-----A-AG-----AAA-----AG-----GAGCTACAGCCACT-----G-A-----
B.AU_MBC200    -C-----C-GG-----T-G-----CA-----GAGCCAGCAAGAGGGGAGAAAGAAACAAGCT-----AG-----
B.GB.MAMC      -CA-A-T-----G-----A-----GAGCCACGAGCT-----GG-----
B.CN_RL42      -G-----G-----G-----A-----A-----G-----G-A-----
B.DE_D31       -T-----G-----A-----A-----GAGCCT-----GG-----
B.DE_HAN2      -T-----AAA-----A-----A-----G-----G-----
B.ES_89SP061   -C-----A-ACC-----T-----A-A-----A-----G-----
B.NL_ACH320A   -C-----GT-G-----G-G-G-----T-----A-A-----A-----TAG-----
B.US_DHL23     -C-----A-GG-----T-----G-A-----A-----GAGCCAGCAGCT-----G-----
B.US_SF2       -C-----G-G-G-----G-----A-----A-----G-----G-----
B.US_NY5CG     -C-----A-AAA-----A-----AA-A-C-----A-----GAGCCAGCAGCAGGGAGTA-GAG-----GG-A-A-----
B.US_WC1PR18  -C-----C-----G-----A-----A-----A-----GAGCCAGCAGCAAGAGAGTAGAGCAAGCT-----GG-A-A-----
B.US_YU2       -CA-----CC-----C-----T-----C-----GAGCCAGCAGCAGATAGGCTGAGACAACT-----GG-A-A-----
B.US_JRCSF     -C-----CC-----C-----T-----GAGCCAGCAGCAGATAGGCTGAGACAACT-----GG-A-A-----
B.US_JRFL      -C-----C-----C-----T-----G-----G-----A-----A-----AGACCAGCAGTAGAGGGGAGAACT.....A-A-A-----
B.US_MN        -C-----TG-A-GG-----A-G-----A-----A-----GAGCCAGCAGCT-----G-----
B.US_BCSG3     -C-----GT-G-----G-A-----G-----A-----CA-AA-----G-----
B.US_WEAU160   -C-----AGA-GG-----G-G-A-----D-----D-----G-----
B.US_RF        -C-----GG-----G-G-A-----D-----D-----G-----
B.US_WR27      -Gg-9gtg--a--tgcagtatagttggat--gcta--A-aa--g--c-a-????????????????????????????????????????-a-ag---
CONSENSUS C    -GAA-A-----T-----ACAG-G-C-----G-A-----A-----ACAGCAGCAGCAGGGAGTA-GAG-----GG-A-A-----
C.BW_96BW0402 -G-----C-A-AG-----A-----AA-A-C-----A-----GAGCCAGCAGCAGGGAGTA-GAG-----GG-A-A-----
C.BW_96BW1104 -C-----A-AAA-----A-----AA-A-C-----A-----GAGCCAGCAGCAGGGAGTA-GAG-----GG-A-A-----
C.BW_96BW15C02 -T-----C-A-AG-----T-----G-----A-----A-----GAGCCAGCAGCAGGGAGTA-GAG-----GG-A-A-----
C.BW_96BW0502 -G-----T-C-----C-----C-----T-----G-----A-----A-----A-----A-----AGACCAGCAGTAGAGGGGAGAACT.....A-A-A-----
C.BW_96BW16B01 -G-----C-----C-----C-----C-----C-----A-----T-----A-----A-----A-----GAGCCAGCAGCT-----G-----
C.BW_96BW1210 -G-----C-AT-----T-----CCAG-A-----G-A-----A-----A-----A-----GAGCCAGCAGCT-----G-----
C.ET_ETH2220   -G-----T-C-CA-AG-----GA-A-----A-----C-----A-----G-----G-----
C.IN_301999    -T-C-CA-AG-----GA-A-----A-----G-----A-----A-----G-----G-----
C.IN_21068     -C-CA-AG-----G-A-----A-----G-----A-----A-----G-----G-----
C.IN_301905    -T-C-CA-AG-----GA-A-----A-----G-----A-----A-----G-----G-----
C.IN_301904    -T-C-CA-AG-----GA-A-----A-----G-----A-----A-----G-----G-----
C.IN_93IN101   -T-C-CA-AG-----GAGA-A-----G-----A-----A-----G-----G-----
C.IN_11246     -T-C-CA-AG-----G-----G-----A-----A-----A-----A-----G-----G-----
D.UG_94UG114  -A-----T-AG-----G-----A-----A-----A-----A-----A-----A-----G-----G-----
D.CD_NDK       -A-----CD.ELI-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----G-----G-----
D.KE_MB2059    -A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----G-----G-----
D.CD_Z226      -GA-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----G-----G-----
D.CD_84ZR085   -A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----G-----G-----
F1_BR_93BR020  -A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----G-----G-----
F1_FT_FIN9363   -A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----G-----G-----
F1_BE_VI850     -C-----A-AG-----G-----G-----A-----G-----A-----C-----C-----CCTCCAGCCT-----GG-----
F1.FR_MF411     -C-----A-AG-----G-----G-----A-----G-----A-----C-----A-----C-----CCTA-----AG-----
F2.CM_MF255C   -A-AG-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----G-----G-----
F2.CM_MF257C   -A-AG-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----G-----G-----
8893          .....GAGCCAGCAGCAGATAGGTTGGGAGCAG
8905          -aFAG-A-A-
8944          -A-AG-A-A-
8077          -A-AG-A-A-
8110          -A-AG-A-A-
8087          -A-AG-A-A-
8351          -CCAG-A-A-
8155          -CCAG-A-A-
8109          -CCAG-A-A-
8308          -A-AG-A-A-
8271          -ACAG-A-A-
8254          -A-AG-A-A-
8316          -G-A-A-
8911          -atg-g
8912          -g
8939          -G-A-
8447          -AG
8886          -GG
8906          -GG
8246          -G-A-
8420          -GG
8355          -G
8471          -G
8902          -G
8967          -G
8977          -TAG
8910          -G
8830          -A-
8462          -G
8884          -G
8910          -TAG
8212          -TAG
8913          -T
8425          -G
8900          -AG
8438          -G
8971          -GG
8926          -ag
8248          -A-C
8393          -A
8262          -A
8371          -A
8360          -A
8394          -A
8346          -A
8271          -CT
8290          -A
8322          -A
8285          -A
8204          -A
8246          -G-AG-A-A-
8868          -A-A-
8216          -C
8211          -AG
8397          -G-A-
8429          -A-T-----T-----
8863          -T-----T-----
8886          -G-A-
8438          -G
8206          -GG
8192          -CCTCCAGCCT-----GG-----
8170          -AG
8060          -AG
8940          -GAA-----GAA-----
8073          -AG-----AG-----

```


HIV-1/SIVcpz complete genomes

B_FR_HXB2R 9017
CONSENSUS A
A.UG_U455 8211
A.SE_UGSE8891 8468
A.SE_TZSE8538 8395
A.SE_UGSE6594 8205
A.KE_Q2317 8469
A.SE_UGSE7535 8273
A.SE_SOSE7253 8227
A.UG_92UG037 8426
A.SE_SE8131 8395
A2_CD_97CDKTB48 8378
A2_CY_94CY017 8443
CONSENSUS B 8925
B.AU_MBC925 9036
B.AU_MBC200 9066
B.GA_OYI 8571
B.GB_MANC 9010
B.GB_CAMI 9030
B.CN_RL42 8370
B.DE_D31 8544
B.DE_HAN2 8479
B.ES_89SP061 8595
B.NL_ACH320A 9026
B.TM_LM49 8991
B.US_DHI123 9004
B.US_SF2 9034
B.US_NY5CG 8954
B.US_WCIPR18 8886
B.US_YU2 9008
B.US_JRC5F 9034
B.US_JRFL 8336
B.US_MN 9037
B.US_BCSG3 8549
B.US_WEAU160 9018
B.US_RF 8568
B.US_WR27 8395
CONSENSUS C 8353
C.BW_92BR025 8375
C.BW_96BW0402 8520
C.BW_96BM1104 8386
C.BW_96BW15C02 8498
C.BW_96BW01B03 8487
C.BW_96BW05I02 8521
C.BW_96BW16B01 8470
C.BW_96BW12110 8398
C.ET_ETH2220 8417
C.IN_21068 8449
C.IN_30199 8388
C.IN_301905 8331
C.IN_301904 8973
C.IN_93JN101 8995
C.IN_11246 8343
C.IN_11246 8335
D.UG_94UG114 8524
D.CD_NDK 8556
D.CD_ELI 8990
D.KE_MB2059 9013
D.CD_Z226 8565
D.CD_84ZR085 8330
F1_BR_93BR020 8316
F1_FT_FIN9363 8394
F1_BE_VI850 8184
F1_FR_MF411 8161
F2_CM_MF255C 8194
F2_CM_MF257C

HIV-1/SIVcpz Complete Genomes

Table with columns: Accession ID, Strain Name, and Nucleotide Sequence. Includes strains such as B.FR.HXB2R, G.BE.DRCBL, G.ING.92NG083, G.SE.SE6165, G.FI.HH87932, H.BE.VI991, H.BE.VI997, H.CF.90CF056, J.SE.SE91733, J.SE.SE92809, K.CM.MP535C, K.CD.EOTB11C, O.CM.ANT70, O.CM.MP5180, O.SN.MP1299, O.SN.MP1300, N.CM.YBF30, N.CM.YBF106, CONSENSUS 01, 01.AE.TH.93TH253, 01.AE.TH.CM240, 01.AE.CF.90CF402, 01.AE.TH.92TH022, 01.AE.TH.95TNLH047, 01.AE.CF.11697, 01.AE.TH.93TH057, 01.AE.TH.93TH065, 01.AE.CF.4071, CONSENSUS 02, 02.AG.SE.7812, 02.AG.GH.G829, 02.AG.DJ.DJ263, 02.AG.DJ.DJ264, 02.AG.ING.IBNG, 02.AG.SN.MP1211, 02.AG.SN.MP1213, 02.AG.CM.MB807, 03.AB.RU.98001, 03.AB.RU.KALI53-2, 04.CPX.GR.97PVNY, 04.CPX.GR.97PVCH, 04.CPX.CY.94CY0323, 05.DF.BE.VI1310, 05.DF.BE.VI961, 06.CPX.ML.95ML84, 06.CPX.AU.BPP90, 06.CPX.ML.95ML127, 06.CPX.SN.97SEI078, 10.CD.BET061, 10.CD.BEL110, 10.CD.BEL1071, 11.CPX.CM.MB818, 11.CPX.FR.MP1298, 11.CPX.FR.MP1307, 11.CPX.GR.GR17, CPZ.GA.SIVCPZGAB, CPZ.US.SIVCPZUS, CPZ.US.SIVCPZCAM3, CPZ.CM.SIVCPZCAM5, CPZ.CD.SIVCPZANT.

HIV-1/SIVcpz complete genomes

Nef premature end in HXB2 <-|

ACACRAGGCTACTTC...CCTGATTAGCAGAACTACACACAGCCAGCCAGGG...GTCAGATATCCACTGACCTTTGGATGGTGCTACAGCTTAGTACAGCCAGCATAGATA...GAAGAGGCCA

9265
9268
9270
9272
9274
9276
9278
9280
9282
9284
9286
9288
9290
9292
9294
9296
9298
9300
9302
9304
9306
9308
9310
9312
9314
9316
9318
9320
9322
9324
9326
9328
9330
9332
9334
9336
9338
9340
9342
9344
9346
9348
9350
9352
9354
9356
9358
9360
9362
9364
9366
9368
9370
9372
9374
9376
9378
9380
9382
9384
9386
9388
9390
9392
9394
9396
9398
9400
9402
9404
9406
9408
9410
9412
9414
9416
9418
9420
9422
9424
9426
9428
9430
9432
9434
9436
9438
9440
9442
9444
9446
9448
9450
9452
9454
9456
9458
9460
9462
9464
9466
9468
9470
9472
9474
9476
9478
9480
9482
9484
9486
9488
9490
9492
9494
9496
9498
9500
9502
9504
9506
9508
9510
9512
9514
9516
9518
9520
9522
9524
9526
9528
9530
9532
9534
9536
9538
9540
9542
9544
9546
9548
9550
9552
9554
9556
9558
9560
9562
9564
9566
9568
9570
9572
9574
9576
9578
9580
9582
9584
9586
9588
9590
9592
9594
9596
9598
9600
9602
9604
9606
9608
9610
9612
9614
9616
9618
9620
9622
9624
9626
9628
9630
9632
9634
9636
9638
9640
9642
9644
9646
9648
9650
9652
9654
9656
9658
9660
9662
9664
9666
9668
9670
9672
9674
9676
9678
9680
9682
9684
9686
9688
9690
9692
9694
9696
9698
9700
9702
9704
9706
9708
9710
9712
9714
9716
9718
9720
9722
9724
9726
9728
9730
9732
9734
9736
9738
9740
9742
9744
9746
9748
9750
9752
9754
9756
9758
9760
9762
9764
9766
9768
9770
9772
9774
9776
9778
9780
9782
9784
9786
9788
9790
9792
9794
9796
9798
9800

B.FR.HXB2R
G.BE.DRCBL
G.NG.92NG083
G.SE.SE6165
G.FI.HH87932
H.BE.VI991
H.BE.VI997
H.CF.90CF056
J.SE.SE91733
J.SE.SE92809
K.CM.MP535C
K.CD.FOPE11C
O.CM.AN70
O.CM.MP5180
O.SN.MP1299
O.SN.MP1300
N.CM.YBF30
N.CM.YBF106
CONSENSUS 01
01.AE.TH.93TH253
01.AE.CF.90CF402
01.AE.TH.92TH022
01.AE.TH.95TH047
01.AE.CF.11697
01.AE.TH.93TH057
01.AE.TH.93TH065
01.AE.CF.4071
CONSENSUS 02
02.AG.SE.7812
02.AG.GH.G829
02.AG.DJ.DJ263
02.AG.DJ.DJ264
02.AG.NG.IBNG
02.AG.SN.MP1211
02.AG.SN.MP1213
02.AG.CM.ME807
03.AB.RU.98001
03.AB.RU.KALI53-2
04.CPX.GR.97PMVY
04.CPX.GR.97PVCH
04.CPX.CY.94CY0323
05.DF.BE.VI1310
05.DF.BE.VI961
06.CPX.ML.95ML84
06.CPX.AU.BFP90
06.CPX.ML.95ML127
06.CPX.SN.97SEL078
10.CD.BEL061
10.CD.BEL110
10.CD.BEL1071
11.CPX.CM.MP818
11.CPX.FR.MP1298
11.CPX.FR.MP1307
11.CPX.GR.GR17
CPZ.GA.SIVCPZGAB
CPZ.US.SIVCPZUS
CPZ.CM.SIVCPZCAM3
CPZ.CM.SIVCPZCAM5
CPZ.CD.SIVCPZANT

Nef premature end in HXB2 <-|

T_O_G_Y_F_P_D_S_Q_N_Y_T_P_G_P_G_V_R_Y_P_L_T_F_G_W_C_Y_K_L_V_P_V_E_P_D_K_I_E_E_A_A

Nef premature end in HXB2 <-|

9925
9926
9927
9928
9929
9930
9931
9932
9933
9934
9935
9936
9937
9938
9939
9940
9941
9942
9943
9944
9945
9946
9947
9948
9949
9950
9951
9952
9953
9954
9955
9956
9957
9958
9959
9960
9961
9962
9963
9964
9965
9966
9967
9968
9969
9970
9971
9972
9973
9974
9975
9976
9977
9978
9979
9980
9981
9982
9983
9984
9985
9986
9987
9988
9989
9990
9991
9992
9993
9994
9995
9996
9997
9998
9999

9925
9926
9927
9928
9929
9930
9931
9932
9933
9934
9935
9936
9937
9938
9939
9940
9941
9942
9943
9944
9945
9946
9947
9948
9949
9950
9951
9952
9953
9954
9955
9956
9957
9958
9959
9960
9961
9962
9963
9964
9965
9966
9967
9968
9969
9970
9971
9972
9973
9974
9975
9976
9977
9978
9979
9980
9981
9982
9983
9984
9985
9986
9987
9988
9989
9990
9991
9992
9993
9994
9995
9996
9997
9998
9999

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz complete genomes

Table with 3 columns: Accession Number (e.g., B_FR_HXB2R), Genomic Region (e.g., gag), and Nucleotide Sequence (e.g., ATAAAGGAGAACACAGCTTGTACACCTGTGACCCCTGGATGGATGATGACCCGAGAGAAAGTGTGAGAGTGGAGTTTGCACCCGCTCAGCATTTTCACAGTGGCCGAGAGGTGCATCC). The table lists 257 different HIV-1/SIVcpz complete genomes.

Accession	Sequence
B. FR. HXB2R	ATAAAGGAGAACACAGCTTTTACACCCCTGTGACCCGCGATGCGATGGGATGATGCGCGGAGAGAGAAAGTTGAGTGGAGTTTGGACAGCCCGCTAGCATTTTCATCGTGGCCCGAGAGCTGCATCC
G. BE. DRCEI	-----G-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
G. NG. 92NG083	-----T-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
G. SE. SE6165	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
G. FI. HH87932	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
H. BE. VI991	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
H. BE. VI997	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
H. CF. 90CF056	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
J. SE. SE91733	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
J. SE. SE92809	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
K. CM. MP535C	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
K. CD. FOE111C	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
O. CM. ANT70	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
O. CM. MP5180	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
O. SN. MP1299	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
O. SN. MP1300	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
N. CM. YBF106	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
N. CM. YBF30	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
CONSENSUS 01	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
01 AE. TH. 93TH253	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
01 AE. TH. CM240	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
01 AE. CF. 90CF402	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
01 AE. TH. 92TH022	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
01 AE. TH. 95THH047	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
01 AE. CF. 11697	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
01 AE. TH. 93TH057	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
01 AE. TH. 93TH065	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
01 AE. CF. 4071	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
CONSENSUS 02	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
02 AG. SE. 7812	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
02 AG. GH. G829	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
02 AG. DJ. DJ264	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
02 AG. NG. I8NG	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
02 AG. SN. MP1211	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
02 AG. SN. MP1213	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
02 AG. CM. MP807	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
03 AB. RU. 98001	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
03 AB. RU. KALI53-2	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
04 CPX. GR. 97PVNY	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
04 CPX. GR. 97PVCH	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
04 CPX. CY. 94CY0323	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
05 DF. BE. V11310	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
05 DF. BE. V1961	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
06 CPX. ML. 95ML84	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
06 CPX. AU. BPP90	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
06 CPX. ML. 95ML127	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
06 CPX. SN. 97SE1078	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
10 CD. BEL061	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
10 CD. BEL110	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
10 CD. BEL1071	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
11 CPX. CM. MP818	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
11 CPX. FR. MP1298	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
11 CPX. FR. MP1307	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
11 CPX. GR. GR17	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
CPZ. GA. SIVCPZGAB	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
CPZ. US. SIVCPZUS	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
CPZ. CM. SIVCPZCAM3	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
CPZ. CM. SIVCPZCAM5	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
CPZ. CD. SIVCPZANT	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----
Nef	-----A-----TC-A-----CA-CT-A-----A-----G-GAA-TG-----TA-----G-----CGAGA-----T-A-----

HIV-1/SIVcpz complete genomes

B.FR.HXB2R	GGGACTTCTC	TCF-1 alpha binding	- <- Nef cds end	NF-kappa-B II	9447
CONSENSUS A	---c-acAAGAACTGCTGACATCGAGCTTGCTACAAGGGACTTTCGGCT	8639
A.UG.U455	---T-A	???????	c-c-CA-aag--c-GACa?	---CA-AG--GAC	8906
A.SE.UGSE8891	---T-A	AAAAGACTG--AG	---CA-AG--GAC-G	---CA-AG--GAC-G	8626
A.SE.TZSEF8538	---T-A	---AG	---CA-AG--GAC	---CA-AG--GAC	8658
A.SE.UGSEF6594	---T-A	---AG	---CA-AG--GAC	---CA-AG--GAC	8635
A.KE.Q2317	---T-A	---G	---CA-AG--T-GACT	---CA-AG--T-GACT	8899
A.SE.UGSE7535	---T-A	---AG	---CA-AG--GAC	---CA-AG--GAC	8703
A.SE.SOSE7253	---TT-A	---AG	---CA-AG--GAC	---CA-AG--GAC	8657
A.UG.92UG037	---T-A	---AG	---CA-GAG--GACT	---CA-GAG--GACT	8856
A.SE.SE8131	---T-A	---AG	---A-T-CA-CAG--GACT	---A-T-CA-CAG--GACT	8825
A2.CD.97CDKTB48	---T-A	---AG	---CA-AGA--GAC-CAGAAGTTCGACAA	---CA-AGA--GAC-CAGAAGTTCGACAA	8873
A2.CT.94CY017	---T-A	---AG	---CA-AG--GAGG	---CA-AG--GAGG	8873
CONSENSUS B	ggagctactac	???????	ctgacatcagagctttctacaa?gggactttccgct	9363
B.AU.MBC925	---A	g	---G	---A	9466
B.AU.MBC200	---A	---G	---A-C	---T	9496
B.GA.OYI	---A	---AG	---T	---G	9001
B.GB.MANC	---A	---AG	---G	---T	9440
B.GB.CAML	---T-A	---AG	---T	---T	9460
B.CN.RL42	---CA	---AG	---T-T	---T-T	8800
B.DE.D31	---A	AAAAAAGT--AG	---A	---T-T	8982
B.DE.HAN2	---A	---AG	---C-T	---T-T	891
B.ES.89SP061	---A	---G	---CT	---T	9025
B.NL.ACH320A	---A	---G	---Y	---T	9456
B.IW.LM49	---A	---Y	---Y	---T	9423
B.US.DH123	---A	---AG	---C	---AT	9434
B.US.SF2	---T-A	---AG	---T	---T	9464
B.US.WC1PR18	---A	---G	---A-C	---T	9016
B.US.YU2	---A	---G	---T	---T	9435
B.US.JRCSF	---A	---G	---C	---T	9464
B.US.JRFL	---A	---G	---A	---AT	8766
B.US.MN	---A	---C	---A	---T	9466
B.US.BCSG3	---T	---AG	---T	---T	8979
B.US.WEAU160	---A	---AG	---T	---T	9447
B.US.RF	---A	---AG	---T	---T	8998
B.US.WR27	---A	---AG	---A	---TTCTAC-A	8926
CONSENSUS C	---g-aT-Ac???	---AG	---c-ca-aa??????????999actttccgct??	8784
C.BR.92BR025	---T-A	---AG	---T-CA-AG-ACAA-G-CTGCTGA.CACAGCT	8518
C.BW.96BW0402	---T-A	---AG	---CA-AGGGACTTTCGGCT	8953
C.BW.96BW1104	---A	---AG	---CA-AGGAACTTTCGGCT	8817
C.BW.96BW15C02	---T-A	---AG	---CA-AGGGACTTTCGGCT	8931
C.BW.96BW01B03	---T-AT	---AG	---CA-AGGGACTTTCGGCT	8920
C.BW.96BW0502	---T-A	---AG	---CA-GAGGGACTTTCGGCT	8963
C.BW.96BW16B01	---T-A	---AG	---CA-GAGGGACTTTCGGCT	8903
C.BW.96BW1210	---T-A	---AG	---CA-AGGGACTTTCGGCT	8831
C.ET.FTH2220	---T-A	---AG	---CA-CGGGACTTTCGGCT	8850
C.IN.301999	---TT-A	---AG	---CA-CGGGACTTTCGGCT	8982
C.IN.21068	---TT-A	---AG	---CG-AGGGACTTTCGGCT	8924
C.IN.301905	---TT-A	---AG	---CA-AGGGACTTTCGGCT	8764
C.IN.301904	---TT-A	---AG	---CA-AGGGACTTTCGGCT	8906
C.IN.93IN101	---TT-A	---AG	---CA-AGGGACTTTCGGCT	9415
C.IN.11246	---TT-A	---AG	---CA-AGGGACTTTCGGCT	8768
D.UG.94UG114	---A	---AG	---CTTTCTACAA	8765
D.CD.NDK	---T-A	---AG	---ATTTCTACAA	8954
D.CD.ELI	---T-A	---A	---CTTTCTACAA	8986
D.CD.84ZR085	---D	CAAGACTGAG	---CATCTAAA	8975
F1.BR.93BR020	---A	CGAGACTGAG-T	---CA-A	---GAC-CAGAAGA	8780
F1.FI.FIN9363	---T-A	CGAGACTGAG	---CA-A	---GAC-CAGAAGA	8767
F1.BE.VI850	---T-A	CGAGACTGAG	---CA-A	---GAC-CAGAAGA	8745
F1.FR.ME411	---T-AT	C-A	---CA-A	---GAC-CAGAAGA	8745
F2.CM.ME255C	---TT-A	C-AG	---C-AG	---GAC-CAGAAGA	8555
F2.CM.ME257C	---A	---AG	---AG	---GAC-CAGAAGA	8588

Strain	Genomic Region	Sequence
B.FR.HXB2R	GGAGTACTTC	GGAGTACTTC
G.BE.DRCBL	-----A	-----A
G.NG.92NG083	-----A	-----A
G.SE.SEG165	-----A	-----A
G.FI.HH87932	-----A	-----A
H.BE.VI991	-----A	-----A
H.BE.VI997	-----A	-----A
H.CF.90CF056	-----A	-----A
J.SE.SE91733	-----A	-----A
J.SE.SE92809	-----A	-----A
K.CM.MP535C	-----A	-----A
K.CD.EQPE11C	-----A	-----A
O.CM.ANT70	-----A	-----A
O.CM.MPF5180	-----A	-----A
O.SN.MP1299	-----A	-----A
O.SN.MP1300	-----A	-----A
N.CM.YBF30	-----A	-----A
N.CM.YBF106	-----A	-----A
CONSENSUS 01	-----A	-----A
01.AE.TH.93TH253	-----A	-----A
01.AE.CF.90CF402	-----A	-----A
01.AE.TH.92TH022	-----A	-----A
01.AE.TH.95TH047	-----A	-----A
01.AE.CF.11697	-----A	-----A
01.AE.TH.93TH057	-----A	-----A
01.AE.TH.93TH065	-----A	-----A
01.AE.CF.4071	-----A	-----A
CONSENSUS 02	-----A	-----A
02.AG.SE.7812	-----A	-----A
02.AG.GH.G829	-----A	-----A
02.AG.DJ.DJ263	-----A	-----A
02.AG.DJ.DJ264	-----A	-----A
02.AG.NG.18NG	-----A	-----A
02.AG.SN.MP1211	-----A	-----A
02.AG.SN.MP1213	-----A	-----A
02.AG.CM.ME807	-----A	-----A
03.AB.RU.98001	-----A	-----A
03.AB.RU.KALI53-2	-----A	-----A
04.CPX.GR.97PVMY	-----A	-----A
04.CPX.GR.97PVGH	-----A	-----A
04.CPX.CY.94CY0323	-----A	-----A
05.DF.BE.VI1310	-----A	-----A
05.DF.BE.VI961	-----A	-----A
06.CPX.ML.95ML84	-----A	-----A
06.CPX.AU.BFP90	-----A	-----A
06.CPX.MI.95MI127	-----A	-----A
06.CPX.SN.97SEL078	-----A	-----A
10.CD.BEL061	-----A	-----A
10.CD.BEL110	-----A	-----A
10.CD.BEL071	-----A	-----A
11.CPX.CM.MP818	-----A	-----A
11.CPX.FR.MP1298	-----A	-----A
11.CPX.FR.MP1307	-----A	-----A
11.CPX.FR.MP1307	-----A	-----A
11.CPX.GR.GR17	-----A	-----A
CPZ.GA.SIVCPZGAB	-----A	-----A
CPZ.US.SIVCPZUS	-----A	-----A
CPZ.CM.SIVCPZCAM3	-----A	-----A
CPZ.CM.SIVCPZCAM5	-----A	-----A
CPZ.CD.SIVCPZANT	-----A	-----A
B.FR.HXB2R	-----A	-----A

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz complete genomes

	NF-kappa-B_I	SPI-III	SPI-II	SPI-I	TATA_box	
B.FR.HXB2R	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
CONSENSUS A	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
A.UG.U455	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
A.SE.UGSE8891	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
A.SE.TZSE8538	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
A.SE.UGSE6594	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
A.KE.Q2317	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
A.SE.UGSE7535	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
A.SE.SOSE7253	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
A.UG.92UG037	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
A.SE.SE8131	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
A2.CD.97CDKTB48	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
A2.CD.94CY017	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
CONSENSUS B	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.AU.MBC925	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.AU.MBC200	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.GA.OYI	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.GB.MANC	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.GB.CAML	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.CN.RL42	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.DE.D31	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.DE.HAN2	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.ES.89SP061	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.NL.ACH320A	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.IW.LM49	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.US.DH123	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.US.SF2	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.US.WCIPR18	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.US.YU2	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.US.JRCSF	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.US.JRFL	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.US.MN	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.US.BCSG3	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.US.WEAU160	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.US.RF	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
B.US.WR27	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
CONSENSUS C	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.BR.92BR025	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.BW.96BW0402	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.BW.96BW1104	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.BW.96BW15C02	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.BW.96BW01B03	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.BW.96BW0502	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.BW.96BW16B01	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.BW.96BW1210	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.ET.FTH2220	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.IN.301999	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.IN.21068	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.IN.301905	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.IN.301904	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.IN.93IN101	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
C.IN.11246	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
D.UG.94UG114	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
D.CD.NDK	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
D.CD.ELI	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
F1.BR.93BR020	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
F1.FI.FIN9363	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT
F1.BE.VI850	GGGACTTTCC	AGGAGG	CTGGCCCTGG	CCGGGACTGGGAGTGGC	GAGCCCTCAGATCCTGGCA	TATAAGCAGCTGCTTTT

	NF-kappa-B_I	SPI-III	SPI-I	SPI-II	SPI-I	TATA_box	
B.FR.HXB2R	.GGGACTTTCC	.AGGAGG.CGTGGCTGG	.GCCGGACTGGGAGTGGC	.GAGCCCTCAGATCCCTGCA	TATAA	GCAGCGTCTGT	9528
G.BE.DRCBL		.AG-A	.A-G	.T-A	.AG	.C-C	9515
G.NG.92NG083		.AG-A	.A-G	.T-A	.AG	.C-C	8880
G.SE.SE6165		.G-T	.A-AGT	.T-A	.AG	.C-C	8972
G.FI.HH87932		.G-T	.A-AGT	.T-A	.AG	.C-C	8930
H.BE.VI991	.T	.G-T	.A-AGT	.T-A	.AG	.C-C	8946
H.BE.VI997		.G-T	.A-AGT	.T-A	.AG	.C-C	8845
H.CF.90CF056		.G-T	.A-AGT	.T-A	.AG	.C-C	8846
J.SE.SE91733		.AG-T	.G	.T-A	.AG	.C-C	8837
J.SE.SE92809		.AG	.GT	.T-A	.AG	.C-C	9649
O.CM.AN70	AAAG.ACTGCTGACACTGG	.AG.T	.GGAG-A-A	.GG	.TTC	.C	9645
O.CM.MP5180	.ACTGCTGACACTGG	.AGCT	.GGAG-ATAA	.GG	.TTC	.G	9711
O.SN.MP1299	AGAG.ACTGCTGACACTGG	.AGTGG	.A.G	.TT-G	.G	.C	9681
O.SN.MP1300	AGAG.GCTGCTGACACTGG	.AGTGG	.A.G	.TT-G	.G	.C	9127
N.CM.YBF30	TGACAAGGACTTATACTTG	.G	.T	.TTG	.TT	.AG	9045
N.CM.YBF106	.GGGACTTTCCGCCA		.T	.TTG	.TT	.AG	9452
CONSENSUS_01	?	.AG-g	.A-AAG	.TGT-T-G	.AG	.C	9530
01.AE.TH.93TH253		.AG-g	.gTgtggcgggg	.gcgagggttgggagtggc	.T	.C	9452
01.AE.TH.CM240		.AG-T	.G	.AGT	.T	.A	9530
01.AE.CF.92CF402		.AG-T	.G	.AGT	.A	.C	9095
01.AE.TH.92TH022		.AG-T	.G	.AGT	.A	.C	9653
01.AE.TH.95THLH047		.AG-T	.CCA-G-AG	.GT	.T-A	.C	9427
01.AE.CF.11697		.AG	.T	.T	.T-A	.C	9495
01.AE.TH.93TH057		.AG	.T	.G	.T-A	.A	9480
01.AE.TH.93TH065		.AG	.T	.G	.T-A	.A	9465
01.AE.CF.4071		.AG-A	.T	.G	.T-A	.A	9457
CONSENSUS_02	?????????????	.?G-gagg	.tgtgggtt	.A-AGT	.C	.C	8700
02.AG.SE.7812	A	.G-T	.TT	.A-AGT	.C	.C	8953
02.AG.GH.G829		.G-T	.TT	.A-AGT	.C	.C	8723
02.AG.DJ.DJ263		.G-T	.TT	.A-AGT	.C	.C	8925
02.AG.DJ.DJ264		.G-T	.TT	.A-AGT	.C	.C	8884
02.AG.NG.IBNG		.G-T	.TT	.A-AGT	.C	.C	9055
02.AG.SN.MP1211		.AG	.T	.T	.T-A	.C	8754
02.AG.SN.MP1213		.G	.T	.TT	.T-A	.C	8770
02.AG.CM.M807	CAA	.GG	.T	.TT	.T-A	.C	8730
03.AB.RU.98001	A	.AG	.T	.TT	.T-A	.C	8658
03.AB.RU.KAL153-2		.AG	.T	.TT	.T-A	.G	9247
04.CPZ.GR.97PMY	C	.AG	.T	.C	.T-A	.C	9598
04.CPZ.CY.94CY0323	C	.AG	.T	.C	.T-A	.C	8943
05.DF.BE.VII1310		.G	.S-AA	.GC-GACT	.TCA	.G	8974
06.CPZ.ML.95ML84		.G-T	.T-G	.GC-GACT	.TCA	.G	8915
06.CPZ.AU.BFP90		.G-T	.T-G	.G	.T-A	.C	8697
06.CPZ.ML.95ML127		.G-T	.T-A	.G	.T-A	.C	9530
06.CPZ.SN.97SE1078		.AG	.Y	.G	.T-A	.C	9618
10.CD.BEL061		.AATGGG CGTT	.CCA-A-T	.A-A	.T-A	.C	8981
10.CD.BEL101		.ACTGGGCGTT	.CCA-G-T	.T	.T-A	.C	8996
10.CD.BEL101		.ACTGGGCGTT	.CCA-A-T	.T	.T-A	.C	8927
11.CPZ.CM.ME818		.AG	.T	.G	.T-A	.C	9519
11.CPZ.FR.MP1298		.AG	.T	.G	.T-A	.C	9546
11.CPZ.FR.MP1307		.AG	.T	.T	.T-A	.C	9578
11.CPZ.GR.GR17		.AG	.T	.T	.T-A	.C	8829
CPZ.GA.SIVCPZGAB		.G	.GA-AG	.GTC-GAG	.GTC	.T	9621
CPZ.US.SIVCPZUS		.AA	.GG-GTC	.AG-GC-ACA	.C	.T	9594
CPZ.CM.SIVCPZCAM3		.A	.GG-TTC	.AGG	.C	.T	8983
CPZ.CM.SIVCPZCAM5		.A	.GGAG-TC	.AGG	.C	.T	9261
CPZ.CD.SIVCPZANT	AT	.G-CA-TA	.T-G-G-C	.-G	.AGTGGCTTTC	.-G-CT-G	8985

3' LTR U3 end // 3' LTR repeat start | ->
TGCCTGTACTGGG.TCTCTC.TGGTTAGACAGATCTGAGCCTGGAGCTCTCTGGC...TAACTAG.GGAACCCACTGCTTAA.GCCTCAATAAA..GCTTCCCTTGGAG.TGCTTCAAGTAGTGTGTG 9647
Secondary structure region start | ->
3' LTR poly-A signal R repeat end // 3' LTR U5 start
B.FR.HXB2R
G.BE.DRCBL
G.NG.92NG083
G.SE.SE6165
G.FI.HH87932
H.BE.VI991
H.BE.VI997
H.CF.90CF056
J.SE.SE91733
J.SE.SE92809
O.CM.ANT70
O.CM.MVF5180
O.SN.MPI299
O.SN.MPI300
N.CM.YBF30
CONSENSUS 01
01.AE.TH.93TH253
01.AE.TH.CM240
01.AE.CF.90CF402
01.AE.CF.11697
01.AE.TH.93TH057
01.AE.TH.93TH065
01.AE.CF.4071
CONSENSUS 02
02.AG.SE.7812
02.AG.GH.8829
02.AG.DJ.DJ263
02.AG.DJ.DJ264
02.AG.NG.IBNG
02.AG.SN.MPI211
02.AG.SN.MPI213
02.AG.CM.MP807
03.AB.RU.98001
03.AB.RU.KA1153-2
04.CFX.GR.97FVWY
04.CFX.CY.94CY0323
05.DF.BE.VII1310
05.DF.BE.VI961
06.CFX.AU.BFP90
06.CFX.ML.95ML127
06.CFX.SN.97SE1078
10.CD.BFL061
10.CD.BFL110
10.CD.BFL071
11.CFX.CM.MP818
11.CFX.FR.MP1298
11.CFX.FR.MP1307
11.CFX.FR.MP1307
11.CFX.FR.MP1307
11.CFX.FR.GRI17
CPZ.GA.SIVCPZGAB
CPZ.US.SIVCPZUS
CPZ.CM.SIVCPZCAM3
CPZ.CD.SIVCPZANT

```

B.FR.HXB2R
CONSENSUS A
A.UG.U455
A.KE.Q2317
CONSENSUS B
B.AU.MBC925
B.AU.MBC200
B.GB.MANC
B.TW.LM49
B.US.DH123
B.US.SF2
B.US.YU2
B.US.MN
B.US.WEAU160
CONSENSUS C
C.IN.93IN101
G.BE.DRCBL
O.CM.MVP5180
O.SN.MP1299
O.SN.MP1300
CONSENSUS 01
01.AE.TH.93TH253
01.AE.TH.CM240
01.AE.CF.90CF402
01.AE.CF.11697
01.AE.TH.93TH057
01.AE.TH.93TH065
01.AE.CF.4071
CONSENSUS 02
02.AG.SE.7812
02.AG.NG.IBNG
06.CPX.ML.95ML84
06.CPX.AU.BFP90
06.CPX.ML.95ML127
06.CPX.SN.97SE1078
11.CPX.CM.MP818
11.CPX.FR.MP1298
11.CPX.FR.MP1307
CPZ.GA.SIVCPZGAB
CPZ.US.SIVCPZUS
CPZ.CM.SIVCPZCAM3
    
```

Extensive secondary structure in this region

See Rizvi J Virology 67: 2681-8 (1993)

```

3' ITR U5 end <-|
CCGCTCTGTGGTGTGCTAGACTGGTAACCTAGAGATCCCTCAGACCCCTTTAGTCAGTCGTGGAAAATCCTCTAGCA
????????????????????????????????????????????????????????????????????????????
-----A-T-----AC-C--A-T-----AA-----
????????????????????????????????????????????????????????????????????????????
-----AC-C--A-G--AA-----
-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
-----A-----G-----T-----G-----
????????????????????????????????????????????????????????????????????????????
-----A-----AC--CTAGATAGTGTA-----
-----ATC-G-TCAACCC--GTGTCTAG-GA-C
-----ATC-G-TCGACCC--GTAICTAG-GA-C
-----ATC-G-TCAACCC--GTAICTAG-GA-C
cccgTctgttA.ggactctgg??????????????????????????????????????????????????????
--A-----GT-A-G-----T-AC--CTAGACTGAGTA-----
-----GT-A-G-----AC--CTAGACTGAGTA-----
-----GA-A-G-----A-T-----A-----
-----GT-A-G-----A-----
-----GA-A-G-----A-----
-----GA-A-G-----G-CGCA--C
????????????????????????????????????????????????????????????????????????????
-----A-TGTTG-----
-----T-----AC--CTAGACTGAGTA-----
-----T-----AC.GCTAGGC-GTGTA-----
-----G--TT-----AC--CTAGG--GTGTA-----
-----CG-CTG--G-GG--T--A-C--A-----A-----C-G-G-----T-T-----
-----A-----C-----C-----
-----T-----A-----C-----TTAAA-TA-AGTCAA-----
-----TA-A--CA--A-----A-TC-A-AA--AG-TG--A-----
-----TA-CA--CA--A-----A..C-A-TA-A--AG-TG--A-----
    
```

9719
8797
9178
9169
9551
9739
9767
9655
9694
9704
9736
9706
9738
9719
9680
9680
9705
9793
9859
9829
9593
9720
9203
9843
9628
9645
9612
9597
8777
9063
9201
8697
9775
9719
9808
9710
9736
9768
9811
9781
9170