

II

HIV-1/SIVcpz Complete Genomes

Introduction	279
Table of features annotated in the HIV-1/SIVcpz Complete Genome Alignments	281
Table of HIV-1/SIVcpz Complete Genome Alignments	283
Notes on full-length HIV-1/SIVcpz Complete Genome Alignments	286
Nucleotide Alignments of HIV-1/SIVcpz Complete Genomes	302

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 March 2002 there were 344 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 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 468 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 286. 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 the programs MASE 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	302			
TCF-1 α	305			
NF- κ B-I, II	308			
Spl, II, III	308			
TATA Box	308			
5' LTR U3 end\ /R repeat begins	310			
mRNA start site	310			
Poly-A signal	310			
5' LTR repeat end\ /U5 start	310			
5' LTR U5 end	312			
Lys tRNA primer binding site	312			
Gag binding loops	314			
	316	Gag-Pol start		
	322	p17\ /p24		
	334	p24\ /p2		
	334	p2\ /p7		
	334	gag-pol TF		
	336	p7\ /p1		
ribosome slip site	336			
	338	p1\ /p6		
	342			gag-pol TF\ / pol protease
	342	Gag end		
	346			protease\ / p51 RT
	366			p51 RT\ /p15 RNase H
	372			p15 RNaseH\ /p31 integrase
	384	Vif start		Pol end
	392	Vpr start		
	394	Vif end		
	398		Tat ex1 start	Vpr end
	400			Rev ex1 start
	400		Tat ex1 end	Rev ex1 end
	402		Vpu start	
	404			Env signal peptide start
	406		Vpu end	gp120 Env start
	420			V3 loop start
	422			V3 loop end
Rev res. element start	430			
Stem I, IIA, IIIB	432			gp120\ /gp41
Stem IIB', IIC, IIC', IIA'	434			
Stem III-IV, III-IV', V, V', I'	434			
Stem I', Rev res. element end	436			
	442	Tat ex2 start	Rev ex2 start	
	444	Tat ex2 end		
	446		Rev ex2 end	
	448			Env end
	450	Nef start		
3' LTR U3 start	454			
TCF-1 α binding	460	Nef end		
NF- κ -BII	460			
NF- κ -BI, SPI-III, II, I	462			
TATA box	462			
poly-A signal	464			
3' LTR U3 end\ /repeat start	464			
3' LTR repeat end' LTR U5 start	464			
3' LTR U5 end	466			

References

- [Carr96] J.K. Carr, M.O. Salminen, C.Koch, D.Gotte, A.W. Artenstein, P.A. Hegerich, 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**:5935–43, 1996.
- [Charpentier97] B.Charpentier, F.Schultz, M.Rosbash. A dynamic *in vivo* view of the HIV-1 Rev-RRE interaction. *J Mol Biol* **266**:950–962, 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**:4053–62, 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**:7013–29, 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**:6087–91, 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**:6143–52, 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**:14902–8, 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**:8657–65, 1997.
- [Reil93] H.Reil, H.Kollmus, U.H. Weidle, H. Hauser. A heptanucleotide sequence mediates ribosomal frameshifting in mammalian cells. *J Virol* **67**:5579–84, 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**:77–80, 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**:5608–13, 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>.

A.BY.97BL006	AF193275	Belarus	Liitsola, K	<i>ARHR</i> 16 (11):1047–53 (2000)
A.SE.SE7535	AF069671	Sweden	Laukkanen, T	<i>AIDS</i> 13 (14):1819–1826 (1999)
A.SE.SE8538	AF069669	Sweden	Laukkanen, T	<i>AIDS</i> 13 (14):1819–1826 (1999)
A.SE.SE8891	AF069673	Sweden	Laukkanen, T	<i>AIDS</i> 13 (14):1819–1826 (1999)
A.SE.UGSE8131	AF107771	Sweden	Laukkanen, T	<i>AIDS</i> 13 (14):1819–1826 (1999)
A1.KE.Q23-17	AF004885	Kenya	Poss, M	<i>J Virol</i> 72 (10):8240–51 (1998)
A1.SE.SE7253	AF069670	Sweden	Laukkanen, T	<i>AIDS</i> 13 (14):1819–1826 (1999)
A1.UG.92UG037	U51190	Uganda	Gao, F	<i>J Virol</i> 70 (3):1651–1657 (1996)
A1.UG.U455	M62320	Uganda	Oram, JD	<i>ARHR</i> 6 (9):1073–1078 (1990)
A2.CD.97CDKTB48	AF286238	Congo	Gao, F	<i>ARHR</i> 17 (8):675–688 (2001)
A2.CY.94CY017-41	AF286237	Cyprus	Gao, F	<i>ARHR</i> 17 (8):675–688 (2001)
B.AU.MBC200	AF042100	Australia	Oelrichs, RB	<i>ARHR</i> 14 (9):811–4 (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, AAG	PhD dissertation, University of Cambridge (1989)
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.JRFL	U63632	U.S.A.	O'Brien, WA	<i>Nature</i> 348 :69–73 (1990)
B.US.MNCG	M17449	U.S.A.	Gurgo, C	<i>Virology</i> 164 (2):531–536 (1988)
B.US.P896	U39362	U.S.A.	Kim, FM	<i>J Virol</i> 69 (3):1755–1761 (1995)
B.US.RF	M17451	U.S.A.	Starcich, BR	<i>Cell</i> 45 (5):637–648 (1986)
B.US.SF2	K02007	U.S.A.	van Beveren, CP	in Weiss, RL, et al., (Eds); <i>RNA Tumor Viruses</i> , 2nd Ed., Vol 2: 1124–1141; Cold Spring Harbor (1985)
B.US.WR27	AF286365	U.S.A.	Salminen, MO	<i>Virology</i> 278 (1):103–110 (2000)
B.US.YU2	M93258	U.S.A.	Li, Y	<i>J Virol</i> 65 :3973–3985 (1991)
C.BR.92BR025	U52953	Brazil	Gao, F	<i>J Virol</i> 70 (3):1651–1667 (1996)
C.BR.98BR004	AF286228	Brazil	Rodenburg, CM	<i>ARHR</i> 17 (2):161–168 (2001)
C.BW.96BW0502	AF110967	Botswana	Novitsky, VA	<i>J Virol</i> 73 (5):4427–32 (1999)
C.BW.96BW06J4	AF290028	Botswana	Ndung'u, T	<i>Virology</i> 278 (2):390–399 (2000)
C.BW.96BW1104	AF110969	Botswana	Novitsky, VA	<i>J Virol</i> 73 (5):4427–32 (1999)
C.BW.96BW1210	AF110972	Botswana	Novitsky, VA	<i>J Virol</i> 73 (5):4427–32 (1999)
C.BW.96BW15C02	AF110974	Botswana	Novitsky, VA	<i>J Virol</i> 73 (5):4427–32 (1999)
C.BW.96BW17B03	AF110980	Botswana	Novitsky, VA	<i>J Virol</i> 73 (5):4427–32 (1999)
C.ET.ETH2220	U46016	Ethiopia	Salminen, MO	<i>ARHR</i> 12 (14):1329–1339 (1996)
C.IL.98IS002	AF286233	Israel	Rodenburg, CM	<i>ARHR</i> 17 (2):161–168 (2001)
C.IN.93IN101	AB023804	India	Mochizuki, N	<i>ARHR</i> 15 (14):1321–1324 (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.94IN476	AF286223	India	Rodenburg, CM	<i>ARHR</i> 17 (2):161–168 (2001)

C.IN.98IN012	AF286231	India	Rodenburg, CM	<i>ARHR</i> 17 (2):161–168 (2001)
C.IN.98IN022	AF286232	India	Rodenburg, CM	<i>ARHR</i> 17 (2):161–168 (2001)
C.TZ.98TZ013	AF286234	Tanzania	Rodenburg, CM	<i>ARHR</i> 17 (2):161–168 (2001)
C.TZ.98TZ017	AF286235	Tanzania	Rodenburg, CM	<i>ARHR</i> 17 (2):161–168 (2001)
C.ZA.97ZA012	AF286227	South Africa	Rodenburg, CM	<i>ARHR</i> 17 (2):161–168 (2001)
C.ZM.96ZM651	AF286224	Zambia	Rodenburg, CM	<i>ARHR</i> 17 (2):161–168 (2001)
C.ZM.96ZM751	AF286225	Zambia	Rodenburg, CM	<i>ARHR</i> 17 (2):161–168 (2001)
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):5680–98 (1998)
F1.BE.VI850	AF077336	Belgium	Laukkanen, T	<i>Virology</i> 269 (1):95–104 (2000)
F1.BR.93BR020-1	AF005494	Brazil	Gao, F	<i>J Virol</i> 72 (7):5680–98 (1998)
F1.FI.FIN9363	AF075703	Finland	Laukkanen, T	<i>Virology</i> 269 (1):95–104 (2000)
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)
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):5680–98 (1998)
G.SE.SE6165	AF061642	Sweden	Laukkanen, T	<i>Virology</i> 247 (1):22–31 (1998)
H.BE.VI991	AF190127	Belgium	Janssens, W	<i>AIDS</i> 14 (11):1533–43 (2000)
H.BE.VI997	AF190128	Belgium	Janssens, W	<i>AIDS</i> 14 (11):1533–43 (2000)
H.CF.90CF056	AF005496	C.A.R.	Murphy, E	<i>ARHR</i> 9 (10):997–1006 (1993)
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.EQTB11C	AJ249235	Congo	Triques, K	<i>ARHR</i> 16 (2):139–151 (2000)
K.CM.MP535	AJ249239	Cameroon	Triques, K	<i>ARHR</i> 16 (2):139–151 (2000)
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.95TH022	AB032740	Thailand	Auwanit, W	<i>ARHR</i> 17 (9):867–871 (2001)
01_AE.TH.95TH047	AB032741	Thailand	Auwanit, W	<i>ARHR</i> 17 (9):867–871 (2001)
01_AE.TH.CM240	U54771	Thailand	Laukkanen, T	<i>J Virol</i> 70 (9):5935–5943 (1996)
02_AG.CM.97CM-MP807	AJ286133	Cameroon	Montavon, C	<i>JAIDS</i> 23 (5):363–74 (2000)
02_AG.FR.DJ263	AF063223	France	Laukkanen, T	<i>Virology</i> 247 (1):22–31 (1998)
02_AG.FR.DJ264	AF063224	France	Laukkanen, T	<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.98SE-MP1211	AJ251056	Senegal	Toure-Kane, C	<i>ARHR</i> 16 (6):603–9 (2000)
02_AG.SN.MP1213	AJ251057	Senegal	Toure-Kane, C	<i>ARHR</i> 16 (6):603–9 (2000)
03_AB.RU.KAL153-2	AF193276	Russia	Liitsola, K	<i>AIDS</i> 12 (14):1907–1919 (1998)
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)
07_BC.CN.97CN001	AF286226	China	Rodenburg, CM	<i>ARHR</i> 17 (2):161–168 (2001)
07_BC.CN.98CN009	AF286230	China	Rodenburg, CM	<i>ARHR</i> 17 (2):161–168 (2001)
07_BC.CN.C54A	AX149647	China	Shao, Y	Patent: WO 0136614-A 1
08_BC.CN.97 CNGX6F	AY008715	China	Piyasirisilp, S	<i>J Virol</i> 74 (23):11286–11295 (2000)
08_BC.CN.97 CNGX7F	AY008716	China	Piyasirisilp, S	<i>J Virol</i> 74 (23):11286–11295 (2000)
08_BC.CN.97 CNGX9F	AY008717	China	Piyasirisilp, S	<i>J Virol</i> 74 (23):11286–11295 (2000)
08_BC.CN.98CN006	AF286229	China	Rodenburg, CM	<i>ARHR</i> 17 (2):161–168 (2001)
10_CD.TZ.96TZBF061	AF289548	Tanzania	Kouliniska, IN	<i>ARHR</i> 17 (5):423–431 (2001)
10_CD.TZ.96TZBF071	AF289549	Tanzania	Kouliniska, IN	<i>ARHR</i> 17 (5):423–431 (2001)
10_CD.TZ.96TZBF110	AF289550	Tanzania	Kouliniska, IN	<i>ARHR</i> 17 (5):423–431 (2001)
11_cpx.CM.MP818	AJ291718	Cameroon	Montavon, C	<i>ARHR</i> 18 (3):231–236 (2002)
11_cpx.FR.MP1298	AJ291719	France	Montavon, C	<i>ARHR</i> 18 (3):231–236 (2002)
11_cpx.FR.MP1307	AJ291720	France	Montavon, C	<i>ARHR</i> 18 (3):231–236 (2002)
11_cpx.GR.GR17	AF179368	Greece	Paraskevis, D	<i>ARHR</i> 16 (9):845–855 (2000)
12_BF.AR.ARMA159	AF385936	Argentina	Carr, JK	<i>AIDS</i> 15 (15):F41–7
12_BF.UY.URTR23	AF385934	Uruguay	Carr, JK	<i>AIDS</i> 15 (15):F41–7
12_BF.UY.URTR35	AF385935	Uruguay	Carr, JK	<i>AIDS</i> 15 (15):F41–7
N.CM.YBF106	AJ271370	Cameroon	Ayoubu, A	<i>AIDS</i> 14 (16):2623–2625 (2000)
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 (3):1581–1585 (1994)
O.SN.SEMP1299	AJ302646	Senegal	Toure-Kane, C	<i>ARHR</i> 17 (12):1211–1216 (2001)
O.SN.SEMP1300	AJ302647	Senegal	Toure-Kane, C	<i>ARHR</i> 17 (12):1211–1216 (2001)
CPZ.CD.CPZANT	U42720	Congo	Vanden Haesevelde, M	<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	Souquiere, S	Unpublished
CPZ.GA.CPZGAB	X52154	Gabon	Huet, T	<i>Nature</i> 345 (6273):356–359 (1990)
CPZ.US.CPZUS	AF103818	U.S.A.	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.BY.97BL006	AF193275	Belarus	Liitsola, K	<i>ARHR</i> 16(11):1047–53 (2000) This sequence came from a PBMC sample taken in 1997 from a male IVDU from Svetlogorsk, Byelorussia. The A subtype infection is associated with the A regions of the CRF03_AB, and so related to the A subtype parental strain.
A.SE.SE7535	AF069671	Sweden	Laukkanen, T	<i>AIDS</i> 13(14):1819–1826 (1999) 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.
A.SE.SE8538	AF069669	Sweden	Laukkanen, T	<i>AIDS</i> 13(14):1819–1826 (1999) 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.
A.SE.SE8891	AF069673	Sweden	Laukkanen, T	<i>AIDS</i> 13(14):1819–1826 (1999) 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.
A.SE.UGSE8131	AF107771	Sweden	Laukkanen, T	<i>AIDS</i> 13(14):1819–1826 (1999) 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.
A1.KE.Q23-17	AF004885	Kenya	Poss, M	<i>J Virol</i> 72(10):8240–51 (1998) 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., et al. (<i>ARHR</i> 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.
A1.SE.SE7253	AF069670	Sweden	Laukkanen, T	<i>AIDS</i> 13(14):1819–1826 (1999) 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.
A1.UG.92UG037	U51190	Uganda	Gao, F	<i>J Virol</i> 70(3):1651–1657 (1996) This sequence is from a complete genome PCR amplified from proviral DNA. The patient was a 31 year old asymptomatic female from Entebbe, Uganda. 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 Network, <i>ARHR</i> 10 :1327-1344 (1994)) and comes from an asymptomatic 31-year old female from Entebbe, 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., <i>J Virol</i> 71:7478 (1997) and Rucker, J., et al., <i>J Virol</i> 71: 8999-9007 (1997). See also: Gao F. et al., <i>J Virol</i> 70: 7013-7029 (1996). This sequence was kindly made available prior to publication, and is now published (Gao F., et al., <i>J Virol</i> 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.

- A1.UG.U455 M62320 Uganda Oram, JD *ARHR* 6(9):1073–1078 (1990)
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.
- A2.CD.97CDKTB48 AF286238 Congo Gao, F *ARHR* 17(8):675–688 (2001)
This sample was obtained during a molecular epidemiological study survey in Kinshasa, DRC, in April 1997. This sequence is from a primary culture of virus obtained from a patient suffering from tuberculosis.
- A2.CY.94CY017-41 AF286237 Cyprus Gao, F *ARHR* 17(8):675–688 (2001)
This sequence was obtained from a 35 year old female AIDS patient from Nicosia, Cyprus (HO17). Her husband and child and the husband's sexual partner were also infected with this strain, which may have originated in the UK. The sequence is from a primary isolate.
- B.AU.MBC200 AF042100 Australia Oelrichs, RB *ARHR* 14(9):811–4 (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.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 Sauer mann, 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, HAN/2 and HAN/3 were purified. HAN/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.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 a typical 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. 5 sibling sequences for OYI(397), each 59 bases long, are available (see the 1989 HIV Database Compendium page I-A-181).

B.GB.CAM1 D10112 U.K. McIntosh, AAG PhD dissertation, University of Cambridge (1989)

This sequence is from the British isolate CAM1. It has a defective vpu gene. GenBank accession numbers D10112, D00917 (secondary). The sample date is September 1983 (Abraham Karpas, Nov. 2001 e-mail to Una Smith). Dissertation author McIntosh is Alison Agnes Gibson McIntosh, evidently supervised by Abraham Karpas.

B.GB.MANC U23487 U.K. Zhu, T *Nature* **374**(6522):503–504 (1995)

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. Kindly provided in electronic form by Dr. David Ho, Aaron Diamond AIDS Research Center, New York City.

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.JRFL U63632 U.S.A. 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 (JRBR). 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.A. 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. Another complete genome of the MN isolate is available with accession number AF075719 and it too has defective genes, although not pol nor vpu. A set of V3 sequences from this isolate are available (Accession #s L48364-L48379, Lukashov, V. and Goudsmit, J., AIDS 9:1307-1311 (1995). This sequence was kindly provided in computer readable format by M. Reitz, N.C.I. The isolate MN is available from the NIH AIDS Reagent program, and is X4.

- B.US.P896 U39362 U.S.A. Kim, FM *J Virol* **69**(3):1755–1761 (1995)
89.6 is cloned from a highly cytopathic primary macrophage-tropic and syncytium-inducing isolate. It comes from an AIDS patient living in Philadelphia for 15 years, although the patient was originally from Jamaica. At the time of viral isolation, this 47 year old man had received no antiretroviral therapy and had advanced immunodeficiency. 89.6 was cloned in phage and was replication competent after reconstruction. Other receptors defined for this isolate are:CCR3,CCR2b, CCR8, V28. Second receptor usage for this isolate was defined by Farzan et al, *J Exp Med* 186:405 (1997) and Rucker et al., *J Virol* 71(12):8999-9007 (1997). Observed phenotype of culture may not correspond to the sequence. Also see Kim, F., et al., *J Virol* 69:1755-1761 (1995). The 89.6 env gene was substituted into the SHIV-4 chimeric simian-human virus by (Reimann et al *J Virol* 70:3198 (1996)) see accession numbers U89134 and AF038398. The isolate89.6 is available from the NIH AIDS Reagent program and is listed as SI R5X4.
- B.US.RF M17451 U.S.A. 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 strains. 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., *ARRH* 8:1950 (1992).
- B.US.SF2 K02007 U.S.A. van Beveren, CP in) Weiss,RL, et al., (Eds); *RNA Tumor Viruses*, 2nd Ed., Vol 2: 1124–1141; Cold Spring Harbor (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. It is syncytia-inducing (SI) and uses both coreceptors (R5X4).
- B.US.WR27 AF286365 U.S.A. Salminen, MO *Virology* **278**(1):103–110 (2000)
Another complete genome sequence from this isolate is available with accession number U26546. This one is of higher quality, with fewer N R and Y ambiguity characters in it. This sequence is from a molecular clone which has been proven to be infectious.
- B.US.YU2 M93258 U.S.A. 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. 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. Additional env, nef and ltr region sequences are available from this isolate: U09126, U09132, U51282, and U15121. 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 indels 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.BR.98BR004 AF286228 Brazil Rodenburg, CM *ARHR* 17(2):161–168 (2001)
 This sequence is part of an international survey of subtype C, based mostly on isolates obtained by the UNAIDS and NIAID networks for isolation and characterization of HIV. DNA was isolated from PBMC after short-term coculture. The samples included CRF07 and CRF08 BC recombinants.
- C.BW.96BW0502 AF110967 Botswana Novitsky, VA *J Virol* 73(5):4427–32 (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.96BW06J4 AF290028 Botswana Ndung'u, T *Virology* 278(2):390–399 (2000)
- C.BW.96BW1104 AF110969 Botswana Novitsky, VA *J Virol* 73(5):4427–32 (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):4427–32 (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):4427–32 (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.96BW17B03 AF110980 Botswana Novitsky, VA *J Virol* 73(5):4427–32 (1999)
 This sequence comes from a study of 23 full length C subtype sequences from Gaborone, Botswana, from eight individuals, sampled in 1996. The pol protease region of the 96BW17 genomes clusters with two isolates from Zimbabwe (Z1226 AF083262, and Z84 AF083267) external to the rest of the subtype C clade. Three full length sequences from this patient are available: AF110979-AF110981.
- 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-kB 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.IL.98IS002 AF286233 Israel Rodenburg, CM *ARHR* **17**(2):161–168 (2001)
This sequence is part of an international survey of subtype C, based mostly on isolates obtained by the UNAIDS and NIAID networks for isolation and characterization of HIV. DNA was isolated from PBMC after short-term coculture. The samples included CRF07 and CRF08 BC recombinants.
- 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.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.94IN476 AF286223 India Rodenburg, CM *ARHR* **17**(2):161–168 (2001)
This sequence is part of an international survey of subtype C, based mostly on isolates obtained by the UNAIDS and NIAID networks for isolation and characterization of HIV. DNA was isolated from PBMC after short-term coculture. The samples included CRF07 and CRF08 BC recombinants.
- C.IN.98IN012 AF286231 India Rodenburg, CM *ARHR* **17**(2):161–168 (2001)
This sequence is part of an international survey of subtype C, based mostly on isolates obtained by the UNAIDS and NIAID networks for isolation and characterization of HIV. DNA was isolated from PBMC after short-term coculture. The samples included CRF07 and CRF08 BC recombinants.
- C.IN.98IN022 AF286232 India Rodenburg, CM *ARHR* **17**(2):161–168 (2001)
This sequence is part of an international survey of subtype C, based mostly on isolates obtained by the UNAIDS and NIAID networks for isolation and characterization of HIV. DNA was isolated from PBMC after short-term coculture. The samples included CRF07 and CRF08 BC recombinants.
- C.TZ.98TZ013 AF286234 Tanzania Rodenburg, CM *ARHR* **17**(2):161–168 (2001)
This sequence is part of an international survey of subtype C, based mostly on isolates obtained by the UNAIDS and NIAID networks for isolation and characterization of HIV. DNA was isolated from PBMC after short-term coculture. The samples included CRF07 and CRF08 BC recombinants.
- C.TZ.98TZ017 AF286235 Tanzania Rodenburg, CM *ARHR* **17**(2):161–168 (2001)
This sequence is part of an international survey of subtype C, based mostly on isolates obtained by the UNAIDS and NIAID networks for isolation and characterization of HIV. DNA was isolated from PBMC after short-term coculture. The samples included CRF07 and CRF08 BC recombinants.
- C.ZA.97ZA012 AF286227 South africa Rodenburg, CM *ARHR* **17**(2):161–168 (2001)
This sequence is part of an international survey of subtype C, based mostly on isolates obtained by the UNAIDS and NIAID networks for isolation and characterization of HIV. DNA was isolated from PBMC after short-term coculture. The samples included CRF07 and CRF08 BC recombinants.

- C.ZM.96ZM651 AF286224 Zambia Rodenburg, CM *ARHR* **17**(2):161–168 (2001)
This sequence is part of an international survey of subtype C, based mostly on isolates obtained by the UNAIDS and NIAID networks for isolation and characterization of HIV. DNA was isolated from PBMC after short-term coculture. The samples included CRF07 and CRF08 BC recombinants.
- C.ZM.96ZM751 AF286225 Zambia Rodenburg, CM *ARHR* **17**(2):161–168 (2001)
This sequence is part of an international survey of subtype C, based mostly on isolates obtained by the UNAIDS and NIAID networks for isolation and characterization of HIV. DNA was isolated from PBMC after short-term coculture. The samples included CRF07 and CRF08 BC recombinants.
- 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 the 1995 Compendium (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)
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. Kindly provided prior to publication by J.-C. Chermann, Pasteur Institute, Marseille.
- 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):5680–98 (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 Laukkanen, T *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):5680–98 (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 *Virology* **269**(1):95–104 (2000)
 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 woman 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) (pers. 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):5680–98 (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 Laukkanen, T *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):1533–43 (2000)
This sequence was isolated from a Belgian man who lived with his Belgian wife in the DRC for 20 years. The wife died of AIDS in 1990. The route of infection was unknown for both.
- H.BE.VI997 AF190128 Belgium Janssens, W *AIDS* 14(11):1533–43 (2000)
This virus was isolated from a Belgian man with multiple sexual contacts in the DRC.
- H.CF.90CF056 AF005496 C.A.R. Murphy, E *ARHR* 9(10):997–1006 (1993)
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 sequences is a complete genome reference for subtype K.

- K.CM.MP535 AJ249239 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 subtype K.
- 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. The sequence was 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. This sequence is from the French isolate LAI (formerly BRU) which is also referred to as IIIB. Primary reference: Wain-Hobson S, Sonigo P, Danos O, Cole S, Alizon M. Nucleotide sequence of the AIDS virus, LAV. *Cell*. 1985 Jan;40(1):9-17. Also see: Lukashov VV, Goudsmit J. Increasing genotypic and phenotypic selection from the original genomic RNA populations of HIV-1 strains LAI and MN (NM) by peripheral blood mononuclear cell culture, B-cell-line propagation and T-cell-line adaptation. *AIDS*. 1995 Dec;9(12):1307-11. Wain-Hobson S, Vartanian JP, Henry M, Chenciner N, Cheynier R, Delassus S, Martins LP, Sala M, Nugeyre MT, Guetard D, et al. LAV revisited: origins of the early HIV-1 isolates from Institut Pasteur. *Science*. 1991 May 17;252(5008):961-5.
- 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 recombined 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 Robertson, D., part III pages 25-30, of this compendium, 1997, Gao, F. et al., *J Virol* 70:7013-7029 (1996), and Carr, J. et al., *J Virol* 70:5935-5943 (1996). The prototype of this circulating recombinant form is the sequence CM240.
- 01_AE.TH.95TH022 AB032740 Thailand Auwanit, W *ARHR* **17**(9):867–871 (2001)
 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 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.95TH047 AB032741 Thailand Auwanit, W *ARHR* **17**(9):867–871 (2001)
 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 Laukkanen, T *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 through out 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)).
- 02_AG.CM.97CM-MP807 AJ286133 Cameroon Montavon, C *JAIDS* **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 Laukkanen, T *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, CRF02_AG; IbNG is the prototype, sharing a similar structure with DJ264 and DJ263.
- 02_AG.FR.DJ264 AF063224 France Laukkanen, T *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, CRF02_AG; 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.98SE-MP1211 AJ251056 Senegal Toure-Kane, C *ARHR* 16(6):603–9 (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):603–9 (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 *AIDS* 12(14):1907–1919 (1998)
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 non-recombinant 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 IVDU 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 IVDUs 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 was isolated from patient GR11 (accession AF049292) is from the same patient as 97PVCHAF049292. 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.VI1310 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; VI1310 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-naïve. 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. Commercial sex worker
- 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.
- 07_BC.CN.97CN001 AF286226 China Rodenburg, CM *ARHR* 17(2):161-168 (2001)
This sequence is reported (Feng Gao and Yiing Shao, personal communication to HIV-DB) to be from the same blood sample as the sequences with accession numbers AX149647, AX149672, AX149771 and AX149898. The four AX- entries are 100% identical to each other in the env gene region and 98% identical to this sequence. The other CRF07_BC genome available at this time (Sept 2001) is AF286230 from sample CN-009 and is 97% identical to each of these 5 sequences in the env gene.
- 07_BC.CN.98CN009 AF286230 China Rodenburg, CM *ARHR* 17(2):161-168 (2001)
This sequence is part of an international survey of subtype C, based mostly on isolates obtained by the UNAIDS and NIAID networks for isolation and characterization of HIV. DNA was isolated from PBMC after short-term coculture. The samples included CRF07 and CRF08 BC recombinants.
- 07_BC.CN.C54A AX149647 China Shao, Y Patent: WO 0136614-A 1
- 08_BC.CN.97CNGX6F AY008715 China Piyasirisilp, S *J Virol* 74(23):11286-11295 (2000)
This sequence is the first complete genome available for the CRF08_BC circulating recombinant form. The genome is mostly subtype C, with two regions of subtype B.
- 08_BC.CN.97CNGX7F AY008716 China Piyasirisilp, S *J Virol* 74(23):11286-11295 (2000)
- 08_BC.CN.97CNGX9F AY008717 China Piyasirisilp, S *J Virol* 74(23):11286-11295 (2000)
Although this sequence was initially labeled as subtype C, Rodenburg et al in *AIDS Res Hum Retroviruses* 17(2):161-168 (2001) found that isolate 98CN006 (accession AF286229) is B/C recombinant, and analysis at LANL HIV-DB shows that 97CNGX-9F has the same form, and is closely related to 98CN006. This sequence has thus been re-labeled as a B/C intersubtype recombinant, June 12, 2001.
- 08_BC.CN.98CN006 AF286229 China Rodenburg, CM *ARHR* 17(2):161-168 (2001)
This sequence is part of an international survey of subtype C, based mostly on isolates obtained by the UNAIDS and NIAID networks for isolation and characterization of HIV. DNA was isolated from PBMC after short-term coculture. The samples included CRF07 and CRF08 BC recombinants.
- 10_CD.TZ.96TZBF061 AF289548 Tanzania Koulinska, IN *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.96TZBF071 AF289549 Tanzania Koulinska, IN *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.96TZBF110 AF289550 Tanzania Koulinska, IN *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 Montavon, C *ARHR* 18(3):231-6 (2002)

- 11_cpx.FR.MP1298 AJ291719 France Montavon, C *ARHR* **18**(3):231–6 (2002)
- 11_cpx.FR.MP1307 AJ291720 France Montavon, C *ARHR* **18**(3):231–6 (2002)
- 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 is thought to be a mosaic of A/G/E/J.
- 12_BF.AR.ARMA159 AF385936 Argentina Carr, JK *AIDS* **15**(15):F41–7
This sequence is part of a set of South American isolates from various countries. It is one of the reference strains for CRF12_BF.
- 12_BF.UY.URTR23 AF385934 Uruguay Carr, JK *AIDS* **15**(15):F41–7
This sequence is part of a set of South American isolates from various countries. It is one of the reference strains for CRF12_BF.
- 12_BF.UY.URTR35 AF385935 Uruguay Carr, JK *AIDS* **15**(15):F41–7
This sequence is part of a set of South American isolates from various countries. It is one of the reference strains for CRF12_BF.
- N.CM.YBF106 AJ271370 Cameroon Ayouba, A *AIDS* **14**(16):2623–2625 (2000)
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 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). 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.
- 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. A husband and 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**(3):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.SEMP1299 AJ302646 Senegal Toure-Kane, C *ARHR* **17**(12):1211–1216 (2001)
Sequence from spouse2 of a Senegalese man (patient 98SE-42HALD) with 2 wives. This woman, who previously had lived in Cameroon, was apparently the source of infection of both her husband and his first wife (isolate 99SE-MP1300).
- O.SN.SEMP1300 AJ302647 Senegal Toure-Kane, C *ARHR* **17**(12):1211–1216 (2001)
Sequence from spouse1 of a Senegalese man (patient 98SE-42HALD) with 2 wives. The man was apparently infected by spouse 2 (isolate 99SE-MP1299) who previously had lived in Cameroon.
- CPZ.CD.CPZANT 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. In 2002, B. Hahn et al reported (*Science* 295[5554]: 465 2002) sequencing of virus from another Pan troglodytes schweinfurthii chimp. That sequence is available with accession number AF382822, and is more closely related to this schweinfurthii sequence, than to sequences from other chimpanzees or humans. The Hahn paper labels the CPZ-ANT chimpanzee as ch-No.
- CPZ.CM.CAM3 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 AJ271369 Cameroon Souquiere, S Unpublished
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.CPZGAB 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.CPZUS AF103818 U.S.A. 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 still born 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.

B.FR.HXB2	/ 5' LTR U3 start	129
CONSENSUS_01_AE	TGGAAAGGGCTAATTCACCTCCCAAGGACAAAGATATCCTTTGATCTCTGTGGATCTACACACACAAAGGCTACTTCCCTGATTAGCAGAACTACACACCCAGGGCCAGGGATCAGATATCCACTGACCTTTTG	
01_AE.CF.90CF11697	-t-----g-----c-----tgt-----	72
01_AE.CF.90CF402	-A-GA-----G-----CT-A--G--C--A-T-----G-----tgt-----	66
01_AE.CF.90CF4071	-A-GA-----G-----G-----G-----G-----tgt-----	129
01_AE.TH.93TH057	-A-GA-----G-----G-----G-----G-----tgt-----	54
01_AE.TH.93TH065	-A-GA-----G-----G-----G-----G-----tgt-----	72
01_AE.TH.93TH253	-A-GA-----G-----G-----G-----G-----tgt-----	72
01_AE.TH.95TH022	-A-GA-----G-----G-----G-----G-----tgt-----	129
01_AE.TH.95TH047	-A-GA-----G-----G-----G-----G-----tgt-----	81
01_AE.TH.CM240	-A-GA-----G-----G-----G-----G-----tgt-----	0
CONSENSUS_02_AG	-A-GA-----G-----G-----G-----G-----tgt-----	0
02_AG.CM.97CM-MP807	-A-GA-----G-----G-----G-----G-----tgt-----	0
02_AG.FR.DJ263	-A-GA-----G-----G-----G-----G-----tgt-----	0
02_AG.FR.DJ264	-A-GA-----G-----G-----G-----G-----tgt-----	0
02_AG.GH.G829	-A-GA-----G-----G-----G-----G-----tgt-----	0
02_AG.NG.IBNG	-A-GA-----G-----G-----G-----G-----tgt-----	0
02_AG.SE.SF7812	-A-GA-----G-----G-----G-----G-----tgt-----	0
02_AG.SN.98SE-MP1211	-A-GA-----G-----G-----G-----G-----tgt-----	0
02_AG.SN.MP1213	-A-GA-----G-----G-----G-----G-----tgt-----	0
03_AB.RU.KAL153-2	-A-GA-----G-----G-----G-----G-----tgt-----	114
03_AB.RU.RU98001	-A-GA-----G-----G-----G-----G-----tgt-----	0
04_CPX.CY.94CY032-3	-A-GA-----G-----G-----G-----G-----tgt-----	0
04_CPX.FR.97FVGH	-A-GA-----G-----G-----G-----G-----tgt-----	130
04_CPX.FR.97FVNY	-A-GA-----G-----G-----G-----G-----tgt-----	129
05_DF.BE.VI1310	-A-GA-----G-----G-----G-----G-----tgt-----	0
05_DF.BE.VI1961	-A-GA-----G-----G-----G-----G-----tgt-----	0
CONSENSUS_06_CPX	-A-GA-----G-----G-----G-----G-----tgt-----	129
06_CPX.AU.BFP90	-A-GA-----G-----G-----G-----G-----tgt-----	129
06_CPX.ML.95ML127	-A-GA-----G-----G-----G-----G-----tgt-----	129
06_CPX.ML.95ML84	-A-GA-----G-----G-----G-----G-----tgt-----	0
06_CPX.SN.97SE1078	-A-GA-----G-----G-----G-----G-----tgt-----	129
07_BC.CN.97CN001	-A-GA-----G-----G-----G-----G-----tgt-----	0
07_BC.CN.98CN009	-A-GA-----G-----G-----G-----G-----tgt-----	0
07_BC.CN.C54A	-A-GA-----G-----G-----G-----G-----tgt-----	0
CONSENSUS_08_BC	-A-GA-----G-----G-----G-----G-----tgt-----	0
08_BC.CN.97CNGX6F	-A-GA-----G-----G-----G-----G-----tgt-----	0
08_BC.CN.97CNGX7F	-A-GA-----G-----G-----G-----G-----tgt-----	0
08_BC.CN.97CNGX9F	-A-GA-----G-----G-----G-----G-----tgt-----	0
08_BC.CN.98CN006	-A-GA-----G-----G-----G-----G-----tgt-----	0
10_CD.TZ.96TZBF061	-A-GA-----G-----G-----G-----G-----tgt-----	0
10_CD.TZ.96TZBF071	-A-GA-----G-----G-----G-----G-----tgt-----	0
10_CD.TZ.96TZBF110	-A-GA-----G-----G-----G-----G-----tgt-----	0
CONSENSUS_11_CPX	-A-GA-----G-----G-----G-----G-----tgt-----	129
11_CPX.CM.MP818	-A-GA-----G-----G-----G-----G-----tgt-----	129
11_CPX.FR.MP1298	-A-GA-----G-----G-----G-----G-----tgt-----	129
11_CPX.FR.MP1307	-A-GA-----G-----G-----G-----G-----tgt-----	129
11_CPX.FR.MP1307	-A-GA-----G-----G-----G-----G-----tgt-----	0
11_CPX.FR.GR17	-A-GA-----G-----G-----G-----G-----tgt-----	129
12_BF.AR.ARMA159	-A-GA-----G-----G-----G-----G-----tgt-----	129
12_BF.UY.URTR23	-A-GA-----G-----G-----G-----G-----tgt-----	129
12_BF.UY.URTR35	-A-GA-----G-----G-----G-----G-----tgt-----	129
N.CM.YBF106	-A-GA-----G-----G-----G-----G-----tgt-----	0
N.CM.YBF30	-A-GA-----G-----G-----G-----G-----tgt-----	0
CONSENSUS_0	-A-GA-----G-----G-----G-----G-----tgt-----	125
O.CM.AN170	-A-GA-----G-----G-----G-----G-----tgt-----	130
O.CM.MVP5180	-A-GA-----G-----G-----G-----G-----tgt-----	130
O.SN.SEMP1299	-A-GA-----G-----G-----G-----G-----tgt-----	130
O.SN.SEMP1300	-A-GA-----G-----G-----G-----G-----tgt-----	130
CONSENSUS_CFPZ	-A-GA-----G-----G-----G-----G-----tgt-----	0
CFZ.CD.CFZANT	-A-GA-----G-----G-----G-----G-----tgt-----	0
CFZ.CM.CAM3	-A-GA-----G-----G-----G-----G-----tgt-----	0
CFZ.CM.CAM5	-A-GA-----G-----G-----G-----G-----tgt-----	0
CFZ.GA.CFZGAB	-A-GA-----G-----G-----G-----G-----tgt-----	129
CFZ.US.CFZUS	-A-GA-----G-----G-----G-----G-----tgt-----	129

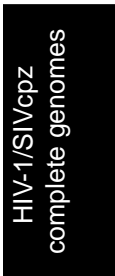
	5' LTR U3 end	/ +1 mRNA start site	5' LTR R repeat begin	Poly-A signal	5' LTR R repeat end	5' LTR U5 start	
B.FR.HXB2	TTTTTGGCTACTGGGTCCTCTCTGGTTAGACCAAGTCTTAAAGCCTCAATA.AAGCCTTGCCTTGGAGTG.CTTCAAGTAGTGTGTGCGCCG						566
CONSENSUS_A-A1							0
A.BY.97BL006							0
A.SE.SE7535							0
A.SE.SE8538							0
A.SE.SE8891							0
A.SE.UGSE8131							0
A1.KE.O23-17							36
A1.SE.SE7253							0
A1.UG.92UG037							0
A1.UG.U455							36
A2.CD.97CDKTB48							0
A2.CY.94CY017.41							0
CONSENSUS_B							529
B.AU.MBC200							563
B.CN.RL42							0
B.DE.D31							112
B.DE.HAN							36
B.ES.89SP061							113
B.GA.OYL							112
B.GB.MANC							512
B.NL.3202A21							567
B.ITW.TWCYS							565
B.US.JRFL							0
B.US.MNCG							565
B.US.RF							80
B.US.SF2							566
B.US.YU2							562
CONSENSUS_C							0
C.BR.92BR025							0
C.BR.98BR004							54
C.BW.96BW0502							567
C.BW.96BW06							56
C.BW.96BW15C02							0
C.ET.ETH2220							0
C.IL.98IS002							0
C.IN.93IN101							566
C.IN.94IN476							0
C.IN.98IN012							0
C.TZ.98TZ013							0
C.TZ.98TZ017							0
C.ZA.97ZA012							0
C.ZM.96ZM651							0
C.ZM.96ZM751							0
CONSENSUS_D							111
D.CD.84ZE085							80
D.CD.FLI							112
D.CD.Z266							566
D.KE.MB2059							566
D.UG.94UG1141							0
CONSENSUS_F1							0
F1.BE.VI850							0
F1.BR.93BR020.1							0
F1.FI.FIN9363							0
F1.FR.MF411							0
F2.CM.MP255							0
F2.CM.MP257							0
CONSENSUS_G							0
G.BE.DRCBL							532
G.FI.HH8793-12.1							0
G.NG.92NG083							0
G.SE.SE6165							0
H.BE.VI991							0
H.BE.VI997							0
H.CF.90CF056							0
J.SE.SE7022							0
J.SE.SE7887							0
K.CD.E0TB11C							0
K.CM.MP535							0

Extensive secondary structure in this region. See Rizvi, J Virol 67:2681-8(1993)		5' LTR U5 end	lys trna primer binding site		
TCTG.TTGTGTGACTCTGGTAACAGAGATCCC.TCAGAC.CCTTTTGTAGTCAGTGTGG.AAAATCTCTA...GCAAGTGGCCGCCGAACAGGG.....ACCTGAAAGCG.A					663
B.FR.HXB2	CONSENSUS_A-A1	-----Y-----		---tc---	10
A.BY.97BL006	A.SE.SE7535			---T---	41
A.SE.SE8538	A.SE.SE8891				0
A.SE.UGSE8131	A1.KE.O23-17				0
A1.SE.SE7253	A1.UG.92UG037				132
A1.UG.U455					0
A2.CD.97CDKTB48	A2.CY.94CY017.41				132
CONSENSUS_B	B.AU.MBC200			TTTT-CG-	13
B.CN.RL42	B.DE.D31			---TC-	34
B.DE.HAN	B.ES.89SP061				210
B.GA.OYL	B.GB.MANC				660
B.NL.3202A21	B.ITW.TWCYS				40
B.US.JRFL	B.US.MNCG				209
B.US.RF	B.US.SF2				133
B.US.YU2					212
CONSENSUS_C	C.BR.92BR025				209
C.BW.98BR004	C.BW.96BW0502				12
C.BW.96BW06	C.BW.96BW15C02				13
C.ET.ETH2220	C.IL.98IS002				150
C.IN.93IN101	C.IN.94IN476				664
C.IN.98IN012	C.TZ.98TZ013				659
C.TZ.98TZ017	C.ZA.97ZA012				12
C.ZM.96ZM651	C.ZM.96ZM751				13
CONSENSUS_D	D.CD.84ZRO85				10
D.CD.FLI	D.CD.Z266				9
D.KE.MB2059	D.UG.94UG1141				10
CONSENSUS_E1	F1.BE.VI850				10
F1.BR.93BR020.1	F1.FI.FIN9363				10
F1.FR.MF411	F2.CM.MP255				10
F2.CM.MP257	CONSENSUS_G				208
G.BE.DRCBL	G.FI.HH8793-12.1				177
G.NG.92NG083	G.SE.SE6165				209
H.BE.VI991	H.BE.VI997				663
H.CF.90CF056	J.SE.SE7022				663
J.SE.SE7887	K.CD.EOTB11C				10
K.CM.MP535					10

Extensive secondary structure in this region.
 See Rizvi, J Virol 67:2681-8(1993)

5' LTR U5 end \ | - Lys tRNA primer binding site - |
 TCTG.TTGTGACTCTGTAAGATAGATCCC.TCAGAC.CCTTTTGTAGTCAGTGTGG.AAAATCTCTA...GCAGTGGCCGCCAAACAGGG.....ACCTGAAAGCG.A

B.FR.HXB2	CONSENSUS_01_AE	663
01_AE.CF.90CF41697Gt-Ag	663
01_AE.CF.90CF402-GA-AG	628
01_AE.CF.90CF401-GT-AG	703
01_AE.TH.93TH057-GA-AG	610
01_AE.TH.93TH065-GT-AG	628
01_AE.TH.93TH253-GT-AG	628
01_AE.TH.95TH022-GT-AG	638
01_AE.TH.95TH047-GT-AG	638
01_AE.TH.CM240-GT-AG	234
CONSENSUS_02_AG-GT-AG	0
02_AG.CM.97CM-MP807-GT-AG	0
02_AG.FR.DJ263-GT-AG	31
02_AG.FR.DJ264-GT-AG	32
02_AG.GH.G829-GT-AG	0
02_AG.NG.IE8G-GT-AG	208
02_AG.SE.SE7812-GT-AG	55
02_AG.SN.96SE-MP1211-GT-AG	0
02_AG.SN.MP1213-GT-AG	0
03_AB.RU.KAL153-2-GT-AG	543
03_AB.RU.RU98001-GT-AG	41
04_cpx.CY.94CY032-3-GT-AG	33
04_cpx.GR.97PVCH-GT-AG	684
04_cpx.GR.97PVMY-GT-AG	688
05_DF.BE.VI1310-GT-AG	39
05_DF.BE.VI961-GT-AG	15
CONSENSUS_06_cpx-GT-AG	682
06_cpx.AU.BEP90-GT-AG	698
06_cpx.ML.95ML127-GT-AG	685
06_cpx.ML.95ML84-GT-AG	0
06_cpx.SN.97SE1078-GT-AG	708
07_BC.CN.97CN001-GT-AG	10
07_BC.CN.98CN009-GT-AG	10
07_BC.CN.C54A-GT-AG	0
CONSENSUS_08_BC-GT-AG	0
08_BC.CN.97CNGX6F-GT-AG	0
08_BC.CN.97CNGX7F-GT-AG	0
08_BC.CN.97CNGX9F-GT-AG	0
08_BC.CN.98CN006-GT-AG	10
10_CD.TZ.96TZBF061-GT-AG	41
10_CD.TZ.96TZBF071-GT-AG	40
10_CD.TZ.96TZBF110-GT-AG	40
CONSENSUS_11_cpx-GT-AG	658
11_cpx.CM.MP818-GT-AG	669
11_cpx.FR.MP1298-GT-AG	663
11_cpx.FR.MP1307-GT-AG	674
11_cpx.GR.GR17-GT-AG	0
12_BF.AR.ARMA159-GT-AG	683
12_BF.UY.URTR23-GT-AG	704
12_BF.UY.URTR35-GT-AG	683
N.CM.YBF106-GT-AG	221
N.CM.YBF30-GT-AG	221
CONSENSUS_0-GT-AG	663
O.CM.ANT70-GT-AG	692
O.CM.MVP5180-GT-AG	666
O.SN.SEMP1299-GT-AG	693
O.SN.SEMP1300-GT-AG	693
CONSENSUS_CPZ-GT-AG	257
CPZ_CD.CPZANT-GT-AG	133
CPZ_CD.CPZANT-GT-AG	133
CPZ_CD.CAM3-GT-AG	337
CPZ_CD.CAM5-GT-AG	681
CPZ_GA.CPZGAB-GT-AG	681
CPZ_US.CPZUS-GT-AG	686



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

Table of HIV-1/SIVcpz complete genomes. Columns include strain identifier (e.g., B.FR.HXB2, CONSENSUS_A-A1), sequence alignment (with dashes for gaps and lowercase 'a' for insertions), and position numbers (770, 103, 146, etc.).

B.FR.HXB2	GGG.....GGCTAGAAG.....	end packaging loops - Gag and Gag-Pol CDS startT-.....A-.....GC-.....C-.....	873
CONSENSUS_01_AET-.....A-.....GC-.....C-.....	882
01_AE.CF.90CF11697T-.....A-.....GC-.....C-.....	828
01_AE.CF.90CF402T-.....A-.....GC-.....C-.....	915
01_AE.CF.90CF4071T-.....A-.....GC-.....C-.....	816
01_AE.TH.93TH057T-.....A-.....GC-.....C-.....	832
01_AE.TH.93TH065T-.....A-.....GC-.....C-.....	831
01_AE.TH.93TH253T-.....A-.....GC-.....C-.....	890
01_AE.TH.95TH022T-.....A-.....GC-.....C-.....	842
01_AE.TH.95TH047T-.....A-.....GC-.....C-.....	842
01_AE.TH.CM240T-.....A-.....GC-.....C-.....	438
CONSENSUS_02_AGt-.....a-.....gC-.....g-.....	107
02_AG.CM.97CM-MP807t-.....a-.....gC-.....g-.....	47
02_AG.FR.DJ263t-.....a-.....gC-.....g-.....	236
02_AG.FR.DJ264t-.....a-.....gC-.....g-.....	237
02_AG.GH.G829t-.....a-.....gC-.....g-.....	107
02_AG.NG.IBNGt-.....a-.....gC-.....g-.....	413
02_AG.SE.SE7812t-.....a-.....gC-.....g-.....	260
02_AG.SN.98SE-MP1211-G-.....T-.....A-.....GC-.....G-.....	85
02_AG.SN.MP1213-G-.....T-.....A-.....GC-.....G-.....	82
03_AB.RU.KAL153-2-G-.....T-.....A-.....GC-.....G-.....	648
03_AB.RU.RU98001-G-.....T-.....A-.....GC-.....G-.....	249
04_cpx.CY.94CY032-3-G-.....T-.....A-.....GC-.....G-.....	239
04_cpx.FR.97FVGH-G-.....T-.....A-.....GC-.....G-.....	890
04_cpx.FR.97FVNY-G-.....T-.....A-.....GC-.....G-.....	893
05_DF.BE.VI1310-G-.....T-.....A-.....GC-.....G-.....	249
05_DF.BE.VI1961-G-.....T-.....A-.....GC-.....G-.....	225
CONSENSUS_06_cpx-G-.....T-.....A-.....GC-.....G-.....	883
06_cpx.AU.BFP90-G-.....T-.....A-.....GC-.....G-.....	903
06_cpx.ML.95ML127-G-.....T-.....A-.....GC-.....G-.....	890
06_cpx.ML.95ML84-G-.....T-.....A-.....GC-.....G-.....	110
06_cpx.SN.97SE1078-G-.....T-.....A-.....GC-.....G-.....	917
07_BC.CN.97CN001-G-.....T-.....A-.....GC-.....G-.....	220
07_BC.CN.98CN009-G-.....T-.....A-.....GC-.....G-.....	220
07_BC.CN.C54A-G-.....T-.....A-.....GC-.....G-.....	78
CONSENSUS_08_BC-G-.....T-.....A-.....GC-.....G-.....	80
08_BC.CN.97CNGX6F-G-.....T-.....A-.....GC-.....G-.....	80
08_BC.CN.97CNGX7F-G-.....T-.....A-.....GC-.....G-.....	86
08_BC.CN.97CNGX9F-G-.....T-.....A-.....GC-.....G-.....	80
08_BC.CN.98CN006-G-.....T-.....A-.....GC-.....G-.....	220
10_CD.TZ.96TZBF061-G-.....T-.....A-.....GC-.....G-.....	252
10_CD.TZ.96TZBF071-G-.....T-.....A-.....GC-.....G-.....	250
10_CD.TZ.96TZBF110-G-.....T-.....A-.....GC-.....G-.....	250
CONSENSUS_11_cpx-G-.....T-.....A-.....GC-.....G-.....	864
11_cpx.CM.MP818-G-.....T-.....A-.....GC-.....G-.....	879
11_cpx.FR.MP1298-G-.....T-.....A-.....GC-.....G-.....	876
11_cpx.FR.MP1307-G-.....T-.....A-.....GC-.....G-.....	885
11_cpx.FR.MP1307-G-.....T-.....A-.....GC-.....G-.....	190
12_BF.AR.ARMA159-G-.....T-.....A-.....GC-.....G-.....	893
12_BF.UY.URTR23-G-.....T-.....A-.....GC-.....G-.....	914
12_BF.UY.URTR35-G-.....T-.....A-.....GC-.....G-.....	893
N_CM.YBF106-G-.....T-.....A-.....GC-.....G-.....	432
N_CM.YBF30-G-.....T-.....A-.....GC-.....G-.....	431
CONSENSUS_0-G-.....T-.....A-.....GC-.....G-.....	894
O_CM.AN170-G-.....T-.....A-.....GC-.....G-.....	925
O_CM.MVP5180-G-.....T-.....A-.....GC-.....G-.....	900
O_SN.SEMP1299-G-.....T-.....A-.....GC-.....G-.....	927
O_SN.SEMP1300-G-.....T-.....A-.....GC-.....G-.....	927
CONSENSUS_CPZ-G-.....T-.....A-.....GC-.....G-.....	483
CPZ_CD.CPZANT-G-.....T-.....A-.....GC-.....G-.....	237
CPZ_CM.CAM3-G-.....T-.....A-.....GC-.....G-.....	263
CPZ_CM.CAM5-G-.....T-.....A-.....GC-.....G-.....	569
CPZ_GA.CPZGAB-G-.....T-.....A-.....GC-.....G-.....	910
CPZ_US.CPZUS-G-.....T-.....A-.....GC-.....G-.....	915
Gag pl7-G-.....T-.....A-.....GC-.....G-.....	Gag

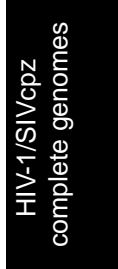
HIV-1/SIVcpz complete genomes

B.FR.HXB2	CAGAAGAACTTAGATCATTTATATAACAGTACCAACCCCTCTATTGTGTGTCATCAAAAGGATAGAGATATAAAGACACCAAGGAAGCTTTTAGACAAGATAGAGGAGGAGCAAAACAAAAGTAAAG...AAAAA	1130
CONSENSUS_A-A1	-----A-----A-----c-----tg-----Ata-----t-g-c-----c-----	461
A.BY.97BL006	-----T-----A-----C-----A-----TG-----C-----T-----A-----ATA-----TN-N-C-----C-----	506
A.SE.SE7535	-----A-----C-----A-----TG-----C-----C-----T-----A-----ATA-----T-G-G-----C-----	362
A.SE.SE8538	-----T-----GG-----C-----C-----TG-----G-----C-----T-----AT-----T-G-C-----C-----	335
A.SE.SE8891	-----C-----A-----A-----CTG-----A-----A-----ATA-----T-G-AC-----C-----	526
A.SE.UGSEF8131	-----A-----A-----A-----C-----TG-----A-----GATA-----T-G-C-----C-----	594
A1.KE.Q23-17	-----T-----T-----A-----C-----TG-----A-----ATA-----T-G-C-----C-----	337
A1.SE.SE7253	-----T-----C-----A-----C-----G-----A-----ATA-----G-G-C-----C-----	500
A1.UG.92UG037	-----C-----A-----C-----G-----TG-----A-----AT-----T-G-AC-----C-----G-----	597
A1.UG.U455	-----GT-----GT-----A-----A-----TG-----A-----A-----T-----C-----C-----G-----	477
A2.CD.97CDKTB48	-----A-----GT-----C-----T-----A-----A-----G-----G-----T-----C-----C-----G-----	493
A2.CY.94CY017.41	-----T-----GT-----C-----G-----A-----G-----TG-----A-----A-----G-----G-----C-----G-----	675
CONSENSUS_B	-----A-----A-----a-----g-----?-----?-----?-----?-----?-----?-----	1129
B.AU.MBC200	-----TC-----T-----A-----T-----A-----G-----G-----G-----G-----G-----G-----	502
B.CN.RL42	-----T-----C-----T-----A-----A-----G-----G-----G-----G-----G-----G-----	676
B.DE.D31	-----A-----A-----A-----A-----G-----G-----G-----G-----G-----G-----	599
B.DE.HAN	-----A-----A-----A-----A-----TG-----A-----G-----G-----G-----G-----	679
B.ES.89SP061	-----A-----A-----A-----A-----G-----G-----G-----G-----G-----G-----	676
B.GA.OYI	-----T-----T-----A-----A-----G-----G-----G-----G-----G-----G-----	1076
B.GB.MANC	-----C-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	1132
B.NL.3202A21	-----GT-----GT-----A-----A-----A-----A-----C-----C-----A-----A-----	1129
B.TW.TWCYS	-----T-----T-----A-----A-----A-----A-----G-----G-----A-----GC-----A-----	454
B.US.JREL	-----A-----A-----A-----A-----A-----A-----G-----G-----A-----G-----A-----	1127
B.US.MNCG	-----A-----A-----A-----A-----AT-----G-----G-----G-----G-----G-----G-----	644
B.US.RF	-----G-----G-----C-----A-----A-----A-----TG-----G-----G-----G-----G-----G-----	1131
B.US.SF2	-----C-----C-----A-----A-----A-----A-----G-----G-----G-----G-----G-----G-----	1128
B.US.YU2	-----G-----G-----C-----C-----G-----T-----A-----G-----a-----g-----G-----G-----	477
CONSENSUS_C	-----g-----c-----c-----c-----T-----T-----A-----G-----a-----g-----G-----G-----	480
C.BR.92BR025	-----A-----G-----GT-----C-----G-----T-----T-----A-----G-----A-----TG-----CG-----C-----C-----	477
C.BR.98BR004	-----A-----G-----GT-----T-----T-----T-----A-----G-----A-----TG-----CG-----C-----C-----	617
C.BW.96BW0502	-----G-----G-----G-----C-----C-----C-----C-----A-----GC-----G-----A-----G-----C-----	1133
C.BW.96BW06	-----TC-----C-----G-----T-----T-----G-----A-----G-----A-----A-----G-----C-----	621
C.BW.96BW15C02	-----G-----A-----A-----T-----C-----G-----T-----A-----CA-----G-----AC-----G-----G-----	510
C.ET.ETH2220	-----G-----A-----A-----T-----C-----G-----T-----A-----CA-----G-----AC-----G-----G-----	476
C.IL.98ILS002	-----G-----G-----A-----A-----C-----C-----C-----A-----CG-----G-----A-----C-----C-----	1130
C.IN.93IN101	-----G-----G-----A-----A-----C-----C-----C-----A-----CG-----G-----A-----C-----C-----	478
C.IN.94IN476	-----G-----G-----A-----A-----C-----C-----C-----A-----CG-----G-----A-----C-----C-----	479
C.IN.98IN012	-----G-----G-----A-----A-----C-----C-----C-----A-----CG-----G-----A-----C-----C-----	476
C.TZ.98TZ013	-----G-----G-----A-----A-----C-----C-----C-----A-----CG-----G-----A-----C-----C-----	478
C.TZ.98TZ017	-----G-----G-----A-----A-----C-----C-----C-----A-----CG-----G-----A-----C-----C-----	479
C.ZA.97ZA012	-----G-----G-----A-----A-----C-----C-----C-----A-----CG-----G-----A-----C-----C-----	477
C.ZM.96ZM651	-----G-----G-----A-----A-----C-----C-----C-----A-----CG-----G-----A-----C-----C-----	477
C.ZM.96ZM751	-----G-----G-----A-----A-----C-----C-----C-----A-----CG-----G-----A-----C-----C-----	477
CONSENSUS_D	-----A-----A-----A-----A-----a-----a-----g-----g-----?-----?-----?-----?	673
D.CD.84ZR085	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----C-----C-----	647
D.CD.FLI	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----C-----C-----	676
D.CD.Z2Z6	-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----AAC-----	1132
D.KE.MB2059	-----G-----A-----A-----A-----A-----CA-----A-----C-----G-----C-----A-----C-----	479
D.UG.94UG1141	-----G-----A-----A-----A-----A-----G-----G-----A-----A-----GC-----AG-----A-----GC-----	462
CONSENSUS_F1	-----G-----?-----?-----GT-----?-----a-----A-----a-----?-----g-----?-----g-----?-----C-----C-----	485
F1.BE.VI850	-----G-----AG-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----C-----C-----	473
F1.BR.93BR020.1	-----G-----C-----A-----A-----GT-----A-----A-----A-----A-----A-----A-----G-----G-----G-----G-----CG-----	471
F1.FI.FIN9363	-----G-----C-----A-----A-----GT-----T-----A-----A-----A-----A-----A-----A-----G-----G-----G-----G-----CG-----	341
F1.FR.MF411	-----G-----G-----TC-----G-----TC-----G-----T-----A-----A-----A-----A-----A-----A-----G-----G-----G-----G-----CG-----	341
F2.CM.MP255	-----G-----G-----A-----C-----C-----G-----T-----GTT-----A-----A-----A-----A-----TG-----G-----A-----G-----T-----C-----C-----	341
F2.CM.MP257	-----G-----G-----A-----C-----C-----G-----T-----GTT-----A-----A-----A-----A-----TG-----G-----A-----G-----T-----C-----C-----	341
CONSENSUS_G	-----G-----ga-----a-----t-----C-----C-----A-----T-----A-----A-----a-----a-----TG-----G-----A-----T-----G-----AC-----C-----G-----	526
G.BE.DRCBL	-----G-----GA-----A-----C-----T-----T-----C-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----C-----C-----	1115
G.FI.HH8793-12.1	-----G-----GA-----A-----T-----T-----C-----T-----A-----C-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----C-----C-----	532
G.ING.92ING083	-----G-----G-----A-----A-----T-----T-----C-----C-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----C-----C-----	496
G.SE.SE6165	-----G-----A-----A-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	533
H.BE.VI991	-----C-----CA-----A-----A-----GT-----A-----A-----A-----A-----TG-----G-----A-----T-----G-----AC-----C-----G-----	516
H.BE.VI997	-----CW-----T-----C-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	457
H.CF.90CF056	-----T-----CT-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	477
J.SE.SE7022	-----A-----A-----A-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	454
J.SE.SE7887	-----C-----A-----A-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	453
K.CD.EQTB11C	-----G-----G-----C-----T-----G-----A-----A-----A-----A-----T-----C-----G-----GG-----A-----AC-----A-----G-----C-----C-----	341
K.CM.MP535	-----GA-----A-----C-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----CG-----	341

Accession	Sequence	Position
B.FR.HXB2	AGCACAGCAAGCAGCAGCT.....GACACAGGA.....	1215
CONSENSUS_01_AE	ga-----g-----g-----GCA-A-----A--TGCA--A--	1173
01_AE.CF.90CF11697	GA-----G-----G-----AG-A--GCA-A-----A--TGCA--A--	1170
01_AE.CF.90CF402	GA-----G-----G-----AG--GCA-A-----A--TGCA--A--	1257
01_AE.CF.90CF4071	GA-----G-----G-----AG--GCA-A-----A--TGCA--A--	1158
01_AE.TH.93TH057	A-----G-----G-----AG--GCA-A-----A--TGCA--A--	1174
01_AE.TH.93TH065	GA-----G-----G-----AG--GCA-A-----A--TGCA--A--	1173
01_AE.TH.93TH253	GAA-----G-----G-----AG--GCA-A-----A--TGCA--A--	1232
01_AE.TH.95TH022	GA-----G-----G-----AG--GCA-A-----A--TGCA--A--	1184
01_AE.TH.95TH047	GA-----G-----G-----AG--GCA-A-----A--TGCA--A--	1184
01_AE.TH.CM240	GA-----G-----G-----AG--GCA-A-----A--TGCA--A--	780
CONSENSUS_02_AG	ga-----g-----g-----gc.???-----A--TGCA--a--	438
02_AG.CM.97CM-MP807	G-----G-----G-----G-----TG-----A--TGCA--A--	377
02_AG.FR.DJ263	GA-----G-----G-----G-----G-----A--TGCA--A--	569
02_AG.FR.DJ264	GA-----G-----G-----G-----G-----A--TGCA--A--	570
02_AG.GH.G829	GA-----G-----G-----G-----G-----A--TGCA--A--	437
02_AG.NG.IBNG	G-T-----GA-----G-----G-----G-----A--TGCA--A--	746
02_AG.SE.SE7812	GA-----G-----G-----G-----G-----A--TGCA--A--	593
02_AG.SN.96SE-MP1211	GA-----G-----G-----G-----G-----A--TGCA--A--G--	418
02_AG.SN.MP1213	GA-----G-----G-----G-----G-----A--TGCA--A--G--	415
03_AB.RU.KAL153-2	GA-C-A-G-----A-----AG--GCA-A-----A--TGCA--A--	990
03_AB.RU.RU98001	GA-C-A-G-----A-----AG--GCA-A-----A--TGCA--A--	591
04_cpx.CV.94CV032-3	GA-----G-----G-----GGT-----GCA-T-----A--TGCA--A--	581
04_cpx.GR.97PVGH	GA-----G-----G-----GGT-----GCA-T-----A--TGCA--A--	1232
04_cpx.GR.97PVNY	GA-----G-----G-----A-T-----GCA-T-----A--TGCA--A--	1229
05_DF.BE.VI1310	G-----A-----A-----A-----GC-----C-----T-----A--	594
05_DF.BE.VI1961	G-----A-----A-----A-----GGC-C-----T-----A--	570
CONSENSUS_06_cpx	G-----g-----C-----A-----G-a-TC-----C-----a--TGCA--A--	1217
06_cpx.AU.BFP90	G-----g-----C-----A-----G-a-TC-----C-----a--TGCA--A--	1240
06_cpx.ML.95ML127	G-----A-----C-----C-----A-----G-a-TC-----C-----a--TGCA--A--	1232
06_cpx.ML.95ML84	G-T-----C-----C-----A-----G-a-TC-----C-----a--TGCA--A--	452
06_cpx.SN.97SE1078	GA-G-----GA-A-----A-----A-----G-a-TC-----C-----a--TGCA--A--	1259
07_BC.CN.97CN001	A-----G-----AAG-AG-----A-----A-----TC-----A--	553
07_BC.CN.98CN009	A-----G-----AAG-AG-----A-----A-----TC-----A--	553
07_BC.CN.C54A	A-----G-----AAG-AG-----A-----A-----TC-----A--	411
CONSENSUS_08_BC	A-----g-----Aa-aaa????????-----ctGAC-ag-----TCC--A--	413
08_BC.CN.97CNGX6F	A-----g-----Aa-aaa-----ctGAC-ag-----TCC--A--	413
08_BC.CN.97CNGX7F	A-----G-----AA-AAA-----CTGAC-AG-----TCC--A--	419
08_BC.CN.97CNGX9F	A-----G-----AA-AAA-----CTGAC-AG-----TCC--A--	413
08_BC.CN.98CN006	A-----AA-A--CAGCAGCAAAAGAG-----CTGAC-G-----TC--A--	565
10_CD.TZ.96TZBF061	G-----A-----A-----A-----GC-----T-----A--TT--A--	594
10_CD.TZ.96TZBF071	G-----A-----A-----A-----GC-----T-----A--TT--A--	592
10_CD.TZ.96TZBF110	G-----A-----A-----A-----GC-----T-----A--TT--A--	592
CONSENSUS_11_cpx	GAA-----G-----A-----?ca?-----T-----c-----a--TGCA--A--?	1198
11_cpx.CM.MP818	GAA-----G-----A-----GCA-T-----T-----c-----a--TGCA--A--?	1221
11_cpx.FR.MP1298	GAA-----A-----G-----A-----GCA-T-----T-----c-----a--TGCA--A--	1218
11_cpx.FR.MP1307	GAA-----G-----A-----A-----GCA-T-----T-----c-----a--TGCA--A--	1227
11_cpx.GR.GR17	GAA-----G-----T-----CA-----CA-----T-----TGCA--A--	526
12_BF.AR.ARMA159	A-----G-----G-----A-----A-----TC-T-----A--	1223
12_BF.UY.URTR23	GA-----G-----A-----A-----A-----TC-T-----A--	1244
12_BF.UY.URTR35	GA-----G-----A-----A-----A-----TC-T-----A--	1223
N.CM.YBF106	AAAG-A-C-A-----GGCAGCGCAGCACT-----T-----A--	780
N.CM.YBF30	AA-A-C-C-A-----GGCAGCGCAGCACT-----T-----A--	779
CONSENSUS_0	?-t?????Ag-????????-----AG-gca-gg-acCg-T-----AaCA--TGCA--A--	1196
O.CM.ANT170	T-G-----AC-C--AAGGAA-----AG-GCA-GG-CGG-T-----ATCA--TGCG--A--	1261
O.SN.MVP5180	T-----G-----A-----A-----AG-CCT-GG-ACA-----AACA--TGCA--A--	1236
O.SN.SEMP1299	-----GTACCGCTAAG-A-----AG-GCA-GG-ACGG-T-----AACA--TGCA--A--	1263
O.SN.SEMP1300	-----GTGCGCTAAG-A-----AG-GCA-GG-ACGG-T-----AGCA--TGCA--A--	1263
CONSENSUS_CPZ	-?-????-c-c-?a?aa?????????cagc?????????????????gat-gggc-?--taG--a--g--	795
CPZ.CD.CPZANT	-G-AGTAGCCA-A-C--AAGCAGAGCATGCTCTGGGGTCTCTGTTAAACAACAGTGTGTCAGCGA--TCTT-G-GGC-AAG-GAG--C-CA-AGT-G-TGCAGGA--AAAT-GCA	627
CPZ.CM.CAM3	-TTGGC--C-CGACGAG-----GAA-GCCT--TAG-C-AG-A-----TGCA--A--A--GC-A	605
CPZ.CM.CAM5	-TTGGC--C-CGACGAA-----GAA-GCAGCA-A-TAG-C-A-----TGCA--A--A--	908
CPZ.GA.CPZGAB	T-A-AGTA-CT--G-AGCCGTGAAGGGGAGCCAGT-----TCTTC-GGCA-T-TGG-----CC-C-A-A--TGCT--A--G--	1276
CPZ.US.CPZUS	CA--G--G--A-AA-AAACAACAGCAACAG-----GG-A-T-----AGC-----AG-A-A--GCT-----A--G--	1263
Gag p17	A_Q_Q_A_A_A_A-----D_T_G-----H_S_N_Q_V_S_Q_N_Y_P_I_V_Q_N_I_Q_G_Q_M	Gag

HIV-1/SIVcpz complete genomes

B_FR_HXB2	CCATGCTAAACACAGTGGGGGACATCAAGCAGCATCAAAATGTTAAAGAGACCATCAATGAGAAAGCTGCAGAAATGGATGATGCATGCAGGCGCTATTGCACCGCCAGATGAG	1475
CONSENSUS_01_AE	TTG-----t-T-----c-G-----A-----a-----g-A-C-----A-----C-----	1433
01_AE_CF_90CF11697	TG-----T-----G-----A-----A-----G-----A-----C-----A-----C-----	1430
01_AE_CF_90CF402	TTG-----C-----T-----C-----G-----A-----A-----T-----G-----A-----C-----A-----C-----	1430
01_AE_CF_90CF4071	TG-----T-----T-----C-----G-----A-----A-----T-----G-----A-----C-----A-----C-----	1517
01_AE_TH_93TH057	TTG-----T-----A-----G-----A-----A-----T-----G-----A-----C-----A-----C-----	1418
01_AE_TH_93TH065	TG-----T-----T-----C-----G-----A-----A-----C-----A-----C-----A-----C-----	1434
01_AE_TH_93TH065	TTG-----T-----T-----C-----G-----A-----A-----C-----A-----C-----A-----C-----	1433
01_AE_TH_93TH065	TTG-----T-----T-----C-----G-----A-----A-----C-----A-----C-----A-----C-----	1432
01_AE_TH_95TH022	TTG-----T-----T-----C-----G-----A-----A-----C-----A-----C-----A-----C-----	1444
01_AE_TH_95TH047	TTG-----T-----T-----C-----G-----A-----A-----C-----A-----C-----A-----C-----	1444
01_AE_TH_CM240	TTG-----T-----T-----C-----G-----A-----A-----C-----A-----C-----A-----C-----	1040
CONSENSUS_02_AG	TTG-----T-----T-----C-----G-----A-----g-----A-----T-----T-----G-----A-----C-----A-----C-----	698
02_AG_CM_97CM-MP807	TTG-----T-----T-----G-----C-----G-----A-----T-----T-----G-----A-----C-----A-----C-----	637
02_AG_FR_DJ263	TTG-----T-----T-----C-----G-----A-----G-----A-----T-----T-----G-----A-----C-----A-----C-----	829
02_AG_FR_DJ264	TTG-----T-----T-----C-----G-----A-----G-----A-----T-----T-----G-----A-----C-----A-----C-----	830
02_AG_GH_G829	TTG-----T-----T-----C-----G-----A-----G-----A-----T-----T-----G-----A-----C-----A-----C-----	697
02_AG_NG_IBNG	TTG-----T-----T-----C-----G-----A-----G-----A-----T-----T-----G-----A-----C-----A-----C-----	1006
02_AG_SE_SE7812	TTG-----T-----T-----C-----G-----A-----G-----A-----T-----T-----G-----A-----C-----A-----C-----	853
02_AG_SN_98SE-MP1211	TTG-----T-----T-----C-----G-----A-----T-----G-----C-----G-----A-----C-----A-----C-----	678
02_AG_SN_MP1213	TTG-----T-----T-----C-----G-----A-----G-----A-----T-----T-----G-----A-----C-----A-----C-----	675
03_AB_RU_KAL153-2	TTG-----G-----T-----T-----C-----G-----T-----G-----T-----T-----G-----A-----C-----A-----G-----	1250
03_AB_RU_RU98001	TTG-----G-----T-----T-----C-----G-----T-----G-----T-----T-----G-----A-----C-----A-----G-----	851
04_cpx_CX_94CY032-3	TTG-----T-----T-----C-----G-----A-----A-----T-----T-----C-----G-----A-----C-----A-----C-----	841
04_cpx_GR_97PVGH	TTG-----T-----T-----C-----G-----A-----A-----T-----T-----C-----G-----A-----C-----A-----C-----	1492
04_cpx_GR_97PVNY	TTG-----T-----T-----C-----G-----A-----A-----T-----T-----C-----G-----A-----C-----A-----C-----	1489
05_DF_BE_V11310	-----T-----C-----G-----A-----A-----T-----T-----G-----A-----C-----A-----G-----	854
05_DF_BE_V1961	-----T-----C-----G-----A-----A-----T-----T-----G-----A-----C-----A-----G-----	830
CONSENSUS_06_cpx	TTG-----G-----T-----C-----G-----T-----T-----C-----?-----b-----A-----?-----C-----C-----A-----	1471
06_cpx_AU_BFP90	TTG-----G-----T-----C-----G-----T-----T-----C-----G-----A-----A-----A-----A-----C-----A-----	1500
06_cpx_ML_95ML127	TTG-----G-----T-----C-----G-----T-----T-----C-----G-----A-----A-----A-----A-----C-----A-----	1492
06_cpx_ML_95ML84	TTG-----G-----T-----C-----G-----T-----T-----C-----G-----A-----A-----A-----A-----C-----A-----	712
06_cpx_SN_97SEL078	TTG-----G-----T-----C-----G-----T-----T-----C-----G-----A-----A-----A-----A-----C-----A-----	1519
07_BC_CN_97CN001	-----T-----A-----A-----A-----A-----T-----T-----A-----A-----C-----A-----A-----	813
07_BC_CN_98CN009	-----T-----A-----A-----A-----A-----T-----T-----A-----A-----C-----A-----A-----	813
07_BC_CN_C54A	-----T-----A-----A-----A-----A-----T-----T-----A-----A-----C-----A-----A-----	671
CONSENSUS_08_BC	-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	673
08_BC_CN_97CNGX6F	-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	673
08_BC_CN_97CNGX7F	-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	679
08_BC_CN_97CNGX9F	-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	673
08_BC_CN_98CN006	-----T-----T-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----	825
10_CD_TZ_96TZBF061	-----T-----G-----C-----A-----A-----A-----A-----G-----C-----A-----A-----G-----	854
10_CD_TZ_96TZBF071	-----T-----G-----C-----A-----A-----A-----A-----G-----C-----A-----A-----G-----	852
10_CD_TZ_96TZBF110	-----T-----G-----C-----A-----A-----A-----A-----G-----C-----A-----A-----G-----	852
CONSENSUS_11_cpx	TTG-----G-----T-----C-----g-----T-----T-----C-----G-----A-----A-----A-----C-----	1454
11_cpx_CM_MP818	TTG-----G-----T-----C-----G-----T-----T-----C-----G-----A-----A-----A-----C-----	1481
11_cpx_FR_MP1298	TTG-----G-----T-----C-----G-----T-----T-----C-----G-----A-----A-----A-----C-----	1478
11_cpx_FR_MP1307	TTG-----G-----T-----C-----G-----T-----T-----C-----G-----A-----A-----A-----C-----	1487
11_cpx_GR_GR17	TTG-----G-----T-----C-----G-----T-----T-----C-----G-----A-----A-----A-----C-----	786
12_BF_AR_ARMA159	-----T-----T-----A-----A-----T-----T-----C-----C-----T-----A-----A-----C-----	1483
12_BF_UY_URTR23	-----T-----T-----A-----A-----T-----T-----C-----C-----T-----A-----A-----C-----	1504
12_BF_UY_URTR35	-----T-----T-----A-----A-----T-----T-----C-----C-----T-----A-----A-----C-----	1483
N_CM_YBF106	-----T-----T-----A-----A-----T-----T-----G-----G-----C-----G-----G-----G-----A-----	1040
N_CM_YBF30	-----T-----T-----A-----A-----T-----T-----G-----G-----C-----G-----G-----G-----A-----	1039
CONSENSUS_0	-----T-----T-----G-----T-----A-----G-----C-----A-----A-----T-----G-----A-----C-----A-----C-----	1449
O_CM_ANT70	-----T-----T-----G-----T-----A-----G-----C-----A-----A-----T-----G-----A-----C-----A-----C-----	1521
O_CM_MVP5180	-----T-----T-----G-----T-----A-----G-----C-----A-----A-----T-----G-----A-----C-----A-----C-----	1496
O_SN_SEMP1299	-----T-----T-----G-----T-----A-----G-----C-----A-----A-----T-----G-----A-----C-----A-----C-----	1523
O_SN_SEMP1300	-----T-----T-----G-----T-----A-----G-----C-----A-----A-----T-----G-----A-----C-----A-----C-----	1523
CONSENSUS_CPZ	-----t-----t-----g-----g-----?-----g-----g-----?-----g-----g-----?-----t-----a-----c-----a-----	1041
CPZ_CD_CFPZANT	-----T-----T-----G-----T-----A-----G-----C-----G-----C-----G-----G-----G-----A-----A-----C-----A-----	867
CPZ_CM_CAM3	-----T-----T-----G-----T-----A-----G-----C-----G-----C-----G-----G-----G-----A-----A-----C-----A-----	865
CPZ_CM_CAM5	-----T-----T-----G-----T-----A-----G-----C-----G-----C-----G-----G-----G-----A-----A-----C-----A-----	1168
CPZ_GA_CFPZGAB	-----T-----T-----G-----T-----A-----G-----C-----G-----C-----G-----G-----G-----A-----A-----C-----A-----	1536
CPZ_US_CFPZUS	-----T-----T-----G-----T-----A-----G-----C-----G-----C-----G-----G-----G-----A-----A-----C-----A-----	1523
Gag p24	T_M_L_N_T_V_G_H_Q_A_A_M_Q_M_L_K_E_T_I_N_E_A_A_E_W_D_R_V_H_P_V_H_A_G_P_I_A_P_G_Q_M_R	Gag



B.FR.HXB2	AGAACCAAGGGAAGTGCATAGCAGAACTACTAGTACCCCTTCAGGAAACAATAGGATGGATGACA.....AATAATCCACCTATCCCACTAGGAGAAAATTTATAAAAGATGGATAAATCCTGGGA	1596
CONSENSUS_A-A1	-----C--A-----	927
A.BY.97BL006	-----R-----	972
A.SE.SE7535	-----A-----	825
A.SE.SE8538	-----C-A-G-----	801
A.SE.SE8891	-----C-A-T-CA-----	801
A.SE.UGSE8131	-----C-A-----	992
A1.KE.Q23-17	-----TCC-C-G-----	1060
A1.SE.SE7253	-----C-A-G-----	803
A1.UG.92UG037	-----C-A-----	966
A1.UG.U455	-----C-G-A-----	1051
A2.CD.97CDKTB48	-----G-G-A-----	961
A2.CY.94CY017.41	-----T-----	953
CONSENSUS_B	-----C-----	1141
B.AU.MBC200	-----A-----	1595
B.CN.RL42	-----T-----	968
B.DE.D31	-----G-----	1142
B.DE.HAN	-----A-----	1065
B.ES.89SP061	-----C-G-----	1145
B.GA.OYI	-----C-----	1142
B.GB.MANC	-----C-----	1542
B.NL.3202A21	-----C-----	1598
B.TW.TWCYS	-----G-----	1595
B.US.JREL	-----C-----	913
B.US.MNCG	-----C-----	1602
B.US.RF	-----C-----	1110
B.US.SF2	-----C-----	1603
B.US.YU2	-----C-----	1594
CONSENSUS_C	-----g-t-g-----	934
C.BR.92BR025	-----C-G-----	940
C.BR.98BR004	-----C-G-----	937
C.BW.96BW0502	-----G-C-G-T-G-----	1074
C.BW.96BW06	-----C-C-----	1590
C.BW.96BW15C02	-----T-G-C-C-----	1078
C.ET.ETH2220	-----GGG-C-----	970
C.IL.98ILS002	-----C-G-T-G-----	933
C.IN.93IN101	-----GG-C-----	1587
C.IN.94IN476	-----G-C-T-G-----	935
C.IN.98IN012	-----G-C-T-G-----	939
C.TZ.98TZ013	-----G-C-G-----	947
C.TZ.98TZ017	-----G-C-G-----	933
C.ZA.97ZA012	-----G-C-T-G-----	939
C.ZM.96ZM651	-----G-C-T-G-----	940
C.ZM.96ZM751	-----C-T-----	934
CONSENSUS_D	-----GC-----	1139
D.CD.84ZR085	-----GC-----	1113
D.CD.FLI	-----G-C-----	1142
D.CD.Z2Z6	-----GC-C-----	1599
D.KE.MB2059	-----GC-----	1598
D.UG.94UG1141	-----GC-----	945
CONSENSUS_F1	-----CA-----	912
F1.BE.VI850	-----?GC-C-----	939
F1.BR.93BR020.1	-----GC-C-----	927
F1.FI.FIN9363	-----G-C-----	931
F1.FR.MF411	-----GC-C-----	795
F2.CM.MP255	-----GGC-C-----	795
F2.CM.MP257	-----GC-C-----	807
CONSENSUS_G	-----C-----	985
G.BE.DRCBL	-----G-A-----	1560
G.FI.HH8793-12.1	-----GC-C-----	998
G.NG.92NG083	-----G-A-----	962
G.SE.SE6165	-----T-G-ACT-----	999
H.BE.VI991	-----G-C-----	988
H.BE.VI997	-----C-C-----	923
H.CF.90CF056	-----GGC-AG-----	943
J.SE.SE7022	-----A-C-----	920
J.SE.SE7887	-----GGC-----	919
K.CD.EQTB11C	-----AC-----	795
K.CM.MP535	-----C-C-----	795

Accession	Sequence	Position
B_FR_HXB2	TAAATAAATAGTAAGAATGTATAGCCCTACCAGCAATTCGGACATAAGAACAAAGGACCAACCCCTTTAGAGACTATGTAGACCGGTTCTATAAAAACCTTAAGAGCGGCAAGCTTCACAGGAGG	1726
CONSENSUS_01_AE	-----G-----T-----G-----A-----A-----C-----C-----G-----A-----A-----A-----	1684
01_AE_CF_90CF11697	GTT-----T-----G-----TA-----G-----TA-----G-----A-----A-----A-----A-----	1681
01_AE_CF_90CF402	GTT-----T-----G-----A-----A-----T-----T-----C-----C-----G-----A-----A-----A-----	1768
01_AE_CF_90CF4071	GTT-----T-----G-----A-----A-----T-----T-----C-----C-----G-----A-----A-----A-----	1669
01_AE_TH_93TH057	-----C-----G-----A-----A-----T-----T-----C-----C-----G-----A-----A-----A-----	1685
01_AE_TH_93TH065	GTT-----T-----G-----A-----A-----T-----T-----C-----C-----G-----A-----A-----A-----	1684
01_AE_TH_93TH253	-----CAA-----G-----A-----A-----T-----T-----C-----C-----G-----A-----A-----A-----	1743
01_AE_TH_95TH022	GTT-----T-----G-----A-----A-----T-----T-----C-----C-----G-----A-----A-----A-----	1695
01_AE_TH_95TH047	GTT-----T-----G-----A-----A-----T-----T-----C-----C-----G-----A-----A-----A-----	1695
01_AE_TH_CM240	GTT-----T-----G-----A-----A-----T-----T-----C-----C-----G-----A-----A-----A-----	1291
CONSENSUS_02_AG	-----g-----A-----A-----t-----t-----g-----t-----a-----a-----a-----?	946
02_AG_CM_97CM-MP807	T-----G-----A-----A-----T-----T-----G-----T-----A-----A-----A-----A-----	888
02_AG_FR_DJ263	GTT-----T-----G-----A-----A-----T-----T-----G-----T-----A-----A-----A-----A-----	1080
02_AG_FR_DJ264	GTT-----T-----G-----A-----A-----T-----T-----G-----T-----A-----A-----A-----A-----	1081
02_AG_GH_G829	-----G-----G-----A-----A-----T-----T-----C-----C-----G-----A-----A-----A-----	948
02_AG_NG_IBNG	G-----G-----G-----A-----A-----T-----T-----C-----C-----G-----A-----A-----A-----	1257
02_AG_SE_SE7812	-----G-----G-----A-----A-----T-----T-----G-----T-----A-----A-----A-----A-----	1104
02_AG_SN_98SE-MP1211	GTT-----T-----G-----A-----A-----T-----T-----G-----T-----A-----A-----A-----A-----	929
02_AG_SN_MF1213	GTT-----T-----G-----A-----A-----T-----T-----G-----T-----A-----A-----A-----A-----	926
03_AB_RU_KAL153-2	-----G-----G-----A-----A-----T-----T-----A-----A-----T-----T-----G-----T-----A-----A-----	1501
03_AB_RU_RU98001	-----G-----G-----A-----A-----T-----T-----A-----A-----T-----T-----G-----T-----A-----A-----	1102
04_CPX_CX_94CY032-3	-----C-----TT-----A-----A-----C-----C-----T-----T-----TG-----C-----A-----A-----A-----C-----	1092
04_CPX_GR_97PVGH	GTT-----T-----G-----A-----A-----T-----T-----G-----T-----A-----A-----A-----A-----	1743
04_CPX_GR_97PVMY	GTT-----T-----G-----A-----A-----T-----T-----G-----T-----A-----A-----A-----A-----	1740
05_DF_BE_VI1310	-----GT-----T-----T-----T-----T-----G-----T-----A-----A-----A-----A-----	1105
05_DF_BE_VI1961	-----GT-----T-----T-----T-----T-----G-----T-----A-----A-----A-----A-----	1081
CONSENSUS_06_CPX	-----G-----A-----A-----C-----G-----T-----C-----A-----T-----T-----A-----A-----	1722
06_CPX_AU_BFP90	GTTG-----T-----A-----A-----T-----T-----C-----G-----T-----A-----A-----A-----A-----	1751
06_CPX_ML_95ML127	-----G-----A-----A-----T-----T-----C-----G-----T-----A-----A-----A-----A-----	1743
06_CPX_ML_95ML84	-----G-----A-----A-----T-----T-----C-----G-----T-----A-----A-----A-----A-----	963
06_CPX_SN_97SEL078	-----G-----A-----A-----T-----T-----C-----G-----T-----A-----A-----A-----A-----	1770
07_BC_CN_97CN001	-----A-----G-----A-----A-----T-----T-----G-----A-----A-----A-----A-----	1064
07_BC_CN_98CN009	-----A-----G-----A-----A-----T-----T-----G-----A-----A-----A-----A-----	1064
07_BC_CN_C54A	-----A-----G-----A-----A-----T-----T-----G-----A-----A-----A-----A-----	922
CONSENSUS_08_BC	-----A-----G-----A-----A-----T-----T-----G-----A-----A-----A-----A-----	923
08_BC_CN_97CNGX6F	-----A-----G-----A-----A-----T-----T-----G-----A-----A-----A-----A-----	924
08_BC_CN_97CNGX7F	-----A-----G-----A-----A-----T-----T-----G-----A-----A-----A-----A-----	930
08_BC_CN_97CNGX9F	-----A-----G-----A-----A-----T-----T-----G-----A-----A-----A-----A-----	924
08_BC_CN_98CN006	-----A-----G-----A-----A-----T-----T-----G-----A-----A-----A-----A-----	1076
10_CD_TZ_96TZBF061	-----GT-----T-----T-----A-----T-----T-----G-----T-----A-----A-----A-----A-----	1105
10_CD_TZ_96TZBF071	-----GT-----T-----T-----A-----T-----T-----G-----T-----A-----A-----A-----A-----	1103
10_CD_TZ_96TZBF110	-----GT-----T-----T-----A-----T-----T-----G-----T-----A-----A-----A-----A-----	1103
CONSENSUS_11_CPX	-----GTT-----T-----T-----A-----A-----C-----C-----t-----t-----TA-----T-----TC-----?	1699
11_CPX_CM_MP818	GTT-----T-----T-----A-----A-----C-----C-----t-----t-----TA-----T-----TC-----?	1732
11_CPX_FR_MP1298	GTT-----T-----T-----A-----A-----C-----C-----t-----t-----TA-----T-----TC-----?	1729
11_CPX_FR_MP1307	GTT-----T-----T-----A-----A-----C-----C-----t-----t-----TA-----T-----TC-----?	1738
11_CPX_GR_GR17	GTT-----T-----T-----A-----A-----C-----C-----t-----t-----TA-----T-----TC-----?	1037
12_BF_AR_ARM159	-----GT-----T-----T-----A-----A-----T-----T-----G-----A-----A-----A-----A-----	1734
12_BF_UY_URTR23	-----GT-----T-----T-----A-----A-----T-----T-----G-----A-----A-----A-----A-----	1755
12_BF_UY_URTR35	-----GT-----T-----T-----A-----A-----T-----T-----G-----A-----A-----A-----A-----	1734
N_CM_YBF106	-----C-----G-----T-----G-----A-----A-----G-----C-----A-----A-----G-----C-----C-----T-----A-----A-----G-----AA-----	1291
N_CM_YBF30	-----C-----G-----T-----G-----A-----A-----G-----C-----A-----A-----G-----C-----C-----T-----A-----A-----G-----AA-----	1290
CONSENSUS_0	?-----C-----?-----g-----A-----A-----C-----AGTG-----CT-----A-----T-----T-----?-----G-----A-----A-----A-----A-----	1691
O_CM_ANT70	C-----C-----G-----A-----A-----C-----T-----AGTG-----CT-----A-----T-----T-----AG-----A-----A-----A-----A-----	1775
O_CM_MVP5180	C-----C-----G-----A-----A-----C-----T-----AGTG-----CT-----A-----T-----T-----AG-----A-----A-----A-----A-----	1750
O_SN_SEMP1299	-----C-----G-----A-----A-----C-----AGTG-----CT-----A-----T-----T-----AG-----A-----A-----A-----A-----	1777
O_SN_SEMP1300	-----C-----G-----A-----A-----C-----AGTG-----CT-----A-----T-----T-----AG-----A-----A-----A-----A-----	1777
CONSENSUS_CPZ	c-----?-----c-----g-----t-----a-----g-----t-----c-----a-----g-----t-----a-----a-----c-----?-----t-----a-----a-----	1282
CPZ_CD_CPZANT	-----GG-----G-----C-----G-----N-----T-----AGT-----A-----A-----G-----A-----G-----C-----T-----G-----TA-----A-----A-----A-----CCT	1147
CPZ_CM_CAM3	C-----T-----C-----G-----A-----G-----A-----G-----A-----G-----A-----G-----A-----G-----A-----A-----A-----A-----	1116
CPZ_CM_CAM5	C-----C-----G-----A-----G-----A-----T-----C-----A-----G-----A-----G-----A-----G-----A-----A-----A-----A-----	1419
CPZ_GA_CPZGAB	C-----C-----G-----T-----G-----A-----T-----C-----T-----C-----G-----A-----A-----G-----A-----A-----A-----A-----	1787
CPZ_US_CPZUS	C-----T-----G-----A-----A-----CT-----T-----CGTTG-----T-----CT-----A-----T-----C-----AG-----G-----T-----A-----G-----A-----A-----T-----CA-----A-----	1774
Gag p24	L_N_K_I_V_R_M_Y_S_P_T_S_I_L_D_I_R_Q_G_P_K_E_P_F_R_D_Y_V_D_R_F_Y_K_T_L_R_A_E_Q_A_S_Q_E	Gag



B_FR_HXB2	TAATAAATTGGATGACAGAAACCTTTGGTCCAAAATCGAATCCAGATTGTAAGACTATTTTTAAAGCAATTTGGGACCGCGCTACACTAGAAGAAATGATGACAGCATGTTCAGGGAGTAGGGAC	1856
CONSENSUS_01_AE	-----C-A-----t-----C-----T-C-----A-----A-----G-----T-----C-----G-----	1814
01_AE_CF_90CF11697	-----AC-A-----T-----C-----T-C-----A-----A-----G-----T-----C-----G-----	1811
01_AE_CF_90CF402	-----T-----C-A-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1898
01_AE_CF_90CF4071	-----C-A-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1799
01_AE_TH_93TH057	-----C-A-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1815
01_AE_TH_93TH065	-----C-A-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1814
01_AE_TH_93TH253	-----C-A-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1873
01_AE_TH_95TH022	-----C-A-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1825
01_AE_TH_95TH047	-----C-A-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1825
01_AE_TH_CM240	-----C-A-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1421
CONSENSUS_02_AG	-----C-----T-----C-----C-----T-----C-----G-----T-----C-----G-----	1076
02_AG_CM_97CM-MP807	-----C-----A-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1018
02_AG_FR_DJ263	-----C-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1210
02_AG_FR_DJ264	-----C-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1211
02_AG_GH_G829	-----C-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1078
02_AG_NG_IBNG	-----C-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1387
02_AG_SE_SE7812	-----C-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1234
02_AG_SN_98SE-MP1211	-----C-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1059
02_AG_SN_MP1213	-----C-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1056
03_AB_RU_KAL153-2	-----G-C-----A-----T-----A-----C-----G-----A-----T-----G-----T-----C-----G-----	1631
03_AB_RU_RU98001	-----G-C-----A-----T-----A-----C-----G-----A-----T-----G-----T-----C-----G-----	1232
04_cpx_CX_94CY032-3	-----G-----C-----C-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1222
04_cpx_GR_97PVGH	-----C-----T-----C-----C-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1873
04_cpx_GR_97PVNY	-----C-----T-----C-----C-----T-----C-----C-----A-----A-----G-----T-----C-----G-----	1870
05_DF_BE_V11310	-----C-----A-----A-----T-----A-----C-----C-----A-----A-----G-----T-----C-----G-----	1235
05_DF_BE_V1961	-----C-----A-----A-----T-----A-----C-----C-----A-----A-----G-----T-----C-----G-----	1211
CONSENSUS_06_cpx	-----C-----C-----?-----C-----C-----G-----A-----G-----A-----G-----T-----C-----G-----	1851
06_cpx_AU_BFP90	-----C-----C-----A-----A-----G-----A-----G-----A-----G-----T-----C-----G-----	1881
06_cpx_ML_95ML127	-----C-----C-----C-----C-----C-----A-----A-----G-----A-----G-----T-----C-----G-----	1873
06_cpx_ML_95ML84	-----C-----C-----C-----C-----C-----A-----A-----G-----A-----G-----T-----C-----G-----	1093
06_cpx_SN_97SEL078	-----C-----C-----A-----T-----A-----C-----C-----C-----A-----A-----G-----T-----C-----G-----	1900
07_BC_CN_97CN001	-----C-----C-----C-----C-----C-----G-----A-----G-----T-----A-----G-----T-----C-----G-----	1194
07_BC_CN_98CN009	-----C-----C-----C-----C-----C-----G-----A-----G-----T-----A-----G-----T-----C-----G-----	1194
07_BC_CN_C54A	-----C-----C-----C-----C-----C-----G-----A-----G-----T-----A-----G-----T-----C-----G-----	1052
CONSENSUS_08_BC	-----C-----C-----A-----A-----T-----A-----G-----T-----G-----T-----C-----G-----	1053
08_BC_CN_97CNGX6F	-----C-----C-----A-----A-----T-----A-----G-----T-----G-----T-----C-----G-----	1054
08_BC_CN_97CNGX7F	-----C-----C-----A-----A-----T-----A-----G-----T-----G-----T-----C-----G-----	1060
08_BC_CN_97CNGX9F	-----C-----C-----A-----A-----T-----A-----G-----T-----G-----T-----C-----G-----	1054
08_BC_CN_98CN006	-----C-----C-----A-----A-----T-----A-----G-----T-----G-----T-----C-----G-----	1206
10_CD_TZ_96TZBF061	-----C-----C-----A-----A-----T-----A-----G-----T-----G-----T-----C-----G-----	1235
10_CD_TZ_96TZBF071	-----C-----C-----A-----A-----T-----A-----G-----T-----G-----T-----C-----G-----	1233
10_CD_TZ_96TZBF110	-----C-----C-----A-----A-----T-----A-----G-----T-----G-----T-----C-----G-----	1233
CONSENSUS_11_cpx	-----?c-----?-----C-----?-----T-----C-----T-----C-----G-----A-----G-----T-----C-----G-----	1825
11_cpx_CM_MP818	-----GGA-----T-----AC-A-----T-----C-----T-----C-----G-----A-----G-----T-----C-----G-----	1862
11_cpx_FR_MP1298	-----C-----C-----AC-A-----T-----C-----T-----C-----G-----A-----G-----T-----C-----G-----	1859
11_cpx_FR_MP1307	-----GGC-----C-----C-----C-----C-----GG-G-----A-----A-----G-----T-----C-----G-----	1868
11_cpx_GR_GR17	-----G-C-----C-----C-----T-----C-----T-----C-----G-----A-----GT-----G-----A-----G-----T-----C-----G-----	1167
12_BF_AR_ARMA159	-----G-----C-----C-----A-----C-----A-----C-----T-----G-----T-----C-----G-----	1864
12_BF_UY_URTR23	-----GG-----C-----A-----C-----A-----C-----G-----T-----G-----T-----C-----G-----	1885
12_BF_UY_URTR35	-----GGG-----C-----A-----C-----A-----C-----G-----T-----G-----T-----C-----G-----	1864
N_CM_YBF106	-----G-----G-----AC-C-A-A-----A-----ACAG-CC-----A-----G-----GA-----CT-----G-----G-----G-----	1421
N_CM_YBF30	-----G-----G-----AC-C-A-A-----A-----ACAG-CC-----A-----G-----GA-----CT-----G-----G-----G-----	1420
CONSENSUS_0	-----C-----C-----T-----T-----C-----ACAG-----G-----t-----A-----g-----gA-----ct-----?	1819
O_CM_ANT70	-----TC-----C-----T-----T-----C-----ACAG-----G-----T-----A-----G-----GA-----CT-----G-----	1905
O_CM_MVP5180	-----C-----C-----T-----T-----G-----T-----A-----ACAA-----C-----G-----AA-----TT-----GT-----C-----A-----G-----	1880
O_SN_SEMP1299	-----C-----C-----T-----T-----C-----ACAG-----G-----T-----A-----G-----GA-----CT-----T-----T-----	1907
O_SN_SEMP1300	-----C-----C-----T-----T-----C-----ACAG-----G-----T-----A-----G-----GA-----CT-----T-----T-----	1907
CONSENSUS_CPZ	-----t-----c-----c-----?-----a-----t-----t-----a-----ac-----c-----g-----c-----Ga-----C-----Ct-----?	1410
CPZ_CD_CPFZANT	-----G-----GCC-----AA-----C-----C-----T-----T-----C-----ACAC-----CC-----G-----T-----A-----GA-----CT-----CT-----	1277
CPZ_CM_CAM3	-----T-----C-----C-----A-----A-----G-----T-----T-----C-----A-----AC-----C-----G-----CC-----GA-----CT-----C-----G-----	1246
CPZ_CM_CAM5	-----T-----C-----C-----T-----A-----T-----T-----C-----A-----C-----G-----C-----G-----CC-----GA-----CT-----C-----G-----	1549
CPZ_GA_CPFZGAB	-----G-----T-----C-----TC-----C-----A-----G-----A-----A-----C-----A-----G-----C-----G-----T-----G-----C-----T-----A-----G-----	1917
CPZ_US_CPFZUS	-----CC-----US-----G-----C-----T-----A-----A-----C-----T-----T-----A-----G-----C-----C-----G-----T-----GA-----C-----C-----G-----	1904
Gag p24	V_K_N_W_M_T_E_T_L_L_V_Q_N_A_N_P_D_C_K_T_I_L_K_A_L_G_P_A_A_T_L_E_F_E_M_M_T_A_C_O_G_V_G_G_P	Gag

HIV-1/SIVcpz complete genomes

Accession	Strain	Genomic Region	Sequence
B.FR.HXB2		Gag p24 Capsid end / Gag p2 start	CGGCCATAGCCAAAGAGTTTGGCTGAGCAATGAGCCCAAGTAACAATCA.....GCTACCAATATGATGAGAGGCAATTTT...AGGAACCAAGAAAGATTTTAAGTGTTC
CONSENSUS_01_AE		Gag p2 end / Nucleocapsid (NC) startat.....A-GG-G-a.....gAA.....c
01_AE.CF.90CF11697			T-AE.CF.90CF11697.....A-GG-G-a.....gAA.....c
01_AE.CF.90CF402			T-AE.CF.90CF402.....A-GG-G-a.....gAA.....c
01_AE.CF.90CF4071			T-AE.CF.90CF4071.....A-GG-G-a.....gAA.....c
01_AE.TH.93TH057			T-AE.TH.93TH057.....A-GG-G-a.....gAA.....c
01_AE.TH.93TH065			T-AE.TH.93TH065.....A-GG-G-a.....gAA.....c
01_AE.TH.93TH253			T-AE.TH.93TH253.....A-GG-G-a.....gAA.....c
01_AE.TH.95TH022			T-AE.TH.95TH022.....A-GG-G-a.....gAA.....c
01_AE.TH.95TH047			T-AE.TH.95TH047.....A-GG-G-a.....gAA.....c
01_AE.TH.CM240			T-AE.TH.CM240.....A-GG-G-a.....gAA.....c
CONSENSUS_02_AG		A-GG-G-a.....gAA.....c
02_AG.CM.97CM-MP807			02_AG.CM.97CM-MP807.....A-GG-G-a.....gAA.....c
02_AG.FR.DJ263			02_AG.FR.DJ263.....A-GG-G-a.....gAA.....c
02_AG.FR.DJ264			02_AG.FR.DJ264.....A-GG-G-a.....gAA.....c
02_AG.GH.G829			02_AG.GH.G829.....A-GG-G-a.....gAA.....c
02_AG.NG.IBNG			02_AG.NG.IBNG.....A-GG-G-a.....gAA.....c
02_AG.SE.SE7812			02_AG.SE.SE7812.....A-GG-G-a.....gAA.....c
02_AG.SN.98SE-MP1211			02_AG.SN.98SE-MP1211.....A-GG-G-a.....gAA.....c
02_AG.SN.MP1213			02_AG.SN.MP1213.....A-GG-G-a.....gAA.....c
03_AB.RU.KAL153-2			03_AB.RU.KAL153-2.....A-GG-G-a.....gAA.....c
03_AB.RU.RU98001			03_AB.RU.RU98001.....A-GG-G-a.....gAA.....c
04_CPX.CV.94CV032-3			04_CPX.CV.94CV032-3.....A-GG-G-a.....gAA.....c
04_CPX.GR.97PVCH			04_CPX.GR.97PVCH.....A-GG-G-a.....gAA.....c
04_CPX.GR.97PVNY			04_CPX.GR.97PVNY.....A-GG-G-a.....gAA.....c
05_DF.BE.VI1310			05_DF.BE.VI1310.....A-GG-G-a.....gAA.....c
05_DF.BE.VI1961			05_DF.BE.VI1961.....A-GG-G-a.....gAA.....c
CONSENSUS_06_CPX		A-GG-G-a.....gAA.....c
06_CPX.AU.BEP90			06_CPX.AU.BEP90.....A-GG-G-a.....gAA.....c
06_CPX.ML.95ML127			06_CPX.ML.95ML127.....A-GG-G-a.....gAA.....c
06_CPX.ML.95ML84			06_CPX.ML.95ML84.....A-GG-G-a.....gAA.....c
06_CPX.SN.97SE1078			06_CPX.SN.97SE1078.....A-GG-G-a.....gAA.....c
07_BC.CN.97CN001			07_BC.CN.97CN001.....A-GG-G-a.....gAA.....c
07_BC.CN.98CN009			07_BC.CN.98CN009.....A-GG-G-a.....gAA.....c
07_BC.CN.C54A			07_BC.CN.C54A.....A-GG-G-a.....gAA.....c
CONSENSUS_08_BC		A-GG-G-a.....gAA.....c
08_BC.CN.97CNGX6F			08_BC.CN.97CNGX6F.....A-GG-G-a.....gAA.....c
08_BC.CN.97CNGX7F			08_BC.CN.97CNGX7F.....A-GG-G-a.....gAA.....c
08_BC.CN.97CNGX9F			08_BC.CN.97CNGX9F.....A-GG-G-a.....gAA.....c
08_BC.CN.98CN006			08_BC.CN.98CN006.....A-GG-G-a.....gAA.....c
10_CD.TZ.96TZBF061			10_CD.TZ.96TZBF061.....A-GG-G-a.....gAA.....c
10_CD.TZ.96TZBF071			10_CD.TZ.96TZBF071.....A-GG-G-a.....gAA.....c
10_CD.TZ.96TZBF110			10_CD.TZ.96TZBF110.....A-GG-G-a.....gAA.....c
CONSENSUS_11_CPX		A-GG-G-a.....gAA.....c
11_CPX.CM.MP818			11_CPX.CM.MP818.....A-GG-G-a.....gAA.....c
11_CPX.FR.MP1298			11_CPX.FR.MP1298.....A-GG-G-a.....gAA.....c
11_CPX.FR.MP1307			11_CPX.FR.MP1307.....A-GG-G-a.....gAA.....c
11_CPX.GR.GR17			11_CPX.GR.GR17.....A-GG-G-a.....gAA.....c
12_BF.AR.ARMA159			12_BF.AR.ARMA159.....A-GG-G-a.....gAA.....c
12_BF.UY.URTR23			12_BF.UY.URTR23.....A-GG-G-a.....gAA.....c
12_BF.UY.URTR35			12_BF.UY.URTR35.....A-GG-G-a.....gAA.....c
N.CM.YBF106			N.CM.YBF106.....A-GG-G-a.....gAA.....c
N.CM.YBF30			N.CM.YBF30.....A-GG-G-a.....gAA.....c
CONSENSUS_0		A-GG-G-a.....gAA.....c
O.CM.ANT70			O.CM.ANT70.....A-GG-G-a.....gAA.....c
O.CM.MP5180			O.CM.MP5180.....A-GG-G-a.....gAA.....c
O.SN.SEMP1259			O.SN.SEMP1259.....A-GG-G-a.....gAA.....c
O.SN.SEMP1300			O.SN.SEMP1300.....A-GG-G-a.....gAA.....c
CONSENSUS_CPZ		A-GG-G-a.....gAA.....c
CPZ.CD.CPZANT			CPZ.CD.CPZANT.....A-GG-G-a.....gAA.....c
CPZ.CM.CAM3			CPZ.CM.CAM3.....A-GG-G-a.....gAA.....c
CPZ.CM.CAM5			CPZ.CM.CAM5.....A-GG-G-a.....gAA.....c
CPZ.GA.CPZGAB			CPZ.GA.CPZGAB.....A-GG-G-a.....gAA.....c
CPZ.US.CPZUS			CPZ.US.CPZUS.....A-GG-G-a.....gAA.....c
Gag p24		A-GG-G-a.....gAA.....c

HIV-1/SIVcpz complete genomes

B_FR_HXB2
 AATTGTGGCAAGAGAGGACACAGCCAGAAATTGAGGGCCCTAGGAAAAAGGGCTTTGGAAATGTGGAAGGAGGACACCAAAATGAAAGATTTGTACT.....GAGAGAGAG.....GCTAAATTTTTT
 2089
 CONSENSUS_01_AE
 01_AE.CF.90CF11697
 01_AE.CF.90CF402
 01_AE.CF.90CF4071
 01_AE.CF.90CF4071
 01_AE.TH.93TH057
 01_AE.TH.93TH065
 01_AE.TH.93TH253
 01_AE.TH.95TH022
 01_AE.TH.95TH047
 01_AE.TH.CM240
 1648
 CONSENSUS_02_AG
 02_AG.CM.97CM-MP807
 02_AG.FR.DJ263
 02_AG.FR.DJ264
 02_AG.GH.G829
 02_AG.NG.IENG
 02_AG.SE.SE7812
 02_AG.SN.98SE-MP1211
 02_AG.SN.MP1213
 03_AB.RU.KAL153-2
 03_AB.RU.RU98001
 04_CPX.CY.94CY032-3
 04_CPX.GR.97PVGH
 04_CPX.GR.97PVNY
 05_DF.BE.V11310
 05_DF.BE.V1961
 CONSENSUS_06_CPX
 06_CPX.AU.BFP90
 06_CPX.ML.95ML127
 06_CPX.ML.95ML84
 06_CPX.SN.97SE1078
 07_BC.CN.97CN001
 07_BC.CN.98CN009
 07_BC.CN.C54A
 CONSENSUS_08_BC
 08_BC.CN.97CNGX6F
 08_BC.CN.97CNGX7F
 08_BC.CN.97CNGX9F
 08_BC.CN.98CN006
 10_CD.TZ.96TZBF061
 10_CD.TZ.96TZBF071
 10_CD.TZ.96TZBF110
 CONSENSUS_11_CPX
 11_CPX.CM.MP818
 11_CPX.FR.MP1298
 11_CPX.FR.MP1307
 11_CPX.GR.GR17
 12_BF.AR.ARM159
 12_BF.UY.URTR23
 12_BF.UY.URTR35
 N_CM.YBF106
 N_CM.YBF30
 CONSENSUS_O
 O_CM.ANT70
 O_CM.MVP5180
 O_SN.SEMP1299
 O_SN.SEMP1300
 CONSENSUS_CPZ
 CPZ_CD.CPZANT
 CPZ_CM.CAM3
 CPZ_CM.CAM5
 CPZ_GA.CPZGAB
 CPZ_US.CPZUS
 1302
 1245
 1437
 1438
 1305
 1614
 1461
 1285
 1283
 1858
 1459
 1458
 2109
 2106
 1471
 1444
 2080
 2111
 2109
 1329
 2136
 1421
 1421
 1279
 1280
 1281
 1287
 1281
 1433
 1471
 1466
 1466
 2045
 2089
 2086
 2095
 1394
 2094
 2115
 2094
 1660
 1662
 2044
 2156
 2131
 2155
 2155
 1637
 1528
 1485
 1788
 2156
 2143
 /Gag-Pol TF
 TF
 TF
 Gag
 Gag p7 Nucleocapsid end / \ p1 start

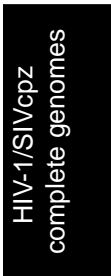


Table with 3 columns: Accession ID (e.g., B.FR.HXB2, CONSENSUS_A-A1), Nucleotide Sequence (e.g., AGAAGAGAGCTTCAGGTCTGGGTAGAGACAACAAC...), and Line Number (e.g., 2240, 1565, 1610, etc.).

Table with columns for accession numbers (e.g., B_FR_HXB2, CONSENSUS_01_AE), sequence identifiers (e.g., AGAAGAGAGCTTCAGGTCTGGGTAGACAAACAAC), and sequence alignments (e.g., --c---?a--59G--ATG--A---t---t---????????????). Includes Gag-Pol and Gag protein domains.

HIV-1/SIVcpz complete genomes

Table with columns for accession numbers, strain names, and nucleotide sequences. Includes labels like 'Gag-Pol TF end', 'Pol Protease start', and 'Gag end'.

HIV-1/SIVcpz complete genomes

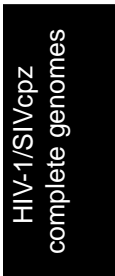
HIV-1/SIVcpz
complete genomes

B.FR.HXB2	GAGTTTCCAGGAAAGATGGAAACCAAAAATGATAGGGGAATTCGGAGGTTTTTATCAAAGTAAGACAGTATGATCAGATCTCATAGAAAATCTGTGGACATAAAGCTATAGGTACAGTATTAGTAGGACCT	2489
CONSENSUS_A-A1	A-A-----A-----A-----g-----t-----T-----R-----A-A-G-----	1814
A.BY.97BL006	A-A-----A-----G-----G-----G-----T-----R-----A-A-G-----R-----G-----	1859
A.SE.SE7535	A-A-----A-----G-----C-----G-----A-----G-----T-----A-----A-A-G-----G-----	1706
A.SE.SE8538	A-A-----A-----G-----C-----G-----A-----G-----T-----A-----A-A-G-----G-----	1688
A.SE.SE8891	A-A-----A-----G-----C-----G-----A-----G-----T-----A-----A-A-G-----G-----	1691
A.SE.UGSF8131	A-A-----A-----G-----C-----G-----A-----G-----T-----A-----A-A-G-----G-----	1861
A1.KE.Q23-17	A-A-----A-----G-----C-----G-----A-----G-----T-----A-----A-A-G-----G-----	1947
A1.SE.SE7253	A-A-----A-----G-----C-----G-----A-----G-----T-----A-----A-A-G-----G-----	1693
A1.UG.92UG037	A-A-----A-----G-----C-----G-----A-----G-----T-----A-----A-A-G-----G-----	1856
A1.UG.U455	A-A-----A-----A-----A-----A-----T-----T-----A-----A-A-GA-----G-----	1935
A2.CD.97CDKTB48	A-A-C-----T-G-A-----A-----A-----G-T-----T-----A-----A-A-GG-----C-----	1827
A2.CY.94CY017.41	A-A-----A-----A-----A-----A-----GCT-----T-----A-----A-A-GG-----C-----	1846
CONSENSUS_B	--a-----	2034
B.AU.MBC200	--A-----	2488
B.CN.RL42	--A-----	1876
B.DE.D31	--A-----C-----C-----C-----	2035
B.DE.HAN	--A-----	1955
B.ES.89SP061	--A-----A-----	2047
B.GA.OYI	--A-----A-----G-----G-----	2435
B.GB.MANC	--A-----G-----	2032
B.NL.3202A21	--A-----G-----C-----C-----	2491
B.TW.TWCYS	--A-----A-----A-----C-----	2488
B.US.JREL	--GA-----	1806
B.US.MNCG	--A-----A-----AC-----G-----	2504
B.US.RF	--A-----A-----G-----C-----CTG-----	2006
B.US.SF2	--A-----A-----C-----C-----	2496
B.US.YU2	--A-----	2487
CONSENSUS_C	a-a-----a-----a-----l-----T-----A-----A-G-----	1812
C.BR.92BR025	C-RA-----AT-----A-----A-----T-----T-----A-----A-G-----C-----	1827
C.BR.98BR004	A-AA-----AT-----A-----A-----T-----T-----A-----A-G-----C-----	1814
C.BW.96BM0502	A-A-----A-----A-----A-----G-T-----T-----A-----A-G-----C-----	1982
C.BW.96BM06	-----T-----A-----A-----CT-----T-----A-----A-G-----G-----A-----	2471
C.BW.96BM15C02	A-A-----A-----A-----A-----T-----T-----A-----A-G-----A-----	1956
C.ET.ETH2220	A-A-----A-----A-----A-----A-----T-----T-----A-----A-G-----C-----	1881
C.IL.98IS002	A-A-----T-----A-----A-----A-----TGT-----T-----GA-A-G-----C-----G-----C-----	1823
C.IN.93IN101	A-A-----A-----A-----A-----A-----CT-----T-----A-----A-G-----C-----	2462
C.IN.94IN476	AGC-----A-----A-----A-----T-----T-----A-----A-G-----C-----	1810
C.IN.98IN012	A-A-----A-----A-----A-----T-----T-----A-----A-G-----C-----	1814
C.TZ.98TZ013	A-A-----A-----G-----A-----A-----CT-----T-----A-----A-G-T-----	1832
C.TZ.98TZ017	A-A-----A-----A-----A-----A-----T-----T-----A-----A-G-----G-----	1843
C.ZA.97ZA012	AGA-----G-----G-----A-----A-----A-----CT-----T-----A-----A-G-----C-----	1823
C.ZM.96ZM651	A-A-----C-A-----A-----A-----A-----CT-G-----T-----A-----A-G-----	1818
C.ZM.96ZM751	A-A-----A-----A-----A-----A-----T-----T-----A-----A-G-----	1829
CONSENSUS_D	a-A-----A-----a-----	2030
D.CD.84ZR085	A-A-----AG-----C-----C-----G-----	2011
D.CD.FLI	--A-----A-----A-----C-----G-----	2035
D.CD.ZZ6	--A-----A-----A-----A-----G-----	2492
D.KE.MB2059	A-A-----A-----C-----A-----AG-----G-----	2491
D.UG.94UG1141	A-A-----A-----G-C-----A-----CT-----	1844
CONSENSUS_F1	A-A-----A-----A-----?-----C-----T-----T-----C-----?	1782
F1.BE.VI850	A-A-----A-----A-----C-----T-----T-----C-----G-----	1826
F1.BR.93BR020.1	A-A-----A-----A-----AGC-----T-----T-----C-----G-----	1817
F1.FI.FIN9363	A-A-----A-----A-----A-----C-----T-----T-----C-----G-----	1818
F1.FR.MF411	AGA-----A-----A-----A-----A-----AC-----T-----T-----C-----G-----	1691
F2.CM.MP255	A-A-----A-----G-A-----T-----T-----A-----A-G-----G-----	1687
F2.CM.MP257	A-A-----A-----A-----AG-TC-----T-----T-----A-----A-----G-----	1695
CONSENSUS_G	A-A-----a-----a-----T-----Ta-----A-A-G-----	1869
G.BE.DRCBL	AGA-----A-----A-----A-----T-----TA-----A-A-GG-----G-----	2453
G.FI.HH8793-12.1	A-AG-----A-----A-----A-----T-----TA-----A-A-G-----G-----	1890
G.NG.92NG083	A-A-----A-----A-----A-----T-----TG-----A-A-G-----G-----	1858
G.SE.SE6165	A-A-----A-----A-----A-----AG-CT-----TA-----A-A-G-----G-GA-----	1886
H.BE.VI991	A-A-----A-----A-----G-AG-GC-----T-----T-----A-A-G-----A-----	1878
H.BE.VI997	A-A-----T-----A-----AG-GC-----AG-GC-----A-----A-G-----G-----	1813
H.CF.90CF056	A-A-----G-----A-----G-AG-GC-----G-----AG-GC-----A-A-G-----	1836
J.SE.SE7022	AGAC-----A-----A-----G-----A-G-G-----TGAG-----A-A-G-----A-----	1804
J.SE.SE7887	AGAC-----C-----A-----A-----CG-G-----CG-----TGAG-----A-A-G-----G-----A-----	1803
K.CD.EQTB11C	A-A-----A-----G-----A-----AG-TGT-----G-----T-----G-----A-----G-----	1685
K.CM.MP535	A-A-----A-----A-----AG-----T-----T-----A-----A-----G-----	1685

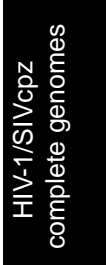
HIV-1/SIVcpz complete genomes

	Protease end / Pol p66 and p51 RT start	
B.FR.HXB2	ACACCTGTCAACATAATGGAAAGAAATCTGTTGACTCAGATGGTTGACCTTTAAATTTCCCATTTAGCCCTATTGAGACTGTACCAGTAGTAATAAATTAAGCCAGGAATGGATGCCCAAAAGTTAAACAAT	2619
CONSENSUS_A-A1	-----A-----T-----C--A--T-----T-----g-----	1944
A.BY.97BL006	-----R-----T-----C--A--T-----A-----T-----G-----	1989
A.SE.SE7535	-----A-----C-----T-----C--A--T-----A-----T-----G-----	1836
A.SE.SE8538	-----A-----C-----T-----C--A--T-----A-----T-----G-----	1818
A.SE.SE8891	-----A-----C-----T-----C--A--T-----A-----T-----G-----	1821
A.SE.UGSE8131	-----A-----C-----T-----C--A--T-----A-----T-----G-----	1991
A1.KE.O23-17	CA-----T-----CA-----T-----C--A--T-----A-----T-----G-----	2077
A1.SE.SE7253	-----A-----C-----T-----C--A--T-----A-----T-----G-----	1823
A1.UG.92UG037	-----G--A--C--T-----T-----A-----A-----A-----A-----GGA-----	1986
A1.UG.U455	-----G--A--C--T-----T-----A-----A-----A-----A-----G-----	2065
A2.CD.97CDKTB48	-----A-----GT-----C-----T-----C--A--T-----A-----T-----G--C-----	1957
A2.CY.94CY017.41	-----A-----GT-----C-----T-----C--A--T-----A-----T-----T-----G-----	1976
CONSENSUS_B	-----T-----a-----	2164
B.AU.MBC200	-----T-----A-----A-----T-----A-----A-----A-----	2618
B.CN.RL42	-----T-----A-----A-----T-----A-----A-----A-----	2006
B.DE.D31	-----T-----A-----A-----T-----A-----A-----A-----	2165
B.DE.HAN	-----T-----A-----A-----T-----A-----A-----A-----	2085
B.ES.89SP061	-----T-----A-----A-----T-----A-----A-----A-----	2177
B.GA.OYI	-----T-----A-----A-----T-----A-----A-----A-----	2162
B.GB.MANC	-----T-----A-----A-----T-----A-----A-----A-----	2565
B.NL.3202A21	-----T-----A-----A-----T-----A-----A-----A-----	2621
B.TM.TWCYS	-----T-----T-----T-----T-----T-----T-----	2618
B.US.JREL	-----T-----T-----T-----T-----T-----T-----	1936
B.US.MNCG	-----T-----T-----T-----T-----T-----T-----	2634
B.US.RF	-----T-----T-----T-----T-----T-----T-----	2136
B.US.SF2	-----T-----T-----T-----T-----T-----T-----	2626
B.US.YU2	-----T-----T-----T-----T-----T-----T-----	2617
CONSENSUS_C	-----A-----C-----A-----AC-----A-----T-----C--A-----A-----G-----	1942
C.BR.92BR025	-----CA-----C--G--A-----AC-----A-----T-----C--A-----A-----G--C-----	1957
C.BR.98BR004	-----T-----A-----C-----A-----A-----T-----C--A-----A-----G--C-----	1944
C.BW.96BW0502	-----T-----A-----C-----A-----A-----T-----C--A-----A-----G--C-----	2112
C.BW.96BW15C02	-----G--A-----C--A-----AC-----A-----T-----C--A-----A-----G-----	2601
C.ET.ETH2220	-----C--CA-----C--AC--A-----C-----A-----T-----C--A-----A-----G-----	2086
C.IL.98ILS002	-----A--A-----C--A-----T--AC-----A-----T-----C--A-----A-----T-----	2011
C.IN.93IN101	-----A-----A-----C--A-----AC-----A-----C-----T-----C--A-----A-----G-----	1953
C.IN.94IN476	-----G--A-----C--A-----C-----A-----A-----C-----A-----A-----G-----	2592
C.IN.98IN012	-----A-----A-----C-----A-----GC-----A-----T-----C--A-----A-----G-----	1940
C.TZ.98TZ013	-----A-----A-----C-----A-----AC-----A-----T-----C--A-----A-----G-----	1944
C.TZ.98TZ017	-----A-----A-----C-----A-----AC-----A-----T-----C--C--A-----T-----GG-----	1962
C.ZA.97ZA012	-----G--A-----C--A-----AC-----A-----T-----C--C--A-----T-----GG-----	1973
C.ZM.96ZM651	-----G--A-----C--A-----AC-----A-----T-----C--C--A-----T-----G-----	1953
C.ZM.96ZM751	-----G--A-----C--C-----AC-----A-----T-----A-----A-----T-----GG--C-----	1948
CONSENSUS_D	-----T-----C-----C-----C-----T-----A-----A-----T-----G-----	1959
D.CD.84ZR085	-----T-----C-----C-----C-----C-----T-----A-----A-----T-----G-----	2160
D.CD.FLI	-----T-----C-----C-----C-----C-----T-----A-----A-----T-----G-----	2141
D.CD.Z2Z6	-----T-----C-----C-----C-----C-----T-----A-----A-----T-----G-----	2165
D.KE.MB2059	-----T-----C-----C-----C-----C-----T-----A-----A-----T-----G-----	2622
D.UG.94UG1141	-----T-----C-----C-----C-----C-----T-----A-----A-----T-----G-----	2621
CONSENSUS_F1	-----T-----A-----A-----T-----A-----A-----T-----A-----G-----	1974
F1.BE.VI850	-----G--A-----A-----A-----A-----T-----A-----A-----T-----G-----	1912
F1.BR.93BR020.1	-----A-----A-----C-----T-----C-----A-----A-----T-----A-----G-----	1956
F1.FI.FIN9363	-----G--A-----A-----A-----A-----T-----A-----A-----T-----G-----	1947
F1.FR.MP411	-----G--A-----A-----A-----A-----T-----A-----A-----T-----G-----	1948
F2.CM.MP255	-----G--A-----A-----A-----A-----T-----A-----A-----T-----G-----	1821
F2.CM.MP257	-----G--A-----A-----A-----A-----T-----A-----A-----T-----G-----	1815
CONSENSUS_G	-----A-----g-----A-----A-----A-----T-----A-----A-----T-----G-----	1827
G.BE.DRCBL	-----A-----T-----G-----A-----A-----A-----T-----A-----A-----T-----G-----	1998
G.FI.HH8793-12.1	-----A-----T-----G-----A-----A-----A-----T-----A-----A-----T-----G-----	2583
G.NG.92NG083	-----A-----T-----G-----A-----A-----A-----T-----A-----A-----T-----G-----	2020
G.SE.SE6165	-----A-----T-----G-----A-----A-----A-----T-----A-----A-----T-----G-----	1988
H.BE.VI991	-----T-----G--A--A-----A--G-----C-----T-----A-----A-----T-----G-----	2016
H.BE.VI997	-----T-----G--A--A-----A--A-----C-----T-----A-----A-----T-----G-----	2008
H.CF.90CF056	-----T-----G--A--A-----A--A-----C-----T-----A-----A-----T-----G-----	1943
J.SE.SE7022	-----CA-----CA-----C-----T-----T-----A-----A-----T-----A-----A-----	1966
J.SE.SE7887	-----CA-----CA-----C-----T-----T-----A-----A-----T-----A-----A-----	1934
K.CD.EOTB11C	-----CA-----CA-----C-----T-----T-----A-----A-----T-----A-----A-----	1933
K.CM.MP535	-----C-----C-----C-----T-----T-----A-----A-----T-----A-----A-----	1815

B_FR_HXB2	ACACCTGTCAACATAATGGAAAGAAATCTGTTGACTCAGATGGTTGACACTTAAATTTCCCATTTAGCCCTATTGAGACTGTACCCAGTAGTAATAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAAT	2619
CONSENSUS_01_AE	-----C-----A-----T-----C-----A-----t-----C-----c-----A-----A-----g-----g-----	2570
01_AE_CF_90CF11697	-----C-----A-----T-----T-----C-----A-----A-----T-----C-----C-----C-----A-----A-----G-----g-----	2568
01_AE_CF_90CF402	-----C-----G-----C-----A-----C-----A-----T-----C-----C-----C-----A-----A-----G-----g-----	2655
01_AE_CF_90CF4071	-----T-----C-----A-----C-----A-----T-----T-----C-----C-----C-----A-----A-----G-----g-----	2556
01_AE_TH_93TH057	-----G-----C-----A-----C-----A-----T-----T-----C-----C-----C-----A-----A-----G-----g-----	2572
01_AE_TH_93TH065	-----G-----C-----A-----C-----A-----T-----T-----C-----C-----C-----A-----A-----G-----g-----	2564
01_AE_TH_93TH253	-----C-----C-----A-----C-----A-----T-----T-----C-----C-----C-----A-----A-----G-----g-----	2606
01_AE_TH_95TH022	-----C-----C-----A-----C-----A-----T-----T-----C-----C-----C-----A-----A-----G-----g-----	2582
01_AE_TH_95TH047	-----C-----C-----A-----C-----A-----T-----T-----C-----C-----C-----A-----A-----G-----g-----	2582
01_AE_TH_CM240	-----C-----C-----A-----C-----A-----T-----T-----C-----C-----C-----A-----A-----G-----g-----	2193
CONSENSUS_02_AG	-----C-----A-----T-----T-----A-----A-----T-----A-----A-----g-----g-----	1831
02_AG_CM_97CM-MP807	-----C-----A-----T-----T-----A-----A-----T-----A-----A-----g-----g-----	1775
02_AG_FR_DJ263	-----G-----C-----A-----A-----T-----T-----A-----A-----C-----G-----G-----g-----	1967
02_AG_FR_DJ264	-----T-----C-----A-----A-----T-----T-----A-----A-----C-----G-----G-----g-----	1968
02_AG_GH_G829	-----C-----C-----A-----A-----T-----T-----A-----A-----G-----G-----G-----g-----	1835
02_AG_NG_IBNG	-----C-----C-----A-----A-----T-----T-----A-----A-----G-----G-----G-----g-----	2144
02_AG_SE_SE7812	-----C-----G-----A-----A-----T-----T-----A-----A-----C-----C-----C-----g-----	1991
02_AG_SN_96SE-MP1211	-----A-----A-----T-----T-----A-----A-----T-----A-----G-----T-----g-----	1816
02_AG_SN_MP1213	-----CA-----A-----T-----T-----C-----CA-----T-----A-----A-----G-----T-----g-----	1813
03_AB_RU_KAL153-2	-----C-----A-----T-----T-----A-----A-----T-----A-----A-----C-----C-----G-----g-----	2388
03_AB_RU_RU98001	-----C-----A-----T-----T-----A-----A-----T-----A-----A-----C-----C-----G-----g-----	1989
04_cpx_CV_94CV032-3	-----C-----CA-----C-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	1985
04_cpx_GR_97PVGH	-----C-----C-----CA-----C-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	2636
04_cpx_GR_97PVNY	-----C-----C-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	2633
05_DF_BE_VI1310	-----T-----T-----C-----C-----A-----A-----T-----A-----A-----C-----C-----G-----g-----	2001
05_DF_BE_VI1961	-----T-----T-----C-----C-----A-----A-----T-----A-----A-----C-----C-----G-----g-----	1974
CONSENSUS_06_cpx	-----C-----CA-----C-----C-----A-----A-----T-----T-----A-----A-----C-----C-----G-----g-----	2600
06_cpx_AU_BFP90	-----A-----C-----CA-----C-----C-----A-----A-----T-----T-----A-----A-----C-----C-----G-----g-----	2647
06_cpx_ML_95ML127	-----C-----C-----CA-----C-----C-----A-----A-----T-----T-----A-----A-----C-----C-----G-----g-----	2636
06_cpx_ML_95ML84	-----C-----C-----CA-----C-----C-----A-----A-----T-----T-----A-----A-----C-----C-----G-----g-----	1862
06_cpx_SN_97SE1078	-----C-----C-----CA-----C-----C-----A-----A-----T-----T-----A-----A-----C-----C-----G-----g-----	2678
07_BC_CN_97CN001	-----C-----C-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----C-----G-----g-----	1951
07_BC_CN_98CN009	-----C-----C-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----C-----G-----g-----	1951
07_BC_CN_C54A	-----C-----C-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----C-----G-----g-----	1809
CONSENSUS_08_BC	-----a-----A-----C-----C-----A-----A-----T-----T-----A-----A-----C-----C-----G-----g-----	1798
08_BC_CN_97CNGX6F	-----A-----A-----C-----C-----A-----A-----T-----T-----A-----A-----C-----C-----G-----g-----	1799
08_BC_CN_97CNGX7F	-----A-----A-----C-----C-----A-----A-----T-----T-----A-----A-----C-----C-----G-----g-----	1805
08_BC_CN_97CNGX9F	-----A-----A-----C-----C-----A-----A-----T-----T-----A-----A-----C-----C-----G-----g-----	1799
08_BC_CN_98CN006	-----A-----A-----C-----C-----A-----A-----T-----T-----A-----A-----C-----C-----G-----g-----	1951
10_CD_TZ_96TZBF061	-----G-----G-----T-----T-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	1998
10_CD_TZ_96TZBF071	-----T-----T-----T-----T-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	1996
10_CD_TZ_96TZBF110	-----C-----T-----T-----T-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	1987
CONSENSUS_11_cpx	-----CA-----A-----C-----t-----C-----t-----A-----T-----C-----C-----C-----G-----g-----	2561
11_cpx_CM_MP818	-----G-----CA-----A-----C-----t-----C-----t-----A-----T-----C-----C-----C-----G-----g-----	2622
11_cpx_FR_MP1298	-----C-----G-----CA-----A-----C-----t-----C-----t-----A-----T-----C-----C-----C-----G-----g-----	2613
11_cpx_FR_MP1307	-----C-----CA-----A-----C-----t-----C-----t-----A-----T-----C-----C-----C-----G-----g-----	2670
11_cpx_GR_GR17	-----G-----CA-----A-----C-----t-----C-----t-----A-----T-----G-----C-----A-----G-----g-----	1921
12_BF_AR_ARM159	-----G-----G-----T-----T-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	2624
12_BF_UY_URTR23	-----T-----T-----T-----T-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	2645
12_BF_UY_URTR35	-----T-----T-----T-----T-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	2642
N_CM_YBF106	-----T-----T-----T-----T-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	2211
N_CM_YBF30	-----T-----T-----T-----T-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	2213
CONSENSUS_0	-----T-----T-----T-----T-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	2538
O_CM_ANI70	-----T-----T-----T-----T-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	2674
O_CM_MVP5180	-----T-----T-----T-----T-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	2649
O_SN_SEMP1299	-----T-----T-----T-----T-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	2673
O_SN_SEMP1300	-----T-----T-----T-----T-----A-----A-----T-----T-----A-----A-----T-----A-----A-----C-----G-----g-----	2673
CONSENSUS_CPZ	-----a-----a-----t-----t-----c-----c-----a-----a-----a-----a-----t-----t-----a-----a-----g-----g-----	2135
CPZ_CD_CFPZANT	-----AC-----A-----A-----T-----A-----T-----G-----T-----ATG-----T-----TT-----A-----A-----T-----A-----A-----T-----A-----A-----g-----g-----	2058
CPZ_CM_CAM3	-----A-----A-----G-----T-----T-----C-----A-----A-----A-----A-----A-----C-----T-----A-----A-----G-----C-----T-----A-----A-----g-----g-----	2030
CPZ_CM_CAM5	-----A-----A-----G-----T-----T-----C-----A-----A-----A-----A-----A-----C-----T-----A-----A-----G-----C-----T-----A-----A-----g-----g-----	2321
CPZ_GA_CFPZGAB	-----A-----A-----T-----T-----C-----G-----T-----G-----C-----A-----A-----A-----A-----C-----C-----G-----A-----T-----T-----G-----A-----g-----	2680
CPZ_US_CFPZUS	-----A-----A-----T-----T-----C-----G-----A-----T-----A-----A-----A-----A-----T-----T-----C-----T-----A-----A-----T-----T-----G-----A-----g-----	2679
Pol Protease	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_P_V_K_L_K_P_G_M_D_G_P_K_V_K_Q	Pol



B_FR_HXB2	GGCCATTGACAGAAAGAAAAATAAAGCAATTAGTAGAAAATTTGTACAGAGATGGAAAGGAAAGGAAAAATTTCAAAAAATTTGGCCTGAAAAATCCATACATACTCCAGTATTTGCCATFAAAGAAAAAGA	2749
CONSENSUS_01_AE	-----AC-----A-----G-----A-----c-----t-----g-----	2700
01_AE_CF_90CF11697	---A---T---AC---A---G---A---A---A---C---A---C---G---A---C---T---T---G---G---	2698
01_AE_CF_90CF402	---A---T---AC---A---G---A---A---A---C---A---C---G---A---C---T---T---G---G---	2785
01_AE_CF_90CF4071	---A---T---AC---A---G---A---A---A---C---A---C---G---A---C---T---T---G---G---	2686
01_AE_TH_93TH057	---A---T---AC---A---G---A---A---A---C---A---C---G---A---C---T---T---G---G---	2702
01_AE_TH_93TH065	---A---T---AC---A---G---A---A---A---C---A---C---G---A---C---T---T---G---G---	2694
01_AE_TH_93TH253	---A---T---AC---A---G---A---A---A---C---A---C---G---A---C---T---T---G---G---	2736
01_AE_TH_95TH022	---A---T---AC---A---G---A---A---A---C---A---C---G---A---C---T---T---G---G---	2712
01_AE_TH_95TH047	---A---T---AC---A---G---A---A---A---C---A---C---G---A---C---T---T---G---G---	2723
01_AE_TH_CM240	---A---T---AC---A---G---A---A---A---C---A---C---G---A---C---T---T---G---G---	1961
CONSENSUS_02_AG	-----AC-----C-----T-----T-----A-----	1905
02_AG_CM_97CM-MP807	---G---G---AC---C---T---T---A-----	2097
02_AG_FR_DJ263	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2098
02_AG_FR_DJ264	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	1965
02_AG_GH_G829	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2274
02_AG_NG_IBNG	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2121
02_AG_SE_SE7812	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	1945
02_AG_SN_98SE-MP1211	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2518
02_AG_SN_MP1213	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2119
03_AB_RU_KAL153-2	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2115
03_AB_RU_RU98001	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2766
04_cpx_CX_94CY032-3	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2763
04_cpx_GR_97PVGH	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2131
04_cpx_GR_97PVNY	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2104
05_DF_BE_VI1310	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2730
05_DF_BE_VI1961	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2777
CONSENSUS_06_cpx	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2766
06_cpx_AU_BFP90	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	1992
06_cpx_ML_95ML127	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2808
06_cpx_ML_95ML84	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2081
06_cpx_SN_97SEL078	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2081
07_BC_CN_97CN001	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2081
07_BC_CN_98CN009	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2081
07_BC_CN_C54A	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	1939
CONSENSUS_08_BC	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	1928
08_BC_CN_97CNGX6F	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	1929
08_BC_CN_97CNGX7F	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	1935
08_BC_CN_97CNGX9F	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	1929
08_BC_CN_98CN006	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2081
10_CD_TZ_96TZBF061	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2128
10_CD_TZ_96TZBF071	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2126
10_CD_TZ_96TZBF110	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2117
CONSENSUS_11_cpx	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2690
11_cpx_CM_MP818	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2752
11_cpx_FR_MP1298	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2743
11_cpx_FR_MP1307	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2800
11_cpx_GR_GR17	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2051
12_BF_AR_ARMA159	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2754
12_BF_UY_URTR23	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2775
12_BF_UY_URTR35	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2772
N_CM_YBF106	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2341
N_CM_YBF30	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2343
CONSENSUS_0	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2668
O_CM_ANT70	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2804
O_CM_MVP5180	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2779
O_SN_SEMP1299	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2803
O_SN_SEMP1300	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2803
CONSENSUS_CPZ	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2264
CPZ_CD_CFPZANT	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2188
CPZ_CM_CAM3	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2160
CPZ_CM_CAM5	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2451
CPZ_GA_CFPZGAB	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2810
CPZ_US_CFPZUS	---G---G---AC---C---C-----G-----A-----G-----G-----G-----	2809
Po1_P51_RT	W_P_L_T_E_E_K_I_K_A_L_V_E_I_C_T_E_M_E_K_E_G_K_I_S_K_I_G_P_E_N_P_Y_N_T_P_V_F_A_I_K_K_K_D	856



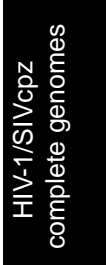
B.FR.HXB2	3008	TGTGGGTGATGCATATTTTTTCAGTTCCCTTAGATGAAGACCTTCAGGAAGTACTGCAATTTACCATACCTAGTATTAACAATGAGACACCGGGATTTAGATATCAATGTAATAATGTCITTCACAGGGATGG
CONSENSUS_A-A1	3333	-----g-c-c-----t-----t-----g-----t-----a-----c-----a-----c-----g-----
A.BY.97BL006	2378	-----A-----T-----A-----T-----A-----N-----N-----C-----T-----A-----G-----R-----A-----C-----g-----
A.SE.SE7535	2225	-----T-----G-----C-----C-----T-----C-----G-----T-----A-----A-----C-----G-----C-----A-----C-----G-----A-----
A.SE.SE8538	2207	-----G-----C-----G-----G-----G-----G-----C-----C-----C-----C-----C-----C-----C-----A-----C-----G-----
A.SE.SE8891	2197	-----A-----G-----C-----T-----G-----A-----T-----A-----T-----A-----C-----G-----C-----A-----C-----G-----
A.SE.UGSEF8131	2380	-----A-----G-----C-----T-----C-----G-----T-----A-----A-----C-----G-----C-----A-----C-----G-----
A1.KE.Q23-17	2466	-----G-----C-----C-----T-----C-----G-----T-----A-----A-----C-----C-----C-----AG-C-G-----AG-C-G-----
A1.SE.SE7253	2212	-----G-----C-----C-----G-----C-----C-----AG-T-A-----A-----C-----G-----C-----A-----C-----G-----
A1.UG.92UG037	2375	-----G-----C-----C-----T-----AG-T-A-----A-----A-----C-----G-----C-----A-----C-----G-----
A1.UG.U455	2454	-----G-----C-----C-----T-----AG-T-A-----A-----A-----G-----C-----G-----AG-C-G-----G-----
A2.CD.97CDKTB48	2346	-----G-----C-----C-----C-----C-----C-----A-----A-----A-----A-----A-----G-----C-----G-----A-----
A2.CY.94CY017.41	2365	-----G-----C-----C-----C-----C-----C-----A-----A-----A-----A-----A-----AG-----G-----
CONSENSUS_B	2553	-----a-----a-----a-----a-----a-----a-----a-----a-----a-----a-----a-----a-----
B.AU.MBC200	3007	-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
B.CN.RL42	2395	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
B.DE.D31	2554	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
B.DE.HAN	2474	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
B.ES.89SP061	2566	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
B.GA.OYI	2551	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
B.GB.MANC	2954	-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----
B.NL.3202A21	3010	-----A-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
B.TW.TWCYS	3007	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
B.US.JREL	3007	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
B.US.MNCG	2325	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
B.US.RF	3023	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
B.US.SF2	2525	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
B.US.YU2	3006	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
CONSENSUS_C	2331	-----g-----T-----g-----T-----a-----a-----c-----c-----c-----c-----c-----c-----
C.BR.92BR025	2346	-----A-----T-----G-----T-----A-----A-----A-----A-----A-----A-----A-----A-----
C.BR.98BR004	2333	-----A-----T-----G-----T-----A-----A-----A-----A-----A-----A-----A-----A-----
C.BW.96BW0502	2501	-----A-----G-----G-----G-----T-----A-----A-----A-----A-----A-----A-----A-----A-----
C.BW.96BW06	2990	-----G-----T-----G-----G-----T-----A-----A-----A-----A-----A-----A-----A-----A-----
C.BW.96BW15C02	2475	-----G-----T-----G-----G-----T-----A-----A-----A-----A-----A-----A-----A-----A-----
C.ET.ETH2220	2400	-----G-----C-----C-----C-----GT-----A-----A-----A-----A-----A-----A-----A-----A-----
C.IL.98ILS002	2342	-----A-----G-----C-----C-----GAG-----A-----A-----A-----A-----A-----A-----A-----A-----
C.IN.93IN101	2981	-----G-----T-----T-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----
C.IN.94IN476	2329	-----G-----T-----G-----G-----A-----A-----A-----A-----A-----A-----A-----A-----
C.IN.98IN012	2333	-----G-----T-----G-----G-----A-----A-----A-----A-----A-----A-----A-----A-----
C.TZ.98TZ013	2351	-----G-----T-----AG-T-----G-----C-----G-----A-----A-----A-----A-----A-----
C.TZ.98TZ017	2351	-----G-----T-----G-----G-----A-----A-----A-----A-----A-----A-----A-----A-----
C.ZA.97ZA012	2342	-----A-----C-----T-----A-----T-----AC-----C-----T-----C-----T-----A-----A-----
C.ZM.96ZM651	2337	-----G-----T-----AG-----AG-----A-----A-----A-----A-----A-----A-----A-----A-----
C.ZM.96ZM751	2348	-----G-----G-----T-----I-----G-----G-----A-----A-----A-----A-----A-----A-----A-----
CONSENSUS_D	2548	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----
D.CD.84ZR085	2530	-----A-----A-----TG-----T-----T-----T-----T-----C-----C-----C-----C-----C-----
D.CD.FLI	2554	-----T-----A-----T-----T-----A-----A-----C-----C-----C-----C-----C-----
D.CD.Z2Z6	3011	-----A-----A-----T-----A-----T-----A-----A-----C-----C-----C-----C-----C-----
D.KE.MB2059	3011	-----T-----C-----C-----A-----T-----A-----A-----C-----C-----C-----C-----C-----
D.UG.94UG1141	2363	-----C-----C-----C-----C-----T-----A-----A-----A-----A-----A-----A-----A-----
CONSENSUS_F1	2299	-----G-----A-----g-----t-----C-----C-----C-----g-----C-----A-----g-----C-----A-----
F1.BE.VI850	2345	-----G-----A-----G-----T-----A-----C-----C-----C-----A-----G-----T-----A-----G-----C-----A-----
F1.BR.93BR020.1	2336	-----G-----A-----G-----T-----A-----C-----C-----C-----AG-----G-----C-----A-----A-----
F1.FI.FIN9363	2337	-----G-----A-----A-----T-----A-----C-----C-----C-----C-----C-----A-----G-----C-----A-----
F1.FR.MF411	2210	-----G-----A-----G-----G-----A-----G-----G-----C-----C-----C-----A-----C-----C-----A-----G-----
F2.CM.MP255	2204	-----G-----G-----A-----G-----G-----A-----C-----C-----G-----C-----C-----A-----A-----
F2.CM.MP257	2216	-----G-----G-----A-----G-----G-----A-----C-----C-----G-----C-----C-----A-----A-----
CONSENSUS_G	2381	-----g-----?-----?-----T-----A-----A-----A-----A-----A-----A-----A-----A-----
G.BE.DRCBL	2965	-----T-----A-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
G.FI.HH8793-12.1	2409	-----G-----C-----AGT-----T-----A-----A-----A-----A-----A-----A-----A-----A-----
G.NG.92NG083	2377	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
G.SE.SE6165	2405	-----G-----A-----A-----T-----A-----A-----T-----A-----A-----T-----A-----A-----T-----
H.BE.VI991	2397	-----G-----G-----C-----T-----C-----C-----C-----C-----C-----C-----C-----C-----
H.BE.VI997	2332	-----C-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
H.CF.90CF056	2355	-----G-----C-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
J.SE.SE7022	2323	-----C-----C-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----
J.SE.SE7887	2322	-----C-----C-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----
K.CD.EQTB11C	2204	-----A-----G-----C-----T-----A-----T-----T-----T-----T-----T-----T-----T-----T-----
K.CM.MP535	2204	-----G-----G-----C-----T-----A-----A-----A-----A-----A-----A-----A-----A-----

B_FR_HXB2	TGTGGTGCATATTTTTCAGTTCCTTAGATGAAGACCTTCAGGAAGTACTGCAATTTACCATACCTAGTATATAACAATGAGACACCGGGATTAGATATCAAGTACAATGTCTCCACAGGATGG	3008
CONSENSUS_01_AE	-----a-----T-----T-----AG-----t-----A-----C-----C-----t-----A-----C-----g-----	2959
01_AE_CF_90CF11697	-----A-----T-----T-----G-----T-----A-----A-----A-----C-----A-----C-----A-----C-----g-----	2957
01_AE_CF_90CF402	-----A-----C-----T-----T-----C-----G-----T-----A-----A-----AG-----C-----AG-----C-----G-----G-----	3044
01_AE_CF_90CF4071	-----A-----T-----C-----T-----C-----AG-----T-----A-----A-----A-----C-----G-----A-----C-----G-----	2944
01_AE_TH_93TH057	-----G-----T-----C-----AG-----T-----A-----A-----G-----A-----C-----G-----A-----C-----T-----	2961
01_AE_TH_93TH065	-----A-----T-----C-----AG-----T-----A-----A-----C-----A-----C-----G-----A-----C-----G-----	2953
01_AE_TH_93TH253	-----A-----T-----T-----AG-----T-----A-----A-----C-----A-----C-----G-----A-----C-----G-----	2995
01_AE_TH_95TH022	-----A-----T-----GAG-----T-----A-----A-----C-----A-----C-----G-----A-----C-----G-----	2971
01_AE_TH_95TH047	-----A-----T-----AG-----T-----A-----A-----A-----C-----A-----C-----G-----A-----C-----G-----	2971
01_AE_TH_CM240	-----A-----T-----AG-----T-----A-----A-----C-----A-----C-----G-----A-----C-----G-----	2582
CONSENSUS_02_AG	-----g-----T-----T-----A-----A-----c-----t-----?-----t-----T-----G-----C-----	2219
02_AG_CM_97CM-MP807	-----G-----G-----T-----T-----A-----A-----c-----T-----T-----A-----C-----C-----C-----	2164
02_AG_FR_DJ263	-----C-----C-----A-----A-----T-----T-----A-----A-----C-----T-----C-----T-----C-----	2356
02_AG_FR_DJ264	-----C-----C-----A-----A-----T-----T-----A-----A-----C-----T-----G-----T-----C-----	2357
02_AG_GH_G829	-----G-----G-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----A-----	2224
02_AG_NG_IBNG	-----G-----G-----A-----A-----T-----T-----A-----A-----T-----T-----A-----A-----C-----	2533
02_AG_SE_SE7812	-----A-----G-----A-----A-----T-----T-----A-----A-----C-----T-----T-----G-----T-----	2380
02_AG_SN_98SE-MP1211	-----G-----G-----C-----C-----A-----A-----T-----T-----A-----A-----C-----T-----T-----G-----	2205
02_AG_SN_MF1213	-----G-----C-----C-----C-----A-----A-----T-----T-----A-----A-----C-----C-----T-----	2202
03_AB_RU_KAL153-2	-----C-----C-----C-----C-----A-----A-----A-----A-----C-----C-----C-----C-----	2777
03_AB_RU_RU98001	-----C-----C-----C-----C-----A-----A-----A-----A-----C-----C-----C-----C-----	2378
04_cpx_CX_94CY032-3	-----G-----G-----CC-----G-----G-----C-----C-----C-----C-----CC-----C-----A-----C-----	2374
04_cpx_GR_97PVGH	-----G-----G-----C-----C-----CG-----G-----C-----C-----C-----C-----CC-----C-----AG-----	3025
04_cpx_GR_97PVNY	-----G-----G-----C-----C-----CG-----G-----C-----C-----C-----C-----CC-----C-----G-----	3022
05_DF_BE_VI1310	-----G-----G-----A-----G-----A-----A-----C-----C-----C-----C-----C-----C-----A-----A-----	2390
05_DF_BE_VI1961	-----A-----C-----G-----G-----A-----A-----C-----C-----C-----C-----A-----A-----C-----C-----	2363
CONSENSUS_06_cpx	-----G-----G-----?-----?-----T-----a-----C-----T-----?-----T-----A-----A-----	2985
06_cpx_AU_BFP90	-----G-----G-----A-----C-----A-----A-----T-----A-----A-----C-----T-----A-----A-----	3036
06_cpx_ML_95ML127	-----G-----G-----A-----A-----A-----A-----T-----T-----A-----A-----C-----T-----A-----A-----	3025
06_cpx_ML_95ML84	-----G-----G-----GC-----A-----A-----T-----A-----A-----C-----T-----T-----T-----A-----	2251
06_cpx_SN_97SE1078	-----A-----G-----C-----C-----A-----A-----T-----T-----A-----A-----C-----T-----C-----T-----	3067
07_BC_CN_97CN001	-----G-----G-----A-----T-----T-----T-----A-----A-----C-----G-----A-----G-----A-----	2340
07_BC_CN_98CN009	-----G-----G-----T-----T-----T-----T-----A-----A-----C-----C-----C-----C-----A-----A-----	2340
07_BC_CN_C54A	-----G-----G-----A-----T-----T-----T-----A-----A-----C-----G-----G-----G-----A-----	2198
CONSENSUS_08_BC	-----C-----C-----T-----T-----A-----A-----G-----G-----G-----G-----G-----G-----	2187
08_BC_CN_97CNGX6F	-----C-----C-----T-----T-----A-----A-----G-----G-----G-----G-----G-----G-----	2188
08_BC_CN_97CNGX7F	-----C-----C-----G-----G-----T-----T-----A-----A-----G-----G-----G-----G-----	2194
08_BC_CN_97CNGX9F	-----C-----C-----T-----T-----A-----A-----G-----G-----G-----G-----G-----G-----	2188
08_BC_CN_98CN006	-----C-----C-----T-----T-----A-----A-----G-----G-----G-----G-----G-----G-----	2340
10_CD_TZ_96TZBF061	-----G-----G-----A-----T-----T-----T-----A-----A-----C-----C-----C-----C-----A-----A-----	2387
10_CD_TZ_96TZBF071	-----G-----G-----T-----T-----T-----T-----A-----A-----C-----C-----C-----C-----A-----A-----	2385
10_CD_TZ_96TZBF110	-----A-----A-----T-----T-----T-----T-----A-----A-----C-----C-----C-----C-----A-----A-----	2376
CONSENSUS_11_cpx	-----g-----C-----Tc-----a-----a-----C-----t-----C-----t-----C-----a-----	2947
11_cpx_CM_MP818	-----C-----C-----Tc-----A-----A-----C-----C-----C-----C-----C-----C-----a-----	3011
11_cpx_FR_MP1298	-----C-----C-----T-----A-----A-----A-----A-----C-----C-----C-----C-----C-----C-----	3002
11_cpx_FR_MP1307	-----G-----C-----C-----C-----Tc-----AG-----A-----A-----C-----C-----C-----C-----A-----	3059
11_cpx_GR_GR17	-----A-----C-----Tc-----AG-----A-----A-----A-----A-----C-----C-----C-----C-----A-----	2310
12_BF_AR_ARMA159	-----G-----C-----C-----G-----CT-----C-----T-----G-----C-----A-----C-----G-----C-----G-----	3013
12_BF_UY_URTR23	-----G-----C-----C-----CA-----G-----G-----C-----C-----C-----C-----A-----G-----C-----	3034
12_BF_UY_URTR35	-----G-----C-----C-----A-----G-----T-----T-----C-----C-----C-----C-----A-----G-----	3031
N_CM_YBF106	-----A-----A-----T-----G-----G-----CA-----T-----T-----A-----A-----T-----T-----C-----T-----	2600
N_CM_YBF30	-----A-----A-----T-----G-----G-----CA-----T-----T-----A-----A-----T-----T-----C-----T-----	2602
CONSENSUS_0	-----A-----A-----T-----C-----Tc-----cCC-----t-----?-----A-----A-----c-----C-----T-----T-----	2922
O_CM_ANT70	-----A-----A-----T-----C-----Tc-----C-----C-----C-----T-----T-----A-----A-----C-----C-----G-----a-----	3063
O_SN_MVP5180	-----A-----A-----T-----C-----Tc-----C-----C-----C-----T-----T-----A-----A-----C-----C-----G-----A-----	3038
O_SN_SEMP1299	-----A-----A-----C-----Tc-----C-----C-----C-----T-----T-----A-----A-----C-----C-----G-----A-----	3062
O_SN_SEMP1300	-----A-----A-----T-----C-----Tc-----C-----C-----C-----T-----T-----A-----A-----C-----C-----G-----A-----	3062
CONSENSUS_CPZ	-----a-----C-----C-----ctg-----a-----?-----?-----C-----g-----t-----C-----A-----A-----t-----a-----	2518
CPZ_CD_CFPZANT	-----A-----C-----C-----CA-----A-----C-----G-----T-----A-----A-----A-----T-----C-----A-----T-----A-----C-----	2447
CPZ_CM_CAM3	-----A-----T-----C-----CTGC-----A-----A-----T-----T-----A-----A-----T-----C-----G-----T-----A-----A-----	2419
CPZ_CM_CAM5	-----A-----T-----C-----CTGC-----A-----A-----T-----T-----A-----A-----C-----A-----T-----A-----R-----A-----	2710
CPZ_GA_CFPZGAB	-----A-----A-----C-----C-----CTG-----C-----G-----A-----T-----T-----A-----A-----T-----C-----A-----T-----	3069
CPZ_US_CFPZUS	-----A-----A-----C-----C-----TTG-----AC-----G-----A-----G-----A-----T-----A-----A-----C-----A-----	3068
Pol p51 RT	V_G_D_A_Y_F_S_V_P_L_D_E_D_F_R_K_Y_T_A_F_T_I_P_S_I_N_N_E_T_P_G_I_R_Y_Q_Y_N_V_L_P_Q_G_W	

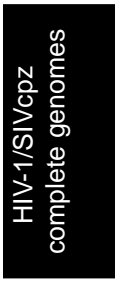
HIV-1/SIVcpz complete genomes

B_FR_HXB2	CAGCATAGAACAAAAATAGAGAGGCTGAGACAAATCTGTTGGAGTGGGGACTTACACACAGACAAAAAAA...CATCAGAAAAGAACCTCCATTCTCTTGGATGGGTTATGAACCTCCATCCATCTGATAAATGG	3266
CONSENSUS_A-A1	-----a---T-a---GcT---A---c---t---T---t---t---t---G...g...t---a---a---c---g---	2591
A_BY_97BL006	A---T---T---G---G---A---C---C---T---T---T---T---T---C---C---T---T---G---G---A---G---A---N---	2636
A_SE_SE7535	A---T---T---G---G---A---C---C---T---T---T---T---T---C---C---T---T---G---G---A---G---A---G---	2483
A_SE_SE8538	A---T---T---G---G---A---C---C---T---T---T---T---T---C---C---T---T---G---G---A---G---A---G---	2465
A_SE_SE8891	A---T---G---G---G---G---A---T---T---G---G---G---G---G---G---T---T---C---C---T---T---C---G---	2455
A_SE_UGSE8131	A---T---G---G---G---G---A---T---T---G---G---G---G---G---G---T---T---C---C---T---T---C---G---	2638
A1_KE_Q23-17	A---T---A---G---G---T---A---C---C---T---T---T---T---T---C---C---T---T---G---G---A---G---	2724
A1_SE_SE7253	A---T---A---G---G---T---A---C---C---T---T---T---T---T---C---C---T---T---G---G---A---G---	2470
A1_UG_92UG037	A---T---A---G---G---T---A---C---C---T---T---T---T---T---C---C---T---T---G---G---A---G---	2633
A1_UG_U455	AT---A---G---G---T---A---C---C---T---T---T---T---T---C---T---C---T---C---T---G---G---A---G---	2712
A2_CD_97CDKTB48	-----G---GT---T---T---A---GGCT---T---A---A---A---A---T---T---T---T---G---A---G---T---C---G---	2604
A2_Cy_94CY017.41	-----GT---AT---A---GGCT---CT---A---AA---A---A---T---T---T---T---T---T---T---G---A---G---T---C---	2623
CONSENSUS_B	-----A---A---C---C---T---T---T---T---T---T---T---T---T---T---G---A---G---T---C---G---	2811
B_AU_MBC200	A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	3265
B_CN_RL42	A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	2653
B_DE_D31	A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	2812
B_DE_HAN	A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	2732
B_ES_89SP061	A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	2824
B_GA_OYI	A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	2809
B_GB_MANC	A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	3212
B_NL_3202A21	A---A---G---GT---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	3268
B_TW_TWCYS	T---G---G---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	3266
B_US_JREL	G---G---G---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	2583
B_US_MNCG	-----AT---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	3281
B_US_RF	T---G---G---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	2783
B_US_SF2	A---A---G---G---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	3273
B_US_YU2	A---A---G---G---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	3264
CONSENSUS_C	-----a---T---A---g---a---a---a---t---t---t---t---t---g...g...t---a---a---c---t---	2589
C_BR_92BR025	A---G---G---T---A---G---G---A---A---A---A---T---T---C---C---C---C---T---G---G---A---G---	2604
C_BR_98BR004	A---GG---A---AT---A---G---G---T---A---A---A---T---T---C---C---C---C---T---G---G---A---G---	2591
C_BW_96BW0502	A---GC---G---C---A---AT---A---G---G---C---C---A---A---A---T---T---C---C---C---C---T---G---	2759
C_BW_96BW06	A---G---G---G---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	3248
C_BW_96BW15C02	A---G---G---G---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---A---	2733
C_ET_ETH2220	A---G---C---C---C---A---T---A---G---A---A---A---T---T---T---T---T---G---G---A---G---T---	2658
C_IL_98ILS002	A---A---C---G---T---T---A---G---A---A---A---T---T---T---T---T---G---G---A---G---T---	2600
C_IN_93IN101	A---G---G---T---T---A---G---A---A---A---T---T---T---T---T---G---G---A---G---T---	3239
C_IN_94IN476	T---G---G---G---T---A---GC---A---A---A---T---T---C---C---T---G---G---A---G---T---	2587
C_IN_98IN012	A---G---G---A---T---A---G---A---A---A---T---T---A---A---A---A---G---G---A---G---T---	2591
C_TZ_98TZ013	A---G---G---A---T---A---G---A---A---A---T---T---A---A---A---A---G---G---A---G---T---	2609
C_TZ_98TZ017	A---G---G---A---T---A---G---A---A---A---T---T---A---A---A---A---G---G---A---G---T---	2620
C_ZA_97ZA012	A---G---G---A---T---A---G---A---A---A---T---T---A---A---A---A---G---G---A---G---T---	2600
C_ZM_96ZM651	A---G---G---A---T---A---G---A---A---A---T---T---A---A---A---A---G---G---A---G---T---	2595
C_ZM_96ZM751	A---A---C---G---T---A---G---G---C---A---A---A---T---T---T---T---G---G---A---G---T---	2606
CONSENSUS_D	-----AT---A---G---G---a---a---a---t---t---t---t---t---t---t---g...g...t---a---a---c---	2806
D_CD_84ZR085	A---AT---A---G---G---A---A---A---A---T---T---G---G---T---T---G---G---A---G---T---G---	2788
D_CD_FLI	A---AT---A---G---G---A---A---A---A---T---T---G---G---T---T---G---G---A---G---T---G---	2812
D_CD_Z26	-----AT---A---G---G---A---A---A---A---T---T---G---G---T---T---G---G---A---G---T---G---	3269
D_KE_MB2059	-----AT---A---GGG---C---A---A---A---T---T---G---G---T---T---G---G---A---G---T---G---	3270
D_UG_94UG1141	-----AT---A---GGG---C---C---A---A---T---T---G---G---T---T---G---G---A---G---T---G---	2621
CONSENSUS_F1	-----T---A---G---G---AC---aA---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2555
F1_BE_VI850	T---A---G---G---AC---aA---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2603
F1_BR_93BR020.1	A---T---A---G---G---AC---AA---T---A---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2594
F1_FI_FIN9363	T---A---G---G---AC---AA---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2595
F1_FR_MP411	G---TG---T---A---GG---AC---AA---CT---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2468
F2_CM_MP255	-----G---G---T---A---G---G---A---A---A---A---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2462
F2_CM_MP257	-----G---G---T---A---G---G---A---A---A---A---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2474
CONSENSUS_G	-----G---G---T---A---G---G---A---A---A---A---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2637
G_BE_DRCBL	-----G---G---T---A---G---G---A---A---A---A---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	3223
G_FI_HH8793-12.1	-----G---G---T---A---G---G---A---A---A---A---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2667
G_NG_92NG083	-----G---G---T---A---G---G---A---A---A---A---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2635
G_SE_SE6165	-----G---G---T---A---G---G---A---A---A---A---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2663
H_BE_VI991	-----GA---A---T---A---G---G---T---T---T---T---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2655
H_BE_VI997	-----A---G---G---T---A---G---G---T---T---T---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2590
H_CF_90CF056	-----A---G---G---T---A---G---G---T---T---T---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2613
J_SE_SE7022	-----G---G---A---A---GG---A---A---A---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2581
J_SE_SE7887	-----G---G---A---A---GG---A---A---A---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2580
K_CD_E0TB11C	-----A---G---G---T---A---G---G---A---A---A---A---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2462
K_CM_MP535	-----A---C---G---G---T---A---G---G---A---A---A---A---T---T---T---T---T---T---T---T---T---G...g...t---a---a---c---	2462

B_FR_HXB2
CONSENSUS_01_AE
01_AE_CF_90CF11697
01_AE_CF_90CF402
01_AE_CF_90CF4071
01_AE_TH_93TH057
01_AE_TH_93TH065
01_AE_TH_93TH253
01_AE_TH_95TH022
01_AE_TH_95TH047
01_AE_TH_CM240
CONSENSUS_02_AG
02_AG_CM_97CM-MP807
02_AG_FR_DJ263
02_AG_FR_DJ264
02_AG_GH_G829
02_AG_NG_IBNG
02_AG_SE_SE7812
02_AG_SN_98SE-MP1211
02_AG_SN_MP1213
03_AB_RU_KAL153-2
03_AB_RU_RU98001
04_CFX_CX_94CY032-3
04_CFX_GR_97PVGH
04_CFX_GR_97PVNY
05_DF_BE_V11310
05_DF_BE_V1961
CONSENSUS_06_CPX
06_CPX_AU_BFP90
06_CPX_ML_95ML127
06_CPX_ML_95ML84
06_CPX_SN_97SEL078
07_BC_CN_97CN001
07_BC_CN_98CN009
07_BC_CN_C54A
CONSENSUS_08_BC
08_BC_CN_97CNGX6F
08_BC_CN_97CNGX7F
08_BC_CN_97CNGX9F
08_BC_CN_98CN006
10_CD_TZ_96TZBF061
10_CD_TZ_96TZBF071
10_CD_TZ_96TZBF110
CONSENSUS_11_CPX
11_CPX_CM_MP818
11_CPX_FR_MP1298
11_CPX_FR_MP1307
11_CPX_GR_GR17
12_BF_AR_ARMA159
12_BF_UY_URTR23
12_BF_UY_URTR35
N_CM_YBF106
N_CM_YBF30
CONSENSUS_0
O_CM_ANT70
O_CM_MVP5180
O_SN_SEMP1299
O_SN_SEMP1300
CONSENSUS_CPZ
CPZ_CD_CFPZANT
CPZ_CM_CAM3
CPZ_CM_CAM5
CPZ_GA_CFPZGAB
CPZ_US_CFPZUS
PoI P51 RT



B_FR_HXB2	TTAGAGAACCAAGCACTAACAGAAAGTAAATACCACTAACACAGAAAGCAGAGCTAGAACTGGCAGAAAACAGAGAGATTCTAAAAAAGACAGATGAGTGATTATGACCCATCAAAAAGACTTAAAT	3526
CONSENSUS_01_AE	C-G-g-T-----G-T-----AT-----T-----G-----g-----ACC-T-G-----a-----G-----	3477
01_AE_CF_90CF11697	C-G-g-T-G-----G-T-G-----AT-----T-----G-----G-----A-----A-----C-----T-----G-----A-----	3475
01_AE_CF_90CF402	C-G-g-T-G-----G-T-G-----AT-----T-----G-----G-----A-----A-----C-----T-----G-----A-----	3562
01_AE_CF_90CF4071	C-G-g-T-G-----G-T-G-----AT-----T-----G-----G-----A-----A-----C-----T-----G-----A-----	3462
01_AE_TH_93TH057	C-G-g-T-G-----G-T-G-----AT-----T-----G-----G-----A-----A-----C-----T-----G-----A-----	3479
01_AE_TH_93TH065	C-G-g-T-G-----G-T-G-----AT-----T-----G-----G-----A-----A-----C-----T-----G-----A-----	3471
01_AE_TH_93TH253	C-G-g-T-G-----G-T-G-----AT-----T-----G-----G-----A-----A-----C-----T-----G-----A-----	3513
01_AE_TH_95TH022	C-G-g-T-G-----G-T-G-----AT-----T-----G-----G-----A-----A-----C-----T-----G-----A-----	3489
01_AE_TH_95TH047	C-G-g-T-G-----G-T-G-----AT-----T-----G-----G-----A-----A-----C-----T-----G-----A-----	3489
01_AE_TH_CM240	C-G-g-T-G-----G-T-G-----AT-----T-----G-----G-----A-----A-----C-----T-----G-----A-----	3100
CONSENSUS_02_AG	C-g-g-g-----G-t-g-----AT-----T-----G-----g-----a-----a-----a-----a-----a-----	2736
02_AG_CM_97CM-MP807	C-g-g-g-----G-t-g-----AT-----T-----G-----g-----a-----a-----a-----a-----a-----	2682
02_AG_FR_DJ263	C-g-g-g-----G-t-g-----AT-----T-----G-----g-----a-----a-----a-----a-----a-----	2874
02_AG_FR_DJ264	C-g-g-g-----G-t-g-----AT-----T-----G-----g-----a-----a-----a-----a-----a-----	2875
02_AG_GH_G829	C-g-g-g-----G-t-g-----AT-----T-----G-----g-----a-----a-----a-----a-----a-----	2742
02_AG_NG_IBNG	C-g-g-g-----G-t-g-----AT-----T-----G-----g-----a-----a-----a-----a-----a-----	3051
02_AG_SE_SE7812	C-g-g-g-----G-t-g-----AT-----T-----G-----g-----a-----a-----a-----a-----a-----	2898
02_AG_SN_98SE-MP1211	C-g-g-g-----G-t-g-----AT-----T-----G-----g-----a-----a-----a-----a-----a-----	2723
02_AG_SN_MF1213	C-g-g-g-----G-t-g-----AT-----T-----G-----g-----a-----a-----a-----a-----a-----	2720
03_AB_RU_KAL153-2	G-G-G-----C-----C-----C-----C-----T-----T-----G-----G-----a-----a-----G-----G-----	3295
03_AB_RU_RU98001	G-G-G-----C-----C-----C-----C-----T-----T-----G-----G-----a-----a-----G-----G-----	2896
04_CFX_CX_94CY032-3	G-g-T-----C-----TAC-----G-----T-----T-----A-----G-----G-----G-----CA-----C-----	2892
04_CFX_GR_97PVGH	G-g-T-----C-----TAC-----G-----T-----T-----A-----G-----G-----G-----CA-----C-----	3543
04_CFX_GR_97PVNY	G-g-T-----C-----TAC-----G-----T-----T-----A-----G-----G-----G-----CA-----C-----	3540
05_DF_BE_V11310	G-----T-----G-----T-----AT-----T-----G-----G-----A-----G-----A-----G-----G-----	2908
05_DF_BE_V1961	G-----T-----G-----T-----AT-----T-----G-----G-----A-----G-----A-----G-----G-----	2881
CONSENSUS_06_CPX	C-G-G-A-----T-C-----A-----T-----?-----G-----A-----?-----G-----A-----g-----	3495
06_CPX_AU_BFP90	C-G-G-A-----T-C-----A-----T-----G-----G-----A-----T-----G-----A-----G-----G-----	3554
06_CPX_ML_95ML127	C-G-G-A-----T-C-----A-----T-----G-----G-----A-----T-----G-----A-----G-----G-----	3543
06_CPX_ML_95ML84	C-G-G-A-----T-C-----A-----T-----G-----G-----A-----T-----G-----A-----G-----G-----	2769
06_CPX_SN_97SEL078	C-G-G-A-----T-C-----A-----T-----T-----G-----A-----A-----G-----G-----A-----G-----	3585
07_BC_CN_97CN001	G-GG-----T-----T-----AT-----T-----G-----A-----A-----A-----C-----G-----G-----	2858
07_BC_CN_98CN009	G-GG-----T-----T-----AT-----T-----G-----A-----A-----A-----C-----G-----G-----	2858
07_BC_CN_C54A	G-GG-----T-----T-----AT-----T-----G-----A-----A-----A-----C-----G-----G-----	2716
CONSENSUS_08_BC	G-GG-----T-----T-----AT-----T-----a-----G-----T-----G-----A-----T-----CA-----A-----g-----	2705
08_BC_CN_97CNGX6F	G-GG-----T-----T-----AT-----T-----a-----G-----T-----G-----A-----T-----CA-----A-----g-----	2706
08_BC_CN_97CNGX7F	G-GG-----T-----T-----AT-----T-----a-----G-----T-----G-----A-----T-----CA-----A-----g-----	2712
08_BC_CN_97CNGX9F	G-GG-----T-----T-----AT-----T-----a-----G-----T-----G-----A-----T-----CA-----A-----g-----	2706
08_BC_CN_98CN006	G-GG-----T-----T-----AT-----T-----a-----G-----T-----G-----A-----T-----CA-----A-----g-----	2858
10_CD_TZ_96TZBF061	G-G-----T-----T-----AT-----T-----G-----T-----G-----A-----A-----A-----G-----G-----	2905
10_CD_TZ_96TZBF071	G-G-----T-----T-----AT-----T-----G-----T-----G-----A-----A-----A-----G-----G-----	2903
10_CD_TZ_96TZBF110	G-G-----T-----T-----AT-----T-----G-----T-----G-----A-----A-----A-----G-----G-----	2894
CONSENSUS_11_CPX	G-g-----T-C-----AT-----G-----GT-----g-----G-----A-----A-----A-----A-----	3454
11_CPX_CM_MP818	G-g-----T-C-----AT-----G-----GT-----g-----G-----A-----A-----A-----A-----	3529
11_CPX_FR_MP1298	G-g-----T-C-----AT-----G-----GT-----g-----G-----A-----A-----A-----A-----	3520
11_CPX_FR_MP1307	G-g-----T-C-----AT-----G-----GT-----g-----G-----A-----A-----A-----A-----	3577
11_CPX_GR_GR17	G-g-----T-C-----AT-----G-----GT-----g-----G-----A-----A-----A-----A-----	2828
12_BF_AR_ARM159	G-G-----C-----A-----A-----A-----A-----G-----A-----G-----A-----A-----G-----	3531
12_BF_UY_URTR23	G-G-----C-----A-----A-----A-----A-----G-----A-----G-----A-----A-----G-----	3552
12_BF_UY_URTR35	G-G-----C-----A-----A-----A-----A-----G-----A-----G-----A-----A-----G-----	3549
N_CM_YBF106	G-TT-----G-----CACT-----T-----C-----AT-----A-----A-----G-----A-----A-----G-----	3118
N_CM_YBF30	G-TT-----G-----CACT-----T-----C-----AT-----A-----A-----G-----A-----A-----G-----	3120
CONSENSUS_0	c-gT-----g-----GTA-----g-----GTA-----g-----T-----A-----A-----g-----?-----A-----a-----	3428
O_CM_ANT70	c-gT-----g-----GTA-----g-----GTA-----g-----T-----A-----A-----g-----?-----A-----a-----	3581
O_CM_MVP5180	c-gT-----g-----GTA-----g-----GTA-----g-----T-----A-----A-----g-----?-----A-----a-----	3556
O_SN_SEMP1299	c-gT-----g-----GTA-----g-----GTA-----g-----T-----A-----A-----g-----?-----A-----a-----	3580
O_SN_SEMP1300	c-gT-----g-----GTA-----g-----GTA-----g-----T-----A-----A-----g-----?-----A-----a-----	3580
CONSENSUS_CPZ	a-a-----t-----?-----a-g-----c-g-----Aa-g-----T-----A-----a-----t-----a-----a-----g-----c-----c-----	3018
CPZ_CD_CPZANT	a-a-----t-----?-----a-g-----c-g-----Aa-g-----T-----A-----a-----t-----a-----a-----g-----c-----c-----	2965
CPZ_CM_CAM3	A-A-----G-----T-----CA-----G-----CA-----G-----C-----G-----AA-----G-----T-----A-----A-----G-----T-----R-----G-----T-----G-----C-----C-----	2937
CPZ_CM_CAM5	A-A-----G-----T-----CA-----G-----CA-----G-----C-----G-----AA-----G-----T-----A-----A-----G-----T-----R-----G-----T-----G-----C-----C-----	3228
CPZ_GA_CPZGAB	A-A-----G-----T-----T-----G-----T-----T-----C-----CC-----AT-----T-----A-----A-----T-----G-----AG-----GCAC-----G-----A-----C-----T-----G-----G-----	3587
CPZ_US_CPZUS	A-A-----G-----T-----T-----TG-----AAAT-----T-----C-----T-----G-----AA-----G-----GT-----A-----A-----A-----A-----G-----G-----C-----C-----	3586
Po1_P51_RT	L_R_G_T_K_A_L_T_E_V_I_P_L_T_E_F_A_E_L_E_L_A_E_N_R_E_I_L_K_E_P_V_H_G_V_Y_Y_D_P_S_K_D_L_I	Po1

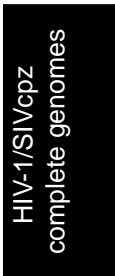


HIV-1/SIVcpz
complete genomes

B.FR.HXB2	AGCAGAAATACAGAGCGGGGCAATGGACATATCAAATTTTATCAAGAGCCATTTAAATAATCTGAAAACAGAAAATATGCAAGAATGAGGGGTGCCACACTAATGATGTAACAATAAACA	3656
CONSENSUS_A-A1	-----A-A-A-A-----A-----a-----Aa-TC-T-----g	2981
A.BY.97BL006	-----A-A-A-A-----A-----G-----A-ARR-TC-T-----R	3026
A.SE.SE7535	-----A-A-A-A-----A-----A-----AA-TC-T-----G	2873
A.SE.SE8538	-----A-A-A-A-----A-----C-----A-TC-T-----G	2855
A.SE.SE8891	-----A-A-A-A-----A-----T-----A-TC-T-----G	2845
A.SE.UGSE8131	-----A-A-A-A-----A-----G-----AA-TC-T-----GT	3028
A1.KE.Q23-17	-----A-A-A-A-----A-----T-----AA-TC-T-----G	3114
A1.SE.SE7253	-----A-A-A-A-----A-----G-----AA-TC-T-----G	2860
A1.UG.92UG037	-----A-A-A-A-----A-----T-----AA-TC-T-----GG	3023
A1.UG.U455	-----A-A-A-A-----A-----A-----AA-TC-T-----G	3102
A2.CD.97CDKTB48	-----A-A-A-A-----A-----T-----A-----A-G-A-TCCA-----C	2994
A2.CY.94CY017.41	-----A-A-A-A-----A-----G-A-C-----A-G-----TCCA-----A	3013
CONSENSUS_B	-----T-----g-----	3200
B.AU.MBC200	-----T-----C-----A-----G-----G-----	3655
B.CN.RL42	-----G-----A-----G-----G-----	3043
B.DE.D31	-----G-----A-----G-----G-----	3202
B.DE.HAN	-----G-----A-----G-----G-----	3122
B.ES.89SP061	-----G-----A-----G-----G-----	3214
B.GA.OYI	-----G-----G-----G-----A-----A-----G-----	3199
B.GB.MANC	-----G-----C-----G-----G-----C-----	3602
B.NL.3202A21	-----G-----T-----T-----G-----C-----	3656
B.TW.TWCYS	-----C-----G-----C-----G-----G-----	2973
B.US.JREL	-----C-----C-----G-----G-----G-----	3671
B.US.MNCG	-----C-----G-----G-----G-----G-----	3173
B.US.RF	-----C-----G-----G-----G-----G-----	3663
B.US.SF2	-----C-----G-----G-----G-----G-----	3654
B.US.YU2	-----C-----G-----G-----G-----G-----	2979
CONSENSUS_C	-----T-A-----t-A-----C-----C-----A-----C-----g-g-----g	2994
C.BR.92BR025	-----A-A-----GAA-----C-----C-----A-----C-----A-----AC-----G-----G-----	2981
C.BR.98BR004	-----A-A-----G-----C-----C-----A-----C-----A-----AC-----T-----G-----G-----	2981
C.BW.96BW0502	-----A-A-----T-A-----G-----C-----C-----A-----C-----A-----AC-----G-----G-----	3149
C.BW.96BW06	-----A-A-----A-----GAT-----C-----C-----A-----C-----C-----AC-----G-----G-----	3638
C.BW.96BW15C02	-----A-A-----T-A-----T-----C-----C-----C-----C-----A-----AC-----G-----G-----	3123
C.ET.ETH2220	-----A-A-----A-----T-----T-----C-----C-----A-----C-----A-----GAG-----AC-----G-----G-----	3048
C.IL.98ILS002	-----A-A-----T-A-----G-----T-----C-----C-----A-----C-----A-----AC-----G-----G-----	2990
C.IN.93IN101	-----A-A-----G-A-----G-----C-----C-----A-----C-----A-----G-----G-----G-----	3629
C.IN.94IN476	-----A-A-----T-A-----T-----C-----C-----A-----C-----A-----AC-----T-----G-----G-----	2977
C.IN.98IN012	-----A-A-----A-----T-----A-----C-----C-----A-----C-----A-----A-----AC-----GG-----	2981
C.TZ.98TZ013	-----A-A-----A-----G-----A-----C-----C-----A-----C-----A-----TC-----G-----G-----	2999
C.TZ.98TZ017	-----A-A-----T-----T-----A-----C-----C-----A-----C-----A-----AC-----G-----G-----	3010
C.ZA.97ZA012	-----A-A-----G-----T-----A-----C-----C-----A-----C-----A-----AC-----G-----G-----	2990
C.ZM.96ZM651	-----A-A-----T-----T-----A-----C-----C-----A-----C-----A-----ACA-----G-----G-----	2985
C.ZM.96ZM751	-----A-A-----T-----A-----A-----G-----C-----C-----A-----C-----A-----AC-----G-----G-----	2995
CONSENSUS_D	-----A-A-----A-----?-----A-----A-----A-----A-----G-----G-----	3195
D.CD.84ZR085	-----A-A-----A-----T-----T-----A-----A-----A-----A-----G-----G-----	3178
D.CD.FLI	-----A-A-----A-----C-----C-----A-----A-----A-----A-----G-----G-----	3202
D.CD.ZZ6	-----A-A-----A-----C-----C-----A-----A-----A-----A-----G-----G-----	3659
D.KE.MB2059	-----A-TAA-----AAA-----T-----A-A-----AC-----G-----A-----A-----G-----	3660
D.UG.94UG1141	-----A-A-----AT-----A-----A-----A-----A-----G-----G-----A-----A-----	3011
CONSENSUS_F1	-----A-A-----A-----g-----A-----A-----A-----A-----TC-----	2943
F1.BE.VI850	-----A-A-----AG-C-----G-----GA-C-----A-----A-----G-----TCG-----	2993
F1.BR.93BR020.1	-----A-A-----A-----G-----A-----A-----A-----A-----TC-----G-----	2984
F1.FI.FIN9363	-----A-A-----A-----A-----A-----C-----G-----A-----A-----TC-----G-----	2985
F1.FR.MF411	-----C-G-----A-A-----A-----C-----A-----C-----A-----CA-----TC-----A-----G-----	2858
F2.CM.MP255	-----A-A-----A-----T-----AT-----A-----A-----A-----T-----A-----G-----A-----TC-----	2852
F2.CM.MP257	-----A-A-----A-----A-----C-----A-----CA-----C-----A-----CA-----G-----A-----TC-----G-----	2864
CONSENSUS_G	-----G-----?-----a-----A-----t-----?-----Ac-----	3020
G.BE.DRCBL	-----G-----A-----A-----A-----T-----G-----G-----A-----Ac-----G-----G-----TC-----G-----	3613
G.FI.HH8793-12.1	-----G-----A-----A-----AAA-----A-----C-----A-----C-----A-----G-----AC-----G-----G-----	3056
G.NG.92NG083	-----G-----A-----A-----C-----A-----A-----AC-----A-----A-----G-----TC-----G-----	3025
G.SE.SE6165	-----G-----A-----A-----T-----A-----A-----AC-----A-----A-----G-----TC-----G-----	3053
H.BE.VI991	-----A-----A-----C-----A-----G-----C-----C-----A-----TC-----G-----	3045
H.BE.VI997	-----A-----A-----C-----A-----C-----C-----A-----A-----AAA-----G-----G-----	2980
H.CF.90CF056	-----G-----A-----A-----A-----G-----G-----A-----A-----AAC-----A-----A-----	3003
J.SE.SE7022	-----G-G-----A-----A-----TG-A-----G-----G-----C-----C-----A-----G-----A-----G-----T-----G-----	2971
J.SE.SE7887	-----G-G-----A-----A-----TG-A-----G-----G-----C-----C-----A-----A-----G-----A-----G-----G-----	2970
K.CD.EQTB11C	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----TC-----	2852
K.CM.MP535	-----A-----A-----A-----T-----A-----A-----CA-----A-----ATC-----A-----G-----	2852

B.FR.HXB2	GAGGCAGTGC	AAAATAAACACACAGAAAGCATAGTAATATGGGAAAG	...ACTCTAAATTTAAACTGCC	...ATACAAAAGAAACATGGGAAACA	...TGTTGGACAGAGTATTGGCAACCCACCTGGAT	3781
CONSENSUS_A-A1	--A-Tg	-----9G-9gt--tg-A-?	-----a	-----Tg--c---	3781
A.BY.97BL006	--CA-TG	-----G-GG	-----G-A-C-	-----R	-----R-Tg--c---	3106
A.SE.SE7535	--A-TG	-----GG-GG	-----G-A-	-----A-G	-----R-Tg--c---	3153
A.SE.SE8538	--A-TG	-----GG-GGT	-----G-C	-----A-G	-----R-Tg--c---	2998
A.SE.SE8891	--A-TG	-----GG-GT	-----G-T	-----A-G	-----R-Tg--c---	2980
A.SE.UGSEF8131	--A-T	-----G-GGT	-----G-A	-----A-G	-----R-Tg--c---	2970
A1.KE.Q23-17	--A-TG	-----GG-GGT	-----G-A	-----A-C	-----Tg--c---	3153
A1.SE.SE7253	--A-TG	-----GG-GGT	-----G-A	-----A-C	-----Tg--c---	3239
A1.UG.92UG037	--A-TG	-----GG-GGT	-----G-A	-----A-C	-----Tg--c---	2985
A1.UG.U455	--A-TG	-----GG-GT	-----TC--G-A	-----A	-----Tg--c---	3148
A2.CD.97CDKTB48	--A--	-----G	-----T	-----A	-----G--G--G-T	3227
A2.CY.94CY017.41	--A--	-----G	-----T	-----A	-----G--G--G-T	3119
CONSENSUS_B	-----A	-----g	-----a	-----a	-----g	3133
B.AU.MBC200	-----A	-----G	-----A	-----A	-----G	3325
B.CN.RL42	-----A	-----G	-----A	-----A	-----G	3780
B.DE.D31	-----A	-----G	-----A	-----A	-----G	3168
B.DE.HAN	-----A	-----G	-----A	-----A	-----G	3327
B.ES.89SP061	-----A	-----G	-----A	-----A	-----G	3247
B.GA.OYI	-----A	-----G	-----A	-----A	-----G	3339
B.GB.MANC	-----A	-----G	-----A	-----A	-----G	3324
B.NL.3202A21	-----A	-----G	-----A	-----A	-----G	3727
B.TW.TWCYS	-----A	-----G	-----A	-----A	-----G	3783
B.US.JREL	-----A	-----G	-----A	-----A	-----G	3781
B.US.MNCG	-----A	-----G	-----A	-----A	-----G	3098
B.US.RF	-----A	-----G	-----A	-----A	-----G	3796
B.US.SF2	-----A	-----G	-----A	-----A	-----G	3288
B.US.YU2	-----A	-----G	-----A	-----A	-----G	3779
CONSENSUS_C	-----A	-----g	-----a	-----a	-----g	3104
C.BR.92BR025	-----A	-----G	-----A	-----A	-----G	3119
C.BR.98BR004	-----A	-----G	-----A	-----A	-----G	3106
C.BW.96BW0502	-----A	-----G	-----A	-----A	-----G	3274
C.BW.96BW06	-----A	-----G	-----A	-----A	-----G	3763
C.BW.96BW15C02	-----A	-----G	-----A	-----A	-----G	3248
C.ET.ETH2220	-----A	-----G	-----A	-----A	-----G	3173
C.IL.98ILS002	-----A	-----G	-----A	-----A	-----G	3115
C.IN.93IN101	-----A	-----G	-----A	-----A	-----G	3754
C.IN.94IN476	-----A	-----G	-----A	-----A	-----G	3101
C.IN.98IN012	-----A	-----G	-----A	-----A	-----G	3106
C.TZ.98TZ013	-----A	-----G	-----A	-----A	-----G	3124
C.TZ.98TZ017	-----A	-----G	-----A	-----A	-----G	3135
C.ZA.97ZA012	-----A	-----G	-----A	-----A	-----G	3115
C.ZM.96ZM651	-----A	-----G	-----A	-----A	-----G	3110
C.ZM.96ZM751	-----A	-----G	-----A	-----A	-----G	3120
CONSENSUS_D	-----A	-----g	-----a	-----a	-----g	3319
D.CD.84ZR085	-----A	-----G	-----A	-----A	-----G	3303
D.CD.FLI	-----A	-----G	-----A	-----A	-----G	3327
D.CD.Z2Z6	-----A	-----G	-----A	-----A	-----G	3784
D.KE.MB2059	-----A	-----G	-----A	-----A	-----G	3785
D.UG.94UG1141	-----A	-----G	-----A	-----A	-----G	3136
CONSENSUS_F1	-----A	-----G	-----A	-----A	-----G	3066
F1.BE.VI850	-----A	-----G	-----A	-----A	-----G	3121
F1.BE.VI850	-----A	-----G	-----A	-----A	-----G	3121
F1.BR.93BR020.1	-----A	-----G	-----A	-----A	-----G	3109
F1.FI.FIN9363	-----A	-----G	-----A	-----A	-----G	3110
F1.FR.MF411	-----A	-----G	-----A	-----A	-----G	2993
F2.CM.MP255	-----A	-----G	-----A	-----A	-----G	2977
F2.CM.MP257	-----A	-----G	-----A	-----A	-----G	2989
CONSENSUS_G	-----A	-----g	-----a	-----a	-----g	3143
G.BE.DRCBL	-----A	-----G	-----A	-----A	-----G	3738
G.FI.HH8793-12.1	-----A	-----G	-----A	-----A	-----G	3181
G.NG.92NG083	-----A	-----G	-----A	-----A	-----G	3150
G.SE.SE6165	-----A	-----G	-----A	-----A	-----G	3178
H.BE.VI991	-----A	-----G	-----A	-----A	-----G	3170
H.BE.VI997	-----A	-----G	-----A	-----A	-----G	3105
H.CF.90CF056	-----A	-----G	-----A	-----A	-----G	3128
J.SE.SE7022	-----A	-----G	-----A	-----A	-----G	3096
J.SE.SE7887	-----A	-----G	-----A	-----A	-----G	3095
K.CD.EQTB11C	-----A	-----G	-----A	-----A	-----G	2977
K.CM.MP535	-----A	-----G	-----A	-----A	-----G	2977

B_FR_HXB2
 GAGGCAGTGCAGAAAAATAAACACAGAAAGCATAGTAATATGGGAAAG...ACTCTAAATTTAAACTGCC...ATACAAAAGGAAACATGGGAAACA...TGTTGGACAGAGTATTGGCAACCCACCTGGAT
 3781
 -A-TG-----g-----g-A-----g-A-----gA-----tg-----G-T
 3732
 -A-TG-----GG-GG-----C-C-----C-G-----G-A-----G-T-----G-G-----G-T
 3730
 -A-TG-----G-G-----C-C-----C-G-----G-A-----G-T-----G-G-----G-T
 3717
 -A-TG-----G-G-----C-C-----C-G-----G-A-----G-T-----G-G-----G-T
 3734
 -A-TG-----G-G-----C-C-----C-G-----G-A-----G-T-----G-G-----G-T
 3734
 -A-TG-----G-G-----C-C-----C-G-----G-A-----G-T-----G-G-----G-T
 3768
 -A-TG-----G-G-----C-C-----C-G-----G-A-----G-T-----G-G-----G-T
 3768
 -A-TG-----G-G-----C-C-----C-G-----G-A-----G-T-----G-G-----G-T
 3744
 -A-TG-----G-G-----C-C-----C-G-----G-A-----G-T-----G-G-----G-T
 3744
 -A-TG-----G-G-----C-C-----C-G-----G-A-----G-T-----G-G-----G-T
 3355
 -T-----G-99-T-----C-----C-----C-----G-----G-----G-----G-----G-T
 2990
 -T-----G-GG-T-----C-----C-----C-----G-----G-----G-----G-----G-T
 2937
 -T-----G-GGTT-----C-----C-----C-----G-----G-----G-----G-----G-T
 3129
 -T-----G-G-----C-----G-----GC-----G-----G-----G-----G-----G-T
 3130
 -A-T-----G-G-----C-----G-----GC-----G-----G-----G-----G-----G-T
 2997
 -T-----G-G-----C-----G-----G-----G-----G-----G-----G-----G-T
 3306
 -T-----GG-GG-T-----C-----C-----C-----G-----G-----G-----G-----G-T
 3153
 -T-----GG-GG-T-----C-----C-----C-----G-----G-----G-----G-----G-T
 2978
 -T-----GC-G-T-TG-----C-----C-----C-----G-----G-----G-----G-----G-T
 2975
 -G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3550
 -G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3151
 -A-----A-----G-----G-----G-----G-----G-----G-----G-----G-T
 3147
 -A-T-----A-----G-----G-----G-----G-----G-----G-----G-----G-T
 3798
 -A-----A-----G-----G-----G-----G-----G-----G-----G-----G-T
 3795
 -G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3163
 -G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3136
 -A-----?-----G-----G-----G-----G-----G-----G-----G-----G-T
 3749
 -A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3809
 -A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3798
 -A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3024
 -A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3840
 -T-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3113
 -T-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 2971
 -T-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 2960
 -T-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 2961
 -T-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 2967
 -T-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 2961
 -T-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3113
 -G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3160
 -G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3158
 -T-----A-----G-----G-----G-----G-----G-----G-----G-----G-T
 3149
 -A-Tg-a-----g-ttg-----?-----a-----g-----g-----g-----g-----g-----g
 3705
 -A-TGA-A-----TG-CTG-GC-----G-----G-----G-----G-----G-----G-----G
 3784
 -AAT-----A-----G-----TTG-----G-----G-----G-----G-----G-----G-----G
 3775
 -A-TG-----G-----G-----TTG-----G-----G-----G-----G-----G-----G-----G
 3832
 -A-TG-----G-----G-----TTG-----G-----G-----G-----G-----G-----G-----G
 3083
 -A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3786
 -G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3807
 -G-----G-----G-----G-----G-----G-----G-----G-----G-----G-T
 3804
 -CA-TG-----G-GG-A-----T-----T-----T-----A-----A-----G-----G-----G-----G-----G-T
 3373
 -A-TG-----AAGG-----G-GG-A-----T-----T-----T-----A-----A-----G-----G-----G-----G-----G-T
 3375
 -A-T-C-C-G-GG-GT-TCA-----gCT-----a-T-C-----A...TTG-----G-----G-----G-----G-----G-T
 3676
 -A-T-A-C-G-GG-GT-TCA-----GCT-----A-T-C-----A...TTG-----G-----G-----G-----G-----G-T
 3836
 -A-T-C-C-G-GG-GT-TCA-----GCT-----A-T-C-----G...TTA-----C-GG-----A...G-TACT-GA-----T
 3811
 -A-T-C-C-G-GG-GT-TCA-----GCT-----A-T-C-----A...TTG-----G-----G-----G-----G-----G-T
 3835
 -A-T-C-C-G-GG-GT-TCA-----GCT-----A-T-C-----A...TTG-----G-----G-----G-----G-----G-T
 3835
 -a-t-----G-t-t-----t-----a-----a-----g-T-a-?-g-?-?-g-?-?-g-?-?-g-?-?-g-?-?-g-?-?-g-?-?-g-?-?
 3255
 -GA-T-----GGATT-----T-----T-----C-----G-----G-----G-----G-----G-----G-----G-T
 3220
 -T-----A-----G-----T-----C-----A-----C-----T-----A-----C-----T-----A-----G-----G-----G-----G-----G-T
 3192
 -TG-----G-----T-----T-----C-----A-----A-----C-----G-----T-----A-----T-----G-----T-----G-----G-----G-----G-----G-T
 3483
 -A-----A-----TG-T-----T-----C-----A-----A-----A-----C-----G-----T-----A-----T-----G-----G-----G-----G-----G-T
 3842
 -A-TG-----G-T-----T-----G-----T-----C-----T-----GTA-----A-----G-----G-----G-----G-----G-T
 3841
 E_A_V_O_K_I_T_T_E_S_I_V_I_W_G_K...T_P_K_F_K_L_P...I_O_K_E_T_W_E_T...W_W_T_E_Y_W_O_A_T_W_I
 P_O_I



HIV-1/SIVcpz
 complete genomes

HIV-1/SIVcpz complete genomes

(RT is heterodimer of p51 and p66)
 p51 end p66 RT continue V Pol p15 RNase H start

B.FR.HXB2	TCCTGAGTGGGAGTTTGTAAATACCCCTCCCTAGTGAATAATGTTACAGTATTAGAGAAAGAACCCCATAGTAGGAGCAGAAACCTTCTATGTAGTGGGCGACTAAACGGAGACTAAATATTAGGAAAA	3911
CONSENSUS_A-A1	-----tC-A-A-C-----g-t-----c-t-----a-----Gc-----	3236
A.BY.97BL006	-----A-A-----TC-A-A-C-----G-----G-T-----A-----G-T-----GA-----	3283
A.SE.SE7535	-----A-C-----TC-A-A-C-----A-----G-T-----G-T-----C-T-----GC-----	3128
A.SE.SE8538	-----A-C-----TC-A-A-C-----C-----C-----C-T-----T-----GC-----	3110
A.SE.SE8891	-----A-A-----TC-A-A-C-----C-----C-----C-T-----T-----A-----GC-----	3100
A.SE.UGSE8131	-----T-----TC-A-A-C-----C-----C-----C-T-----G-T-----C-T-----GC-----	3283
A1.KE.O23-17	-----A-A-----TC-A-A-C-----A-----G-----A-----C-----T-----A-----GC-----	3369
A1.SE.SE7253	-----A-A-----TC-A-A-C-----C-----C-----C-T-----G-T-----C-T-----A-----GC-----	3115
A1.UG.92UG037	-----A-----TC-A-A-C-----C-----C-----C-----C-----C-T-----G-T-----GC-----	3278
A1.UG.U455	-----A-----C-----C-----C-----C-----C-----C-----C-----T-----A-----GC-----	3357
A2.CD.97CDKTB48	-----TC-A-A-C-----A-C-----C-----C-----C-----T-----T-----A-----C-----	3249
A2.CY.94CY017.41	-----TC-A-A-C-----G-----C-----C-----C-----C-----T-----T-----A-----C-----	3268
CONSENSUS_B	-----C-----C-----T-----T-----T-----T-----T-----T-----t-----	3455
B.AU.MBC200	-----C-----C-----T-----T-----T-----T-----T-----T-----T-----T-----	3910
B.CN.RL42	-----A-C-----T-----T-----T-----T-----T-----T-----T-----T-----	3298
B.DE.D31	-----C-----C-----G-----C-----C-----C-----C-----C-----T-----T-----	3457
B.DE.HAN	-----C-----C-----C-----C-----C-----C-----C-----C-----T-----T-----	3377
B.ES.89SP061	-----C-----C-----A-----A-----A-----A-----T-----T-----T-----T-----	3469
B.GA.OYL	-----A-----C-----A-----A-----C-----C-----T-----T-----C-----T-----	3454
B.GB.MANC	-----C-----C-----A-----A-----A-----A-----T-----T-----C-----T-----	3857
B.NL.3202A21	-----C-----C-----G-----G-----T-----T-----T-----T-----T-----T-----	3913
B.TW.TWCYS	-----C-----C-----A-----A-----T-----T-----T-----T-----T-----T-----	3911
B.US.JRFL	-----G-----C-----C-----C-----T-----T-----T-----T-----T-----C-----	3228
B.US.MNCG	-----C-----C-----A-----A-----A-----A-----T-----T-----T-----T-----AA-----	3926
B.US.RF	-----C-----C-----A-----A-----A-----A-----T-----T-----T-----T-----	3428
B.US.SF2	-----C-----C-----A-----A-----A-----A-----T-----T-----T-----T-----	3918
B.US.YU2	-----C-----C-----A-----A-----A-----A-----T-----T-----T-----T-----	3909
CONSENSUS_C	-----T-----C-----A-----A-----C-----g-----C-----A-----T-----a-----a-----	3234
C.BR.92BR025	-----A-----C-----C-----A-----C-----C-----C-----T-----T-----A-----T-----A-----G-----	3249
C.BR.98BR004	-----A-----T-----C-----A-----A-----C-----C-----T-----T-----A-----A-----A-----G-----	3236
C.BW.96BW0502	-----A-----T-----C-----A-----A-----C-----C-----T-----T-----A-----A-----A-----G-----	3404
C.BW.96BW06	-----A-----C-----C-----A-----A-----G-----G-----T-----T-----A-----A-----C-----C-----	3893
C.BW.96BW15C02	-----A-----C-----C-----A-----A-----C-----C-----T-----T-----A-----A-----A-----A-----	3378
C.ET.ETH2220	-----A-----C-----C-----A-----C-----C-----C-----T-----T-----A-----A-----A-----C-----	3303
C.IL.98IS002	-----A-----C-----A-----A-----C-----C-----C-----C-----T-----T-----A-----A-----C-----	3245
C.IN.93IN101	-----T-----C-----A-----A-----C-----C-----T-----T-----A-----A-----A-----G-----	3884
C.IN.94IN476	-----A-----C-----A-----A-----C-----A-----C-----A-----T-----T-----A-----C-----A-----	3231
C.IN.98IN012	-----A-----C-----C-----A-----A-----C-----C-----T-----T-----A-----A-----A-----G-----	3236
C.TZ.98TZ013	-----T-----C-----C-----A-----A-----C-----C-----T-----T-----A-----A-----A-----C-----	3254
C.TZ.98TZ017	-----T-----C-----C-----C-----A-----C-----C-----T-----T-----A-----A-----A-----G-----	3265
C.ZA.97ZA012	-----T-----C-----C-----A-----A-----C-----C-----T-----T-----A-----A-----A-----G-----	3245
C.ZM.96ZM651	-----T-----C-----C-----A-----A-----C-----C-----T-----T-----A-----A-----A-----G-----	3240
C.ZM.96ZM751	-----T-----C-----C-----A-----A-----C-----C-----T-----T-----A-----A-----A-----G-----	3250
CONSENSUS_D	-----a-----C-----T-----a-----a-----G-----G-----T-----T-----T-----T-----T-----A-----	3449
D.CD.84ZRO85	-----A-----C-----T-----T-----A-----A-----G-----G-----T-----T-----T-----T-----T-----AA-----	3433
D.CD.FLI	-----A-----C-----T-----T-----A-----A-----G-----G-----T-----T-----T-----T-----T-----AA-----	3457
D.CD.Z266	-----A-----C-----T-----T-----A-----A-----G-----G-----T-----T-----T-----T-----T-----AA-----	3914
D.KE.MB2059	-----A-----C-----T-----T-----A-----A-----G-----G-----T-----T-----T-----T-----T-----AA-----	3915
D.UG.94UG1141	-----A-----C-----T-----T-----A-----A-----G-----G-----T-----T-----T-----T-----T-----AA-----	3266
CONSENSUS_F1	-----?-----C-----TC-----a-----C-----C-----T-----A-----C-----G-----?-----T-----T-----A-----C-----AA-----	3194
F1.BE.VI850	-----A-----C-----TC-----C-----TC-----A-----C-----G-----C-----C-----C-----T-----T-----A-----C-----AA-----	3251
F1.BR.93BR020.1	-----C-----C-----TC-----A-----C-----T-----T-----A-----C-----G-----C-----C-----T-----T-----A-----C-----AA-----	3239
F1.FI.FIN9363	-----A-----C-----TC-----A-----C-----T-----T-----A-----A-----C-----G-----C-----C-----T-----T-----A-----C-----AA-----	3240
F1.FR.MF411	-----A-----C-----TC-----A-----C-----T-----T-----A-----A-----C-----G-----C-----C-----T-----T-----A-----C-----AA-----	3113
F2.CM.MP255	-----A-----C-----TC-----A-----A-----A-----C-----T-----T-----A-----A-----C-----C-----C-----	3107
F2.CM.MP257	-----A-----C-----TC-----A-----A-----A-----C-----T-----T-----A-----A-----C-----C-----C-----	3119
CONSENSUS_G	-----C-----C-----TC-----a-----a-----T-----?-----A-----C-----C-----CC-----T-----T-----A-----A-----G-----	3272
G.BE.DRCBL	-----A-----C-----TC-----C-----C-----C-----C-----T-----T-----A-----C-----CC-----T-----T-----A-----A-----G-----	3868
G.FI.HH8793-12.1	-----T-----C-----C-----TC-----A-----A-----G-----T-----G-----A-----C-----C-----C-----T-----T-----A-----A-----G-----	3311
G.NG.92NG083	-----C-----C-----TC-----A-----C-----C-----C-----T-----T-----A-----A-----C-----C-----T-----T-----A-----A-----G-----	3280
G.SE.SE6165	-----C-----C-----TC-----A-----A-----G-----T-----G-----A-----C-----CC-----T-----T-----A-----A-----G-----	3308
H.BE.VI991	-----C-----C-----ATC-----A-----T-----T-----T-----A-----A-----C-----G-----A-----T-----T-----A-----A-----G-----	3300
H.BE.VI997	-----A-----C-----C-----ATC-----A-----T-----T-----T-----A-----A-----C-----G-----A-----T-----T-----A-----A-----	3235
H.CF.90CF056	-----A-----C-----C-----ATC-----A-----T-----T-----T-----A-----A-----C-----G-----A-----T-----T-----A-----A-----	3258
J.SE.SE7022	-----C-----C-----TC-----C-----TC-----A-----A-----A-----A-----G-----A-----G-----T-----T-----A-----A-----G-----	3226
J.SE.SE7887	-----C-----C-----TC-----C-----TC-----A-----A-----A-----A-----G-----A-----G-----T-----T-----A-----A-----G-----	3225
K.CD.EOTB11C	-----C-----A-----C-----C-----T-----T-----C-----A-----C-----C-----T-----T-----A-----A-----CAG-----	3107
K.CM.MP535	-----C-----A-----C-----T-----T-----C-----A-----C-----C-----T-----T-----C-----T-----A-----AAG-----	3107

(RT is heterodimer of p51 and p66)
p51 end p66 RT continue / Pol p15 RNase H start

Table with columns for accession numbers (e.g., B.FR.HXB2, CONSENSUS_01_AE), sequence alignments (A, C, G, T, -, ?), and position numbers (3911, 3862, 3860, 3947, 3847, 3864, 3856, 3898, 3874, 3875, 3485, 3119, 3067, 3259, 3260, 3127, 3436, 3283, 3108, 3105, 3680, 3281, 3277, 3928, 3925, 3293, 3266, 3878, 3939, 3928, 3154, 3970, 3243, 3243, 3101, 3090, 3091, 3097, 3091, 3243, 3290, 3288, 3279, 3834, 3914, 3905, 3962, 3213, 3916, 3937, 3934, 3503, 3505, 3803, 3966, 3941, 3965, 3965, 3389, 3350, 3322, 3613, 3972, 3971, Pol). The table shows sequence alignments for various HIV-1/SIVcpz strains and consensus sequences.

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz complete genomes

B.FR.HXB2
 CONSENSUS_A-A1
 A.BY.97BL006
 A.SE.SE7535
 A.SE.SE8538
 A.SE.SE8891
 A.SE.UGSE8131
 A1.KE.Q23-17
 A1.SE.SE7253
 A1.UG.92UG037
 A1.UG.U455
 A2.CD.97CDKTB48
 A2.CY.94CY017.41
 CONSENSUS_B
 B.AU.MBC200
 B.CN.RL42
 B.DE.D31
 B.DE.HAN
 B.ES.89SP061
 B.GA.OYI
 B.GB.MANC
 B.NL.3202A21
 B.TW.TWCYS
 B.US.JREL
 B.US.MNCG
 B.US.RF
 B.US.SF2
 B.US.YU2
 CONSENSUS_C
 C.BR.92BR025
 C.BR.98BR004
 C.BW.96BW0502
 C.BW.96BW15C02
 C.ET.ETH2220
 C.IL.98ILS002
 C.IN.93IN101
 C.IN.94IN476
 C.IN.98IN012
 C.TZ.98TZ013
 C.TZ.98TZ017
 C.ZA.97ZA012
 C.ZM.96ZM651
 C.ZM.96ZM751
 CONSENSUS_D
 D.CD.84ZR085
 D.CD.FLI
 D.CD.Z2Z6
 D.KE.MB2059
 D.UG.94UG1141
 CONSENSUS_F1
 F1.BE.VI850
 F1.BR.93BR020.1
 F1.FI.FIN9363
 F1.FR.MF411
 F2.CM.MP255
 F2.CM.MP257
 CONSENSUS_G
 G.BE.DRCBL
 G.FI.HH8793-12.1
 G.NG.92NG083
 G.SE.SE6165
 H.BE.VI991
 H.BE.VI997
 H.CF.90CF056
 J.SE.SE7022
 J.SE.SE7887
 K.CD.EQTB11C
 K.CM.MP535
 B.FR.HXB2
 CGAGGATGTTACTAATAGAGGAAACAAAAAGTTGTCAACCTTAAGTACACA...ACAAATCAGAAGACTGAGTTTACAGCAATTTATCTAGCTTTTCAGGATTCGGGATTAGAAGTAAACATAGTA...AC 4039
 ---g---c---g---c---Tt---T---G---G---?---G---CA---G---A---A---T---cc---T---?---?---?---A---C---3363
 ---G---G---CT---A---G---TC---T---G---CA---G---A---A---T---CC---T---G---T---G---T---A---3413
 ---G---C---C---G---C---G---G---TT---T---A---A---AG---T---AC---T---CT---C---A---C---G---T---3426
 ---G---C---G---C---G---C---G---T---T---G---G---G---A---A---AG---T---AC---G---T---CC---A---C---G---T---3439
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3411
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3411
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3497
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3243
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3406
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3485
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3377
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3396
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3583
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---4038
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3426
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3585
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3505
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3597
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3582
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---4041
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---4039
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3356
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---4054
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3556
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---4046
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---4037
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3362
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3377
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3364
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---5332
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---4021
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3506
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3431
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3373
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---4012
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3359
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3364
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3382
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3392
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3373
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3368
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3378
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3576
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3561
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3585
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---4042
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---4043
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3394
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3319
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3379
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3367
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3241
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3235
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3247
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3400
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3996
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3439
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3408
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3436
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3428
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3363
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3386
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3354
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3353
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3235
 ---G---C---G---C---G---G---AG---G---A---T---T---G---G---G---A---A---A---T---CC---C---A---C---3235

B_FR_HXB2
CONSENSUS_01_AE
01_AE_CF_90CF11697
01_AE_CF_90CF402
01_AE_CF_90CF4071
01_AE_TH_93TH057
01_AE_TH_93TH065
01_AE_TH_93TH253
01_AE_TH_95TH022
01_AE_TH_95TH047
01_AE_TH_CM240
CONSENSUS_02_AG
02_AG_CM_97CM-MP807
02_AG_FR_DJ263
02_AG_FR_DJ264
02_AG_GH_G829
02_AG_NG_IBNG
02_AG_SE_SE7812
02_AG_SN_98SE-MP1211
02_AG_SN_MF1213
03_AB_RU_KAL153-2
03_AB_RU_RU98001
04_CPX_CX_94CY032-3
04_CPX_GR_97PVGH
04_CPX_GR_97PVNY
05_DF_BE_V11310
05_DF_BE_V1961
CONSENSUS_06_CPX
06_CPX_AU_BFP90
06_CPX_ML_95ML127
06_CPX_ML_95ML84
06_CPX_SN_97SEL078
07_BC_CN_97CN001
07_BC_CN_98CN009
07_BC_CN_C54A
CONSENSUS_08_BC
08_BC_CN_97CNGX6F
08_BC_CN_97CNGX7F
08_BC_CN_97CNGX9F
08_BC_CN_98CN006
10_CD_TZ_96TZBF061
10_CD_TZ_96TZBF071
10_CD_TZ_96TZBF110
CONSENSUS_11_CPX
11_CPX_CM_MP818
11_CPX_FR_MP1298
11_CPX_FR_MP1307
11_CPX_GR_GRL17
12_BF_AR_ARM159
12_BF_UY_URTR23
12_BF_UY_URTR35
N_CM_YBF106
N_CM_YBF30
CONSENSUS_0
O_CM_ANT70
O_CM_MVP5180
O_SN_SEMP1299
O_SN_SEMP1300
CONSENSUS_CPZ
CPZ_CD_CPFZANT
CPZ_CM_CAM3
CPZ_CM_CAM5
CPZ_GA_CPFZGAB
CPZ_US_CPFZUS
Pol p15 RNase H

HIV-1/SIVcpz complete genomes

B_FR_HXB2
AGACTCACAATATGCATTAGGAATCAATT...CAAGCACAAACAGATCAAAAGTGAATCAGAGTTAGTCAATAAATAGAGCAGGTTAATAAAAAAGGAAAGGTTCTATCTG...GCATGGTACCAGCACACA 4167

CONSENSUS_01_AE
01_AE_CF_90CF11697
01_AE_CF_90CF402
01_AE_CF_90CF4071
01_AE_TH_93TH057
01_AE_TH_93TH065
01_AE_TH_93TH253
01_AE_TH_95TH022
01_AE_TH_95TH047
01_AE_TH_CM240

CONSENSUS_02_AG
02_AG_CM_97CM_MP807
02_AG_FR_DJ263
02_AG_FR_DJ264
02_AG_GH_G829
02_AG_NG_IBNG
02_AG_SE_SE7812
02_AG_SN_98SE_MP1211
02_AG_SN_MP1213

03_AB_RU_KAL153-2
03_AB_RU_RU98001

04_CFX_CX_94CY032-3
04_CFX_GR_97PVGH
04_CFX_GR_97PVNY

05_DF_BE_V11310
05_DF_BE_V1961

CONSENSUS_06_CPX
06_CPX_AU_BFP90
06_CPX_ML_95ML127
06_CPX_ML_95ML84
06_CPX_SN_97SEL078

07_BC_CN_97CN001
07_BC_CN_98CN009

07_BC_CN_C54A

CONSENSUS_08_BC
08_BC_CN_97CNGX6F
08_BC_CN_97CNGX7F
08_BC_CN_97CNGX9F
08_BC_CN_98CN006

10_CD_TZ_96TZBF061
10_CD_TZ_96TZBF071
10_CD_TZ_96TZBF110

CONSENSUS_11_CPX
11_CPX_CM_MP818
11_CPX_FR_MP1298
11_CPX_FR_MP1307
11_CPX_GR_GR17

12_BF_AR_ARM159
12_BF_UY_URTR23
12_BF_UY_URTR35

N_CM_YBF106
N_CM_YBF30

CONSENSUS_0
O_CM_ANT70
O_CM_MVP5180
O_SN_SEMP1299
O_SN_SEMP1300

CONSENSUS_CPZ
CPZ_CD_CFPZANT
CPZ_CM_CAM3
CPZ_CM_CAM5
CPZ_GA_CFPZGAB
CPZ_US_CFPZUS

Po1_p15_RNase H



HIV-1/SIVcpz
complete genomes

B.FR.HXB2 Pol p66 RT end \ Pol p31 Integrase start
AAGGAAATTGGAGAAATGAACAAGTAGATAAAATTAGTCAGTCTGGG...ATCAGAAAGTACTATTATTTTGTAGATGAATAGATAAGCCCAAGAT...GAACATGAGAAATAT...CACAGTAATTGGAGCAAT 4294

CONSENSUS_01_AE 4245
01_AE.CF.90CF11697 4243
01_AE.CF.90CF402 4243
01_AE.CF.90CF4071 4230
01_AE.TH.93TH057 4230
01_AE.TH.93TH065 4247
01_AE.TH.93TH253 4239
01_AE.TH.95TH022 4281
01_AE.TH.95TH047 4257
01_AE.TH.CM240 4257
01_AE.TH.CM240 3868

CONSENSUS_02_AG 3502
02_AG.CM.97CM-MP807 3502
02_AG.FR.DJ263 3450
02_AG.FR.DJ264 3642
02_AG.GH.G829 3643
02_AG.NG.IBNG 3510
02_AG.SE.SE7812 3819
02_AG.SN.96SE-MP1211 3666
02_AG.SN.MP1213 3491
02_AG.SN.MP1213 3488

03_AB.RU.KAL153-2 4063
03_AB.RU.RU98001 3664

04_CPX.CV.94CV032-3 3660
04_CPX.GR.97PVGH 4311
04_CPX.GR.97PVNY 4308

05_DF.BE.VI1310 3676
05_DF.BE.VI1961 3649

CONSENSUS_06_CPX 4259
06_CPX.AU.BFP90 4322
06_CPX.ML.95ML127 4311
06_CPX.ML.95ML84 3537
06_CPX.SN.97SE1078 4353

07_BC.CN.97CN001 3626
07_BC.CN.98CN009 3626
07_BC.CN.C54A 3484

CONSENSUS_08_BC 3473
08_BC.CN.97CNGX6F 3474
08_BC.CN.97CNGX7F 3480
08_BC.CN.97CNGX9F 3474
08_BC.CN.98CN006 3626

10_CD.TZ.96TZBF061 3673
10_CD.TZ.96TZBF071 3671
10_CD.TZ.96TZBF110 3662

CONSENSUS_11_CPX 4207
11_CPX.CM.MP818 4287
11_CPX.FR.MP1298 4288
11_CPX.FR.MP1307 4345
11_CPX.GR.GR17 3596

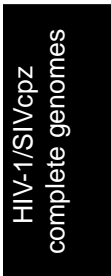
12_BF.AR.ARMA159 4299
12_BF.UY.URTR23 4320
12_BF.UY.URTR35 4317

N.CM.YBF106 3886
N.CM.YBF30 3888

CONSENSUS_O 4179
O.CM.ANT170 4179
O.CM.MVP5180 4349
O.SN.SEMP1299 4324
O.SN.SEMP1300 4348

CONSENSUS_CPZ 4348
CPZ.CD.CPZANT 3763
CPZ.CM.CAM3 3733
CPZ.CM.CAM5 3996
CPZ.GA.CPZGAB 4355
CPZ.U5.CPZUS 4354

K_G_I_G_G_N_E_Q_V_D_K_L_V_S_A_G...I_R_K_V_L_F_L_D_G_I_D_K_A_Q_D...E_H_F_K_Y...H_S_N_W_R_A_M Pol
Pol p66 RT end \ Pol p31 Integrase start



B_FR_HXB2
 GGCTAGTGAATTTAAACCTGCCACCTGTAGTAGCAAAA...GAAATAGTAGCCAGCTGTGATAAATGT...CAGCTAAAGGAGAGAGCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGAT 4421

CONSENSUS_01_AE
 01_AE_CF_90CF11697
 01_AE_CF_90CF402
 01_AE_CF_90CF4071
 01_AE_TH_93TH057
 01_AE_TH_93TH065
 01_AE_TH_93TH253
 01_AE_TH_95TH022
 01_AE_TH_95TH047
 01_AE_TH_CM240

CONSENSUS_02_AG
 02_AG_CM_97CM-MP807
 02_AG_FR_DJ263
 02_AG_FR_DJ264
 02_AG_GH_G829
 02_AG_NG_IBNG
 02_AG_SE_SE7812
 02_AG_SN_98SE-MP1211
 02_AG_SN_MP1213

03_AB_RU_KAL153-2
 03_AB_RU_RU98001

04_CPX_CX_94CY032-3
 04_CPX_GR_97PVGH
 04_CPX_GR_97PVNY

05_DF_BE_V11310
 05_DF_BE_V1961

CONSENSUS_06_CPX
 06_CPX_AU_BFP90
 06_CPX_ML_95ML127
 06_CPX_ML_95ML84
 06_CPX_SN_97SEL078

07_BC_CN_97CN001
 07_BC_CN_98CN009
 07_BC_CN_C54A

CONSENSUS_08_BC
 08_BC_CN_97CNGX6F
 08_BC_CN_97CNGX7F
 08_BC_CN_97CNGX9F
 08_BC_CN_98CN006

10_CD_TZ_96TZBF061
 10_CD_TZ_96TZBF071
 10_CD_TZ_96TZBF110

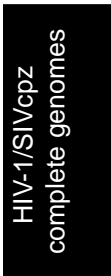
CONSENSUS_11_CPX
 11_CPX_CM_MP818
 11_CPX_FR_MP1298
 11_CPX_FR_MP1307
 11_CPX_GR_GR17

12_BF_AR_ARMA159
 12_BF_UY_URTR23
 12_BF_UY_URTR35
 N_CM_YBF106
 N_CM_YBF30

CONSENSUS_0
 O_CM_ANT70
 O_CM_MVP5180
 O_SN_SEMP1299
 O_SN_SEMP1300

CONSENSUS_CPZ
 CPZ_CD_CPFZANT
 CPZ_CM_CAM3
 CPZ_CM_CAM5
 CPZ_GA_CPFZGAB
 CPZ_US_CPFZUS

PoI_p31 Integrase



HIV-1/SIVcpz complete genomes

Table with columns for sample ID (e.g., B.FR.HXB2, CONSENSUS_A-A1), sequence coordinates (4681-4070), and the corresponding nucleotide sequence. The table lists various HIV-1 and SIVcpz strains and their complete genome sequences.

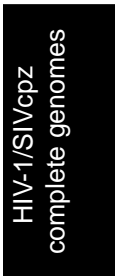
4681	B_FR_HXB2	GGCCAGTAAACAAATACATACATACTGACAAATGGCAGCAATTTCCACGGTGTCTACGGTTAGGGCCCGCTTGTGGTGGGGGGAATCAACAGGAAATTTGGAAATTCCTTACAATCCCAAAAGTCAAGGAGTAGT	4681
4632	CONSENSUS_01_AE	-----GT-----C-A-----C-T-----A-C-g-a--AA-a-----CAATG-Cga-----g-c-----	4632
4630	01_AE_CF_90CF11697	-----GGTC-----C-A-----A-C-----G-TA-G-AA-A-----CAATG-C-A-----G-----	4630
4617	01_AE_CF_90CF402	-----GGTC-----C-A-----A-C-----G-A-----CAATG-C-A-----C-----	4617
4617	01_AE_CF_90CF4071	-----GGT-----C-A-----A-----A-----CAATG-C-A-----G-----	4617
4634	01_AE_TH_93TH057	-----GT-----C-A-----C-T-----A-C-G-T-----AA-A-----CAATG-CGA-----G-----	4634
4626	01_AE_TH_93TH065	-----GT-----C-A-----C-T-----A-C-G-A-----AA-G-----CAATG-CGA-----G-----	4626
4668	01_AE_TH_93TH253	-----GT-----C-A-----C-T-----A-C-G-A-----AA-A-----CAATG-CGA-----G-----	4668
4644	01_AE_TH_95TH022	-----GT-G-----C-A-----C-T-----A-C-G-A-----AA-A-----CAATG-CG-----C-----	4644
4644	01_AE_TH_95TH047	-----GT-G-----C-A-----C-T-----A-C-G-A-----AA-A-----CAATG-CGA-----G-----	4644
4255	01_AE_TH_CM240	-----GT-----C-A-----C-T-----A-C-G-A-----AA-A-----CAATG-C-A-----G-----	4255
3889	CONSENSUS_02_AG	-----g-----GT-----C-a-----A-----G-A-A-a-----AAATG-CA-a-----c-----	3889
3837	02_AG_CM_97CM-MP807	-----G-----GT-----C-a-----A-----G-A-A-a-----AAATG-CA-a-----c-----	3837
4029	02_AG_FR_DJ263	-----G-----GT-----C-a-----A-----A-----A-----AAATG-CA-a-----c-----	4029
4030	02_AG_FR_DJ264	-----G-----GT-----C-a-----A-----A-----A-----AAATG-CA-a-----c-----	4030
3897	02_AG_GH_G829	-----G-----GT-----C-a-----A-----A-----A-----AAATG-CA-a-----c-----	3897
4206	02_AG_NG_IBNG	-----G-----GT-----C-a-----A-----A-----A-----AAATG-CA-a-----c-----	4206
4053	02_AG_SE_SE7812	-----G-----GT-----C-a-----A-----A-----A-----AAATG-CA-a-----c-----	4053
3878	02_AG_SN_98SE-MP1211	T-G-CAGT-----GT-G-----C-A-----G-A-A-A-----A-----CAAG-T-CA-CA-G-A-----G-A-T-----	3878
3875	02_AG_SN_MP1213	-----G-----GT-----C-G-----T-----A-----G-A-A-A-----A-----AAATG-----C-----	3875
4450	03_AB_RU_KAL153-2	-----T-----A-----A-----A-----G-----A-----A-----T-----A-----G-----	4450
4047	03_AB_RU_RU98001	-----GT-----A-----A-----T-----A-----A-----G-----A-----A-----G-----	4047
4698	04_CX_CX_94CY032-3	-----G-----TG-----G-A-----C-----CC-----A-----G-----A-----A-----A-----AT-----C-----	4698
4695	04_CX_GR_97PVNY	-----T-----T-----A-----A-----G-----A-----G-----A-----A-----A-----ATG-----C-----	4695
4063	05_DF_BE_V11310	-----TG-----A-----T-----A-----T-----A-----G-----G-----A-----A-----A-----T-----C-----	4063
4037	05_DF_BE_V1961	-----TGG-----A-----T-----A-----T-----A-----G-----A-----A-----T-----A-----	4037
4643	CONSENSUS_06_CX	-----GTG-----C-----A-----T-----A-----A-----AA-----A-----AAAT?-----CA-----A-----	4643
4709	06_CX_AU_BFP90	-----GTG-----C-----A-----T-----A-----A-----AA-----A-----AAAT-----CA-----A-----T-----	4709
4698	06_CX_ML_95ML127	-----GTG-----C-----A-----T-----A-----G-----A-----AA-----A-----AAAT-----CA-----A-----	4698
3924	06_CX_ML_95ML84	-----GTG-----C-----A-----T-----A-----A-----G-----A-----AA-----A-----AAATG-----CA-----A-----	3924
4740	06_CX_SN_97SEL078	-----GTG-----C-----A-----T-----A-----G-----A-----G-----AA-----A-----AAATG-----CA-----A-----	4740
4013	07_BC_CN_97CN001	-----C-----GT-----A-----A-----T-----T-----A-----A-----G-----A-----A-----A-----T-----C-----	4013
4013	07_BC_CN_98CN009	-----C-----GT-----A-----A-----T-----T-----A-----A-----G-----A-----A-----A-----T-----C-----	4013
3871	07_BC_CN_C54A	-----C-----GT-----A-----A-----T-----T-----A-----A-----G-----A-----A-----A-----T-----C-----	3871
3860	CONSENSUS_08_BC	-----C-----GT-----T-----T-----A-----A-----G-----A-----A-----A-----T-----C-----	3860
3861	08_BC_CN_97CNGX6F	-----C-----GT-----T-----T-----A-----A-----G-----A-----A-----A-----T-----C-----	3861
3867	08_BC_CN_97CNGX7F	-----C-----GT-----T-----T-----A-----A-----G-----A-----A-----A-----T-----C-----	3867
3861	08_BC_CN_97CNGX9F	-----C-----GT-----T-----T-----A-----A-----G-----A-----A-----A-----T-----C-----	3861
4013	08_BC_CN_98CN006	-----C-----GT-----T-----T-----A-----A-----G-----A-----A-----A-----T-----C-----	4013
4060	10_CD_TZ_96TZBF061	-----GT-----G-----A-----A-----T-----T-----A-----A-----C-----G-----A-----A-----	4060
4058	10_CD_TZ_96TZBF071	-----GT-----G-----A-----A-----T-----T-----A-----A-----C-----G-----A-----A-----	4058
4049	10_CD_TZ_96TZBF110	-----GT-----G-----A-----A-----T-----T-----A-----A-----C-----G-----A-----A-----	4049
4589	CONSENSUS_11_CX	-----gT-----A-----t-----A-----C?-----g-----A-----AA-----A-----AaaT-----C-----A-----g-----	4589
4684	11_CX_CM_MP818	-----TG-----A-----T-----A-----C-----G-----A-----AA-----A-----AAAT-----C-----A-----G-----	4684
4732	11_CX_FR_MP1298	-----GT-----A-----T-----A-----C-----G-----A-----AA-----A-----AAAT-----C-----A-----G-----	4732
3983	11_CX_FR_MP1307	-----GT-----A-----T-----A-----C-----G-----A-----AA-----A-----AAAT-----C-----A-----G-----	3983
4686	12_BF_AR_ARM159	-----G-----T-----A-----T-----A-----C-----T-----A-----A-----A-----T-----C-----A-----	4686
4707	12_BF_UY_URTR23	-----G-----T-----A-----T-----A-----C-----G-----A-----A-----A-----T-----C-----A-----	4707
4273	N_CM_YBF106	-----GT-----T-----C-----T-----A-----A-----A-----A-----A-----T-----A-----A-----C-----G-----	4273
4275	N_CM_YBF30	-----GT-----T-----C-----T-----A-----A-----A-----A-----A-----T-----A-----A-----C-----G-----	4275
4554	CONSENSUS_0	-----T-----t-----gt-----?-----c-----t-----aa-----a-----?-----aa-----a-----c-----c-----g-----	4554
4736	O_CM_ANT70	-----T-----T-----GT-----C-----C-----T-----T-----AA-----A-----A-----TA-----G-----A-----T-----A-----	4736
4711	O_CN_MVP5180	-----T-----T-----GT-----C-----C-----T-----T-----AA-----A-----A-----AG-----A-----A-----A-----	4711
4735	O_SN_SEMP1299	-----T-----T-----GT-----C-----C-----T-----T-----AA-----A-----A-----AG-----A-----A-----A-----	4735
4735	O_SN_SEMP1300	-----T-----T-----GT-----C-----C-----T-----T-----AA-----A-----A-----AG-----A-----A-----A-----	4735
4139	CONSENSUS_CPZ	-----t-----gt-----?-----c-----t-----aa-----a-----?-----aa-----a-----c-----c-----g-----	4139
4120	CPZ_CD_CFPZANT	-----T-----GT-----C-----C-----T-----T-----AA-----A-----A-----AG-----A-----A-----A-----	4120
4095	CPZ_CM_CAM3	-----T-----GT-----C-----C-----T-----T-----AA-----A-----A-----AG-----A-----A-----A-----	4095
4383	CPZ_CM_CAM5	-----T-----T-----C-----A-----T-----T-----CCCA-----T-----AA-----G-----A-----C-----A-----T-----	4383
4742	CPZ_GA_CFPZGAB	-----T-----T-----C-----A-----T-----T-----ATCT-----C-----T-----AA-----AG-----A-----A-----T-----C-----	4742
4741	CPZ_US_CFPZUS	-----T-----T-----C-----A-----T-----T-----ATCT-----C-----T-----AA-----AG-----A-----A-----T-----C-----	4741

Pol p31 Integrase W_P_V_K_T_I_H_T_D_N_G_S_N_F_T_T_G_A_T_V_R_A_A_C_W_A_G_I_K_Q_E_F_G_I_P_Y_N_P_O_S_Q_G_V_V



Table with columns for accession numbers (e.g., B_FR_HXB2, CONSENSUS_01_AE), sequence alignments (AGAACTCTATGAAT...), and positions (4810, 4761, 4759, 4846, 4746, 4763, 4753, 4797, 4773, 4384, 4018, 3966, 4158, 4159, 4026, 4335, 4182, 4007, 4004, 4579, 4180, 4176, 4827, 4824, 4192, 4166, 4771, 4838, 4827, 4053, 4869, 4142, 4189, 4187, 4178, 4718, 4813, 4804, 4861, 4112, 4815, 4836, 4833, 4402, 4404, 4680, 4865, 4840, 4864, 4864, 4267, 4249, 4224, 4512, 4871, 4870). The table shows sequence alignments for various HIV-1/SIVcpz strains, with gaps indicated by dashes.

Po1 p31 Integrase



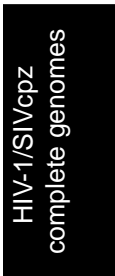
49 39

B_FR_HXB2
 CAGTGCAGGGGAAAGTATAGCATATAGCAACAGACATACAAACTAAAGAAATACAAAACAAAATTTACAAAA...ATTCAAAAATTTTCGGGTTTATTACAGGGACAGAGAAATCCACITTTGGAAAGG
 ---A-----A-----A-----Y-----C-----C-----A-----G-C-A-
 01_AE_CF_90CF11697
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 01_AE_CF_90CF402
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 01_AE_CF_90CF4071
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 01_AE_TH_93TH057
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 01_AE_TH_93TH065
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 01_AE_TH_93TH253
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 01_AE_TH_95TH022
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 01_AE_TH_95TH047
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 01_AE_TH_CM240
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 CONSENSUS_02_AG
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 02_AG_CM_97CM_MP807
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 02_AG_FR_DJ263
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 02_AG_FR_DJ264
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 02_AG_GH_G829
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 02_AG_NG_IBNG
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 02_AG_SE_SE7812
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 02_AG_SN_98SE_MP1211
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 02_AG_SN_MP1213
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 03_AB_RU_KAL153-2
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 03_AB_RU_RU98001
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 04_CPX_CX_94CY032-3
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 04_CPX_GR_97PVGH
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 04_CPX_GR_97PVNY
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 05_DF_BE_VI1310
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 05_DF_BE_VI1961
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 CONSENSUS_06_CPX
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 06_CPX_AU_BFP90
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 06_CPX_ML_95ML127
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 06_CPX_ML_95ML84
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 06_CPX_SN_97SEL078
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 07_BC_CN_97CN001
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 07_BC_CN_98CN009
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 07_BC_CN_C54A
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 CONSENSUS_08_BC
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 08_BC_CN_97CNGX6F
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 08_BC_CN_97CNGX7F
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 08_BC_CN_97CNGX9F
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 08_BC_CN_98CN006
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 10_CD_TZ_96TZBF061
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 10_CD_TZ_96TZBF071
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 10_CD_TZ_96TZBF110
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 CONSENSUS_11_CPX
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 11_CPX_CM_MP818
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 11_CPX_FR_MP1298
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 11_CPX_FR_MP1307
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 11_CPX_GR_GR17
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 12_BF_AR_ARMA159
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 12_BF_UY_URTR23
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 12_BF_UY_URTR35
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 N_CM_YBF106
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 N_CM_YBF30
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 CONSENSUS_0
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 O_CM_ANT70
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 O_CM_MVP5180
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 O_SN_SEMP1299
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 O_SN_SEMP1300
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 CONSENSUS_CPZ
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 CPZ_CD_CFPZANT
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 CPZ_CM_CAM3
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 CPZ_CM_CAM5
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 CPZ_GA_CFPZGAB
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 CPZ_US_CFPZUS
 ---A-----A-----A-----C-----C-----A-----G-C-A-
 Pol p31 Integrase
 S_A_G_E_R_I_V_D_I_I_A_T_D_I_Q_T_K_E_L_Q_K_Q_I_T_K_I_Q_N_F_R_V_Y_R_D_S_R_N_P_L_W_K_G

4890
4888
4975
4875
4892
4882
4926
4902
4902
4513
4147
4095
4287
4288
4155
4464
4311
4136
4130
4708
4309
4305
4956
4953
4321
4296
4900
4967
4956
4182
4998
4271
4271
4129
4118
4119
4125
4119
4271
4318
4316
4307
4846
4942
4933
4990
4241
4944
4965
4962
4531
4533
4805
4994
4969
4993
4993
4395
4378
4353
4641
5000
4999
Pol

HIV-1/SIVcpz complete genomes

B.FR.HXB2 / Vif start
 ACCAGCAAAGCTCCTCTGGAAGGTCGAGGCGAGTGTAAATACAAGATAAATAGTGCATATAAAAGTAGTCCCAAGAAAGAAAGCAAAAGATCATTAGGGATTATGGAAAAACAGATGGCAGGTGATGATTGT
 -----a--A-----C-----T-----T-----a-----
 01.AE.CF.90CF11697 -----A--A-----C-----T-----T-----A-----
 01.AE.CF.90CF402 -----A--A-----G-----C-----T-----A-----
 01.AE.CF.90CF4071 -----A--A-----C-----T-----T-----A-----
 01.AE.TH.93TH057 -----A--A-----C-----T-----C-----A-----
 01.AE.TH.93TH065 -----A--A-----C-----T-----A-----A-----
 01.AE.TH.93TH253 -----A--A-----C-----T-----A-----A-----
 01.AE.TH.95TH022 -----A--A-----C-----T-----A-----A-----
 01.AE.TH.95TH047 -----C--A--A-----C-----T-----A-----A-----
 01.AE.TH.CM240 -----A--A-----C-----T-----A-----A-----
 CONSENSUS_02_AG -----a--A-----g--C-----T-----G-----A-----A-----
 02.AG.CM.97CM-MP807 -----A--A-----G-----C-----T-----G-----A-----T-----A-----A-----
 02.AG.FR.DJ263 -----A--A-----G-----C-----T-----G-----A-----A-----AA-----
 02.AG.FR.DJ264 -----A--A-----G-----C-----T-----G-----A-----A-----AA-----
 02.AG.GH.G829 -----A--A-----G-----C-----T-----G-----A-----A-----C-----
 02.AG.NG.IBNG -----A--A-----G-----C-----T-----G-----A-----A-----C-----
 02.AG.SE.SE7812 -----A--A-----G-----C-----T-----G-----A-----G-----G-----
 02.AG.SN.96SE-MP1211 -----A--A-----G-----C-----T-----G-----A-----A-----C-----
 02.AG.SN.MP1213 GA -----A--A-----G-----C-----T-----G-----A-----A-----C-----
 03.AB.RU.KAL153-2 -----A--A-----G-----C-----AC--T-----A-----A-----
 03.AB.RU.RU98001 -----A--A-----G-----C-----AC--T-----A-----A-----
 04.cpx.CY.94CY032-3 -----A--A-----G-----C-----C-----T-----C-----T-----T-----C-----C-----
 04.cpx.GR.97PVCH -----A--A-----G-----C-----C-----T-----G-----A-----T-----T-----C-----C-----
 04.cpx.GR.97PVNY -----A--A-----G-----C-----C-----T-----G-----A-----A-----A-----C-----C-----
 05.DF.BE.VI1310 -----A--A-----G--A-----C-----C-----C-----T-----A-----G-----A-----
 05.DF.BE.VI1961 -----A--A-----C-----C-----C-----G-----G-----A-----A-----
 CONSENSUS_06_cpx -----A--A-----C-----C-----C-----G-----G-----A-----A-----
 06.cpx.AU.BFP90 -----A--A-----C-----C-----a--b--A-----A-----A-----
 06.cpx.MI.95MI127 -----A--A--T-----C-----G-----A-----G-----A-----A-----
 06.cpx.MI.95MI184 -----A--A-----C-----T-----G-----A-----G-----A-----AA-----
 06.cpx.SN.97SE1078 -----A--A-----C-----C-----A-----T-----G-----A-----A-----
 07.BC.CN.97CN001 -----C--A--A-----G-----A-----G-----G-----A-----A-----A-----C-----
 07.BC.CN.98CN009 -----C--A--A-----G-----A-----G-----G-----A-----A-----A-----C-----
 07_BC.CN.C54A -----C--A--A-----A-----G-----G-----A-----A-----A-----C-----
 CONSENSUS_08_BC -----C--A--A-----G-----A-----G-----G-----A-----A-----A-----C-----
 08.BC.CN.97CNGX6F -----C--A--A-----G-----A-----G-----G-----A-----A-----A-----C-----
 08.BC.CN.97CNGX7F -----C--A--A-----G-----A-----G-----G-----A-----A-----A-----C-----
 08.BC.CN.97CNGX9F -----C--A--A-----G-----A-----G-----G-----A-----A-----A-----C-----
 08_BC.CN.98CN006 -----C--A--A-----G-----A-----G-----G-----A-----A-----A-----C-----
 10.CD.TZ.96TZBF061 -----C--A--A-----A-----C-----G-----A-----G-----G-----T-----A-----A-----C-----
 10.CD.TZ.96TZBF071 -----T-----T-----G-----A-----G-----G-----A-----G-----T-----A-----A-----C-----
 10.CD.TZ.96TZBF110 -----C--A--A-----C-----A-----G-----G-----T-----A-----A-----C-----
 CONSENSUS_11_cpx -----A--a-----g--C-----T-----?-----A-----
 11.cpx.CM.MP818 -----A--A-----G-----C-----T-----G-----A-----
 11.cpx.FR.MP1298 -----A--A-----G-----C-----T-----G-----A-----
 11.cpx.FR.MP1307 -----A--G-----C-----T-----G-----A-----
 11.cpx.GR.GR17 -----A--A-----A-----G-----C-----T-----A-----
 12.BF.AR.ARMA159 -----A-----C-----CG-----A-----G-----A-----C-----
 12.BF.UY.URTR23 -----A-----C-----C-----G-----A-----G-----A-----
 12.BF.UY.URTR35 -----A--G-----C-----C-----G-----A-----G-----A-----
 N.CM.YBF106 -----C--A--T--G-----A-----G-----C-----T-----G-----A-----A-----A-----
 N.CM.YBF30 -----C--A--T--G-----A-----G-----C-----T-----G-----A-----A-----A-----
 CONSENSUS_0 -----G--C-----G--a--a-----T--g--A-----G--A--C--A--C--A--?-----A-----
 O.CM.ANT70 -----G--C-----G--A-----G--A-----T--G--A-----G--A--C--A--G-----A-----
 O.CM.MVP5180 -----G--C-----G--A-----G--A-----T--G--A-----G--A--C--A--G-----A-----
 O.SN.SEMP1299 -----G--C-----G--A-----G--A-----T--G--A-----G--A--C--A--G-----A-----
 O.SN.SEMP1300 -----G--C-----G--A-----G--A-----T--G--A-----G--A--C--A--G-----A-----
 CONSENSUS_CPZ -----?--c--?--a--t--G-----?--ca--g--aG--aG--aG--t-----a--c-----g-----a-----c-----g-----a-----a-----
 CPZ.CD.CPZANT -----C--A--T--G-----C--CA--CC--AGAG--A--T--G-----C--TC--G-----T--A--A--G--A-----A--A--A--AGA-----GA-----
 CPZ.CM.CAM3 -----T--C--GA--T--G-----G--TA--G--GAGAG--AG--T-----A--A--C-----G--A-----A--A--A--T-----
 CPZ.CM.CAM5 -----T--C--GA--T--G-----G--TA--G--GAGAG--AG--T-----A--A--C-----G--T-----A--C--A-----G--A-----
 CPZ.GA.CPZGAB -----G-----CCT--A--G-----G--C--G--C--AG--G--AC--G-----G-----G-----A-----T--A--AA--C-----
 CPZ.US.CPZUS -----C--A--T--G-----CA--G--A--CGAG--GG--T-----A--C-----
 Pol p31 Integrase P_A_K_L_L_W_K_G_E_G_A_V_V_I_Q_D_N_S_D_I_K_V_V_P_R_R_K_A_K_I_I_R_R_D_Y_G_K_Q_M_A_G_D_D_C
 Vif \ M_E_N_R_W_Q_V_M_I_V



B.FR.HXB2	Pol p31 Integrase end \	GTGGCAAGTAGACAGGAGGATTAGAACATGTAATGTTTTCAGGAAAAGCTAGGGGATGGTTTATATAGACATCACTATGAAAGCCCTCATCCAAAGAAATAAGTTTCA	5199
CONSENSUS_A-A1		-----g-----C-----T-----C-----aa-----a-----t-----a-----a-----g-----	4523
A.BY.97BL006		-----T-----T-----C-----AA-----A-----AG-----A-----G-----AG-----G-----	4576
A.SE.SE7535		-----G-----C-----C-----G-----C-----AA-----A-----AAAT-----A-----TAG-----G-----	4416
A.SE.SE8538		-----G-----T-----C-----A-----C-----AA-----A-----AAAAT-----T-----AG-----G-----	4398
A.SE.SE8891		-----G-----T-----C-----A-----C-----CAA-----AA-----T-----C-----AG-----G-----	4388
A.SE.UGSE8131		-----G-----T-----C-----A-----C-----AA-----A-----GCA-----T-----C-----AG-----G-----	4570
A1.KE.Q23-17		-----G-----T-----C-----C-----CAA-----A-----AAA-----G-----AG-----A-----G-----	4657
A1.SE.SE7253		-----G-----T-----C-----C-----A-----AA-----AT-----C-----AGGA-----A-----C-----	4403
A1.UG.92UG037		-----G-----G-----C-----C-----C-----G-----AA-----T-----AGG-----A-----G-----C-----	4566
A1.UG.U455		-----A-----G-----C-----CAA-----T-----CAA-----T-----TAGA-----T-----G-----	4645
A2.CD.97CDKTB48		-----G-----C-----C-----C-----G-----AA-----A-----AG-----C-----G-----AGA-----G-----	4539
A2.CY.94CY017.41		-----G-----C-----C-----C-----A-----A-----AA-----T-----G-----C-----A-----AGAA-----	4556
CONSENSUS_B		-----a-----a-----a-----a-----a-----a-----a-----a-----a-----a-----	4743
B.AU.MBC200		-----A-----A-----A-----A-----A-----A-----GC-----A-----A-----A-----	5198
B.CN.RL42		-----G-----C-----C-----A-----A-----A-----A-----T-----A-----A-----A-----	4586
B.DE.D31		-----G-----G-----C-----A-----GA-----AA-----A-----A-----A-----A-----	4745
B.DE.HAN		-----A-----A-----TG-----C-----A-----AAC-----G-----A-----A-----A-----	4665
B.ES.89SP061		-----A-----A-----A-----A-----A-----A-----A-----TA-----A-----G-----	4757
B.GA.OYI		-----A-----A-----A-----A-----A-----A-----A-----A-----A-----G-----	4742
B.GB.MANC		-----A-----A-----A-----A-----A-----A-----C-----A-----A-----A-----	5145
B.NL.3202A21		-----G-----AAG-----AA-----GA-----A-----TA-----TA-----	5201
B.TM.TWCYS		-----G-----G-----A-----AA-----AA-----C-----A-----TA-----	5200
B.US.JREL		-----C-----AC-----A-----AA-----A-----A-----A-----A-----	4516
B.US.MNCG		-----A-----A-----AA-----AA-----C-----A-----A-----A-----	5214
B.US.RF		-----A-----A-----A-----AA-----AA-----A-----A-----A-----	4716
B.US.SF2		-----A-----A-----AA-----AA-----A-----TA-----A-----G-----	5206
B.US.YU2		-----G-----G-----A-----A-----A-----T-----	5197
CONSENSUS_C		-----g-----a-----a-----a-----g-----at-----c-----T-----AGA-----a-----g-----	4522
C.BR.92BR025		-----A-----A-----T-----T-----G-----C-----A-----G-----T-----T-----AGA-----A-----	4536
C.BR.98BR004		-----A-----T-----T-----A-----G-----C-----T-----A-----A-----T-----G-----AGA-----A-----	4524
C.BW.96BW0502		-----G-----G-----AA-----C-----A-----AAA-----G-----A-----C-----C-----AGA-----A-----G-----	4692
C.BW.96BW15C02		-----G-----T-----C-----G-----T-----A-----C-----AA-----G-----AT-----AGA-----A-----G-----	5181
C.ET.ETH2220		-----G-----T-----G-----T-----A-----A-----G-----A-----AT-----T-----T-----TAGA-----A-----G-----	4666
C.IL.98ILS002		-----A-----A-----C-----A-----A-----G-----AT-----G-----T-----C-----AGA-----AGG-----	4591
C.IN.93IN101		-----G-----T-----G-----G-----CA-----AA-----G-----A-----C-----T-----G-----AGA-----A-----G-----	4533
C.IN.94IN476		-----G-----A-----A-----G-----A-----G-----AA-----T-----AGA-----A-----G-----	5172
C.IN.98IN012		-----G-----T-----T-----A-----A-----G-----T-----C-----A-----T-----AGA-----A-----G-----	4519
C.TZ.98TZ013		-----G-----A-----A-----A-----G-----A-----G-----T-----C-----A-----T-----AGA-----A-----G-----	4524
C.TZ.98TZ017		-----G-----A-----A-----C-----T-----GT-----A-----A-----G-----T-----AGA-----A-----G-----	4542
C.ZA.97ZA012		-----G-----A-----C-----A-----A-----A-----A-----A-----C-----C-----G-----T-----AGA-----A-----G-----	4553
C.ZM.96ZM651		-----G-----T-----T-----A-----A-----GAT-----A-----GAT-----C-----G-----TAGG-----A-----G-----	4533
C.ZM.96ZM751		-----G-----G-----G-----A-----C-----AA-----AT-----T-----AGA-----GG-----	4528
CONSENSUS_D		-----A-----A-----AAA-----G-----A-----G-----TA-----C-----T-----AGA-----A-----	4538
D.CD.84ZR085		-----AC-----T-----C-----A-----AA-----aa-----aa-----c-----?-----c-----A-----	4735
D.CD.FLI		-----A-----A-----ACA-----A-----ACA-----C-----C-----C-----A-----	4721
D.CD.Z2Z6		-----CA-----CA-----CA-----CA-----CA-----C-----C-----C-----A-----	4745
D.KE.MB2059		-----CA-----G-----AA-----G-----CAA-----C-----G-----C-----C-----A-----	5202
D.UG.94UG1141		-----A-----A-----AA-----G-----AA-----AA-----AA-----CT-----AA-----A-----	4554
CONSENSUS_F1		-----T-----c-----c-----AA-----C-----AA-----?-----T-----?-----AGG-----a-----	4462
F1.BE.VI850		-----G-----G-----AC-----AC-----A-----A-----C-----AA-----C-----AGG-----A-----G-----	4539
F1.BR.93BR020.1		-----G-----G-----AC-----AC-----A-----A-----C-----AA-----C-----AGG-----A-----G-----	4527
F1.FI.FIN9363		-----G-----T-----AC-----AC-----A-----A-----C-----AAA-----T-----AGA-----A-----	4528
F1.FR.MP411		-----G-----G-----C-----C-----C-----C-----C-----AAAAT-----T-----T-----C-----A-----	4401
F2.CM.MP255		-----G-----G-----C-----C-----G-----A-----G-----AA-----AA-----AGG-----A-----G-----	4395
F2.CM.MP257		-----G-----G-----C-----C-----C-----A-----AA-----G-----GCA-----AGG-----G-----	4407
CONSENSUS_G		-----G-----c-----c-----c-----AA-----?-----A-----C-----?-----AGG-----?-----G-----	4553
G.BE.DRCBL		-----G-----T-----T-----A-----AA-----AA-----C-----TAGG-----C-----A-----G-----	5156
G.FI.HH8793-12.1		-----G-----T-----T-----C-----AA-----A-----C-----A-----C-----AGG-----A-----G-----	4599
G.NG.92NG083		-----G-----T-----T-----C-----AA-----AA-----C-----AGG-----A-----G-----	4568
G.SE.SE6165		-----G-----G-----C-----C-----G-----C-----AA-----A-----C-----C-----AGG-----G-----	4596
H.BE.VI991		-----A-----A-----C-----C-----C-----AA-----AAAA-----G-----TA-----A-----C-----AG-----C-----	4588
H.BE.VI997		-----G-----G-----AC-----A-----A-----AA-----ACA-----G-----T-----T-----C-----G-----	4523
H.CF.90CF056		-----G-----G-----AC-----C-----A-----A-----A-----T-----T-----A-----A-----G-----	4546
J.SE.SE7022		-----G-----AC-----AC-----A-----AA-----ACA-----C-----G-----T-----G-----A-----	4514
J.SE.SE7887		-----G-----AC-----AC-----A-----AA-----AAA-----C-----G-----G-----C-----AA-----A-----	4513
K.CD.EOTB11C		-----G-----AT-----T-----A-----AA-----G-----ACA-----CG-----A-----C-----AA-----A-----	4391
K.CM.MP535		-----G-----AT-----T-----C-----A-----AA-----ACC-----A-----AGG-----A-----	4395

HIV-1/SIVcpz complete genomes

B.FR.HXB2	GCACACAAGTAGACCTGAACACTAGCAGACCAACTAAATTCATCTGTATTACTTTCAGACTGCTTTTTCAGACTCTGCTATAAGAAAGCCCTATTATTAGGACACATAGTTAGCCCTAGGTTGTGAATATCAAGCAGG	5453
CONSENSUS_A-A1	-----a-t-t-c-g-g-----c-t-t-----c-g-a-a-a-----g	4776
A.BY.97BL006	-----A-----A-----C-----G-----A-----A-----G-----	4830
A.SE.SE7535	GA-----T-----C-----G-----C-----G-----A-----C-----AG-----	4670
A.SE.SE8538	-----A-----T-----C-----G-----A-----A-----AG-----	4652
A.SE.SE8891	-----A-----T-----C-----G-----A-----A-----AG-----	4642
A.SE.UGSE8131	-----A-----T-----C-----G-----A-----A-----AG-----	4824
A1.KE.Q23-17	-----A-----ACT-----C-----G-----A-----A-----G-----AG-----	4911
A1.SE.SE7253	-----A-----T-----C-----G-----A-----A-----A-----G-----AG-----	4657
A1.UG.92UG037	-----A-----T-----C-----G-----C-----G-----A-----A-----G-----AG-----	4820
A1.UG.U455	-----T-----T-----C-----C-----C-----T-----A-----G-----GA-----A-----A-----	4899
A2.CD.97CDKTB48	-----T-----A-----T-----G-----T-----C-----C-----G-----A-----A-----G-----A-----	4993
A2.CY.94CY017.41	AT-----A-----T-----T-----C-----C-----T-----C-----G-----A-----A-----G-----A-----	4780
CONSENSUS_B	-----c-----c-----t-----t-----a-----t-----a-----t-----t-----	4995
B.AU.MBC200	-----C-----C-----G-----A-----A-----T-----A-----T-----G-----	5452
B.CN.RL42	-----GC-----T-----T-----A-----T-----GT-----	4840
B.DE.D31	-----GC-----C-----T-----T-----A-----GT-----G-----	4999
B.DE.HAN	-----A-----C-----T-----T-----A-----GT-----	4919
B.ES.89SP061	-----T-----A-----C-----T-----A-----T-----	5011
B.GA.OYI.	-----GC-----ACA-----T-----A-----A-----C-----A-----T-----	4996
B.GB.MANC	-----T-----GC-----T-----T-----A-----T-----A-----TG-----	5399
B.NL.3202A21	-----T-----GC-----C-----A-----A-----A-----G-----TGT-----	5455
B.TW.TWCYS	-----A-----C-----A-----T-----A-----A-----G-----T-----	5454
B.US.JREL	-----C-----C-----C-----T-----T-----A-----T-----G-----T-----	4770
B.US.MNCG	-----C-----C-----C-----C-----C-----C-----A-----A-----T-----TT-----	5468
B.US.RF	-----GC-----C-----C-----T-----T-----A-----A-----C-----A-----C-----	4970
B.US.SF2	-----GC-----C-----C-----T-----T-----A-----A-----T-----T-----G-----	5460
B.US.YU2	-----C-----C-----T-----T-----A-----A-----T-----T-----G-----	5451
CONSENSUS_C	-----GC-----G-----d-----A-----C-----T-----T-----G-----C-----a-----a-----t-----	4776
C.BR.92BR025	-----GC-----G-----G-----A-----C-----T-----T-----G-----C-----G-----A-----A-----T-----	4790
C.BR.98BR004	-----GC-----G-----G-----A-----AC-----T-----T-----G-----C-----AG-----A-----A-----C-----	4778
C.BW.96BW0502	-----A-----GC-----G-----G-----C-----A-----C-----I-----T-----G-----C-----G-----A-----A-----C-----	4946
C.BW.96BW06	-----A-----GC-----G-----G-----C-----A-----A-----C-----I-----T-----G-----C-----G-----A-----A-----C-----	5435
C.BW.96BW15C02	-----GC-----G-----G-----C-----C-----A-----T-----T-----G-----A-----A-----A-----C-----	4920
C.ET.ETH2220	-----GC-----G-----G-----C-----C-----A-----C-----T-----T-----G-----A-----A-----A-----C-----	4845
C.IL.98ILS002	-----A-----AGC-----G-----A-----C-----T-----T-----G-----G-----A-----A-----T-----	4787
C.IN.93IN101	-----A-----A-----GC-----G-----C-----A-----C-----T-----T-----G-----G-----A-----A-----C-----	5426
C.IN.94IN476	-----A-----A-----GC-----G-----C-----A-----C-----T-----T-----G-----G-----A-----A-----C-----	4773
C.IN.98IN012	-----A-----A-----GC-----G-----C-----A-----C-----T-----T-----G-----G-----A-----A-----C-----	4778
C.TZ.98TZ013	-----A-----AGC-----G-----C-----G-----A-----C-----T-----T-----G-----G-----A-----A-----C-----	4796
C.TZ.98TZ017	-----GC-----G-----G-----T-----A-----C-----T-----T-----G-----G-----A-----A-----C-----	4807
C.ZA.97ZA012	-----GC-----G-----G-----T-----A-----C-----T-----T-----G-----G-----A-----A-----C-----	4877
C.ZM.96ZM651	-----GC-----G-----G-----T-----A-----C-----T-----T-----G-----G-----A-----A-----C-----	4782
C.ZM.96ZM751	-----GC-----G-----G-----T-----A-----C-----T-----T-----G-----G-----A-----A-----C-----	4792
CONSENSUS_D	-----GC-----G-----G-----T-----T-----A-----T-----T-----G-----A-----A-----C-----C-----	4988
D.CD.84ZR085	-----GC-----G-----G-----T-----T-----A-----T-----T-----G-----A-----A-----C-----C-----	4975
D.CD.FLI	-----GC-----G-----G-----T-----T-----A-----T-----T-----G-----A-----A-----C-----C-----	4975
D.CD.ZZ6	-----GC-----G-----G-----T-----T-----A-----T-----T-----G-----A-----A-----C-----C-----	4999
D.KE.MB2059	-----GC-----G-----G-----T-----T-----A-----T-----T-----G-----A-----A-----C-----C-----	5456
D.UG.94UG1141	-----GC-----G-----G-----T-----T-----A-----T-----T-----G-----A-----A-----C-----C-----	5457
CONSENSUS_F1	A-----A-----GC-----G-----G-----T-----T-----A-----C-----I-----T-----T-----G-----A-----A-----AG-----TAT-----	4808
F1.BE.VI850	-----G-----A-----GC-----G-----G-----T-----T-----A-----a-----C-----G-----A-----A-----A-----C-----	4713
F1.BR.93BR020.1	-----G-----A-----GC-----G-----G-----T-----T-----A-----a-----C-----G-----A-----A-----A-----C-----	4793
F1.FI.FIN9363	-----G-----A-----GC-----G-----G-----T-----T-----A-----a-----C-----G-----A-----A-----A-----C-----	4781
F1.FR.MF411	-----G-----A-----GC-----G-----G-----T-----T-----A-----a-----C-----G-----A-----A-----A-----C-----	4782
F2.CM.MP255	-----G-----A-----GC-----G-----G-----T-----T-----A-----a-----C-----G-----A-----A-----A-----C-----	4655
F2.CM.MP257	-----G-----A-----GC-----G-----G-----T-----T-----A-----a-----C-----G-----A-----A-----A-----C-----	4652
F2.CM.MP257	-----G-----A-----GC-----G-----G-----T-----T-----A-----a-----C-----G-----A-----A-----A-----C-----	4664
CONSENSUS_G	t-----?-----t-----c-----t-----g-----c-----t-----t-----?-----	4803
G.BE.DRCBL	-----T-----T-----C-----C-----C-----T-----A-----T-----A-----T-----A-----C-----G-----A-----A-----C-----	5410
G.FI.HH8793-12.1	-----A-----T-----C-----C-----C-----T-----A-----T-----A-----T-----A-----C-----G-----A-----A-----C-----	4853
G.ING.92NG083	-----T-----A-----A-----CAC-----T-----G-----T-----G-----A-----A-----AGA-----	4822
G.SE.SE6165	-----A-----G-----C-----C-----T-----G-----T-----G-----A-----A-----G-----	4850
H.BE.VI991	-----A-----A-----C-----G-----G-----T-----T-----C-----T-----G-----A-----A-----TA-----	4842
H.BE.VI997	-----G-----GCT-----G-----ACCC-----T-----T-----A-----C-----GGGA-----A-----GAG-----	4777
H.CF.90CF056	-----G-----G-----GC-----G-----G-----T-----T-----A-----C-----G-----A-----A-----G-----	4800
J.SE.SE7022	-----A-----A-----GC-----G-----G-----T-----T-----A-----C-----G-----A-----A-----A-----C-----	4768
J.SE.SE7887	-----A-----A-----GC-----G-----G-----T-----T-----A-----C-----G-----A-----A-----A-----C-----	4767
K.CD.EQTB11C	-----A-----T-----GC-----G-----G-----T-----T-----A-----A-----T-----A-----A-----C-----C-----	4645
K.CM.MP535	-----A-----A-----C-----C-----C-----T-----T-----A-----A-----T-----T-----G-----T-----G-----	4649

B_FR_HXB2	GCACACAAGTAGCAGCACTAAGTAACTAATTCATCTGTATTAATTTTCAGACTCTGCTATAAGAAAGCCCTTATTAGGACACATAGTTAGCCCTAGTGTGTAATATCAAGCAGG	5453
CONSENSUS_01_AE	-----a-----T-----c-----G-----AG-----C-----G-----A-----a-----Ag-----A-----g-----C-----t-----	5404
01_AE_CF_90CF11697	-----T-----T-----C-----G-----AG-----C-----A-----T-----A-----A-----G-----AG-----C-----A-----	5402
01_AE_CF_90CF402	-----A-----T-----C-----G-----G-----C-----A-----T-----A-----A-----A-----A-----A-----A-----C-----A-----	5489
01_AE_CF_90CF4071	-----A-----T-----T-----G-----G-----G-----G-----A-----A-----A-----A-----A-----A-----A-----A-----	5389
01_AE_TH_93TH057	-----A-----T-----T-----G-----G-----G-----G-----A-----A-----A-----A-----A-----A-----A-----A-----	5406
01_AE_TH_93TH065	-----A-----T-----T-----G-----G-----G-----G-----A-----A-----A-----A-----A-----A-----A-----A-----	5396
01_AE_TH_93TH253	-----A-----T-----T-----G-----G-----G-----G-----A-----A-----A-----A-----A-----A-----A-----A-----	5440
01_AE_TH_95TH022	-----A-----T-----T-----G-----G-----G-----G-----A-----A-----A-----A-----A-----A-----A-----A-----	5416
01_AE_TH_95TH047	-----A-----T-----T-----G-----G-----G-----G-----A-----A-----A-----A-----A-----A-----A-----A-----	5416
01_AE_TH_CM240	-----A-----T-----T-----G-----G-----G-----G-----A-----A-----A-----A-----A-----A-----A-----A-----	5027
CONSENSUS_02_AG	-----A-----t-----C-----C-----C-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----	4658
02_AG_CM_97CM-MP807	-----A-----T-----C-----C-----C-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----	4609
02_AG_FR_DJ263	-----A-----T-----C-----C-----C-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----	4801
02_AG_FR_DJ264	-----A-----T-----C-----C-----C-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----	4802
02_AG_GH_G829	-----A-----T-----C-----C-----C-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----	4669
02_AG_NG_IBNG	-----A-----T-----C-----C-----C-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----	4978
02_AG_SE_SE7812	-----GA-----T-----C-----C-----C-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----	4825
02_AG_SN_98SE-MP1211	-----A-----A-----T-----C-----C-----C-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----	4650
02_AG_SN_MP1213	-----A-----A-----T-----C-----C-----C-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----	4644
03_AB_RU_KAL153-2	-----T-----T-----A-----C-----C-----T-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----	5222
03_AB_RU_RU98001	-----T-----T-----A-----C-----C-----T-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----	4823
04_CPX_CX_94CY032-3	-----G-----T-----C-----G-----A-----C-----T-----T-----A-----A-----C-----G-----A-----A-----A-----A-----	4819
04_CPX_GR_97PVGH	-----T-----T-----C-----G-----A-----C-----T-----T-----A-----A-----C-----G-----A-----A-----A-----A-----	5470
04_CPX_GR_97PVNY	-----T-----T-----A-----C-----G-----C-----T-----T-----A-----A-----C-----G-----A-----A-----A-----A-----	5467
05_DF_BE_VI1310	-----A-----A-----AGC-----G-----T-----T-----T-----A-----A-----C-----T-----A-----A-----A-----A-----	4835
05_DF_BE_VI1961	-----A-----A-----AGC-----G-----T-----T-----T-----A-----A-----C-----T-----A-----A-----A-----A-----	4810
CONSENSUS_06_CPX	-----a-----aGT-----G-----A-----C-----T-----T-----a-----a-----C-----G-----A-----a-----A-----A-----?	5407
06_CPX_AU_BFP90	-----A-----AGTA-----G-----A-----A-----C-----T-----T-----A-----A-----C-----G-----A-----A-----A-----A-----	5481
06_CPX_ML_95ML127	-----A-----GT-----G-----A-----C-----T-----T-----A-----A-----C-----G-----A-----A-----A-----A-----	5470
06_CPX_ML_95ML84	-----G-----AGT-----G-----A-----C-----T-----T-----A-----A-----C-----G-----A-----A-----A-----A-----	4696
06_CPX_SN_97SE1078	-----A-----AGT-----G-----A-----C-----T-----T-----A-----A-----C-----G-----A-----A-----A-----A-----	5512
07_BC_CN_97GN001	-----C-----A-----A-----GC-----G-----T-----T-----T-----G-----G-----A-----A-----A-----A-----	4785
07_BC_CN_98CN009	-----C-----A-----A-----GC-----G-----T-----T-----T-----G-----G-----A-----A-----A-----A-----	4785
07_BC_CN_C54A	-----C-----A-----A-----GC-----G-----T-----T-----T-----G-----G-----A-----A-----A-----A-----	4643
CONSENSUS_08_BC	-----A-----A-----A-----GC-----G-----T-----T-----T-----G-----G-----A-----A-----A-----A-----	4632
08_BC_CN_97NGX6F	-----A-----A-----A-----GC-----G-----T-----T-----T-----G-----G-----A-----A-----A-----A-----	4633
08_BC_CN_97NGX7F	-----A-----A-----A-----GC-----G-----T-----T-----T-----G-----G-----A-----A-----A-----A-----	4639
08_BC_CN_97NGX9F	-----A-----A-----A-----GC-----G-----T-----T-----T-----G-----G-----A-----A-----A-----A-----	4633
08_BC_CN_98CN006	-----AT-----A-----A-----GC-----G-----T-----T-----T-----G-----G-----A-----A-----A-----A-----	4785
10_CD_TZ_96TZBF061	-----T-----T-----A-----A-----GC-----G-----T-----T-----T-----G-----G-----A-----A-----A-----A-----	4832
10_CD_TZ_96TZBF071	-----T-----T-----A-----A-----GC-----G-----T-----T-----T-----G-----G-----A-----A-----A-----A-----	4830
10_CD_TZ_96TZBF110	-----T-----T-----A-----A-----GC-----G-----T-----T-----T-----G-----G-----A-----A-----A-----A-----	4821
CONSENSUS_11_CPX	-----A-----A-----GC-----g-----A-----C-----?-----T-----T-----g-----a-----C-----G-----A-----A-----A-----?	5352
11_CPX_CM_MP818	-----A-----A-----GC-----G-----A-----AC-----T-----T-----A-----A-----C-----GG-----A-----A-----A-----A-----	5456
11_CPX_FR_MP1298	-----A-----A-----GC-----G-----A-----C-----T-----T-----A-----A-----C-----G-----CA-----A-----A-----A-----	5447
11_CPX_FR_MP1307	-----A-----A-----GC-----G-----A-----C-----T-----T-----A-----A-----C-----G-----CA-----A-----A-----A-----	5504
11_CPX_GR_GR17	-----A-----A-----GC-----G-----A-----C-----T-----T-----A-----A-----C-----G-----A-----A-----A-----A-----	4755
12_BF_AR_ARM159	-----G-----A-----GC-----G-----G-----C-----A-----A-----T-----T-----A-----A-----C-----G-----A-----A-----	5458
12_BF_UY_URTR23	-----G-----A-----GC-----G-----G-----C-----A-----A-----T-----T-----A-----A-----C-----G-----A-----A-----	5479
12_BF_UY_URTR35	-----G-----A-----GC-----G-----G-----C-----A-----A-----T-----T-----A-----A-----C-----G-----A-----A-----	5476
N_CM_YBF106	AG-----T-----T-----A-----G-----A-----A-----TGT-----T-----A-----T-----A-----CT-----C-----GC-----A-----GG-----C-----	5045
N_CM_YBF30	AG-----T-----T-----A-----G-----A-----A-----TGT-----T-----A-----T-----A-----CT-----C-----GC-----A-----GG-----C-----	5047
CONSENSUS_0	aa-----GA-----T-----AC-----AG7A-----G-----A-----?-----T-----T-----?-----C-----A-----C-----g-----g-----a-----	5295
O_CM_ANT70	AA-----GA-----T-----AC-----AGGA-----G-----A-----C-----T-----T-----CACA-----A-----CA-----A-----AG-----C-----A-----	5083
O_CM_MVP5180	AA-----GA-----T-----AC-----AGGA-----G-----A-----C-----T-----T-----CACA-----A-----CA-----A-----AG-----C-----A-----	5483
O_SN_SEMP1299	AA-----GA-----T-----AC-----AG-----A-----G-----A-----A-----A-----C-----T-----T-----C-----A-----C-----G-----G-----	5507
O_SN_SEMP1300	CA-----GA-----T-----AC-----AG-----A-----G-----A-----A-----ACA-----T-----T-----C-----A-----C-----G-----G-----	5507
CONSENSUS_CPZ	?-----t-----t-----?-----Acg-----ta-----c-----tc-----T-----t-----g-----ct-----c-----c-----A-----tg-----t-----a-----	4892
CPZ_CD_CPZANT	T-----t-----A-----AC-----T-----T-----A-----G-----A-----C-----AAA-----T-----T-----CCC-----T-----T-----C-----A-----	4898
CPZ_CM_CAM3	T-----t-----GA-----G-----TA-----A-----C-----TC-----A-----T-----T-----G-----CT-----C-----C-----G-----A-----	4867
CPZ_CM_CAM5	AA-----G-----T-----AGC-----TAG-----T-----G-----C-----CTCTC-----A-----T-----T-----C-----G-----AA-----	5155
CPZ_GA_CPZGAB	T-----G-----G-----T-----CTTAC-----TAG-----T-----G-----C-----CTCTC-----A-----T-----T-----C-----G-----AA-----	5517
CPZ_US_CPZUS	AG-----A-----T-----GA-----G-----A-----G-----A-----A-----C-----C-----C-----C-----A-----TG-----T-----	5513
Vif	S_T_Q_V_D_P_E_L_A_D_Q_L_I_H_L_I_Y_Y_F_D_C_F_S_D_S_A_I_R_K_A_L_L_G_H_I_V_S_P_R_C_E_Y_Q_A_G	

HIV-1/SIVcpz complete genomes

```

B.FR.HXB2          ACATAAAC...AAGGTAGGATCTCAATAACTTTGGCACTAGCAGCAITTAATA.....ACACCAAAAAAGATA...AAGCCACCTTTGGCTAGTGTTCAGAAACTCAGACAGGATAGATGGAAACAAG / Vpr CDS start 5568
CONSENSUS_A-A1    -----t-a-----gAA-----g-c-----g-----t-a-----
A.BY.97BL006      -----G-----G-----C-G-GA-----G-----G-----N-A-----R-----
A.SE.SE7535       -----T-A-----GAA-----G-----G-----T-T-AG-----R-----
A.SE.SE8538       -----C-----GAA-----C-----G-----T-A-----C-----
A.SE.SE8891       -----T-A-----GAA-----G-----G-----GT-AG-----C-----
A.SE.UGSE8131     -----T-A-----GAA-----G-----G-----GT-AG-----C-----
A1.KE.O23-17      -----T-A-----GAA-----G-----C-G-C-----T-T-A-----A-----
A1.SE.SE7253      -----T-A-----GAA-----G-----C-G-C-----T-T-AG-----G-----
A1.UG.92UG037     -----T-A-----AA-----TC-G-G-----T-AG-----G-----
A1.UG.U455        -----T-----T-GAA-----C-----C-G-GC-----A-----T-AG-----
A2.CD.97CDKTB48  -----G-----T-----T-GAG-----G-----G-----G-----T-AGT-----
A2.CY.94CY017.41 -----G-----T-----T-GAA-----G-----G-----G-----T-AGT-----
CONSENSUS_B      -----g-----g-----g-----g-----g-----g-----g-----g-----
B.AU.MBC200       -----G-----A-----C-----C-G-----C-----G-----C-----G-----
B.CN.RL42         -----G-----A-----C-----C-G-----C-----G-----C-----G-----
B.DE.D31          -----G-----G-----C-----C-G-----C-----G-----C-----G-----
B.DE.HAN          -----G-----A-----C-----G-----G-----A-----A-----C-----
B.ES.89SP061     -----G-----A-----G-----G-----G-----A-AGATA-----A-----
B.GA.OYI          -----G-----G-----G-----C-----C-----A-----A-----
B.GB.MANC        -----G-----A-----A-----CA-----GT-----A-----
B.NL.3202A21     -----G-----A-----A-----A-----T-----A-----
B.TM.TWCYS       -----G-----C-----A-----A-----T-----A-----
B.US.JREL        -----G-----C-----A-----C-----C-----A-----
B.US.MNCG        -----G-----C-----C-----C-----C-----A-----
B.US.RF          -----G-----C-----C-----C-----C-----A-----
B.US.SF2         -----G-----A-----A-----C-----C-----AA-----
CONSENSUS_C      -----T-----C-----GA-----g-----A-----g-----t-AGT-----
C.BR.92BR025     -----TC-----GA-----G-----G-----C-----A-----T-AGT-----
C.BR.98BR004     -----TC-----GA-----G-----G-----C-----G-----T-AGT-----
C.BW.96BW0502    -----G-----G-----GA-----G-----C-----G-----T-AGT-----
C.BW.96BW15C02   -----G-----G-----GA-----G-----C-----G-----T-AGT-----
C.ET.ETH2220     -----T-----G-----GA-----G-----C-----G-----T-AGT-----
C.IL.98IS002     -----T-----G-----GA-----G-----C-----G-----T-AGT-----
C.IN.93IN101     -----T-----C-----GA-----C-----G-----A-----T-AGT-----
C.IN.94IN476     -----T-----G-----GA-----C-----A-----G-----T-AGT-----
C.IN.98IN012     -----T-----G-----GA-----C-----C-----A-----T-AGT-----
C.TZ.98TZ013     -----G-----G-----GA-----C-----G-----G-----T-AGT-----
C.TZ.98TZ017     -----T-----G-----GA-----A-----A-----G-----G-----A-----
C.ZA.97ZA012     -----T-----G-----GA-----G-----AG-----GC-----A-----
C.ZM.96ZM651     -----T-----G-----GA-----G-----G-----G-----GA-----T-AGT-----
C.ZM.96ZM751     -----T-----C-----GA-----G-----G-----G-----G-----T-AGT-----
CONSENSUS_D      -----CT-----g-L-A-----?-----g-----g-----g-----a-----A-----
D.CD.84ZR085     -----CT-----G-T-A-----A-----G-----G-----C-----A-----G-----A-----
D.CD.FLI         -----C-----C-----G-----T-----G-----C-----C-----G-----G-----A-----
D.CD.Z2Z6        -----CT-----G-T-----A-----G-----G-----G-----GT-A-----A-----
D.KE.MB2059      -----T-----T-----A-----GT-----A-----A-----C-----GG-----G-----A-----
D.UG.94UG1141    -----T-----T-----A-----G-----G-----GG-----A-----A-----C-----GG-----G-----A-----
CONSENSUS_E      -----C-----L-----A-----A-----g-T-----B-----C-----C-----C-----C-----
E1.BE.VI850      -----C-----T-----A-----A-----C-----C-----C-----C-----C-----C-----
E1.BR.93BR020.1  -----T-----T-----A-----A-----G-----T-----G-----G-----C-----G-----C-----
E1.FI.FIN9363    -----C-----T-----A-----A-----G-----T-----G-----G-----C-----G-----C-----
E1.FR.MP411      -----C-----T-----A-----A-----G-----T-----G-----G-----C-----G-----C-----
F2.CM.MP255      -----C-----C-----T-----GA-----C-----C-----C-----G-----C-----G-----
F2.CM.MP257      -----C-----C-----T-----A-----A-----T-----T-----G-----A-----C-----G-----
CONSENSUS_G      -----T-----T-----T-----t-gAA-----?-----G-----C-----g-----g-----T-----A-----a-----
G.BE.DRCBL       -----T-----T-----T-----AA-----T-----G-----G-----C-----G-----G-----G-----T-----A-----G-----
G.FI.HH8793-12.1 -----T-----T-----G-----GAA-----G-----G-----C-----C-----C-----G-----T-----A-----A-----
G.NG.92NG083     -----T-----TC-----TCGAA-----G-----G-----C-----G-----A-----G-----G-----T-----AG-----A-----
G.SE.SE6165      -----T-----C-----T-----T-----GAA-----T-----G-----T-----G-----GT-----G-----G-----T-----AG-----A-----
H.BE.VI991       -----GA-----T-----T-----GA-----G-----G-----G-----C-----G-----GA-----AGT-----
H.BE.VI997       -----C-----C-----G-----T-----GA-----G-----T-----G-----C-----AA-----G-----AGT-----
H.CF.90CF056     -----A-----C-----C-----T-----A-----G-----G-----G-----G-----GA-----G-----AGT-----
J.SE.SE7022      -----T-----TC-----T-----A-----A-----G-----G-----G-----G-----CA-----AGT-----
J.SE.SE7887      -----T-----TC-----T-----A-----A-----G-----G-----G-----G-----CA-----AGT-----
K.CD.E0TB11C    -----C-----C-----CT-----G-----T-----T-----A-----G-----T-----G-----CC-----AGT-----
K.CM.MP535       -----C-----C-----CT-----G-----T-----T-----CA-----G-----G-----G-----G-----CC-----AGT-----

```


HIV-1/SIVcpz complete genomes

Table with 2 columns: Accession ID (e.g., B.FR.HXB2, CONSENSUS_A-A1, A.BY.97BL006) and Nucleotide sequence (e.g., TGAACAATTATGGGGATACTTTGGGCAGGAGTGGGAAGCCATAAATAAGAAATTTTCGAAACAACCTGCTGTTATCCATTTTTCAGAAATTTGGGTGCACATAGCAGAGA...). The table lists various HIV-1 and SIVcpz strains and their corresponding genomic sequences.

B_FR_HXB2	TGAAACTTATGGGGATACTTGGCGAGGAGTGGGAGGAGCCATAAAGAAATTTTCAGAAATTTGGGTGCGACATAGCAGA...ATAGGCGTT...ACTCGA.....	5814
CONSENSUS_01_AE	A-C-a	5764
01_AE_CF_90CF11697	A-G-T	5762
01_AE_CF_90CF402	A-G-T	5849
01_AE_CF_90CF4071	A-G-T	5747
01_AE_TH_93TH057	A-G-T	5766
01_AE_TH_93TH065	A-G-T	5756
01_AE_TH_93TH253	A-G-T	5801
01_AE_TH_95TH022	A-G-T	5776
01_AE_TH_95TH047	A-G-T	5773
01_AE_TH_CM240	A-G-T	5381
CONSENSUS_02_AG	A-G-T	5018
02_AG_CM_97CM-MP807	A-G-T	4969
02_AG_FR_DJ263	A-G-T	5161
02_AG_FR_DJ264	A-G-T	5162
02_AG_GH_G829	A-G-T	5029
02_AG_NG_IBNG	A-G-T	5338
02_AG_SE_SE7812	A-G-T	5185
02_AG_SN_98SE-MP1211	A-G-T	5010
02_AG_SN_MF1213	A-G-T	9004
03_AB_RU_KAL153-2	A-G-T	5582
03_AB_RU_RU98001	A-G-T	5183
04_cpx_CX_94CY032-3	A-G-T	5189
04_cpx_GR_97PVGH	A-G-T	5840
04_cpx_GR_97PVNY	A-G-T	5837
05_DF_BE_VI1310	A-G-T	5195
05_DF_BE_VI1961	A-G-T	5170
CONSENSUS_06_cpx	A-G-T	5761
06_cpx_AU_BFP90	A-G-T	5841
06_cpx_ML_95ML127	A-G-T	5830
06_cpx_ML_95ML84	A-G-T	5056
06_cpx_SN_97SE1078	A-G-T	5872
07_BC_CN_97CN001	A-G-T	5148
07_BC_CN_98CN009	A-G-T	5148
07_BC_CN_C54A	A-G-T	5006
CONSENSUS_08_BC	A-G-T	4995
08_BC_CN_97CNGX6F	A-G-T	5006
08_BC_CN_97CNGX7F	A-G-T	4992
08_BC_CN_97CNGX9F	A-G-T	4996
08_BC_CN_98CN006	A-G-T	5145
10_CD_TZ_96TZBF061	A-G-T	5192
10_CD_TZ_96TZBF071	A-G-T	5190
10_CD_TZ_96TZBF110	A-G-T	5181
CONSENSUS_11_cpx	A-G-T	5707
11_cpx_CM_MP818	A-G-T	5816
11_cpx_FR_MP1298	A-G-T	5807
11_cpx_FR_MP1307	A-G-T	5858
11_cpx_GR_GR17	A-G-T	5115
12_BF_AR_ARM159	A-G-T	5818
12_BF_UY_URTR23	A-G-T	5839
12_BF_UY_URTR35	A-G-T	5836
N_CM_YBF106	A-G-T	5405
N_CM_YBF30	A-G-T	5407
CONSENSUS_0	A-G-T	5649
O_CM_ANT70	A-G-T	5871
O_CM_MVP5180	A-G-T	5853
O_SN_SEMP1299	A-G-T	5877
O_SN_SEMP1300	A-G-T	5877
CONSENSUS_CPZ	A-G-T	5233
CPZ_CD_CPZANT	A-G-T	5261
CPZ_CM_CAM3	A-G-T	5227
CPZ_CM_CAM5	A-G-T	5515
CPZ_GA_CPZGAB	A-G-T	5877
CPZ_US_CPZUS	A-G-T	5886
B_FR_HXB2	TGAAACTTATGGGGATACTTGGCGAGGAGTGGGAGGAGCCATAAAGAAATTTTCAGAAATTTGGGTGCGACATAGCAGA...ATAGGCGTT...ACTCGA.....	5814
E_T_Y_G_D_T_W_A_G_V_E_A_I_I_R_I_L_Q_Q_L_L_F_I_H_F_R_I_G_C_A_T_H_S_R_I_G_V_T_R	frameshift in HXB2	Vpr

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz complete genomes

Table with columns for sample ID, sequence alignment, and position. Includes sample names like B.FR.HXB2, A.BY.97BL006, and various consensus sequences. The alignment shows nucleotide differences between samples and the reference sequence. Position numbers are listed on the right side of the alignment.

Table with columns for accession numbers (e.g., B.FR.HXB2, CONSENSUS_01_AE), sequence alignment (Tat, Rev exon 1 start, intron start), and reverse sequence (Rev, Tat). The table lists 100+ different HIV-1/SIVcpz strains and their corresponding genomic sequences.

HIV-1/SIVcpz complete genomes

B.FR.HXB2	...GTAGTACATGTAACGCA...	...CTTATCAATAATAGCAATAAATAGCAATAGTTGTGTGTCATAGTAATCATAGAA...	6148
CONSENSUS_A-A1	TactA-A-A	T-act	6148
A.BY.97BL006	TACTA-A-A	T-AC	5478
A.SE.SE7535	TAAT	T-ATT	5374
A.SE.SE8538	TAGTG	T-TCT	5371
A.SE.SE8891	TAAT	T-ACT	5502
A.SE.UGSE8131	TAAT	T-AGT	5615
A1.KE.O23-17	TAATA	T-TCT	5527
A1.SE.SE7253	TAATA	T-TAG	5598
A1.UG.92UG037	TAAT	T-TG	5497
A1.UG.U455	TAAT	T-TG	5514
A2.CD.97CDKTB48	TAATC-A	T-TG	5687
A2.CY.94CY017.41	TAAT	T-TT	6146
CONSENSUS_B	T-TT	T-TT	5534
B.AU.MBC200	T-TT	T-TT	5693
B.CN.RL42	T-TT	T-TT	5613
B.DE.D31	T-TT	T-TT	5708
B.DE.HAN	T-TT	T-TT	5690
B.ES.89SP061	T-TT	T-TT	6152
B.GA.OYI	T-TT	T-TT	5464
B.GB.MANC	T-TT	T-TT	5664
B.NL.3202A21	T-TT	T-TT	6157
B.TM.TWCYS	T-TT	T-TT	6148
B.US.JREL	T-TT	T-TT	5464
B.US.MNCG	T-TT	T-TT	5664
B.US.RF	T-TT	T-TT	6157
B.US.SF2	T-TT	T-TT	6148
B.US.YU2	T-TT	T-TT	5483
CONSENSUS_C	T-TT	T-TT	5506
C.BR.92BR025	T-TT	T-TT	5485
C.BR.98BR004	T-TT	T-TT	5652
C.BW.96BW0502	T-TT	T-TT	6151
C.BW.96BW06	T-TT	T-TT	5629
C.BW.ETHW15C02	T-TT	T-TT	5560
C.ET.97ET2220	T-TT	T-TT	5512
C.IL.98IL002	T-TT	T-TT	6120
C.IN.93IN101	T-TT	T-TT	5479
C.IN.94IN476	T-TT	T-TT	5472
C.IN.98IN012	T-TT	T-TT	5502
C.TZ.98TZ013	T-TT	T-TT	5523
C.TZ.98TZ017	T-TT	T-TT	5483
C.ZA.97ZA012	T-TT	T-TT	5488
C.ZM.96ZM651	T-TT	T-TT	5508
C.ZM.96ZM751	T-TT	T-TT	5679
CONSENSUS_D	T-TT	T-TT	5669
D.CD.84ZR085	T-TT	T-TT	5693
D.CD.FLI	T-TT	T-TT	6150
D.CD.Z2Z6	T-TT	T-TT	6150
D.KE.MB2059	T-TT	T-TT	5502
D.UG.94UG1141	T-TT	T-TT	5396
CONSENSUS_F1	T-TT	T-TT	5497
F1.BE.VI850	T-TT	T-TT	5482
F1.BR.93BR020.1	T-TT	T-TT	5486
F1.FI.FIN9363	T-TT	T-TT	5354
F1.FR.MP411	T-TT	T-TT	5346
F2.CM.MP255	T-TT	T-TT	5358
F2.CM.MP257	T-TT	T-TT	5495
CONSENSUS_G	T-TT	T-TT	6114
G.BE.DRCBL	T-TT	T-TT	5557
G.FI.HH8793-12.1	T-TT	T-TT	5516
G.NG.92NG083	T-TT	T-TT	5554
G.SE.SE6165	T-TT	T-TT	5530
H.BE.VI991	T-TT	T-TT	5465
H.BE.VI997	T-TT	T-TT	5491
H.CF.90CF056	T-TT	T-TT	5471
J.SE.SE7022	T-TT	T-TT	5470
J.SE.SE7887	T-TT	T-TT	5346
K.CD.EOTB11C	T-TT	T-TT	5350
K.CM.MP535	T-TT	T-TT	

B.FR.HXB2 / Vpu CDS start (ACG start in HXB2) ...GATAGTACATGTAACGCAA... CTTATACCAATAGTAGCAATAGTACGCAATTAATAAATAGCAATAGTTGCTGCTCCATAGTAAATCATAGAA... 6148



HIV-1/SIVcpz complete genomes

B.FR.HXB2 / Env start
CONSENSUS_01_AE
01_AE.CF.90CF11697
01_AE.CF.90CF402
01_AE.CF.90CF4071
01_AE.TH.93TH057
01_AE.TH.93TH065
01_AE.TH.93TH253
01_AE.TH.95TH022
01_AE.TH.95TH047
01_AE.TH.CM240
01_AE.TH.CM242
02_AG.CM.97CM-MP807
02_AG.FR.DJ263
02_AG.FR.DJ264
02_AG.GH.G829
02_AG.GH.IBNG
02_AG.SE.SE7812
02_AG.SN.98SE-MP1211
02_AG.SN.MP1213
03_AB.RU.KAL153-2
03_AB.RU.RU98001
04_CFX.CV.94CV032-3
04_CFX.GR.97PVGH
04_CFX.GR.97PVNY
05_DF.BE.VI1310
05_DF.BE.VI1961
CONSENSUS_06_CPX
06_CFX.AU.BFP90
06_CFX.ML.95ML127
06_CFX.ML.95ML84
06_CFX.SN.97SE1078
07_BC.CN.97CN001
07_BC.CN.98CN009
07_BC.CN.C54A
CONSENSUS_08_BC
08_BC.CN.97CNGX6F
08_BC.CN.97CNGX7F
08_BC.CN.97CNGX9F
08_BC.CN.98CN006
10_CD.TZ.96TZBF061
10_CD.TZ.96TZBF071
10_CD.TZ.96TZBF110
CONSENSUS_11_CPX
11_CFX.CM.MP818
11_CFX.FR.MP1298
11_CFX.FR.MP1307
11_CFX.GR.GR17
12_BF.AR.ARMA159
12_BF.UY.URTR23
12_BF.UY.URTR35
N.CM.YBF106
N.CM.YBF30
CONSENSUS_0
O.CM.AN170
O.CM.AN170
O.SN.SEMP1289
O.SN.SEMP1290
O.SN.SEMP1300
CONSENSUS_CPZ
CPZ.CD.CPZANT
CPZ.CM.CAM3
CPZ.CM.CAM5
CPZ.GA.CPZGAB
CPZ.US.CPZUS
Env signal
Env
Vpu

HIV-1/SIVcpz complete genomes

Env signal peptide end \ / Env gp120 start
Vpu end \

Table with columns for accession number (e.g., B.FR.HXB2), strain name (e.g., CONSENSUS_A-A1), and sequence alignment. The alignment shows nucleotide sequences for various HIV-1/SIVcpz strains, with positions 6376 to 5575 marked at the top. Consensus sequences are indicated by dashes.

B_FR_HXB2 Env signal peptide end // Env gp120 start Vpu end
 ACTTTGGAGATGGGGTGGAGATGGGCACCATGCTCTTTGGGATGTTGATGCTGTAGTGTGAC...GAAATAATTTGGGTCACAGTCTATTATGGGTACCTGTTGGAAGGAAAGCAACCCACC
 6376
 6377
 6378
 6379
 6380
 6381
 6382
 6383
 6384
 6385
 6386
 6387
 6388
 6389
 6390
 6391
 6392
 6393
 6394
 6395
 6396
 6397
 6398
 6399
 6400
 6401
 6402
 6403
 6404
 6405
 6406
 6407
 6408
 6409
 6410
 6411
 6412
 6413
 6414
 6415
 6416
 6417
 6418
 6419
 6420
 6421
 6422
 6423
 6424
 6425
 6426
 6427
 6428
 6429
 6430
 6431
 6432
 6433
 6434
 6435
 6436
 6437
 6438
 6439
 6440
 6441
 6442
 6443
 6444
 6445
 6446
 6447
 6448
 6449
 6450
 6451
 6452
 6453
 6454
 6455
 6456
 6457
 6458
 6459
 6460
 6461
 6462
 6463
 6464
 6465
 6466
 6467
 6468
 6469
 6470
 6471
 6472
 6473
 6474
 6475
 6476
 6477
 6478
 6479
 6480
 6481
 6482
 6483
 6484
 6485
 6486
 6487
 6488
 6489
 6490
 6491
 6492
 6493
 6494
 6495
 6496
 6497
 6498
 6499
 6500
 6501
 6502
 6503
 6504
 6505
 6506
 6507
 6508
 6509
 6510
 6511
 6512
 6513
 6514
 6515
 6516
 6517
 6518
 6519
 6520
 6521
 6522
 6523
 6524
 6525
 6526
 6527
 6528
 6529
 6530
 6531
 6532
 6533
 6534
 6535
 6536
 6537
 6538
 6539
 6540
 6541
 6542
 6543
 6544
 6545
 6546
 6547
 6548
 6549
 6550
 6551
 6552
 6553
 6554
 6555
 6556
 6557
 6558
 6559
 6560
 6561
 6562
 6563
 6564
 6565
 6566
 6567
 6568
 6569
 6570
 6571
 6572
 6573
 6574
 6575
 6576
 6577
 6578
 6579
 6580
 6581
 6582
 6583
 6584
 6585
 6586
 6587
 6588
 6589
 6590
 6591
 6592
 6593
 6594
 6595
 6596
 6597
 6598
 6599
 6600
 6601
 6602
 6603
 6604
 6605
 6606
 6607
 6608
 6609
 6610
 6611
 6612
 6613
 6614
 6615
 6616
 6617
 6618
 6619
 6620
 6621
 6622
 6623
 6624
 6625
 6626
 6627
 6628
 6629
 6630
 6631
 6632
 6633
 6634
 6635
 6636
 6637
 6638
 6639
 6640
 6641
 6642
 6643
 6644
 6645
 6646
 6647
 6648
 6649
 6650
 6651
 6652
 6653
 6654
 6655
 6656
 6657
 6658
 6659
 6660
 6661
 6662
 6663
 6664
 6665
 6666
 6667
 6668
 6669
 6670
 6671
 6672
 6673
 6674
 6675
 6676
 6677
 6678
 6679
 6680
 6681
 6682
 6683
 6684
 6685
 6686
 6687
 6688
 6689
 6690
 6691
 6692
 6693
 6694
 6695
 6696
 6697
 6698
 6699
 6700
 6701
 6702
 6703
 6704
 6705
 6706
 6707
 6708
 6709
 6710
 6711
 6712
 6713
 6714
 6715
 6716
 6717
 6718
 6719
 6720
 6721
 6722
 6723
 6724
 6725
 6726
 6727
 6728
 6729
 6730
 6731
 6732
 6733
 6734
 6735
 6736
 6737
 6738
 6739
 6740
 6741
 6742
 6743
 6744
 6745
 6746
 6747
 6748
 6749
 6750
 6751
 6752
 6753
 6754
 6755
 6756
 6757
 6758
 6759
 6760
 6761
 6762
 6763
 6764
 6765
 6766
 6767
 6768
 6769
 6770
 6771
 6772
 6773
 6774
 6775
 6776
 6777
 6778
 6779
 6780
 6781
 6782
 6783
 6784
 6785
 6786
 6787
 6788
 6789
 6790
 6791
 6792
 6793
 6794
 6795
 6796
 6797
 6798
 6799
 6800
 6801
 6802
 6803
 6804
 6805
 6806
 6807
 6808
 6809
 6810
 6811
 6812
 6813
 6814
 6815
 6816
 6817
 6818
 6819
 6820
 6821
 6822
 6823
 6824
 6825
 6826
 6827
 6828
 6829
 6830
 6831
 6832
 6833
 6834
 6835
 6836
 6837
 6838
 6839
 6840
 6841
 6842
 6843
 6844
 6845
 6846
 6847
 6848
 6849
 6850
 6851
 6852
 6853
 6854
 6855
 6856
 6857
 6858
 6859
 6860
 6861
 6862
 6863
 6864
 6865
 6866
 6867
 6868
 6869
 6870
 6871
 6872
 6873
 6874
 6875
 6876
 6877
 6878
 6879
 6880
 6881
 6882
 6883
 6884
 6885
 6886
 6887
 6888
 6889
 6890
 6891
 6892
 6893
 6894
 6895
 6896
 6897
 6898
 6899
 6900
 6901
 6902
 6903
 6904
 6905
 6906
 6907
 6908
 6909
 6910
 6911
 6912
 6913
 6914
 6915
 6916
 6917
 6918
 6919
 6920
 6921
 6922
 6923
 6924
 6925
 6926
 6927
 6928
 6929
 6930
 6931
 6932
 6933
 6934
 6935
 6936
 6937
 6938
 6939
 6940
 6941
 6942
 6943
 6944
 6945
 6946
 6947
 6948
 6949
 6950
 6951
 6952
 6953
 6954
 6955
 6956
 6957
 6958
 6959
 6960
 6961
 6962
 6963
 6964
 6965
 6966
 6967
 6968
 6969
 6970
 6971
 6972
 6973
 6974
 6975
 6976
 6977
 6978
 6979
 6980
 6981
 6982
 6983
 6984
 6985
 6986
 6987
 6988
 6989
 6990
 6991
 6992
 6993
 6994
 6995
 6996
 6997
 6998
 6999
 7000
 7001
 7002
 7003
 7004
 7005
 7006
 7007
 7008
 7009
 7010
 7011
 7012
 7013
 7014
 7015
 7016
 7017
 7018
 7019
 7020
 7021
 7022
 7023
 7024
 7025
 7026
 7027
 7028
 7029
 7030
 7031
 7032
 7033
 7034
 7035
 7036
 7037
 7038
 7039
 7040
 7041
 7042
 7043
 7044
 7045
 7046
 7047
 7048
 7049
 7050
 7051
 7052
 7053
 7054
 7055
 7056
 7057
 7058
 7059
 7060
 7061
 7062
 7063
 7064
 7065
 7066
 7067
 7068
 7069
 7070
 7071
 7072
 7073
 7074
 7075
 7076
 7077
 7078
 7079
 7080
 7081
 7082
 7083
 7084
 7085
 7086
 7087
 7088
 7089
 7090
 7091
 7092
 7093
 7094
 7095
 7096
 7097
 7098
 7099
 7100
 7101
 7102
 7103
 7104
 7105
 7106
 7107
 7108
 7109
 7110
 7111
 7112
 7113
 7114
 7115
 7116
 7117
 7118
 7119
 7120
 7121
 7122
 7123
 7124
 7125
 7126
 7127
 7128
 7129
 7130
 7131
 7132
 7133
 7134
 7135
 7136
 7137
 7138
 7139
 7140
 7141
 7142
 7143
 7144
 7145
 7146
 7147
 7148
 7149
 7150
 7151
 7152
 7153
 7154
 7155
 7156
 7157
 7158
 7159
 7160
 7161
 7162
 7163
 7164
 7165
 7166
 7167
 7168
 7169
 7170
 7171
 7172
 7173
 7174
 7175
 7176
 7177
 7178
 7179
 7180
 7181
 7182
 7183
 7184
 7185
 7186
 7187
 7188
 7189
 7190
 7191
 7192
 7193
 7194
 7195
 7196
 7197
 7198
 7199
 7200
 7201
 7202
 7203
 7204
 7205
 7206
 7207
 7208
 7209
 7210
 7211
 7212
 7213
 7214
 7215
 7216
 7217
 7218
 7219
 7220
 7221
 7222
 7223
 7224
 7225
 7226
 7227
 7228
 7229
 7230
 7231
 7232
 7233
 7234
 7235
 7236
 7237
 7238
 7239
 7240
 7241
 7242
 7243
 7244
 7245
 7246
 7247
 7248
 7249
 7250
 7251
 7252
 7253
 7254
 7255
 7256
 7257
 7258
 7259
 7260
 7261
 7262
 7263
 7264
 7265
 7266
 7267
 7268
 7269
 7270
 7271
 7272
 7273
 7274
 7275
 7276
 7277
 7278
 7279
 7280
 7281
 7282
 7283
 7284
 7285
 7286
 7287
 7288
 7289
 7290
 7291
 7292
 7293
 7294
 7295
 7296
 7297
 7298
 7299
 7300
 7301
 7302
 7303
 7304
 7305
 7306
 7307
 7308
 7309
 7310
 7311
 7312
 7313
 7314
 7315
 7316
 7317
 7318
 7319
 7320
 7321
 7322
 7323
 7324
 7325
 7326
 7327
 7328
 7329
 7330
 7331
 7332
 7333
 7334
 7335
 7336
 7337
 7338
 7339
 7340
 7341
 7342
 7343
 7344
 7345
 7346
 7347
 7348
 7349
 7350
 7351
 7352
 7353
 7354
 7355
 7356
 7357
 7358
 7359
 7360
 7361
 7362
 7363
 7364
 7365
 7366
 7367
 7368
 7369
 7370
 7371
 7372
 7373
 7374
 7375
 7376
 7377
 7378
 7379
 7380
 7381
 7382
 7383
 7384
 7385
 7386
 7387
 7388
 7389
 7390
 7391
 7392
 7393
 7394
 7395
 7396
 7397
 7398
 7399
 7400
 7401
 7402
 7403
 7404
 7405
 7406
 7407
 7408
 7409
 7410
 7411
 7412
 7413
 7414
 7415
 7416
 7417
 7418
 7419
 7420
 7421
 7422
 7423
 7424
 7425
 7426
 7427
 7428
 7429
 7430
 7431
 7432
 7433
 7434
 7435
 7436
 7437
 7438
 7439
 7440
 7441
 7442
 7443
 7444
 7445
 7446
 7447
 7448
 7449
 7450
 7451
 7452
 7453
 7454
 7455
 7456
 7457
 7458
 7459
 7460
 7461
 7462
 7463
 7464
 7465
 7466
 7467
 7468
 7469
 7470
 7471
 7472
 7473
 7474
 7475
 7476
 7477
 7478
 7479
 7480
 7481
 7482
 7483
 7484
 7485
 7486
 7487
 7488
 7489
 7490
 7491
 7492
 7493
 7494
 7495
 7496
 7497
 7498
 7499
 7500
 7501
 7502
 7503
 7504
 7505
 7506
 7507
 7508
 7509
 7510
 7511
 7512
 7513
 7514
 7515
 7516
 7517
 7518
 7519
 7520
 7521
 7522
 7523
 7524
 7525
 7526
 7527
 7528
 7529
 7530
 7531
 7532
 7533
 7534
 7535
 7536
 7537
 7538
 7539
 7540
 7541
 7542
 7543
 7544
 7545
 7546
 7547
 7548
 7549
 7550
 7551
 7552
 7553
 7554
 7555
 7556
 7557
 7558
 7559
 7560
 7561
 7562
 7563
 7564
 7565
 7566
 7567
 7568
 7569
 7570
 7571
 7572
 7573
 7574
 7575
 7576
 7577
 7578
 7579
 7580
 7581
 7582
 7583
 7584
 7585
 7586
 7587
 7588
 7589
 7590
 7591
 7592
 7593
 7594
 7595
 7596
 7597
 7598
 7599
 7600
 7601
 7602
 7603
 7604
 7605
 7606
 7607
 7608
 7609
 7610
 7611
 7612
 7613
 7614
 7615
 7616
 7617
 7618
 7619
 7620
 7621
 7622
 7623
 7624
 7625
 7626
 7627
 7628
 7629
 7630
 7631
 7632
 7633
 7634
 7635
 7636
 7637
 7638
 7639
 7640
 7641
 7642
 7643
 7644
 7645
 7646
 7647
 7648
 7649
 7650
 7651
 7652
 7653
 7654
 7655
 7656
 7657
 7658
 7659
 7660
 7661
 7662
 7663
 7664
 7665
 7666
 7667
 7668
 7669
 7670
 7671
 7672
 7673
 7674
 7675
 7676
 7677
 7678
 7679
 7680
 7681
 7682
 7683
 7684
 7685
 7686
 7687
 7688
 7689
 7690
 7691
 7692
 7693
 7694
 7695
 7696
 7697
 7698
 7699
 7700
 7701
 7702
 7703
 7704
 7705
 7706
 7707
 7708
 7709
 7710
 7711
 7712
 7713
 7714
 7715
 7716
 7717
 7718
 7719
 7720
 7721
 7722
 7723
 7724
 7725
 7726
 7727
 7728

B.FR.HXB2	TCTATTTTGTGCATCAGATGCTAAAGCATATGATGATAGAGGTACATAAATGTTTGGGCCACACATGCTGTGTTACCCACAGACCCCAAGAGGATGTGTTGTTAAATGTCACAGAAAATTTTAAAC	6506
CONSENSUS_A-A1	C-----a-g-----c-----t-----c-----a-----cat-----a-----g-g-----	5833
A.BY.97BL006	C-----A-----A-----C-----T-----T-----A-----T-----A-----G-----a-----t-----A-----G-----g-----	5886
A.SE.SE7535	C-----G-----AAAG-----C-----T-----T-----A-----A-----T-----A-----G-----a-----t-----A-----G-----g-----	5729
A.SE.SE8538	C-----A-----G-----A-----G-----C-----T-----T-----A-----CAT-----AA-----A-----G-----a-----t-----A-----G-----g-----	5726
A.SE.SE8891	CT-----A-----A-----G-----G-----C-----T-----G-----T-----CAT-----C-----A-----G-----a-----t-----A-----G-----g-----	5701
A.SE.UGSEF8131	C-----C-----G-----A-----A-----G-----C-----T-----T-----A-----T-----A-----G-----a-----t-----A-----G-----g-----	5857
A1.KE.Q23-17	CT-----G-----A-----A-----G-----C-----T-----T-----A-----CAT-----AC-----A-----G-----a-----t-----A-----G-----g-----	5970
A1.SE.SE7253	C-----C-----A-----G-----AA-----G-----C-----T-----T-----A-----G-----CAT-----AA-----A-----G-----a-----t-----A-----G-----g-----	5719
A1.UG.92UG037	CT-----G-----A-----G-----A-----G-----C-----T-----G-----T-----A-----A-----G-----a-----t-----A-----G-----g-----	5882
A1.UG.U455	C-----G-----A-----G-----A-----G-----C-----T-----T-----A-----AT-----C-----A-----G-----a-----t-----A-----G-----g-----	5951
A2.CD.97CDKTB48	C-----C-----AAAG-----C-----A-----A-----A-----A-----R-----A-----A-----C-----C-----A-----A-----R-----A-----A-----G-----	5849
A2.CY.94CY017.41	C-----A-----G-----A-----G-----A-----A-----A-----A-----R-----A-----A-----C-----C-----A-----A-----R-----A-----A-----G-----	5866
CONSENSUS_B	-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	6041
B.AU.MBC200	-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----C-----	6501
B.CN.RL42	-----T-----C-----C-----A-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5889
B.DE.D31	-----T-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	6048
B.DE.HAN	-----T-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----	5968
B.ES.89SP061	-----C-----G-----G-----C-----C-----C-----T-----T-----A-----C-----A-----A-----G-----a-----t-----A-----G-----g-----	6063
B.GA.OYI	-----C-----C-----C-----GG-----C-----T-----T-----A-----A-----C-----A-----A-----G-----a-----t-----A-----G-----g-----	6045
B.GB.MANC	-----C-----C-----C-----GG-----C-----T-----T-----A-----A-----C-----A-----A-----G-----a-----t-----A-----G-----g-----	6448
B.NL.3202A21	-----C-----C-----C-----GG-----C-----T-----T-----A-----A-----C-----A-----A-----G-----a-----t-----A-----G-----g-----	6507
B.TW.TWCYS	-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	6503
B.US.JREL	-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5819
B.US.MNCG	-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	6517
B.US.RF	-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	6021
B.US.SF2	-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	6511
B.US.YU2	-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	6503
CONSENSUS_C	-----g-a-a-g-----c-----t-----t-----g-----a-----t-----a-----a-----a-----a-----a-----a-----a-----a-----a-----	5836
C.BR.92BR025	-----C-----G-----A-----G-----A-----G-----C-----T-----G-----T-----A-----A-----G-----T-----A-----A-----G-----a-----t-----A-----G-----	5861
C.BR.98BR004	-----C-----C-----A-----G-----A-----G-----C-----T-----G-----T-----A-----A-----T-----A-----A-----G-----a-----t-----A-----G-----	5840
C.BW.96BW0502	-----C-----A-----G-----G-----G-----G-----C-----T-----T-----G-----G-----T-----A-----T-----A-----A-----G-----a-----t-----A-----G-----	6007
C.BW.96BW06	-----C-----G-----A-----G-----G-----G-----C-----T-----T-----G-----G-----T-----A-----T-----A-----A-----G-----a-----t-----A-----G-----	6500
C.BW.96BW15C02	-----C-----G-----A-----G-----G-----G-----C-----T-----T-----G-----G-----T-----A-----T-----A-----A-----G-----a-----t-----A-----G-----	5975
C.ET.ETH2220	-----C-----G-----A-----G-----G-----G-----C-----T-----T-----G-----G-----T-----A-----T-----A-----A-----G-----a-----t-----A-----G-----	5915
C.IL.98ILS002	-----C-----G-----G-----A-----G-----G-----C-----T-----T-----G-----G-----T-----A-----T-----A-----A-----G-----a-----t-----A-----G-----	5867
C.IN.93IN101	-----C-----C-----C-----G-----G-----A-----G-----C-----T-----T-----G-----G-----T-----A-----T-----A-----A-----G-----a-----t-----A-----G-----	6475
C.IN.94IN476	-----C-----C-----C-----G-----G-----A-----G-----C-----T-----T-----G-----G-----T-----A-----T-----A-----A-----G-----a-----t-----A-----G-----	5834
C.IN.98IN012	-----C-----A-----G-----G-----A-----A-----A-----C-----T-----T-----G-----G-----T-----A-----T-----A-----A-----G-----a-----t-----A-----G-----	5821
C.TZ.98TZ013	-----C-----G-----G-----A-----A-----A-----C-----T-----T-----G-----G-----T-----A-----T-----A-----A-----G-----a-----t-----A-----G-----	5857
C.TZ.98TZ017	-----C-----G-----G-----A-----A-----A-----C-----T-----T-----G-----G-----T-----A-----T-----A-----A-----G-----a-----t-----A-----G-----	5878
C.ZA.97ZA012	-----C-----A-----T-----G-----A-----A-----G-----C-----T-----T-----G-----G-----T-----A-----T-----A-----A-----G-----a-----t-----A-----G-----	5848
C.ZM.96ZM651	-----C-----T-----G-----A-----A-----G-----C-----T-----T-----G-----G-----T-----A-----T-----A-----A-----G-----a-----t-----A-----G-----	5843
C.ZM.96ZM751	-----T-----G-----G-----A-----A-----G-----C-----T-----T-----G-----G-----T-----A-----T-----A-----A-----G-----a-----t-----A-----G-----	5863
CONSENSUS_D	-----t-----a-----A-----C-----A-----C-----t-----a-----c-----t-----a-----c-----t-----a-----c-----t-----a-----c-----t-----	6032
D.CD.84ZR085	-----T-----A-----A-----C-----A-----C-----T-----A-----A-----C-----A-----A-----C-----A-----A-----C-----A-----A-----C-----	6027
D.CD.FLI	-----T-----A-----A-----C-----A-----C-----T-----A-----A-----C-----A-----A-----C-----A-----A-----C-----A-----A-----C-----	6048
D.CD.Z2Z6	-----T-----A-----A-----C-----A-----C-----T-----A-----A-----C-----A-----A-----C-----A-----A-----C-----A-----A-----C-----	6505
D.KE.MB2059	-----G-----A-----AG-----C-----A-----C-----T-----A-----A-----T-----G-----A-----AAC-----A-----C-----C-----C-----	6506
D.UG.94UG1141	-----G-----A-----AG-----C-----A-----C-----T-----A-----A-----T-----G-----A-----AAC-----A-----C-----C-----C-----	5857
CONSENSUS_F1	?-----A-----?-----?-----?-----c-----?-----?-----t-----c-----aa-----a-----?-----t-----c-----aa-----a-----?-----t-----	5741
F1.BE.VI850	?-----A-----?-----?-----?-----c-----?-----?-----t-----t-----c-----AA-----A-----A-----T-----TC-----AA-----A-----G-----T-----	5852
F1.BR.93BR020.1	?-----T-----A-----A-----A-----C-----C-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5837
F1.FI.FIN9363	?-----T-----A-----A-----A-----A-----A-----C-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5838
F1.FR.MF411	GT-----G-----A-----G-----A-----A-----G-----C-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5709
F2.CM.MP255	-----G-----CC-----A-----A-----A-----A-----C-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5707
F2.CM.MP257	-----A-----G-----A-----A-----A-----A-----C-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5713
CONSENSUS_G	-----t-----?-----AG-----T-----aAg-----?-----c-----t-----t-----a-----acta-----?-----?-----A-----acta-----?-----?-----	5839
G.BE.DRCLB	-----T-----C-----AG-----T-----AAG-----A-----C-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	6466
G.FI.HH8793-12.1	-----T-----AG-----T-----AG-----A-----C-----T-----T-----G-----G-----A-----AATA-----AG-----A-----A-----A-----A-----A-----	5912
G.ING.92ING083	C-----C-----T-----AG-----T-----AAA-----C-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5871
G.SE.SE6165	-----T-----C-----T-----AG-----G-----T-----AAG-----C-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5909
H.BE.VI991	-----C-----C-----AAG-----A-----A-----T-----A-----G-----C-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5884
H.BE.VI997	-----C-----G-----GC-----AAG-----C-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----	5811
H.CF.90CF056	-----C-----G-----AAG-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5846
J.SE.SE7022	-----T-----C-----T-----AG-----T-----AAA-----C-----T-----T-----T-----A-----A-----G-----G-----A-----G-----A-----G-----	5826
J.SE.SE7887	-----T-----C-----T-----AG-----T-----AAA-----C-----T-----T-----T-----A-----A-----G-----G-----A-----G-----A-----G-----	5825
K.CD.EQTB11C	-----C-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5701
K.CM.MP535	-----C-----G-----AG-----G-----AG-----C-----T-----T-----C-----T-----C-----T-----C-----T-----C-----T-----C-----	5705

Table with columns for accession numbers (e.g., B.FR.HXB2, CONSENSUS_A-A1), nucleotide sequences, and corresponding line numbers (6712 to 5908). The sequences are aligned across the lines, showing variations between different HIV-1/SIVcpz strains.

Env gp120

Env gp120
V V I R S V N F T D N A K T I I V Q L N T S V E I N C T R P N N N T R K R I R I O R
- - V3 loop start
- - V3 tip

- B.FR.HXB2
CONSENSUS_01_AE
01_AE.CF.90CF11697
01_AE.CF.90CF402
01_AE.CF.90CF4071
01_AE.TH.93TH057
01_AE.TH.93TH065
01_AE.TH.93TH253
01_AE.TH.95TH022
01_AE.TH.95TH047
01_AE.TH.CM240
CONSENSUS_02_AG
02_AG.CM.97CM-MP807
02_AG.FR.DJ263
02_AG.FR.DJ264
02_AG.GH.G829
02_AG.NG.IBNG
02_AG.SE.SE7812
02_AG.SN.98SE-MP1211
02_AG.SN.MP1213
03_AB.RU.KAL153-2
03_AB.RU.RU98001
04_Cpx.CV.94CV032-3
04_Cpx.FR.97FPVGH
04_Cpx.FR.97FPVNY
05_DF.BE.VI1310
05_DF.BE.VI1961
CONSENSUS_06_CPX
06_CPX.AU.BEP90
06_CPX.ML.95ML127
06_CPX.ML.95ML84
06_CPX.SN.97SE1078
07_BC.CN.97CN001
07_BC.CN.98CN009
07_BC.CN.C54A
CONSENSUS_08_BC
08_BC.CN.97CNGX6F
08_BC.CN.97CNGX7F
08_BC.CN.97CNGX9F
08_BC.CN.98CNG006
10_CD.TZ.96TZBF061
10_CD.TZ.96TZBF071
10_CD.TZ.96TZBF110
CONSENSUS_11_CPX
11_CPX.CM.MP818
11_CPX.FR.MP1298
11_CPX.FR.MP1307
11_CPX.FR.MP1307
11_CPX.FR.MP1307
12_BF.AR.ARMAL59
12_BF.UY.URTR23
12_BF.UY.URTR35
N.CM.YBF106
N.CM.YBF310
CONSENSUS_07
O.CM.AN170
O.CM.AN170
O.SN.SEMP1289
O.SN.SEMP1290
O.SN.SEMP1300
CONSENSUS_CPZ
CPZ.CD.CPZANT
CPZ.CM.CAM3
CPZ.CM.CAM5
CPZ.GA.CPZGAB
CPZ.US.CPZUS

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz complete genomes

B.FR.HXB2	V3 loop tip -ATAGGAAATATGACAGACACCAATTTGTAACATTAGTAGACAAAATGGAATTAACACTTTTAAAAACAGATAGCT.....AGCAAA	7268
CONSENSUS_A-11	V3 loop end -	-----aAcA-C-CTA-G-C-TG-T-C-----aTa-----GG--A-----TG-C-----t-t-----TG-C-----a-----aac	6584
A.BY.97BL006		-----A-A-T-CTA-G-C-TG-T-C-----ATA-----GGG-C-A-A-----T-----TG-C-----GC-R-----C-A-A-----AG-	6654
A.SE.SE7535		-----ACA-T-CTA-G-C-TG-T-C-----ATA-----GGG-C-A-A-----T-----TG-C-----A-T-----GCA-----C-A-G-----TA	6509
A.SE.SE8538		-----ACA-T-CTA-G-C-TG-T-C-----ATA-----GGG-C-A-A-----T-----TG-C-----A-T-----GCA-----C-A-G-----TA	6491
A.SE.SE8891		-----ACA-T-CTA-G-C-TG-T-C-----ATA-----GGG-C-A-A-----T-----TG-C-----A-T-----GCA-----C-A-G-----TA	6448
A.SE.UGSE8131		-----ACA-T-CTA-G-C-TG-T-C-----ATA-----GGG-C-A-A-----T-----TG-C-----A-T-----GCA-----C-A-G-----TA	6655
A1.KE.Q23-17		-----ACA-T-CTA-G-C-TG-T-C-----ATA-----GGG-C-A-A-----T-----TG-C-----A-T-----GCA-----C-A-G-----TA	6705
A1.SE.SE7253		-----ACA-T-CTA-G-C-TG-T-C-----ATA-----GGG-C-A-A-----T-----TG-C-----A-T-----GCA-----C-A-G-----TA	6478
A1.UG.92UG037		-----ACA-T-CTA-G-C-TG-T-C-----ATA-----GGG-C-A-A-----T-----TG-C-----A-T-----GCA-----C-A-G-----TA	6650
A1.UG.U455		-----ACA-T-CTA-G-C-TG-T-C-----ATA-----GGG-C-A-A-----T-----TG-C-----A-T-----GCA-----C-A-G-----TA	6716
A2.CD.97CDKTB48		-----ACAG-C-CTA-----ATA-----GGG-A-----ATA-----GGG-C-A-A-----T-----TG-C-----GC-----C-A-G-----TA	6593
A2.CY.94CY017.41		-----ACA-C-CTA-----ATA-----GGG-A-----ATA-----GGG-C-A-A-----T-----TG-C-----GC-----C-A-G-----TA	6646
CONSENSUS_B		-----A-A-----TA-----g-----ata-----g-----ATA-----TG-----A-----T-----TG-----C-----t-????-aa-----	6792
B.AU.MBC200		-----G-A-----GGTA-----C-----C-----ATA-----G-----ATA-----G-----CA-----C-----G-----A-----A-----A-----A-----	7284
B.CN.RL42		-----G-A-----GGTA-----C-----C-----ATA-----G-----ATA-----G-----CA-----C-----G-----A-----A-----A-----A-----	6630
B.DE.D31		-----C-----A-----TA-----G-----C-----G-----ATA-----G-----ATA-----G-----CA-----C-----G-----A-----A-----A-----A-----	6810
B.DE.HAN		-----C-----A-----TA-----G-----C-----G-----ATA-----G-----ATA-----G-----CA-----C-----G-----A-----A-----A-----A-----	6742
B.ES.89SP061		-----C-----A-----TA-----G-----C-----G-----ATA-----G-----ATA-----G-----CA-----C-----G-----A-----A-----A-----A-----	6849
B.GA.OYI		-----AG-----CA-----GT-----C-----A-----GCC-----ATA-----C-----G-----A-----A-----C-----G-----G-----A-----G-----	6822
B.GB.MANC		-----A-----T-----TAC-----GC-----A-----ATA-----G-----G-----A-----A-----C-----G-----G-----A-----C-----T-----	7278
B.NL.3202A21		-----T-----TA-----T-----C-----G-----ATA-----G-----G-----A-----A-----C-----G-----G-----A-----C-----T-----	7222
B.TM.TWCYS		-----T-----TA-----T-----C-----G-----ATA-----G-----G-----A-----A-----C-----G-----G-----A-----C-----T-----	7251
B.US.JREL		-----T-----TA-----T-----C-----G-----ATA-----G-----G-----A-----A-----C-----G-----G-----A-----C-----T-----	6572
B.US.MNCG		-----T-----TA-----T-----C-----G-----ATA-----G-----G-----A-----A-----C-----G-----G-----A-----C-----T-----	7294
B.US.RF		-----T-----TA-----T-----C-----G-----ATA-----G-----G-----A-----A-----C-----G-----G-----A-----C-----T-----	6822
B.US.SF2		-----T-----TA-----T-----C-----G-----ATA-----G-----G-----A-----A-----C-----G-----G-----A-----C-----T-----	7282
B.US.YU2		-----G-----TA-----C-----G-----ATA-----G-----A-----A-----C-----G-----A-----C-----G-----A-----G-----A-----	7253
CONSENSUS_C		-----aCa-a-----CTA-g-----C-----g-----c-----ATA-----g-----c-----A-----a-----t-----a-----t-----a-----a-----ag-	6579
C.BR.92BR025		-----ACA-A-----CTA-g-----C-----g-----c-----ATA-----g-----c-----A-----a-----t-----a-----t-----a-----a-----ag-	6617
C.BR.98BR004		-----ACA-A-----CTA-g-----C-----g-----c-----ATA-----g-----c-----A-----a-----t-----a-----t-----a-----a-----ag-	6611
C.BW.96BW0502		-----ACA-A-----CTA-g-----C-----g-----c-----ATA-----g-----c-----A-----a-----t-----a-----t-----a-----a-----ag-	6763
C.BW.96BW15C02		-----ACA-A-----CTA-g-----C-----g-----c-----ATA-----g-----c-----A-----a-----t-----a-----t-----a-----a-----ag-	6740
C.ET.ETH2220		-----ACA-A-----CTA-g-----C-----g-----c-----ATA-----g-----c-----A-----a-----t-----a-----t-----a-----a-----ag-	6668
C.IN.98IS002		-----ACA-A-----CTA-g-----C-----g-----c-----ATA-----g-----c-----A-----a-----t-----a-----t-----a-----a-----ag-	6608
C.IN.93IN101		-----G-----ACA-A-----CTA-g-----C-----AACC-----ATA-----G-----C-----A-----G-----C-----A-----C-----G-----G-----G-----	7240
C.IN.94IN476		-----G-----ACA-A-----CTA-g-----C-----AACC-----ATA-----G-----C-----A-----G-----C-----A-----C-----G-----G-----G-----	6587
C.TZ.98TZ013		-----ACA-A-----CTA-g-----C-----G-----ATA-----G-----C-----A-----G-----C-----A-----C-----G-----G-----G-----	6625
C.TZ.98TZ017		-----ACA-A-----CTA-g-----C-----G-----ATA-----G-----C-----A-----G-----C-----A-----C-----G-----G-----G-----	6628
C.ZA.97ZA012		-----TCA-A-----CTA-g-----C-----G-----ATA-----G-----C-----A-----G-----C-----A-----C-----G-----G-----G-----	6604
C.ZM.96ZM651		-----ACA-A-----CTA-g-----C-----G-----ATA-----G-----C-----A-----G-----C-----A-----C-----G-----G-----G-----	6641
C.ZM.96ZM751		-----ACA-A-----CTA-g-----C-----G-----ATA-----G-----C-----A-----G-----C-----A-----C-----G-----G-----G-----	6628
CONSENSUS_D		-----?-----Ca-----C-----CTA-----C-----????-----aCa-a-----?-----La-----Ag-----T-----g-----g-----a-----a-----g-----	6768
D.CD.84ZR085		-----A-----CA-----GC-----CTA-----C-----A-----GT-----T-----ACA-----C-----A-----G-----TA-----AG-----T-----G-----T-----C-----CG-----G-----	6807
D.CD.FLI		-----T-----CA-----T-----C-----CTA-----T-----C-----AGATCA-----G-----TC-----TA-----AG-----T-----T-----A-----NATC-----G-----A-----C-----AG-----	6810
D.CD.Z2Z6		-----TT-----CA-----GC-----CTA-----C-----CAAAATA-----G-----A-----T-----TA-----AG-----T-----T-----A-----NATC-----G-----A-----C-----AG-----	7257
D.KE.MB2059		-----C-----C-----CT-----C-----CAAAATA-----G-----A-----T-----TA-----AG-----T-----T-----A-----NATC-----G-----A-----C-----AG-----	7250
D.UG.94UG1141		-----CA-----C-----CT-----C-----AAAGTA-----G-----A-----A-----GG-----A-----G-----G-----A-----G-----C-----G-----G-----	6592
CONSENSUS_F1		-----ACA-----?-----TA-----g-----C-----G-----ATA-----G-----c-----a-----A-----G-----G-----a-----C-----?-----g-----aAg-----	6461
F1.BE.VI850		-----ACA-A-----CTA-G-----C-----GC-----ATA-----G-----C-----A-----A-----G-----G-----C-----A-----C-----G-----G-----GT-----AG-----	6587
F1.BR.93BR020.1		-----AC-----T-----TA-----G-----C-----G-----ATA-----G-----C-----A-----A-----G-----G-----C-----A-----C-----G-----G-----GT-----AG-----	6605
F1.FI.FIN9363		-----ACA-T-----TA-----G-----C-----G-----ATA-----G-----C-----A-----A-----G-----G-----C-----A-----C-----G-----G-----GT-----AG-----	6606
F1.FR.MP411		-----ACA-----?-----TA-----g-----C-----G-----ATA-----G-----c-----a-----A-----A-----G-----T-----G-----A-----A-----A-----C-----C-----G-----A-----	6474
F2.CM.MP255		-----ACA-A-----CT-G-----C-----G-----ATA-----GG-----A-----A-----G-----T-----ATA-----A-----A-----CTG-----G-----A-----A-----TG-----AG-----	6463
F2.CM.MP257		-----AC-----?-----TA-----g-----C-----G-----ATA-----G-----c-----a-----A-----A-----G-----G-----a-----C-----?-----g-----aAg-----	6478
CONSENSUS_G		-----ACA-----g-----CTA-----g-----C-----TG-----c-----ata-----g-----c-----a-----A-----A-----G-----G-----a-----C-----?-----g-----aAg-----	6566
G.BE.DRCBL		-----ACA-----GC-----CTA-----G-----C-----GG-----C-----ATA-----G-----G-----C-----A-----A-----G-----G-----C-----G-----G-----GT-----CCAG-----	7231
G.FI.HH8793-12.1		-----ACA-----GC-----CTA-----G-----C-----GG-----C-----ATA-----G-----G-----C-----A-----A-----G-----G-----C-----G-----G-----GT-----CCAG-----	6677
G.NG.92NG083		-----ACA-----G-----CTA-----G-----C-----TG-----T-----ATA-----G-----C-----A-----A-----G-----G-----C-----A-----A-----G-----G-----GT-----CCAG-----	6627
G.SE.SE6165		-----ACA-A-----G-----CTA-G-----C-----TGCC-----ATA-----G-----C-----A-----AG-----T-----TG-----C-----A-----AG-----AG-----GG-----C-----C-----GA-----TG-----	6695
H.BE.VI991		-----G-----ACA-----CTA-----G-----C-----TG-----C-----ATA-----G-----C-----A-----A-----G-----T-----T-----TG-----G-----G-----C-----CA-----G-----ATC-----	6667
H.BE.VI997		-----G-----CA-----CTA-----G-----C-----TG-----C-----ATA-----G-----C-----A-----A-----G-----T-----T-----TG-----G-----G-----C-----CA-----G-----CAC-----	6597
H.CF.90CF056		-----G-----AC-----CTA-----G-----C-----TG-----C-----ATA-----G-----C-----A-----A-----G-----T-----T-----TG-----G-----G-----C-----CA-----G-----CAC-----	6596
J.SE.SET022		-----ACA-----TGC-----CTAG-----C-----GG-----ATA-----G-----A-----GA-----T-----T-----TA-----G-----A-----G-----T-----G-----CG-----TAG-----	6585
J.SE.SET887		-----ACA-----TGC-----CTAG-----C-----GG-----ATA-----G-----A-----GA-----T-----T-----TA-----G-----A-----G-----T-----G-----CG-----TAG-----	6578
K.CD.EOTB11C		-----A-----A-----CTA-----G-----C-----TG-----C-----ATA-----G-----C-----A-----A-----G-----G-----A-----G-----G-----C-----AG-----AAG-----	6487
K.CM.MP535		-----A-----A-----CTA-----G-----C-----TG-----T-----ATA-----G-----C-----A-----A-----G-----G-----A-----G-----G-----C-----AG-----AAG-----	6473

Table with columns for sequence identifier (e.g., B.FR.HXB2, CONSENSUS_01_AE), V3 loop tip, V3 loop end, and the corresponding nucleotide sequence. The table lists various HIV-1 and SIVcpz strains and their V3 loop regions.

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz complete genomes

Table with 3 columns: Accession ID (e.g., B.FR.HXB2, CONSENSUS_A-A1), Nucleotide Sequence (e.g., ...TTAAGACACAAATTTGGAAATAAT...), and Position (e.g., 7380, 6692, 6618).

B_FR_HXB2
CONSENSUS_01_AE
01_AE_CF_90CF11697
01_AE_CF_90CF402
01_AE_CF_90CF4071
01_AE_TH_93TH057
01_AE_TH_93TH065
01_AE_TH_93TH253
01_AE_TH_95TH022
01_AE_TH_95TH047
01_AE_TH_CM240
CONSENSUS_02_AG
02_AG_CM_97CM_MP807
02_AG_FR_DJ263
02_AG_FR_DJ264
02_AG_GH_G829
02_AG_NG_IBNG
02_AG_SE_SE7812
02_AG_SN_98SE_MP1211
02_AG_SN_MP1213
03_AB_RU_KAL153-2
03_AB_RU_RU98001
04_CX_CX_94CY032-3
04_CX_GR_97PVGH
04_CX_GR_97PVNY
05_DF_BE_V11310
05_DF_BE_V1961
CONSENSUS_06_CPX
06_CX_AU_BFP90
06_CX_ML_95ML127
06_CX_ML_95ML84
06_CX_SN_97SEL078
07_BC_CN_97CN001
07_BC_CN_98CN009
07_BC_CN_C54A
CONSENSUS_08_BC
08_BC_CN_97CNGX6F
08_BC_CN_97CNGX7F
08_BC_CN_97CNGX9F
08_BC_CN_98CN006
10_CD_TZ_96TZBF061
10_CD_TZ_96TZBF071
10_CD_TZ_96TZBF110
CONSENSUS_11_CPX
11_CX_CM_MP818
11_CX_FR_MP1298
11_CX_FR_MP1307
11_CX_GR_GR17
12_BF_AR_ARM159
12_BF_UY_URTR23
12_BF_UY_URTR35
N_CM_YBF106
N_CM_YBF30
CONSENSUS_0
O_CM_ANI70
O_CM_MVP5180
O_SN_SEMP1299
O_SN_SEMP1300
CONSENSUS_CPZ
CPZ_CD_CFPZANT
CPZ_CM_CAM3
CPZ_CM_CAM5
CPZ_GA_CFPZGAB
CPZ_US_CFPZUS
Env gp120

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz
complete genomes

	Rev Responsive Element (RRE)	Stem I
B.FR.HXB2AATGAGTCGGAGATCTTCAGACCTCGAGGAGGAGATATGGGACAAATTGGGAAAGTGAATTATATAAATATAAAGTAGTAAATAATTTGAACCATTTAGGAGTAGCACCCACCAAGGCA	7727
CONSENSUS_A-A1??-g-a-aAaT-a-c	7020
A.BY.97BL006NGAGCAAT-C	7098
A.SE.SE7535G-ACAAAT-A-C	6968
A.SE.SE8538ACAAAT-A-C	6944
A.SE.SE8891AAAGT-A-C	6892
A.SE.UGSE8131AAT-CAA-TAAT-A-C	7105
A1.KE.O23-17TAAAT-A-C	7164
A1.SE.SB7253G-ACAAAT-A-C	6922
A1.UG.92UG037G-CAGT-A-C	7088
A1.UG.U455CAA-AAA-T-C	7157
A2.CD.97CDKTB48G-ATTAAT-C	7037
A2.CY.94CY017.41GGGACTAAT-C	7099
CONSENSUS_B	?????pac?-c-c-a	7232
B.AU.MBC200G-CAC-GT-C	7749
B.CN.RL42AGC-GCCCA-C	7077
B.DE.D31GAC-C-A-C	7251
B.DE.HANAGT-G-ACCA-A	7186
B.ES.89SP061ACC-G-ACCA-T-G	7302
B.GA.OYIACG-G-AT-A-C	7263
B.GB.MANCGGG-CACCA-C	7684
B.NL.3202A21AAG-C-G-A-T	7731
B.ITW.TWCYSG-GT-G	7698
B.US.JRFLG-A	7016
B.US.MNCGGACAG-C-CA	7747
B.US.RFACT-ACTA-A-T	7269
B.US.SF2ACT-CA-G	7729
B.US.YU2-C-G-A-T	7688
CONSENSUS_C	?????????aaa?-a-ca-a	7018
C.BR.92BR025C-CA-A-A-G	7064
C.BW.98BR004-CA-A-A-G	7082
C.BW.96BM0502CCA-CA-A-A-G	7210
C.BW.96BM06AAT-CAGT-CA	7697
C.BW.96BM15C02AAT-TCACCA-A-ACA	7184
C.ET.ETH2220CAT-GCAGAAA-A	7106
C.IL.98IS002ACA-C-A-A-CA	7058
C.IN.93IN101GATACAG-A-A-A-T	7702
C.IN.94IN476G-AGCA-A-A	7040
C.IN.98IN012AGTCCAGAAA-T-A-A-CA	7063
C.TZ.98TZ013AAC-CA-G-AAT-CA	7081
C.TZ.98TZ017-C-CAA-G-A	7033
C.ZA.97ZA012-AACAG-A	7024
C.ZM.96ZM651AGCAG-GACAC-A-CA	7082
C.ZM.96ZM751ACAG-GACAC-A-CA	7072
CONSENSUS_D	???-g-c-aAT-C	7206
D.CD.84ZE085-C-C-AAT-C-T	7248
D.CD.FLIG-ACTAA-C-T	7263
D.CD.Z266G-TCTAA-C	7720
D.KE.MB2059G-C-AAT-C	7697
D.UG.94UG1141AGT-GCC-AAT-C	7045
CONSENSUS_F1	??????-g-a-La-T	6862
F1.BE.VI850-G-A-CATT-C	7004
F1.BR.93BR020.1G-ACTAAT-C	7031
F1.FI.FIN9363AATPAAT-G--T-T-C	7023
F1.FR.MF411GATFACT-AG-A-T-CT	6894
F2.CM.MP255G-C-G-AAA-TC	6886
F2.CM.MP257-G-C-C-C	6919
CONSENSUS_G	?????????gca-?-?-agt-?	6978
G.BE.DRCBLGCAGTAGT-AAA	7666
G.FI.HH8793-12.1AATGCA-G-A-T-C	7097
G.NG.92NG083GAC-G-ACAGAG-C	7053
G.SE.SE6165AATACAAATACA-G--AGT	7145
H.BE.VI991CAA-TAAT-TT-C	7126
H.BE.VI997TAAAT-TT-C	7041
H.CF.90CF056GCA--AAAT-T-C	7046
J.SE.SE7022GAG---GCA-T-C	7041
J.SE.SE7887GGCAGT...GAG---GCA-T-AC	7034
K.CD.EOTB11CAATFACT-GGACTGAG-C	6910
K.CM.MP535C-TAAT-C	6887

Accession	Sequence	Rev Responsive Element (RRE)
B.FR.HXB2AATGAGTCCGAGATCTTCAGACCTCGAGGAGGAGATATGAGGGAATAATGAGGAAAGTGAATATATAAATATAAAGTAAAGCAATTGAAACCAATTAGGAGTAGCACCCACCAAGGCA	7727
CONSENSUS_01_AE	t?????.....C-----A-a-A-----C-----C-----a-----g-----	7637
01_AE.CF.90CF11697	T.....ACT-CA-AAAAT-----C-----A-----A-----A-----G-----	7679
01_AE.CF.90CF402	TCAG.....ACT-CA-AAAAT-----C-----A-----A-----A-----G-----	7790
01_AE.CF.90CF4071	TGAGACC.....AT-A-----C-----C-----C-----A-----G-----	7653
01_AE.TH.93TH057	GAGT.....AA-----C-----A-----A-----A-----G-----	7664
01_AE.TH.93TH065	TACGAAT.....AA-----C-----A-----A-----A-----G-----	7649
01_AE.TH.93TH253	T.....ACG--A-TGAG-----C-----A-----A-----A-----G-----	7711
01_AE.TH.95TH022	T.....CAATTAAT-----C-----A-----A-----A-----T-----	7675
01_AE.TH.95TH047	G.....AGTAGT-----C-----A-----A-----A-----G-----	7684
01_AE.TH.CM240CG--TAA-----C-----A-----C-----A-----G-----	7280
CONSENSUS_02_AGg-acaagt-----C-----g-----g-----C-----t-----cgt-----	6889
02_AG.CM.97CM-MP807	T.....g-acaagt-----C-----g-----g-----C-----t-----cgt-----	6910
02_AG.FR.DJ263G-ACAAAT-----C-----G-----G-----C-----T-----G-----	7102
02_AG.FR.DJ264G-ACAAAT-----C-----G-----G-----C-----T-----G-----	7061
02_AG.GH.G829G-ACAAAT-----C-----G-----G-----C-----T-----G-----	6922
02_AG.NG.IBNGG-ACAAAT-----C-----G-----G-----C-----T-----G-----	7233
02_AG.SE-SE7812G-ACAAAT-----C-----G-----G-----C-----T-----G-----	7114
02_AG.SN.96SE-MP1211GTGAT-----C-----G-----G-----C-----T-----G-----	6929
02_AG.SN.MP1213G-ACAAAT-----C-----G-----G-----C-----T-----G-----	6909
03_AB.RU.KAL153-2TAA-----T-----T-----A-----T-----G-----	7447
03_AB.RU.RU98001TCA-----T-----T-----A-----T-----G-----	7042
04_cpx.CV.94CY032-3A-T-----C-----C-----G-----G-----A-----A-----AT-----	7097
04_cpx.GR.97PVCHTAAAT-----C-----C-----G-----G-----A-----A-----GG-----	7826
04_cpx.GR.97PVNYGAT-C-GCAAT-----C-----C-----G-----G-----T-----G-----	7775
05_DF.BE.VI1310GATAGTCC-C-CA-A-----A-----C-----C-----C-----A-----G-----	7133
05_DF.BE.VI961ACT-----TA-A-----CA-----C-----C-----C-----G-----	7084
CONSENSUS_06_cpx?-----?-----a?cagtt-----?-----?-----A-----C-----A-----?-----	7577
06_cpx.AU.BEP90GAT-GCA-TAGT-----C-----C-----G-----G-----A-----C-----A-----TG-----	7717
06_cpx.ML.95ML127GAT-GCA-TAGT-----C-----C-----G-----G-----A-----C-----A-----TG-----	7717
06_cpx.ML.95ML84C-ACCAGT-----A-----A-----G-----G-----A-----C-----A-----G-----	6931
06_cpx.SN.97SE1078A-T-A-----A-----A-----G-----G-----A-----C-----GA-----	7762
07_BC.CN.97CN001	A.....TA-A-----CA-----A-----G-----G-----A-----G-----T-----CA-----	7049
07_BC.CN.98CN009	A.....TA-A-----CA-----A-----G-----G-----A-----G-----T-----CA-----	7049
07_BC.CN.C54A	A.....TA-A-----CA-----A-----G-----G-----A-----G-----T-----CA-----	6907
CONSENSUS_08_BC?-----?-----a-----La-A-----A-----A-----G-----G-----TGCA-----	6887
08_BC.CN.97CNGX6F	A.....A-TA-A-----A-----A-----A-----A-----G-----G-----TGCA-----	6906
08_BC.CN.97CNGX7F	A.....A-TA-A-----A-----A-----A-----A-----G-----G-----TGCA-----	6900
08_BC.CN.97CNGX9F	A.....A-TA-A-----A-----A-----A-----A-----G-----G-----TGCA-----	6888
08_BC.CN.98CN006AATAAA-CAA-GTA-----A-----A-----G-----G-----A-----G-----TGCA-----	7061
10_CD.TZ.96TZBF061	TAAT.....AAT-G-C--AAT-----C-----C-----C-----C-----A-----TC-----	7078
10_CD.TZ.96TZBF071	TAAT.....ACT-G-C--AAT-----C-----C-----C-----C-----C-----TC-----	7111
10_CD.TZ.96TZBF110	TAAT.....AGTAATC-GAGTGGG-----A-----A-----GG-----C-----TC-----	7092
CONSENSUS_11_cpx?????-?-?-c?c?A?l-----C-----AC-----G-----G-----?-----g-----	7510
11_cpx.CM.MP818G-ACAAAT-----C-----T-----AC-----G-----G-----A-----C-----G-----	7705
11_cpx.FR.MP1298CCACCAGT-----C-----T-----AC-----G-----G-----A-----C-----G-----	7721
11_cpx.FR.MP1307GAC-CA-TAGT-----C-----A-----A-----G-----G-----A-----C-----T-----G-----	7759
11_cpx.GR.GR17GATACC-T-G-AAA-----C-----AC-----G-----G-----A-----C-----T-----G-----	6994
12_BF.AR.ARMA159	ACT.....C-A-T-A-C-----G-----A-----AA-----G-----G-----C-----G-----	7692
12_BF.UY.URTR23GCTACT--A-A-T-----T-----G-----A-----AA-----G-----G-----C-----A-----	7725
12_BF.UY.URTR35	TCAT.....GACTGGG--CG-A-T-----C-----G-----A-----AA-----G-----G-----	7750
N.CM.YBF106A-----GAAACA--AG-ATAT--CTC-----A-C-----GTTA-TCTC-----CAA--G-G-----G-C-----T-GT-A-----CA-----AGGT--A--T	7272
N.CM.YBF30A-----GAAACA--AG-ATAT--CTC-----A-C-----GTTA-TCTC-----CAA--G-G-----G-C-----T-GT-A-----CA-----AGGT--A--T	7272
O.CM.ANT70?????aac--cagc?Aa--cc-CA--T-----Aat--G-----AA--T-TA-----C-----G-tC-----?-----C-----G-tC-----g?AA-----T-TA-T-----g-----t-A--aATT	7402
O.CM.MVP5180AGC-----CA-CAAT--TA-CA--T-----AAT--G-----AA--T-TA-----C-----G-tC-----C-----G-tC-----G--AA-----T-TA-T-----G-----ACGTATT	7769
O.SN.SEMP1299AGC-----CA-CAAT--TA-CA--T-----AAT--G-----AA--T-TA-----C-----G-tC-----C-----G-tC-----G--AA-----T-TA-T-----G-----T-A--A--AATT	7783
O.SN.SEMP1300AGC-----CA-CAAT--TA-CA--T-----AAT--G-----AA--T-TA-----C-----G-tC-----C-----G-tC-----G--AA-----T-TA-T-----G-----T-A--A--AATT	7819
CONSENSUS_CPZ?????-?-?-a?taataKa--cag-gta?-----aC-----?-----?-----?-----?-----?-----g-----g-----t-gc-a-----ca-----A-----t-----a-----t	6879
CPZ_CD.CPZANTAACT--AA--TT--AGG-GTC-----CT-C-A-----TAGCA.....C-A-----GCG--G--G--CC--GG--CC-G-----G-----G-----RT--C-----G-----T-----A--CA--GN	7150
CPZ_CD.CAM3AACAGG--TCACAA--ATC--CGG--ATAC--AC-----A-C-----GTA-----TTG-----GCT-----GG-----C-----G-----T-----GC-----A-----CA-----A-----G-----T-----A-----T	7137
CPZ_CD.CAM5AAC-G--TATA--TATA--CAG-GTAC--TC-----G-----A-----A-----TTG-----GCT-----GG-----C-----G-----T-----GC-----A-----CA-----A-----A-----T-----T	7443
CPZ_GA.CPZGABAAC-G--TATA--TATA--CAG-GTAC--TC-----G-----A-----A-----TTG-----GCT-----GG-----C-----G-----T-----GC-----A-----CA-----A-----A-----T-----T	7749
CPZ_US.CPZUS	C.....GAT-G-AG-AATATA-CAG-G-TG--AC-----A-----C-----CTTTA-----A-----A-----C-----C-----G-----G-----G-----C-----GC-----A-----TA-----G-----AGGT-----C-----	7729
Env gp120N E S E I F R R P G G G D M R D N W R S E L Y K Y K V V K I E F L G V A P T K A	Env

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz complete genomes

Table with columns: Accession ID (e.g., B.FR.HXB2, CONSENSUS_A-A1), Genomic coordinates (Env gp120 end to Env gp41 start), and Nucleotide sequence. Includes structural annotations like RRE, Stem I, Stem IIA, and Stem IIB.

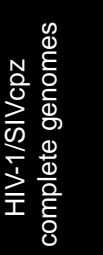
	Stem IIB' !!!	Stem IIC	Stem IIC'	Stem IIA'	Stem III-IV	Stem III-IV	Stem V	Stem V'
B.FR.HXB2	CGCTGACGCTACAGCCAGACAAATTAATTTGCTGTGATATAGTGCAGCAGCAGACAAATTTGCTGTGAGGGGCTATTGAGGGGCAACAGCATCTGTTGCACTCACAGTCTCTGGGGCATCAAGCAGGCTCCAGGCAAG							
CONSENSUS_01_AE	---	---	---	---	---	---	---	---
01_AE.CF.90CF11697	---	---	---	---	---	---	---	---
01_AE.CF.90CF402	---	---	---	---	---	---	---	---
01_AE.CF.90CF4071	---	---	---	---	---	---	---	---
01_AE.TH.93TH057	---	---	---	---	---	---	---	---
01_AE.TH.93TH065	---	---	---	---	---	---	---	---
01_AE.TH.93TH253	---	---	---	---	---	---	---	---
01_AE.TH.95TH022	---	---	---	---	---	---	---	---
01_AE.TH.95TH047	---	---	---	---	---	---	---	---
01_AE.TH.CM240	---	---	---	---	---	---	---	---
CONSENSUS_02_AG	---	---	---	---	---	---	---	---
02_AG.CM.97CM-MP807	---	---	---	---	---	---	---	---
02_AG.FR.DJ263	---	---	---	---	---	---	---	---
02_AG.FR.DJ264	---	---	---	---	---	---	---	---
02_AG.GH.GB29	---	---	---	---	---	---	---	---
02_AG.NG.IBNG	---	---	---	---	---	---	---	---
02_AG.SE-SE7812	---	---	---	---	---	---	---	---
02_AG.SN.98SE-MP1211	---	---	---	---	---	---	---	---
02_AG.SN.MP1213	---	---	---	---	---	---	---	---
03_AB.RU.KAL153-2	---	---	---	---	---	---	---	---
03_AB.RU.RU98001	---	---	---	---	---	---	---	---
04_CPX.CV.94CY032-3	---	---	---	---	---	---	---	---
04_CPX.GR.97PVCH	---	---	---	---	---	---	---	---
04_CPX.GR.97PVMY	---	---	---	---	---	---	---	---
05_DF.BE.VI1310	---	---	---	---	---	---	---	---
05_DF.BE.VI961	---	---	---	---	---	---	---	---
CONSENSUS_06_CPX	---	---	---	---	---	---	---	---
06_CPX.AU.BEP90	---	---	---	---	---	---	---	---
06_CPX.ML.95ML127	---	---	---	---	---	---	---	---
06_CPX.ML.95ML84	---	---	---	---	---	---	---	---
06_CPX.SN.97SE1078	---	---	---	---	---	---	---	---
07_BC.CN.97CN001	---	---	---	---	---	---	---	---
07_BC.CN.98CN009	---	---	---	---	---	---	---	---
07_BC.CN.C54A	---	---	---	---	---	---	---	---
CONSENSUS_08_BC	---	---	---	---	---	---	---	---
08_BC.CN.97CNGX6F	---	---	---	---	---	---	---	---
08_BC.CN.97CNGX7F	---	---	---	---	---	---	---	---
08_BC.CN.97CNGX9F	---	---	---	---	---	---	---	---
08_BC.CN.98CN006	---	---	---	---	---	---	---	---
10_CD.TZ.96TZBF061	---	---	---	---	---	---	---	---
10_CD.TZ.96TZBF071	---	---	---	---	---	---	---	---
10_CD.TZ.96TZBF110	---	---	---	---	---	---	---	---
CONSENSUS_11_CPX	---	---	---	---	---	---	---	---
11_CPX.CM.MP818	---	---	---	---	---	---	---	---
11_CPX.FR.MP1298	---	---	---	---	---	---	---	---
11_CPX.FR.MP1307	---	---	---	---	---	---	---	---
11_CPX.GR.GR17	---	---	---	---	---	---	---	---
12_BF.AR.ARMA159	---	---	---	---	---	---	---	---
12_BF.UY.URTR23	---	---	---	---	---	---	---	---
12_BF.UY.URTR35	---	---	---	---	---	---	---	---
N.CM.YBF106	---	---	---	---	---	---	---	---
N.CM.YBF30	---	---	---	---	---	---	---	---
CONSENSUS_0	---	---	---	---	---	---	---	---
O.CM.ANT70	---	---	---	---	---	---	---	---
O.CM.MVP5180	---	---	---	---	---	---	---	---
O.SN.SEMP1259	---	---	---	---	---	---	---	---
O.SN.SEMP1300	---	---	---	---	---	---	---	---
CONSENSUS_CPFZ	---	---	---	---	---	---	---	---
CPFZ.CD.CPZANT	---	---	---	---	---	---	---	---
CPFZ.CM.CAM3	---	---	---	---	---	---	---	---
CPFZ.CM.CAM5	---	---	---	---	---	---	---	---
CPFZ.GA.CPZGAB	---	---	---	---	---	---	---	---
CPFZ.US.CPZUS	---	---	---	---	---	---	---	---
Env gp41	TLTLTVQARQLLSGLIVQQQNNLILRLAIEAQQLLQLLTLTVWGLIKQLLQAR							

HIV-1/SIVcpz complete genomes

Accession/Strain	Genomic Region	Sequence	Position
B.FR.HXB2	Stem I'	AAATCCTGGCTGTGGAAGATACATAAAGGATCAACAGCTCCTGGGGATTGGGGTTCCTGGAAAACTCATTTGGCACCACCTGCTGGCCTTTGGAATGTGGAGTAAT.....	8072
CONSENSUS_01_AE	RRE end -	8072
01_AE.CF.90CF11697		7982
01_AE.CF.90CF402		8024
01_AE.CF.90CF4071		8135
01_AE.TH.93TH057		7998
01_AE.TH.93TH065		8009
01_AE.TH.93TH253		7994
01_AE.TH.95TH022		8056
01_AE.TH.95TH047		8020
01_AE.TH.CM240		7625
CONSENSUS_02_AG		7233
02_AG.CM.97CM-MP807		7252
02_AG.FR.DJ263		7447
02_AG.FR.DJ264		7406
02_AG.GH.G829		7267
02_AG.NG.IBNG		7578
02_AG.SE.SE7812		7462
02_AG.SN.98SE-MP1211		7274
02_AG.SN.MP1213		7254
03_AB.RU.KAL153-2		7792
03_AB.RU.RU98001		7387
04_CPX.CV.94CY032-3		7442
04_CPX.FR.97FVCH		8171
04_CPX.FR.97FVNY		8123
05_DF.BE.VI1310		7478
05_DF.BE.VI961		7429
CONSENSUS_06_CPX		7919
06_CPX.AU.BEP90		8120
06_CPX.ML.95ML127		8065
06_CPX.ML.95ML84		7276
06_CPX.SN.97SE1078		8110
07_BC.CN.97CN001		7394
07_BC.CN.98CN009		7394
07_BC.CN.C54A		7252
CONSENSUS_08_BC		7232
08_BC.CN.97CNGX6F		7251
08_BC.CN.97CNGX7F		7245
08_BC.CN.97CNGX9F		7233
08_BC.CN.98CN006		7406
10_CD.TZ.96TZBF061		7423
10_CD.TZ.96TZBF071		7456
10_CD.TZ.96TZBF110		7437
CONSENSUS_11_CPX		7847
11_CPX.CM.MP818		8050
11_CPX.FR.MP1298		8069
11_CPX.FR.MP1307		8107
11_CPX.FR.MP1307		7339
12_BF.AR.ARMA159		8037
12_BF.UY.URTR23		8070
12_BF.UY.URTR35		8095
N.CM.YBF106		7575
N.CM.YBF30		7620
CONSENSUS_0		7747
O.CM.ANT70		8126
O.CM.MVP5180		8143
O.SN.SEMP1259		8179
O.SN.SEMP1300		8149
CONSENSUS_CPZ		7217
CPZ.CD.CPZANT		7522
CPZ.CM.CAM3		7485
CPZ.CM.CAM5		7791
CPZ.GA.CPZGAB		8112
CPZ.US.CPZUS		8077
Env gp41		Env

HIV-1/SIVcpz complete genomes

B_FR_HXB2 ... AGCACTGGATGGAGTGGACAGAGAAAATTACACAGCTTAAATACACTCTTTAATTTGAAGAATCGCAAAACCCAGCAAGAAA ... 8187



HIV-1/SIVcpz complete genomes

Table with columns for genome identifiers (e.g., B.FR.HXB2, A.BY.97BL006), sequence coordinates (8441 to 7601), and nucleotide sequences. Includes a legend for Tat Rev intron end and premature stop in HXB2 Tat.

Table with columns for accession numbers (e.g., B.FR.HXB2, CONSENSUS_01_AE), sequence identifiers (e.g., Tat Rev intron end, Premature stop), and nucleotide sequences (e.g., TTTTGTGTA...).

HIV-1/SIVcpz complete genomes

Accession	Sequence	Position
B.FR.HXB2	AACTTCTGGG.....CGCAGGGGGTGGGAAGCTCAAAATATTGGTGAATCTCTACAGTATTGGAGTTCAGGAACATAAGAAATAGTCTGTTAGCTTGGCTCAATGCCACAGC	8680
CONSENSUS_01_AE	---CsCAGAGTCTcaaggactga-ac	8680
01_AE.CF.90CF11697	---CACAGCTCTCAAGGACTGA-AC	8653
01_AE.CF.90CF402	---CACAGAGTCTACAGGACTGA-AC	8764
01_AE.CF.90CF4071	---CACAGAGTCTCAGGACTGA-AC	8627
01_AE.TH.93TH057	---CACAGAGTCTCAAGGACTGA-AC	8638
01_AE.TH.93TH065	---CGCAGAGTCTCAAGGACTGA-AC	8623
01_AE.TH.93TH253	---CACAGAGTCTCAAGGACTGA-AC	8688
01_AE.TH.95TH022	---CACAGAGTCTCAAGGACTGA-AC	8649
01_AE.TH.95TH047	---CACAGAGTCTCAAGGACTGA-AC	8658
01_AE.TH.CM240	---CACAGAGTCTCAAGGACTGA-AC	8254
CONSENSUS_02_AG	---CACagagctcaaggactga-act	7858
02_AG.CM.MP807	---CACAGAGTCTCAAGGACTGA-ACT	7881
02_AG.FR.DJ263	---CACAGAGCTCAAGGACTGA-ACT	8076
02_AG.FR.DJ264	---CACAGAGCTCAAGGACTGA-ACT	8035
02_AG.GH.G829	---CACAGAGTCTCAAGGACTGA-ACT	7875
02_AG.NG.IBNG	---CACAGAGTCTCAAGGACTGA-ACT	8207
02_AG.SE.SF7812	---CACAGAGTCTCAAGGACTGA-ACT	8070
02_AG.SN.985E-MP1211	---CACAGAGTCTCAAGGACTGA-ACT	7906
02_AG.SN.MP1213	---CACAGAGTCTCAAGGACTGA-ACT	7862
03_AB.RU.KAL153-2ATA	8400
03_AB.RU.RU98001ATA	7995
04_CPX.CV.94CV032-3ATA	8053
04_CPX.FR.97FVGHATA	8779
04_CPX.FR.97FVNYATA	8734
05_DF.BE.VI1310ATA	8086
CONSENSUS_06_CPXATA	8524
06_CPX.AU.BFP90ATA	8728
06_CPX.ML.95ML127ATA	8673
06_CPX.ML.95ML84ATA	7884
06_CPX.SN.97SE1078ATA	8718
07_BC.CN.97CN001ATA	8023
07_BC.CN.98CN009ATA	8032
07_BC.CN.C54AATA	7881
CONSENSUS_08_BCATA	7861
08_BC.CN.97CNX6FATA	7880
08_BC.CN.97CNX7FATA	7874
08_BC.CN.97CNX9FATA	7862
08_BC.CN.98CN006ATA	8035
10_CD.TZ.96TZBF061ATA	8031
10_CD.TZ.96TZBF071ATA	8061
10_CD.TZ.96TZBF110ATA	8045
CONSENSUS_11_CPXATA	8440
11_CPX.CM.MP818ATA	8658
11_CPX.FR.MP1298ATA	8680
11_CPX.FR.MP1307ATA	8715
11_CPX.FR.MP1307ATA	7953
12_BF.AR.ARMA159ATA	8645
12_BF.UY.URTR23ATA	8678
12_BF.UY.URTR35ATA	8703
N.CM.YBF106ATA	8210
N.CM.YBF30ATA	8252
CONSENSUS_0ATA	8350
O.CM.AN170ATA	8749
O.SN.MP95189ATA	8775
O.SN.SEMP1250ATA	8817
O.SN.SEMP1300ATA	8787
CONSENSUS_CPZATA	7805
CPZ_CD.CPZANTATA	8139
CPZ_CM.CAM3ATA	8117
CPZ_CM.CAM5ATA	8423
CPZ_GA.CPZGABATA	8741
CPZ_US.CPZUSATA	8709
RevATA	8709
Env gp41ATA	8709

HIV-1/SIVcpz complete genomes

B_FR_HXB2	8796	CAATAGCAGTAGCTGAGGGGACAGATAGGTTATAGAAGTGTACAAAGGAGCTTGTAGAGCTATTGCCACATACCTAGAAAGTAAGACAGGGCTTGGAAAGGATTTTCTATAA.G.....
CONSENSUS_01_AE	8725	A-----G-G-T-----c-----g-----c-----t-----c-----a-----gc-----c-----
01_AE_CF_90CF11697	8769	AG-----G-G-T-----G-----GA-----G-----C-----AT-----C-----A-----GC-----C-----
01_AE_CF_90CF402	8880	A-----G-G-T-----A-----A-----G-----C-----T-----T-----C-----A-----GC-----C-----
01_AE_CF_90CF4071	8743	A-----G-G-T-----C-----A-----C-----A-----T-----T-----C-----A-----GC-----C-----
01_AE_TH_93TH057	8754	A-----G-G-T-----C-----C-----G-----C-----T-----A-----C-----C-----A-----GC-----C-----
01_AE_TH_93TH065	8739	A-----G-G-T-----C-----C-----C-----C-----C-----T-----A-----G-----C-----C-----
01_AE_TH_93TH253	8804	AG-----T-----C-----G-----T-----C-----C-----C-----T-----A-----G-----C-----C-----
01_AE_TH_95TH022	8765	A-----T-----C-----G-----T-----C-----C-----C-----T-----A-----G-----C-----C-----
01_AE_TH_95TH047	8774	A-----G-G-T-----A-----A-----G-----C-----T-----A-----G-----C-----C-----
01_AE_TH_CM240	8370	A-----C-----G-G-T-----C-----C-----G-----C-----T-----A-----G-----C-----C-----
CONSENSUS_02_AG	7971	A-----a-----c-----t-----a-----?-----a-----?-----a-----?-----a-----t-----a-----c-----c-----
02_AG_CM_97CM-MP807	7997	A-----A-----C-----T-----A-----G-----A-----G-----A-----G-----A-----G-----A-----G-----C-----
02_AG_FR_DJ263	8192	AT-----A-----T-----T-----C-----A-----A-----G-----A-----T-----G-----A-----T-----C-----T-----
02_AG_FR_DJ264	8151	A-----C-----A-----C-----A-----A-----T-----G-----A-----T-----G-----A-----T-----C-----T-----
02_AG_GH_G829	7991	A-----T-----C-----A-----C-----A-----C-----A-----G-----G-----C-----T-----T-----T-----
02_AG_NG_IBNG	8323	A-----T-----A-----C-----A-----C-----A-----G-----A-----T-----G-----A-----T-----G-----
02_AG_SE_SE7812	8186	A-----T-----A-----C-----A-----C-----A-----G-----A-----T-----G-----A-----T-----G-----
02_AG_SN_98SE-MP1211	8022	AG-----A-----C-----T-----A-----A-----A-----A-----G-----A-----T-----A-----A-----G-----
02_AG_SN_MP1213	7978	A-----AC-----A-----C-----T-----A-----A-----A-----A-----G-----A-----T-----A-----A-----G-----
03_AB_RU_KAL153-2	8516	A-----G-----G-----T-----A-----A-----G-----A-----A-----T-----T-----A-----G-----A-----G-----
03_AB_RU_RU98001	8111	A-----G-----G-----T-----A-----A-----G-----A-----A-----T-----T-----A-----G-----A-----G-----
04_cpx_CY_94CY032-3	8169	A-----A-----A-----A-----A-----C-----C-----C-----G-----A-----G-----A-----T-----T-----A-----
04_cpx_GR_97PVGH	8895	A-----A-----A-----A-----A-----C-----C-----C-----G-----A-----G-----A-----T-----T-----A-----
04_cpx_GR_97PVNY	8851	AG-----A-----A-----C-----C-----G-----G-----A-----A-----T-----T-----A-----A-----G-----
05_DF_BE_V11310	8202	AG-----T-----A-----A-----C-----C-----T-----T-----G-----A-----A-----G-----A-----T-----A-----
05_DF_BE_V1961	8153	AG-----T-----A-----A-----C-----C-----T-----T-----G-----A-----A-----G-----A-----T-----A-----
CONSENSUS_06_cpx	8636	A-----A-----C-----T-----?-----?-----?-----T-----T-----A-----?-----A-----a-----t-----a-----?
06_cpx_AU_BFP90	8844	A-----A-----C-----T-----G-----C-----A-----A-----A-----T-----T-----T-----T-----A-----G-----
06_cpx_ML_95ML127	8789	A-----A-----C-----T-----A-----A-----A-----A-----T-----T-----G-----G-----T-----T-----A-----
06_cpx_ML_95ML84	8000	A-----A-----C-----T-----A-----A-----A-----A-----T-----T-----G-----G-----T-----T-----A-----
06_cpx_SN_97SE1078	8834	A-----A-----T-----T-----A-----A-----G-----A-----T-----T-----C-----T-----T-----A-----G-----
07_BC_CN_97CN001	8139	A-----A-----A-----A-----T-----T-----A-----C-----T-----A-----A-----C-----T-----A-----A-----
07_BC_CN_98CN009	7997	A-----A-----A-----A-----T-----T-----A-----C-----T-----A-----A-----C-----T-----A-----A-----
07_BC_CN_C54A	7977	A-----A-----A-----A-----A-----A-----C-----A-----A-----T-----C-----A-----A-----G-----
CONSENSUS_08_BC	7996	A-----A-----A-----A-----A-----A-----C-----A-----A-----T-----C-----A-----A-----G-----
08_BC_CN_97CNGX6F	7990	A-----A-----A-----A-----A-----A-----C-----A-----A-----T-----C-----A-----A-----G-----
08_BC_CN_97CNGX7F	7978	A-----A-----A-----A-----A-----A-----C-----A-----A-----T-----C-----A-----A-----G-----
08_BC_CN_97CNGX9F	8151	A-----A-----A-----A-----A-----A-----C-----A-----A-----T-----C-----A-----A-----G-----
08_BC_CN_98CN006	8147	A-----A-----A-----A-----A-----A-----C-----A-----A-----T-----C-----A-----A-----G-----
10_CD_TZ_96TZBF061	8177	AG-----G-----C-----T-----Z-----9-----6-----T-----Z-----B-----F-----0-----6-----1-----
10_CD_TZ_96TZBF071	8161	AG-----G-----C-----T-----Z-----9-----6-----T-----Z-----B-----F-----0-----7-----1-----
10_CD_TZ_96TZBF110	8551	AG-----G-----C-----T-----Z-----9-----6-----T-----Z-----B-----F-----1-----1-----0-----
CONSENSUS_11_cpx	8810	A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
11_cpx_CM_MP818	8069	A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
11_cpx_FR_MP1298	8761	A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
11_cpx_FR_MP1307	8794	A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
11_cpx_GR_GR17	8819	A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
12_BF_AR_ARMAL59	8326	A-----T-----G-----T-----G-----T-----G-----T-----G-----T-----G-----T-----G-----T-----G-----
12_BF_UY_URTR23	8368	A-----T-----G-----T-----G-----T-----G-----T-----G-----T-----G-----T-----G-----T-----G-----
12_BF_UY_URTR35	8865	A-----T-----G-----T-----G-----T-----G-----T-----G-----T-----G-----T-----G-----T-----G-----
N_CM_YBF106	8891	T-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
N_CM_YBF30	8933	T-----T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
O_CM_ANTI70	8903	AG-----G-----C-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
O_CM_ANTI70	7915	AG-----G-----C-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----
O_SN_SEMP1299	8256	O-----S-----N-----S-----E-----M-----P-----1-----2-----9-----9-----
O_SN_SEMP1300	8233	O-----S-----N-----S-----E-----M-----P-----1-----3-----0-----0-----
CONSENSUS_CPZ	8869	CPZ-----G-----A-----C-----P-----Z-----G-----A-----C-----P-----Z-----G-----A-----
CPZ_CD_CFPZANT	8823	CPZ-----G-----A-----C-----P-----Z-----G-----A-----C-----P-----Z-----G-----A-----
CPZ_CM_CAM3	8869	CPZ-----G-----A-----C-----P-----Z-----G-----A-----C-----P-----Z-----G-----A-----
CPZ_CM_CAM5	8823	CPZ-----G-----A-----C-----P-----Z-----G-----A-----C-----P-----Z-----G-----A-----
CPZ_GA_CFPZGAB	8823	CPZ-----G-----A-----C-----P-----Z-----G-----A-----C-----P-----Z-----G-----A-----
CPZ_US_CFPZUS	8823	CPZ-----G-----A-----C-----P-----Z-----G-----A-----C-----P-----Z-----G-----A-----
Env gp41		I-----A-----V-----A-----E-----G-----T-----D-----R-----V-----I-----E-----V-----V-----O-----G-----A-----C-----R-----A-----I-----R-----H-----I-----P-----R-----R-----I-----R-----O-----G-----L-----E-----R-----I-----L-----L-----S-----
Env gp41, gp160 end /		Env gp41, gp160 end /

HIV-1/SIVcpz complete genomes

B.FR.HXB2	/ Nef start	ATGGTGGCAAGTGTGTCATTTGGATTGCGCTTCTGTAAGGAAAGAAATGAGCAGCT.....GAGCCAGCAGCAGATAGGG	8884
CONSENSUS_01_AE		ca-ag-g	A-ag-a-?????????	AG-A
01_AE.CF.90CF11697		AA-AG-G	A-G-A	8857
01_AE.CF.90CF402		T-CA-AG-G	A-G-A-A-CCTGTAGCA	AG-A
01_AE.CF.90CF4071		CA-AG-G	A-G-A-A-CCTGTAGCA	AG-A
01_AE.TH.93TH057		CA-AG-G	A-A-A-CCTGTAGCA	AG-A
01_AE.TH.93TH065		C-CA-AG-G	A-A-A-CCTGTAGCA	AG-A
01_AE.TH.93TH253		CA-AG-G	A-A-A-CCTGTAGCA	AG-A
01_AE.TH.95TH022		AAA	A-A-A-CCTGTAGCA	AG-A
01_AE.TH.95TH047		G-CT	A-A-A-CCTGTAGCA	AG-A
01_AE.TH.CM240		AA-T	A-A-A-CCTGTAGCA	AG-A
CONSENSUS_02_AG		c-ca-ag-g	a-a-c????????	acag-A
02_AG.CM.97CM-MP807		CC-AG-G	A-A-A	8058
02_AG.FR.DJ263		C-CA-AG-G	A-A-A	8085
02_AG.FR.DJ264		AG-T	A-A-A	8280
02_AG.GH.G829		C-CT-AG-G	A-A-A	8239
02_AG.NG.IBNG		C-CA-AG-G	A-A-A	8079
02_AG.SE.SE7812		C-CA-AG-G	A-A-A	8411
02_AG.SN.98SE-MP1211		C-CT-AG-G	A-A-A	8274
02_AG.SN.MP1213		C-CA-AG-G	A-A-A	8119
03_AB.RU.KAL153-2		CA-AG-G	A-A-A	8096
03_AB.RU.RU98001		CA-AG-G	A-A-A	8607
04_Cpx.CY.94CY032-3		CA-AG-G	A-A-A	8202
04_Cpx.GR.97PVGH		CT-AG-G	A-A-A	8293
04_Cpx.GR.97PVNY		AAAT-A	A-A-A	9007
05_DF.BE.VI1310		AG-AG-G	A-A-A	8948
05_DF.BE.VI1961		GA-AG-G	A-A-A	8290
CONSENSUS_06_cpx		ca-ag	a-a-c	8253
06_cpx.AU.BFP90		G	A	8729
06_cpx.ML.95ML127		CA-AG-G	A-A-A	8929
06_cpx.ML.95ML84		CC-AG-G	A-A-A	8886
06_cpx.SN.97SE1078		C-CA-AG-G	A-A-A	8097
07_BC.CN.97CN001		CA-AG-G	A-A-A	8446
07_BC.CN.98CN009		G	A	8227
07_BC.CN.C54A		CA-AG-G	A-A-A	8277
CONSENSUS_08_BC		CA-AG-G	A-A-A	8085
08_BC.CN.97CNGX6F		CA-AG-G	A-A-A	8065
08_BC.CN.97CNGX7F		G	A	8084
08_BC.CN.97CNGX9F		CA-AG-G	A-A-A	8078
08_BC.CN.98CN006		CA-AG-G	A-A-A	8066
10_CD.TZ.96TZBF061		CA-AG-G	A-A-A	8239
10_CD.TZ.96TZBF071		A	A	8235
10_CD.TZ.96TZBF110		T-CTG	A-A-A	8280
CONSENSUS_11_cpx		CA-AG-G	A-A-A	8249
11_cpx.CM.MP818		GAG-A-T	A-A-A	8643
11_cpx.FR.MP1298		G	A	8868
11_cpx.FR.MP1307		CA-AG-G	A-A-A	8904
11_cpx.GR.GR17		CA-AG-G	A-A-A	8926
12_BF.AR.ARMA159		C-A-AG	A-A-A	8163
12_BF.UY.URTR23		A-T	A-A-A	8849
12_BF.UY.URTR35		A-AG-A	A-A-A	8894
N.CM.YBF106		G-C-CC-AG-A	A-A-A	8907
N.CM.YBF30		G-C-CC-AG-A	A-A-A	8471
CONSENSUS_0		CA-AG-G	A-A-A	8571
O.CM.AN170		G-AG-A	A-A-A	8971
O.CM.MP5180		GAATGCA-AGC	A-A-A	8994
O.SN.SEMP1299		GA-TAT-T-AG	A-A-A	9033
O.SN.SEMP1300		GA-AT-T-AG	A-A-A	9003
CONSENSUS_CPZ		?-G-G-g	a-ga-???????	7998
CPZ_CD.CPZANT		TC-AGT-G	A-A-A	8338
CPZ_CM.CAM3		CC-AG-G	A-A-A	8321
CPZ_CM.CAM5		C-G-G	A-A-A	8627
CPZ_GA.CPZGAB		C-G-A	A-A-A	8957
CPZ_US.CPZUS		CA-AG-A	A-A-A	8917
Nef		M_G_K_W_S_K_S_S_V_I_G_W_P_T_V_R_E_R_M_R_R_A	E_P_A_A_D_R	Nef

HIV-1/SIVcpz complete genomes

B.FR.HXB2	ACCTCAGGTACCTTTAAGACCAATGACTTTACAAGCAGCTGTAGATCTTACCCACCTTTTAAAGAAAAGGGGGGACTGGAAGGGCTGAATTCACCTCCCAAGAGACCAAGATATC...CITGTCTCTGTGGA	9136
CONSENSUS_01_AE	G-G-g-----c-----c-----T-T-g-----T-T-T-----TT-----TT-----T-----T-----A-G-a-----G-----CT-A--G	9059
01_AE.CF.90CF11697	G-A-A-----G-----C-----T-T-g-----T-T-T-----TT-----TT-----T-----T-----A-G-a-----G-----CT-A--G	9103
01_AE.CF.90CF402	G-G-----G-----T-----T-----G-----T-----G-----T-----G-----T-----G-----T-----A-G-a-----G-----CT-A--G	9238
01_AE.CF.90CF4071	G-G-----G-----T-----T-----G-----T-----G-----T-----G-----T-----G-----T-----A-G-a-----G-----CT-A--G	9065
01_AE.TH.93TH057	G-G-----G-----T-----T-----G-----T-----G-----T-----G-----T-----G-----T-----A-G-a-----G-----CT-A--G	9088
01_AE.TH.93TH065	G-G-----G-----T-----T-----G-----T-----G-----T-----G-----T-----G-----T-----A-G-a-----G-----CT-A--G	9073
01_AE.TH.93TH253	G-G-----G-----T-----T-----G-----T-----G-----T-----G-----T-----G-----T-----A-G-a-----G-----CT-A--G	9138
01_AE.TH.95TH022	G-A-----G-----T-----T-----G-----T-----G-----T-----G-----T-----G-----T-----A-G-a-----G-----CT-A--G	9090
01_AE.TH.95TH047	G-G-----G-----T-----T-----G-----T-----G-----T-----G-----T-----G-----T-----A-G-a-----G-----CT-A--G	9102
01_AE.TH.CM240	G-G-----G-----T-----T-----G-----T-----G-----T-----G-----T-----G-----T-----A-G-a-----G-----CT-A--G	8704
CONSENSUS_02_AG	g-a-----g-----g-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----A-G-a-----g-----CT-A--G	8309
02_AG.CM.97CM-MP807	G-A-A-----G-----G-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----A-G-a-----G-----CT-A--G	8337
02_AG.FR.DJ263	G-A-----G-----G-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----A-G-a-----G-----CT-A--G	8532
02_AG.FR.DJ264	G-A-----G-----G-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----A-G-a-----G-----CT-A--G	8491
02_AG.GH.G829	G-G-----G-----G-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----A-G-a-----G-----CT-A--G	8331
02_AG.NG.IRNG	G-A-----G-----G-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----A-G-a-----G-----CT-A--G	8663
02_AG.SE.SE7812	G-G-----G-----G-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----A-G-a-----G-----CT-A--G	8526
02_AG.SN.96SE-MP1211	G-A-----G-----G-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----A-G-a-----G-----CT-A--G	8371
02_AG.SN.MP1213	G-A-----G-----G-----T-----T-----C-----C-----C-----C-----C-----C-----C-----C-----A-G-a-----G-----CT-A--G	8354
03_AB.RU.KAL153-2	G-G-----G-----G-----A-----T-----G-----T-----C-----C-----C-----C-----C-----C-----C-----C-----A-G-a-----A-----8858	8858
03_AB.RU.RU98001	G-G-----G-----G-----A-----T-----G-----T-----C-----C-----C-----C-----C-----C-----C-----C-----A-G-a-----A-----8454	8454
04_cpx.CV.94CV032-3	G-A-----G-----G-----C-----T-----A-----G-----T-----C-----C-----C-----C-----C-----C-----C-----C-----A-G-a-----G-----8548	8548
04_cpx.FR.97FVGH	G-A-----G-----G-----C-----T-----A-----G-----T-----C-----C-----C-----C-----C-----C-----C-----C-----A-G-a-----G-----9262	9262
04_cpx.FR.97FVNY	G-A-----G-----G-----C-----T-----A-----G-----T-----C-----C-----C-----C-----C-----C-----C-----C-----A-G-a-----G-----9203	9203
05_DF.BE.VI1310	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8545	8545
05_DF.BE.VI1961	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8505	8505
CONSENSUS_06_cpx	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8976	8976
06_cpx.AU.BFP90	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9181	9181
06_cpx.ML.95ML127	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9138	9138
06_cpx.ML.95ML84	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8349	8349
06_cpx.SN.97SE1078	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9204	9204
07_BC.CN.97CN001	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8479	8479
07_BC.CN.98CN009	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8479	8479
07_BC.CN.C54A	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8337	8337
CONSENSUS_08_BC	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8317	8317
08_BC.CN.97CNGX6F	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8336	8336
08_BC.CN.97CNGX7F	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8330	8330
08_BC.CN.97CNGX9F	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8318	8318
08_BC.CN.98CN006	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8491	8491
10_CD.TZ.96TZBF061	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8490	8490
10_CD.TZ.96TZBF071	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8535	8535
10_CD.TZ.96TZBF110	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8504	8504
CONSENSUS_11_cpx	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8885	8885
11_cpx.CM.MP818	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9120	9120
11_cpx.FR.MP1298	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9153	9153
11_cpx.FR.MP1307	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9178	9178
11_cpx.FR.GR17	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8415	8415
12_BF.AR.ARMA159	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9101	9101
12_BF.UY.URTR23	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9446	9446
12_BF.UY.URTR35	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9159	9159
N.CM.YBF106	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8726	8726
N.CM.YBF30	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8806	8806
CONSENSUS_0	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9223	9223
O.CM.ANT70	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9246	9246
O.CM.MVP5180	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9285	9285
O.SN.SEMP1299	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9255	9255
O.SN.SEMP1300	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8234	8234
CONSENSUS_CPZ	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8590	8590
CPZ.CD.CPZANT	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8579	8579
CPZ.CM.CAM3	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----8882	8882
CPZ.CM.CAM5	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9209	9209
CPZ.GA.CPZGAB	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----9175	9175
CPZ.US.CPZUS	-----T-----T-----A-----G-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----T-----A-G-a-----T-----Nef	Nef

HIV-1/SIVcpz complete genomes

B.FR.HXB2	TCTACCACACAAAGGCTACTTC...CCTGATFAGCAGAACTACACACACAGGCCAGGG...GTCAGATATCCACTGACCTTTGGATGGTGTACAAGCTAGTACCAGTTGAGCCAGATAGATA...GAAGAG	9261
CONSENSUS A-A1	-----G-----T-----C-----A--A-----t-----t-----C-----G--AG-----ga	8566
A.BY.97BL006	-----G-----T-----T-----C-----A--A-----R--G--R-----C-----C--G--AG-----	8656
A.SE.SE7535	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----	8517
A.SE.SE8538	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----	8472
A.SE.SE8891	-----G-----T-----T-----C-----A--A-----G-----T-----C-----A--AG--AG-----GA	8439
A.SE.UGSE8131	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8639
A1.KE.O23-17	-----G-----T-----T-----C-----A--A-----G-----T-----C-----C--G--AG-----GA	8713
A1.SE.SE7253	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8471
A1.UG.92UG037	-----G-----T-----T-----C-----A--A-----G-----T-----C-----C--G--AG-----GA	8670
A1.UG.U455	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8712
A2.CD.97CDKTB48	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8622
A2.CY.94CY017.41	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8687
CONSENSUS B	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8767
B.AU.MBC200	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	9310
B.CN.RL42	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8614
B.DE.D31	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8788
B.DE.HAN	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8717
B.ES.89SP061	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8839
B.GA.OYI	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8815
B.GB.MANC	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	9254
B.NL.3202A21	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	9270
B.TM.TWCYS	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	9237
B.US.JREL	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8590
B.US.MNCG	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8590
8612	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8590
B.US.SF2	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	9278
B.US.YU2	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	9249
CONSENSUS C	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8575
C.BR.92BR025	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8619
C.BR.98BR004	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8640
C.BW.96BW0502	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8765
C.BW.96BW15C02	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	9285
C.ET.ETH2220	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8742
C.IL.98ILS002	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8661
C.IN.93IN101	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8613
C.IN.94IN476	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	9239
C.IN.98IN012	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8598
C.TZ.98TZ013	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8621
C.TZ.98TZ017	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8660
C.ZA.97ZA012	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8621
C.ZM.96ZM651	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8582
C.ZM.96ZM751	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8657
D.CD.84ZR085	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8809
D.CD.FLI	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8800
D.UG.94UG1141	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8579
CONSENSUS F1	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8371
F1.BE.VI850	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8538
F1.BR.93BR020.1	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8574
F1.FI.FIN9363	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8562
F1.FR.MP411	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8488
F2.CM.MP255	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8405
F2.CM.MP257	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8438
CONSENSUS G	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8512
G.BE.DRCHL	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	9264
G.FI.HH8793-12.1	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8661
G.NG.92NG083	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8611
G.SE.SEG165	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8703
H.BE.VI991	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8664
H.BE.VI997	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8579
H.CF.90CF056	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8584
J.SE.SE7022	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8569
J.SE.SE7887	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8559
K.CD.E0TB11C	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8444
K.CM.MP535	-----G-----T-----T-----C-----A--A-----A-----T-----C-----C--G--AG-----GA	8448

Accession	Strain	Genomic Coordinates
B.FR.HXB2	Nef premature end in HXB2	9226-9300
CONSENSUS_01_AE	01_AE.CF.90CF11697	9184-9226
01_AE.CF.90CF402	01_AE.CF.90CF402	9184-9226
01_AE.CF.90CF4071	01_AE.CF.90CF4071	9184-9226
01_AE.TH.93TH057	01_AE.TH.93TH057	9184-9226
01_AE.TH.93TH065	01_AE.TH.93TH065	9184-9226
01_AE.TH.93TH253	01_AE.TH.93TH253	9184-9226
01_AE.TH.95TH022	01_AE.TH.95TH022	9184-9226
01_AE.TH.95TH047	01_AE.TH.95TH047	9184-9226
01_AE.TH.CM240	01_AE.TH.CM240	9184-9226
CONSENSUS_02_AG	02_AG.CM.97CM-MP807	8433-8829
02_AG.FR.DJ263	02_AG.FR.DJ263	8433-8829
02_AG.FR.DJ264	02_AG.FR.DJ264	8433-8829
02_AG.GH.G829	02_AG.GH.G829	8433-8829
02_AG.NG.IBNG	02_AG.NG.IBNG	8433-8829
02_AG.SE.SE7812	02_AG.SE.SE7812	8433-8829
02_AG.SN.96SE-MP1211	02_AG.SN.96SE-MP1211	8433-8829
02_AG.SN.MP1213	02_AG.SN.MP1213	8433-8829
03_AB.RU.KAL153-2	03_AB.RU.KAL153-2	8968-8829
03_AB.RU.RU98001	03_AB.RU.RU98001	8968-8829
04_CPX.CV.94CV032-3	04_CPX.CV.94CV032-3	8670-8630
04_CPX.FR.97FVGH	04_CPX.FR.97FVGH	8670-8630
04_CPX.FR.97FVNY	04_CPX.FR.97FVNY	8670-8630
05_DF.BE.VI1310	05_DF.BE.VI1310	9099-9233
05_DF.BE.VI1961	05_DF.BE.VI1961	9099-9233
CONSENSUS_06_CPX	06_CPX.AU.BFP90	9233-8474
06_CPX.AU.BFP90	06_CPX.AU.BFP90	9233-8474
06_CPX.ML.95ML1127	06_CPX.ML.95ML1127	9233-8474
06_CPX.ML.95ML84	06_CPX.ML.95ML84	9233-8474
06_CPX.SN.97SE1078	06_CPX.SN.97SE1078	9233-8474
07_BC.CN.97CN001	07_BC.CN.97CN001	8604-8604
07_BC.CN.98CN009	07_BC.CN.98CN009	8604-8604
07_BC.CN.C54A	07_BC.CN.C54A	8604-8604
CONSENSUS_08_BC	08_BC.CN.97CNGX6F	8442-8455
08_BC.CN.97CNGX6F	08_BC.CN.97CNGX6F	8442-8455
08_BC.CN.97CNGX7F	08_BC.CN.97CNGX7F	8442-8455
08_BC.CN.97CNGX9F	08_BC.CN.97CNGX9F	8442-8455
08_BC.CN.98CN006	08_BC.CN.98CN006	8616-8616
10_CD.TZ.96TZBF061	10_CD.TZ.96TZBF061	8615-8660
10_CD.TZ.96TZBF071	10_CD.TZ.96TZBF071	8615-8660
10_CD.TZ.96TZBF110	10_CD.TZ.96TZBF110	8615-8660
CONSENSUS_11_CPX	11_CPX.CM.MP818	9006-9245
11_CPX.CM.MP818	11_CPX.CM.MP818	9006-9245
11_CPX.FR.MP1298	11_CPX.FR.MP1298	9006-9245
11_CPX.FR.MP1307	11_CPX.FR.MP1307	9006-9245
11_CPX.FR.MP1307	11_CPX.FR.MP1307	9006-9245
11_CPX.FR.GR17	11_CPX.FR.GR17	9006-9245
12_BF.AR.ARMA159	12_BF.AR.ARMA159	9256-9271
12_BF.UY.URTR23	12_BF.UY.URTR23	9256-9271
12_BF.UY.URTR35	12_BF.UY.URTR35	9256-9271
N.CM.YBF106	N.CM.YBF106	8794-8851
N.CM.YBF30	N.CM.YBF30	8794-8851
CONSENSUS_0	O.CM.AN170	8922-9333
O.CM.AN170	O.CM.AN170	8922-9333
O.CM.MVP5180	O.CM.MVP5180	8922-9333
O.SN.SEMP1299	O.SN.SEMP1299	8922-9333
O.SN.SEMP1300	O.SN.SEMP1300	8922-9333
CONSENSUS_CPZ	CPZ.CD.CPZANT	8706-9007
CPZ.CD.CPZANT	CPZ.CD.CPZANT	8706-9007
CPZ.CM.CAM3	CPZ.CM.CAM3	9007-9334
CPZ.CM.CAM5	CPZ.CM.CAM5	9007-9334
CPZ.GA.CPZGAB	CPZ.GA.CPZGAB	9330-9330
CPZ.US.CPZUS	CPZ.US.CPZUS	9330-9330
Nef	Nef	9300-9300

	TCF-1 alpha binding	Nef end \	NF-kappa-B II	
B.FR.HXB2	ATCCGGAGTACTTC	????????AAGAAGCTGCTGACATCGAGCTTGCTACAAGGGACTTTCCGCTG
CONSENSUS_A-A1	---t-a	????????	---CA-ag	---GAC-?
A.BY.97BL006	---T-A	---CA	---GAC-T
A.SE.SE7535	---T-A	---CA	---GAC-T
A.SE.SE8538	---T-A	---CA	---GAC-T
A.SE.SE8891	---T-A	---CA	---GAC-T
A.SE.UGSE8131	---GG-A	---A	---T-CA-CAG
A1.KE.O23-17	---T-A	---G	---CA-AG
A1.SE.SE7253	---TT-A	---AG	---CA-AG
A1.UG.U455	---T-AT	---AAAGACTG	---AG
A2.CD.97CDKTB48	---T-A	---AG	---CA-AGA
A2.CY.94CY017.41	---T-A	---A2	---CA-AG
CONSENSUS_B	---A	????????	---A	---T
B.AU.MBC200	---A	---A	---T
B.CN.RL42	---CA	---T	---T
B.DE.D31	---A	---AAAACTG	---AG
B.DE.HAN	---A	---AG	---CT
B.ES.89SP061	---A	---G	---CT
B.GA.OYI	---A	---AG	---T
B.GB.MANC	---A	---AG	---T
B.NL.3202A21	---A	---G	---T
B.TW.TWCYS	---A	---Y	---T
B.US.JRFL	---A	---G	---T
B.US.MNCG	---A	---G	---T
B.US.RF	---A	---AG	---T
B.US.SF2	---A	---AG	---T
B.US.YU2	---A	---A	---C
CONSENSUS_C	---t-a	????????	---AG	---CA-a
C.BR.92BR025	---T-A	---AG	---T
C.BR.98BR004	---T-A	---AG	---T
C.BW.96BW0502	---T-A	---AG	---T
C.BW.96BW06	---T-A	---AG	---T
C.BW.96BW15C02	---A	---A	---T
C.ET.FTH2220	---T-A	---AG	---T
C.LL.98LS002	---T-A	---AG	---T
C.IN.93IN101	---TT-A	---AG	---T
C.IN.94IN476	---T-A	---AG	---T
C.IN.98IN012	---C	---T-A	---A
C.TZ.98TZ013	---T-A	---AG	---T
C.TZ.98TZ017	---T-A	---AG	---T
C.ZA.97ZA012	---A	---AG	---T
C.ZM.96ZM651	---T-A	---AG	---T
C.ZM.96ZM751	---T-A	---AG	---T
D.CD.84ZR085	---T-A	---D	---C
D.CD.ELI	---T-A	---A	---C
D.UG.94UG1141	---T-A	---AG	---C
CONSENSUS_F1	---t-a	---C-AG
F1.BE.VI850	---T-A	---C-AG
F1.BR.93BR020.1	---A	---C-AG
F1.FI.FIN9363	---T-A	---CGAG
F1.FR.MF411	---T-AT
F2.CM.MP255	---TT-A
F2.CM.MP257	---A
CONSENSUS_G	---A	---ag
G.BE.DRCBL	---A	---AG
G.FI.HH8793-12.1	---T-AT
G.NG.92NG083	---A	---AG
G.SE.SE6165	---A	---AG
H.BE.VI991	---T-A	---AG
H.BE.VI997	---T-A	---AG
H.CF.90CF056	---A	---AG
J.SE.SE7022	---TT-A	---AG
J.SE.SE7887	---TT-A	---AG
K.CD.E0TB11C	---A	---AG
K.CM.MP535	---C	---A
9448
8752
8854
8704
8659
8627
8826
8900
8658
8857
8907
8823
8874
8951
9497
8801
8983
8892
9026
9002
9441
9457
9424
8767
9467
8999
9465
9436
8751
8819
8817
8944
9431
8919
8840
8789
9415
8797
8836
8795
8759
8813
8773
8975
8987
8766
8575
8746
8781
8768
8578
8555
8588
8699
9415
8850
8799
8892
8866
8765
8766
8756
8746
8594
8598

HIV-1/SIVcpz Complete Genomes

	TCF-1 alpha binding	Nef end \	NF-kappa-B II	
B.FR.HXB2	ATCCGGAGTACTTC.....AAGAAGCTGCTGACATCGAGCTTGCTACAA..GGGACTTTCCCGCTG..	9448
CONSENSUS_01_AE	---A---AT	---AG---aa---AG---L---act?????????	???????????????	9369
01_AE.CF.90CF11697	G---A---A	---GA---AG---T---ACCA	9414
01_AE.CF.90CF402	---C---A	---AA---AG---C---A---G---CTGCTGACAAAGGAGTTTCTAACCTA	9572
01_AE.CF.90CF4071	---A---A	---AA---AG---TA---ACTA	9376
01_AE.TH.93TH057	---A---AT	---AA---AG---T---ACTA	9399
01_AE.TH.93TH065	---A---T---AT	---AA---AG---T---ACTA	9384
01_AE.TH.93TH065	G---A---T---AT	---AA---AG---T---ACTTCTAACCTA	9449
01_AE.TH.95TH022	---C---A	---AA---AG---T---A---GA	9401
01_AE.TH.95TH047	---C---A	---AA---AG---T---A---GA	9413
01_AE.TH.CM240	---A	---AG---AG---AGTCTA.CTA	9014
CONSENSUS_02_AG	---T---A	---CA---AG---GAC-?	---?---??	8618
02_AG.CM.97CM.MP807	---T---A	---CA---GAG---GACCA	8649
02_AG.FR.DJ263	---T---AT	---CA---AG---GAC-AA	8845
02_AG.FR.DJ264	---T---AT	---CA---AG---GAC-AA	8804
02_AG.GH.GB29	---T---A	---CA---AA---GACA	8643
02_AG.NG.IBNG	---T---A	---CA---AG---GACA	8975
02_AG.SE.SE7812	---TT---A	---CA---AG---GACT-ACA	8859
02_AG.SN.96SE.MP1211	---TT---A	---CA---AG---GAC-AA	8684
02_AG.SN.MP1213	---TT---A	---CA---AG---GAC-AAC	8674
03_AB.RU.KAL153-2	---T---A	---CA---C---GAC-TGAAGTTGCTGACAG	9167
03_AB.RU.RU98001	---T---A	---CA---C---GAC-TGAAGTTGCTGACAG	8778
04_CPX.CV.94CY032-3	---TT---A	---CA---AG---GAC-AACC	8862
04_CPX.GR.97PVCH	---T---A	---CA---AGG	9543
04_CPX.GR.97PVNY	---T---A	---CA---AGCC	9517
05_DF.BE.VI1310	---A---T---A	---CAAGCTGAG	8894
05_DF.BE.VI1961	GC---A---T---A	---CAAGCTGAGGAC-CAGGAAGTGTGACAA	8836
CONSENSUS_06_CPX	---TT---A	---a?-?gc-?a?-a-Gaagtttct? ????	???????????????	9286
06_CPX.AU.BEP90	---C---TT---A	---AG---CTGC---AC---A---GAAATTCTA.ATG	9505
06_CPX.ML.95ML127	---TT---A	---AA---AG---T---ACG	9450
06_CPX.ML.95ML84	---TT---A	---AA---AG---T---ACG	8661
06_CPX.SN.97SE1078	---TT---AY	---AG---G-CTCACAAAG-TGCTGACAAAGAAATTTCTGACA	9537
07_BC.CN.97CN001	---TT---A	---CA---AG	8780
07_BC.CN.98CN009	---TT---A	---CA---AG	8780
07_BC.CN.C54A	---TT---A	---AGG	8638
CONSENSUS_08_BC	---TT---A	---AG?	8617
08_BC.CN.97CNGX6F	---TT---A	---AG?	8636
08_BC.CN.97CNGX7F	---TT---A	---AG?	8630
08_BC.CN.97CNGX9F	---TT---A	---AGG	8619
08_BC.CN.98CN006	---TT---A	---AGGGGACTTTACTGCTGACAGAA	8815
10_CD.TZ.96TZBF061	---T---A	---AG---T	8791
10_CD.TZ.96TZBF071	---T---A	---A	8836
10_CD.TZ.96TZBF110	---T---A	---AGT	8805
CONSENSUS_11_CPX	---c-t---A?aaagac?-ag?????????????????	9192
11_CPX.CM.MP818	---C---T---A	---AAAGACT-AG	9438
11_CPX.FR.MP1298	---C---T---A	---AG	9465
11_CPX.FR.MP1307	---GG---AT	---AAGGACT	9497
11_CPX.GR.GR17	---C---T---A	---AAAGACT-AGGACATAAAAAGTTGCTGACA	8748
12_BF.AR.ARMAL59	---A---GG---ATAAAGACTGAGGACACAGAAAGAACTTAA	9433
12_BF.UY.URTR23	---T---A	---CAAGACTGAGGACACAGAAAGAACTTAA	9499
12_BF.UY.URTR35	---T---A	---AAAGACTGAGGACACAGAAAGAACTTAA	9491
N.CM.YBF106	---C---T---AAA-A-G	8966
N.CM.YBF30	---TT---AAA-TGCTGA	9026
CONSENSUS_0	---C---A---CTcacaagActTAAaactgttacc?????????AGcaga	9104
O.CM.ANT70	---C---A---CTAGATTGCTGACACTGTAGCAAA	9554
O.CM.MVP5180	---C---A---CTCCTCAG	9549
O.SN.SEMP1259	---C---A---CTAACAAGGACTTAAACTGCTGACCTG-AG-TAGCAGA	9615
O.SN.SEMP1300	---C---A---CTCTCARGACTTAAACTGTTGACATG-AG-TAGCAGA	9585
CONSENSUS_CFX	---t---u?????????????????g?ac?tAa?	8520
CFZ.CD.CPZANT	---C---T---AAGACT-AGA	8916
CFZ.CM.CAM3	---T---AGGG-A-ATAA-ACCGCTGACAGACACTTGTGGTCTA	8903
CFZ.CM.CAM5	---T---ATGGAGACCAACCACCGCTGCGAGTTAA	9205
CFZ.GA.CPZGAB	---AAAAGACTGACTTTAGACTGCCCATCCGACTAA	9539
CFZ.US.CPZUS	---AAAAGACTGACTTTAGACTGCCCATCCGACTAA	9503
B.FR.HXB2	ATCCGGAGTACTTC.....AAGAAGCTGCTGACATCGAGCTTGCTACAA..GGGACTTTCCCGCTG..	9448
Nef	H_P_E_Y_F_____K_N_C_S	_____Nef end /		

HIV-1/SIVcpz complete genomes

HIV-1/SIVcpz
complete genomes

	NF-kappa-B I	SPI-III	SPI-II	SPI-I	TATA box	
B.FR.HXB2	GGGACTTTC	AGGGAGCGTGGCTGGGGAGTGGG	AGT	AGT	ATCCCTGCATATAAGCAGCTGCTTTTGGCTG	9534
CONSENSUS_A-A1	R	TT	AGT	AGT	t-a	8839
A.BY.97BL006	AGRR	TT	RAGT	RRR	T-A	8941
A.SE.SE7535	AG	TT	AGT	T	T-A	8791
A.SE.SE8538	AG	TT	AGT	T	T-A	8746
A.SE.SE8891	AG	TT	AGT	T	T-A	8714
A.SE.UGSE8131	AG	TT	AGT	T	T-A	8913
AL.KE.O23-17	AG	TT	AGT	T	T-A	8987
AL.SE.SE7253	AG	TT	AGT	T	T-A	8745
AL.UG.92UG037	AG	TT	AGT	T	T-A	8944
AL.UG.U455	G	TT	AGT	T	T-A	8994
A2.CD.97CDKTB48	G	TGA	AGT	AGT	ATG-TGCATATA-GCAGCTGCT	8909
A2.CY.94CY017.41	T	TG	AGT	AGT	T-A	8960
CONSENSUS_B	?	?	???????????	?	g	9037
B.AU.MBC200	AG	A	C	GT	G	9583
B.CN.RL42	A	A	C	GT	G	8887
B.DE.D31	A	A	C	GT	G	9069
B.DE.HAN	A	A	C	GT	G	8989
B.ES.89SP061	A	A	C	GT	G	9112
B.GA.OYI	AG	A	C	GT	G	9088
B.GB.MANC	AG	A	C	GT	G	9528
B.NL.3202A21	AG	A	C	GT	G	9544
B.ITW.TWCYS	G	C	R	G	A	9509
B.US.JRFL	T	C	C	C	G	8853
B.US.MNCG	T	C	C	C	G	9553
B.US.RF	T	C	C	C	G	9085
B.US.SF2	T	C	C	C	G	9550
B.US.YU2	A	A	C	GT	G	9521
CONSENSUS_C	actgggg.cgttcca	t	T	?	c-a	8848
C.BR.92BR025	G	A	C	C	A	8903
C.BR.98BR004	ACTGGG.CGTTCCA	T	T	C	A	8915
C.BW.96BW0502	ATAAGAGGACTTCCGCT	T	T	C	A	9056
C.BW.96BW06	ACTGGG.CGTTCCA	T	T	C	A	9528
C.BW.96BW15C02	GGACTTCCACTA	T	T	C	A	9010
C.ET.ETH2220	G	A	C	GT	G	8934
C.IL.98IS002	ACTGGG.CGTTCCA	T	T	C	A	8885
C.IN.93IN101	ACTGGG.CGTTCCA	T	T	C	A	9512
C.IN.94IN476	GGG.CGTTCCA	T	T	C	A	8858
C.IN.98IN012	ACTGGG.CGTTCCA	T	T	C	A	8894
C.TZ.98TZ013	ACCAGG.CGTGAC	T	T	C	A	8933
C.TZ.98TZ017	ACTGGG.CGTTCCA	T	T	C	A	8911
C.ZA.97ZA012	ACTGGG.CGTTCCA	T	T	C	A	8857
C.ZM.96ZM651	ACTG.GGTTCCA	T	T	C	A	8909
C.ZM.96ZM751	ACTGGG.CGTTCCA	T	T	C	A	8871
D.CD.ELI	AA	AG	AA	AG	T	9073
D.UG.94UG1141	AA	AG	AA	AG	T	8851
F1.BE.VI850	A	G	CG	CCAGA	TCA	8832
F1.BR.93BR020.1	A	G	TG	CCAGA	TCA	8867
F1.FI.FIN9363	A	G	CG	CCAGA	TCA	8854
CONSENSUS_G	????????????????	C	A	g	AG	8785
G.BE.DRCBL	GACAAAGGACTTCCGCC	C	A	g	AG	9521
G.FI.HH8793-12.1	AG	C	A	A	T	8936
G.NG.92NG083	G	A	A	A	T	8886
G.SE.SE6165	G	A	A	A	T	8978
H.BE.VI991	T	T	T	A	T	8952
H.BE.VI997	G	C	T	GA	T	8851
H.CF.90CF056	GA	A	T	GA	T	8852
J.SE.SE7022	AG	T	GT	T	A	8843
J.SE.SE7887	AG	A	GT	T	A	8833

3' LTR U3 end / 3' LTR repeat start / 3' LTR Poly-A signal R repeat end / 5' LTR U5 start / Secondary structure region

Sequence alignment for HIV-1/SIVcpz complete genomes. Includes labels for structural elements like TAR, stem, bulge, loop, and secondary structure regions. Lists accession numbers and strain identifiers on the left and right margins.

3' LTR Poly-A signal R repeat end \ 5' LTR U5 start

3' LTR Poly-A signal R repeat end \ 5' LTR U5 start

3' LTR U3 end \ 3' LTR repeat start

B.FR.HXB2 TACTGGTCTCTCTGGTTAGACAGATCTGACGCTGGGAGCTCTCTGGC.TAACTAGGGAACCCACTGCTTAAAGCCTTAAGCTTGAAGTCTTCAAGTAGTGTGCCCTGCTGTTGGT.G 9662

Secondary structure region stem bulge loop stem

Consensus and sequence alignment for HIV-1/SIVcpz complete genomes. Includes labels like CONSENSUS_01_AE, 01_AE.CF.90CF11697, etc., and a large block of nucleotide sequence with secondary structure annotations.



Extensive secondary structure in this region
See Rizvi, J Virol 67:2681-8(1993) 3' LTR U5 end \

```

ACTCTGGTAACTAGAGATC.CCTCAGACCCCTTTTGTAGTGTGAGAAA.TCTCTAGCA
-----AC-C--A-G---AA-----
-----AC-A--A-T---AA-----
-----cCcTtt-tc-g-----
-----A-----
-----T-----G-----
-----B.US.SF2-----A-----
-----B.US.YU2-----
-----C.BW.96BW06-----G-----A-----
-----C.IN.93IN101-----
-----G.BE.DRCBL-----AC-C--AT-----AA-----
-----01_AE.CF.90CF402-----
-----01_AE.CF.90CF4071-----AC-C--A-T-A--AA-----
-----01_AE.TH.93TH057-----A-T-----
-----01_AE.TH.93TH065-----
-----01_AE.TH.93TH253-----T-AC-C--A-T-A--AA-----
GACTCT-GT-AC
GA.-CTGGTAACT--GA.T-C-TC-GA-CAC-CTAG-CA--T--A-----
GA.-CTGGTAACT--GA.T-C-TC-GA-CACGCTAGCG--T-----
GA.-CTGGTAACT--GA.T-C-TC-GA-CAC-CTAG-AG--T--A-----
GA.-TTGG-ACCT--AAT--CT-A-A--TAG-CCG-G--G-A-T-T-----
GA.-CTGGTAACT--GAT--CTCAGA--TAGTCCGTGT-G-A-----
GA.-CTGGTAACT--GAT--CTCAGA--C-CTAGTCCGTGT-G-A-----
GA.-CTGGTAACT--GAT--CTCAGA--A-CTAGTCCGTGT-G-A-----
GA.-CTGGTAACT--GAT--CTCAGA--TAGTCCGTGT--A-----
GA.-CTGGTAACT--GAT--CTCAGA--A-TAGTCCGTGT-G-A-----
...-CCTGGTG-CT--GA.T-
...-CCTGGTA-CT--GA.T-
...-CCTGGTA-CT--GA.T-
-A.-CTGGTAACT--GA.T-C-TC-GAAC-A-TA-A--AG-TG--A-----
CA.-CTGGTAACT--GA.T-C-TC-GAT-AAATTA--CAA-G-A-----
CPZ.US.CPZUS  GA.-CTGGTAACT--GA.T-C-TC-GAAC-C-TAAA--AG-TG--A-----

```

B.FR.HXB2
A1.KE.Q23-17
A1.UG.U455
CONSENSUS B
B.AU.MBC200
B.IW.TWCYS
B.US.MNCG
B.US.SF2
B.US.YU2
C.BW.96BW06
C.IN.93IN101
G.BE.DRCBL
01_AE.CF.90CF402
01_AE.CF.90CF4071
01_AE.TH.93TH057
01_AE.TH.93TH065
01_AE.TH.93TH253
02_AG.NG.IBNG
06_CPX.AU.BFP90
06_CPX.ML.95ML127
06_CPX.SN.97SE1078
11_CPX.CM.MP818
11_CPX.FR.MP1298
11_CPX.FR.MP1307
12_BF.AR.ARMAL159
12_BF.UY.URTR23
12_BF.UY.URTR35
O.CM.MVP5180
O.SN.SEMP1299
O.SN.SEMP1300
CPZ.CM.CAM3
CPZ.GA.CPZGAB
CPZ.US.CPZUS

9719
9169
9178
9563
9767
9694
9738
9736
9706
9713
9680
9705
9843
9597
9645
9612
9720
9201
9775
9719
9808
9710
9736
9768
9704
9770
9762
9793
9859
9829
9170
9811
9781