

# III

## HIV-2/SIV Complete Genomes

### Contents

---

III-1	Introduction . . . . .	151
III-2	Annotated Features . . . . .	153
III-3	Sequences . . . . .	155
III-4	Alignments . . . . .	157

---

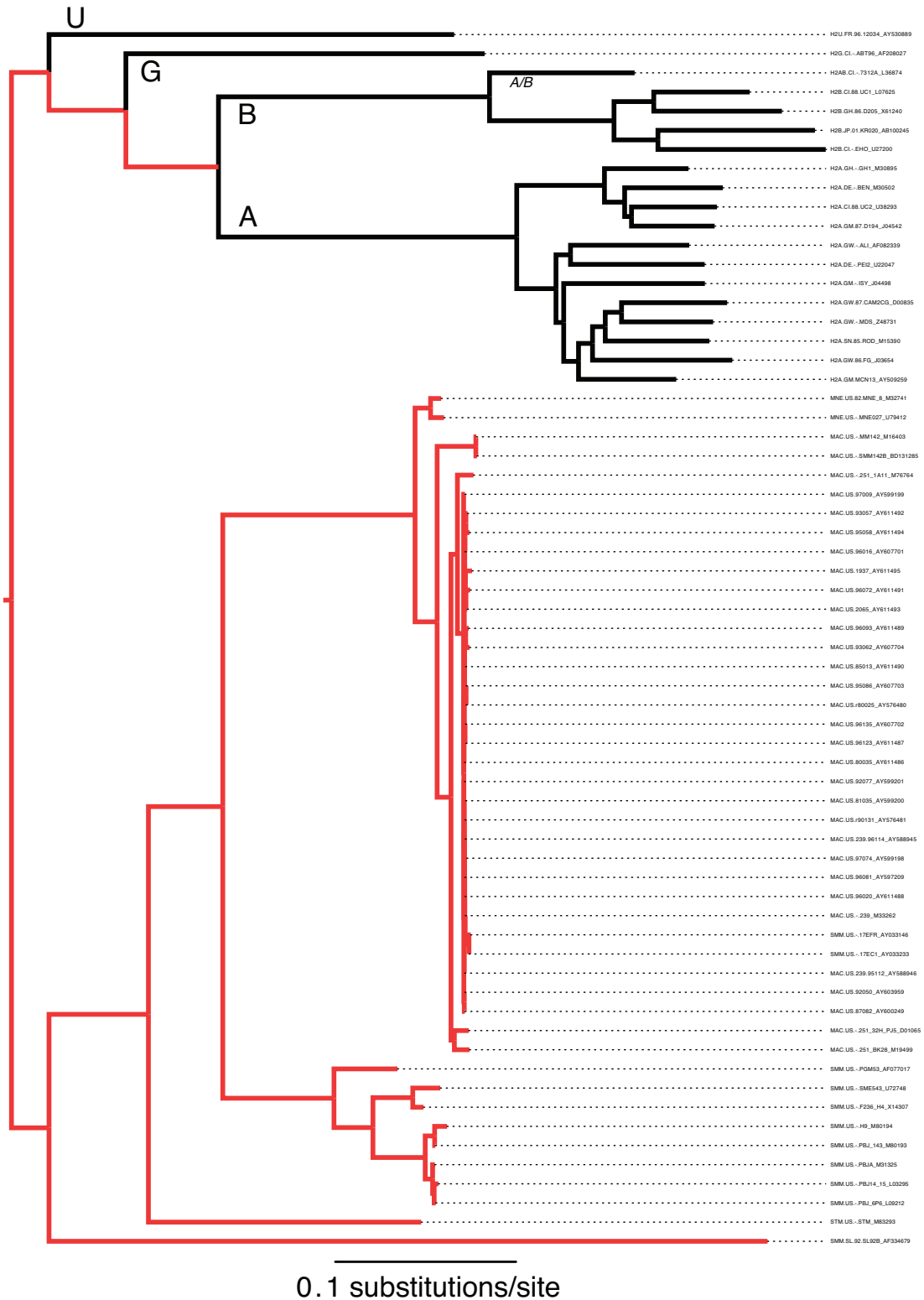
### III-1 Introduction

The HIV-2/SIVsmm alignment contains a non-redundant set of all HIV-2, SIVsmm, and related species (SIVmne, SIVstm and SIVmac). Because relatively to HIV-1 few HIV-2 complete genomes have been determined, this alignment contains all but a few duplicate sequences available in the database. This 2007 compendium alignment has not changed since the 2005 compendium. For a full HIV-2/SIVsmm alignment, the 2007 web-alignment has some additional sequences.

As with HIV-1/SIVcpz, this alignment was generated by an iterative process between automated alignment using HMMER and manual editing using MASE, BioEdit and Se-Al. As in previous years, the alignment presented 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. Figure III-1 shows a maximum likelihood tree based on the alignment with known recombinants removed.

The MAC239 sequence was chosen as master sequence in this set.

HIV-2/SIV Complete Genomes



**Figure III-1:** Rooted maximum likelihood tree of HIV-2 and SIVmm (and SIVmac and SIVstm) lineages in the compendium alignment. SIVmm lineages are in red and HIV-2 in black. Believed cross species transmissions to humans are indicated by letters. Note that one HIV-2 A/B recombinant is included in the tree. The root is placed at the earliest possible point before HIV-2 lineages occur to reflect the origin in SIVmm. The tree was calculated under a general-time-reversible model with site rate variation (). Cosmetic changes were done using FigTree and Adobe Illustrator.

### III-2 Annotated Features

Features of HIV-2 annotated in the alignment that follows.

<b>Feature</b>	<b>Location</b>	<b>Page</b>
5' LTR U3 start	1	157
TCF-1 alpha	309-324	158
NF- $\kappa$ -B-II	415-424	159
NF- $\kappa$ -B-I	441-449	160
TATA Box	488-493	160
TAR element start	518	160
5' LTR U3 end	520	160
+1 mRNA start site	521	160
5' LTR R repeat begin	521	160
TAR element end	642	161
Poly-A signal	671-676	161
5' LTR R repeat end	694	161
5' LTR U5 start	695	161
5' LTR U5 end	818	162
Lys tRNA primer binding site	821-839	162
Gag and Gag-Pol start	1053	164
Gag p17 Matrix end	1457	166
Gag p24 Capsid start	1458	166
Gag p24 Capsid end	2144	170
Gag p2 Spacer start	2145	170
Gag p2 Spacer end	2195	170
Gag p8 Nucleocapsid start	2196	170
Gag p8 Nucleocapsid end	2351	171
Gag p1 Spacer start	2352	171
Gag-Pol -1 ribosomal slip site	2354	171
Pol start	2354	171
Gag p1 spacer end	2393	172
Gag p6 start	2394	172
Pol Protease start	2555	173
Gag p6 end	2585	173
Gag end	2585	173
Pol Protease end	2851	175
Pol p66 and p51 RT start	2852	175
Pol RT end	4168	182
Pol p15 RNase H start	4169	182
Pol RNase H end	4528	185
Pol p31 Integrase start	4529	185
Vif start	5340	189
Pol, Gag-Pol, and p31 Integrase end	5410	190
Vpx start	5812	192
Vif end	5984	193
Vpx end	6150	194
Vpr start	6151	194
Tat exon 1 start	6302	195
Vpr end	6456	196
Rev exon 1 start	6528	196
Tat Rev exon 1 end	6597	197
Tat Rev intron start	6598	197

Feature	Location	Page
Env signal peptide start	6604	197
Env signal peptide end	6669	197
Env gp120 start	6670	197
V1 loop start	6733	198
V1 loop end	6774	198
V2 loop start	6904	199
V2 loop end	7074	200
V3 loop start	7183	200
V3 loop end	7392	202
V4 loop start	7534	203
V4 loop end	7635	203
V5 loop start	7786	204
V5 loop end	7980	205
Rev Responsive Element (RRE) start	8123	206
Env gp120 end	8178	207
Env gp41 start	8179	207
RRE end	8479	208
Tat Rev intron end	8805	210
Tat Rev exon 2 start	8806	210
Tat end	8902	211
Rev end	9059	212
Nef start	9077	212
Env gp41, gp160 end	9243	213
Premature stop in SMM239	9355	214
3' LTR U3 start	9462	214
TCF-1 alpha binding	9770-9785	216
Nef end	9868	217
NF- $\kappa$ -B-II	9876-9884	217
NF- $\kappa$ -B-I	9891-9900	217
TATA box	9949-9955	218
TAR element start	9980	218
3' LTR U3 end	9978	218
3' LTR repeat start	9979	218
TAR element end	10103	219
Poly-A signal	10132-10137	219
3' LTR R repeat end	10155	219
3' LTR U5 start	10156	219
3' LTR U5 end	10279	220

## III-3 Sequences

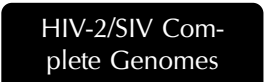
Sequences included in the HIV-2/SIV complete genome alignment.

Name	Accession	Country	Author	Reference
MAC.US.-.239	M33262	United States	Kestler, H	<i>Science</i> <b>248</b> (4959); 1109-12 (1990)
H2A.GM.MCN13	AY509259	Gambia	Schmitz, C	<i>J Virol</i> <b>78</b> (4):2006-2016 (2004)
H2A.CI.88.UC2	U38293	Cote D'Ivoire	Barnett, SW	<i>Virology</i> <b>222</b> (1); 257-61 (1996)
H2A.DE.-.BEN	M30502	Germany	Kirchhoff, F	<i>Virology</i> <b>177</b> (1); 305-11 (1990)
H2A.DE.-.PEI2	U22047	Germany	Talbot, R	<i>PNAS USA</i> <b>90</b> (9); 4226-30 (1993)
H2A.GH.-.GH1	M30895	Ghana	Hasegawa, A	<i>ARHR</i> <b>5</b> (6); 593-604 (1989)
H2A.GM.-.ISY	J04498	Gambia	Franchini, G	<i>PNAS USA</i> <b>86</b> (7); 2433-7 (1989)
H2A.GM.87.D194	J04542	Gambia	Kuehnel, H	<i>PNAS USA</i> <b>86</b> (7); 2383-7 (1989)
H2A.GW.-.ALI	AF082339	Guinea- Bissau	Azevedo-Pereira, JM	Unpublished (1998)
H2A.GW.-.MDS	Z48731	Guinea- Bissau	Becker, M	Unpublished (1995)
H2A.GW.86.FG	J03654	Guinea- Bissau	Zagury, JF	<i>PNAS USA</i> <b>85</b> (16); 5941-5 (1988)
H2A.GW.87.CAM2CG	D00835	Guinea- Bissau	Tristem, M	<i>J Gen Virol</i> <b>1991</b> Mar;72( Pt 3):721-4
H2A.SN.85.ROD	M15390	Senegal	Clavel, F	<i>Nature</i> <b>324</b> (6098); 691-5 (1986)
H2AB.CI.-.7312A	L36874	Cote D'Ivoire	Gao, F	Unpublished
H2B.CI.-.EHO	U27200	Cote D'Ivoire	Rey-Cuille, MA	<i>Virology</i> <b>202</b> (1); 471-6 (1994)
H2B.CI.88.UC1	L07625	Cote D'Ivoire	Barnett, SW	<i>J Virol</i> <b>67</b> (2); 1006-14 (1993)
H2B.GH.86.D205	X61240	Ghana	Kreutz, R	<i>ARHR</i> <b>8</b> (9); 1619-29 (1992)
H2B.JP.01.KR020	AB100245	Japan	Oka, S-I	<i>ARHR</i> <b>19</b> (11); 1045-9 (2003)
H2G.CI.-.ABT96	AF208027	Cote D'Ivoire	Brennan, CA	<i>ARHR</i> <b>13</b> (5); 401-4 (1997)
H2U.FR.96.12034	AY530889	France	Damond, F	<i>ARHR</i> <b>20</b> (6); 666-72 (2004)
MAC.US.-.251_1A11	M76764	United States	Planelles, V	<i>ARHR</i> <b>7</b> (11); 889-98 (1991)
MAC.US.-.251_32H_PJ5	D01065	United States	Rud, EW	<i>J Gen Virol</i> <b>1994</b> Mar;75( Pt 3):529-43
MAC.US.-.251_BK28	M19499	United States	Franchini, G	<i>Nature</i> <b>328</b> (6130); 539-43 (1987)
MAC.US.-.MM142	Y00277	United States	Chakrabarti, L	<i>Nature</i> <b>328</b> (6130); 543-7 (1987)
MAC.US.-.SMM142B	BD131285	United States	Alizon, M	Patent: JP 2002030099-A 2 29-JAN-2002; INSTITUT PASTEUR
MAC.US.1937	AY611495	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.2065	AY611493	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.239.95112	AY588946	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.239.96114	AY588945	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.80035	AY611486	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.81035	AY599200	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.85013	AY611490	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.87082	AY600249	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.92050	AY603959	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.92077	AY599201	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.93057	AY611492	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)

Name	Accession	Country	Author	Reference
MAC.US.93062	AY607704	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.95058	AY611494	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.95086	AY607703	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.96016	AY607701	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.96020	AY611488	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.96072	AY611491	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.96081	AY597209	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.96093	AY611489	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.96123	AY611487	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.96135	AY607702	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.97009	AY599199	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.97074	AY599198	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.r80025	AY576480	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MAC.US.r90131	AY576481	United States	O'Connor, DH	<i>J Virol</i> <b>78</b> (24):14012-14022 (2004)
MNE.US.-.MNE027	U79412	United States	Kimata, JT	<i>J Virol</i> <b>72</b> (1); 245-56 (1998)
MNE.US.82.MNE_8	M32741	United States	Kimata, JT	<i>J Virol</i> <b>72</b> (1); 245-56 (1998)
SMM.SL.92.SL92B	AF334679	Sierra Leone	Chen, Z	<i>J Virol</i> <b>70</b> (6); 3617-27 (1996)
SMM.US.-.17EC1	AY033233	United States	Anderson, MG	<i>Virology</i> <b>195</b> (2); 616-26 (1993)
SMM.US.-.17EFR	AY033146	United States	Flaherty, MT	<i>J Virol</i> <b>71</b> (8); 5790-8 (1997)
SMM.US.-.F236_H4	X14307	United States	Hirsch, VM	<i>Nature</i> <b>339</b> (6223); 389-92 (1989)
SMM.US.-.H9	M80194	United States	Courgnaud, V	<i>J Virol</i> <b>66</b> (1); 414-9 (1992)
SMM.US.-.PBJ14_15	L03295	United States	Dewhurst, S	<i>Nature</i> <b>345</b> (6276); 636-40 (1990)
SMM.US.-.PBJA	M31325	United States	Dewhurst, S	<i>Nature</i> <b>345</b> (6276); 636-40 (1990)
SMM.US.-.PBJ_143	M80193	United States	Courgnaud, V	<i>J Virol</i> <b>66</b> (1); 414-9 (1992)
SMM.US.-.PBJ_6P6	L09212	United States	Novembre, FJ	<i>J Virol</i> <b>67</b> (5); 2466-74 (1993)
SMM.US.-.PGM53	AF077017	United States	Novembre, FJ	<i>J Virol</i> <b>72</b> (11); 8841-51 (1998)
SMM.US.-.SME543	U72748	United States	Hirsch, VM	<i>J Virol</i> <b>71</b> (2); 1608-20 (1997)
STM.US.-.STM	M83293	United States	Novembre, FJ	<i>Virology</i> <b>186</b> (2); 783-7 (1992)

5' LTR U3 start

MAC.US.-.239	TGGAAAGGGATTATTACAGTCAAGA . AGACATAGAATCTTAGACATATACTTAGAAAAAGGAAGGCATCATACCAGATTGGCAGGATTACACCTCAGGACCAGGAATTAGATACCCAAAGACATTTGGCTGGCTATGGAAATAGTCCCTGTAATGTATCAGATGA	169
H2A.GM.MCN13	.....G-T.....AGGGAT.....C.....T-G.....C.....G.....GG-A.....A.....T-TCAT-G.....G-G-G.....T-TGC-C-G-G.....C-A-A.....CA-G.....C-A.....	0
H2A.CI.88.UC2	.....G-T.....AGG.....G.....C.....T-G.....C.....G.....G-A.....A.....T-TCAT-G.....G-A-G.....T-TAC-C-G-G.....GC-A-A.....G-AC-C.....C-A.....	169
H2A.DE.-.BEN	.....C.....G-T.....AGG.....G.....C.....T-G.....C.....G.....G-A.....A.....T-TCAT-G.....G-A-G.....T-TAC-C-G-G.....GC-A-A.....G-AC-C.....C-A.....	169
H2A.DE.-.PEI2	.....T.....G.....AG.....AG.....A.....G.....A.....G.....G-A.....T.....A.....C.....T-TCAT-G.....G-A-G.....TTC.....G.....GC.....A-A.....G-C-CC.....C-A-G	170
H2A.GH.-.GHI	.....T.....G.....AG.....AG.....A.....G.....A.....G.....G-A.....T.....A.....C.....T-TCAT-G.....G-A-G.....TTC.....G.....GC.....A-A.....G-C-CC.....C-A-G	0
H2A.GM.-.ISY	.....T.....G.....AG.....AG.....A.....G.....A.....G.....G-A.....T.....A.....C.....T-TCAT-G.....G-A-G.....TTC.....G.....GC.....A-A.....G-C-CC.....C-A-G	0
H2A.GM.87.D194	.....G-T.....AG.....A.....C.....T-TCAT-G.....G-A.....G.....G-T-TTC.....G.....G.....GC-A-A.....G-CC.....C-A.....	169
H2A.GW.-.ALI	.....G-T.....AG.....A.....C.....T-TCAT-G.....G-A.....G.....G-T-TTC.....G.....G.....GC-A-A.....G-CC.....C-A.....	0
H2A.GW.-.MIS	.....G-T.....AG.....A.....C.....T-TCAT-G.....G-A.....G.....G-T-TTC.....G.....G.....GC-A-A.....G-CC.....C-A.....	0
H2A.GW.86.FG	.....C-G-T.....A.....C-G-T.....A.....G-A-TG.....A-C-T-TAGT-G.....G-A.....T-TTC.....G.....GC-A-A.....G-ACC.....C-A.....	169
H2A.GW.87.CAM2C6	.....C-G-T.....A.....C-G-T.....A.....G-A-TG.....A-C-T-TAGT-G.....G-A.....T-TTC.....G.....GC-A-A.....G-ACC.....C-A.....	0
H2A.SN.85.ROD	.....C-G-T.....A.....C-G-T.....A.....G-A-TG.....A-C-T-TAGT-G.....G-A.....T-TTC.....G.....GC-A-A.....G-ACC.....C-A.....	0
H2AB.CI.-.7312A	.....T.....T.....A.....G-A.....T-C.....T-T-G.....T-GAG.....T-AG-TGGT-GA.....A-C.....A-AT-G.....G-A.....ACAC.....G.....C.....G-A-GG-A-GG.....CA-T	169
H2B.CI.-.EHO	.....TC-T.....AG.....G-A.....AC.....C.....T.....TG-GT-T-GA.....A-C.....ACAT-G.....G-A.....T-C.....TTC.....G.....GC.....G-AA.....A-GAT-CA.....	169
H2B.CI.88.UC1	.....C-T.....AG.....AC.....C.....TC.....G-T-G.....TG-T-T-GA.....AA-C-T-A-AT-G.....G-A-G-T-G.....G.....GC.....G-A-A-G-A-C.....G.....	169
H2B.GH.86.D205	.....G-T.....AG.....C.....A.....C.....T-T-G.....T.....TG-GT-T-GC.....AA-C-T-TCAT-G.....G-A-GC-T-C-ATAC.....T.....G.....GC-G-A-A-G-G-GC.....CA-C	169
H2B.JP01.KR020	.....G-T.....AG.....C.....A.....C.....T-T-G.....T.....TG-GT-T-GC.....AA-C-T-TCAT-G.....G-A-GC-T-C-ATAC.....T.....G.....GC-G-A-A-G-G-GC.....CA-C	0
H2G.CI.-.ABT96	.....G-T.....AG.....C.....A.....C.....T-T-G.....T.....TG-GT-T-GC.....AA-C-T-TCAT-G.....G-A-GC-T-C-ATAC.....T.....G.....GC-G-A-A-G-G-GC.....CA-C	0
H2U.FR.96.12034	.....G-T.....AG.....C.....A.....C.....T-T-G.....T.....TG-GT-T-GC.....AA-C-T-TCAT-G.....G-A-GC-T-C-ATAC.....T.....G.....GC-G-A-A-G-G-GC.....CA-C	0
MAC.US.-.251_1A11	.....T.....G.....G.....C.....	169
MAC.US.-.251_32H_PJ5	.....T.....G.....G.....C.....	169
MAC.US.-.263_BK28	.....T.....G.....G.....C.....	169
MAC.US.-.MM142	.....T.....G.....G.....C.....	0
MAC.US.-.SMM142B	.....T.....G.....G.....C.....	0
MAC.US.1937	.....T.....G.....G.....C.....	0
MAC.US.2065	.....T.....G.....G.....C.....	0
MAC.US.239.95112	.....T.....G.....G.....C.....	0
MAC.US.239.96114	.....T.....G.....G.....C.....	0
MAC.US.80035	.....T.....G.....G.....C.....	0
MAC.US.81035	.....T.....G.....G.....C.....	0
MAC.US.85013	.....T.....G.....G.....C.....	0
MAC.US.87082	.....T.....G.....G.....C.....	0
MAC.US.92050	.....T.....G.....G.....C.....	0
MAC.US.92077	.....T.....G.....G.....C.....	0
MAC.US.93057	.....T.....G.....G.....C.....	0
MAC.US.93062	.....T.....G.....G.....C.....	0
MAC.US.95058	.....T.....G.....G.....C.....	0
MAC.US.95086	.....T.....G.....G.....C.....	0
MAC.US.96016	.....T.....G.....G.....C.....	0
MAC.US.96020	.....T.....G.....G.....C.....	0
MAC.US.96072	.....T.....G.....G.....C.....	0
MAC.US.96081	.....T.....G.....G.....C.....	0
MAC.US.96093	.....T.....G.....G.....C.....	0
MAC.US.96093	.....T.....G.....G.....C.....	0
MAC.US.96123	.....T.....G.....G.....C.....	0
MAC.US.96135	.....T.....G.....G.....C.....	0
MAC.US.97009	.....T.....G.....G.....C.....	0
MAC.US.97074	.....T.....G.....G.....C.....	0
MAC.US.r80025	.....T.....G.....G.....C.....	0
MAC.US.r90131	.....T.....G.....G.....C.....	0
MNE.US.-.MNE027	.....T.....G.....G.....C.....	0
MNE.US.82.MNE_8	.....T.....G.....G.....C.....	0
SMM.SL.92.SL92B	.....T.....G.....G.....C.....	0
SMM.US.-.17EC1	.....T.....G.....G.....C.....	169
SMM.US.-.17EFR	.....T.....G.....G.....C.....	169
SMM.US.-.F236_H4	.....A-AG.....A.....T-G.....G.....A-A.....AA.....A-G.....T-T-CAC-A-G.....C.....A-G.....C.....	169
SMM.US.-.H9	.....A-AG.....A.....T-G.....G.....A-A.....AA.....A-G.....T-T-CAC-A-G.....C.....A-G.....C.....	0
SMM.US.-.PB114_15	.....A-AG.....A.....T-G.....G.....A-A.....AA.....A-G.....T-T-CAC-A-G.....C.....A-G.....C.....	0
SMM.US.-.PB1A	.....A-AG.....A.....T-G.....G.....A-A.....AA.....A-G.....T-T-CAC-A-G.....C.....A-G.....C.....	0
SMM.US.-.PB1_143	.....AA.....A-G.....C.....T-T-TTC.....A-C.....G-G.....C.....	95
SMM.US.-.PB1_6P6	.....AA.....A-G.....C.....T-T-TTC.....A-C.....G-G.....C.....	95
SMM.US.-.PGM53	.....A-AG-A.....A.....T-G-A-G.....A-A.....AA.....A-T-G.....C.....T-T-TAC.....G-C.....G-A-G-C.....	169
SMM.US.-.SME543	.....A-AG-A.....A.....T-G-A-G.....A-A.....AA.....A-T-G.....C.....T-T-TAC.....G-C.....G-A-G-C.....	169
STM.US.-.STM	.....AG.....	28



TCF-1 alpha

MAC.US.-.239	GGCACAG...GAGGATGAGGAG.....CATTATTTAATGCATCCAGCTCAAACCTCCAGTGGGATGACCCTGGGGAGAGGTTCTAGCATGGAAGTTTGATCCAACCTCGCCCTACACTTATGAGGCATATGTTAGATACCCAGAAGATTTGGAAGCAAGT	324
H2A.GM.MCN13		0
H2A.CI.88.UC2	.....GAG-T-ACC.....A-C-GC-C-A-A-AAG-AGA-AT-ATACAT-G-ACA-TG-CG-A-CT-C-TG-TGAG-CA-C-TCAC-CT-A-GCA	324
H2A.DE-BEN	.....GAG-A-CC.....A-C-GC-C-A-A-AAG-AGACAT-TGAGCAT-G-ACAT-TG-C-G-CT-C-TG-T-AC-CA-C-TCAC-CTG-A-GCA	324
H2A.DE-PEI2	T.....A-C.....C-GC-C-A-A-AAG-GG-CT-CAT-G-AACAT-ATG-G-C-T-GG-TGAG-AC-T-AA-C-A-GTAT	316
H2A.GH-GH1		0
H2A.GM-ISKY		0
H2A.GM.87.D194		0
H2A.GW-AL1	A-GG...GAG-CAC-ACT.....C-GCC-GC-A-C-TA-AAG-AG-CAT-A-GCAT-G-ACAT-TT-GA-C-T-AG-TC-TGA-CA-A-C-T-A-CT-C-G-A-GTA	324
H2A.GW-MIS		0
H2A.GW.86.FG		0
H2A.GW.87.CAM2CG	--G-G--CAC-ACTGACACTGAGACT-C-GC-T-A-C-A-AAG-AG-CAT-ATGCAT-G-ACA-G-TG-CT-C-TG-CTT-AG-C-T-AC-C-A-GCA	336
H2A.SN.85.ROD		0
H2AB.CI-.7312A	AA-C-G-GAG-A-G-ACC.....G-C-G-C-A-A-G-CG-GCA-T-CAT-GAGA-C-T-TC-C-T-CCTC-A-TT-AG-C-TCAAC-G-T-GTATC	326
H2B.CI-EHO	AC-G-...-C-G-AACC.....G-C-GG-A-A-G-C-GCA-CCAC-A-ACC-T-TC-C-CT-CCTC-A-A-TGAC-T-C-TCAGC-G-T-G-GTATC	324
H2B.CI.88.UC1	--A-AG--GA-CG-AACC.....AG-G-C-GG-C-A-A-G-TC-TTCA-ATCCAC-G-ACC-T-C-CG-C-CCTC-AC-TGAC-TA-C-TCAAC-G-GTATC	324
H2B.GH.86.D205	-A-C-GAGAG--G-AACC.....GCC-C-G-A-A-G-TC-TCA-ATCCAT-G-AC-TATC-C-T-CCTC-A-TGA-T-T-TCAA-G-TT-C-GTA-C	327
H2B.JP01.KR020		0
H2G.CI-ABT96		0
H2U.FR.96.12034		0
MAC.US.-.251.1A11		324
MAC.US.-.251.32H.PJ5	.....AG-G.....G-A.....C-A.....C	324
MAC.US.-.251.BK28	.....AG-G.....G-A.....G-A.....C	323
MAC.US.-.SMM142		0
MAC.US.-.SMM142B		0
MAC.US.1937		0
MAC.US.2065		0
MAC.US.239.95112		0
MAC.US.239.96114		0
MAC.US.80035		0
MAC.US.81035		0
MAC.US.85013		0
MAC.US.87082		0
MAC.US.92050		0
MAC.US.92077		0
MAC.US.93057		0
MAC.US.93062		0
MAC.US.95058		0
MAC.US.95086		0
MAC.US.96016		0
MAC.US.96020		0
MAC.US.96072		0
MAC.US.96081		0
MAC.US.96093		0
MAC.US.96123		0
MAC.US.96135		0
MAC.US.97009		0
MAC.US.97074		0
MAC.US.r80025		0
MAC.US.r90131		0
MNE.US.-.MNE027		0
MNE.US.82.MNE_8		0
SMM.SL.92.SL92B		0
SMM.US.-.17EC1		324
SMM.US.-.17EFR		324
SMM.US.-.F236_H4	A-T-...-A-C-ACA.....GCC-GG--G-A-G-AT-----C-----A-G-GAAT-A-T-T-GC-A--T-A-AG-----T-T	324
SMM.US.-.H9		0
SMM.US.-.PB114_15		79
SMM.US.-.PB1A		79
SMM.US.-.PB1_143		0
SMM.US.-.PB1_6P6	A-T-...-A-C-ACA.....C-G-A-T-----C-A-G-AGT-A-T-T-AC-A--T--GAGC-----T-TC	250
SMM.US.-.PGM53	A-T-A--A-C-ACA.....C-G-A-T-----C-A-G-AGT-A-T-T-AC-A--T-A--C-----T-T	250
SMM.US.-.SME543	A-C-...-A-C-ACA.....GCC-GG--A-G-CAT-----C-A-G-GAAT--T-T-GC-A--T-A-AG-----T-T	324
STM.US.-.STM		28





	<u>NF-κB-I</u>	<u>TATA Box</u>	+1 mRNA start site TAR element start 5' LTR U3 end 5' LTR R repeat begin	
MAC.US.-.239	.....GGGGATGTACG.GGGAGGTACTGGGAGGAGCCGGTCGGG..AACGCCCACTTTCTTGATGT.ATAAATCACTGCATTTCCGCTCTGTA.TTCAGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCTCCAGCACT			573
H2A.GM.MCN13	.....			55
H2A.CI.88.UC2	.....C--A--C-AA-G-AGGGACAT-G-AG-AGCTGGT--GG--T-A-A--AC--G-AC-C-T-CC--A--T-C-----C-----A-----			610
H2A.DE.-.BEN	.....C--A--C-AA-G-AGGGACAT-G-AG-AGCTGGT--GG--T-A-A--AC--G-AC-C-T-C-T--AT-----C-----A-----			610
H2A.DE.-.PEI2	.....C--A--C-AA-G-AGGGAC-T-G-AG-AACCGGT--GG--T-A-A--C-----G-AC-C-TGC-T-AT-----C-----			590
H2A.GH.-.GHI	.....			55
H2A.GM.-.ISY	.....			55
H2A.GM.87.D194	.....			55
H2A.GW.-.ALI	.....C--A--C-A--G-AAGGCAT-G-AG-AGCTGGT--GG--T-A-A--C.C-----G-AC-C-TGC-T-AT-----			606
H2A.GW.-.MDS	.....			55
H2A.GW.86.FG	.....			54
H2A.GW.87.CAM2CG	.....C--AC-A--G-AGGGATAT-G-AG-AGCTGGT--GG--T-A-A-C-T.CTGT--TCAC-C-TGC-T-AT--C-----G.C			621
H2A.SN.85.ROD	.....			55
H2AB.CI.-.7312A	.....C--A--A--T-G-AGGGACAT-G-AG-AGTTGGT--GG--T-AAGCCT.CTGT--G-AC-C-T-C-T-AT-----G			618
H2B.CI.-.EHO	.....C--A--A--T-G-AGGGACAT-G-AG-AGCA.G--AGG--TCAAGACT.CTGT--G-AC-C-T-GT-AT-----C-----			613
H2B.CI.88.UC1	.....A-C--A--A--T-G-AGGGACAT-G-AG-AGTTGGTT--GG--T-AAA-CT.CTGT--G-AC-C-T-C-T-AT-----C-----			613
H2B.GH.86.D205	.....A-C--A--A--T-G-AGGAGCAA-G-AG-GGTTGGTT--GG--T-AAA-CT.CTGT--G-AC-C-T-C--AT-----			616
H2B.JP01.KR020	.....			0
H2G.CI.-.ABT96	.....			0
H2U.FR.96.12034	.....			106
MAC.US.-.251.1A11	.....			572
MAC.US.-.251.32H.PJ5	.....			573
MAC.US.-.251.BK28	.....T-----AG-----A-----A-----G-----			561
MAC.US.-.MM142	.....			55
MAC.US.-.SMM142B	.....			55
MAC.US.1937	.....			0
MAC.US.2065	.....			0
MAC.US.239.95112	.....			0
MAC.US.239.96114	.....			0
MAC.US.80035	.....			0
MAC.US.81035	.....			0
MAC.US.85013	.....			0
MAC.US.87082	.....			0
MAC.US.92050	.....			0
MAC.US.92077	.....			0
MAC.US.93057	.....			0
MAC.US.93062	.....			0
MAC.US.95058	.....			0
MAC.US.95086	.....			0
MAC.US.96016	.....			0
MAC.US.96020	.....			0
MAC.US.96072	.....			0
MAC.US.96081	.....			0
MAC.US.96093	.....			0
MAC.US.96123	.....			0
MAC.US.96135	.....			0
MAC.US.97009	.....			0
MAC.US.97074	.....			0
MAC.US.r80025	.....			0
MAC.US.r90131	.....			0
MNE.US.-.MNE027	.....			55
MNE.US.82.MNE.8	.....			55
SMM.SI.92.SI.92B	.....			573
SMM.US.-.17EC1	.....			573
SMM.US.-.17EER	.....			570
SMM.US.-.F236_H4	.....C--C-T-----T-CT-----AT-CTC-----CA-----A-----			56
SMM.US.-.H9	.....			350
SMM.US.-.PB114_15	.....T-CT-----G-AT-CTC-----CA-----A-----			350
SMM.US.-.PB1A	.....CT-----G-AT-CTC-----CA-----A-----			56
SMM.US.-.PB1_143	.....			56
SMM.US.-.PB1_6P6	CAGCAGGGACTTTCCACAAA.....T-CT-----G-AT-CTC-----CA-----A-----			543
SMM.US.-.PGM53	.....T-----A-T-T-----C-----AT-TC-----CA-----G-C-----			499
SMM.US.-.SME543	.....C--C-T-----T-CT-----C-----AT-CTC-----CA-----G-----			571
STM.US.-.STM	.....C--A-AG-----T-G-----TT-A-T-C-----CA-C-T-A-A-----A-----			246









MAC.US.-.239
Gag
H2A.GM.MCN13
H2A.CI.88.UC2
H2A.DE.-.BEN
H2A.DE.-.PEI2
H2A.GH.-.GH1
H2A.GM.-.ISY
H2A.GM.87.D194
H2A.GW.-.AL1
H2A.GW.-.MDS
H2A.GW.86.FG
H2A.GW.87.CAM2CG
H2A.SN.85.ROD
H2AB.CI.-.7312A
H2B.CI.-.EHO
H2B.CI.88.UC1
H2B.GH.86.D205
H2B.JP01.KR020
H2G.CI.-.AB196
H2U.FR.96.12034
MAC.US.-.251.1A11
MAC.US.-.251.32H.PI5
MAC.US.-.251.BK28
MAC.US.-.MM142
MAC.US.-.SMM142B
MAC.US.1937
MAC.US.2065
MAC.US.239.95112
MAC.US.239.96114
MAC.US.80035
MAC.US.81035
MAC.US.85013
MAC.US.87082
MAC.US.92050
MAC.US.92077
MAC.US.93057
MAC.US.93062
MAC.US.95058
MAC.US.95086
MAC.US.96016
MAC.US.96020
MAC.US.96072
MAC.US.96081
MAC.US.96093
MAC.US.96123
MAC.US.96135
MAC.US.97009
MAC.US.97074
MAC.US.r80025
MAC.US.r90131
MNE.US.-.MNE027
MNE.US.82.MNE.8
SMM.SL.92.SL92B
SMM.US.-.17EC1
SMM.US.-.17EFR
SMM.US.-.F236\_H4
SMM.US.-.H9
SMM.US.-.PB114\_15
SMM.US.-.PB1A
SMM.US.-.PBJ\_143
SMM.US.-.PBJ\_6P6
SMM.US.-.PGM53
SMM.US.-.SME543
STM.US.-.STM

GCCTGTTGGAGAACAAGAGGATGTCAAAAATCCTTTGGTCTTAGCTCCATTAGTGCCACAGGCTCAGAAAAATTTAAAAAGCCCTTTATAACTGTCTCGCTCATCTGGTGCATTCACGCGAGAAGAGAAAGTGAACACACTGAGGAAGCAAAACAGATAGTGCAG 1367
S L L E N K E G C Q K I L S V L A P L V P T G S E N L K S L Y N T V C V I W C I H A E E K V K H T E E A K Q I V Q
A-----TCA-----T-C-G-T--A-A-T--G-----A-G-T-----T-----A-----G-T-A-G-----A-A--A-- 862
-----TCA-----T-C-G-T--CA-A-T--A-----A-G-----T-----C-----T-G-----G-T-A-G-----A-AC--A-- 1417
-----TCA-----T-C-G-T--CAGA-T--A-----A-G-----T-----C-----T-G-----G-T-A-G-----GA-AC--CA-- 1417
-----TCA-----T-C-----TA-A-A-T--A-----A-G-----T-----T-G-----G-T-A-G-----A-A--A-- 1385
-----TCA-----T-C-G-T--CA-A-T--A-----A-G-----T-----C-A-----T-G-----G-T-A-G-----A-AC--A-- 858
-----TCA-----T-C-----TA--A-T--AA-----A-G-----T-----AC-----T-G-----G-T-A-G-----AGA--AGG-- 861
-----ATCA-----T-C-G-T--AAA-T--AA-----A-G-----T-----C-----AG-T-G-----G-T-A-G-----GA-AC--CA-- 861
-----TCA-----T-C-G-T--AAA-T--AA-----A-G-----T-----T-----AG-T-G-----G-T-A-G-----A-A--A-- 1410
-----ATCA-----G-T-C-G-T--A-A-T--A-----A-G-----T-----T-TC-----A-----G-----G-T-A-G-----A-T-C-CA-G- 861
-----TCA-----T-C-G-T--A-A-T--A-----A-G-----T-----T-----A-----A-----G-T-A-G-----A-A--A-- 861
-----ATCA-----T-C-G-T--AAA-T--A-----A-G-----T-----T-----A-----A-----G-T-A-G-----AGA--CA-T- 1426
-----TCA-----G-T-----T--A-A-T--A-----A-G-----T-----T-----A-----A-----G-T-A-G-----A-A--G- 860
A-----ATCA-----G-T-A-----A-----G-----T-----T-----T-A-----A-----G-T-C-----A-C--CA-- 1399
-G-----GATCA-----C-GG-T--GGAAA-----GA-----A-----T-----T-----C-----TT-C-G-----G-T-----A--CA-- 1394
-----ATCA-----C-----T--A-----A-----A-----T-----T-----AC-T-G-----G-T-----A-----CA-- 1398
-----ATCA-----C-----T--AAA-----A-----G-----T-----T-----TC-----TT-C-G-----G-T-A-----A--CA-- 1393
-----TCA-----C-----T--A-----AA-----A-----A-----T-----T-----AC-TC-G-----G-T-----A-T--T--A--CA-- 339
-----ATCA-----T-C-----T--A-A-----A-----A-----G-----T-----TT-G-----C--C-----A-T-A-----AG----- 780
-----GC-----TT-A-----G-----T-----T-----G-----T-----TT-G-----C--C-----A-T-A-----G-----A-- 896
-----A----- 1366
----- 1368
----- 1355
----- 865
----- 865
-----R-----G-----A-----M-----R----- 473
----- 473
----- 463
----- 469
----- 473
-----R----- 474
-----R----- 407
----- 444
----- 467
----- 474
----- 472
----- 459
-----R----- 461
----- 474
-----T----- 466
-----A----- 478
----- 476
----- 474
----- 363
----- 459
-----T----- 342
-----D-----T----- 342
----- 330
-----C-----G----- 473
----- 847
----- 847
-----A-G-----G-----T-AG-A-AC-----A-----G-----T-G-T-G-----AG-T-TC-G-----AT-----G-A-----T-GA-A-C-A- 807
-----T-----T-----T-A-T-----T-G-T-----C-G-----T-----T-----A-----A----- 1367
-----T-----T-----T-A-T-----T-G-T-----G-----C-----C-----C-T-----T-----A-----A----- 1363
-----T-----T-----T-A-T-----T-G-T-----G-----C-----C-----C-T-----T-----A-----A----- 849
-----T-----T-----T-A-T-----T-G-T-----G-----C-----C-----C-T-----T-----A-----A----- 1143
-----T-----T-----T-A-T-----T-G-T-----G-----C-----C-----C-T-----T-----A-----A----- 1142
-----T-----T-----T-A-T-----T-G-T-----G-----C-----C-----C-T-----T-----A-----A----- 849
-----C-----T-----TA-A-T-----T-GG-T-----G-----C-----C-----C-T-----T-----A-----A----- 1335
-----T-----T-----T-A-T-----T-G-T-----G-----G-----T-----T-----A-----A----- 1293
-----T-----T-----T-A-T-----T-G-T-----G-----G-----T-----T-----A-----A----- 1364
-----TCA-----G-----TA-AA-----AA-----T-G-----T-----G-T-----T-----T-----C-----T-----G-AG-G-AA- 1023













MAC.US.-.239
Pol
Gag
H2A.GM.MCN13
H2A.CI.88.UC2
H2A.DE.-BEN
H2A.DE.-PEI2
H2A.GH.-GH1
H2A.GM.-J5Y
H2A.GM.87.D194
H2A.GW.-ALI
H2A.GW.-MDS
H2A.GW.86.FG
H2A.GW.87.CAM2CG
H2A.SN.85.ROD
H2AB.CI.-7312A
H2B.CI.-EHO
H2B.CI.88.UC1
H2B.GH.86.D205
H2B.JP01.KR020
H2G.CI.-ABT96
H2U.FR.96.12034
MAC.US.-.251.1A11
MAC.US.-.251.32H.P15
MAC.US.-.251.BK28
MAC.US.-.SMM142
MAC.US.-.SMM142B
MAC.US.19037
MAC.US.2065
MAC.US.239.95112
MAC.US.239.96114
MAC.US.80035
MAC.US.81035
MAC.US.85013
MAC.US.87082
MAC.US.92050
MAC.US.92077
MAC.US.93057
MAC.US.93062
MAC.US.95058
MAC.US.95086
MAC.US.96016
MAC.US.96020
MAC.US.96072
MAC.US.96081
MAC.US.96093
MAC.US.96123
MAC.US.96135
MAC.US.97009
MAC.US.97074
MAC.US.r80025
MAC.US.r90131
MNE.US.-.MNE027
MNE.US.82.MNE.8
SMM.SL.92.SL.92B
SMM.US.-.17EC1
SMM.US.-.17EFC
SMM.US.-.F236\_H4
SMM.US.-.H9
SMM.US.-.PBJ14\_15
SMM.US.-.PBJA
SMM.US.-.PBJ\_143
SMM.US.-.PBJ\_6P6
SMM.US.-.PGM53
SMM.US.-.SME543
STM.US.-.STM

Sequence alignment table with columns for amino acid positions (A, Q, Q, R, L, G, P, R, K, P, I, K, C, W, N, C, G, K, E, G, H, S, A, R, Q, C, R, A, P, R, R, Q, G, C, W, K, C, G, K, M, D, H, V, M, A, K, C, P, D, R, Q, A, G, F, L, F, G) and corresponding nucleotide sequences for various HIV-2/SIV strains.

Gag-Pol -1 ribosomal slip site
Pol start
Gag p8 Nucleocapsid end Gag p1 Spacer start

Strain	Sequence	Position
MAC.US.-.239	CTTGGTCCATGGGAAAGAGCCCGCAATTTCCCATGGCTCAAGTGCATCAGGGGCTGATGCCAACTGCTCC...CCAGAGGACCCAGCTGGATGCTCTAAAGAATACATGCAGTTGGGCAAGCAGCAGAG...	2494
Pol	R P W S M G K E A P Q F P H G S S A S G A D A N C S P R G P S C G S A K E L H A V G Q A A	
Gag	L G P W G K K P R N F P M A Q V H Q G L M P T A P P E D P A V D L L K N Y M Q L G K Q R	
H2A.GM.MCN13	A-G-C-CC-----C-G-C---C-CG-----A-CA---A-A-CCC---TA-T---A---C-A--GG---T-T---CAA-A-AAGA--A-AGAACAGAGAGAGACCATACAAGGAGGTGAC	2016
H2A.CI.88.UC2	A-G-CT-C-----C---C---CT-C-----A-A-A-CCC---CA-T---G-A-C-C-T-G---A-TT---CA-G-GAA---GGAGCAGAGGAGAGACCATACAAGGAGGTGAC	2571
H2A.DE.-BEN	T-G-C-C-----T-C-G-A-C-CC-C-----A-A-A-TCC---G-CA-T---A-C-A--T-GG--GA-T---CAA-G-GAA---GGAGCAGAGGAGAGACCATACAAGGAGGTGAC	2571
H2A.DE.-PEI2	A---C-----T-C-G-A-C-G--C-CG-----CA-A-A-CCC---CA---A-CA-C---G-G-T---CAA-G-G-A---AAGAGCAGAAAATGAGACCATACAAGGAGGTGAC	2539
H2A.GH.-GH1	A-G-----T-C-G---C-CT-C-C-----A-A-A-CCC---CA-T---A---AT-GG-A-GA-T---CA-G-GAG---GGAGCAGAGGAGAGACCATACAAGGAGGTGAC	2015
H2A.GM.-J57	A---A-----C-G-TC-AGTTG-----A-CA---A-A-CCC---AT-T---A---C-A--GG---G---CAA-G-GAA-A---AGAGCAGAGACAAGACCATACAAGGAGGTGAC	2012
H2A.GM.87.D194	A-G-C--C-----C---G-C-CT-C-----A-A-A-CCC---ATA-T---A---C-T-GG---A-T---CAA-G-GAA---AGAGCAGAGAGAGACCATACAAGGAGGTGAC	2015
H2A.GW.-ALI	A---C-----T-C-G-AA-C-G--CT-CG---AT-A-CA---A-A-CCC---CA-G---A-C-C---GG-C-G-T---CA-G-GAA---AGAGCAGAGAGAGACCATACAAGGAGGTGAC	2564
H2A.GW.-MDS	A---C-C-----C-G-C-G-T-CG-----A-CA-A-A-CCC---CA-T---A---C-A--GG---A-T---CAA-G-GAA---GGAGCAGAGAGAGACCATACAAGGAGGTGAC	2015
H2A.GW.86.FG	A---C-AC-----G---C-GC-C--C-CG---A-CA-A-A-CCC---TT-T---A---C-A--GG---A-T---CAA-A-AAG--A-GGAGCAGAGAGAGACCCTACAAGGAGGTGAC	2009
H2A.GW.87.CAM2CG	A---C-C-----C-G-C---T-CA---A-CA-A-A-CCC---TT-T---A---A---A-GG---A-T---CAA-G-AAG---GGAGCAGAGACAGACCATACAAGGAGGTGAC	2580
H2A.SN.85.ROD	A-G-C-T-----C-G-C---T-CG---A-CA-A-A-CCC---T-T---A---A---A-GG---A-T---CAA-G-AAGA---AGAGCAGAGAGAGACCATACAAGGAGGTGAC	2014
H2AB.CI.-7312A	---A-C-----T-C---C---C---A-A-AG---T---G-CCCGATGA-CACAG---AG-GCA-GACA-ATC---GGAGCG--A-CA-CT-CGCCC-CTGCAGATCCCGCAGTGGAGATGCTGAAAAGCTACATGCA	2553
H2B.CI.-EHO	T-C-C-C-----T---G---C-G-CA-CC---A-AG---T---G-CCCGATGA-CCCAG---TTC-GCATGACA-CTC---GGTGGC--T-CA-CT-CACCC-CTGCAGATCCAGCAGAGGAGATGCTAAAGAACTACATGCA	2542
H2B.CI.88.UC1	T-A-A-C-----T-C---A-C---C---AG-CA-T---A-CCCGATG-CCCAG---AG-GCATGACA-CTCG-GGGCGG-CA-CA-CT-CGCCC-CTGCAGATCCAGCAGTGGAGATGCTGAAAAGCTACATGAA	2552
H2B.GH.86.D205	T-A-A-C-----T-C---A-C---C---AG-CA-T---A-CCCGATGA-CCCAG---AG-GCATGACA-CTCG-GGGCGG-CA-CA-CT-CGCCC-CTGCAGATCCAGCAGTGGAGATGCTGAAAAGCTACATGCA	2547
H2B.JP.01.KR.020	T---C-C-----T-C---TG-A-C---CA-CC---A-AC---T---A-ACCAATGA-CCCAG---AGAATGACA-CTC---GGTGCA---CA-CT-CGCCC-CTGCAGATCCAGCAGTGGAGATGTTGAAGGACTACATGCA	1693
H2G.CI.-ABT96	T---A-A-T-----T-C---A-C---C-C-T---CT-T-A-A---A-GAT---A-A-C-A-G---T---RC---G-G-A---A-A	1916
H2U.FR.96.12034	---C-T---A---A---T---C---C---C-CG---A---CC---A-C---T---CA-A---A---G-CCCGACAGCC-CTCCA-CGG-C-CAGCAGTGA...TCTGTAAGAAGCTACATGCA	2053
MAC.US.-.251_1A11	C	2493
MAC.US.-.251_32H_P15	C	2495
MAC.US.-.251_BK28	C	2482
MAC.US.-.MM142	C-T	1992
MAC.US.-.SMM142B	C-T	1992
MAC.US.1937	C	1600
MAC.US.2065	C	1600
MAC.US.239.95112	C	1590
MAC.US.239.96114	C	1596
MAC.US.80035	C	1600
MAC.US.81035	C	1601
MAC.US.85013	C	1534
MAC.US.87082	C	1571
MAC.US.92050	C	1594
MAC.US.92077	C	1601
MAC.US.93057	C	1599
MAC.US.93062	C	1586
MAC.US.95058	C	1588
MAC.US.95086	C	1457
MAC.US.96016	C	1601
MAC.US.96020	C	1593
MAC.US.96072	C	1605
MAC.US.96081	C	1603
MAC.US.96093	C	1601
MAC.US.96123	C	1490
MAC.US.96135	C	1586
MAC.US.97009	C	1469
MAC.US.97074	C	1469
MAC.US.r80025	C	1457
MAC.US.r90131	C	600
MNE.US.-.MNE027	T---C---A---CA---A---A	1974
MNE.US.82.MNE_8	T---C---A---C---A	1974
SMM.SL.92.SL.92B	---C-----T---C---G-GAC-.AACATCAT-CA-GT-A-A-.T---A-CCCAGG-AG-C-G-G-TC-AG-AAA-CAC--AG-G-A-A.	1937
SMM.US.-.17EC1	C	2494
SMM.US.-.17EFR	C	2494
SMM.US.-.F236_H4	C-GA-C---A---T---A-G-T---A-A-GAA	2493
SMM.US.-.H9	C-GA-C---CA---Y---T---A-G-T---A-G-GAAG	1979
SMM.US.-.PB114_15	C-GA-C---CA---Y---T---A-G-T---A-A-GAAG	2273
SMM.US.-.PB1A	C-GA-C---CA---T---A-G-T---A-A-GAAG	2272
SMM.US.-.PB1_143	C-GA-C---CA---Y---T---A-G-T---A-G-GAAG	1979
SMM.US.-.PB1_6P6	C-GA-C---CA---T---A-G-T---A-A-GAAG	2465
SMM.US.-.PGM53	C-GA-C---C---T---A-G-G-T---A-A-GAA	2423
SMM.US.-.SME543	T---C---C-GA-C---A---T---A-G-T---A-A-GAA	2494
STM.US.-.STM	T---C---A-C---C-A-A-C---C---C---T---A---C---G-GA-GT---C---A	2153

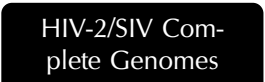


MAC.US.-.239	AGGAATAGAGTTAGGTCACATTATACCCAAAAATAGTAGGAGGAATAGGAGGTTTTATTAATACTAAAGAATACAAAAATGTAGAAATAGAAGTTTTAGGCAAAAGGATTAAGGGACAATCATGACAGGGGACACCCCGATTAACTTTTTGGTAGAAATTTGCTAA	2825
Pol	T G I E L G P H Y T P K I V G G I G G F I N T K E Y K N V E I E V L G K R I K G T I M T G D T P I N I F G R N L	
H2A.GM.MCN13	..G.....GAGCA...GT...G.....G...A...C...A...C...G...T...A...A...T...A...G...A...GG...CC...C...A...T...A...C...C...C...A...T...G...2353	
H2A.CI.88.LUC2	.....GGACA...C...T...G...G...G...A...C...A...C...C...T...G...A...A...AC...AAT...AG...A...G...CC...C...A...A...T...A...C...C...C...A...T...G...2358	
H2A.DE.-.BEN	..A.....GGACA...C...T...G...G...G...A...C...A...C...C...T...G...A...A...AC...AAT...AG...A...G...CC...C...A...A...T...A...C...C...C...A...T...G...2908	
H2A.DE.-.PEI2	.....G...GAGCA...GT...G...G...G...A...C...A...C...C...G...T...A...A...AC...AAT...AG...A...G...CC...C...A...T...T...A...C...C...C...A...T...G...2876	
H2A.GH.-.GH1	.....C...GGACA...CGTT...G...G...G...A...C...A...C...C...GATT...A...A...AC...AAT...AG...A...G...CC...C...A...A...T...A...C...C...C...A...T...G...2352	
H2A.GM.-.ISY	.....GAGCA...GT...G...G...G...A...C...A...C...C...T...G...A...A...AG...GC...AAT...A...G...A...G...CC...C...A...T...T...A...C...C...C...A...T...G...2349	
H2A.GM.87.D194	.....GGACA...C...T...G...G...G...A...C...A...C...C...T...G...A...A...AG...GC...AAT...AG...A...G...CC...C...A...A...T...A...C...C...C...A...T...G...2352	
H2A.GW.-.ALI	.....GAGCA...C...T...G...G...G...A...C...A...C...C...T...G...G...A...A...AC...AAT...AG...A...G...CC...C...A...T...A...C...T...C...A...T...G...2901	
H2A.GW.-.MDS	.....A...AAGCA...GT...G...G...G...A...C...A...C...C...G...T...G...A...A...CC...T...G...A...GG...CC...C...A...T...T...A...C...T...C...A...T...G...2352	
H2A.GW.86.FG	.....GAGCA...GT...G...G...G...A...C...A...C...C...G...T...G...A...A...CC...T...G...A...G...CC...C...A...T...T...A...C...C...C...A...T...G...2346	
H2A.GW.87.CAM2CG	.....A...GAGCA...G...G...G...G...A...C...A...C...C...G...T...G...A...A...CC...T...G...A...GG...CC...T...A...C...C...A...C...C...A...T...G...2917	
H2A.SN.85.ROD	.....GAACA...G...G...G...G...A...C...A...C...C...G...T...G...A...A...CC...T...G...A...ACGG...CC...C...A...C...C...A...C...C...A...T...G...2354	
H2AB.CI.-.7312A	.....A...AAGCA...C...G...G...G...C...A...A...G...G...CA...A...AG...A...A...G...A...GGTCA...T...A...A...A...A...T...C...A...CT...2890	
H2B.CI.-.EHO	..G...A...CAGCA...C...T...G...G...A...A...A...C...T...A...A...AG...A...AG...A...G...CA...G...A...A...A...A...C...A...TT...2879	
H2B.CI.88.UC1	.....A...AGCA...C...G...G...G...C...A...A...C...C...G...G...A...A...AG...G...A...AG...A...GG...CA...T...A...A...A...A...A...C...A...TT...2889	
H2B.GH.86.D205	..G...A...AGCA...C...G...G...G...C...A...A...C...C...G...G...A...A...AG...G...A...AG...A...GG...CA...T...A...A...T...A...A...C...A...TT...2884	
H2B.JP01.KR020	.....A...AGTG...C...G...G...G...C...A...A...C...C...G...G...A...A...AG...A...A...A...G...CA...T...A...A...T...A...A...C...A...T...2030	
H2G.CI.-.ABT96	.....A...AAGTA...GT...G...G...G...A...A...A...C...C...A...A...G...G...A...A...AC...G...CT...TG...T...A...T...A...T...C...A...G...CA...TT...G...2235	
H2U.FR.96.12034	..G...TA...A...AA...T...T...CT...G...A...A...G...C...A...A...TTC...C...A...C...A...A...G...CA...G...T...C...A...A...A...C...T...G...C...TT...G...2390	
MAC.US.-.251.1A11	.....T...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...2824	
MAC.US.-.251.32H.PI5	.....T...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...2826	
MAC.US.-.251.BK28	.....A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...2801	
MAC.US.-.MM142	.....A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...2311	
MAC.US.-.SMM142B	.....A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...2311	
MAC.US.1937	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1931	
MAC.US.2065	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1931	
MAC.US.239.95112	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1921	
MAC.US.239.96114	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1927	
MAC.US.80035	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1931	
MAC.US.81035	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1932	
MAC.US.85013	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1865	
MAC.US.87082	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1902	
MAC.US.92050	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1925	
MAC.US.92077	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1932	
MAC.US.93057	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1930	
MAC.US.93062	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1917	
MAC.US.95058	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1919	
MAC.US.95086	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1788	
MAC.US.96016	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1932	
MAC.US.96020	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1924	
MAC.US.96072	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1936	
MAC.US.96081	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1934	
MAC.US.96093	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1932	
MAC.US.96123	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...821	
MAC.US.96135	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1917	
MAC.US.97009	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1800	
MAC.US.97074	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1800	
MAC.US.r80025	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1788	
MAC.US.r90131	.....Y...A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...1931	
MNE.US.-.MNE027	.....A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...2293	
MNE.US.82.MNE.8	.....A...G...A...G...T...C...A...A...G...T...C...A...A...A...C...T...G...C...TT...G...A...2293	
SMM.SL.92.SL92B	..G.....GGG...C...A...G...T...G...G...T...A...C...A...A...T...G...CA...A...A...GGTA...A...G...AA...C...A...A...T...A...T...C...A...T...G...2256	
SMM.US.-.17EC1	.....GGG...C...A...G...T...G...G...T...A...C...A...A...T...G...CA...A...A...GGTA...A...G...AA...C...A...A...T...A...T...C...A...T...G...2825	
SMM.US.-.17EFR	.....GGG...C...A...G...T...G...G...T...A...C...A...A...T...G...CA...A...A...GGTA...A...G...AA...C...A...A...T...A...T...C...A...T...G...2825	
SMM.US.-.F236_H4	..G...A...G...A...T...C...C...C...C...T...G...A...A...A...C...GGTA...G...A...T...G...A...T...A...T...C...C...2812	
SMM.US.-.H9	..G...A...G...A...T...C...C...C...C...T...G...A...A...A...A...GTA...G...A...T...A...T...A...T...C...C...2298	
SMM.US.-.PB114_15	..G...T...A...G...A...T...C...C...C...C...T...G...A...A...A...A...GTA...G...A...T...A...T...A...T...C...C...2592	
SMM.US.-.PB1A	..G...T...A...G...A...T...C...C...C...C...T...G...A...A...A...A...GTA...G...A...T...A...T...A...T...C...C...2591	
SMM.US.-.PBJ_143	..G...T...A...G...A...T...C...C...C...C...T...G...A...A...A...A...GTA...G...A...T...A...T...A...T...C...C...2298	
SMM.US.-.PBJ_6P6	..G...T...A...G...A...T...C...C...C...C...T...G...A...A...A...A...GTA...G...A...T...A...T...A...T...C...C...2784	
SMM.US.-.PGM53	..G...AC...A...T...C...C...C...C...G...T...G...A...A...A...C...GGTA...G...A...T...A...T...A...T...C...C...T...2742	
SMM.US.-.SME543	..G...A...G...A...T...C...C...C...C...T...G...GA...A...C...GGTA...G...A...T...G...A...T...A...T...C...C...2813	
STM.US.-.STM	..G...G...C...G...T...A...C...GG...T...A...A...G...TT...G...A...C...G...A...G...AA...A...A...T...T...A...A...C...G...CC...AT...G...2472	



Pol Protease end\_Pol p66 and p51 RT start

MAC.US.-239	CAGCTCTGGGGATGCTCTAAATTTCCATAGCTAAAGTAGACCTGTAAGGCTGAAAGCCAGGAAAGGATGGACAAAATTAAGCAGTGGCCATTATCAAAAGAAAAGATAGTTGCATTAAAGAGAAATCTGTGAAAAG.ATGGAAAAGGATGGTCAGTTGGA	2522
Pol	L T A L G M S L N F P I A K V E P V K V A L K P G K D G P K L K Q W P L S K E K I V A L R E I C E # K M E K D G Q L	
H2A.GM.MCN13	CT A C AT A AG C C A A AA A AATG A C G A C A A AG C A GA A C A	2522
H2A.CI.88.UC2	CT A C AT A AG C C GGA AA AA A G A C GGC A A CC A AG AA C AG G A A G C AC A	3077
H2A.DE.BEN	CT A C AT A AG T C GA AA A G G A A G A G A AA C A G A A G C C A	3077
H2A.DE.PE12	CT A C AT C A AG C C G C GA AATAC G A A A G A GA A TC A AG C A A GA A C C A	3045
H2A.GH.GH1	CT A C AT A A C C G A AA AA A G G A A G C GA A CC A G A AA C G C A A G C AC A	2521
H2A.GM.ISY	CT A C AT A AG C C A A A AA A A A G A G CAA GA A C A G A AA C A A GA A C A	2518
H2A.GM.87.D194	A CT A C AT CC A AG C C GT AA AA A G G A A G GGC A A CC A A AA C A T A A G C AC A	2521
H2A.GW.AL1	CT A C AT CC A AG T C GA AA A G A GA A A C G A A A AG C A A A C GA A C A	3070
H2A.GW.MDS	CT A C AT C A AG C C GA A AA A AATGC G A C GA A A A A C A G A A GA A C C A	2521
H2A.GW.86.FG	T A C AT CC G AG T C GA A AA A AATGC G A A A C A A T A A A A C A A A C C A	2515
H2A.GW.87.CAM2CG	CT A C AT C A AG C C GA A AA A AATGC G A A G C GA A C GA A AA C A A G A A A C C A	3086
H2A.SN.85.ROD	CT A C AT C A AG C C AA A AATGC G A A C GA A C A A A AA C A A A C C A	2523
H2AB.CI.-7312A	ACA CT A C A C AG G A A A A G G CAG T A C GA A TC T C G CC T G A G G C A	3059
H2B.CI.-EHO	ATAGCT A C A C AG A GGA A A A CAG T A A A G A C GA A CC C G A C C C A A A G A A	3048
H2B.CI.88.UC1	ATA AT A C A C AG A GA A A AAAC T G A A G A C GA A TC C T G CC C AG A A C A	3058
H2B.JH.86.D205	ATA CT C A T C AG G A G A A T AG A T A G GA C GA A TC C GG C A CC C A A G A A A	3053
H2B.JP01.KR020	ATA CT C A T C AG A A A A G AAAC A T A A C GA A TC T C G CC T A G AA A G A AC A	2200
H2G.CI.-ABT96	TTAAG A C A A T A G A A AAGGC A T C T G A C G A CC T G A CAA C C G T C A C A G A	2404
H2U.FR.96.12034	AA C C A GA A CA TA C TT GG GAC A C A GA T T G AT AA A G A G A G A A G A A	2559
MAC.US.-251.1A11	C A T A T	2993
MAC.US.-251.32H.P15	T C G A T	2995
MAC.US.-251.BK28	C G A A T	2970
MAC.US.-MM142	C G GTCGC	2480
MAC.US.-SMM142B	C G GTCGC	2480
MAC.US.1937	A	2100
MAC.US.2065		2100
MAC.US.239.95112		2090
MAC.US.239.96114		2096
MAC.US.80035		2100
MAC.US.81035		2101
MAC.US.85013		2034
MAC.US.87082	Y	2071
MAC.US.92050		2094
MAC.US.92077		2101
MAC.US.93057	G	2099
MAC.US.93062	Y Y C G K K	2086
MAC.US.95058		2088
MAC.US.95086		1957
MAC.US.96016		2101
MAC.US.96020		2093
MAC.US.96072		2105
MAC.US.96081		2103
MAC.US.96093	C	2101
MAC.US.96123		1990
MAC.US.96135		2086
MAC.US.97009		1969
MAC.US.97074	C	1969
MAC.US.r80025		1957
MAC.US.r90131		2100
MNE.US.-MNE027	A C G A A A G	2462
MNE.US.82.MNE.8	A C A A G	2462
SMM.SL.92.SL92B	AT A G C T AC AG G A G ATA AC AAAG G GAG T G G A A A C T G CAA C G C G A G A A	2425
SMM.US.-17EFC1		2994
SMM.US.-17EFR		2994
SMM.US.-F236_H4	A C T C C G G A AA AC A A A A GA GC A A A C A	2981
SMM.US.-H9	A C C C G G R A AA AC A A A A GA GC G A A G A C A	2467
SMM.US.-PB114_15	A T C C G A AA AC A A A A GA GC A G A C A	2761
SMM.US.-PB1A	A T C C G A AA AC A A A A GA GC A G A C A	2760
SMM.US.-PBJ_143	A C C C G G R A AA AC A A A A GA GC A G A C A	2467
SMM.US.-PBJ_6P6	A T C C G A AA AC A A A A GA GC G A A G A C A	2953
SMM.US.-PGM53	A C CT C C G G G A AA AC A A C A GA GC A A C A	2911
SMM.US.-SME543	A C T C C G G G A AA AC A G A A GA GC A A C A	2982
STM.US.-STM	T AT AG A A G GA A A A A A A A A A A G C A G A G A G C A	2641









MAC.US.-239  
 Pol  
 H2A.GM.MCN13  
 H2A.CI.88.UC2  
 H2A.DE.-BEN  
 H2A.DE.-PEI2  
 H2A.GH.-GH1  
 H2A.GM.-ISY  
 H2A.GM.87.D194  
 H2A.GW.-AL1  
 H2A.GW.-MDS  
 H2A.GW.86.FG  
 H2A.GW.87.CAM2C3  
 H2A.SN.85.ROD  
 H2AB.CI.-7312A  
 H2B.CI.-EHO  
 H2B.CI.88.UC1  
 H2B.GH.86.D205  
 H2B.JP01.KR020  
 H2G.CI.-ABT96  
 H2U.FR.96.12034  
 MAC.US.-251.1A11  
 MAC.US.-251.32H.PI5  
 MAC.US.-251.BK28  
 MAC.US.-MM142  
 MAC.US.-SMM142B  
 MAC.US.1937  
 MAC.US.2065  
 MAC.US.239.95112  
 MAC.US.239.96114  
 MAC.US.80035  
 MAC.US.81035  
 MAC.US.85013  
 MAC.US.87082  
 MAC.US.92050  
 MAC.US.92077  
 MAC.US.93057  
 MAC.US.93062  
 MAC.US.95058  
 MAC.US.95086  
 MAC.US.96016  
 MAC.US.96020  
 MAC.US.96072  
 MAC.US.96081  
 MAC.US.96093  
 MAC.US.96123  
 MAC.US.96135  
 MAC.US.97009  
 MAC.US.97074  
 MAC.US.r80025  
 MAC.US.r90131  
 MNE.US.-MNE027  
 MNE.US.82.MNE.8  
 SMM.SL.92.SL92B  
 SMM.US.-17EC1  
 SMM.US.-17EFR  
 SMM.US.-F236\_H4  
 SMM.US.-H9  
 SMM.US.-PB114\_15  
 SMM.US.-PB1A  
 SMM.US.-PBJ\_143  
 SMM.US.-PBJ\_6P6  
 SMM.US.-PGM53  
 SMM.US.-SME543  
 STM.US.-STM

AGAGAAATCCAAAAATGCCCATTTCAATGGATGGGGTACGAATTGTGGCCAAACAAAATGGAAGTTGCCAAAGATAGAGTTGCCACAAAGAGAGACTGGACAGTGAATGATATACAGAAAGTGTAGGAGTATTAATTTGGCAGCTCAAAATTTATCCAGGTATAA 3674  
 E E K F Q K D P P F Q W M G Y E L W P T K W K L Q K I E L P O R R E T W T V N D I Q K L V G V L N W A A Q I Y P G  
 T---G-----G---C-T--AC-C-----C-T-C-----T-G---C---G---C---C---A---TGTA-----A---C---C---A---G---T---C---C---A---C---C---G--- 3202  
 G---G-----C-T-G-----C-T-C-----T-G---C---G---AC-C-AC---C-G-A---A---A---C---C---A---AC---G---A---C---G---A---C---T---A--- 3203  
 T---G-----C---ACA-----C-T-C-A-----C---C---A---C---A---C---A---AGTA-----C---C---A---C---C---A---C---T---CC-----A---C---C---G--- 3204  
 T---G-----C-T-G-----C-T-C-A-----T---C---G---AT-C-AC---C-G-A---A---TA-----C---C---C---A---C---G---G---G---A---C---C---A--- 3205  
 C-A-G-----G-C-T-AC-----C-T-C-----C---A---A---C---A---C---AG-AGTA-----T---C---C---AC---G---T---CC-----C---G---A---C---C---A--- 3206  
 T---G-----G-C-T-G-----C-T-C-----T---AC-G-A---C---A---A---T---G-A---A---TA-----C---C---C---A---AC---G---T---G---C---G---G---G--- 3207  
 T---G-----C-T-ACA-----C-T-G-----T---C---G---A---C---A---C---G-A---AGTA-----C---C---A---ACAT---G---T---CC-----A---C---C---A--- 3208  
 T-A-G-----C-T-AC-GC-----C-T-C-----T---C---G---A---C---A---C---A---AGTA-----C---C---A---AC---G---T---CC-----A---C---C---A--- 3209  
 T---G-----G-C-T-A-GC-----C-T-A-----T---C---A---C---C---C---A---AGTA-----C---C---C---A---C---G---T---CC-----A---C---C---G--- 3210  
 H2A.GW.87.CAM2C3  
 T---G-----C-T-AC-GC-----C-T-C-A-----T---G---A---C---A---C---AG-GTG-----C---C---A---C---G---T---CC-----A---C---C---G--- 3211  
 H2A.SN.85.ROD  
 T---G-----C-T-AC-C-----C-T-C-A-----T---G---A---C---A---C---A---A---TA-----C---C---C---A---C---G---T---CC-----A---C---C---G--- 3212  
 H2AB.CI.-7312A  
 -A-----G-C-T-C-G-----T-T-GC-C---AG-G---C---A---C---AA-A---G---G-GTT-----G-C-C-T---AC-G---G-----C---TC-C-A-C--- 3213  
 H2B.CI.-EHO  
 -A-G-----C-T-CA-----T-T-GC-C---AG-AC---A---C---C-A---G---A---GTT-----A-C-T---G-----G-----C---TC-G-G-T--- 3214  
 H2B.CI.88.UC1  
 -A-G-----C-T-G-CA-----T-T-GC-C---A-GG-AC---A---C---AC---G---AG-GTT-----G-C-T-A-AC---G-----C-----C---T-T-A-T--- 3215  
 H2B.GH.86.D205  
 -A-G-----C-T-G-CA-----T-T-GC-C---A-G-AC---A---C---AC---G---A---AGTT-----CA-T-A-AC---G-----C-----C---T-T-A-T--- 3216  
 H2B.JP01.KR020  
 -A-G-----G-C-----CA-----T-T-GC-C---A-G-AC---A---C---C---G---A---GA-----A-----T---AC-G-----C-----C---T-T-G-T--- 3217  
 H2G.CI.-ABT96  
 T-A-G-----C-T-A-----A-T-R-----A-G-AC-A-----TA-CC-MG-A---T-----A-----T-A-AC---G---C-G-Y-----GC-A---A-C--- 3218  
 H2U.FR.96.12034  
 C-A-----T-G---C-A-CACA-----C-T-C-A---T-A-G---A---G-A-T---C-C-G-----TGTT-----A-C-T-----C-----A-G-----A--- 3219  
 MAC.US.-251.1A11  
 -A-----G-----G-----G-----R----- 3220  
 MAC.US.-251.32H.PI5  
 -A-----G-----G-----R----- 3221  
 MAC.US.-251.BK28  
 -A-----G-----G-----R----- 3222  
 MAC.US.-MM142  
 -A-----G-----G-----R----- 3223  
 MAC.US.-SMM142B  
 -A-----G-----G-----R----- 3224  
 MAC.US.1937  
 -A-----G-----G-----R----- 3225  
 MAC.US.2065  
 -A-----G-----G-----R----- 3226  
 MAC.US.239.95112  
 -A-----G-----G-----R----- 3227  
 MAC.US.239.96114  
 -A-----G-----G-----R----- 3228  
 MAC.US.80035  
 -A-----G-----G-----R----- 3229  
 MAC.US.81035  
 -A-----G-----G-----R----- 3230  
 MAC.US.85013  
 -A-----G-----G-----R----- 3231  
 MAC.US.87082  
 -A-----G-----G-----R----- 3232  
 MAC.US.92050  
 -A-----G-----G-----R----- 3233  
 MAC.US.92077  
 -A-----G-----G-----R----- 3234  
 MAC.US.93057  
 -A-----G-----G-----R----- 3235  
 MAC.US.93062  
 -A-----G-----G-----R----- 3236  
 MAC.US.95058  
 -A-----G-----G-----R----- 3237  
 MAC.US.95086  
 -A-----G-----G-----R----- 3238  
 MAC.US.96016  
 -A-----G-----G-----R----- 3239  
 MAC.US.96020  
 -A-----G-----G-----R----- 3240  
 MAC.US.96072  
 -A-----G-----G-----R----- 3241  
 MAC.US.96081  
 -A-----G-----G-----R----- 3242  
 MAC.US.96093  
 -A-----G-----G-----R----- 3243  
 MAC.US.96123  
 -A-----G-----G-----R----- 3244  
 MAC.US.96135  
 -A-----G-----G-----R----- 3245  
 MAC.US.97009  
 -A-----G-----G-----R----- 3246  
 MAC.US.97074  
 -A-----G-----G-----R----- 3247  
 MAC.US.r80025  
 -A-----G-----G-----R----- 3248  
 MAC.US.r90131  
 -A-----G-----G-----R----- 3249  
 MNE.US.-MNE027  
 -G-----T-----A-----A----- 3250  
 MNE.US.82.MNE.8  
 -G-----T-----A-----A----- 3251  
 SMM.SL.92.SL92B  
 -C-G-T---GA---A---A-G---A-TTTG-A-AC-C-AG-G---A-A-G-A---AC-C-TG---AG---CGA-----A---C---A-AC---G---TC-----A-G-A---G-T--- 3252  
 SMM.US.-17EC1  
 -A-----G-----G-----R----- 3253  
 SMM.US.-17EFR  
 -A-----G-----G-----R----- 3254  
 SMM.US.-F236\_H4  
 -G---G---C-G---A-T---C---AC-G-A---A---A---A---A---A---GC---A---A--- 3255  
 SMM.US.-H9  
 -G---G---C-G---A-T---C---AC-G-A---A---A---A---A---A---GC---A---A--- 3256  
 SMM.US.-PB114\_15  
 -G---G---C-G---A-T---C---AC-G-A---A---A---A---A---A---GC---A---A--- 3257  
 SMM.US.-PB1A  
 -G---G---C-G---A-T---C---AC-G-A---A---A---A---A---A---GC---A---A--- 3258  
 SMM.US.-PBJ\_143  
 -G---G---C-G---A-T---C---AC-G-A---A---A---A---A---A---GC---A---A--- 3259  
 SMM.US.-PBJ\_6P6  
 -G---G---C-G---A-T---C---AC-G-A---A---A---A---A---A---GC---A---A--- 3260  
 SMM.US.-PGM53  
 -G---G---C-G---A-T---C---AC-G-A---A---A---A---A---A---GC---A---A--- 3261  
 SMM.US.-SME543  
 -G---G---C-G---A-T---C---AC-G-A---A---A---A---A---A---GC---A---A--- 3262  
 STM.US.-STM  
 T-G-G---T-----A-T-G-A-----G-AC---A---A-A-T-G---CGTT-----C-A-AC-G---G-C---C-----A---A-T--- 3263

MAC.US.-239	AAACCAAACATCTCTGTAGGTTAATTAGAGGAAAAATGACTCTAACAGAGGAAGTTCAGTGGACTGAGATGGCAGAAGCAGAATATGAGGAAAAATAAATAATTTCTCAGTCAGGAACAAGAAGGATGTTATACCAAGGCAAGCCATTAGAAGCCACGGTAAATAAG	3844
Pol	I K T K H L C R L I R G K M T L T E E V Q W T E M A E A E Y E E N K I I L S Q E Q E G C Y Y Q E G K P L E A T V I	
H2A.GM.MCN13	---G---CT-A---A-C-----A-C---A---A---A---AC-A---G-G---G-TA-A-G-C-G-T---CT-A-C---G---CAC-----AA-AGAG-----A-A---CCA---A	3372
H2A.CI.88.LUC2	---G---CT-G---C-----A-C---A---G---A---T-A---G---GCTA-A-G-C-G-T---C-AGAC-----G---CAC-----AG-AGA---G-A---A-CCA---A	3927
H2A.DE.-BEN	---G---CT-A---AC-----A-C---A---G---A---AC-A---G---GCTA-A-G-C-T---CT-G-C-----G---G---AA-AGA---G-A---AA-CCA---A	3927
H2A.DE.-PEI2	-G---CT-A---C-----A-C-G-A-A---A---AC-A---G---G---CTA-A-G-C-T---CT-G-C-----G-G---C-----AA-GA-----A-A---CCA---A	3895
H2A.GH.-GH1	---G---T-G---C---C-A---A-C---A---G---A---T-A---G---GCTA-A-C---C-T---CT-A-C-----G---CAC-----AA-AGA---G-A---AA-CCA---A	3371
H2A.GM.-JISY	-G---CT-A---A-C-----G-A-CC-G-A-A---A---AT-A---G---GCTG-----C-T---CT-A-C-----G-G---CAC-----G-AA-AGAG-----A-A---TCA---A	3368
H2A.GM.87.D194	---G---G---T-A---AA-G-----A-C---A---G---A---T-A---G---CTA-A---G---C-T---CT-A-----G---CC-C-T-G---AAG-AGA-C-----A-A---C---A	3371
H2A.GW.-AL1	---G---G---CT-A---AC-----A-C---A---G---A---AC-A---G---CTA---G---C-G-T---CT-A-----G---GCAC-----AA-6AG-----A-A---CCA---A	3920
H2A.GW.-MDS	-G-A---T-A---AC-----A-C-G-A-A---A---A---A---G---CTA-A---C-G-G-T---CT-G-C-----G-CAC-----G-AA-AGA-----A-A---CCA---A	3371
H2A.GW.86.FG	-G---C-A---C-----A-C---A---G-A-A---A---AT-A---G---GCTA---C-G-T---CT-A-C---A---CAC-----G---AA-AAAG-----A-A---CCAG---A	3365
H2A.GW.87.CAM2CG	-G---CT-G---A---A---A-C---A---G-A-A---A---AT-A---G---CTA-A---C-G-T---CT-A-C---A---G-CAC-C---G---AA-GAG-----A-A---CCA---A	3936
H2A.SN.85.ROD	-G---CT-A---C-----A-C---A---A---A---AT-A---G---GCTA-A---C-G-T---CT-A-C---A---G-CAC-----AA-AGAG-----A-A---CCA---A	3373
H2AB.CI.-7312A	---A-GG-CA-A-C-AC-G-C-AG---G-C-G---G-G---A-AC-A---G-GATGC-----C-C-AG-A---G---C-C-A---GGTA-TC-----A-A---GCAG---A	3909
H2B.CI.-EHO	-G---GG---A-A---AAC---A-G---G-C---A-G-A-A---A---AT---G---TCC---C---C-AGAA---A-G-G---CC---A---GGTA-T-----A-A---GCAG---A	3898
H2B.CI.88.UC1	-G-A-GG-CA-A- AAC-----G---G-C---A-A-A---A---AT-A---G-G-TAC---G-C---C-T-AGAA-----CC-C-A-G---GGTA-C-----A-A---CAG---A	3908
H2B.GH.86.D205	-G-A-GG-CA-A-C-AC-----G---G---C---A---A---A---AC-A---G-GCTAC---G---C---CT-AGAA-----CC-C-A-G---A-GTA-GC-----A-A---CAG---A	3900
H2B.JP01.KR020	---A-GG---A-A-C-AC---C-G---CT-G---A---A---A---AT-A---G---T-C-----CT-AGAA---A---G-G---A-C-TA---G-A-TA-TC-----A-A---CAG---A	3050
H2G.CI.-ABT96	---A-A-C-G-Y-AA-----T-R---CT-G-G-A---A---A---AT---TGCA---A---C-G-T---C-ARAA---R-G---CGCC---A---RAG---T-R-----T---CAG---A	3254
H2U.FR.96.12034	-G-A-----T---AA-----AG-G---A-C---A-G-A---A---A---T-CA-G---G-T---T-G---G-GGC---AGG---AAG-AAATC-----A-A---GC-T---A	3409
MAC.US.-251_1A11	-----C-----T-----	3843
MAC.US.-251_32H_P15	-----T-----	3845
MAC.US.-251_BK28	-----G-----	3820
MAC.US.-MM142	-----A-----	3330
MAC.US.-SMM142B	-----A-----	3330
MAC.US.1937	-----	2950
MAC.US.2065	-----	2950
MAC.US.239_95112	-----	2940
MAC.US.239_96114	-----	2946
MAC.US.80035	-----	2950
MAC.US.81035	-----	2951
MAC.US.85013	-----	2884
MAC.US.87082	-----	2921
MAC.US.92050	-----	2944
MAC.US.92077	-----	2951
MAC.US.93057	-----	2949
MAC.US.93062	-----	2936
MAC.US.95058	-----	2938
MAC.US.95086	-----	2807
MAC.US.96016	-----	2951
MAC.US.96020	-----	2943
MAC.US.96072	-----	2955
MAC.US.96081	-----	2953
MAC.US.96093	-----	2951
MAC.US.96123	-----	2840
MAC.US.96135	-----	2936
MAC.US.97009	-----	2819
MAC.US.97074	-----	2819
MAC.US.r80025	-----	2807
MAC.US.r90131	-----	2950
MNE.US.-MNE027	-----G-----	3312
MNE.US.82.MNE_8	-----G-----	3312
SMM.SL.92.SL92B	---A-A-C-G-C-AAA-G-C---G---G---A---A---AT-A---TG-CA-G---G-T---A-A-A-G-G---A-A---AG---ATG-A-GC-----A---TC-----	3275
SMM.US.-17EC1	-----	3844
SMM.US.-17EFR	-----	3844
SMM.US.-F236_H4	-G-T-----T-C-AC---C-----T---A-G-----G---A---C-G-----A---C-G-----C---A-A---A---G-A-A-----	3831
SMM.US.-H9	-G-T-----T-C-A---C-----T---A-G-----G---A---C-G-----C---A---C-G-----C-T-G-G-A---C---G-AC-A-----	3317
SMM.US.-PB114_15	-G-T-----T-C-A---C-----T---A-G-----G---A---C-G-----C---A---C-G-----C-T-G-G-A---C---G-A-T-----	3611
SMM.US.-PB1A	-G-T-----T-C-A---C-----T---A-G-----G---A---C-G-----C---A---C-G-----C-T-G-G-A---C---G-A-A-----	3610
SMM.US.-PB1_143	-G-T-----T-C-A---C-----T---A-G-----G---A---C-G-----C---A---C-G-----C-T-G-G-A---C---G-AC-A-----	3317
SMM.US.-PB1_6P6	-G-T-----T-C-A---C-----T---A-G-----G---A---C-G-----C---A---C-G-----C-T-G-G-A---C---G-A-A-----	3803
SMM.US.-PGM53	-G-T-----T-C-AA---C-----T---A-G-----G---A---C-G-----C---A---G-G-----C---G-A---C---G-A-A-----	3761
SMM.US.-SME543	-G-T-----T-C-AC---C-----T---A-G-----G---A---C-G-----C---A---C-G-----C---G-A-A---A---G-A-A-----	3832
STM.US.-STM	---T-----G-C-AA---C---T-G---G-A-G---T---G-A---A---A---G---CA---C---T-----G-G---GCAA-C---G---A---TC---G-A-A---G-G---A	3491







MAC.US.-239
H2A.GM.MCN13
H2A.CI.88.LUC2
H2A.DE.-BEN
H2A.DE.-PEI2
H2A.GH.-GH1
H2A.GM.-ISY
H2A.GM.87.D194
H2A.GW.-AL1
H2A.GW.-MDS
H2A.GW.86.FG
H2A.GW.87.CAM2C6
H2A.SN.85.ROD
H2AB.CI.-7312A
H2B.CI.-EHO
H2B.CI.88.UC1
H2B.GH.86.D205
H2B.JP01.KR020
H2G.CI.-ABT96
H2U.FR.96.12034
MAC.US.-251.1A11
MAC.US.-251.32H.1P5
MAC.US.-251.BK28
MAC.US.-MM142
MAC.US.-SMM142B
MAC.US.1937
MAC.US.2065
MAC.US.239.95112
MAC.US.239.96114
MAC.US.80035
MAC.US.81035
MAC.US.85013
MAC.US.87082
MAC.US.92050
MAC.US.92077
MAC.US.93057
MAC.US.93062
MAC.US.95058
MAC.US.95086
MAC.US.96016
MAC.US.96020
MAC.US.96072
MAC.US.96081
MAC.US.96093
MAC.US.96123
MAC.US.96135
MAC.US.97009
MAC.US.97074
MAC.US.r80025
MAC.US.r90131
MNE.US.-MNE027
MNE.US.82.MNE.8
SMM.SL.92.SL92B
SMM.US.-17EC1
SMM.US.-17EFR
SMM.US.-F236\_H4
SMM.US.-H9
SMM.US.-PB114\_15
SMM.US.-PB1A
SMM.US.-PBJ\_143
SMM.US.-PBJ\_6P6
SMM.US.-PGM53
SMM.US.-SME543
STM.US.-STM

Table with 48 columns representing amino acid positions (S, C, N, K, Q, S, K, E, G, K, A, G, Y, I, T, D, R, G, K, D, K, V, K, V, L, E, Q, T, T, N, Q, Q, A, E, L, E, A, F, L, M, A, L, T, D, S, G, P, K, A, N, I, I, V, D, S, Q, Y) and rows for various HIV/SIV sequences. The table shows sequence alignments with gaps and specific amino acid residues.

MAC.US.-.239	TATGGGAATAATAACAGGATGCCCTACAGAATCAGAGAGCAGGCTAGTTAATCAAAATAAGAAGAAATGATTAAGAAAGTCAGAAATTTATGTAGCATGGGTACCAGCACACAAGGTATAGGAGGAAACCAAGAAATAGACCACCTAGTTAGTCAAGGGATTAGACAAG	4523
Pol	V M G I I T G C P T E S E S R L V N Q I I E E M I K K S E I Y V A W V P A H K G I G G N Q E I D H L V S Q G I R	
H2A.GM.MCN13	A---G---G---G---CCAG-A-----T-AA-A-----C---G-G---A-----GA-C---C-----T---C---C---T---C-----G-T-G---GG---T---GT---A---G---C---C-----	4051
H2A.CI.88.LUC2	A---G---G---G---A---CCAA-A---G---T---AA-A---A---C---C---T---C---A---G---AGA---C---G---C-----T---A---C---C---T---C-----G---G---TT---A---G---C---C-----	4006
H2A.DE.-.BEN	A---G---G---G---CCAG-A---G---A---AT---AA-A---G---C---G---C-----A---G---GA---C---G---C-----T---C---C---T---C-----G---G---TT---A---G---C---C-----	4006
H2A.DE.-.PEI2	A---G---G---G---CCAG-A-----T-AA-A---A---C---C-----A---G---GA---C---C---C-----T---C---C---T---C-----T---G---G---T---TT---A---G---C---C-----	4574
H2A.GH.-.GH1	A---G---G---GT---CCAA-A---G---A---T---AA-A---A---C---C---T---G---C---A---G---AGA---C---G---C---C---G---C---C---T---C-----G---G---TT---A---G---C---C-----	4050
H2A.GM.-.ISY	A---G---G---CCAA---GG-T-----T-AA-A---A---A---T---G---A---G---GA---C---C---C-----T---C---G---C---C-----T---G---G---T---A---G---C---C-----	4047
H2A.GM.87.D194	A---G---G---G---CCAA-A---G---A---AT---AA-A---A---C---C---T---G---C---A---G---AGA---C---G---C---T---C---C---T---C-----T---G---G---TT---A---G---C---C-----	4050
H2A.GW.-.ALI	A---G---G---G---CCAG-A---G---A---AT---AA-A---A---C---C---T---G---C---A---G---GA---C---C---C-----T---G---G---TT---A---G---C---C-----	4599
H2A.GW.-.MDS	A---G---GGT---CCAA-A---G---A---AT---AA-A---A---G---C---G---A---G---GA---C---C---C-----T---G---G---T---T---A---G---C---C-----	4050
H2A.GW.86.FG	A---G---G---G---CCAA-A---G---A---AT---AA-A---A---G---C---G---A---G---AGA---C---C---C---T---G---C---C---C-----G---G---T---TT---A---G---C---C-----	4044
H2A.GW.87.CAM2CQ	A---G---G---GG---CCAA-A---G---A---AT---AA-A---A---C---C-----A---G---GA---C---C---T---C---C---C---T---C-----G---G---T---TT---A---G---C---C-----	4615
H2A.SN.85.ROD	A---G---C---GTG---A---CCAA-A---G---A---T---AAA---G---C---G---C---A---GA---C---C---T---C---C---C---T---C-----G---G---T---TT---A---G---C---C-----	4052
H2AB.CI.-.7312A	C---G---G---T---CAG-A---A---A---TCACC-A---A---A---G---T---C---G---G---T---C---G---G---T---AC---T---T---G---G---G---A---A---G---G---G---A---G---G---G---	4588
H2B.CI.-.EHO	C---G---G---T---C---CAA-A---A---A---ATCACC-A---A---A---GAG---T---C---G---G---A---A---A---G---T---G---AC---G---T---T---T---G---G---G---A---G---G---A---G---G---A---G---G---	4577
H2B.CI.88.UC1	C---G---G---T---CAG-A---A---A---G---TCACCA-A---A---A---G---T---C---G---G---A---A---G---T---G---AC---T---T---T---G---G---G---A---A---A---G---G---A---G---G---	4587
H2B.GH.86.D205	C---G---G---T---C---CAG-A---A---A---ATCACC-AA---AGCAA---T---C---G---G---A---A---G---T---G---AC---G---T---T---T---G---G---G---A---A---A---C---G---	4579
H2B.JP01.KR020	C---G---T---C---CAG-A---A---A---ATCACC-AA---A---GAC---G---T---C---G---G---A---A---C---G---G---T---T---AC---G---T---T---G---G---A---C---CA---3729	3729
H2G.CI.-.ABT96	C---R---C---G---RCAA-A---G---A---ATC---CC-T---A---A---T---R---AGARCTC-A---G---G---T---T---G---A---A---T---T---G---G---RT---A---C---G---A---G---A---G---A---G---	3933
H2U.FR.96.12034	A---G---G---G---CAG-A---A---A---TC---CCA---A---G---A---C---A---C---T---A---GG---AGA---C---G---G---T---GC---T---T---T---G---A---T---T---C---T---G---T---TT---G---G---G---A---G---G---	4085
MAC.US.-.251.1A11	-----G-----C-----A-----C-----A-----C---T---A---GG---AGA---C---G---G---T---GC-----T---T---T---G---A---T---T---C---T---G---T---TT---G---G---G---A---G---G---	4522
MAC.US.-.251.32H.P15	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	4524
MAC.US.-.251.BK28	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	4499
MAC.US.-.MM142	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	4009
MAC.US.-.SMM142B	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	4009
MAC.US.1937	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3629
MAC.US.2065	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3629
MAC.US.239.95112	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3619
MAC.US.239.96114	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3625
MAC.US.80035	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3629
MAC.US.81035	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3630
MAC.US.85013	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3563
MAC.US.87082	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3600
MAC.US.92050	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3623
MAC.US.92077	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3630
MAC.US.93057	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3628
MAC.US.93062	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3615
MAC.US.95058	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3617
MAC.US.95086	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3486
MAC.US.96016	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3630
MAC.US.96020	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3622
MAC.US.96072	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3634
MAC.US.96081	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3632
MAC.US.96093	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3630
MAC.US.96123	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3519
MAC.US.96135	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3615
MAC.US.97009	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3498
MAC.US.97074	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3498
MAC.US.r80025	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3486
MAC.US.r90131	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3629
MNE.US.-.MNE027	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3991
MNE.US.82.MNE_8	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	3991
SMM.SL.92.SL92B	CT---AAC-----CAG-AT---G---TTCAGATA---GGCA---T-T---C---C-AG-C-G-AGA-C-G-A-CA---G---T---T-T---C-----G---A-T-GG---GTT-----A-A---G---3954	3954
SMM.US.-.17EC1	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	4523
SMM.US.-.17EFR	-----G-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----A-----C-----	4523
SMM.US.-.F236_H4	C---G---G---TCAG-C-T---A---T---AT---A---C---G---G---GA---C---C---T---T---A---C---T---G---T---G---C---G---A---4510	4510
SMM.US.-.H9	C---G---G---TCAG-C-T---A---T---AT---A---C---G---G---GA---C---C---T---T---A---C---T---G---T---G---C---G---A---4510	4510
SMM.US.-.PB114_15	C---G---G---TCAG-C-T---A---T---AT---A---C---G---G---GA---C---C---T---T---A---C---T---G---T---G---C---G---A---4510	3996
SMM.US.-.PB1A	C---G---G---TCAG-C-T---A---T---AT---A---C---G---G---GA---C---C---T---T---A---C---T---G---T---G---C---G---A---4510	4290
SMM.US.-.PB1_143	C---G---G---TCAG-C-T---A---T---AT---A---C---G---G---GA---C---C---T---T---A---C---T---G---T---G---C---G---A---4510	4289
SMM.US.-.PB1_6P6	C---G---G---TCAG-C-T---A---T---AT---A---C---G---G---GA---C---C---T---T---A---C---T---G---T---G---C---G---A---4510	3996
SMM.US.-.PGM53	C---G---G---TCAG-C-T---A---T---AT---A---C---G---G---GA---C---C---T---T---A---C---T---G---T---G---C---G---A---4510	4482
SMM.US.-.SME543	C---G---G---TCAG-C-T---A---T---AT---A---C---G---G---GA---C---C---T---T---A---C---T---G---T---G---C---G---A---4510	4440
STM.US.-.STM	C---G---G---TCAG-C-T---A---T---AT---A---C---G---G---GA---C---C---T---T---A---C---T---G---T---G---C---G---A---4510	4511
	G-----C-----CAG-C-----G---A---T---A---C---G-----G-----GA-C-----T---T---A-----T---G---GG---T---TT---G---A---G---A-----G---4170	4170

Pol RNase H end\_Pol p31 Integrase start

Table of HIV-2/SIV sequence alignments. Columns include accession numbers (e.g., MAC.US.-239, H2A.GM.MCN13), amino acid positions (Q-V-L-F-#-L-E-K-I-E-P-A-Q-E-E-H-D-K-Y-H-S-N-V-K-E-L-V-F-K-F-G-L-P-R-I-V-A-R-Q-I-V-D-T-C-D-K-C-H-Q-K-G-E-A-I-H-G-Q-A), and sequence characters (A, C, G, T, -, R, Y). A column of alignment positions (4692, 4220, 4775, 4773, 4743, 4219, 4216, 4219, 4768, 4219, 4213, 4784, 4221, 4757, 4746, 4756, 4748, 3898, 4102, 4254, 4691, 4683, 4668, 4178, 4178, 3798, 3798, 3788, 3794, 3798, 3799, 3732, 3769, 3792, 3799, 3797, 3784, 3786, 3655, 3799, 3791, 3803, 3801, 3799, 3688, 3784, 3667, 3667, 3655, 3798, 4160, 4160, 4123, 4692, 4680, 4165, 4459, 4458, 4165, 4651, 4609, 4680, 4339) is on the right.

MAC.US.-239	TTTCAGATCTAGGGACTTGGCAAATGGATTGTACCCATCTAGAGGGAAAAATAATCATAGTTGCAGTACATGTAGCTAGTGGATTATAGAAAGCAGAGGTAATTCACAAAGAGACAGGAAGACAGACAGCACTATTCTGTAAAATTGGCAGGCAGATGGCCTATTACAC	4862
Pol	N S D L G T W Q M D C T H L E G K I I T I V A V H V A S G F I E A E V I P Q E T G R Q T A L F L L K L A G R W P T	
H2A.GM.MCN13	G--AT--C-----C-C-A-CT--A--G-C--T--A-----T--A-----T--A-----A--C--C--G--AT--G-----C--C--A-----C--CA-T--G--A--A--	4390
H2A.CI.88.UC2	G--AA--CGT-----C-C-A-CT--A--G-C--T--A-----T--A-----T--A-----A--C--C--G--AT--G-----C--C--A-----C--CA-T--G--A--A--	4945
H2A.DE.-BEN	G--AA--CGT-----C-AC-A-CT--A--G-C--T--A-----T--A-----T--A-----A--C--C--G--AT--G-----C--C--A-----C--CA-T--G--A--A--	4945
H2A.DE.-PEI2	G--A-G-C-----C-C-A-CT--A--G-C--T--A-----T--A--C-G-T-----A--T--C--G--A--G--G--A-----C--C--A-----C--CA-T--G--A--A--	4913
H2A.GH.-GH1	G--AA--TGTC-----C-C-A-CT--A--G-C--T--A-----T--A-----T--A-----A--T--C--G--AT--G-----G--C--C--A-----C--CA-T--G--A--A--	4389
H2A.GM.-ISY	G--A-C-C-----C-C-A-CT--A--G-C--T--A-----T--A-----T--A-----A--C--C--G--AT--G-----G--A-----C--C--A-----C--CA-T--G--A--A--	4386
H2A.GM.87.D194	G--AA--CGT-----C-C-A-CT--A--G-C--T--A-----G--T--A-----T--A-----A--C--C--G--AT--G-----G--A-----C--C--A-----C--CA-T--G--A--A--	4389
H2A.GW.-AL1	G--A-C-----C-C-A-CT--A--G-C--T--A--G-----T--A-----T--A-----A--C--C--G--AT--G-----G--A-----C--C--A-----C--CA-T--G--A--A--	4938
H2A.GW.-MDS	G--CT--C-----C-C-T--T--A--G-C--T--A--G-----T--A-----T--A-----A--C--C--G--AT--G-----G--A-----C--C--A-----C--TA-T--G--A--A--	4389
H2A.GW.86.FG	G--A-C-----C-C-A-T--A--G-C--T--A-----T--A-----T--A-----G--A--T--C--G--AT--G-----G--A-----G--CT-AC-G--C--CA-T--G--A--A--	4383
H2A.GW.87.CAM2CG	G--AG-G-C-C-----C-C-A-T--A--G-C--T--A-----T--A-----T--A-----A--C--C--G--AT--G-----G--A-----C--C--A--G--C--TA--G--A--A--	4954
H2A.SN.85.ROD	G--A-C-----C-C-A-T--A--G-C--T--A-----T--A-----T--A-----A--C--C--G--AT--G-----G--A-----C--C--A--G--C--A--T--G--A--A--	4391
H2AB.CI.-7312A	G--A-A-A-----C-A-T--A--G-G-C-A-A-A-----C-----C-G-T-----A--C-A-A-----A--T--C--C-----GC-A-----C-----	4927
H2B.CI.-EHO	A-A-A-----C-A-T--A--GG-T-A-A-----G--T-----C-----C-----T--G-----A--A--C--A-----A--T--C--C-----GC-CA-----C-----	4916
H2B.CI.88.UC1	G--A-A-A-----G-C-A-CT--A--GG-C-A-A-----G--T-----C-----C-----T--G-----A--A--C--A-----A--T--C--C-----CA-CA-----C-----	4926
H2B.GH.86.D205	G--C-A-A-----G-C-A-CT--A--T--A-----G--C-----C-----G--T-----A--A--C--A-----A--T--C--C--AC--G--C--CA-----C-----	4918
H2B.JP01.KR020	G--A-A-----C-C-A-CT--A--T--A-----G--C-----C-----T--C-----T--C-----A--A--C--A-----T--C--C-----CA-CA-----C-----	4068
H2G.CI.-ABT96	G--AT-G-T-----C-C-T--A--G-----T--A-----T--G--C-----C--TG-----A--C-----A--G--A-----T--C-----C--A--G-----C-----	4272
H2U.FR.96.12034	A--AG--A-----C-----CT--G-TG--C-G-----G--C-----C--T-----A--G--C--G--G--A--T--T--G--A-----C--G--C--AC--GC-A--A--G-----C--T-----	4424
MAC.US.-251_1A11	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	4861
MAC.US.-251_32H_P15	-----C-----A-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3863
MAC.US.-251_BK28	-----C-----A-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	4838
MAC.US.-MM142	-----C-----A-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	4348
MAC.US.-SMM142B	-----C-----A-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	4348
MAC.US.1937	-----Y-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3968
MAC.US.2065	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3968
MAC.US.239.95112	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3958
MAC.US.239.96114	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3964
MAC.US.80035	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3968
MAC.US.81035	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3969
MAC.US.85013	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3902
MAC.US.87082	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3939
MAC.US.92050	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3962
MAC.US.92077	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3969
MAC.US.93057	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3967
MAC.US.93062	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3954
MAC.US.95058	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3956
MAC.US.95086	-----C-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3825
MAC.US.96016	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3969
MAC.US.96020	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3961
MAC.US.96072	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3973
MAC.US.96081	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3971
MAC.US.96093	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3969
MAC.US.96123	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3858
MAC.US.96135	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3954
MAC.US.97009	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3837
MAC.US.97074	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3837
MAC.US.r80025	-----C-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3825
MAC.US.r90131	-----C-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	3968
MNE.US.-MNE027	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	4130
MNE.US.82.MNE_8	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	4330
SMM.SL.92.SL92B	G--GG--C-A--G--C-C--CT--A--GG-T--A--G--C--T--A--TAGG-A--T-G--C--C--A-A--TCA--A-AG--	4293
SMM.US.-17EC1	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	4862
SMM.US.-17EFR	-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	4862
SMM.US.-F236_H4	G--A-----C--G--C--A--C--T-----G-----T--G--A--C--G--G--A--A--G--G--A--T-----C--C-----	4850
SMM.US.-H9	G--R-RA-----C--A--C--C--T-----G--G-----G--A--C--G--RG--A--A--G--C--A--CA-----C--C-----	4335
SMM.US.-PB114_15	G--A-----C--A--C--C--T-----G-----G--A--C--G--G--A--A--G--C--A--CA-----C--C-----	4629
SMM.US.-PB1A	G--A-----C--A--C--C--T-----G-----G--A--C--G--G--A--A--G--C--A--CA-----C--C-----	4628
SMM.US.-PB1_143	G--R-RA-----C--A--C--C--T-----G--G-----G--A--C--GRGG--A--A--G--C--A--CA-----C--C-----	4335
SMM.US.-PB1_6P6	G--A-----C--A--C--C--T-----G-----G--A--C--G--G--A--A--G--C--A--CA-----C--C-----	4821
SMM.US.-PGM53	G--A-----C--A--C--C-----A-----G-----G--A--C--G--G--A--G--A--G--C--A--TA-----C--C-----	4779
SMM.US.-SME543	G--A-----C--G--C--A--C--T-----G-----T--G--A--C--G--G--A--A--G--G--GC-A--T-----C--C-----	4850
STM.US.-STM	G--GT-----G--C--A--A--C--G--T--A-----G--T--G--A--C--G--G--A--G-----T--G--G--G--A--A-----G--C-----	4509

MAC.US.-239 Pol	ATCTACACACAGATAATGGTCTA	ACTTTGCTTCGCAAGAAGTAAAGATGGTTGCATGGTGGGCGAGGATAGAGCACACCTTTGGGGTACCATAACAATCCACAGAGTCAGGGAGTAGTGGAAAGCAATGAATCACCACCTGAAAAATCAAATAGATAGAATCAGGGAAACA	5032
	T H L H T D N G A N F A S Q E V K M V A W W A G I E H T F G V P Y N P Q S Q G V V E A M N H L K N Q I D R I R E		
H2A.GM.MCN13	CT-G	CA A-G A-A	4560
H2A.CI.88.UC2	CT-G	CA A-G A-A	4560
H2A.DE.-BEN	CT-G	CC C CA A-G G-A	5115
H2A.DE.-PEI2	C-G	C CA A-G G-A	5113
H2A.GH.-GH1	CT-G	CA A-G G-A	5083
H2A.GM.-ISY	CT-G	CA A-G G-A	4559
H2A.GM.87.D194	CT-G	CC C CA A-G G-A	4559
H2A.GW.-AL1	CT-G	G TC CA C A-G G-A	5108
H2A.GW.-MDS	CT-G	C CA A-G G-A	4559
H2A.GW.86.FG	C-G	C CA A-G G-A	4553
H2A.GW.87.CAM2CB	CT-G	T CA A-G G-A	5124
H2A.SN.85.ROD	CT-G	T CA A-G G-A	4561
H2AB.CI.-7312A	CT-G	C CA A A-T G-T	5097
H2B.CI.-EHO	C-G	C C CA A T-G A-CA	5086
H2B.CI.88.UC1	C-G	C C C CA C A C-G CG C	5096
H2B.GH.86.D205	C-G	C C C CA C A C-AGT A C	5088
H2B.JP01.KR020	C-G	C C C CA A C-G A-CA	4238
H2G.CI.-ABT96	C-G	T C CA C A-G G-A C	4442
H2U.FR.96.12034	T-G	C CA A-G A-A	4594
MAC.US.-251_1A11			5031
MAC.US.-251_32H_P15		C	5023
MAC.US.-251_BK28		C	5008
MAC.US.-MM142		G	4518
MAC.US.-SMM142B		G	4518
MAC.US.1937			4138
MAC.US.2065			4138
MAC.US.239.95112			4128
MAC.US.239.96114			4134
MAC.US.80035			4138
MAC.US.81035			4139
MAC.US.85013			4072
MAC.US.87082			4109
MAC.US.92050			4132
MAC.US.92077			4139
MAC.US.93057			4137
MAC.US.93062			4124
MAC.US.95058			4126
MAC.US.95086		Y	3995
MAC.US.96016			4139
MAC.US.96020			4131
MAC.US.96072			4143
MAC.US.96081			4141
MAC.US.96093			4139
MAC.US.96123			4028
MAC.US.96135			4124
MAC.US.97009			4007
MAC.US.97074			4007
MAC.US.r80025			3995
MAC.US.r90131			4138
MNE.US.-MNE027		C A	4500
MNE.US.82.MNE_8		C	4500
SMM.SL.92.SL92B		G T C C C T CA C C G G G A TT AG G A GT A G C A A C G CTTA T T G AA C A A	4463
SMM.US.-17EC1			5032
SMM.US.-17EFR			5032
SMM.US.-F236_H4		G T T C T CA A A G A C T A G G T T C A G C T TT A CC G T A	5020
SMM.US.-H9		G T T C T CA A A G A T R T A G R T T C AR G G C T T A CC T A G 4505	4505
SMM.US.-PB114_15		G T T C T CA A A G A T T A G T T C A G G C C T T A CC T A G 4799	4799
SMM.US.-PB1A		G T T C T CA A A G A T T A G T T C A G G C C T T A CC T A G 4798	4798
SMM.US.-PBJ_143		G T T C T CA A A G A T R T A G R T T C AR G G C C T T A CC T A G 4505	4505
SMM.US.-PBJ_6P6		G T T C T CA A A G A T T A G T T C A G G C C T T A CC T A G 4991	4991
SMM.US.-PGM53		G T T C T CA A A G A C A T G C A G T T C A A C C T T A CC G T A G 4949	4949
SMM.US.-SME543		G T T C T CA A A G A C A T A G A G T T C A A C C T T A CC G T A G 5020	5020
STM.US.-STM		C G G C CA C G G A C G A A A A A T C A C C A C T T A G C C G C T A T 4679	4679



MAC.US.-.239  
Poi  
Vif  
H2A.GM.MCN13  
H2A.CI.88.UC2  
H2A.DE.-BEN  
H2A.DE.-PEI2  
H2A.GH.-GH1  
H2A.GM.-J57  
H2A.GM.87.D194  
H2A.GW.-ALI  
H2A.GW.-MDS  
H2A.GW.86.FG  
H2A.GW.87.CAM2CG  
H2A.SN.85.ROD  
H2AB.CI.-7312A  
H2B.CI.-EHO  
H2B.CI.88.UC1  
H2B.GH.86.D205  
H2B.JP01.KR020  
H2G.CI.-ABT96  
H2U.FR.96.12034  
MAC.US.-.251.A111  
MAC.US.-.251.32H.P15  
MAC.US.-.251.BK28  
MAC.US.-.SMM142  
MAC.US.-.SMM142B  
MAC.US.1937  
MAC.US.2065  
MAC.US.239.95112  
MAC.US.239.96114  
MAC.US.80035  
MAC.US.81035  
MAC.US.85013  
MAC.US.87082  
MAC.US.92050  
MAC.US.92077  
MAC.US.93057  
MAC.US.93062  
MAC.US.95058  
MAC.US.95086  
MAC.US.96016  
MAC.US.96020  
MAC.US.96072  
MAC.US.96081  
MAC.US.96093  
MAC.US.96123  
MAC.US.96135  
MAC.US.97009  
MAC.US.97074  
MAC.US.r80025  
MAC.US.r90131  
MNE.US.-.MNE027  
MNE.US.82.MNE\_8  
SMM.SL.92.SL.92B  
SMM.US.-.17EC1  
SMM.US.-.17EFR  
SMM.US.-.F236\_H4  
SMM.US.-.H9  
SMM.US.-.PBI14\_15  
SMM.US.-.PBIA  
SMM.US.-.PBJ\_143  
SMM.US.-.PBJ\_6P6  
SMM.US.-.PGM53  
SMM.US.-.SME543  
STM.US.-.STM

Sequence alignment table with columns for amino acid positions (V, Y, Y, R, E, G, R, D, Q, L, W, K, G, P, G, E, L, L, W, K, G, E, G, A, V, I, L, K, V, G, T, D, I, K, V, V, P, R, R, R, K, A, K, I, I, K, D, Y, G, G, G, K, E, V, D, S, S) and corresponding nucleotide sequences for various HIV strains. Includes a 'Vif start' label above the sequence.

Strain	Sequence	Position
MAC.US.-.239	ACATGGAGGATACCGAGAGGCTAGA...GAGGTGGCATAGCCTCATAAAATCTGAAATATAAACTAAAGCTACAAAAGTTTGCATGTGCCCATTTTAAAGTCGGATGGCATGGTGACCTGCAGCAGAGTAATCTTCCCCTACAGGAAGAACCCATTT	5539
Pol	S H M E D T G E A R E V A	
Vif	T W R T P E R L E R W H S L I K Y L K Y K T K D L Q K V C Y V P H F K V G W A W W T C S R V I F P L Q E G S H L	
H2A.GM.MCN13	...T...G...G...A...G...-...AA...-...-...-...TG...C...-...-...A...-...C...G...A...-...-...G...-...G...-...C...-...-...CCA...-...A...-...A...-...T...-...G...-...-...A...-...-...T...-...A...-...G...-...AAC...-...T...-...C...-...5067	5067
H2A.CI.88.UC2	...C...G...G...A...G...-...AA...-...-...-...TG...C...-...C...A...-...C...G...A...-...C...G...GGG...-...GC...-...T...-...CCAC...-...G...-...-...T...-...G...-...-...A...-...G...-...A...-...T...-...C...-...5062	5062
H2A.DE.-.BEN	...-...C...-...G...-...G...-...A...-...G...-...AA...-...GC...-...TG...-...C...-...G...-...C...-...C...G...-...A...-...C...G...GG...-...GC...-...T...-...CCAC...-...G...-...-...T...-...G...-...-...A...-...G...-...AA...-...T...-...C...-...5062	5062
H2A.DE.-.PEI2	...-...T...-...G...-...A...-...G...-...A...-...G...-...-...-...-...TG...-...C...-...G...-...C...-...A...-...C...G...-...A...-...-...GG...-...GG...-...GC...-...-...T...-...CCAC...-...G...-...-...T...-...G...-...-...A...-...G...-...AAT...-...T...-...CC...-...5590	5590
H2A.GH.-.GH1	TGGA...GT...CC...GG...AG...A...G...-...-...-...-...T...-...TG...-...C...-...G...-...C...-...A...-...C...G...-...A...-...G...-...C...-...G...GG...-...GC...-...-...T...-...CCAC...-...G...-...G...-...-...T...-...G...-...-...A...-...T...-...A...-...G...-...A...-...T...-...C...-...5066	5066
H2A.GM.-.J57	...-...C...-...G...-...G...-...A...-...G...-...-...AA...-...-...-...-...TG...-...C...-...G...-...C...-...A...-...C...G...-...A...-...G...-...C...-...G...-...GC...-...-...T...-...CCAC...-...G...-...G...-...-...T...-...G...-...-...A...-...T...-...A...-...G...-...AAC...-...T...-...C...-...5063	5063
H2A.GM.87.D194	...-...C...-...G...-...G...-...A...-...G...-...-...-...-...TG...-...C...-...G...-...C...-...A...-...C...G...-...A...-...CT...-...G...GG...-...GC...-...-...T...-...CAC...-...A...-...-...T...-...G...-...-...A...-...G...-...A...-...T...-...C...-...5066	5066
H2A.GW.-.ALI	...-...T...-...G...-...G...-...A...-...G...-...-...A...-...-...TG...-...C...-...G...-...C...-...A...-...G...-...C...-...G...-...A...-...-...CT...-...G...-...GC...-...-...T...-...CCA...-...G...-...-...T...-...G...-...-...A...-...T...-...A...-...G...-...A...-...T...-...C...-...5615	5615
H2A.GW.-.MDS	...-...T...-...G...-...G...-...A...-...G...-...-...A...-...-...TG...-...C...-...G...-...T...-...A...-...-...C...-...A...-...-...G...-...G...-...GC...-...-...C...-...CCA...-...G...-...-...T...-...G...-...-...A...-...T...-...GA...-...G...-...AAC...-...T...-...C...-...5066	5066
H2A.GW.86.FG	...-...TT...-...G...-...G...-...A...-...G...-...-...A...-...-...TG...-...GG...-...C...-...G...-...CT...-...-...C...-...G...-...A...-...CT...-...G...-...G...-...A...-...-...T...-...CA...-...G...-...G...-...-...T...-...G...-...-...A...-...T...-...GA...-...GAAC...-...T...-...C...-...5060	5060
H2A.GW.87.CAM2CG	...-...C...-...G...-...G...-...A...-...G...-...AA...-...G...-...-...A...-...-...TG...-...C...-...G...-...A...-...C...-...AGG...-...G...-...G...-...C...-...A...-...-...T...-...CCA...-...A...-...-...T...-...G...-...-...A...-...T...-...AGA...-...G...-...AC...-...T...-...G...-...C...-...5631	5631
H2A.SN.85.ROD	...-...C...-...G...-...G...-...A...-...G...-...-...AA...-...-...TG...-...C...-...G...-...A...-...C...-...A...-...G...-...G...-...C...-...A...-...-...T...-...CCA...-...G...-...-...T...-...G...-...-...A...-...T...-...A...-...G...-...AAC...-...T...-...C...-...5068	5068
H2AB.CI.-.7312A	...-...G...-...-...A...-...A...-...GC...-...-...AA...-...C...-...T...-...AG...-...C...-...G...-...GC...-...G...-...AGG...-...AT...-...G...-...C...-...C...-...CT...-...C...-...CCAC...-...G...-...C...-...-...T...-...T...-...-...A...-...C...-...A...-...G...-...GCATGCC...-...5604	5604
H2B.CI.-.EHO	...-...G...-...-...ATGC...-...-...AA...-...C...-...T...-...A...-...T...-...C...-...GC...-...GG...-...A...-...CT...-...G...-...C...-...C...-...CT...-...T...-...CCA...-...A...-...-...T...-...T...-...-...A...-...T...-...C...-...GA...-...A...-...GCA...-...C...-...5593	5593
H2B.CI.88.UC1	...-...-...-...-...ATGC...-...-...AA...-...C...-...T...-...GG...-...C...-...G...-...C...-...GC...-...G...-...A...-...GT...-...G...-...C...-...C...-...CT...-...C...-...T...-...CCA...-...G...-...-...T...-...T...-...-...A...-...T...-...C...-...A...-...A...-...GCAT...-...C...-...5603	5603
H2B.GH.86.D205	...-...-...-...-...A...-...GC...-...-...AA...-...C...-...T...-...G...-...T...-...G...-...-...T...-...G...-...AGG...-...GT...-...G...-...C...-...C...-...CT...-...C...-...T...-...CCAC...-...A...-...-...T...-...T...-...A...-...A...-...T...-...C...-...A...-...CA...-...GCATGGC...-...5595	5595
H2B.JP.01.KR20	...-...-...-...-...G...-...C...-...-...AA...-...C...-...AG...-...C...-...-...C...-...G...-...A...-...AT...-...GG...-...C...-...CACT...-...T...-...T...-...CCA...-...A...-...-...T...-...T...-...-...A...-...C...-...A...-...A...-...GCAT...-...C...-...4745	4745
H2G.CI.-.ABT96	...-...-...-...-...A...-...GC...-...A...-...G...-...-...A...-...-...T...-...G...-...-...TCT...-...A...-...G...-...T...-...A...-...C...-...G...-...CGGT...-...T...-...T...-...CAC...-...A...-...T...-...T...-...T...-...G...-...A...-...T...-...ACAA...-...A...-...GCA...-...-...4949	4949
H2U.FR.96.12034	...-...-...-...-...G...-...G...-...AGA...-...T...-...-...A...-...-...CTC...-...G...-...C...-...C...-...G...-...C...-...A...-...G...-...G...-...A...-...G...-...T...-...CAC...-...A...-...-...T...-...T...-...-...G...-...A...-...T...-...T...-...AA...-...TCA...-...5100	5100
MAC.US.-.251.A11	.....	5538
MAC.US.-.251.32H.P15	.....	5540
MAC.US.-.251.BK28	.....	5515
MAC.US.-.MM142	.....	5025
MAC.US.-.SMM142B	.....	5025
MAC.US.1937	.....	4645
MAC.US.2065	.....	4645
MAC.US.239.95112	.....	4635
MAC.US.239.96114	.....	4641
MAC.US.80035	.....	4645
MAC.US.81035	.....	4646
MAC.US.85013	.....	4579
MAC.US.87082	.....	4616
MAC.US.92050	.....	4639
MAC.US.92077	.....	4646
MAC.US.93057	.....	4644
MAC.US.93062	.....	4631
MAC.US.95058	.....	4633
MAC.US.95086	.....	4502
MAC.US.96016	.....	4646
MAC.US.96020	.....	4638
MAC.US.96072	.....	4650
MAC.US.96081	.....	4648
MAC.US.96093	.....	4646
MAC.US.96123	.....	4535
MAC.US.96135	.....	4631
MAC.US.97009	.....	4514
MAC.US.97074	.....	4514
MAC.US.r80025	.....	4502
MAC.US.r90131	.....	4645
MNE.US.-.MNE027	.....	5007
MNE.US.82.MNE.8	.....	5007
SMM.SL.92.SL.92B	GT...CAGTAGA...AA...TGGA...C...C...TG...T...TC...AT...G...A...G...G...G...G...A...CC...A...CAC...A...A...AC...AGC...TCT...G...G...T...CT...GG...G...T...C...4973	4973
SMM.US.-.17EC1	.....	5539
SMM.US.-.17EFR	.....	5539
SMM.US.-.F236_H4	...T...C...C...C...C...G...T...C...T...CA...A...T...T...T...AG...T...AG...CT...5527	5527
SMM.US.-.H9	...T...G...A...C...C...C...C...C...G...C...T...A...CA...A...T...T...GT...AG...T...AGGCT...5012	5012
SMM.US.-.PB114_15	...T...G...A...C...C...C...C...C...G...C...T...A...CA...A...T...T...T...A...T...A...GCT...5306	5306
SMM.US.-.PB1A	...T...G...A...C...C...C...C...C...G...C...T...A...CA...A...T...T...T...A...T...A...GCT...5305	5305
SMM.US.-.PB1_143	...T...G...A...C...C...C...C...C...G...C...T...A...CA...A...T...T...GT...AG...T...AGGCT...5012	5012
SMM.US.-.PB1_6P6	...T...A...G...A...C...C...C...C...G...C...T...A...CA...A...T...T...T...A...T...A...GCT...5498	5498
SMM.US.-.PGM53	...T...A...G...A...C...CT...T...G...G...C...G...C...T...A...CA...A...T...T...AGA...T...A...TCT...5456	5456
SMM.US.-.SME543	...T...A...G...A...C...C...C...C...G...C...G...T...C...CA...A...T...T...AG...T...AG...CT...5527	5527
STM.US.-.STM	...C...A...G...G...C...T...A...G...C...G...AGC...CA...CCA...T...G...T...T...G...T...T...CT...G...A...G...A...GCA...CC...5186	5186



MAC.US.-239
H2A.GM.MCN13
H2A.CI.88.UC2
H2A.DE.-BEN
H2A.DE.-PEI2
H2A.GH.-GH1
H2A.GM.-JSY
H2A.GM.87.D194
H2A.GW.-AL1
H2A.GW.-MDS
H2A.GW.86.FG
H2A.GW.87.CAM2C6
H2A.SN.85.ROD
H2AB.CI.-7312A
H2B.CI.-EHO
H2B.CI.88.UC1
H2B.GH.86.D205
H2B.JP01.KR020
H2G.CI.-ABT96
H2U.FR.96.12034
MAC.US.-251.1A11
MAC.US.-251.32H.PI5
MAC.US.-251.BK28
MAC.US.-MM142
MAC.US.-SMM142B
MAC.US.1937
MAC.US.2065
MAC.US.239.95112
MAC.US.239.96114
MAC.US.80035
MAC.US.81035
MAC.US.85013
MAC.US.87082
MAC.US.92050
MAC.US.92077
MAC.US.93057
MAC.US.93062
MAC.US.95058
MAC.US.95086
MAC.US.96016
MAC.US.96020
MAC.US.96072
MAC.US.96081
MAC.US.96093
MAC.US.96123
MAC.US.96135
MAC.US.97009
MAC.US.97074
MAC.US.r80025
MAC.US.r90131
MNE.US.-MNE027
MNE.US.82.MNE.8
SMM.SL.92.SL.92B
SMM.US.-17EC1
SMM.US.-17EFR
SMM.US.-F236\_H4
SMM.US.-H9
SMM.US.-PB114\_15
SMM.US.-PBJ14
SMM.US.-PBJ\_143
SMM.US.-PBJ\_6P6
SMM.US.-PGM53
SMM.US.-SME543
STM.US.-STM

AGAAGTACAAGGGTATTGGCATTGACACCGAAAAAGGGTGGCTCAGTACTTATGCAGTGAGGATAACCTGGTACTCAAAGAACTTTGGACAGATGTAACACAACTATGCAGACATTTTACTGCATAGCACTTATTTCCCTTGCCTTACAGCGGGAGAAGTGAGAA 5709
E V Q G Y W H L T P E K G W L S T Y A V R I T W Y S K N F W T D V T P N Y A D I L L H S T Y F P C F T A G E V R
--A--G-CA--A-CC-A--G-----A-----TCCT--T--A--A--G--T-----TA-GG-A-GG-C-----T-C-G-G-----TCCC-A-A-A-----A-T-A-----5237
--GA--G-CA--A-CC-A--G-----A-----TCCT--C-T--A--AC-----TA-G-A--G-C-----T-C-G-G-----C-A-A-----C-----TG-----G-A-T-A-----5238
--A--G-CA--A-CC-A--G-----A-----TCCT--C-----A--AT-----TA-G-A--G-C-----T-C-G-G-----C-A-A-----C-----T-----G-A-T-A-----5239
--GA--G-CA--A-CC-A--G-----A-----TCCT--C-----A--A-----TA-G-A--GG-C-----T-C-G-G-----TCCC-A-A-A-----T-----G-A-T-A-----5260
--GA--G-CA--A-CC-A--G-----A-----TCCT--C-T--A--A-----TA-G-A--GG-C-----T-C-G-G-----C-A-A-----C-----T-----G-A-T-A-----5236
--GA--G-CA--A-CC-A--G-----A-----TCCT--T--A--A--G--T-----TA-G-A--G-C-----T-C-G-G-----G-AGGG-C-----T-----G-A-T-A-----5233
--GA--G-CA--A-CC-A--G-----A-----TCCT--C-T--A--T-----TA-G-A--G-C-----T-C-G-G-----TCCC-A-A-A-----C-----T-----G-A-T-A-----5236
--GA--G-CA--A-CC-A--G-----A-----TCCT--C-----A--G-----TA-G-A--G-C-----T-C-G-G-----TCCC-A-A-A-----G-----T-----G-A-T-A-----5236
--GA--G-CA--A-CC-A--G-----A-----TCCT--A-----C-G--T-----TA-G-A--A-C-----T-C-G-G-----TCCC-A-A-A-----C-----T-----G-A-T-A-----5236
G--GA--CA--A-C-A-----A-----TCCT--C-CT--A--A-----TA-G-A--G-C-----T-C-G-G-----TCCC-A-A-A-----T-----T-----G-A-T-A-----5230
--GA--G-CA--A-C-A-----A-----TCCT--T--A--A--G--T-----TA-G-A--G-C-----C-C-G-G-----C-A-A-----T-----T-----A-G-A-----5801
--GA--G-CA--A-C-A-----A-----TCCT--T--A--A-----TA-G-A--G-C-----T-C-G-G-----TG-CC-A-A-----T-----T-----A-T-A-----5238
--G-C-----A-C-C-G--GG-A-TCT-G--T-C--T--AT--A--TGA-G--G--AC-----TG-TGTG-----CAGC--T--G-G-----T-----CAAT--A--G--5774
--C-----A-C-A-CC-C--GG-A-TCT-G--T-C--T--A--AC--A--TGA-G--G--AT-----T--TG-TGTA--CGA-----G-GT-----T--CG-----TAAT--A--G--5763
--G-C-----A-C-C-G--GG-A-TCT-G--CT-C--T--A--AC--A--TAA-G--G--AT-----G--TG-TGTG-----CAGC-----G-GT-----T-----CAAT--A--G--5773
--C-----A-A-CC-A--C--GG-A-TCT-G--CT-C--T--A--AC--A--TGA-G--G--AT-----G--TG-TGTG-----CAGC-----G-GT-----T-----T-----CAAT--A--G--5765
--GA--CA--A-CC--C-G--GG-A-TC--G--G-C-----T--A--AC--A--TAG--A--GT--C-AT-----G--TG-TGTA--CACC--T--G-GT-----T-----TAAT--A--G--4915
G--A-----A-C-C-T--G--A--T-A-CCAG-----A--AT--A--GA--A--AT-----G--G-GACA--TCAA--T--G-AT-----TGA-----T--T-A-----5119
G--G--G--CC--A--C-----G--G--A--ATCCT-----A--T--T--A--GAGGT-----TG-T-GC--CAA-----G-G-G-----T-----T-----G-----5270

MAC.US.-.239  
 Vif  
 H2A.GM.MCN13  
 H2A.C1.88.UC2  
 H2A.DE.-.BEN  
 H2A.DE.-.PEI2  
 H2A.GH.-.GH1  
 H2A.GM.-.J57  
 H2A.GM.87.D194  
 H2A.GW.-.ALI  
 H2A.GW.-.MDS  
 H2A.GW.86.FG  
 H2A.GW.87.CAM2CG  
 H2A.SN.85.ROD  
 H2AB.C1.-.7312A  
 H2B.C1.-.EHO  
 H2B.C1.88.UC1  
 H2B.GH.86.D205  
 H2B.JP.01.KR.020  
 H2G.C1.-.ABT96  
 H2U.FR.96.12034  
 MAC.US.-.251.A111  
 MAC.US.-.251.32H.P15  
 MAC.US.-.251.BK28  
 MAC.US.-.MM142  
 MAC.US.-.SMM142B  
 MAC.US.1937  
 MAC.US.2065  
 MAC.US.239.95112  
 MAC.US.239.96114  
 MAC.US.80035  
 MAC.US.81035  
 MAC.US.85013  
 MAC.US.87082  
 MAC.US.92050  
 MAC.US.92077  
 MAC.US.93057  
 MAC.US.93062  
 MAC.US.95058  
 MAC.US.95086  
 MAC.US.96016  
 MAC.US.96020  
 MAC.US.96072  
 MAC.US.96081  
 MAC.US.96093  
 MAC.US.96123  
 MAC.US.96135  
 MAC.US.97009  
 MAC.US.97074  
 MAC.US.r80025  
 MAC.US.r90131  
 MNE.US.-.MNE027  
 MNE.US.82.MNE\_8  
 SMM.SL.92.SL.92B  
 SMM.US.-.17EC1  
 SMM.US.-.17EER  
 SMM.US.-.F236\_H4  
 SMM.US.-.H9  
 SMM.US.-.PBI14\_15  
 SMM.US.-.PBIA  
 SMM.US.-.PBJ\_143  
 SMM.US.-.PBJ\_6P6  
 SMM.US.-.PGM53  
 SMM.US.-.SME543  
 STM.US.-.STM

GGGCCATCAGGGGAGAACAACCTGTCTTGTCTGACAGGTTCCCGAGAGCTCATAAGTACCAGGTACCAAGCCTACAGTACTAGCACTGAAAGTAGTAAGC.....GATGTGACATCCCAAGGGAGAGAAATCCCACTGGAAACAGTGGAGAAGAGACAAATAGGAGAGGC	5873
<sup>Ypx start</sup> M S D P R E R I P P G N S G E E T T G E A	
R A I R G E O L L S C C R F P R A H K Y Q V P S L O Y L A L K V V S . . . . . D T V R S Q G E N P T W K Q W R R D N R R G	
A-----A-G-A-GT-AT-C-----AT-AT-CCA-C-----GAGC-----GTCA-C-A-TT-G-CT-AGTG-----GCAAG-----CAAA-A-A-C-----A-A-CGG-A-C-A-----C-----T-C-A-----	5404
A-----A-G-A-GT-AT-C-----AC-AT-CCA-----AGCA-----TCA-T-A-C-C-----AGTG-----GCAA-----CAAA-A-A-C-----A-A-GG-G-A-----C-----T-----	5959
A-----A-G-A-GT-AT-C-----AC-AT-CCA-----AGCA-----TCA-T-A-C-C-----AGT-----CAA-----CAA-A-C-----A-A-GG-A-G-A-----C-----T-----	5959
A-----A-G-A-GT-A-----C-----AT-A-CCA-C-----GA-CTA-----GTTA-C-A-TTC-G-CT-AGT-----GCAA-----CAAA-G-C-----AA-A-C-G-A-----A-----G-----C-----TA-T-----	5927
A-----A-G-A-GT-AT-C-----AC-AT-CCA-----AGTA-----TCA-T-A-C-C-----AGTG-----GCAA-----CAAA-A-C-----A-A-GG-A-G-A-----C-----T-----	5403
A-----A-G-A-GT-AT-C-----AT-AT-CCA-C-----C-A-----GTCA-C-A-TTC-G-CT-AGT-----CAG-----CAAA-A-A-C-----A-A-C-----A-----C-----A-----T-----C-----A-----	5400
A-----A-G-A-GT-AT-C-----AC-A-----AGCA-----TCA-T-A-C-C-----AGTG-----GCAA-----CAAA-G-C-----A-A-GG-G-A-----A-----C-----T-----	5403
A-----A-G-A-GT-AT-C-----AT-AT-CCAG-C-----CA-----GTCA-C-A-TTC-G-CT-AGTG-----GCAA-----CAAA-G-A-C-----A-A-C-G-A-----A-----C-----T-----C-----	5952
A-----A-G-A-GT-AT-C-----AT-AT-CCAG-C-----A-CT-----GCA-T-A-TTC-G-CT-AGTG-----GCAA-----CAAA-G-----C-----A-----C-G-A-----A-----C-----C-----AGC-----	5403
A-----A-G-A-T-AT-C-----AA-A-----CC-G-----GA-C-----GTCA-T-A-TTC-G-CT-AGTG-----GCAA-----CAAA-A-----A-----A-C-G-G-A-----A-----C-----C-----G-----T-----C-----A-----	5397
A-----A-G-A-GT-AT-C-----AA-AT-CC-----G-A-C-----TCA-T-A-TTC-G-CT-AGTG-----GCAA-----CAAA-A-----C-----A-----CCG-A-----A-----C-----C-----T-----C-----A-----	5968
A-----A-G-A-GT-AT-C-----AT-AT-CC-----GAGC-----GTCA-T-A-TTC-G-CT-AGTG-----GCAA-----CAAA-A-----C-----A-----C-G-A-----A-----C-----C-----T-----C-----	5405
A-----A-GA-AT-C-A-----AC-A-----ATC-----CG-GGG-----TT-----TTC-----CT-A-G-T-----CAAGAAGGAA-AAATG-----G-G-A-----A-----C-----AC-----G-----	5944
A-----A-GA-AT-CCA-----AC-A-----ATC-----CAGG-----TT-----TTC-----C-A-G-T-----CAAGAAGGAA-A-ATG-----G-A-A-----A-----C-----C-----A-----G-----	5933
A-----A-GA-AT-C-A-----AC-AT-ATC-----CG-AGGG-----T-----TTC-----C-A-GG-CA-----CAGGAAGGAA-A-ATG-----G-G-A-----A-----C-----AC-----A-----A-----	5943
A-----A-GA-AT-C-A-----AC-AT-ATC-----CG-AGGG-----T-----TTC-----C-A-GG-C-----CAGGAAGGAA-AAATG-----G-G-A-----A-----C-----AC-----A-----G-----	5935
A-----A-GA-AT-C-A-----AC-A-----ATC-----GAGGA-----TT-----TTC-----C-A-GG-T-----GCAAGAAGGA-GAAATG-----G-G-G-A-----A-----C-----C-----A-----A-----	5085
A-----G-GA-AT-C-----AC-AT-A-C-----AAGG-G-T-----T-G-A-TTC-----CT-AC-----GCAA-----AAA-GACATG-----A-----G-M-----A-A-----C-----AC-----G-----C-----	5286
A-----A-----GT-AT-C-----AC-T-----ATCG-----GAC-G-----TCTT-----A-TTC-----C-G-C-T-GCAA-----GG-----A-A-----A-----A-----TTCA-A-----A-----TG-----	5434
-----T-----C-----A-----G-----T-----G-----T-----T-----	5872
-----T-----C-----A-----G-----T-----G-----T-----T-----G-----A-----	5874
-----A-----A-----G-----T-----A-----A-----G-----T-----T-----C-----	5849
-----G-----A-----A-----C-----T-----G-----T-----C-----	5359
-----G-----A-----A-----C-----T-----G-----T-----C-----	5359
-----G-----A-----A-----C-----T-----G-----T-----C-----	4979
-----G-----A-----A-----C-----T-----G-----T-----C-----	4979
-----G-----A-----A-----C-----T-----G-----T-----C-----	4979
-----G-----A-----A-----C-----T-----G-----T-----C-----	4980
-----G-----A-----A-----C-----T-----G-----T-----C-----	4913
-----G-----A-----A-----C-----T-----G-----T-----C-----	4950
-----G-----A-----A-----C-----T-----G-----T-----C-----	4973
-----G-----A-----A-----C-----T-----G-----T-----C-----	4980
-----G-----A-----A-----C-----T-----G-----T-----C-----	4978
-----G-----A-----A-----C-----T-----G-----T-----C-----	4965
-----G-----A-----A-----C-----T-----G-----T-----C-----	4967
-----G-----A-----A-----C-----T-----G-----T-----C-----	4836
-----G-----A-----A-----C-----T-----G-----T-----C-----	4980
-----G-----A-----A-----C-----T-----G-----T-----C-----	4972
-----G-----A-----A-----C-----T-----G-----T-----C-----	4984
-----G-----A-----A-----C-----T-----G-----T-----C-----	4982
-----G-----A-----A-----C-----T-----G-----T-----C-----	4980
-----G-----A-----A-----C-----T-----G-----T-----C-----	4869
-----G-----A-----A-----C-----T-----G-----T-----C-----	4965
-----G-----A-----A-----C-----T-----G-----T-----C-----	4848
-----G-----A-----A-----C-----T-----G-----T-----C-----	4848
-----G-----A-----A-----C-----T-----G-----T-----C-----	4836
-----G-----A-----A-----C-----T-----G-----T-----C-----	4979
-----G-----A-----A-----C-----T-----G-----T-----C-----	5341
-----G-----A-----A-----C-----T-----G-----T-----C-----	5341
AA-----A-----G-GG-T-A-C-A-----TG-C-A-G-GT-----C-----C-C-CTAGT-C-A-----CTG-G-C-A-G-----TTA-----CAAA-A-----C-A-----A-----A-----A-----	5310
-----A-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5873
-----A-----G-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5873
A-----GA-T-----A-----A-----AA-T-----TC-----A-C-----AGTC-----G-----A-----A-----A-----	5861
-----A-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5346
-----A-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5640
-----A-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5639
-----A-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5346
-----A-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5832
-----GA-T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5790
-----GA-T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----	5861
A-----A-GT-A-----A-----A-----A-----A-----T-----TT-A-----G-----GAAC-----A-----A-----A-G-C-----G-----AG-A-----	5520

MAC.US.-.239  
Vif  
H2A.GM.MCN13  
H2A.CI.88.UC2  
H2A.DE.-.BEN  
H2A.DE.-.PEI2  
H2A.GH.-.GH1  
H2A.GM.-.J51  
H2A.GM.87.D194  
H2A.GW.-.ALI  
H2A.GW.-.MDS  
H2A.GW.86.FG  
H2A.GW.87.CAM2CG  
H2A.SN.85.ROD  
H2AB.CI.-.7312A  
H2B.CI.-.EHO  
H2B.CI.88.UC1  
H2B.GH.86.D205  
H2B.JP.01.KR020  
H2G.CI.-.ABT96  
H2U.FR.96.12034  
MAC.US.-.251.IA11  
MAC.US.-.251.32H.P15  
MAC.US.-.251.BK28  
MAC.US.-.MM142  
MAC.US.-.SMM142B  
MAC.US.1937  
MAC.US.2065  
MAC.US.239.95112  
MAC.US.239.96114  
MAC.US.80035  
MAC.US.81035  
MAC.US.85013  
MAC.US.87082  
MAC.US.92050  
MAC.US.92077  
MAC.US.93057  
MAC.US.93062  
MAC.US.95058  
MAC.US.95086  
MAC.US.96016  
MAC.US.96020  
MAC.US.96072  
MAC.US.96081  
MAC.US.96093  
MAC.US.96123  
MAC.US.96135  
MAC.US.97009  
MAC.US.97074  
MAC.US.r80025  
MAC.US.r90131  
MNE.US.-.MNE027  
MNE.US.82.MNE.8  
SMM.SL.92.SL92B  
SMM.US.-.17EC1  
SMM.US.-.17EFR  
SMM.US.-.F236\_H4  
SMM.US.-.H9  
SMM.US.-.PBI14\_15  
SMM.US.-.PBIA  
SMM.US.-.PBJ\_143  
SMM.US.-.PBJ\_6P6  
SMM.US.-.PGM53  
SMM.US.-.SME543  
STM.US.-.STM

Sequence alignment table with columns for amino acid positions (L, R, M, A, K, Q, N, S, R, G, D, K, Q, R, G, G, K, P, P, T, K, G, A, N, F, P, G, L, A, K, V, L, G, I, L, A, F) and corresponding nucleotide sequences (CTTCGAATGGCTAAACAGAACAGTAGAGGAGATAAACAGAGAGGGCGTAAACCCACTACCAAGGGAGCTAATTTCCAGGTTTGGCAAAGGCTTGGGAATCTGGCATGATGAACAAGGGATGTCACCAAGCTATGTAAAATACAGATACTTGTGTTTAATACAAAAGG 6043). The table shows various mutations and insertions across different HIV-2/SIV strains.



MAC.US.-239  
Vpr  
Tat exon 1  
H2A.GM.MCN13  
H2A.CI.88.UC2  
H2A.DE.-BEN  
H2A.DE.-PEI2  
H2A.GH.-GH1  
H2A.GM.-J57  
H2A.GM.87.D194  
H2A.GW.-ALI  
H2A.GW.-MDS  
H2A.GW.86.FG  
H2A.GW.87.CAM2CG  
H2A.SN.85.ROD  
H2AB.CI.-7312A  
H2B.CI.-EHO  
H2B.CI.88.UC1  
H2B.GH.86.D205  
H2B.JP01.KR020  
H2G.CI.-ABT96  
H2U.FR.96.12034  
MAC.US.-251\_1A11  
MAC.US.-251\_32H\_P15  
MAC.US.-251\_BK28  
MAC.US.-MM142  
MAC.US.-SMM142B  
MAC.US.1937  
MAC.US.2065  
MAC.US.239.95112  
MAC.US.239.96114  
MAC.US.80035  
MAC.US.81035  
MAC.US.85013  
MAC.US.87082  
MAC.US.92050  
MAC.US.92077  
MAC.US.93057  
MAC.US.93062  
MAC.US.95058  
MAC.US.95086  
MAC.US.96016  
MAC.US.96020  
MAC.US.96072  
MAC.US.96081  
MAC.US.96093  
MAC.US.96123  
MAC.US.96135  
MAC.US.97009  
MAC.US.97074  
MAC.US.r80025  
MAC.US.r90131  
MNE.US.-MNE027  
MNE.US.82.MNE\_8  
SMM.SL.92.SL.92B  
SMM.US.-17EC1  
SMM.US.-17EFR  
SMM.US.-F236\_H4  
SMM.US.-H9  
SMM.US.-PBI14\_15  
SMM.US.-PBIA  
SMM.US.-PBJ\_143  
SMM.US.-PBJ\_6P6  
SMM.US.-PGM53  
SMM.US.-SME543  
STM.US.-STM

Genomic alignment table for Tat exon 1 start. Reference sequence: GGATGAATGGGTAGTGGAGGTTCTGGAAAGAACTGAAAGAAGAAGCTTTAAACATTTTGTATCCTCGCTTGTCTAAGTGCACCTTGGTAATCATATCTATAATAGACATGGAGACACCTTGGAGGAGCAGGAGAACTCATTAGAACTCTCAACGAGCGCTTCTCATGCATT 6367. The table lists various HIV-1 sequence variants and their corresponding amino acid translations. Variants include MAC.US.-239, H2A.GM.MCN13, H2A.CI.88.UC2, H2A.DE.-BEN, H2A.DE.-PEI2, H2A.GH.-GH1, H2A.GM.-J57, H2A.GM.87.D194, H2A.GW.-ALI, H2A.GW.-MDS, H2A.GW.86.FG, H2A.GW.87.CAM2CG, H2A.SN.85.ROD, H2AB.CI.-7312A, H2B.CI.-EHO, H2B.CI.88.UC1, H2B.GH.86.D205, H2B.JP01.KR020, H2G.CI.-ABT96, H2U.FR.96.12034, MAC.US.-251\_1A11, MAC.US.-251\_32H\_P15, MAC.US.-251\_BK28, MAC.US.-MM142, MAC.US.-SMM142B, MAC.US.1937, MAC.US.2065, MAC.US.239.95112, MAC.US.239.96114, MAC.US.80035, MAC.US.81035, MAC.US.85013, MAC.US.87082, MAC.US.92050, MAC.US.92077, MAC.US.93057, MAC.US.93062, MAC.US.95058, MAC.US.95086, MAC.US.96016, MAC.US.96020, MAC.US.96072, MAC.US.96081, MAC.US.96093, MAC.US.96123, MAC.US.96135, MAC.US.97009, MAC.US.97074, MAC.US.r80025, MAC.US.r90131, MNE.US.-MNE027, MNE.US.82.MNE\_8, SMM.SL.92.SL.92B, SMM.US.-17EC1, SMM.US.-17EFR, SMM.US.-F236\_H4, SMM.US.-H9, SMM.US.-PBI14\_15, SMM.US.-PBIA, SMM.US.-PBJ\_143, SMM.US.-PBJ\_6P6, SMM.US.-PGM53, SMM.US.-SME543, STM.US.-STM. The table shows nucleotide differences and amino acid changes relative to the reference sequence.

MAC.US.-.239  
Vpr

Tat exon 1  
Rev exon 1  
H2A.GM.MCN13  
H2A.CL88.UC2  
H2A.DE.-BEN  
H2A.DE.-PEI2  
H2A.GH.-GH1  
H2A.GM.-JBY  
H2A.GM.87.D194  
H2A.GW.-ALI  
H2A.GW.-MDS  
H2A.GW.86.FG  
H2A.GW.87.CAM2CG  
H2A.SN.85.ROD  
H2AB.CL.-.7312A  
H2B.CL.-EHO  
H2B.CL88.UC1  
H2B.GH.86.D205  
H2B.JP01.KR020  
H2C.CL.-ABT96  
H2U.FR.96.12034  
MAC.US.-.251.1A11  
MAC.US.-.251.32H.PJ5  
MAC.US.-.251.BK28  
MAC.US.-.MM142  
MAC.US.-.SMM142B  
MAC.US.1937A  
MAC.US.2065  
MAC.US.239.95112  
MAC.US.239.96114  
MAC.US.80035  
MAC.US.81035  
MAC.US.85013  
MAC.US.87082  
MAC.US.92050  
MAC.US.92077  
MAC.US.93057  
MAC.US.93062  
MAC.US.95058  
MAC.US.95086  
MAC.US.96016  
MAC.US.96020  
MAC.US.96072  
MAC.US.96081  
MAC.US.96093  
MAC.US.96123  
MAC.US.96135  
MAC.US.97009  
MAC.US.97074  
MAC.US.r80025  
MAC.US.r90131  
MNE.US.-.MNE027  
MNE.US.82.MNE.8  
SMM.SL.92.SL92B  
SMM.US.-.17EC1  
SMM.US.-.17EFR  
SMM.US.-.F236.H4  
SMM.US.-.H9  
SMM.US.-.PB114.15  
SMM.US.-.PB1A  
SMM.US.-.PB1.143  
SMM.US.-.PB1.6P6  
SMM.US.-.PGM53  
SMM.US.-.SME543  
STM.US.-.STM

Sequence alignment table with columns for Vpr end and Rev exon 1 start. Rows correspond to the sample IDs listed on the left. The table contains nucleotide sequences (A, C, G, T) and their corresponding positions (6079 to 6185).







Accession	Sequence	Position
MAC.US.-.239	.....ATA	6541
Env	.....ATA	6541
H2A.GM.MCN13	.....ATA	6541
H2A.CI.88.UC2	.....ATA	6541
H2A.DE.-BEN	.....ATA	6541
H2A.DE.-PEI2	.....ATA	6541
H2A.GH.-GH1	.....ATA	6541
H2A.GM.-JSY	.....ATA	6541
H2A.GM.87.D194	.....ATA	6541
H2A.GW.-AL1	.....ATA	6541
H2A.GW.-MDS	.....ATA	6541
H2A.GW.86.FG	.....ATA	6541
H2A.GW.87.CAM2CG	.....ATA	6541
H2A.SN.85.ROD	.....ATA	6541
H2AB.CI.-7312A	.....ATA	6541
H2B.CI.-EHO	.....ATA	6541
H2B.CI.88.UC1	.....ATA	6541
H2B.GH.86.D205	.....ATA	6541
H2B.JP01.KR020	.....ATA	6541
H2G.CI.-ABT96	.....ATA	6541
H2U.FR.96.12034	.....ATA	6541
MAC.US.-.251_1A11	.....ATA	7022
MAC.US.-.251_32H_P15	.....ATA	7021
MAC.US.-.251_BK28	.....ATA	7005
MAC.US.-.MM142	.....ATA	6518
MAC.US.-.SMM142B	.....ATA	6515
MAC.US.1937	.....ATA	6126
MAC.US.2065	.....ATA	6126
MAC.US.239.95112	.....ATA	6116
MAC.US.239.96114	.....ATA	6122
MAC.US.80035	.....ATA	6126
MAC.US.81035	.....ATA	6126
MAC.US.85013	.....ATA	6060
MAC.US.87082	.....ATA	6097
MAC.US.92050	.....ATA	6120
MAC.US.92077	.....ATA	6127
MAC.US.93057	.....ATA	6125
MAC.US.93062	.....ATA	6112
MAC.US.95058	.....ATA	6114
MAC.US.95086	.....ATA	5983
MAC.US.96016	.....ATA	6127
MAC.US.96020	.....ATA	6119
MAC.US.96072	.....ATA	6131
MAC.US.96081	.....ATA	6129
MAC.US.96093	.....ATA	6127
MAC.US.96123	.....ATA	6016
MAC.US.96135	.....ATA	6112
MAC.US.97009	.....ATA	5995
MAC.US.97074	.....ATA	5995
MAC.US.r80025	.....ATA	5983
MAC.US.r90131	.....ATA	6126
MNE.US.-.MNE027	.....ATA	6491
MNE.US.82.MNE_8	.....ATA	6491
SMM.SL.92.SL92B	.....ATA	6462
SMM.US.-.17EC1	.....ATA	7029
SMM.US.-.17EFR	.....ATA	7029
SMM.US.-.F236_H4	.....ATA	6973
SMM.US.-.H9	.....ATA	6494
SMM.US.-.PB114_15	.....ATA	6803
SMM.US.-.PB1A	.....ATA	6802
SMM.US.-.PB1_143	.....ATA	6509
SMM.US.-.PB1_6P6	.....ATA	6995
SMM.US.-.PGM53	.....ATA	6960
SMM.US.-.SME543	.....ATA	7024
STM.US.-.STM	.....ATA	6662



MAC.US.-239	ACAAAGGAAATAACT...GGTAATGAAAGTAGATGTTACATGAACCCTGTAACACTTCTGTTATCCAAGAGCTTTGTGACAAACATTATTGGGTGCTATTAGATTAGGTTATGTGCACCTCCAGGTTATGCTTTGCTTAGATGTAATGACACAAATTTCCAGCT	7354
Env	Q G N N T . G N E S R C Y M N H C N T S V I Q E S C D K H Y W D A I R F R Y C A P P G Y A L L R C N D T N Y S G	
H2A.GM.MCN13	-GG-AATGGC-C-A-...GA-C-...T-C--A-G-C-AC--A-...G-C-...G-G-A-C-...T-CC-A-A-...C-...T-C-...C-...G-...	6866
H2A.CI.88.UC2	CAGCACA-C-G-A-...C-CAA-C-CG-C-...GA-T-C--A-A-C-A-G-A-A-...T-G-C-...T-G-AG-A-C-CA-...A-G-C-T-...C-A-A-...C-...T-C-...C-...G-...	7384
H2A.DE.-BEN	CA-C-...CA-CAG-...G-C-C-...GA-T-C--A-A-C-A-G-A-A-...T-G-C-...T-G-AG-A-C-CA-...A-G-C-T-...C-A-A-...C-...T-C-...C-...G-...	7436
H2A.DE.-PEI2	CAC-ACT--GGG-C-...AGCC-C-A-...T-C--A-A-C-AC--A-...T-G-C-...G-G-A-C-...A-G-...T-ATGC-...A-...C-...T-C-...C-...G-...	7377
H2A.GH.-GH1	-TC-AAT--C-G-AA-...A-GGA-AC-...T-C--A-A-C-AC--A-...T-G-C-...A-AG-A-C-...A-G-...T-CC-A-A-...C-...T-C-...C-...G-...	6853
H2A.GM.-ISY	-TC-AT--C-G-A-...AC-...CG-AA-...T-C--A-A-C-AC--A-...T-G-C-...A-G-...A-C-...A-G-...T-TC-...A-G-C-...T-C-...C-...G-...	6850
H2A.GM.87.D194	CA-GACA-...CGGA-...AC-G-C-C-...GA-T-C--A-A-C-A-...A-...T-G-C-...G-AG-A-C-...A-G-...T-CC-A-A-...C-...T-C-...C-...G-...	6844
H2A.GW.-AL1	-C-TTT--C-C-CACA-...CC-G-C-C-G-...T-C--A-A-C-AC--A-...T-G-C-...A-G-...A-C-...A-T-...C-CC-A-A-...C-...T-C-...C-...G-...	7417
H2A.GW.-MDS	-G-AT--C-C-...CACA-...CC-C-C-G-C-...G-T-C--A-A-C-AC--A-...T-G-C-...A-G-...A-C-...A-G-...T-C-...A-A-...C-...T-C-...C-...G-...	6910
H2A.GW.86.FG	CA-T-...CTCA-...G-C-G-C-AG-...T-C--A-A-C-AC--A-...T-G-C-...G-G-...A-C-...A-G-...T-C-...A-A-...C-...T-C-...C-...G-...	6841
H2A.GW.87.CAM2CQ	TA-TA-C-CAG-TCAGACCACA-...G-CC-CG-...T-C--A-A-C-AC--A-...T-G-C-...G-G-...A-C-...A-G-...T-CA-A-A-...C-...T-C-...G-...C-...	7439
H2A.SN.85.ROD	GAC-...T-GC-...ACA-...C-G-CCCAG-...T-C--A-A-C-AC--A-...T-G-C-...A-G-...A-C-...A-G-...T-CC-AT-A-...C-...T-C-...C-...G-...	6879
H2AB.CI.-7312A	GTC-...G-CAGT-A-GGGAG-GAC-...C-...T-T-C--A-A-C-AC--A-A-C-...G-C-C-...A-G-...A-C-...A-G-...T-C-...A-...C-...C-...G-...	7418
H2B.CI.-EHO	-A-G-...GGG-CA-GG-...A-...C-A-...T-A-AAC-...T-A-AA-...A-A-...A-A-...AG-T-A-...T-C-...A-T-...A-...C-...T-T-A-...	7398
H2B.CI.88.UC1	CGGC-...C-G-C-...AGC-CC-T-...GAACT-...T-C-A-G-...A-A-...T-...C-...CAGCT-A-...C-...T-C-...A-...C-...A-...C-...C-...	7408
H2B.GH.86.D205	TA-T-...CC-GG-AGTAT-...ACC-C-...C-T-A-GAAC-...C-AA-AA-...A-A-...CAGCT-A-G-...T-C-...G-G-TT-C-A-...C-...T-C-...	7403
H2B.JP01.KR020	TGGG-...-C-...-C-...CC-T-A-AAC-...A-AA-...A-...CCAAT-A-...C-C-...C-T-...TA-G-...C-...T-T-...C-...T-T-...C-...	6526
H2G.CI.-ABT96	-G-...TCC-...T-AGAGT-AG-...G-...C-T-...GG-...T-A-A-C-T-G-A-A-...T-...CCAAT-A-...A-...A-...T-G-...G-T-...A-...T-T-...C-...A-...	6767
H2U.FR.96.12034	-G-...CT-...AGT-AA-G-...A-...C-...C-...C-...C-...A-...C-A-...A-...G-...C-...C-...T-A-...T-...T-...C-...A-...	6896
MAC.US.-251_1A11	-G-...	7350
MAC.US.-251_32H_P15	-G-...	7352
MAC.US.-251_BK28	-G-...	7336
MAC.US.-MM142	-G-...	6849
MAC.US.-SMM142B	-G-...	6846
MAC.US.1937	-G-...	6460
MAC.US.2065	-G-...	6460
MAC.US.239.95112	-G-...	6450
MAC.US.239.96114	-G-...	6456
MAC.US.80035	-G-...	6460
MAC.US.81035	-G-...	6461
MAC.US.85013	-G-...	6394
MAC.US.87082	-G-...	6431
MAC.US.92050	-G-...	6454
MAC.US.92077	-G-...	6461
MAC.US.93057	-G-...	6459
MAC.US.93062	-G-...	6446
MAC.US.95058	-G-...	6448
MAC.US.95086	-G-...	6317
MAC.US.96016	-G-...	6461
MAC.US.96020	-G-...	6453
MAC.US.96072	-G-...	6465
MAC.US.96081	-G-...	6463
MAC.US.96093	-G-...	6461
MAC.US.96123	-G-...	6350
MAC.US.96135	-G-...	6446
MAC.US.97009	-G-...	6329
MAC.US.97074	-G-...	6329
MAC.US.r80025	-G-...	6317
MAC.US.r90131	-G-...	6460
MNE.US.-MNE027	-T-...G-...-AAG-...	6825
MNE.US.82.MNE_8	-T-...G-...-AA-...	6822
SMM.SL.92.SL92B	G-...AGGA-TGAAAGC-...	6790
SMM.US.-17EC1	-G-...	7354
SMM.US.-17EFR	-G-...	7354
SMM.US.-F236_H4	G-GTCAGCG-T-...AA-G-G-A-...-C-T-T-...CAG-T-A-C-...G-...A-C-...G-...G-...TT-...T-...	7304
SMM.US.-H9	-AAT-GC-TGAAAT-AA-C-C-A-C-T-...T-CAG-T-A-C-...G-...Y-A-...G-...G-...T-C-...	6831
SMM.US.-PB114_15	-AAT-GC-T-...AA-C-C-A-C-T-...T-CAG-T-A-C-...G-...A-...G-...TT-...	7134
SMM.US.-PB1A	-AAT-GC-T-...AA-C-C-A-C-T-...T-CAG-T-A-C-...G-...A-...G-...TT-...	7133
SMM.US.-PBJ_143	-AAT-GC-TGAAAT-AA-C-C-A-C-T-...T-CAG-T-A-C-...G-...Y-A-...G-...G-...T-C-...	6846
SMM.US.-PBJ_6P6	-AAT-GC-T-...AA-C-C-A-C-T-...T-CAG-T-A-C-...G-...A-...G-...TT-...	7326
SMM.US.-PGM53	-A-T-C-CT-G-...A-A-...-A-C-T-...T-CAG-T-A-C-...G-...A-C-...G-...T-...	7294
SMM.US.-SME543	G-GCCAGCA-TGGAAGT-AG-...-...-C-G-T-...CAG-T-G-A-C-...G-...A-C-...G-...TT-...T-...	7346
STM.US.-STM	-AATGTC-CTGGCAA-AG-...-...-C-...C-...GA-T-C-T-CAG-A-T-...A-C-...-...-CT-A-G-...A-T-...-C-...C-...T-...A-...	6993





V5 loop start

MAC.US.-239	TGTCAAACTCCAGGTACTGGA...	ACTAACAATACTGATAAAATCAATTTGACG...	GCTCTGGA...	GGAGGAGATCCGGAAGTACCTTCATGTGGACAATTCAGAGGAGAGTTCCTCTACTGTAAAATGAATGGTTTCTAAATTGGGTAGAAGATAGGA	7846
Env	V K H P R Y T G	T N N T D K I N L T	A P G	G G D P E V T F M W T N C R G E F L Y C K M N W F L N W V E D R	
H2A.GM.MCN13	CA	AA	AAAG	CTC	7364
H2A.CI.88.UC2	CAGG	AA	GTA	CTC	7882
H2A.DE.BEN	A	AA	GGA	CTC	7931
H2A.DE.PEJ2	G	AA	AAA	TTC	7875
H2A.GH.GH1	A	AA	AGA	CTC	7348
H2A.GM.ISY	CA	AA	AAA	ACTC	7348
H2A.GM.87.D194	CA	CGGA	ATA	TTC	7342
H2A.GW.AL1	A	AA	AGA	CTC	7912
H2A.GW.MDS	A	AA	AAA	CTC	7405
H2A.GW.86.FG	CG	AA	GAGA	CTC	7342
H2A.GW.87.CAM2CG	CGG	AA	AGA	TTC	7937
H2A.SN.85.ROD	CA	AA	AAA	CTC	7380
H2AB.CI.-7312A	A	AA	ACA	CTC	7913
H2B.CI.-EHO	AAA	T	AG	AC	7890
H2B.CI.88.UC1	CA	AA	GTA	TTC	7903
H2B.GH.86.D205	CA	AA	AAA	TTC	7904
H2B.JP01.KR020	CA	G	AAA	CTC	7021
H2G.CI.-ABT96	G	AA	GAA	TTC	7262
H2U.FR.96.12034	C	T	AA	T	7388
MAC.US.-251.1A11			A		7842
MAC.US.-251.32H.P15			A		7844
MAC.US.-251.BK28			A		7828
MAC.US.-MM142		G	A		7341
MAC.US.-SMM142B			A		7338
MAC.US.1937					6952
MAC.US.2065		R		AR	6952
MAC.US.239.95112					6942
MAC.US.239.96114					6948
MAC.US.80035					6952
MAC.US.81035					6953
MAC.US.85013				A	6886
MAC.US.87082				R	6923
MAC.US.92050	R				6946
MAC.US.92077					6953
MAC.US.93057		G			6951
MAC.US.93062					6938
MAC.US.95058					6940
MAC.US.95086					6809
MAC.US.96016					6953
MAC.US.96020					6945
MAC.US.96072		G			6957
MAC.US.96081					6955
MAC.US.96093		G			6953
MAC.US.96123					6842
MAC.US.96135					6938
MAC.US.97009			R	A	6821
MAC.US.97074					6821
MAC.US.r80025			R	A	6809
MAC.US.r90131					6952
MNE.US.-MNE027			C		7317
MNE.US.82.MNE_8				T	7314
SMM.SL.92.SL.92B	A	C	G	A	7282
SMM.US.-17EC1		C	G	A	7846
SMM.US.-17EFR			A		7846
SMM.US.-F236_H4	G		A	T	7796
SMM.US.-H9	G	AG	A	T	7323
SMM.US.-PB114_15	G		A	T	7626
SMM.US.-PB1A	G		A	T	7625
SMM.US.-PBJ_143	G	AG	A	T	7338
SMM.US.-PBJ_6P6	G		A	T	7818
SMM.US.-PGM53	G	A	A	T	7786
SMM.US.-SME543	G		A	T	7838
STM.US.-STM	G	A	A	T	7485

MAC.US.-239 Env	ATACAGCTAAC	.....	CAGAAGCCAAAGGAAACAGCATAAAAGGAATTACGTGCCATGTGCATATAGACAAATATCAACACTTGGCATAAAGTAGGCAAAAATGTTTATTTGCCCTCAAGAGAGGGAGACCTCACGTGTAACCTCCACAGTGACCAGTCTC	8001
	N T A N	.....	Q K P K E Q H K R N Y V P C H I R Q I T N T W H K V G K N V Y L P P R E G D L T C N S T V T S L	
H2A.GM.MCN13	CACACCGC	-T	.....	7489
H2A.CI.88.LUC2	CG.GCCAA	-A	.....	8016
H2A.DE..BEN	-CCA-A-ACGG	.....	.....	8056
H2A.DE..PEI2	CACACC-AC-G	.....	.....	8006
H2A.GH.-GH1	CG-ATCAG-CA	.....	.....	7482
H2A.GM.-JISY	CGGGTCAAC-G	.....	.....	7479
H2A.GM.87.D194	CG-ACCAA-CA	.....	.....	7476
H2A.GW.-AL1	CGGGTCAGG-A	.....	.....	8046
H2A.GW.-MDS	C-GATAAGCCA	.....	.....	7539
H2A.GW.86.FG	CGGGTCAG-A	.....	.....	7476
H2A.GW.87.CAM2CG	C-ATA-G-CA	.....	.....	8071
H2A.SN.85.ROD	CACACCGC	-T	.....	7505
H2AB.CI.-7312A	CGGGTCAG-CA	.....	.....	8047
H2B.CI.-EHO	C-GGGCTC-G	.....	.....	8021
H2B.CI.88.UC1	CAGGTA-A-CT	.....	.....	8037
H2B.GH.86.D205	CA-ATA-A-CT	.....	.....	8038
H2B.JP01.KR020	CAGG-AT--G	.....	.....	7152
H2G.CI.-ABT96	--TGA-CT	.....	.....	7417
H2U.FR.96.12034	GC-TG-AGGGA	.....	.....	7543
MAC.US.-251.1A11	.....	.....	.....	7997
MAC.US.-251.32H.P15	.....	.....	.....	8099
MAC.US.-251.BK28	.....	.....	.....	7983
MAC.US.-MM142	.....	.....	.....	7496
MAC.US.-SMM142B	.....	.....	.....	7493
MAC.US.1937	.....	.....	.....	7109
MAC.US.2065	.....	.....	.....	7107
MAC.US.239.95112	.....	.....	.....	7097
MAC.US.239.96114	.....	.....	.....	7103
MAC.US.80035	.....	.....	.....	7107
MAC.US.81035	.....	.....	.....	7108
MAC.US.85013	.....	.....	.....	7041
MAC.US.87082	.....	.....	.....	7078
MAC.US.92050	.....	.....	.....	7101
MAC.US.92077	.....	.....	.....	7108
MAC.US.93057	.....	.....	.....	7106
MAC.US.93062	.....	.....	.....	7075
MAC.US.95058	.....	.....	.....	7095
MAC.US.95086	.....	.....	.....	6964
MAC.US.96016	.....	.....	.....	7108
MAC.US.96020	.....	.....	.....	7100
MAC.US.96072	.....	.....	.....	7112
MAC.US.96081	.....	.....	.....	7110
MAC.US.96093	.....	.....	.....	7108
MAC.US.96123	.....	.....	.....	6997
MAC.US.96135	.....	.....	.....	7093
MAC.US.97009	.....	.....	.....	6976
MAC.US.97074	.....	.....	.....	6970
MAC.US.r80025	.....	.....	.....	6964
MAC.US.r90131	.....	.....	.....	7107
MNE.US.-MNE027	.....	.....	.....	7478
MNE.US.82.MNE_8	.....	.....	.....	7475
SMM.SL.92.SL92B	.....	.....	.....	7452
SMM.US.-17EC1	.....	.....	.....	8001
SMM.US.-17EFR	.....	.....	.....	8001
SMM.US.-F236_H4	.....	.....	.....	7966
SMM.US.-H9	.....	.....	.....	7490
SMM.US.-PB114_15	.....	.....	.....	7793
SMM.US.-PB1A	.....	.....	.....	7792
SMM.US.-PB1_143	.....	.....	.....	7505
SMM.US.-PB1_6P6	.....	.....	.....	7985
SMM.US.-PGM53	.....	.....	.....	7956
SMM.US.-SME543	.....	.....	.....	8008
STM.US.-STM	.....	.....	.....	7655

V5 loop end\_





Env gp120 end Env gp41 start

Accession	Sequence	Position
MAC.US.-.239	CTCAAGAAATAAAGAGGGGCTTTGTGCTAGGGTTCTGGGTTTTCTCGAACGGCAGGTTCTGCAATGGCGCGGGCTGTGACGCTGACCCGAACTTTATTGGCTGGGATAGTGCAGCAACAGCAACAGCTGTTGGACGTGGTCAAGAGACAACAAG	8332
Env	S R N K R G V F V L G F L A T A G S A M G A A S L T L T A Q S R T L L A G I V Q Q Q Q L L D V V K R R Q Q	
H2A.GM.MCN13	AATG -G -C -G -T -G -C -A -C -A -T -G -T -G -C -C -G -814	7814
H2A.CI.88.UC2	AG -G -C -T -G -C -G -A -G -A -T -G -T -G -C -C -G -8341	8341
H2A.DE.-BEN	AGTG -G -C -T -G -C -G -A -G -A -T -G -T -G -C -C -G -8387	8387
H2A.DE.-PEI2	ACAG -G -C -T -G -C -AG -C -C -T -G -TC -G -CT -AC -G -8331	8331
H2A.GH.-GH1	AGTG -G -C -T -G -C -G -A -G -A -T -G -T -G -C -C -G -8707	7807
H2A.GM.-ISY	AGGG -C -G -T -GC -C -A -A -G -A -G -GC -TC -T -G -T -G -C -CG -8707	7807
H2A.GM.87.D194	AGTG -G -C -T -G -C -G -A -G -A -T -G -T -G -T -G -C -C -8701	7801
H2A.GW.-AL1	AGGG -G -C -T -G -C -A -G -A -G -A -AG -T -A -T -G -C -C -8371	8371
H2A.GW.-MDS	GGG -C -C -T -G -C -A -A -A -G -C -T -A -T -G -C -C -8764	7864
H2A.GW.86.FG	CAG -C -C -T -G -C -A -A -A -AGT -C -T -A -T -G -C -C -8798	7798
H2A.GW.87.CAM2CG	GGG -CC -C -T -GC -A -A -A -AGT -A -C -T -A -C -T -G -C -C -8399	8399
H2A.SN.85.ROD	GGG -C -C -T -G -C -A -A -A -AGT -A -C -T -A -C -T -G -C -C -8353	7833
H2AB.CI.-.7312A	GGGG -C -T -A -C -A -G -A -A -AG -C -T -G -T -G -C -C -8378	8378
H2B.CI.-EHO	AC -G -G -T -AC -T -A -C -T -G -C -T -G -C -G -CG -A -8349	8349
H2B.CI.88.UC1	AC -G -G -T -AA -G -A -C -T -T -AA -C -T -A -G -C -C -8374	8374
H2B.GH.86.D205	AC -G -G -T -AA -G -A -C -T -T -AA -T -A -G -C -C -CTG -A -8363	8363
H2B.JP01.KR020	AC -G -G -T -ACCG -T -A -T -T -A -C -T -G -C -G -C -A -8480	7480
H2G.CI.-ABT96	AC -G -A -T -A -T -T -A -T -T -A -A -T -G -G -G -T -8748	7748
H2U.FR.96.12034	AC -G -G -G -A -A -T -C -T -A -T -A -G -C -C -A -8771	7871
MAC.US.-.251.1A11	.....T.....	8328
MAC.US.-.251.32H.PI5	.....A.....	8320
MAC.US.-.251.BK28	.....G.....	8314
MAC.US.-.MM142	.....CG.....	7827
MAC.US.-.SMM142B	.....CG.....	7824
MAC.US.1937	.....R.....	7440
MAC.US.2065	.....R.....	7438
MAC.US.239.95112	.....	7428
MAC.US.239.96114	.....	7434
MAC.US.80035	.....	7438
MAC.US.81035	.....	7439
MAC.US.85013	.....	7372
MAC.US.87082	.....	7409
MAC.US.92050	.....	7432
MAC.US.92077	.....	7439
MAC.US.93057	.....	7437
MAC.US.93062	.....	7406
MAC.US.95058	.....	7426
MAC.US.95086	.....	7295
MAC.US.96016	.....	7439
MAC.US.96020	.....	7431
MAC.US.96072	.....	7443
MAC.US.96081	.....	7441
MAC.US.96093	.....	7439
MAC.US.96123	.....	7328
MAC.US.96135	.....	7424
MAC.US.97009	.....G.....	7307
MAC.US.97074	.....	7301
MAC.US.r80025	.....	7295
MAC.US.r90131	.....	7438
MNE.US.-.MNE027	.....A.....	7809
MNE.US.82.MNE.8	.....A.....	7806
SMM.SL.92.SL92B	.....AG.....A.....	7780
SMM.US.-.17EC1	.....	8332
SMM.US.-.17EFR	.....	8332
SMM.US.-.F236_H4	.....G.....G.....G.....T.....	8297
SMM.US.-.H9	.....G.....A.....CG.....T.....G.....G.....T.....	7821
SMM.US.-.PB114_15	.....G.....A.....CG.....T.....G.....G.....T.....	8124
SMM.US.-.PBJ1	.....G.....A.....CG.....T.....G.....G.....T.....	8123
SMM.US.-.PBJ_143	.....G.....A.....CG.....T.....G.....G.....T.....	7836
SMM.US.-.PBJ_6P6	.....G.....A.....CG.....T.....G.....G.....T.....	8316
SMM.US.-.PGM53	.....G.....A.....CG.....T.....G.....G.....T.....	8284
SMM.US.-.SME543	.....G.....A.....CG.....T.....G.....G.....T.....	8339
STM.US.-.STM	.....C.....G.....C.....G.....A.....	7983







MAC.US.-.239  
Env  
Tat exon 2  
Rev exon 2  
H2A.GM.MCN13  
H2A.CI.88.UC2  
H2A.DE.-.BEN  
H2A.DE.-.PEI2  
H2A.GH.-.GHI  
H2A.GM.-.JBY  
H2A.GM.87.D194  
H2A.GW.-.ALI  
H2A.GW.-.MDS  
H2A.GW.86.FG  
H2A.GW.87.CAM2CG  
H2A.SN.85.ROD  
H2AB.CI.-.7312A  
H2B.CI.-.EHO  
H2B.CI.88.UC1  
H2B.GH.86.D205  
H2B.JP01.KR020  
H2G.CI.-.ABT96  
H2U.FR.96.12034  
MAC.US.-.251.1A11  
MAC.US.-.251.32H.PJ5  
MAC.US.-.251.BK28  
MAC.US.-.MM142  
MAC.US.-.SMM142B  
MAC.US.19374  
MAC.US.2065  
MAC.US.239.95112  
MAC.US.239.96114  
MAC.US.80035  
MAC.US.81035  
MAC.US.85013  
MAC.US.87082  
MAC.US.92050  
MAC.US.92077  
MAC.US.93057  
MAC.US.93062  
MAC.US.95058  
MAC.US.95086  
MAC.US.96016  
MAC.US.96020  
MAC.US.96072  
MAC.US.96081  
MAC.US.96093  
MAC.US.96123  
MAC.US.96135  
MAC.US.97009  
MAC.US.97074  
MAC.US.r80025  
MAC.US.r90131  
MNE.US.-.MNE027  
MNE.US.82.MNE.8  
SMM.SL.92.SL92B  
SMM.US.-.17EC1  
SMM.US.-.17EFR  
SMM.US.-.F236.H4  
SMM.US.-.H9  
SMM.US.-.PBJ14.15  
SMM.US.-.PBJA  
SMM.US.-.PBJ.143  
SMM.US.-.PBJ.6P6  
SMM.US.-.PGM53  
SMM.US.-.SME543  
STM.US.-.STM

Tat end  
TGCCAAC CAGAGAAGGCAAGAAAGAGACGGTGGAGAAGCGGTGGCAACAGCTCTGGCCTTGGCAGATAGAATATATTCTTCTGATCCGCCAACTGATACGCCTCTTGACTTGGCTATTTCAGCAACTGCAGAACCTTGCTATCGAGAGTATACCAGATCCTCCAA 9000  
L P T R E G K K E R D G G E K G G N S S W P W Q I E Y I H F L I R Q L I R L L T W L F S N C R T L L S R V Y Q I L L Q  
C Q P E K A K E T V E K A V A T A P G L G R \*  
A N Q R R Q R K R R W R R W Q Q L L A L A D R I Y S F P D P P T D T P L D L A I Q Q L Q N L A I E S I P D P P  
A--G--AA-C--GA--T--A-CAA-T--AG--G--AG--A--A--T--G-A--A--C--GA-C--T--C--G 8482  
A--GG--AC--AA-C--G--CC--C-C-A--T--GA--T--C--A-C-C-A--GC--A--TG--CT--T--G--A--T--GGA--A--CGCAAC--C--C--G--G 9009  
A--G--AC--A-C--GA--TC--GC-A-A--TT--GA--TG--CA--A-C--G-G-G--A--T--CT--TCG--A--A--T--GA--A--C--AGAAC--C--C--G--G 9055  
A--GA--AA-C--CGA--CC--A-CA--AT--AG--G--CA--C--A--A--G--T--CG--A--GT--AGGA--A--C--GAGC--T--C--C--A--G 8999  
A--G--AC--A-C--GA--ACA--C-A--T--TT--GA--TG--A--C--A-C--A--A--T--CT--CG--A--AG--T--GA--A--C--CGAAC--C--C--A--G 8475  
A--GA--AA-C--GA--T--GA-C-A--T--A-G--A--TG--C--C--A--A--G--GGA--A--C--C--T--C--G--G 8475  
A--G--AC--AA-C--GA--CC--C-A-A--TTT--G--TG--CAC--A-C--C-A--A--A--T--CT--CG--A--A--G--GGG--A--C--AGAAC--C--C--G--G 8469  
A--G--G--AA-C--GA--T--CAA--T--GG--A--TG--C--C--C--T--A--A--TG--AGCT--TCG--G--A--T--GGA--A--C--GA-C--C--C--G--G 9039  
A--GA--AA-C--GA--T--GCCAA--AG--A--T--C--C--C--A--A--TG--T--CA--A--G--A--G--GGA--A--C--GA-C--T--C--C-- 8532  
A--G--AA-C--GA--T--AGCAA--AG--A--T--C--C--A--A--TG--T--CG--A--A--T--GGA--A--C--GA-C--C--C-- 8466  
A--G--AA-C--GA--T--GCCAA--AG--A--T--C--C--A--A--TG--T--CG--A--A--T--G--A--A--C--GA-C--T--C-- 9067  
A--G--AC--AA-C--GA--T--AGCAA--AG--A--A--C--C--C--A--A--G--T--CA--A--A--T--GGA--A--C--GAGC--T--C-- 8501  
A--G--AC--AA-C--GA--GA--A-C-A--G--T--A--T--C--C--C--C--A--T--G--GGAA--T--A--GA--GG--A--G--T--AGACC--T--A--C-- 9046  
A--G--AC--AA-C--GA--GA--A-CAA--AG--T--A--T--C--C--C--C--C--T--GGGA--T--A--GG--A--AG--T--AGACC--T--A--C-- 9017  
A--G--AC--AA-C--GAC--A--A-C-AA--G--CTT--A--T--C--CGC--C--T--T--GGAA--T--G--A--GG--G--A--G--T--AGACC--C--A-- 9042  
A--G--AC--AA-C--GA--GA--A-C-A--G--A--T--A--C--C--C--C--T--AC--G--G--GGAA--A--T--A--GG--A--G--T--AGACC--T--A--C-- 9031  
A--G--AC--AA-C--GA--GA--C-A--AG--A--T--A--CAG--C--C--C--T--GGAA--A--T--A--G--G--A--G--T--AGA-C--T--ACC-- 8148  
W--A--AA-T--G--AA--A-CAAA--T--A--T--M--G--AGGGAA--A--A--G--A--C--CYT--CT-- 8428  
A--A--G--AA-T--GA--G--C--CA--T--A--T--AG--A--A--A--G--GGA--C--GA-T-G--G--G--8539  
G--8996  
G--8998  
G--8982  
A--A--A--G--GCA--C--8495  
A--A--A--G--GCA--C--8492  
G--R--C--8492  
G--M--8106  
G--8096  
G--R--8102  
G--8106  
G--8107  
G--85013  
G--Y--8077  
G--T--8100  
G--8107  
G--8105  
G--G--8074  
G--8094  
G--7963  
G--8107  
G--8099  
G--T--8111  
G--8109  
G--8107  
G--G--7996  
G--R--8092  
G--7975  
G--7969  
G--7963  
G--8106  
A--A--C--G--CA--T--G--G--A--T--C--8477  
A--A--C--G--GCA--A--G--A--T--A--C--8474  
G--A--A--C--G--A--A-C-A--AT--A-TG--G--GA-A--T--A--A--T-TG-CTCG--AGC--C--CC--A--8448  
G--9000  
C--A--AG--G--CA--A--G--A--T--A--GT--G--G--GAT--G--T--GAGC--8965  
A--AG--G--CA--A--A--T--R--G--GT--G--GAT--G--T--GANC--G--A--G--8489  
A--AG--G--GCA--A--A--T--R--G--GT--G--GAT--G--T--GAAC--G--A--C--G--8792  
A--AG--G--GCA--A--A--T--R--G--GT--G--GAT--G--T--GAAC--G--A--C--G--8791  
A--AG--G--GCA--A--A--T--R--G--GT--G--GAT--G--T--GANC--G--A--C--G--8504  
A--AG--G--CA--A--A--T--G--GT--G--GAT--G--T--GANC--G--A--C--G--8984  
A--AG--G--CA--A--A--T--G--GT--G--GAT--G--T--G--A--G--AT--C--T--A--8952  
C--A--AG--G--A--CA--A--G--A--T--A--G--GGAT--G--T--GA--C--9007  
A--A--A--C--GA--GA--CA--A--A--T--A--A--C--G--AG--G--A--A--T--GG-T-GT--C--GA--C--C--C--8651

	Rev end	Nef start	
MAC.US.-.239	CCAACTCCAGAGGCTCTGCGACCTACAGAGGATT	CGA.GAAGTCCTCAGGACTGAAGCTACCTACAATATGGTGGAGCTATTTCCATGAGGCGGTCCAGGCGTCTGGAGATCTGCGACAGAGACTTTGCCGGCGCTGGGAGAC.TTATGGGAGCT	9168
Nef	P I L Q R L S A T L Q R I R # E V L R T E L T Y L Q Y G W S Y F H E A V Q A V W R S A T E T L A G A W G D # L W E T	M G G A I S M R R S R P S G D L R Q R L L R A R G E T # Y G R L	
Rev exon 2	T N T P E A L C D P T E D S # R S P Q D *		
H2A.GM.MCN13	--G-TT---T-TCAGAG-G-A-GACAGCA-CA-CTGG-A-A--CGCAG---T-G-G---CGAG-GGA--A-GA--T---T-GCA--GG-A--G-----A-----A-----G-G-G-A		8650
H2A.CI.88.UC2	--TG-CTC---AT-G.....ACAGCA-CA-G-CTGG-G-ACT-A-GGC-G---G-----CGAG-GGA--A-A-TT---A-AA-GCA-GA--G-----A-G-GG---TA-G-A		9165
H2A.DE.-BEN	--TG-CTC---T-A.....ACAGCA-CA-G-CTGG-G-ACT-A-GGC-G-C-A-G---CGAG-GGA--A-A-TT---A-AT-GC--GA-A--G-----TG---AG-A		9211
H2A.DE.-PEI2	--T-CT---T-TCAGAGAG-A-ACA-CA-CA-G-CTGG-G-ACT-AC-A-AG---G-----CGAG-GGA--A-A-T-C---T-C-TGCA-GA-A--G-----A-----A-----G-G-G-A		9167
H2A.GH.-GH1	--TG-CTC---AT-A.....ACAGCA-CA-G-CTGG-G-ACT-CA-GGAG---G-G-----G-CGAG-GGA--A-A-AT---A-AT-GCA-AGA-A--G-----A-A-A--G-GG---T-CAG-A		8631
H2A.GM.-JBY	--G-CT---G-T-TCAGAGAG-G-GACA-CA-CA-G-CTGG-G-ACT-A-GGC-G---G-G---CGAG-GGA--A-A-T---A-C-GC--GGT-A-A-G-----A-A-A--G-A-GG---G-G-G		8610
H2A.GM.87.D194	--TG-CTC---T-A.....ACAGCA-CA-G-CTGG-G-ACT-A-GGC-G---G-G---CGAG-GGA--A-A-T---GA-AT-GCA-GA--G-----A-----A-----A-G-GG---T-AG-A		8625
H2A.GW.-ALI	--G-CT---G-T-TCAGAGAG-G-GACA-CA-CA-G-CTGG-G-ACT-A-GCAG---G-G---CGAG-GGA--A-A-T---G---T-TGCA-GAT--G-----A-AAA-A-C-A-----G-G-G-A		9207
H2A.GW.-MDS	--C-CT---AT-A.....ACAGCA-CA-G-CTGG-G-ACT-AG-A-AG---T---CGAG-GGA--A-A-T---T-GC--GGT-A-A-G-----A-A-CA-G-G---G-G-G-A		8679
H2A.GW.86.FG	--G-CT---G-T-CAGAGAG-A-ACAGCA-CA-CTGG-G-CT-A-GCAG---G-G---CGAG-GGA--A-A-T---A-C-TGCA-GA-A-A-G-----A-----G-A-----G-AG-G-G		8634
H2A.GW.87.CAM2CG	--C-C---AT-A.....ACAGCA-CA-G-CTGG-G-CT-A-G-AG---T---CGAG-GGA--A-A-T---CTGCA-GG-A-A-G-----A-----G-A-----G-GA-TG		9214
H2A.SN.85.ROD	--TC-CTA---AT-A.....ACAGCA-CA-G-CTGG-G-ACT-AG-ACAG---T-T-G---CGAG-GGA--A-A-AT---C-GC--GG-A-A-G-----CA-G-G---G-AG-GTA		8648
H2AB.CI.-.7312A	--GCT-----C-CCA---CTCCTGT-TG-G---C---ATCG--GG--A-A-A--A-A-CAGC-G-GG-A-GG-----A-A-AG-A-GAC-----G-9181		
H2B.CI.-EHO	--G-----C-CCG---CT-CC-CTG-A---C-G---ATC--GG--A-A-AA--A-A-CAGCC--GG---GG-----A--GC-A-GAC--CG--GAGTC		9152
H2B.CI.88.UC1	--GC-----C-CCG---CT-TC--TG-A---C-G---ATC--GG--A-A-AA--A-A-CAAC--GG---G-----AA-A-AG-AA--C-C---A-AG--9177		
H2B.GH.86.D205	--A-CTCAACC-AT-----C-CCA---CTCCCGG-TG-G---C-G---ATC--GG--A-A-C-G-GA-G-CAGCA-GG-A-GG---A-----A-T-CG-A-AGAC-----G-9175		
H2B.JP01.KR020	--T-G-T---C-CAG---ACT-CC--TG-A---C---ATCG--GG--A-A-AA--A-A-CAGC--GG---GG---A-AAA-A-AG-AA-G-C-C-----G-8283		
H2G.CI.-.ABT96	--GC-GT---A-AA---CAG---C-AGCA-A---CAT--C-CTA--GCAG-T--TT-CAGC---TCC--GGC--A-A-TG-ACC-A-CACT--GG---CA-----A-AA-A--A-C-----A-A-8596		
H2U.FR.96.12034	--GT-GT---CA---CGT--GCACCT--A-AG-T-----CAG--A-A-A-AT---A-GT-G--A-G---CG---AGCC-A-A-G-A---A-C-----A-A-8686		
MAC.US.-.251.LA11			9164
MAC.US.-.251.32H.PJ5			9166
MAC.US.-.251.BK28			9150
MAC.US.-.MM142			8663
MAC.US.-.SMM142B			8660
MAC.US.1937A			8276
MAC.US.2065			8274
MAC.US.239.95112			8270
MAC.US.239.96114			8274
MAC.US.80035			8275
MAC.US.81035			8208
MAC.US.85013			8245
MAC.US.87082			8268
MAC.US.92050			8275
MAC.US.92077			8273
MAC.US.93057			8242
MAC.US.93062			8262
MAC.US.95058			8131
MAC.US.95086			8275
MAC.US.96016			8267
MAC.US.96020			8279
MAC.US.96072			8277
MAC.US.96081			8275
MAC.US.96093			8164
MAC.US.96123			8260
MAC.US.96135			8143
MAC.US.97009			8137
MAC.US.97074			8131
MAC.US.r80025			8274
MAC.US.r90131			8645
MNE.US.-.MNE027			8642
MNE.US.82.MNE.8			8616
SMM.SL.92.SL.92B			9168
SMM.US.-.17EC1			9168
SMM.US.-.17EFR			9168
SMM.US.-.F236.H4			9133
SMM.US.-.H9			8657
SMM.US.-.PB114.15			8960
SMM.US.-.PB1A			8959
SMM.US.-.PB1.143			8672
SMM.US.-.PB1.6P6			9152
SMM.US.-.PGM53			9111
SMM.US.-.SME543			9175
STM.US.-.STM			8819

MAC.US.-.239  
 Env  
 Nef  
 H2A.GM.MCN13  
 H2A.CI.88.UC2  
 H2A.DE.-BEN  
 H2A.DE.-PEI2  
 H2A.GH.-GH1  
 H2A.GM.-J57  
 H2A.GM.87.D194  
 H2A.GW.-ALI  
 H2A.GW.-MDS  
 H2A.GW.86.FG  
 H2A.GW.87.CAM2C6  
 H2A.SN.85.ROD  
 H2A.BN.CI.-7312A  
 H2B.CI.-EHO  
 H2B.CI.88.UC1  
 H2B.GH.86.D205  
 H2B.JP01.KR205  
 H2G.CI.-ABT96  
 H2U.FR.96.12034  
 MAC.US.-.251.1A11  
 MAC.US.-.251.32H.P15  
 MAC.US.-.251.BK28  
 MAC.US.-.MM142  
 MAC.US.-.SMM142B  
 MAC.US.1937  
 MAC.US.2065  
 MAC.US.239.95112  
 MAC.US.239.96114  
 MAC.US.80035  
 MAC.US.81035  
 MAC.US.85013  
 MAC.US.87082  
 MAC.US.92050  
 MAC.US.92077  
 MAC.US.93057  
 MAC.US.93062  
 MAC.US.95058  
 MAC.US.95086  
 MAC.US.96016  
 MAC.US.96020  
 MAC.US.96072  
 MAC.US.96081  
 MAC.US.96093  
 MAC.US.96123  
 MAC.US.96135  
 MAC.US.97009  
 MAC.US.97074  
 MAC.US.r80025  
 MAC.US.r90131  
 MNE.US.-.MNE027  
 MNE.US.82.MNE.8  
 SMM.SL.92.SL.92B  
 SMM.US.-.17EC1  
 SMM.US.-.17EFR  
 SMM.US.-.F236\_H4  
 SMM.US.-.H9  
 SMM.US.-.PB114\_15  
 SMM.US.-.PB1A  
 SMM.US.-.PB1\_143  
 SMM.US.-.PB1\_6P6  
 SMM.US.-.PGM53  
 SMM.US.-.SME543  
 STM.US.-.STM

Env gp41, gp160 end  
 CTTAGGAGAGGTGGAAGATGGACTCGCAATCCCGAGGAGGTAGACAAAGGCTTAGCTCCTCTTGTGA...GGGACAGAAA.....TACAATCAGGGACAGTATATGAATACTCCATGGAGAAAACCGCTGAAGAGAGAGAAAAATAGCATACAGAAAAAC 9329  
 L R R G G R W I L A I P R R L D K G L E L L L \*  
 L G E V E D G Y S Q S P G G L D K G L E S L S C E G Q K . . . Y N Q G Q Y M N T P W R N P A E E R E K L A Y R K  
 -GCA-C-GATC-G-GG-A---T-G---A-A-A-C-G-G-AGCA-A-G-C-A---CGG.....TC-G---G-C-T---C---C---AA---AG-G-G---GA-TTG---AGC---8811  
 G-GCA-C-CATC-G-GG-A---C---G---A---C-G-G-AGCA-AA-G-C-C---TCGG.....TC-G---G-C-T---C---C---AATG-AG-C-G---GAGT---AGC-G---9326  
 GCGC-AC-CATC-G-GG-A---G-T-A-A-A-C-G-G-AGCA-A-G-C-C---C-G.....TC-G---G-C---C-GC---AAC-A-A-C-G---GATTG-T-GC-G---9372  
 A-GG-AC-GATC-C-GG-A---T-G-T-A-A-C-G-G-GCA-A-TG-C-C---G-CGG.....TC-A---G-C-T---C---CT---A-C-GA-G-GGG-AC-TTG---AGC---9328  
 G-GCAAC-C-TC-G-GG-A---G---A-A-C-C-G-G-AGCA-AA-G-C-C---C-G.....TCGG---G-T-TC---C---C---G-AAT-AG-C-G---AAGTT---AGC-G---8792  
 GG-AC-ATC---GG---G-T-AC-A-C-G-G-AGCA-AA-TG-C-C---G.....TC-G-A---G-C-T---C---C---AAC-A-AG---GA-T-G---GC---8771  
 GCGCAAC-CATC-G-GG-A---C---G---A-A-C-G-G-AGCA-AA-G-C-C---C-G.....TC-G---G---TC---C---C---A-C-ATAG-GC-G---AATT---T-AGC-G---8786  
 G-GCA-T-G-TC-G-GA---G---A-A-C-G-G-GCA-AA-TG-C-C---A.....CGG.....TC-G---G-T-T---C---C---AAC-A-A-G---G-GA-TTG---AGC---9368  
 T-GG-AC-AT---G-GG-A---GG---A-A-C---GCA-A-G-C-C---C-G.....TC-A-A-GG-C-TC---C---C---AAC-AG-G---GATT---T---C---8840  
 -GCAAC-ATC-G-GG---T-G-AC-A-C-G-G-AGCA-A-G-C-C---CG.....TC-G---G-T-T-G-A---C---A-C-A-G-G---GA-TTG---A-C-G---8795  
 T-ACAAC-CAT---GG-A---C-T-GG---A-A-C-G-G-AGCA-A-G-C-C---CG.....TC-G---GG-C-T---C---GC---A-C-GA-G---AAT---GC-G---9375  
 T-GGAAC-ATC-G-GG-A---GG-T-A-A-C-G-G-AGCA-AA-G-C-C---G-C-G.....TC-G---G-C-T---AGG---A-C-A-G-G---AATTG---GC---8809  
 -C---GACG-CG-GG-A-CA---G---A-A-C-G-G-AGCA-A-G-C-C---G-A.....CCCG---GA-C-AG-G-GGG.....C---9294  
 -C---CG-C-G-GA-CA-T---A---C---GC-A-G-C-C---T-A.....AG-CTTTG-GA-C-AG-G-G-GG.....9265  
 -C-A---GACA-CGGA-GCA-CA---G---A---A-G-C---GC.....ATGC---AGA-C-AG-G-G-GGG.....C---9290  
 -C---CG-CG-GCA-CA---G---A---C---A-A-G-C---A.....ATGC---GG-GG-C-AG-G-G-GGG.....C---9288  
 -C-A---ACA-CG-GGCA-CG---G---C---A-C-A-G-C---GC.....TGCT---GA-C-C-AG-G-G-GGG.....C---8396  
 GG-A-G-T---G---A-G-A-A-C-G---A-CA-C-T---CG---T-A-T---C---AAC-A-A---C-G---GY-G---8757  
 GG-C-G-T---C-T-GA-T-A-A---C-G---AT---G-G-T---A-CT---GGCAAAGGC-T-G-G-A-G---TC---A-T---C---AAC-G-A---T---CA---T-AGC---8853  
 .....C---T---A---A---G---C---9325  
 .....C---T---A---G---9327  
 .....C---T---G---G---A---A---9311  
 -CAA-G---C---T---C-A---G-A-C---GA---C---8824  
 -CAA-G---C---T---C-A---G-A-C---GA---C---8821  
 .....R.....8437  
 .....R.....8435  
 .....R.....8425  
 .....R.....8431  
 .....R.....8435  
 .....T.....8436  
 .....T.....8369  
 .....T.....8406  
 .....A.....8429  
 .....T.....8436  
 .....T.....8434  
 .....R.....8403  
 .....R.....8292  
 .....R.....8436  
 .....RR.....8428  
 .....RR.....8440  
 .....R.....8438  
 .....R.....8436  
 .....R.....8325  
 .....R.....8421  
 .....R.....8304  
 .....R.....8298  
 .....R.....8292  
 .....R.....8435  
 -GG-A-G-T---T---C---A---T---T---T---A---G---G---8806  
 -GG-A---T---T---C---C---T---T---A---G---G---8803  
 -GG-AC-G-T---G---CGCA---A---C---G-C---A---TG-TAAC-AGGGAT---T-T---GTA-GG-A-T---C-C---T-AC-A-C---GCA-CAG-T-GC-G---8780  
 .....A.....9329  
 -GG-A-G-T---G---T---C-C-C-G---CCT---T-G-G---A-TC---C-C-G---A---A-C---GT-T-C---9294  
 -GG-A-G-T---GG-A---MR-AC-C-C-G---CCT---TTG-G---T-A-TC---C-T-G---AAC-AG-C---GT-T-C---8818  
 -GG-A-G-T---GG-A---A-AC-C-C-G---CCT---TTG-G---T-A-T---C-T-G---AAC-A-C---AT-T-C---9121  
 -GG-A-G-T---GG-A---A-AC-C-C-G---CCT---TTG-G---T-A-T---C-T-G---AAC-A-C---AT-T-C---9120  
 -GG-A-G-T---GG-A---MR-AC-C-C-G---CCT---TTG-G---T-A-TC---C-T-G---AAC-A-C---AT-T-C---8833  
 -GG-A-G-T---GG-A---A-AC-C-C-G---CCT---TTG-G---T-A-T---C-T-G---AAC-A-C---AT-T-C---9313  
 -GG-A-G-T---G---G---G---A---G-A-TG-A-C-T---GG---G-G---A-TC---C-C-T---AG---G-AA---T-C---9272  
 -GG-A-G-T---G---T---C-CG-C-G---G---A---G---G---CCT---T-G-G---A-TC---C-C-G---T-A-C-A-A-C---GT-T-A-C---9336  
 -G-A-G-T---G-A---CGGG-C---A---C-C-C-G---A---A---CCT---G---T-G-A-T-A-T---C-T-A---G-A-C-A-A-C-CT-GC-A---T-C-G---8980











		TATA box	3' LTR U3 end	TAR element start 3' LTR repeat start			
MAC.US.-.239	CTGGGGAGGAGCCGGTCGGGAACGCC	ACTTCTTGATGT	. . . ATAAATAC	TACTCTGTA	TTCAGTCGCTCTGCGGAGGCTGGCAGATTGAGCCCTGGGAGTTCTCCAGCACT	AGCAGGTAGAGCTGGGTGTTCCCTGCTAGACTCTCA	10071
H2A.GM.MCN13	-AT	-T-A-A-TC	G-AC-C-	TGC-T-AT	C	T-C	9589
H2A.CI.88.UC2	-AT	-T-A-A-AC	G-AC-C-	T-CC-A-	C	A	10105
H2A.DE.-BEN	-AT	-T-A-A-AC	G-AC-C-	T-C-T-AT	C	A	10151
H2A.DE.-PEI2	-GT	-A-A-C	G-AC-C-	TGC-T-AT	C	A	10086
H2A.GH.-GH1	-AT	-T-A-A-AC	G-AC-C-	TGC-TA-AT	C	A	9479
H2A.GM.-ISY	-AT	-T-A-A-TC	G-AC-C-	TGC-T-AT	C	A	9549
H2A.GM.87.D194	-AT	-T-A-AT-CTC	G-AC-C-	T-C-T-AT	C	A	9472
H2A.GW.-AL1	-AT	-T-A-A-C-C	G-AC-C-	TGC-T-AT	C	A	10146
H2A.GW.-MDS	-AT	-T-A-A-C-TC	G-AC-C-	TGC-T-AT	C	A	9525
H2A.GW.86.FG	-AT	-TT-A-A-TC	G-AC-C-	TGC-T-AT	C	A	9345
H2A.GW.87.CAM2CG	TAT	-T-A-A-C-TC	TCAC-C-	TGC-T-AT	C	A	10165
H2A.SN.85.ROD	-AT	-T-A-AT-CTC	AC-C-	T-GC-T-AT	C	G	9589
H2AB.CI.-.7312A	-AT	-TT-G	T-AAGCCTC	G-AC-C-	T-C-T-AT		10076
H2B.CI.-.EHO	-AT	-A-GA-	TCAAGACTC	G-AC-C-	T-GT-AT	C	10041
H2B.CI.88.UC1	-AT	-TT-T	T-AAA-CTC	G-AC-C-	T-C-T-AT	C	10066
H2B.GH.86.D205	-AA	-G-TT-T	T-AAA-CTC	G-AC-C-	T-C-AT	C	10070
H2B.JP01.KR020	-AT	-T-G	T-AAACCTC	G-AC-C-	TT-C-ACA-T	AC	9171
H2G.CI.-.ABT96	TAT	-GA-T-T	-AAAA-CTC	AG		C	9499
H2U.FR.96.12034	GCAA	-GA-GAG-CA-G	T-AAATGATC	G-AC-C-	TC-GCA		9600
MAC.US.-.251_1A11		A		T-A	G		10066
MAC.US.-.251_32H_PJ5		A					10069
MAC.US.-.251_BK28		A					10041
MAC.US.-.SMM142		T-A			C		9563
MAC.US.-.SMM142B		T-A					9560
MAC.US.1937							9054
MAC.US.2065							9054
MAC.US.239.95112							9044
MAC.US.239.96114							9048
MAC.US.80035							9054
MAC.US.81035							9055
MAC.US.85013							8986
MAC.US.87082			AT				9028
MAC.US.92050							9048
MAC.US.92077							9055
MAC.US.93057							9051
MAC.US.93062							9020
MAC.US.95058							9042
MAC.US.95086							8909
MAC.US.96016							9053
MAC.US.96020							9047
MAC.US.96072							9057
MAC.US.96081							9044
MAC.US.96093							9053
MAC.US.96123							8944
MAC.US.96135							9038
MAC.US.97009							8867
MAC.US.97074							8858
MAC.US.r80025							8909
MAC.US.r90131							9052
MNE.US.-.MNE027		T	A				9548
MNE.US.82.MNE_8		A					9545
SMM.SL.92.SL.92B	TCT	-C-G-GGAACG	C-AAACTC	AGT	AC-C-TT-GC	A-C-A-C-A-G-T-T-T-A-A-A	9531
SMM.US.-.17EC1							10071
SMM.US.-.17EFR							10071
SMM.US.-.F236_H4		T-CT	AT-CTC	CA			10033
SMM.US.-.H9		T-CT	AT-CTC	CA	A	G	9554
SMM.US.-.PB114_15		T-CT	G-AT-CTC	CA	A		9879
SMM.US.-.PB1A		T-CT	G-AT-CTC	CA	A		9878
SMM.US.-.PB1_143		T-CT	AT-CTC	CA	A		9591
SMM.US.-.PB1_6P6		T-CT	G-AT-CTC	CA	A-A		10071
SMM.US.-.PGM53		T-CT	AT-CTC	CA			10014
SMM.US.-.SME543		T-CT	AT-CTC	CA			10080
STM.US.-.STM		A-T-G	TT-A-TC	CA-C-T-A-A	A		9725

	TAR element end_	Poly-A signal	3' LTR R repeat end_	3' LTR U5 start				
MAC.US.-.239	CCAGCACTTGCCGGTCTGGGCAGAGTGACTCCACGCTTGCTGCTTAAA.	.GCCCTCTCAATAAAG.	CTGCCAT.	TTTAGAAGTAAGC.	TAGTGTGTGTTCCCATCTCTCCTAGCCGCCCTGGTCAACTCGGTACTCAA.T.AATA.AGAAGACCTGGTCTGTT10233			
H2A.GM.MCN13	--G-TG-CA	C-G	A-A-T	G	C-TTA	T	9712	
H2A.CI.88.UC2	-TG-C-CA	C-G	A	T	A	C-TTA	T	10266
H2A.DE.-.BEN	T-A	C-G	A	T	G	C-TTA	T	10313
H2A.DE.-.PEI2	T-A	C-G	A	T	G	C-TTA	T	10172
H2A.GH.-.GH1	..TG-..C-..C-G	..AGA-CTC-T	..G	..C	..TTA	..T	9479	
H2A.GM.-.ISY	..TG-..C-..C-G	..AGA-CTC-T	..G	..C	..TTA	..T	9636	
H2A.GM.87.D194	..TG-..CA-..C-G	..A	..T	..A	..C-G-TTA-AG	..T	9472	
H2A.GW.-.ALJ	..TG-..CA-..C-G	..A	..T	..A	..C-G-TTA-AG	..T	10307	
H2A.GW.-.MIS	..TG-..C-..C-G	..A	..T	..A	..C-TTA	..T	9525	
H2A.GW.86.FG	..TG-..C-..C-G	..A	..T	..A	..C-TTA	..T	9431	
H2A.GW.87.CAM2CG	..TG-..CA-..C-G	..A	..T	..A	..C-TTA	..T	10326	
H2A.SN.85.ROD	..TG-..CA-..C-G-C	..AA-C-T	..G	..C	..TTA	..T	9670	
H2AB.CI.-.7312A	..C-G	A-T	A	A	C-TCA	T	10236	
H2B.CI.-.EHO	..C-G	A-T	A	TA	C-ACA	T	10196	
H2B.CI.88.UC1	..C-G	A-T	A	A	C-TTG	T	10225	
H2B.GH.86.D205	..A-A-T	C-G	A	T	A	C-TA	T	10223
H2B.JP01.KR020	..C-G	A-T	TA	A	C-TTG	T	9330	
H2G.CI.-.ABT96	..C-G	C-AC-A	T	AA	A	GT	9599	
H2U.FR.96.12034	..C-G	A-T	A	A	C-TCA	T	9607	
MAC.US.-.251_1A11	..G	A				T	10228	
MAC.US.-.251_32H_PJ5	..A	G				GG	10231	
MAC.US.-.2613_BK28	..G	A				GG	9203	
MAC.US.-.SMM142	..G	A				GG	9645	
MAC.US.-.SMM142B	..G	A					9599	
MAC.US.1937	..G	A					9054	
MAC.US.2065	..G	A					9054	
MAC.US.239.95112	..G	A					9044	
MAC.US.239.96114	..G	A					9048	
MAC.US.80035	..G	A					9054	
MAC.US.81035	..G	A					9055	
MAC.US.85013	..G	A					8986	
MAC.US.87082	..G	A					9028	
MAC.US.92050	..G	A					9048	
MAC.US.92077	..G	A					9055	
MAC.US.93057	..G	A					9051	
MAC.US.93062	..G	A					9020	
MAC.US.95058	..G	A					9042	
MAC.US.95086	..G	A					8909	
MAC.US.96016	..G	A					9053	
MAC.US.96020	..G	A					9047	
MAC.US.96072	..G	A					9057	
MAC.US.96081	..G	A					9044	
MAC.US.96093	..G	A					9053	
MAC.US.96123	..G	A					8944	
MAC.US.96135	..G	A					9038	
MAC.US.97009	..G	A					8867	
MAC.US.97074	..G	A					8858	
MAC.US.r80025	..G	A					8909	
MAC.US.r90131	..G	A					9052	
MNE.US.-.MNE027	..T	G	A		C	C	T	9664
MNE.US.82.MNE_8	..G	A		T				9627
SMM.SL.92.SL92B	..ACA-AT-T-CG	A	G	A	C	T	C	9613
SMM.US.-.17EC1	..G	A						10233
SMM.US.-.F236_H4	..G	A		A	A	T	T	10195
SMM.US.-.H9	..RG	A	G	G				9637
SMM.US.-.PB114_15	..G	A			C		T	9996
SMM.US.-.PB1A	..G	A	AG		C		T	9996
SMM.US.-.PB1_143	..RG	A	AG		C		T	9674
SMM.US.-.PB1_6P6	..G	A			C		T	10232
SMM.US.-.PGM53	..G	A	GC		A		T	10176
SMM.US.-.SME543	..G	A			A		T	10243
STM.US.-.STM	..A	T	G	T	TAT	A	G	9878

