

## IV

---

**Primate Lentivirus Complete Genomes**


---

Introduction .....	405
PLV phylogenetic tree .....	407
Table of sequences in HIV-1/HIV-2/SIV complete genome alignments .....	409
Nucleotide alignments of PLV complete genomes .....	411

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

HIV-1: One sequence per subtype from M, N, and O groups and CRFs 1-4

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

SIV: Sequences from viruses infecting Chimpanzees, *Pan troglodytes troglodytes* and *P.t. schweinfurtii* (cpz); the subspecies of African green monkeys, also known as Savannah monkeys, *Chlorocebus* (previously *Cercopithecus*) *aethiops aethiops* (gri), *C.a. pygerythrus* (ver), *C.a. tantalus* (tan) and *C.a. sabaues* (sab); Drill-mangabeys including Sooty mangabeys, *Cercocebus atys* (smm) and Macaques infected with sooty mangabey virus (mac, stm, mne), and Red-capped mangabey, *C. torquatus* (rcm); Mandrills, *Mandrillus sphinx*s (mnd-1, mnd-2) and Drills, *M. leucophaeus* (drl); L'Hoest monkeys, *Cercopithecus l'hoesti l'hoesti* (lst) and Sun-tailed monkey, *C.l'h. solatus* (sun); Mona monkeys, *C. mona mona* (mon) and Dent's monkey, *C.m. denti* (den); De Brazza's monkey, *C. neglectus* (deb); Greater Spot-nosed monkey, *C. nictitans nictitans* (gsn); Moustached monkey, *C. cephus cephus* (mus); Syke's monkey, *C. albogularis* (syk); and Guereza colobus monkeys, *Colobus guereza* (col).

There are currently 907 primate lentivirus sequences greater than 5,000 bases long in our database, and the majority (884) of them are longer than 7,000 bases. The 907 longer than 5,000 bases are mostly HIV-1 M group viruses:

803 are HIV-1 (775 M, 25 O, 3 N group)

25 are HIV-2

70 are SIVs (19 SIVsmm, 8 SIVmac, 7 SIVcpz)

7 are SHIV

2 are synthetic DNA

For the Primate lentivirus complete genome DNA alignment, one of each of the major subtypes of the HIV-1 M group was chosen, plus one N group and 2 O group sequences. Six of the 7 SIVcpz were included, SIVcpz-CAM5 was not included because it is quite similar to CAM3. Two HIV-2 A group and 2 HIV-2 B group were included along with the single genomes sequenced to date for HIV-2 G and HIV-2 U. Five SIVsmm genomes (mac, smm, mne, stm) were included in this alignment and more are available in the HIV-2/SIVsmm alignment. For the rest of the SIVs, all complete genomes were included. The principle in deciding which sequences to include in this alignment was to provide a representative sampling of all primate lentiviruses which have had complete or nearly complete genomes sequenced. Many HIV-1, HIV-2, SIVsmm and SIVmac genomes were not included, but they can be found in the HIV-1/SIVcpz and HIV-2/SIVsmm alignments.

Table 1 lists the sequence name, accession number, isolation country, first author, and publication for each of the sequences in this alignment. These genomes represent the full genome breadth of diversity discovered to date in the primate lentivirus lineage. There are, however, more sequences from shorter and

partial genome and gene sequences available from some more simian species, but these are not included in this alignment because they are too short. They are available through our search interface and some are in the protein alignments in section VII. Non-primate lentiviruses such as equine infectious anemia virus (EIAV) and caprine arthritis/encephalitis virus (CAEV) are more distantly related to each other and to the primate lentiviruses, and are not included in this alignment.

The tree in Figure 1 is a disclaimer of the apparent phylogenetic relationships between the PLVs. As has been pointed out by several authors, several of the PLV sequences appear to be recombinant. For instance, SIVrcm has been shown to display discordant clustering in different genomic regions (Beer *et al.*, 2001), SIVcpz has been described to have hybrid origins (Bailes *et al.*, 2003), and SIVsab was described as mosaic already ten years ago (Jin *et al.*, 1994). As can be seen in Figure 1, trees based on *gag*, *pol* and *env* gene sequences differ in their cladistic classification of the PLVs. The topology shown is derived from the complete genome alignment in this section and the thin line arrows indicate alternative clusterings as inferred in *gag*, *pol* and *env* gene trees. Note that this is not intended to be an exhaustive analysis of potential recombination, but merely a display of the problematic relationships among PLVs. The direction of the arrows is in most cases arbitrary and does not indicate direction of potential crossover events or cross species transmissions. Although important, similar problems within the HIVs, such as the recombinant structure of HIV-1 group N genomes (Roques *et al.*, 2004) and the multiphyletic histories of the CRFs, are not shown in Figure 1. The resulting network structure is likely the result of multiple reasons, including true recombination between SIVs, lack of appropriate parental representatives, effects of mutational rate differences, possible convergent evolution, and inadequate phylogenetic reconstruction methods. For instance, early attempts to describe mosaic structures of some SIVs has been revised when virus from other simian species have become available. Also note that even though bootstrap support is strong for many clusters, sequences within those clusters are involved in conflicting topologies, and is thus not a method to reveal uncertainties resulting from any of these problems. Therefore, any tree containing PLV sequences should be interpreted with caution. Recently, several reviews and analyses on the PLVs have been published (*e.g.*, Bailes *et al.*, 2002; Clewley, 2004; Foley, 2000; Peeters and Cournaud, 2002; Salemi *et al.*, 2003).

The tree in Figure 1, as well as the separate *gag*, *pol* and *env* gene trees and the gapstripped alignments the trees were calculated from are available at our website, <http://www.hiv.lanl.gov>.

The alignment was based on the previous version published last year, by a hidden Markov model using HMMER and manual editing using the programs BioEdit and Se-Al. It is codon-aligned, meaning that the correct translation reading frame has been maintained as much as possible; in the case of overlapping frames, obviously one will not translate correctly. We have tried to “reset” the alignment so that each new gene starts in the first frame; this means in some places empty columns have been inserted. The annotation is unchanged from last year’s compendium. The HIV-2/SIVsmm vpx gene is postulated to be a duplication of the vpr gene (Tristem *et al.*, 1990) and thus there may be two alternative alignments of this region of the genome, as there are for the duplicated stem-loops of the TAR element.

The annotation is based on known protein coding regions in HIV-1 and on annotations found in SIV sequence database entries. The protein cleavage sites that create Gag p17, Gag p24 and other mature peptides from the Gag and Gag-Pol precursor polyproteins have been experimentally determined for HIV-1 and at least one strain of HIV-2, the study of analogous cleavages in SIV polyproteins have not been published. Two representative genomes have been translated; HIV-1 subtype B strain HXB2, and SMM-239. The translations are provided as a visual aid for finding landmarks in the genomes.

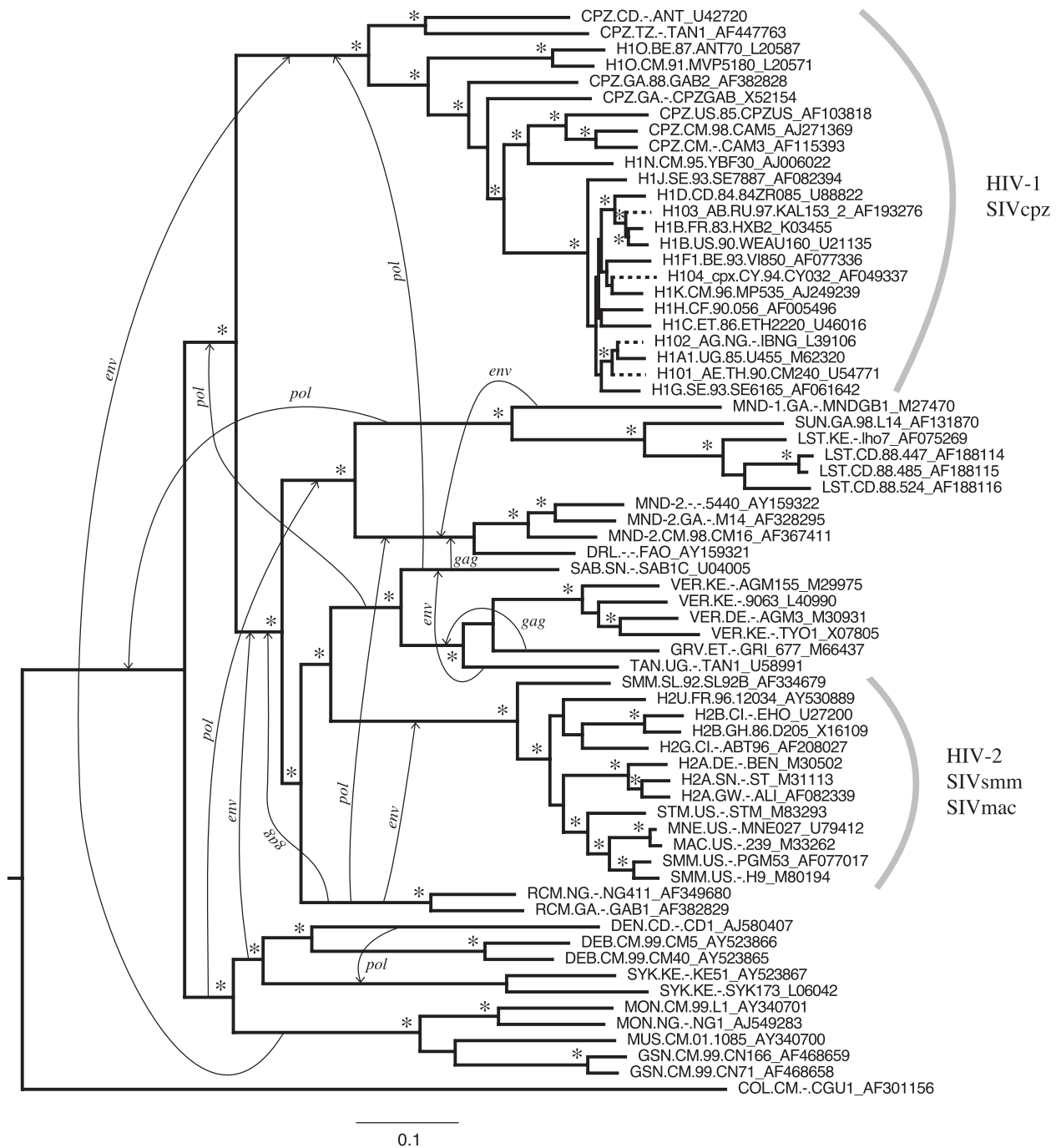


Figure 1. Genetic relationships among Primate Lentiviruses (PLVs). The tree shows the inferred relationships in the following PLV complete genome alignment. The thin line arrows indicate the main differences inferred from the gag, pol and env gene fragments from the same alignment. Dashed tree branches indicate inferred positions for known recombinant HIV-1 circulating forms. The resulting network should therefore not be interpreted as a singular phylogenetic history of the PLVs. Asterisks indicate bootstrap support  $\geq 70\%$  in the complete alignment. The scale bar

at the bottom shows genetic distance according to a F84 model with relative rates. The tree shown was constructed from a gapstripped version of the DNA alignment in this section of the compendium, leaving 5327 unambiguously aligned positions. A neighbor-joining tree, calculated using DNADIST (F84 model) and NEIGHBOR in the PHYLIP package, was used as a guide to calculate relative site rates. The relative site rates were estimated using a generalized version of DNARates (Korber et al., 2000) with eight categories (0.133; 0.287; 2.373; 6.826; 19.634; 56.472; and 115.653). Finally, these rates were given to fastDNAML (version 1.2.2) and a maximum likelihood tree was inferred from the gapstripped alignment (transition/transversion ratio = 2). The tree space was searched by both Jumble and Global Rearrangements (G33). Bootstrap support was calculated using SEQBOOT with the above 8 rate categories on 1000 resampled sets, and DNADIST and NEIGHBOR in PHYLIP (version 3.6.2). The separate gag, pol and env alignments were created using a new version of GeneCutter (see [www.hiv.lanl.gov](http://www.hiv.lanl.gov)) and trees were calculated the same way as in the complete alignment with optimized site rates (not shown). The Colobus SIVcol sequence, which is the most divergent PLV sequence, was used as outgroup to visualize the tree using TreeEdit (version 1.0) and Adobe Illustrator.

## References

- Bailes, E., Chaudhuri, R. R., Santiago, M. L., Bibollet-Ruche, F., Hahn, B. H., and Sharp, P. M. (2002). The evolution of primate lentiviruses and the origins of AIDS. In *The molecular epidemiology of human viruses* (T. Leitner, Ed.). Kluwer Academic Publishers, Boston.
- Bailes, E., Gao, F., Bibollet-Ruche, F., Courgnaud, V., Peeters, M., Marx, P., Hahn, B., and Sharp, P. M. (2003). Hybrid origin of SIV in Chimpanzees. *Science* **300**, 1713.
- Beer, B., Foley, B., Kuiken, C., Tooze, Z., Goeken, R. M., Brown, C. R., Hu, J., St. Claire, M., Korber, B., and Hirsch, V. M. (2001). Characterization of novel simian immunodeficiency viruses from Red-Capped mangabeys from Nigeria (SIVrcmNG409 and -NG411). *Journal of Virology* **75**, 12014–12027.
- Clewley, J. P. (2004). Enigmas and paradoxes: the genetic diversity and prevalence of the primate lentiviruses. *Current HIV Research* **2**, 113–125.
- Foley, B. (2000). An overview of the molecular phylogeny of lentiviruses. In *HIV Sequence Compendium* (C. Kuiken, B. Foley, B. Hahn, P. Marx, F. McCutchan, J. Mellors, J. I. Mullins, J. Sodroski, S. Wolinsky, and B. Korber, Eds.), pp. 35–43. Theoretical Biology & Biophysics, LANL, Los Alamos, NM.
- Jin, M. J., Hui, H., Robertson, D. L., Muller, M. C., Barre-Sinoussi, F., Hirsch, V. M., Allan, J. S., Shaw, G. M., Sharp, P. M., and Hahn, B. (1994). Mosaic genome structure of simian immunodeficiency virus from West African green monkeys. *EMBO Journal* **13**, 2935–2947.
- Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B. H., Wolinsky, S., and Bhattacharya, T. (2000). Timing the ancestor of the HIV-1 pandemic strains. *Science* **288**, 1789–1796.
- Peeters, M., and Courgnaud, V. (2002). Overview of primate lentiviruses and their evolution in non-human primates in Africa. In *HIV Sequence Compendium* (C. Kuiken, B. Foley, E. Freed, B. Hahn, P. Marx, F. McCutchan, J. Mellors, S. Wolinsky, and B. Korber, Eds.), pp. 2–23. Theoretical Biology & Biophysics, LANL, Los Alamos, NM.
- Roques, P., Robertson, D. L., Souquiere, S., Apetrei, C., Nerrienet, E., Barre-Sinoussi, F., Muller-Trutwin, M., and Simon, F. (2004). Phylogenetic characteristics of three new HIV-1 N strains and implications for the origin of group N. *AIDS* **18**, 1371–1381.
- Salemi, M., De Oliveira, T., Courgnaud, V., Moulton, V., Holland, B., Cassol, S., Switzer, W. M., and Vandamme, A.-M. (2003). Mosaic genomes of the six major primate lentivirus lineages revealed by phylogenetic analyses. *Journal of Virology* **77**, 7202–7213.
- Tristem, M., Marshall, C., Karpas, A., Petrik, J., and Hill, F. (1990). Origin of vpx in lentiviruses. *Nature* **347**, 341–342.

**Table 1: Table of Primate Lentivirus (PLV) Sequences in the Complete Genome Alignments**

Name	Accession	Country	Author	Reference
H1B.FR.83.HXB2	K03455	France	Wong-Staal, F	<i>Nature</i> <b>313</b> (6000):277–84 (1985)
H1A1.UG.85.U455	M62320	Uganda	Oram, JD	<i>ARHR</i> <b>6</b> (9):1073–8 (1990)
H1B.US.90.WEAU160	U21135	U.S.A.	Tozser, J	<i>FEBS Lett</i> <b>281</b> (1–2); 77–80 (1991)
H1C.ET.86.ETH2220	U46016	Ethiopia	Salminen, MO	<i>ARHR</i> <b>12</b> (14):1329–39 (1996)
H1D.CD.84.84ZR085	U88822	D.R.C.	Gao, F	<i>J Virol</i> <b>72</b> (7):5680–98 (1998)
H1F1.BE.93.VI850	AF077336	Belgium	Laukkanen, T	<i>Virology</i> <b>269</b> (1):95–104 (2000)
H1G.SE.93.SE6165	AF061642	Sweden	Laukkanen, T	<i>Virology</i> <b>247</b> (1):22–31 (1998)
H1H.CF.90.056	AF005496	C.A.R.	Murphy, E	<i>ARHR</i> <b>9</b> (10):997–1006 (1993)
H1J.SE.93.SE7887	AF082394	Sweden	Laukkanen, T	<i>ARHR</i> <b>15</b> (3):293–7 (1999)
H1K.CM.96.MP535	AJ249239	Cameroon	Peeters, M	<i>ARHR</i> <b>16</b> (2):139–51 (2000)
H101_AE.TH.90.CM240	U54771	Thailand	Laukkanen, T	<i>J Virol</i> <b>70</b> (9):5935–43 (1996)
H102_AG.NG.-.IBNG	L39106	Nigeria	Howard, TM	<i>ARHR</i> <b>10</b> (12):1755–7 (1994)
H1N.CM.95.YBF30	AJ006022	Cameroon	Simon, F	<i>Nat Med</i> <b>4</b> (9):1032–7 (1998)
H1O.BE.87.ANT70	L20587	Belgium	Vanden Haesevelde, M	<i>J Virol</i> <b>68</b> (3):1586–96 (1994)
H1O.CM.91.MVP5180	L20571	Cameroon	Gurtler, LG	<i>J Virol</i> <b>68</b> (3):1581–5 (1994)
CPZ.CD.-.ANT	U42720	D.R.C.	Vanden Haesevelde, M	<i>Virology</i> <b>221</b> (2):346–50 (1996)
CPZ.CM.-.CAM3	AF115393	Cameroon	Corbet, S	<i>J Virol</i> <b>74</b> (1):529–34 (2000)
CPZ.GA.88.GAB2	AF382828	Gabon	Gao, F	Unpublished
CPZ.GA.-.CPZGAB	X52154	Gabon	Huet, T	<i>Nature</i> <b>345</b> (6273):356–9 (1990)
CPZ.US.85.CPZUS	AF103818	U.S.A.	Gao, F	<i>Nature</i> <b>397</b> (6718):436–41 (1999)
CPZ.TZ.-.TAN1	AF447763	Tanzania	Santiago, ML	<i>J Virol</i> <b>77</b> (3):2233–2242 (2003)
H2A.GW.-.ALI	AF082339	Guinea-bissau	Azevedo-Pereira, JM	Unpublished
H2A.DE.-.BEN	M30502	Germany	Kirchhoff, F	<i>Virology</i> <b>177</b> (1):305–11 (1990)
H2A.SN.-.ST	M31113	Senegal	Kumar, P	<i>J Virol</i> <b>64</b> (2):890–901 (1990)
H2B.GH.86.D205	X16109	Ghana	Dietrich, U	<i>Nature</i> <b>342</b> (6252):948–50 (1989)
H2B.CI.-.EHO	U27200	Cote d'ivoire	Rey-Cuille, MA	<i>Virology</i> <b>202</b> (1):471–6 (1994)
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):666–672 (2004)
MAC.US.-.239	M33262	U.S.A.	Kestler, H	<i>Science</i> <b>248</b> (4959):1109–12 (1990)
SMM.SL.92.SL92B	AF334679	Sierra Leone	Chen, Z	<i>J Virol</i> <b>70</b> (6):3617–27 (1996)
SMM.US.-.H9	M80194	U.S.A.	Courgnaud, V	<i>J Virol</i> <b>66</b> (1):414–9 (1992)
STM.US.-.STM	M83293	U.S.A.	Novembre, FJ	<i>Virology</i> <b>186</b> (2):783–7 (1992)
SAB.SN.-.SAB1C	U04005	Senegal	Jin, MJ	<i>EMBO J</i> <b>13</b> (12):2935–47 (1994)
TAN.UG.-.TAN1	U58991	Uganda	Stivahtis, GL	<i>Virology</i> <b>228</b> (2):394–9 (1997)
VER.KE.-.AGM155	M29975	Kenya	Johnson, PR	<i>J Virol</i> <b>64</b> (3):1086–92 (1990)
VER.KE.-.9063	L40990	Kenya	Hirsch, VM	<i>J Virol</i> <b>69</b> (2):955–67 (1995)
VER.DE.-.AGM3	M30931	Germany	Baier, M	<i>Virology</i> <b>176</b> (1):216–21 (1990)
VER.KE.-.TYO1	X07805	Kenya	Fukasawa, M	<i>Nature</i> <b>333</b> (6172):457–61 (1988)
COL.CM.-.CGU1	AF301156	Cameroon	Courgnaud, V	<i>J Virol</i> <b>75</b> (2):857–66 (2001)
DEN.CD.-.CD1	AJ580407	D.R.C.	Saragosti, S	Unpublished
GRV.ET.-.GRI_677	M66437	Ethiopia	Fomsgaard, A	<i>Virology</i> <b>182</b> (1):397–402 (1991)
GSN.CM.99.CN166	AF468659	Cameroon	Courgnaud, V	<i>J Virol</i> <b>76</b> (16):8298–309 (2002)
GSN.CM.99.CN71	AF468658	Cameroon	Courgnaud, V	<i>J Virol</i> <b>76</b> (16):8298–309 (2002)
DRL.-.x.FAO	AY159321		Hu, J	<i>J Virol</i> <b>77</b> (8):4867–4880 (2003)
RCM.GA.-.GAB1	AF382829	Gabon	Gao, F	<i>Science</i> <b>300</b> (5626):1713 (2003)
RCM.NG.-.NG411	AF349680	Nigeria	Beer, BE	<i>J Virol</i> <b>75</b> (24):12014–27 (2001)
MND-2.-.5440	AY159322		Hu, J	<i>J Virol</i> <b>77</b> (8):4867–4880 (2003)
MND-2.CM.98.CM16	AF367411	Cameroon	Takehisa, J	<i>ARHR</i> <b>17</b> (12):1143–54 (2001)
MND-2.GA.-.M14	AF328295	Gabon	Souquiere, S	<i>J Virol</i> <b>75</b> (15):7086–96 (2001)
MNE.US.-.MNE027	U79412	U.S.A.	Kimata, JT	<i>J Virol</i> <b>72</b> (1):245–56 (1998)

LST.CD.88.447	AF188114	D.R.C.	Beer, BE	<i>J Virol</i> <b>74</b> (8):3892–8 (2000)
LST.CD.88.485	AF188115	D.R.C.	Beer, BE	<i>J Virol</i> <b>74</b> (8):3892–8 (2000)
LST.CD.88.524	AF188116	D.R.C.	Beer, BE	<i>J Virol</i> <b>74</b> (8):3892–8 (2000)
LST.KE.-.lho7	AF075269	Kenya	Hirsch, VM	<i>J Virol</i> <b>73</b> (2):1036–45 (1999)
SUN.GA.98.L14	AF131870	Gabon	Beer, BE	<i>J Virol</i> <b>73</b> (9):7734–44 (1999)
MND-1.GA.-.MNDGB1	M27470	Gabon	Tsujimoto, H	<i>Nature</i> <b>341</b> (6242):539–41 (1989)
MON.CM.99.L1	AY340701	Cameroon	Courgnaud, V	<i>J Virol</i> <b>77</b> (23):12523–12534 (2003)
MON.NG.-.NG1	AJ549283	Nigeria	Barlow, KL	<i>J Virol</i> <b>77</b> (12):6879–88 (2003)
MUS.CM.01.1085	AY340700	Cameroon	Courgnaud, V	<i>J Virol</i> <b>77</b> (23):12523–12534 (2003)
DEB.CM.99.CM40	AY523865	Cameroon	Bibollet-Ruche, F	<i>J Virol</i> <b>78</b> (14):7748–7762 (2004)
DEB.CM.99.CM5	AY523866	Cameroon	Bibollet-Ruche, F	<i>J Virol</i> <b>78</b> (14):7748–7762 (2004)
SYK.KE.-.KE51	AY523867	Kenya	Bibollet-Ruche, F	<i>J Virol</i> <b>78</b> (14):7748–7762 (2004)
SYK.KE.-.SYK173	L06042	Kenya	Hirsch, VM	<i>J Virol</i> <b>67</b> (3):1517–28 (1993)





```

/ 5' LTR U3 start
H1B.FR.83.HXB2 .TGGAAGGGCTAATTCACCTCCCAACGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTG 129
H1A1.UG.85.U455 ..... 0
H1B.US.90.WEAU160 ..... 129
H1C.ET.86.ETH2220 ..... 0
H1D.CD.84.84ZR085 ..... 0
H1F1.BE.93.VI850 ..... 0
H1G.SE.93.SE6165 ..... 0
H1H.CF.90.056 ..... 0
H1J.SE.93.SE7887 ..... 0
H1K.CM.96.MP535 ..... 0
H1O1.AE.TH.90.CM240 ..... 0
H1O2.AG.NG.-.IBNG ..... 0
H1N.CM.95.YBF30 ..... 0
H1O.BE.87.ANT70 C-----T---T---T---TAA---GC---A---G---T---G-G-TA---T-G-A-T-----G-----A---A-C---G-TC-----A--- 130
H1O.CM.91.MVP5180 C---T---T---T---T---TAAG---GC---A---G---C---A-T---T-G-A-T-----G---TGT-----G-A---ACCT---TC-----A--- 130
CPZ.CD.-.ANT ..... 0
CPZ.CM.-.CAM3 ..... 0
CPZ.GA.88.GAB2 ..... 0
CPZ.GA.-.CPZGAB ..... 129
CPZ.US.85.CPZUS ..... 129
CPZ.TZ.-.TAN1 ..... 0
H2A.GW.-.ALI ..... 129
H2A.DE.-.BEN ..... 129
H2A.SN.-.ST ..... 0
H2B.GH.86.D205 ..... 129
H2B.CI.-.EHO ..... 129
H2G.CI.-.ABT96 ..... 0
H2U.FR.96.12034 ..... 0
MAC.US.-.239 ..... 129
SMM.SL.92.SL92B ..... 0
SMM.US.-.H9 ..... 0
STM.US.-.STM ..... 0
SAB.SN.-.SAB1C ..... 129
TAN.UG.-.TAN1 ..... 129
VER.KE.-.AGM155 ..... 129
VER.KE.-.9063 ..... 129
VER.DE.-.AGM3 ..... 0
VER.KE.-.TYO1 ..... 0
COL.CM.-.CGU1 ..... 0
DEN.CD.-.CD1 ..... 0
GRV.ET.-.GRI_677 ..... 129
GSN.CM.99.CN166 ..... 0
GSN.CM.99.CN71 ..... 0
DRL.-.-.FAO ..... 0
RCM.GA.-.GAB1 ..... 0
RCM.NG.-.NG411 ..... 0
MND-2.-.-.5440 ..... 0
MND-2.CM.98.CM16 ..... 0
MND-2.GA.-.M14 ..... 0
MNE.US.-.MNE027 ..... 0
LST.CD.88.447 ..... 0
LST.CD.88.485 ..... 0
LST.CD.88.524 ..... 0
LST.KE.-.lho7 ..... 129
SUN.GA.98.L14 ..... 129
MND-1.GA.-.MNDGB1 ..... 0
MON.CM.99.L1 ..... 0
MON.NG.-.NG1 ..... 0
MUS.CM.01.1085 ..... 0
DEB.CM.99.CM40 ..... 0
DEB.CM.99.CM5 ..... 0
SYK.KE.-.KE51 ..... 0
SYK.KE.-.SYK173 ..... 26

```



PLV  
complete genomes



H1B.FR.83.HXB2	GATGGTGTACAAGCTAGTACCAGTTGAGCCAGAGAAGTTAGAAGAA...GCCAACAAAGGAGAGAACACCAGCTTGTACACCTGTGAGC...CTGCATGGAATGGATGACCCGGAGAGAGAAGT	250
H1A1.UG.85.U455	.....	0
H1B.US.90.WEAU160	-----T-----G-A-----G-----T-TG-----A-----A-----G-----AT---A-----	250
H1C.ET.86.ETH2220	.....	0
H1D.CD.84.84ZR085	.....	0
H1F1.BE.93.VI850	.....	0
H1G.SE.93.SB6165	.....	0
H1H.CF.90.056	.....	0
H1J.SE.93.SE7887	.....	0
H1K.CM.96.MP535	.....	0
H101.AE.TH.90.CM240	.....	0
H102.AG.NG.-.IBNG	.....	0
H1N.CM.95.YBF30	.....	0
H10.BE.87.ANT70	-----TG-TT--A-----GTCAGA--AG-GC--AG-CTA-GAG-T-C-T-T---GGG-T-ATC-CC-G-T-A-CAT-T...GCC-----T-T--A--TA-AC-T-A---A-	254
H10.CM.91.MVP5180	-----TG-TT--A-G-----GTCAGA--AG-GC--GAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC---T-A-CTT-T...AAT-----GCT--G--TG-AC-CG-G--GA-	254
CPZ.CD.-.ANT	.....	0
CPZ.CM.-.CAM3	.....	0
CPZ.GA.88.GAB2	.....	0
CPZ.GA.-.CPZGAB	-----T--A---G--CC-GACAGAG--C-G---C---TG-----T---A-T-C---G-T--CA-AT-T...-A---G---A--TGA--C-A--G--	250
CPZ.US.85.CPZUS	-GC-----TT--T---C--TC-CACAGA--G-AG---GCGG...-T-----T---C-A-TAC-C-G---CA-T---A-----A--TGAAC-TG-C---	250
CPZ.TZ.-.TAN1	.....	0
H2A.GW.-.ALI	-G---CTG-GG-----A--T...GTCCTC-C-----GGG-GG-CACT---CTCA-T--C--C-----A--ACAA...ACAAGCA-GCAT-----A-C-TG-G--GAC	244
H2A.DE.-.BEN	-G---CTG-GG-----T---A-A...CTC-C-C---G...AG-GG--AT--GC-A-T---AG-----A-CACAA...ACAAGCA--CAT-----TGA-C-TG-G--GAC	244
H2A.SN.-.ST	.....	0
H2B.GH.86.D205	-T---CTG-GG---G-----A---T---CCAGC--CGACC...CGAG-GG-G-AG--A-C-CATT--C-AA-G---G-CACAG...A-CTCCTC-TG-----ATCC-TG-G--GAC	247
H2B.CI.-.EHO	-C---CTA-GG-----G---A-AA-T...-T-A---C---C-AG-GG-C-AG--A-C-CATT-TC--G-G--T--A-CACAG...ACCTCC-C-TG-----CC-CGA--GAC	244
H2G.CI.-.ABT96	.....	0
H2U.FR.96.L2034	.....	0
MAC.US.-.239	-C---CTA-GG--AT---C--T-AA-T...GTA-C---T-G...-AC-GG-G-AT--G-GCATTAT--AA-G--T--A-CTCAA...ACTTCCCAGTG-----TTG-G---G--	244
SMM.SL.92.SL92B	.....	0
SMM.US.-.H9	.....	0
STM.US.-.STM	.....	0
SAB.SN.-.SAB1C	-----TT--A-G---T-G--C...CT--C--G--G...T-G--TA-T---CA-T--C--C-G--T---CTCAA...G--GCCTATGA-----G-CTG--A--GAC	244
TAN.UG.-.TAN1	-C-TT---TTG--T---G-G-A-T...GTC-CCC-G-----C-GG-T-AG...GACA-T-T---C-C-----CCCAG...A-TG-AT-GGAAAGC-----TTG-A--GAC	241
VER.KE.-.AGM155	-C-TC---T-----G---CC...CT-CAT-G-----AG-A-C-T-T--A-GGCATT---G-----A-CACAAGTGCATGAAG-CCCTGAT-G-ATAAATC-TG-----A-	250
VER.KE.-.9063	-C-TT---T-----G-A---T-CAT---G...-AG-G-CCT-T---GACATTA---GCG--T---CACAAATAAAAGA-G-CCCCGAT-GCAT-AGTC-TG-----GAC	250
VER.DE.-.AGM3	.....	0
VER.KE.-.TY01	.....	0
COL.CM.-.CGU1	.....	0
DEN.CD.-.CD1	.....	0
GRV.ET.-.GRI_677	-C-TC---TT-----C---G-C...CT-CAT-G--G...-ACG--CT-T---GACA-T-TC--A-G--T--A-CACAGATGGGGGAAG--CCTGAT-GAATAGATC-TG-----	250
GSN.CM.99.CN166	.....	0
GSN.CM.99.CN71	.....	0
DRL.-.-.FAO	.....	0
RCM.GA.-.GAB1	.....	0
RCM.NG.-.NG411	.....	0
MND-2.-.-.5440	.....	0
MND-2.CM.98.CM16	.....	0
MND-2.GA.-.M14	.....	0
MNE.US.-.MNE027	.....	0
LST.CD.88.447	.....	0
LST.CD.88.485	.....	0
LST.CD.88.524	.....	0
LST.KE.-.lho7	-----TG-GGC--T--A-----G-CT...TGTTG-TGAGT-TA---ATCC-TC--AT---CACAAAT--C-T-G--TT-AAGCCAG...-GGG-TCC---A---CTG-G---GAG	247
SUN.GA.98.L14	-C---CTA-GG-----C-----CACC...ATTG-TGA--CAGG...-ATCCT--TCATCCTTGTCAAGC-C-AC-G--TT-CAGCCAG...-AAGGG-TT-AT--A---TTG-G---GAG	247
MND-1.GA.-.MNDGB1	.....	0
MON.CM.99.L1	.....	0
MON.NG.-.NG1	.....	0
MUS.CM.01.1085	.....	0
DEB.CM.99.CM40	.....	0
DEB.CM.99.CM5	.....	0
SYK.KE.-.KE51	.....	0
SYK.KE.-.SYK173	-----TA-GGG-A---T--AA-----TAGAG-G-TGC...TTGG-GT-T-AG---T--ACT-C-CC-C--T--A-CCTCT...GGGAC-GG---GTT-GAAC-TG-G--GAA	138





	- TCF-1 alpha -	
H1B.FR.83.HXB2	GTTAGAGTGGAGGTTTGACAGCCGCCCTAGCATTTCATCATCATGCCCCGAGAGCTGCATCCGGAGTAC...TTCAAGAACTGC.....TGACA.....	334
H1A1.UG.85.U455	.....	0
H1B.US.90.WEAU160	A--AT--A-----AAA-----G-A-----T-AG-----TGACA.....	334
H1C.ET.86.ETH2220	.....	0
H1D.CD.84.84ZR085	.....	0
H1F1.BE.93.VI850	.....	0
H1G.SE.93.SB6165	.....	0
H1H.CF.90.056	.....	0
H1J.SE.93.SE7887	.....	0
H1K.CM.96.MP535	.....	0
H101.AE.TH.90.CM240	.....	0
H102.AG.NG.-.IBNG	.....	0
H1N.CM.95.YBF30	.....	0
H10.BE.87.ANT70	AC-GAT--A--T--ATCT--GCAACACC--TG-T--TATGATAACT--C-A--CT--C--G.....G-CTAAAACTGC.....	343
H10.CM.91.MVP5180	AC--A-A--CA--T--ATCAT--GC--AACA--T--A--TGC-AAA--C-A--CT--CCC-G.....T-ACTGACACTGC.....	343
CPZ.CD.-.ANT	.....	0
CPZ.CM.-.CAM3	.....	0
CPZ.GA.88.GAB2	.....	0
CPZ.GA.-.CPZGAB	C-G-TC--C-C-----A-G-G-C--AAGA--T-T--A--A-AA-----A--AG-----TTTAG.....	335
CPZ.US.85.CPZUS	C--ATC--CA-----CTGAA--G--TCGGAGA--GA--TAA-----A-----CG-----	329
CPZ.TZ.-.TAN1	.....	0
H2A.GW.-.ALI	A---TT---A-----CCTAAG--G--TCA-G--T--AA---TTTATT--A--C--A--G-A...TGG-T--AAG...TCAGGCCTGCCAGAAGATGAG--GA--GGCAAGACTGAAAGCAAGA	368
H2A.DE.-.BEN	A---T---CA-----TC-ATG--G--C-A-A-CT--A---TTCACT--T-C--A--AG-G...TGG-C--AAG...TCAGGATTGCCAGAGAAAGAA--GA--GGCAAAACTGAAAGCAAGA	368
H2A.SN.-.ST	.....	0
H2B.GH.86.D205	TC-TATC--CA-----TTC--T--G---A-G--T-TG---TTTCA-TAG-TT--A--AG-G...GG-T--CAG...TCAGGGCTACCAGAGGAGGAG--GA--GGCTAGACTAAAAGCAAGA	371
H2B.CI.-.EHO	CC-T-TC--CA-----TC--T-----A-G-CT-TG-----TTCAGCAG-TTC--A--G-G...TGG-T-TCAG...TCAGGAATGCCAGAGAAAGAG--GA--GGCTAAACTGAGAGCAAGA	368
H2G.CI.-.ABT96	.....	0
H2U.FR.96.12034	.....	0
MAC.US.-.239	TC--CA--A-----TCCAACT--G--C-ACAC-T-TGA---ATAT-TTAGAT-C--A--AG-GT...GG-AGCAA-T...CAGGCCTGTGAGAGGAAGA-GTTAGAAGAAGGCTAACCCG... ..	364
SMM.SL.92.SL92B	.....	0
SMM.US.-.H9	.....	0
STM.US.-.STM	.....TGGAAGGATTTATTACAGTGAGA.GAAGGCTAACCCG... ..	37
SAB.SN.-.SAB1C	CC-G-T---A-----CCA-TA-----G-GG-CT-TG-T--TT-GAG-----C--A--C-G...G-GCC-TCAGCA...CAAGGCTAACCCAAAACCACATC-TATG.....	352
TAN.UG.-.TAN1	A-G-T---AA-----AA--CA-T-C-A--AGTTC--TACA-C---CAG--T...GGA...ATGCATGGTCTGTGTTAAGAGGAAGT-GCCTAACCCGAGGCTTGCG	347
VER.KE.-.AGM155	A-G-CA---A-----T-----CAATGT---TGTT--TACG-C--CTCAAGG...GAGT-CTTTACAGACTTATATTCAACAGTTGGTACAG--A-CTAGCCGACCACAGGCTTG	362
VER.KE.-.9063	C-G-TC---A-----C-ATGT-A--AGTGC--TATG-C--AA-CAGA...CAGT-TTTTGAAGACATGCATGCAATAGTGAAGAGGAAGT-GCTAACCCGAGGCTTGTTG.	361
VER.DE.-.AGM3	.....	0
VER.KE.-.TY01	.....	0
COL.CM.-.CGU1	.....	0
DEN.CD.-.CD1	.....	0
GRV.ET.-.GRI_677	C-G-TC---A-----CCGAAGT-G--GG-GG-GT--CGC.....CATGTTTAAAGG-C-TGCA...GAACATGCAAAGCGCTAG.....	335
GSN.CM.99.CN166	.....	0
GSN.CM.99.CN71	.....	0
DRL.-.-.FAO	.....	0
RCM.GA.-.GAB1	.....	0
RCM.NG.-.NG411	.....	0
MND-2.-.-.5440	.....	0
MND-2.CM.98.CM16	.....	0
MND-2.GA.-.M14	.....	0
MNE.US.-.MNE027	.....	0
LST.CD.88.447	.....	0
LST.CD.88.485	.....	0
LST.CD.88.524	.....	0
LST.KE.-.lho7	-C-TATC--CAC--CA--TCCAATGT-G--G-GG-CTTT--A---TCA-AAA--G--T-C-A-G...A-AC-A--TACT...GCTTTCGCTTTTACTGCAAGA-GA-GTAGCCAAAACCG... ..	365
SUN.GA.98.L14	AC-TATC---CA----TCC-ACT--G--C-A-G--TTT-G---AT-C--AA--C--T--AG-G...-----C-TGTT...ACTAGTCTGCTGAGTGGAGGTG-AG-TTAACCCGAAAACCG... ..	365
MND-1.GA.-.MNDGB1	.....	0
MON.CM.99.L1	.....	0
MON.NG.-.NG1	.....	0
MUS.CM.01.1085	.....	0
DEB.CM.99.CM40	.....	0
DEB.CM.99.CM5	.....	0
SYK.KE.-.KE51	.....	0
SYK.KE.-.SYK173	CC-CAT----AC--CA--CC--A--G--C-A-ACA-CAGGCTGGGAGAT-GCC-GCTT-C---TA...GAG-GAC-AACA...GGAAAACCACAAGAGCTGCAG-C-GCACTTAGTAAGAACATCTCC	262





PLV  
complete genomes



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

	NF-kappa-B-II	NF-kappa-B-I	
H1B.FR.83.HXB2	TCGAGCTTGCT...ACA...AGGGACTTTCCG	CTGGGGAC	368
H1A1.UG.85.U455			0
H1B.US.90.WEAU160	-----T--T-----A		368
H1C.ET.86.ETH2220			0
H1D.CD.84.84ZR085			0
H1F1.BE.93.VI850			0
H1G.SE.93.SE6165			0
H1H.CF.90.056			0
H1J.SE.93.SE7887			0
H1K.CM.96.MP535			0
H1O1.AE.TH.90.CM240			0
H1O2.AG.NG.-.IBNG			0
H1N.CM.95.YBF30	TGACCTGAAGATTGCTGACTG-	...A-----A...GCAAAGACTGCTGACTGTC	404
H1O.BE.87.ANT70		-----A...GACTGCTGACTGTC	375
H1O.CM.91.MVP5180			0
CPZ.CD.-.ANT			0
CPZ.CM.-.CAM3			0
CPZ.GA.88.GAB2			0
CPZ.GA.-.CPZGAB	...ACTGGCGCATGCGC.ACAAGAAC---G---TCTGC-----AA		386
CPZ.US.85.CPZUS	TGGCTGTAACCGCGCAGGCGC.AATA-AAC---G---TG..A-----TAA...GGGACTTTCCAA		396
CPZ.TZ.-.TAN1			0
H2A.GW.-.ALI	GGGATACCATTAGT.AAGAACAGGAACAGCT.GATTTGG...TCAGGGCAGGAAGTAACTACTG...AAA-C.AGCTGA.GACTGC-----A...GAA--GG		464
H2A.DE.-.BEN	GGGATACCATATAGTGA...ATAACAGGAACACCATACT...TGGTCAAGGCAGGAAGTAGCTA.CTAAGAA-C.AGCTGA.GGCTGC-----A...GAA--GG		467
H2A.SN.-.ST			0
H2B.GH.86.D205	GGGATACCCTACAGATTAGGCAAGAGACAGCAGCATAAACA...GGAACCTAGCTGACTGCACAA.GAAGGAA-CTAGCAGA.CACTGC-----A...AAA-AGG		474
H2B.CI.-.EHO	GGAATACCCTACAGAGTAGACAGGAAATAACAGCACAAAGA...GGAACCTAGCTAACACTGCATAGAGAAGGAA-CTAGCTGA.TACTGC-----A...GAA--GG		471
H2G.CI.-.ABT96			0
H2U.FR.96.12034			0
MAC.US.-.239	...AAGA...GGCCTTCTTAACATGGCTGACAAGAAGGAA-CTCGCTGA...ACAGC-----A...CAA--GG		432
SMM.SL.92.SL92B			0
SMM.US.-.H9			0
STM.US.-.STM	...AAGA...GGCCTTCTTAAGATGGCTGACAAGAAGGAA-CAAGCTGA.GACAGC-----A...CAA--GG		105
SAB.SN.-.SAB1C	...GAGTTGTCATGGTGATGACAT-AAGAAC---G---TGAG-----A...GG		405
TAN.UG.-.TAN1	GT...TAAGCTAA...CTAGGCCGTTGCCTAGGAGATG.ACATAAAGAAG---G...TCAGC-----A...413		413
VER.KE.-.AGM155	...CG...GTTCCTGGTTGCCTAGGAGATGACAT-AAGAAC---G...-----A...G-AC----		425
VER.KE.-.9063	...GTTAAGCCGTTGCCAGGGAGATGACAT-T--AAC---G...---G...-----A...A-----		421
VER.DE.-.AGM3			0
VER.KE.-.TY01			0
COL.CM.-.CGU1			0
DEN.CD.-.CD1			0
GRV.ET.-.GRI_677	...T-TCAGCA--T...TGCGGTT-----CC...A----		372
GSN.CM.99.CN166			0
GSN.CM.99.CN71			0
DRL.-.-.FAO			0
RCM.GA.-.GAB1			0
RCM.NG.-.NG411			0
MND-2.-.-.5440			0
MND-2.CM.98.CM16	...CGTCGGGTGGGG-A-TGG		18
MND-2.GA.-.M14			0
MNE.US.-.MNE027			0
LST.CD.88.447			0
LST.CD.88.485			0
LST.CD.88.524			0
LST.KE.-.lho7	...AAACCGCAGCACATCCTCTGTCAGCCCGGTTG...CTAAGGCAACCGGGCTAGCGCATGCGCAATG-----G...-GAGC-----A...GGACGGG-G---G		469
SUN.GA.98.L14	...CATCCTCTTGCA...TCGCCTAGGCAACGGGGCTAGC.GCAT...GC--GC...TAGAGT-----		427
MND-1.GA.-.MNDGB1			0
MON.CM.99.L1			0
MON.NG.-.NG1			0
MUS.CM.01.1085			0
DEB.CM.99.CM40			0
DEB.CM.99.CM5			0
SYK.KE.-.KE51			0
SYK.KE.-.SYK173	TAGGAGACTCCATGGTGACAAGCTCGGCCAGGGGA...-----C-GGGC...GGGG-AA-G-		319





PLV  
complete genomes



	TAR element stem	TAR element bulge loop	TAR element stem	
H1B.FR.83.HXB2	TTAGA.....	CCAGATCTGAGCCTGGGAGCTCTCTGGGCTA	ACTAGGG.AACCCACTGCTTA.....	AGCC.T..... 525
H1A1.UG.85.U455	.....	.....	.....	0
H1B.US.90.WEAU160	.....	.....	.....G.....	525
H1C.ET.86.ETH2220	.....	.....	.....	0
H1D.CD.84.84ZR085	.....	.....	.....CG.G.....	39
H1F1.BE.93.VI850	.....	.....	.....	0
H1G.SE.93.SE6165	.....	.....	.....	0
H1H.CF.90.056	.....	.....	.....	0
H1J.SE.93.SE7887	.....	.....	.....	0
H1K.CM.96.MP535	.....	.....	.....	0
H101.AE.TH.90.CM240	.....	.....G.....	.....A.....	72
H102.AG.NG.-.IBNG	.....	.....	.....G.GGA.....	54
H1N.CM.95.YBF30	C-G.....	.....TA.....	.....G-AG-A.-C-ACTG.....	85
H1O.BE.87.ANT70	GAG.....	.....G.....	.....C-CTAGCT.....	560
H1O.CM.91.MVP5180	GAG.....	.....G.....	.....C-CTAGCT.....	533
CPZ.CD.-.ANT	.....	.....	.....	0
CPZ.CM.-.CAM3	.....	.....	.....	0
CPZ.GA.88.GAB2	.....	.....	.....	0
CPZ.GA.-.CPZGAB	C-G.....	.....T.....	.....A.-C-GTGA.....	544
CPZ.US.85.CPZUS	C.....	.....AA.....	.....C.....	550
CPZ.TZ.-.TAN1	A-CTG.....	.....	.....TAGTGGCT.....	66
H2A.GW.-.ALI	GAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCACTA.	G--G-.A-----T-T-C--CTAG.	CTCA.CCAGTG--TGGCCGGCACTGGGCAGA...	GGCTCCACGCTTG. 683
H2A.DE.-.BEN	GAGAGGCTGGCAGATCGAGCCCTGAGAGGTTCTCT.CCAGCACTA.	G--G-.A-----T-T-C--CTGG.	CTCA.CCAGT--TGGCCGGTACTGGGCAGA...	GGCTCCACGCTTG. 687
H2A.SN.-.ST	GAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCACTA.	G--G-.A-----T-T-C--CTAG.	CTCA.CCAGTG--TGGCCGGCACTGGGCAGA...	GGCTCCACGCTTG. 132
H2B.GH.86.D205	GAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCACTA.	G--G-.A-----T-T-C--CTAG.	CTCA.CCAG--AGCCAGTGTGGGCAGA...	GGCTCCACGCTTG. 693
H2B.CI.-.EHO	GAGAGGCTGGCAGATCGAGCCCTGGGAGGTTCTCT.CCAGCACTA.	G--G-.A-----T-T-C--CTAG.	CTCA.CCAG--TGGCCGGTGTGGGCAGA...	GGCTCCACGCTTG. 690
H2G.CI.-.ABT96	.....	.....	.....	55
H2U.FR.96.12034	GAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCACTG.	G--G-.A-----T-T-C--CTAG.	CTCA.CCAGT--TGGCCGGTACTGGGCAGAGT.	GACTCCACGCTTG. 184
MAC.US.-.239	GAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCACTA.	G--G-.A-----T-T-C--CTAG.	CTCA.CCAG--TGGCCGGTGTGGGCAGAGT.	GACTCCACGCTTG. 651
SMM.SL.92.SL92B	GAGAGGCTGGCAGATAGAGCCCGAGAGGTTCTCT.GCAGCACTTA.GT-G..	-----T-T--ACTA.-	CTCAACCAG--ACAGCCATTGTGGGTAGA...	GGCTCCACGCTTG. 132
SMM.US.-.H9	GAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCACTA.	G--G-.A-----T-T-C--CTGG.	CTCA.CCAG--TGGCCGGTGTGGGCAGAGT.	RGCTCCACGCTTG. 134
STM.US.-.STM	GAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCACTA.	G--G-.A-----A-T-T-C--CTAG.	CTCA.CCAG--TGGCCAGTGTGGGTAGAGT.	GGCTCCACGCTTG. 324
SAB.SN.-.SAB1C	GAG..GCTGCCAGATTGAGCC.TGGG.TGTTCTCTGGTAAGTCTGAA	---C-T-----T-T-G---GG-AGGC-A-C.CTAAG-GCA--	CTGGTGAAGTCTTG-TA..GACCCCTCGCTGGC	614
TAN.UG.-.TAN1	GG-----T-G-----T-G-----TG-.G-CTAAC.CTGGTCTG-	CTCCAGGGGTAAG-----C.....	574	
VER.KE.-.AGM155	GG-----C-T-----T-T-G---T-.G-CTAAC.CTGGT.TG--	CA.CCAGGGGTAAGGA-T..C.....	589	
VER.KE.-.9063	GG-----C-T-----T-T-G---T-.G-CTAAC.CTGGT.TG--	CA.CCAGGGGTAAGGA-T..C.....	585	
VER.DE.-.AGM3	GG-----C-T-----T-T-G---T-.G-CTAAC.CTGGT.TG--	CA.CCAGGGGTAAGGA-T..C.....	81	
VER.KE.-.TY01	GG-----C-T-----T-T-G---T-.G-CTAAC.CTGGT.TG--	CA.CCAGGGGTAAGGA-T..C.....	81	
COL.CM.-.CGU1	C-CTT.....GCCTGGCTGCCTGAGCCCTGAGT	.....G--TG-CA---T-AGCCT.TG--TGAGT--G-GTAGCTTG.	TTGGTTGAAGGGCTTGA	100
DEN.CD.-.CD1	.....	.....GG--T-CTTGCAGT--CTCC-GCGCCTGGGAGC--TGAGAGCCCTTAGCAG	54	
GRV.ET.-.GRI_677	GG-----C-A-----T-T-G---T-.G-CTAAC.CCGGT.TG--	CA.CCAGGGGTAAGGA-T..C.....	551	
GSN.CM.99.CN166	.....	.....T--TCG-CTT.....CAC.-G-GGC--GGAGCCCTTGGAGCT--GT..AAAGCGGCCTGGG.	56	
GSN.CM.99.CN71	.....	.....T--CTC-GTCTC.....CAC.-G-GGC--GGAGCCCTTGGAGCT--GT..AGAGCGGCCTGGG.	56	
DRL.-.-.FAO	.....	.....	0	
RCM.GA.-.GAB1	.....	.....	0	
RCM.NG.-.NG411	.....	.....	0	
MND-2.-.-.5440	.....	.....	0	
MND-2.CM.98.CM16	-AGCCCTGGGTGGTTTCGAGGCCCTGGCTTGTAGCTCTGGGAGCT-GTC-GG--TT---AAGA.GGCTCAT-.....GGGG--C-C-TGCTTGG..TCTT-GTAGACCTCTAGCAGTCT	253		
MND-2.GA.-.M14	-AGCCCTGGGCGGTTTCGAGGCCCTGGCTTGTAGCTCTGGGAGCT-GTC-GG--TT-TG-AA--GTCTC-CTA-.....GTGAG-ACTC-TGCTTGG..TTTT-GTAGACCTCTAGCAATCT	179		
MNE.US.-.MNE027	GAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCACTA.	G--G-.A-----T-T-C--CTAG.	CTCA.CCAG--TGGCCGGTGTGGGCAGAGT.	GGCTCCACGCTTG. 133
LST.CD.88.447	.....	.....	0	
LST.CD.88.485	.....	.....	0	
LST.CD.88.524	.....	.....	0	
LST.KE.-.lho7	.....TCCTT.GCTCTGGGAC-AGC-GTC-CT-----T.GAGCACTGG.T-AGCTA.-G-AGT-A--CCTGGGAGGCTACACTTA..GCACCCACCTGGGC	612		
SUN.GA.98.L14	GA..GGCTAGGCGACG.A.CCCT.....GAGAGGTTCTCA--A----GG--T-TCG.CTCCT-AGTGAG--TGC-GAGCACCTCACCC-AG..CCTGATCAGCTTG.	598		
MND-1.GA.-.MNDGB1	-CTCTGAG.CAGATCCCTTAGAGCAAGGACCAGAGTCTCT.....GA-T-A---TCTGAG-ACCTC-.CTCG--CTGATCACCTCG-GGTAGTGGAACTC-T..-GCTTCTTG...	140		
MON.CM.99.L1	.....	.....	-GCG-CTCGT-T-CACAGCGCCTGGGAGCCCTTGGAGCTAGGTAAAGCGGC	52
MON.NG.-.NG1	.....	.....	0	
MUS.CM.01.1085	.....	.....	-G-A--AGCTCCGACGCGCCTGGGAGCCCTCGAACTA	39
DEB.CM.99.CM40	.....CGGGAGTCG-GCCTTGC-A-TC--GGTGA-TAGAG-GGCTGGGAG-C--TGC-CAC..TGGGATAGC-TAGCTCCAGCGCCTGG	84		
DEB.CM.99.CM5	.....CGGGAGTCG-GCCTAGC-A-TC--GGTGA-TAGAG-GGCTGGGAG-C--TGCCTGC.TAGTGACT-C-TAGCTCCAGTGGCCTGG	85		
SYK.KE.-.KE51	C-CTCTCAGCGCCTGGGAGCCC.TAGGAGAC.....TAGAG-GCCTGGGAG-C--TGC-GT..AAAAGGCTAGTAAGACAAGGACTGATC	104		
SYK.KE.-.SYK173	C-CTCTCAGCGCCTGGGAGCCC.TAGGAGAC.....TGTA--G-GGC-C-GGAGCCCTGGCAGT.....AAAGGCTAGTGAG.	449		





PLV  
complete genomes



see Zhang, *J. Virol* **72**(7):5664-5471(1998)  
5' LTR U5 end \ / Lys tRNA binding site  
AAAA.TCTCTAGCAG..TGCC

H1B.FR.83.HXB2	TGTGCCCGT.CTGTT.GT.....GTGACTCT.GGTAACCTAGAGATCCCTCAGACCCCTTTTAGTCA.GTGTGG.....	639
H1A1.UG.85.U455	-----AT-----AC-A--A-T-----A-----	108
H1B.US.90.WEAU160	-----A-----	639
H1C.ET.86.ETH2220	-----TT-----AGC-----C-G-----	17
H1D.CD.84.84ZR085	-----TT-----AGC-----C-G-----	153
H1F1.BE.93.VI850	.....GGCTT.....	0
H1G.SE.93.SE6165	.....GGCTT.....	22
H1H.CF.90.056	.....	0
H1J.SE.93.SE7887	.....	0
H1K.CM.96.MP535	.....	0
H101.AE.TH.90.CM240	-----G.T-----AG-----AC-C--A-T.-A-A-----C-----	185
H102.AG.NG.-.IBNG	-----A-T-----AC-C--A-T.-A-A-----	168
H1N.CM.95.YBF30	-----A-CA-C-----A-----C-----A-C-AGAC.T.-A--A-----	197
H1O.BE.87.ANT70	-----T-A-CTG.--CAACC-----GT-----T-AC--AGAC.T.-AAGCAG-----	669
H1O.CM.91.MVP5180	-----T-A-CTG.--CAACC-----GT-----T-AC--AGAC.T.-AAGCAG-----	642
CPZ.CD.-.ANT	.....	0
CPZ.CM.-.CAM3	.....	8
CPZ.GA.88.GAB2	.....	0
CPZ.GA.-.CPZGAB	-----TA-CTAT-CA-----AC-----T-----A-A-T.-TA-CA-----AGG-----	658
CPZ.US.85.CPZUS	-----TA-C.-T-CA-----A-----C-CT-A-AA--A-----GGTG-----	662
CPZ.TZ.-.TAN1	-----A-TTCA-ACCGCGT....C--C-----CTGGGG-----T-----TTGTA..GTGGCTAAGTA-----C-----	188
H2A.GW.-.ALI	---T---A-CT.C-CCTAGTCGCCGCC--GTCA...TCGGTGTTCACCTG.AGTA--AAGACCCTGGT.C--TAGGACCCCTTCTGCTTTGGGA..AACGGAGGCAGG-----C-----GT-----	857
H2A.DE.-.BEN	---T---A-CT.C-CCTAGTCGCCGCC--GTCA...TCGGTGTTCACCTGAGT-AC.AAGACCCTGGT.C--TAGGACCCCTTCTGCTTTGGGA..ATCCAAGGCAGG-----C-----GT-----	862
H2A.SN.-.ST	---CT---A-CT.C-CCTAGTCGCCGCC--GTCA...TCGGTGTTCACCTGAAAGTA--AAGACCCTGGT.C--TAGGACCCCTTCTGCTTTGGGA..AACCAAGGCAGG-----C-----GT-----	308
H2B.GH.86.D205	---T---A-CT.C-CCTAGTCGC.GCC--GTCAC...TCGGTGTCTCC-CTGA...TAGACCCTGGT.C--TAGGACCCCTTCTGCTTTGGGA..AACCAAGGCAGG-----C-----AT-----	859
H2B.CI.-.EHO	---T---A-CT.C-CCTAGTCGCCGCC--GTCA...TCGGTGTCTCC-CTAG...AAACCCTGGT.C--TAGGACCCCTTCTGCTTTGGGA..AACCAAGGCAGG-----C-----AT-----	858
H2G.CI.-.ABT96	---T---A-CT.C-CCTAGTCGCCGCC--GTCA...TCGGTGTCTCC-CTAG...AAACCCTGGT.C--TAGGACCCCTTCTGCTTTGGGA..AACCAAGGCAGG-----C-----AT-----	230
H2U.FR.96.12034	---T---A-CT.C-CCTAGTCGCCGCC--GTCA...TCGGAG-TTC-C-TTAT...AACCCTGGTAC--TAGGACCCCTTCTGCTTTGGGA..AACCAAGGCAGG-----C-----GT-----	355
MAC.US.-.239	---T---A-CT.C-CCTAGCCGCCGCC--GTCAA.CTCGGTACTCA--AATAAG...AAGACCCTGGT.C--TAGGACCCCTTCTGCTTTGGGA..AACCAAGGCAGG-----C-----AT-----	825
SMM.SL.92.SL92B	---T-T-A-CT.C-CCTAGTCGCCGCC--GTCA...TCGGGTTCGCCGATAA...GAG.CCCTGGTAC--CAGAACCCTCTACTCTTGGGGCAACTCCTGAGTG---A---G---TT---	303
SMM.US.-.H9	---T---A-CT.C-CCTAGMCGCCGCC--GTCA...CTCGGTACTC--A-ATAAG...AAGSACCCTGGT.C--TAGGACCCCTTCTGCTTTGGGA..AACCAAGGCAGG-----C-----GT-----	308
STM.US.-.STM	---T---A-CT.C-CCTAGTCGCCGCC--GTCAA.CTCGGTACTC...AAGACCCTGGT.C--TAGGACCCCTTCTGCTTTGGGC..TACTGAGGCAGA-----C-----AT-----	489
SAB.SN.-.SAB1C	-----A-TT.A--CC-CAGGACAACCCCT...GGTTACTAAGGA..TCCTGATAGAA-----CT-----	772
TAN.UG.-.TAN1	A-----TC.A-TGC...CGC-C.TTCCT--G-A---GTTA-TGG...GGTT...TCTTA..T..CCAGGCTGAG-----C-----	710
VER.KE.-.AGM155	A-----TCAT.TAAT-CCTCACTC...TGAA-G.--AG-AGTTCCT-A-TG...GGTTCTCTCTCA..AACCCAGGCAG-G--A--C-----	729
VER.KE.-.9063	A-----TCAT.TGAC-CCTCACTC...AAG-AGG.--A-C-GTTCCT-A-TG...GGTTCTCTCTCT..GACCCAGGCAGG-G--A--C-----	728
VER.DE.-.AGM3	A-----TCAT.TGAC-CCTCACTCTTGA--GGG.TAA--TTCCT-A-TG...GGTTCTCTCTCA..ACCCAGGCAG-G--A--C-----	223
VER.KE.-.TYO1	A-----TCAT.TGAC-CCTCACTCTC-TGAA-G.--A-T-.TTCCT-A-TG...GGTTCTCTCTCT..GACCCAGGCAG-G--A--C-----	223
COL.CM.-.CGU1	AC-AGTTTCA-GTCCGCCTTCTC...GCCCT...C...AGAGCGGGCCCTTTTGGCTG...TCCGGGAAGG--GGC--T-----	268
DEN.CD.-.CD1	.....ATCTTG--TT-GCTC-T-A--AGTGGG.T-CCA-GAGGTTCCGGTGTGAGCCT..TACGCAAGGTGA---C-C--C-----	230
GRV.ET.-.GRI_677	CTCATTGC-G-GCCGA-CCTCTAGAG---AC--CTC-T--.T-G-T--T-CTGT-----CCAGGTGGG-G--A--C-----	692
GSN.CM.99.CN166	CA-TGGGTCGAG--CG-GCTTG...AGT..CCGGTGGACACC-G--CA-C-----	201
GSN.CM.99.CN71	CA-TGGGTCGGA-CG-GCTTGG...AGT..CGCGGTGGATACC-G--CA-C-----	209
DRL.-.-.FAO	.....	4
RCM.GA.-.GAB1	.....	4
RCM.NG.-.NG411	.....	4
MND-2.-.-.5440	.....	4
MND-2.CM.98.CM16	-ACT-G-C-AT.T--GTCTTTGGT...GAA-CTCTGGTTA-T-GAGA-CCCTCAG-.TC-GT-GCAGAGATTTGATATCTAGT..GGCAGTAGAGAA-----C-C-----	453
MND-2.GA.-.M14	-ACT-GTCCAT.T--GTCTTTGAT...GAA-CTCTGGTTA-T-GAGA-CCCTCAG-.TC-GT-GCAGAGATTTGATATCTAGT..GAGAGTAGAGAA-----C-----	375
MNE.US.-.MNE027	---T---A-CT.C-CCTAGTCGCCGCC--GTCAA.CTCGGTACTCA--AATAAG...AAGACCCTGGT.C--TAGGACCCCTTCTGCTTTGGGA..AACCAAGGCAGG-----C-----AT-----	306
LST.CD.88.447	.....	0
LST.CD.88.485	.....	0
LST.CD.88.524	.....	0
LST.KE.-.lh07	---CT-ATAGT.AG-CTCTTCTCTCTC-C-CTG.-T-C-G-GATCTCT-ACT--GAGA--GGAG-C.T---GATCCGGG-C.GGG--TC-----	795
SUN.GA.98.L14	---T---GT-C-C.TCTTCTCTCTCT...A.AACCCTGTC...AC-GCTC-T.ACTA--GAGCAATTACTGGCCGCTA..GTGATCCGGGGCTG-C.GGC-----	774
MND-1.GA.-.MNDGB1	-TAT--AT-GTCCGCTCTCTCT...AAACC-GT-GT.--TCTCATTTAGAGAACAGAGGACT..TCTAGTTAACCTT-A.GC--TT--	278
MON.CM.99.L1	.....CT-G-GAGTCTG--C-CATC-GTGC-ATCAGGAGCGGGCTTGACCCGGGGAGGGGACAGCAGAC-GC.C--TCTC-----	225
MON.NG.-.NG1	.....T-AGG-AAGTCTG--CC--T---T-CACACTGAGAGCCTATCTCAGCGT..GGTGTGAGTCTGG---CAC--C-----	211
MUS.CM.01.1085	.....T-AGG-AAGTCTG--CC--T---T-CACACTGAGAGCCTATCTCAGCGT..GGTGTGAGTCTGG---CAC--C-----	211
DEB.CM.99.CM40	-C-----TCATCA--GACTCCAGGGAA-TGGG-C-AGCGGGCT--GCT...AAGCG-G-GGTC-----T-----	245
DEB.CM.99.CM5	-C-----TCATCA--GACTCCAGGGGTC-GGAGCTCAAGCGGCT--GCT...TGGGC-GG-GAAC-----T-----	247
SYK.KE.-.KE51	GAGTG-T-C-A-C-GTTCTGCCGGCTCTC-CAGAA-ACG-AC-G-AA-...G-G-CTT-----	247
SYK.KE.-.SYK173	-C--TT-TGCTG-G-CACCTCCCGAA-GTG--A.--G-G-GC-----CTT-----	607



H1B.FR.83.HXB2	GCCCCAACAGG.G.ACCTGA...A.....AGCGAAAGG..GA..AA.....CC.....AGA...GGAG.....CTCTCT.CGACGCA...GGACTCGGCTT..	704
H1A1.UG.85.U455	-----TC-----T..T.....-A-----T-----G	171
H1B.US.90.WEAU160	-----C-----	705
H1C.ET.86.ETH2220	-----G-----T...G-----A-----G	84
H1D.CD.84.84ZR085	-----T..AG-----A-----G	219
H1F1.BE.93.VI850	-----T-----T..AG-----A-----GG-----G	56
H1G.SE.93.SB6165	-----T-----T..T...-CAGGGAC.TCGAAAGC...GAAAGTT-----A-----T-----G	109
H1H.CF.90.056	-----T-----T-----T-----T..A-----A-----A-----G	52
H1J.SE.93.SE7887	-----T-----T-----T-----T-----T-----G	25
H1K.CM.96.MP535	-----T-----T-----T-----T-----T-----G	0
H101.AE.TH.90.CM240	-----C-TC-----T..T...-TAGGGAC.TCGAAAGC...GAAAGTT-----A-----T-----G	273
H102.AG.NG.-.IBNG	-----A...T...C...G-TA-T-----AC.TCGAAAGC...GAAAGTT-----A-----A-----G	248
H1N.CM.95.YBF30	-----T-----T..AG-----G-----CTGA...A-----CG	265
H1O.BE.87.ANT70	-----T-----GTG-A...TG...A-----G...A-AA...AA-CTC...-AC...-G-----A	738
H10.CM.91.MVP5180	-----GCG-----AGTG-A...TG...G-A-----G...A-AA...AA-CTC...-AC...-G-----A	712
CPZ.CD.-.ANT	-----T-----GG-----TAC-----A-C...-G...GA...-CA...-TC..	45
CPZ.CM.-.CAM3	-----T-----TTG-A...T-G...-CTGA...-TC...-CG...	78
CPZ.GA.88.GAB2	-----T-----CA-----G...GAGGAA-CTGA...-C...-GC...-A...	58
CPZ.GA.-.CPZGAB	-----T-----TA-----AG...GCT...A-T...-C-GA...-C...	723
CPZ.US.85.CPZUS	-----T-----A-----TAC-----GG...GCTGAA-C...T...-GA...-G...-C...	729
CPZ.TZ.-.TAN1	-----T-----G...-AGC-GG...AAGC...-GGC-C...TGGA-GCA...-G-----C...	245
H2A.GW.-.ALI	-----T-----GAA-CT-A...-GTCTAG..GAACACGG...CTGAGTGAA...G-C...A-TAAG...GG-GGC.A-GAA-AA.CC--GAC-GAGT	947
H2A.DE.-.BEN	-----T-----GAG-CT-A...-GCCCTG..GAACCTCGG...CTGAGTGAA...G-C...A-TAAG...GG-GGC.A-GAA-AA.CC--GAC-GAGT	952
H2A.SN.-.ST	-----T-----GAA-CT-A...-G.CCTTG..GAACACGG...CTGAGTGAA...G-C...A-TAAG...GG-GGC.A-GAA-AA.CC--GAC-GAGT	398
H2B.GH.86.D205	-----G...G...GAA-GT-A...-GTCTTCAA..GAGCACAG...CTGAGTGAA...G-C...A-TAAG...GG-GGC.A-GAATCAA.CC--GAC-GAGA	949
H2B.CI.-.EHO	-----G...G...GAA-GT-A...-GTCTTCG..GGACACGG...CTGAGTGAA...G-C...A-TAAG...GG-GGC.A-GAATCAA.CC--GAC-GAGT	948
H2G.CI.-.ABT96	-----T-----G...G...-AG-GT-A...-GTCCTTT..GAACACGG...CAGAGTGAA...G-C...A-TAAG...GG-GGC.A-GAT-CAA.CC--GAC-GAGT	320
H2U.FR.96.12034	-----T-----T-----GT-A...-GTCTTCTG..AAACTCGG...CTGAGTGAG...G-C...A-TAAG...GG-GGC.A-GAA-AA.CC--GAC-GAGT	443
MAC.US.-.239	-----T-----T-----G-A-GT-A...-G.ACTCCT.GAGTACGG...CTGAGTGAA...G-C...A-TAAG...GG-GGC.A-GAA-CAA.CC--GAC-GAGT	915
SMM.SL.92.SL92B	-----T-----T-----GA-G-CT...CGG...CTGAGTGAA...G-C...A-TAAG...GG-GGC.A-GAATCAA.CCG-GAC-GAGT	378
SMM.US.-.H9	-----T-----T-----G-A-GT-A...-G..CTCCT.GAGTACGG...CTGAGTGAA...G-C...A-CAAG...GG-GGC.A-GAA-AA.CC--GAC-GAGT	397
STM.US.-.STM	-----T-----T-----G-A-GT-A...-G.ACTCCT.GAGTACGG...CTGAGTGAA...G-C...A-TAAG...GG-GGC.A-GAA-AA.CC--GAC-GAGT	579
SAB.SN.-.SAB1C	-----T-----C...TAA--A.CGTCCC.-G..CCGCAGAG.TAGACCAG...CCTGAGGA...TCC...TTT-AC...TGAGTA.A--CGGAC.--AGAAAGGCG	866
TAN.UG.-.TAN1	-----T-----G...TC...ACTTC-CTGAAGGGGAGGAGGCATCA...-TC-AG...T--C-C-G...GGT.AAGTCG--ACGC	801
VER.KE.-.AGM155	-----T-----GT-A..AG..GCACGTAC.AGCTGAGA...AGACGTCCG...-CG...C--AGG...AA.C-G...-GG-TGCG.ACGTGACCCGAGA	813
VER.KE.-.9063	-----T-----GGTCCGAGGACTT-GT-A..AG..GCACGTAC.AGCTGAGA...AGACGTCCG...-CG...C-TAGG...AA.C-G...-GG-TGCG.ACGTGACCAAGA	827
VER.DE.-.AGM3	-----T-----T.T-GT-A..AG..GCACGTAC.AGCTGAGA...AGACGTCCG...-CG...C-AGG...AAGG-G...-GG-TGCG.ACGTGACCAAGA	313
VER.KE.-.TY01	-----T-----G...T-A-GT-T..AG..GCACGTAC.AGCTGAGA...AGGCGTCCG...-CG...C-AGG...AA.G-G...-GG-TGCG.ACG-GACCAAGA	313
COL.CM.-.CGU1	-----G-----T-----A-G.TCC..CT..G-GGGTCG.CGGAAGAA...GAAGAGGAG...-CT...-ATCGC...-CTC.A--A-AGAG.-T-AGTC-AAAA	358
DEB.CD.-.CD1	-----T-----A--GTCCTGA-AC-GGGACTTGAGTACCCAGAGGCTCAGG-CCGAAG-CGGTT-GAACGTACCAC-CTCT-A-CGA--AGG-CGAGGAAGACG	346
GRV.ET.-.GRI_677	-----T-----G...AA--GGCATC-GCACCCAGCCTGAGTTGCT...GAGCGTCCG...-GGGACGACTCAGGTAGGTTGAGAGCCTA-GAGTTTTT-GCTACCT	799
GSN.CM.99.CM166	-----T-----ACAGCCC--.AC..TTCAAGCC.CGAAGGCG...GGACCCAAG...GAT...TG-C-C.T-G-CG-GT.-T-CG--CC	287
GSN.CM.99.CM171	-----G-T-----GCATCCC-T.GAC..TTCAAGCT.CGCAGGCG...GGACCCAAG...GAT...TG-C-C.T-G-CG-GT.-T-CG--CC	295
DRL.-.FAO	-----T-----C...GTCGACCC.CACCTGTG...AGGGACGAA.GGCGGC-GCCATCCGGACCGACC-A-C-GGAA-A-AGAGT-CTAAG..AAGC	102
RCM.GA.-.GAB1	-----T-----G...T-GCT-A...-G-CCTCCG.AGGCTAAG...GTTCGGCG-...G-C...AG-TC...AC-G-G...-G-GT...-AC-TTGAC.	89
RCM.NG.-.NG411	-----T-----G...T-CT-A...-G-CCTCCG.ACGCGGGT...CCGGA...A--TG...ACTGG-.A--ACGCGG.ACC-AGTC-AGG	93
MND-2.-.5440	-----T-----C...AA-GGTTT..AG..GTCGCACC.CACCTGTA...AGGGACGAGCGGCGGC-GCA-TCGGACAGACC-A-C-GG--AA.GAGTTA-T-GAAGAGC	109
MND-2.CM.98.CM16	-----T-----C...AA-GGTTT..GTCATTCC.AACCTGTA...AGGGACGAA.GGCGGC-GCCA-CCGGACCGACC-A-C-GG--A-TGAGTTA--CAA-GAGC	559
MND-2.GA.-.M14	-----T-----C...AA-GGTTT..GTCGCTCC.GACCTGTA...AGGGACGAGCGGCGGC-GCC-TCGGACAGACC-A-C-GA--A-AGAGTTA-TCAA-GAGC	482
MNE.US.-.MNE027	-----T-----G-A-GT-A...-G.ACTCCT.GAGTACGG...CTGAGTGAA...G-C...A-TAAG...GG-GGC.A-GAA-CAA.CC--GGC-GAGT	396
LST.CD.88.447	-----T-----T-----T-----T-----T-----G	0
LST.CD.88.485	-----T-----T-----T-----T-----T-----G	0
LST.CD.88.524	-----T-----T-----T-----T-----T-----G	0
LST.KE.-.lho7	-----T-----G-A--GA...-C..GCTGGCAA.GAACCTCT...GGAGGGAGG...-A...A-C-CG...-AGAAG.A-G-CAGGA.A-CTAG-A-AGA	886
SUN.GA.98.L14	-----A-----G...G-A-G-A...-T..G-CGAAGA.GACGTGGA...GTGAAA-T...TC...-A-AA...G-GAGC.A-CT--TGA.-A-GAA--AAGG	865
MND-1.GA.-.MNDGB1	-----T-----T-----GAGAGGCAC..TG...CACTGA.GGCAGAGC...ACTCCC-T...G--.A-AGC...AGG-TG.AAGGAG-GT...-G-T--GA	368
MON.CM.99.L1	-----T-----G...TAACG-CCC..AGCG-CTCGAGCCCGCAGGCGAACAGATCCG-GAGAAGCCTTGCC--TGCGCCCCA-AAAGCTGCG-AGCAAGA-GTAA--AGAG	339
MON.NG.-.NG1	-----T-----T-----T-----T-----T-----G	0
MUS.CM.01.1085	-----T-----TT...AA-CCC-CGA-GAGGTAGGCGGTGTGACGGAGAGAAGGCC-AGCCGAGTGCCT-CCAAGCCACCGAGAA-AAGTG-TGCGCA-CCA-AA-G-AG	327
DEB.CM.99.CM40	-----T-----G...CTA-CC-GA-CTCG.GCCCCAGGCGGTGTGCGGAGTCTGGAAA-CCTGAGCCGATCTCTTCCCAGCCGGA..AGCTGCG-AGCAAA--TAAGC-GACG	360
DEB.CM.99.CM5	-----T-----AG...C-T-CC-GT-TTTGAGCCCGAAGGCGGAGTGTGAGTCTGGAAA-CCTGAGCCGATCTCTTCCCAGCCGGA..AGCTGCG-AGCATA--TAAGC-GACG	365
SYK.KE.-.KE51	-----T-----C...T-A-CT...-CCCAGAGAGGAGACCCAGTCAAAGGACCCAGAGGAAC-T-CGCACGAAGAAGAGAAGCGGTGTCCGCCCTAT-C-GAT	359
SYK.KE.-.SYK173	-----G-----T-----C...TAG-CC...-CCCGAGAG.GAGGTGAC...GCCAGTCAA...-G...-AGTCC...GGA--C.AAGA-GGAT.C-G--GTC-G-G-C	697







H1B.FR.83.HXB2	...CTA....GAAGGAG..AGAG.....	789
H1A1.UG.85.U455	-----A.....	256
H1B.US.90.WEAU160	-----	788
H1C.ET.86.ETH2220	-----	169
H1D.CD.84.84ZR085	-----	306
H1F1.BE.93.VI850	-----	144
H1G.SE.93.SB6165	-----G.....	192
H1H.CF.90.056	-----	136
H1J.SE.93.SE7887	-----	112
H1K.CM.96.MP535	-----	0
H101.AE.TH.90.CM240	-----	354
H102.AG.NG.-.IBNG	-----	329
H1N.CM.95.YBF30	...AAG...T-G-AGA..GAG.....	347
H1O.BE.87.ANT70	...-C...--CCT...G-GAAGGG.CGA..AGTCTCTAGG.G..GAGGAAG.....	841
H1O.CM.91.MVP5180	...-C...--CCT...G-GAAGGG.CGA..AGTCTCTAGG.G..GAGGAAG.....	816
CPZ.CD.-.ANT	...GC...ATCCT...G-GAAGGT.CGA..AGTCTCTAGGAA..CAGGAAG.....	153
CPZ.CM.-.CAM3	...-G...--CCT...G-GAAGGG.CGA..AGTCTCTAGG.G..GAAGAGG.....	179
CPZ.GA.88.GAB2	...-CCT...--GAAGGG.CGA..AGTCTCTAGGTG..GA.GAGG.....	161
CPZ.GA.-.CPZGAB	...-CCT...--GAAGGG.CGA..AGTCTCTAGG.G..GGAGAAG.....	826
CPZ.US.85.CPZUS	...-G...--CCT...G-GAAGGG.CGA..AGTCTCTAGG.G..GAAGAGG.....	831
CPZ.TZ.-.TAN1	...GC...ACCCT...--GAAGGG.CGA..AGTCTCTAGGTAACAGGGGAA.....	352
H2A.GW.-.ALI	...TA...-T-CCTA..CACCAAAA.T.T..GTAGCC.GAAAG..GGCTTGTATCCTACCTTTAG..ACAGGTAGA..AGATT..GTGGGAG.....	1095
H2A.DE.-.BEN	...TA...-T-CCTA..CACCAAAA.ACT..GTAGCCAGAAAA..GGCTTGTATCCTACCTTTAG..ACAGGTAGA..AGATT..GTGGGAG.....	1102
H2A.SN.-.ST	...TA...-TGCCTA..CACCAAAAT.ACA..GTAGCCAGAA.G..GGCTTGTATCCTACCTTTAG..ACGGGTAGA..AGATT..GTGGGAG.....	547
H2B.GH.86.D205	...TA...-T-CTCC..-A-...TTT..TGCATTGTCTACTCTAAAGAG..GGAGTAGGG..CATAA..GTGGGAG.....	1078
H2B.CI.-.EHO	...TA...-T-CTCC..-A-TTT..GCAATTCTGCTACTCTAGGAC..AGGGTACGG..AATAA..GTGGGAG.....	1079
H2G.CI.-.ABT96	...TA...-TGCACC..-A-AA'TTT.GTA..GCCAAAAAGGGC..TGTGAATCCTACCTGAATAG..TCAGGTAGA..AGT..GTGGGAG.....	465
H2U.FR.96.12034	...TA...-T-GCA-C.-A-AAATA.GA..ACCAA..AGGGT..TGTCTGCCTCTGAGAAAGGAG...AGGTAGT..ATATT..GTGGGAG.....	582
MAC.US.-.239	...TA...-T-GCA-C.-C-AAAAA.GAA..A.....TAG..CTGTCTT.TTATCCAGGAAGG..GGTAATAAG..ATAGA..GTGGGAG.....	1052
SMM.SL.92.SL92B	...T.....CCAAAA.....GGTGTG..AATATCCTACCTACGGGATAG..GTAGAAGTA...AA..GCGGGAG.....	492
SMM.US.-.H9	...TA...-T-GCA-C.-A-A.....GAGTCATAGG..ACTGAGTCCCTACTTTTGTAG...AAAAGAGTA..GGAGA..GTGGGAG.....	534
STM.US.-.STM	...TG...-T-GCA-C.-AGTAGCC.AGA.....GG..GCTGATTGTTCTACTCTTGTAG..GGAAGAGGT..AGAAG..GTGGGAG.....	708
SAB.SN.-.SAB1C	GAGAAG...TC--ACA..G-GCTGAC.AAG..GCTGGCTTAGCA..GTGGCAGAAGGCTAGGTCGAG...CGGGAAGG...AAA.....	1016
TAN.UG.-.TAN1	...GGT...AG-TCTA..G-T-GGCG.ACA..GATCAGCACCTG..GGTGTAGCA.....	901
VER.KE.-.AGM155	...G-...ACTACTC..T-G-CAAG.TAG..GGCAGGCGGACG..GGTACGTA.....	930
VER.KE.-.9063	...A-...ACTACTC..T-G-CAAG.TAG..GGCAGGCGGAC..GGTACGCA.....	941
VER.DE.-.AGM3	...A-...ACTACTC..T-G-CAAG.TAG..GGCAGGCGGACG..GGTACGCA.....	430
VER.KE.-.TY01	...G-...ACTACTC..T-G-CAAG.TAG..GGCAGGCGGT.G..GGTACGCA.....	431
COL.CM.-.CGU1	AGGAAG...TGGCAGA..C--AGAG.CAC..AGAGAGGGAGTG..CTCCAGAGAAGGAGAAGAG...CTCATTAGG...AGGTC..TTCAGGCAGGGAGTAGGCCCCCTGGGCAGGTCGAGGGGG	549
DEN.CD.-.CD1	GG.....	478
GRV.ET.-.GRI_677	...GC.....	896
GSN.CM.99.CN166	GAGGAGTAACA--CTCCCACCAGAGGATAGGGCAGA.....	448
GSN.CM.99.CN71	AAGAAGTAACA--CTCCCACCAGAGGATAGGGCAGA.....	454
DRL.-.-.FAO	GAAAA-GGGAGA-GTA--GCCTGTTGCT.GAT..AAGTGAGCAAAAAGAGCCTA.GGGGTAAGAAAAAG.....	290
RCM.GA.-.GAB1	TCTAGG.....C--AGTCC.TAG..AGAAGGGAGGAA..A.....	211
RCM.NG.-.NG411	TCTAGG.....GT..C--AGTCC.TAG..AGAAGGGAGGAA..A.....	211
MND-2.-.-.5440	TAAGGCCGACGC--A-T-AA--TAAGGG.CTT..CCTCCGAAGGGAGAAGCAGGGGGCAAGAAA.....	293
MND-2.CM.98.CM16	TAAGGCCGAGGC--A-T-AA-AGAGGTC.CTT..GTGAAAGAGGA...CAGGGGGCAAGAAA.....	736
MND-2.GA.-.M14	TAAGGCCGAAGC--A-T-AA--TAAAAG.GGC..CTTCCGAAGGGAGGGCAGGGGGCAAGAAA.....	660
MNE.US.-.MNE027	...TA...-T-GCA-C.-C-AAAAA.GAA..A.....TAG..CTGTCTT.TTATCCAGGAAGG...GATAATAAG..ATAGA..GTGGGAG.....	532
LST.CD.88.447	.....	0
LST.CD.88.485	.....	0
LST.CD.88.524	.....	0
LST.KE.-.lho7	...AGG...T-----TAAGAC.CTC..ACCGGGAGAGTG..GCAGCAGGCCCGGGGCGAGCCT..TGTGAGAAG..TGTGT..AATTCTGATCTAGTGTAACTGGGACACTAGTGGTAGCTT	1071
SUN.GA.98.L14	...AG-C...CT--CT..G-G-GGAG.GTC..CCGCTAGGTTAG..TACTGTCAAGTAGGCAGACTA..GTAGTACGT...GGAAT..AAAATAAGAAGTGTGTGTATAACTCTGTGTCTCGGAAGA	1051
MND-1.GA.-.MNDGB1	A.....	449
MON.CM.99.L1	.....	441
MON.NG.-.NG1	.....	0
MUS.CM.01.1085	.....	446
DEB.CM.99.CM40	.....	484
DEB.CM.99.CM5	.....	487
SYK.KE.-.KE51	.....	431
SYK.KE.-.SYK173	GGGGA-...AT--G--CC.....	794





PLV  
complete genomes



	/ Gag and Gag-Pol start (p17)	
H1B.FR.83.HXB2	ATGGGTGCGAGAGCGTCAGTA...TTAAGCGGGGGAGAATTAGATCGATGGGAAAAAATTCGGTTAAGGCCAGGG	861
Gag	M G A R A S V . L S G G E L D R W E K I R L R P G	p17
H1A1.UG.85.U455	.....AA-A-----TC-----G-----	328
H1B.US.90.WEAU160	.....G-----A-G-----	860
H1C.ET.86.ETH2220	.....-A-----A-C-A-A-----GCC-----AA-----	241
H1D.CD.84.84ZR085	.....-A-----A-G-----GCT-----C-----	378
H1F1.BE.93.VI850	.....-A-----A-----GA-----A-----G-----	216
H1G.SE.93.SE6165	.....-C-----A-----GCT-----G-----	264
H1H.CF.90.056	.....-C-----A-----GCT-----G-----C-----	208
H1J.SE.93.SE7887	.....-A-----T-----A-----GAT-----G-----	184
H1K.CM.96.MP535	.....-A-----A-----GC-----G-----C-----	72
H101.AE.TH.90.CM240	.....-T-----A-----GC-----G-----GC-----	426
H102.AG.NG.-.IBNG	.....-T-----A-----GC-----G-----GC-----	401
H1N.CM.95.YBF30	.....G-----CA-----A-----A-----TC-----TAT-----G-----A-----	419
H1O.BE.87.ANT70	.....-T-----T-----G-----G-----CA-----AA-----CA-----G-----GC-----C-----A-----A-----A-----	913
H1O.CM.91.MVP5180	.....-T-----T-----G-----G-----CA-----A-----TA-----G-----GC-----CG-----A-----A-----A-----	888
CPZ.CD.-.ANT	.....-A-----G-----G-----T-----T-----G-----G-----A-----AGA-----GC-----AC-----GT-----CA-----C-----TC-----C-----T-----	225
CPZ.CM.-.CAM3	.....-C-----CA-----CG-----GCT-----A-----C-----	251
CPZ.GA.88.GAB2	.....-TA-----C-----C-----GCG-----GCC-----G-----A-----C-----C-----A-----	233
CPZ.GA.-.CPZGAB	.....-T-----C-----CA-----A-----C-----C-----G-----G-----C-----T-----C-----	898
CPZ.US.85.CPZUS	.....-T-----C-----G-----CA-----CG-----GCT-----G-----C-----T-----A-----T-----	903
CPZ.TZ.-.TAN1	.....-G-----G-----A-----ATA-----GC-----G-----AC-----TCC-----AA-----C-----T-----AAT-----CA-----A-----	424
H2A.GW.-.ALI	.....-C-----AAC-----C-----C-----G-----A-----A-----A-----GC-----CGA-----TA-----G-----A-----C-----C-----C-----	1167
H2A.DE.-.BEN	.....-C-----AAC-----C-----C-----G-----A-----AA-----A-----GC-----CGA-----TA-----G-----A-----C-----C-----C-----	1174
H2A.SN.-.ST	.....-C-----AAC-----C-----C-----G-----A-----AA-----A-----GC-----CGA-----TA-----G-----A-----C-----C-----C-----	619
H2B.GH.86.D205	.....-C-----GC-----C-----C-----GTCA-----AAGA-----AC-----GA-----TA-----G-----A-----C-----C-----C-----	1150
H2B.CI.-.EHO	.....-C-----GC-----C-----C-----GTCA-----AAGA-----GAC-----GA-----TA-----GG-----A-----C-----C-----C-----	1151
H2G.CI.-.ABT96	.....-C-----AGC-----C-----C-----GTCA-----AA-----A-----GCT-----GA-----TA-----G-----A-----C-----C-----C-----	537
H2U.FR.96.12034	.....-C-----AAC-----C-----C-----GTCA-----AAGA-----G-----GA-----TA-----G-----A-----C-----CAAC-----	653
MAC.US.-.239	.....-C-----T-----AAC-----C-----C-----GTCA-----AAGA-----GC-----GA-----TA-----A-----C-----C-----A-----CAAC-----	1124
Gag	M G V R N S V . L S G K K A D E L E K I R L R P N	p17
SMM.SL.92.SL92B	.....-C-----GC-----C-----C-----GTCA-----AA-----A-----GC-----GA-----TA-----G-----A-----C-----C-----C-----	564
SMM.US.-.H9	.....-T-----AAC-----C-----C-----GTCA-----AAGA-----GC-----GA-----TA-----A-----C-----C-----C-----	606
STM.US.-.STM	.....-C-----AGC-----C-----C-----GTCA-----AA-----A-----GC-----GA-----TA-----G-----A-----C-----C-----C-----	780
SAB.SN.-.SAB1C	.....-TAAC-----C-----T-----TA-----GA-----C-----GC-----TC-----TC-----G-----CA-----C-----CAAC-----	1088
TAN.UG.-.TAN1	.....-A-----G-----CAC-----G-----C-----C-----GTCA-----A-----GA-----T-----G-----CAC-----TT-----G-----A-----C-----C-----C-----GAAC-----	973
VER.KE.-.AGM155	.....-G-----GCTA-----C-----C-----C-----G-----ATA-----A-----C-----GA-----TT-----GC-----T-----A-----AC-----TC-----C-----GAAC-----	1002
VER.KE.-.9063	.....-G-----GCTA-----C-----C-----C-----AA-----A-----C-----C-----A-----TC-----G-----A-----A-----C-----C-----GAAC-----	1013
VER.DE.-.AGM3	.....-G-----GCTA-----C-----C-----C-----ATA-----A-----C-----CAA-----TT-----GC-----T-----A-----AC-----TC-----C-----GACC-----	502
VER.KE.-.TY01	.....-G-----GCTA-----C-----C-----C-----ATA-----A-----C-----C-----A-----TT-----G-----A-----AC-----TC-----C-----GAAC-----	503
COL.CM.-.CGU1	GCCTGCAAGTGGAGTAAGCTGCGGCGCTCCAGGACCCGGTAGAGGGAGGTCATTA...-CAACGAGCA-GG-C-G...-G-GAA-AA-AC-----AGACCTAC-G-----G-C-----C-G-A-AA-----	675
DEN.CD.-.CD1	.....-GAGCG-GAGC-----G-----TCCCA-G-CA-----T-CG-CCC-G-----A-----TT-----G-AA-A-----A-----	553
GRV.ET.-.GRI_677	.....-C-G-G-TCAC-----C-----C-----GTCA-----A-----AGCC-C-----CACG-TC-----G-----G-----A-----C-----C-----T-----GAAC-----	968
GSN.CM.99.CN166	.....-A-----GCAC-----G-----CGATGC-G-CA-----TACTA-G-----G-----A-G-AT-----G-----G-----G-----TC-----C-----A-----AA-----	523
GSN.CM.99.CN71	.....-A-----GCAC-----G-----CGATGC-G-CG-----TACTA-G-----G-----A-G-AT-----G-----G-----GA-TC-----C-----A-----GA-A-----	529
DRL.-.FAO	.....-A-----C-----G-----C-----T-----G-----A-----A-----GC-----T-----GCT-----TT-----GG-----A-----C-----CTCC-----	362
RCM.GA.-.GAB1	.....-C-----T-----G-----GTCA-----AAGA-----GC-----CGC-----TC-----G-----A-----C-----C-----C-----	283
RCM.NG.-.NG411	.....-G-----A-----T-----GTCA-----AAGA-----G-----GCT-----C-----GG-----A-----C-----C-----C-----	283
MND-2.-.5440	.....-C-----C-----G-----C-----T-----G-----A-----A-----G-----GA-----CT-----G-----A-----C-----CTCC-----	365
MND-2.CM.98.CM16	.....-C-----C-----GT-----C-----T-----G-----A-----A-----G-----CGAGCT-----G-----A-----C-----CTCC-----	808
MND-2.GA.-.M14	.....-C-----C-----GT-----C-----T-----G-----A-----AGA-----G-----GA-----CT-----C-----C-----CTCC-----	732
MNE.US.-.MNE027	.....-C-----AAC-----C-----C-----GTCA-----AAGA-----GC-----GA-----TA-----A-----C-----C-----A-----C-----C-----	604
LST.CD.88.447	.....-AT-----G-----GAAC-----T-----C-----C-----TA-----CA-----ATTGAG-GAGA-----TCTGC-GT-----CA-----C-----A-----T-----AC-----	72
LST.CD.88.485	.....-AT-----G-----GAAC-----T-----C-----C-----TA-----CA-----ATTGAG-GAGA-----TCTGC-GT-----CA-----C-----A-----T-----C-----	72
LST.CD.88.524	.....-AT-----G-----GAAC-----C-----C-----C-----TA-----CA-----ATTGAGA-AGAT-----TTTGT-----TG-----CA-----C-----A-----T-----	72
LST.KE.-.lho7	CACT...-AT-----G-----TAAC-----T-----C-----C-----TA-----CA-----ATTGA-A-AGAT-----TTTGT-----GTG-----CA-----C-----A-----T-----	1147
SUN.GA.98.L14	GTGTCTCAAAC...-A-----G-----AAC-----CAC-----G-----GGATA-----A-----TTG-GAGGA-C-----TT-----G-----G-----AGCTC-G-AA-----T-----A-----	1134
MND-1.GA.-.MNDGB1	.....-AATG-GAAC-----T-----CC-----GTTA-----ACT-----T-----G-----AA-----TT-----G-----AA-----A-----A-----AG-----T-----	521
MON.CM.99.L1	.....-G-----GCAC-----G-----CGATGC-T-----T-----TACTA-----CAAG-AT-----G-----GG-----G-----C-----T-----A-----GC-----C-----	516
MON.NG.-.NG1	.....-G-----GCAC-----G-----CGATGC-T-----T-----TACTA-----CAAG-AT-----G-----GG-----G-----C-----T-----A-----GC-----C-----	0
MUS.CM.01.1085	.....-A-----GCAC-----G-----CGATGC-G-CA-----TAACA-G-----G-----A-----G-----AT-----G-----GG-----GA-----C-----C-----A-----AAA-----	521
DEB.CM.99.CM40	.....-G-----G-----GCG-----G-----C-----CCA-----C-----CG-----CT-----CA-----TT-----G-----AA-----A-----G-----A-----AAA-----	559
DEB.CM.99.CM5	.....-G-----G-----GCGA-----G-----CGCCA-----C-----C-----CT-----CA-----TT-----G-----AA-----A-----G-----A-----AAA-----	562
SYK.KE.-.KE51	.....-A-----AG-----G-----GA-----G-----CGATCC-----CA-----TA-----A-----GC-----T-----AA-----AT-----G-----C-----TC-----GAA-----	506
SYK.KE.-.SYK173	.....-A-----AGCG-----G-----G-----C-----ATA-----CA-----TA-----GC-----CA-----AT-----G-----C-----CC-----A-----GAAA-----	869



H1B.FR.83.HXB2	GGAAAGAAAAATATAAAATTAAAACATATAGTATGGCCAGCAGGGAGCTAGAACGATTTCGAGTTAATCTGGCCTGTTAGAACATCAGAAGGCTGTAGACAAATCTGGACAGCTACAACCATCCC	991
Gag	<u>G K K K Y K L K H I V W A S R E L E R F A V N P G L L E T S E G C R Q I L G Q L Q P S</u>	p17
H1A1.UG.85.U455	-----C-----G-C-G-----T-----G--AA--A-C--C-----T-----G-----A--CAG-----AT-----G-T	458
H1B.US.90.WEAU160	-----T-----C-----C-----C-----C-----C-----C-----C-----A--TAA-----AT-----G-T	990
H1C.ET.86.ETH2220	-----C-C--TGC-G-----CC--C-----A-----G--AA--T--C--C--A--T--T-----C-----A-----A-TAA-----G-T	371
H1D.CD.84.84ZR085	-----G-----GC-----A-----G-----T-----C-----T-----A-----A-----A-----G-T	508
H1F1.BE.93.VI850	-----G-----A-G-----C-A-----T-----C-G-----TC-----CA-A-----A-AA-----346	346
H1G.SE.93.SB6165	---G---TC---A---C---G---GA---T---C---C---A---T---G---T---CA---A---A---T---G---T---T	394
H1H.CF.90.056	-----G-----GGC-----C-----G--A--T--C--C--C-----T-----C-----CT--G--A--A--A-----G--G-TA	338
H1J.SE.93.SE7887	---G---GGA---G---C-----G--CA--T--C--C-----TC--GT--G--A-----CA-----A--T-----C-----G--T	314
H1K.CM.96.MP535	-----C-G-----C-----T-----C-C-----C-----T-----G--A-----C-G-----A-AAC--AA--G-----A	202
H101.AE.TH.90.CM240	---GA---GGC-G-----T-----A--T--A-----C--C--A--T--T-----G-----A--CA-----A--A--A-----T--GT--A--T	556
H102.AG.NG.-.IBNG	-----G-C-----C-----G--A-----C-----C-----T-----G--G--A-----C--A-----C--A-----T--G--T	531
H1N.CM.95.YBF30	-----C-G-A-G-----T-----G--A-----TTG--C--A--T--CA-G--C--G--G--C-----GCCA-GT-----AAAT--AT--G-----G-T	549
H10.BE.87.ANT70	TCT--A--G-----G-C-----CT-----A-----G--A-----TG-----AG--AC-----TG-----G--TAA--GAGA--GC--G--ACAG--T--G--G--G--T	1043
H10.CM.91.MVP5180	TCT--A--GGC-----GGC-----T-----G--A-----A-----TG-----T--A-----TG-----TAC--GAG--C--G--ACAG--T--G--G--G--T	1018
CPZ.CD.-.ANT	---C-----G--C-TGA-----C-G--T-----ATC-----GC-G--T--T--GC-C-GCT--CTC--TC-----T--GA-A-GGCTA--CCAT--AT--GAGC--T---A	355
CPZ.CM.-.CAM3	-----G--CA-----T-----G--CA-----TTG--C-----T--AA-G-----GG-----A--GAG--GT--T--AA-G--AT--G-----G--T	381
CPZ.GA.88.GAB2	---GA---G---C-TGA-G-----T-----G--A-----TTG-G-C-----T--AA-G-----AA-----TAT--GT--T--AA-----G-----G--T	363
CPZ.GA.-.CPZGAB	---GA---G---TGA-G-----T-----G--A-----TG-G-C-C-G--AA-G-----GTAAG-----A-----CTA--T--GT--ACA--AT--G--G--G--T	1028
CPZ.US.85.CPZUS	-----G-----G--TGA-G-----T-----T-----G--A-----TTG--C--A--T--AA-G-----G--C-----CTC--T--T--AAA-----T--G--G--T	1033
CPZ.TZ.-.TAN1	---C-G-----TT-A-----C-----CG-A--C--A-----C-G--T-----GA-G-----C--T--CA-G--G--ACGT-----CT--GA-----CA--CCTC-----G--G--T--GG	554
H2A.GW.-.ALI	-----C-GC-----T--G-----GCG-AT--AT--G--CA-----G--T--GGCAGAAA-----G--GT--AA-----T--CCA--AG--T--TAA--GTTT--G-----TAG	1297
H2A.DE.-.BEN	-----G--C-GG-----T--G-----GCG-AT--AT--G--TAA-----G--T--GGCAGAGA-----G--GT--AA-----T--CCA--A--G--T--CA--GTTT--G--T--TAG	1304
H2A.SN.-.ST	-----G-----GGC-----T--G-----GCG-AT--AT--G--CA-----G--T--GGCAGAGA-----G--GT--AA-----G--T--CCA--A-----T--TAC--GTTT--G--T--TAG	749
H2B.GH.86.D205	-----G-----TGC-----G-----GTG-AC--AT--CA-----T--G--T--AGCAGAGA-----G--T--AA-----A--CA--A-----T--TAA--GTCT--GCT--TAG	1280
H2B.CI.-.EHO	-----G--C-TGC-----T--G-----GTG-AT--AT--GA-----T--G--T--GGCAGAGA-----G--G--T--AA-----A--CC--GA-----T--G--AA--GTCT--G--G--TAG	1281
H2G.CI.-.ABT96	-----C--TG-----G--T--G-----GCG-AC--AT--G--CA-----G--T--AGCAGAGA-----G--T--AA-----T--CCA--A-----T--TTC--GTAT--GCT--TAG	667
H2U.FR.96.12034	-----C--TG-----T--G-----GCG-AC--AT--G--CA-----T--G--T--AGCAGAGA-----G--GG--CAA-----A--CA--A-----TT--ATCGGT--T--GCT--TAG	783
MAC.US.-.239	-----G--C-TG--G--G-----G-----GCA-AT--AT--TA-----T--G--T--AGCAGAAA-----G--G--ACAA-----A--CA--A-----TTCGGTCT--GCT--TAG	1254
Gag	<u>G K K K Y M L K H V V W A A N E L D R F G L G E S L L E N K E G C Q K I L S V L G P L</u>	p17
SMM.SL.92.SL92B	---G---G---C-TGC-T--G---TA---GC---A--AT---CA---T--G--TCGGCAGAAA-----G---GC---AA-----A--CA--AG---T--A--C--GTA---GCT---TAA	694
SMM.US.-.H9	---G---G---C-TGC-G--G---GCA-AT--AT--G--CA---T--G--T--AGCAGAAA-----G---G--ACAA-----T--CA--A-----T--TTC--GTTT--GCT---TAG	736
STM.US.-.STM	---G---G---C-TG--G---G---G---GCA-AC--AT--G--CA---T--G--T--GGCAGAAA-----G---GT--AA-----G---CA--A-----TA--AACGGTCT--G---TAG	910
SAB.SN.-.SAB1C	---G---G---G---G---C-----A--A-----T---TAGCC--TCAG--AAAT--C-----G---AAG-----GGTCGTCA---T--TA--TGTA--C--TT--CCTAG	1218
TAN.UG.-.TAN1	-----G---C-----CT--A-----G---AAA-AA-G--G--C--T--GGC-CC-CGAGAAA--C-----AA-----G---CA--A--G--CA--A--G--TAT--GAGC--GCTAG	1103
VER.KE.-.AGM155	-----G--C--A--T-----T--A-----G--A--A--A--G--C--C-----GCC-CC--GAGAAGT--A--G--G---GAG-----T---A--A--G--CA--A--A--GTT--CTCT--CCTAG	1132
VER.KE.-.9063	-----C--A--T--G---T--A-----G--T--AA--A--G--G--C-----GCC-CC--GA--A--GT-----GA--T--A---A--A--G--CA--A--A--GTC--TTTT--T--TAG	1143
VER.DE.-.AGM3	-----G--CC--A--T-----T--A-----G--A--AA--G--G--C-----GCC-CC--GAGA--AT--AC-----T--GA-----A---AGA--G--CA--A--A--GTA--CT--C--GCTAG	632
VER.KE.-.TY01	-----G--CC--A--T-----T--A-----G--A--A--G--G--C-----GCC-CC--GAGA--GT-----G--G---GAG--G--G---A--AG---CA--A--A--GTC--CT--C--CCTAG	633
COL.CM.-.CGU1	AAG--AGG--TGT-----GA-----G--TAG---ATGT---CA---G--AGC--C--GT--TTT--A--TATT--AG-----A--GT--G--AC---AGTGGCT-----T---A--A--G--GACG---TAG	805
DEN.CD.-.CD1	A--T-----G--G--C-TGA--T--G--CC-GA--T---T---AG--A--A--G--G-----T--GCT--AGGAAACAG---C-----C--AT-----A--CCA--ACT--CA--A--AGACCT--GCT---TAG	683
GRV.ET.-.GRI_677	---G--A--G--G--CC--A--T-----T--A-----G--A--AA--AA--G-----T--GGT--AC--GAGAAA--T-----AA-----CA--A--A--GTTT--ACC--G--TGG	1098
GSN.CM.99.CN166	-----G--C--TGA--C--G--CT-----CTC--AA---T--GA---AT--GC--GTCCGAC--CT--TC--T---ATCA--G--G---C---GG---T--A--A--ATTG--CTTC--CCTAG	653
GSN.CM.99.CN71	-----G--G--CGCGA--C--G--CT-----CTC--A---T--GA---A---GC--ATCAGGA--T--TC--T---T--TAA-----G---CA---GG--T---A--GTCACCTCC---TAG	659
DRL.-.-.FAO	---A--G-----GCC--C--G---C--T--C-----T--GCG--A--A--T-----G---GCC--G--GAGA--AT--A-----GTCAG-----G--CCA--A---TT--A---GTA--CTTT--TCTAG	492
RCM.GA.-.GAB1	---G--A--G---C-TGC-G--G---T--G-----T---AAA--A---A--TAA---T--GCT--G--G--GA--CATT-----G--AC---A---GA--A---T--A---GTC--G--TG--TCTAG	413
RCM.NG.-.NG411	-----G---TG-----G--C-----T-----T-----G---CA---T--GCT--G--G--GA--ACTT-----GTG-----AATCCAGA---TA--A--GGTTA--T--TG--TCTAG	413
MND-2.-.-.5440	-----G--CC--G-----G--TA-----T---A--A-----CA---T--GCC--AC--GAGAAGT-----GTGAG-----A--CGA--A---T--TA--CGTA--CTTT--CTAG	495
MND-2.CM.98.CM16	---A--G---CC--GC--G---C---A---T---A--A---CA---T--GCC--C--GAAAAG-----GTAAG-----A--CGAGA---T--TA--CGTA--CTTT--CTAG	938
MND-2.GA.-.M14	-----G---CC--G-----A---T---A--A---T---T--GCC--C--GAAAAA--A-----TCAG-----A--CGAGA---T--TA--CGTA--CTTT--CTAG	862
MNE.US.-.MNE027	---G-----G--C--TG--G--G---G-----GCA-AT--AT--TA---T--G--T--AGCAGACA-----G--G--ACAA-----A---CA--A-----TTCGGCCT--GCT---TAG	734
LST.CD.88.447	A--T--A--G--CT---C--GAAG--G---T---G-----CA--A--AT---TA--G--T--G--T--AGGAG--ACA--T--C---G---G---T---C--AGA---T--T---GTCTGTTGG---CTAT	202
LST.CD.88.485	A--T--A--G--CG---C--GAA--G---T---G-----CA--AA--AT---TA--G--T--GTT--AGGAG--ACA--T--C--G--G---G---T---C--AGAG---T--T---GTCTGTTGG---CTAT	202
LST.CD.88.524	A--T--A--G--CT---C--GAA--G---G---AC-----CG--AA--A---TA--G--T--GGC--AGGAG--AAA---TC-----G---G---AGAG--TT---GTC--GTTGG--T--TAT	202
LST.KE.-.lho7	A--T--A--G--CT---C--AA--G---G---A-----CG--AA--AT---CA---T--G--T--GGGAT--ACAA--C--G-----G-----G---AGA---C--ATCTGCTGCTGG--T--TAT	1277
SUN.GA.98.L14	---G--A--G--CC--CC--GGC--G---G-----G--T--AA--AT---TA--G--T--G--C---CAAG--AA-----AG---GTG-----CAGA--G--T---TC--GTCTGTTGG--T--TGT	1264
MND-1.GA.-.MNDGB1	---T--A---TGT---G---G---CC--CTGT---TGT--AAG--T--AT---TA---T--GCT--ATCGGA--AAA--CC--T-----CAGC---A---GA--A---C--CTC--GTATGTTGG---TAT	651
MON.CM.99.L1	---T---G---G---CCTGA--T--G-----TGC--A-----T--TA-----GTC--CTCGGACTCG--CC--G--G---CAG--C-----AGA--G--TT---A--GTTA--T--TT--TCTA-	646
MON.NG.-.NG1	-----G-----CCTGA--C--G--C--TA-----TTC-----T--TA---T--G--C--CTC--GA--TC---T--G---T--AGG--T--G---CA--A--G--CA--A--AGGTCA--C--TC--C--TAG	651
MUS.CM.01.1085	---C--A--G--G---TG---G---CC--C--C-----G--T--A---A---CA---T--GCC--GCAG--A--CAT--A-----ATAAG-----CAGG---TT--ACTAG--T---G--T--TAG	689
DEB.CM.99.CM40	---C-----G---TTGA---G---CC--C--C-----T---A-----CA---T--GCC--GCAGGA--CAT--A--G---CAA-----A---CA--G--G---CT--AACTAGAT-----T--TAG	692
SYK.KE.-.KE51	---G---GCCG--CCT--G--TCG---T---T-----G--A--A-----G--CA-----GCC--CTCAGACCAG--TC--T---GTAA-----T---GA-----CA--AAA--ACCA--CTT---TAG	636
SYK.KE.-.SYK173	---G---GCCG--CCT--G--TCG---T---T-----AG--A-----T--G--CA-----GCC--TCGGACCAA--CA--G--GT--TAA-----GA--A---T--CAC--GTC---TT---TAG	999

PLV  
complete genomes

H1B.FR.83.HXB2	TTCAGACAGGATCAGAAGAAGCTTAGATCATTATATAATACAGTAGCAACCCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAA.....	1113
Gag	L O T G S E E L R S L Y N T V A T L Y C V H Q R I E I K D T K E A L D K I E E E Q .....	p17
H1A1.UG.85.U455	-C-----A-----GT-----A-----TG-----A-T-A-----AT-----	580
H1B.US.90.WEAU160	-----A-----G-----A-----T-----G-----T-----A-----A-----G-----T-----	1112
H1C.ET.86.ETH2220	-----A-----G-----A-----T-----G-----T-----A-----A-----C-----T-----A-----	493
H1D.CD.84.84ZR085	-A-----G-----AG-----T-----GT-----A-----A-----GC-----G-----C-----C-----A-----	630
H1F1.BE.93.VI850	-----A-----G-----A-----T-----G-----GT-----A-----A-----GC-----G-----C-----C-----A-----	468
H1G.SE.93.SB6165	-C-----A-----G-----A-----T-----G-----GT-----A-----A-----GC-----G-----C-----C-----A-----	516
H1H.CF.90.056	-A-----A-----A-----T-----CT-----GT-----C-----A-----G-----AA-----TG-----G-----T-----G-----T-----ATA-----	460
H1J.SE.93.SE7887	-C-----A-----C-----A-----A-----G-----T-----C-----A-----G-----A-----A-----T-----G-----T-----G-----T-----ATT-----	436
H1K.CM.96.MP535	-A-----A-----GA-----A-----C-----A-----GT-----T-----A-----A-----G-----C-----AC-----A-----	324
H1O1.AE.TH.90.CM240	-CA-----T-----A-----T-----T-----GG-----C-----A-----C-----G-----T-----A-----TA-----	678
H1O2.AG.NG.-.IBNG	-CAG-----A-----T-----A-----T-----GG-----C-----C-----C-----C-----T-----A-----G-----ATA-----	653
H1N.CM.95.YBF30	-CA-----G-----GC-----C-----T-----CG-----TC-----GTT-----T-----C-----AGT-----C-----C-----CA-----AC-----G-----G-----A-----A-----GA-----A-----G-----GAACA	676
H1O.BE.87.ANT70	-CA-----G-----CAGC-----GCAG-----C-----C-----GG-----CG-----A-----T-----GTG-----GG-----T-----CA-----C-----ATATA-----A-----TGG-----T-----GC-----C-----G-----AA-----C-----A-----T-----A-----TAATG-----	1165
H1O.CM.91.MVP5180	-CA-----G-----G-----C-----G-----A-----TC-----C-----GG-----CG-----A-----GTA-----GG-----C-----T-----CA-----C-----AT-----T-----C-----CCG-----T-----AC-----C-----G-----AA-----C-----A-----T-----A-----TAATG-----	1140
CPZ.CD.-.ANT	-AG-A-T-A---CCCT---A-A-T---T-G-T---C-CA-TTGTGTT-G-GG-C-A---A-G-GA-A---AG-AC---CG-TA-A-CAG-GA-AATGA-AGT-----	477
CPZ.CM.-.CAM3	-CA-----GC-----GC-----T-----TC-----GTG-----T-----GG-----T-----A-----G-----A-----TTTCA-----G-----G-----AC-----C-----G-----ACA-----AT-----GA-----A-----CAGTT-----	503
CPZ.GA.88.GAB2	--A-----G-----T-----AC-----G-----C-----T-----TT-----G-----TGTGT-----A-----GG-----T-----GCT-----A-----AG-----G-----T-----AC-----C-----C-----A-----T-----GA-----A-----TTTGT-----	485
CPZ.GA.-.CPZGAB	-CA-----C-----G-----GC-----C-----G-----T-----C-----TC-----G-----GTA-----G-----GG-----CA-----A-----AGTGAC-----CACTG-----G-----C-----AC-----C-----A-----G-----A-----G-----GAACA	1141
CPZ.US.85.CPZUS	-AA-----A-----G-----GC-----GC-----C-----T-----C-----CT-----GGT-----T-----GG-----C-----CAGT-----G-----GACAG-----GG-----G-----T-----GC-----C-----G-----TG-----AC-----GA-----A-----G-----TA-----	1152
CPZ.TZ.-.TAN1	-AG-C-TT-T-TCC---A-C-TT-T-G-T---CA-CTGTGA---C---A---CGC-G-AGA-AGAG-CC---T-GG-A---AG-CA-A-TTG-GA-AATGA-A-T-----	676
H2A.GW.-.ALI	-G-CA-----A-----TT-----A-----AGCC-----T-----T-----CTGCGTAG-----T-----GG-----C-----CGC-----GAAGAGA-----AG-----G-----T-----TG-----A-----G-----AAA-----C-----A-----TA-----C-----GAGA-----TCTAGCGGC	1424
H2A.DE.-.BEN	-A-CA-----G-----A-----TT-----A-----AGCC-----T-----T-----C-----CTGCGT-----A-----T-----GG-----CT-----CGC-----GAAGAGA-----AG-----G-----T-----TG-----A-----AAAGA-----ACTAGC-----C-----GAGA-----TCTAGTGGC	1431
H2A.SN.-.ST	-A-C-----G-----A-----TT-----A-----AGCC-----T-----T-----T-----CTGCGT-----A-----T-----GG-----A-----A-----CGC-----GAAGAGA-----AGCC-----T-----TG-----A-----AAA-----C-----A-----G-----C-----GAGA-----TCTAGTGGC	876
H2B.GH.86.D205	-A-CA-----G-----A-----TT-----A-----AGCC-----T-----T-----C-----CTGCGT-----A-----T-----CC-----CGC-----GAAGAGA-----AG-----G-----T-----AG-----A-----A-----G-----TAGC-----C-----GAGA-----TCTAGCGGC	1407
H2B.CI.-.EHO	-A-CA-----T-----A-----TT-----A-----AGTC-----T-----C-----CTGCGT-----A-----T-----CC-----CGC-----GAAGAGA-----AG-----G-----T-----TG-----A-----AAA-----A-----G-----TAGC-----C-----GCGA-----TCTAGCAGC	1408
H2G.CI.-.ABT96	-A-CA-----T-----A-----TT-----A-----AGAGCC-----T-----T-----CTGCGT-----A-----T-----GG-----T-----CGCCGAACAGAG-----AG-----C-----T-----TG-----A-----AAA-----C-----AGTAG-----G-----C-----GAGA-----TCTAGTGGT	794
H2U.FR.96.12034	--CA-----T-----A-----TT-----A-----AGAGCC-----T-----T-----C-----CTGCGT-----A-----T-----GG-----T-----CGC-----GAAGAGA-----AG-----G-----C-----T-----TG-----A-----AAAGC-----G-----TAG-----C-----GAGA-----TCTAGCGGT	910
MAC.US.-.239	-G-CA-----C-----A-----TT-----A-----AGCC-----T-----T-----CTGCGT-----A-----GG-----CA-----T-----CGC-----GAAGAGA-----AG-----G-----C-----TG-----AAA-----C-----G-----TAG-----G-----C-----GAGA-----CCTAGTGGT	1381
Gag	V P T G S E N L K S L Y N T V G V I W C I H G E E K V K H T E E A K Q I V . Q R H L V V	p17
SMM.SL.92.SL92B	-G-CA-----C-----G-----A-----TT-----A-----AGC-----G-----T-----G-----C-----CTGCGTAG-----T-----GG-----C-----CGC-----GAA-----GA-----AG-----G-----T-----AG-----AAGA-----A-----CAG-----C-----GAGC-----TCTAGTGGT	821
SMM.US.-.H9	--C-----T-----A-----TT-----A-----AGAGCC-----T-----C-----T-----CCGCGT-----T-----GG-----CA-----T-----CGC-----GAAGAGA-----AG-----G-----C-----T-----TG-----A-----AAA-----C-----A-----TAG-----G-----C-----GAGA-----TCTAGTGGT	863
STM.US.-.STM	-G-CA-----T-----A-----TT-----A-----AGAGTC-----T-----T-----C-----CTGCGT-----A-----T-----GG-----A-----C-----CGC-----GAAGAGA-----AG-----G-----C-----T-----TG-----A-----AAAGC-----AGT-----G-----A-----GAGA-----TCTTGTGGT	1037
SAB.SN.-.SAB1C	--CT-----G-----GA-----TT-----A-----TTG-----TC-----C-----T-----TT-----TGCTGCGTGT-----AGC-----A-----T-----CGC-----GAA-----A-----AG-----G-----T-----AG-----AAA-----CG-----AG-----A-----A-----G-----AGTCCAGCAGA	1348
TAN.UG.-.TAN1	AA-CA-----G-----A-----A-----AGTC-----G-----T-----TTGTGCTGCGT-----A-----T-----GG-----A-----T-----CGC-----GAACAGA-----AG-----G-----T-----AG-----AG-----G-----TATTAG-----A-----AC-----CGCTGCCATCTAGT	1233
VER.KE.-.AGM155	AA-CA-----G-----G-----G-----A-----G-----A-----AGTC-----G-----CTG-----GTGCGTAT-----GCT-----C-----C-----GAA-----AGA-----AG-----G-----AG-----CA-----TAG-----AGAC-----TGCTGCCACCTAGT	1262
VER.KE.-.9063	AA-CA-----C-----G-----G-----GCT-----A-----A-----AGTC-----G-----TC-----CTT-----GTGCGTG-----A-----CA-----C-----CGCAACAGA-----AG-----G-----AG-----GG-----CA-----CAG-----AGAC-----TGCTGCCATCTAGT	1273
VER.DE.-.AGM3	AA-CA-----G-----G-----GCT-----A-----A-----AGTC-----G-----T-----CCTT-----GTGCGTAT-----G-----T-----C-----C-----A-----GAT-----AG-----AG-----G-----AG-----CA-----TAG-----AGAC-----TGCTGCCATCTAGT	762
VER.KE.-.TY01	AA-CA-----G-----G-----GCT-----A-----A-----AGTC-----G-----TC-----CTT-----GTGCGTA-----A-----CT-----CA-----GAAACAGA-----AG-----G-----AG-----AG-----CA-----CAG-----AGAC-----C-----CTGCCATCTAGT	763
COL.CM.-.CGU1	-GG-C-----TG-----G-----G-----CC-----GGA-----T-----TGTT-----TGTTGT-----C-----CC-----CAG-----AATGGA-----T-----TG-----G-----TC-----A-----G-----AGA-----A-----G-----AG-----A-----CTT-----TAAAAAGCA	932
DEN.CD.-.CD1	AA-CACA---CAGT---GAGT---G-A-AGCC---TC-GA---TGCTGCCTTG---GG---TGC---CAGGG-AT-G---AG-GGC---T---AG---G---AC---AGAG-AG---AGAA-GAGAAT-GGTGCAA	813
GRV.ET.-.GRI_677	AA-C-----C-----C-----G-----GG-----A-----A-----G-----TC-----G-----T-----TTGTGCTGCGT-----A-----T-----GG-----CA-----T-----CGC-----GAACAGA-----AG-----G-----AG-----G-----G-----TA-----CAG-----TA-----C-----CT-----CCATCTAGT	1228
GSN.CM.99.CN166	AA-CA-----G-----CT-----C-----A-----TC-----C-----T-----GGA-----TC-----GT-----GTG-----A-----C-----GC-----GAC-----C-----G-----GG-----G-----T-----AG-----C-----G-----AAGC-----GC-----AG-----CAG-----AT-----AGATGTCATCTAGA	783
GSN.CM.99.CN71	AA-CA-AT---G---C---TCT---C-A-AGCC---GGA---TC---GT---GTG---A---C---GC---GAC---TG---G---G---T---AG---C---G---AAGA---GC---AG---TC---AT---AG---TGTCATCTAGT	789
DRL.-.FAO	--CT-----G-----GA-----TT-----A-----TT-----TC-----C-----T-----CTGCTGCGTAG-----T-----GG-----C-----T-----CTCG-----AAG-----GA-----AG-----C-----AG-----G-----TC-----A-----AG-----A-----AC-----C-----CTGCCATCTAGT	622
RCM.GA.-.GAB1	--C-----G-----G-----GG-----A-----A-----AGCC-----C-----T-----TTGTGCTGCGTA-----GG-----C-----A-----CA-----GGAAG-----GA-----AG-----G-----AG-----G-----CA-----AG-----GA-----A-----TGCTGCCATCTAGT	543
RCM.NG.-.NG411	--C-----G-----G-----T-----G-----GCT-----A-----A-----CC-----C-----T-----TTGTGCTGCGTA-----TG-----C-----CA-----GGAAG-----GA-----AG-----G-----AG-----G-----GG-----G-----CGC-----TG-----A-----AA-----GC-----TGCCATCTAGT	543
MND-2.-.5440	--CT-----G-----A-----TT-----A-----TT-----GC-----G-----C-----C-----CTGTTGTTG-----A-----T-----GG-----C-----A-----GCG-----AAG-----GA-----AG-----C-----T-----TG-----AAA-----G-----AG-----A-----AC-----CG-----T-----CCACTAGT	625
MND-2.CM.98.CM16	--CT-----G-----A-----TT-----A-----TC-----GC-----G-----C-----C-----CTGCTGCTG-----G-----A-----GG-----C-----A-----GCG-----AAG-----GA-----AG-----C-----T-----AG-----GAA-----G-----AG-----A-----AC-----A-----TCCATCTAGT	1068
MND-2.GA.-.M14	--CT-----G-----G-----A-----T-----CTTG-----GC-----G-----C-----C-----CTGTTGTTG-----A-----GG-----C-----A-----GCG-----AAG-----GA-----AG-----CT-----T-----AG-----A-----G-----AG-----A-----AC-----TGCTTCCATCTAGC	992
MNE.US.-.MNE027	-G-CA-----C-----A-----TT-----A-----AGAGCC-----T-----T-----CTGCGT-----A-----GG-----CA-----T-----CGC-----GAAGAGA-----AG-----G-----C-----TG-----AAA-----C-----G-----TAG-----G-----C-----GAGA-----CCTAGTGGT	861
LST.CD.88.447	ACTCC---C---GA-GA-TT-A-A-G---C-TGTAGGA---T---TGTGTG---G---G---CTGC---TT---G---A---AG---GC---C---A---A---GC---A---A---AAT---A---C---ACTCAGGA	332
LST.CD.88.485	ACTCC---C---A-GA-TT-A-A-G---C-TGTAGGA---T---TGTGTG---G---G---CTGC---TT---G---A---AG---GC---A---A---GC---A---A---AAT---A---C---ACTCAGGA	332
LST.CD.88.524	A-CT---C---TA-A-T---G-A-G-TC---GTAGGG---GTGTG---A---G---CTGT---C---T---G---A---AGAG---C---T---T---AC---AA---TA---T---AG---GA---AAT---GAACAAACC	332
LST.KE.-.lho7	A-GCA-T-T-TA-GA-C-A-AGG-C---GTAGGG---TGTGT-A---GC-CTGT---C-T-G-AG-AGA---TC---TC---CA-CAGT---AG---A---AAT---A---C---GCACCAAGA	1407
SUN.GA.98.L14	A-GCAT-T-T-TA-GA-C-G-AGG-C---GTAGG---TGTGTG---TGCC---TGT---G---C---TCCAG---AC---G---C---A---A---G---CAGACT---C---GCAAAAAA	1394
MND-1.GA.-.MNDGB1	A-G-CCA---TA-T-A-A-G-T-T-GGTAGGG---CTGTGTTG-AGCC---CA-A-CGC-G-T---A-T-GAG---AC-A-T---A-A-T-A-A-TCATAAC-AGAAAGGA	781
MON.CM.99.L1	AG-C-----C-----GTCTA-----C-----A-----TC-----C-----T-----GG-----TC-----CTT-----GGTAT-----G-----CA-----A-----GCCG-----T-----A-----TG-----G-----AG-----AC-----G-----AAGC-----GC-----AG-----A-----AATCAG-----TGTCACCTTAGC	776
MON.NG.-.NG1	.....	0
MUS.CM.01.1085	AA-CAG---G---C---GTCC---C---A---C---C---T---GGA---TC---GT---GTG---GG---A---C---CGC---GA---G---GC---TG---GG---AG---G---AAGA---GC---A---CAGA---C---AG---TGCCATCTAGA	781
DEB.CM.99.CM40	AAGCA---CAGT---GA---TT---G---A---TC---GGC---T---CTGTG---A---GGGATGC---GCC-ACC-CTC-G-GG-T---G---G---CAA-A-GGCAG-GT-CAGGAG-AT-AAGCAGGT	819
DEB.CM.99.CM5	AGGCAT---CAGT---GC---AG---CC---CGGC---T---TGTGT---G---GGGATGC---CGCC-ACC-CTC-G-G-G---T---G---G---AAA-A-GGCAG-G-CTAG-AG-AT-AAGCAGGC	822
SYK.KE.-.KE51	AAA-ACAT---C-----A-----TT-----A-----A-----T-----GGC-----T-----AC-----GTAG---GGCC---C---CGC---GAA---A---CG---GG---T---GG---AC---AAA---C---G---AG---GA---A---G---CTTGCATTTGGGA	766
SYK.KE.-.SYK173	AAGCT-AT---G-----A-----TT-----A-----A-----AGCC-----T-----T-----GGC-----T-----A-----TT-----GGT-----G-----A-----GGCC---T---GCC-AA-AG-AG-G-----TG---C---AAA---C---G---AG---A---A---CATGTAATTGGAA	1129



H1B.FR.83.HXB2	... AACAAAAGTAAGAAAAAGCACAGCAAGCA... GCAGCTGAC... ACAGGA...	1158
Gag	... N K S K K K A Q O A ... A A D ... T G ...	p17
H1A1.UG.85.U455	... -T-G-AC--C--GGA-----G-----A-----	625
H1B.US.90.WEAU160	... -C-----G-----A-----	1157
H1C.ET.86.ETH2220	... -G-----C-C-----A-----G-----G--A-CT... GACA--	538
H1D.CD.84.84ZR085	... -C-----C-----GG-----G-----	675
H1F1.BE.93.VI850	... -C-----C-----GA-----G-----	513
H1G.SE.93.SB6165	... -G-----C-G-----AT-----G-----ATG--T... A-----	561
H1H.CF.90.056	... -C-----C-----A-----T-----	505
H1J.SE.93.SB7887	... -G-AC-AC-GC-G-----A-----A-A-----	481
H1K.CM.96.MP535	... -C-----C-G-----GA-----A-----A-GCAGCT...	360
H101.AE.TH.90.CM240	... -T-G--CC--CG--GA-----G-----G-----	723
H102.AG.NG.-.IBNG	... -T-G--C--C--G-T-----G-----C-----	698
H1N.CM.95.YBF30	GCAC--GCCCGAGCCA--CC--G-AGC--GG-----G-CA... G--ACTGAT...	728
H1O.BE.87.ANT70	... GGG-GC--G--TCTGCG-ACGCCGCTAAG... A--ACACA... -GC...	1207
H1O.CM.91.MVP5180	... GCA-GC--G--TCTGC--AGGCCGCTAAG... A--AAACA... -GC...	1182
CPZ.CD.-.ANT	... -TGC-G-CAC-AGC-G--A--GGAAGTAGCCAAACCGCAAGCAG--GCATGCTTCTGCGGTGCTCCTGTTA-ACAA-CAG... TGGTGTGAGCGA...	570
CPZ.CM.-.CAM3	... -CGGCTCCCGAG--C-G-A-GT-GCCCA... C-GCAGC-G... CAGCAGCAA...	551
CPZ.GA.88.GAB2	... -C-A--GGAA-CAGC--C-----G-----AA-G... CA-A-TGAG...	533
CPZ.GA.-.CPZGAB	... -CGGCA-C-TGG-G--CA--AGCAA-ACTGAAAGTAACT--GAAG-CGTGAAGGGGAGCCAGTCAAGGC-CTAGTGCC... TCT...	1225
CPZ.US.85.CPZUS	... GTGC-GTTG... C-GAA--A--G-A--A-AAA--ACAACAGCAACAG... GAG-C-AGT... GGC...	1209
CPZ.TZ.-.TAN1	... -CTGT-CAG--A--T--CT-CACAGCGA--... TCTAGT-G-CAAGA... CAGAATGCAGGT-A-AAAGAGGAAACAGTGCCACCTAGTGGCAATACAGGAAACACAGGGAGA...	784
H2A.GW.-.ALI	AGAA... AT--AAACA... G--AGAAAATGCCAAGCACAAGTAGACCAACAGCA...	1473
H2A.DE.-.BEN	AGAA... A-T-GAACT... G--AGAAAATGCCAATAACAAGTAGACCAACAGCA...	1480
H2A.SN.-.ST	AGAA... A--AAAAC... A--AAAATGCCAAGTACAAGTAGACCAACAGCA...	925
H2B.GH.86.D205	GGAC... A--AAA-A... -TGCC-GCTACAAAATAAACCAACAGCA...	1447
H2B.CI.-.EHO	GGAC... A--AAA-A... -TGCC-GCTATGAGTAAACCAAGTAAA...	1448
H2G.CI.-.ABT96	AGAA... A-T-GAACT... -A-AAAGTGCCTGCCAAGCAGACCAATAGCA...	843
H2U.FR.96.12034	GGAA... -T-A-CTGCAGAAAAAATGCCAGCCACAAGTAGACCAACAGYT...	959
MAC.US.-.239	GGAA... A--GAACA... -A-ACTATGCCAAAAACAAGTAGACCAACAGCA...	1430
Gag	... E ... T G T ... T E T M P K T S R P T A ...	p17
SMM.SL.92.SL92B	AGAA... AGT-GAACT... G--A-AAATTGCCAGCTCAAAGCAGACCAACAGCT...	870
SMM.US.-.H9	GGAA... A--GAACA... G--ACAAAATGCCAGCAACAAGTAGACCAACAGCA...	912
STM.US.-.STM	AGAG... A-T-GAACT... G--AACAAAATGCCTGCCACAAGTAGACCAACAGCA...	1086
SAB.SN.-.SAB1C	AATG-CAG----GCC-C-GCGA--TCTAGT-GC... CA-A--A-G... GA-CTGCAGGCAAAAAAGAAAATGAGCCAAACAGTG...	1427
TAN.UG.-.TAN1	GGAA--AG--AA-CTGC-GC--G-CATCT-GT... -GCCAGC-G... CA-AATTACAACACAGCTGCG...	1297
VER.KE.-.AGM155	GGAC--AG--AA-CTGC-GTTA-G-CA-CT-GT... -G-CAGC-G... -A-AATAACACAGGAGGAACAGCG...	1329
VER.KE.-.9063	GGAA--GG--AA-GAGC--C--AG-CATCTAGT... -G--AAA-G... -A-A-TAACAGGGAGACAACAGCG...	1340
VER.DE.-.AGM3	GGAG--AG--A--TGC-G--AG-A-TAC-A--AG-ATCT... -GT--CAAAAAGAAAATGACAAGGGAGTAAACAGTG...	841
VER.KE.-.TY01	GGAA--AG--AA-GTGC--C--AGACATCTAGT... -G-CAAA-G... -A-AATGACAAGGAAATAGCAGCG...	830
COL.CM.-.CGU1	GGCC-TG-T-GAA-T-GCC... ..AGC	954
DEN.CD.-.CD1	AAAG--G--C-TAGTT-TTGGCAATGCAG-CCA-GGAGATGATCAGC-TC-AC-GGGA... .GCCCGGGAGGTGGCAGCTCCGCGGAACAAGGAGCCTCTGCG	913
GRV.ET.-.GRI_677	GGAC--A--TGAG--AGC-GCT... -A-AAGAAAATGAGACAACAGCG...	1274
GSN.CM.99.CN166	GGAG-C--GG--AG--GCAGA-TA--AC... CCCCCG	823
GSN.CM.99.CN71	AGGA-CA--GAG--AG--C-CAGA-CA--AC... CCCCCA	829
DRL.-.-.FAO	GGAT--A--TGAG--TGC-GCTT--A-AA-T-A-AATGGGAAACA--GA-ATCTAGCGGCAGG...	686
RCM.GA.-.GAB1	GGAA--AGC-GAA--T-C--C--A-A-AG--AAG... -G--AAACA...	586
RCM.NG.-.NG411	GGAC--AG-TGAG--TGC-GG--A--AG--AA... -G--CACA... GTGAC...	592
MND-2.-.-.5440	GGTTG-A-G-GAG--TGC-GCCT--G-AG--A-AAAGGAGCAACA--GA-ACCTGCTGTTGCG...	689
MND-2.CM.98.CM16	GGCCG-A--GGAA--TGC-GC-T--G-AA--A-CAAAGAGCAATA-TGA-ACCTAGTGGCCG...	1132
MND-2.GA.-.M14	GGCC--AGG-GAA-GTGC-GCCT--G-AA--A-GAGAAAGCAACA--GA-ACCTAGTGGCCG...	1056
MNE.US.-.MNE027	GGAA... A--GAACA... G--A-ACTATGCCAAAAACAAGTAGACCAACAGCA...	910
LST.CD.88.447	AGAA-GGC-G... ..342	342
LST.CD.88.485	AGAC-GGC-G... ..342	342
LST.CD.88.524	TGTG... ..336	336
LST.KE.-.lho7	AGCAGCAGGG... ..1417	1417
SUN.GA.98.L14	TGAAG-A... ..1401	1401
MND-1.GA.-.MNDGB1	AGAA--GC-GGAGG-TG--... ..800	800
MON.CM.99.L1	AGGGG-AC-GG-AG-AC-G----TGCAGCC--GGCCGCT... ..816	816
MON.NG.-.NG1	... ..0	0
MUS.CM.01.1085	GGAG--A--GG-AG-A--G--T--AGCGCAGGCAGCT... ..821	821
DEB.CM.99.CM40	GGAAC-GG-GGAAG-ATT-G--ATGCC-T-ATCCAAGACAGAAA-AGAAG... GAAGTAGAACAGAAAACAACAAAAGCAACAGGAAACAAAGACAGCAG	916
DEB.CM.99.CM5	AGAGC-GG-GGAAG--TT-G-GATGCCATT-ATTGAGAAAGCA-A-AAG... GAGTTAGAGAAAAG... CAGAAAACCTCAGGCAGCAGCAGCAGAC	916
SYK.KE.-.KE51	AGATG-GG--CAGT--C-TCTAGTGGC--AA... -A-AA-AG... -AC-ACACAGCGACATCTAGTGGCAGAGAAGGA... AAAATGCAGCTGCCAGCA	860
SYK.KE.-.SYK173	AGATG--CC-CCAGC--C-TCT-GTGA--AGT... -A-AA-AG... -GTCA-AACATGGCTAGTGAGACATCTAGTGGCCAAAAGGTAGTACAGCAGGAA	1226



PLV  
complete genomes



	Gag p17 \ / Gag p24	
H1B.FR.83.HXB2	.....CACAGCAAT...CAGGTCAGC...CAAAATTAC...CCTATAGTCAGACAACATCCAGGGCCAAATGGTACATCAGGCCATATCACCTAGAACTTTAAATGCATGGGTAAAAGTAGTAGAAG	1270
Gag	.....H S N . . . . . Q V S . . . . . Q N Y . . . . . P I V Q N I Q G Q M V H Q A I S P R T L N A W V K V V E	p24
H1A1.UG.85.U455	.....AG--T.....-C-----A-TGCA-A-----CCA-----C-----T-----G-C-G-----G-----	725
H1B.US.90.WEAU160	.....A--A-C.....-C-A-----A-----C-----C-----G-----A-----	1269
H1C.ET.86.ETH2220	.....GGA-AGG-C.....-T-----T-----T-----G-----C-----G-----G-----	644
H1D.CD.84.84ZR085	.....A---GC.....-T-----A-----C-A-A-----C-----C-----G-C-----A-----	787
H1F1.BE.93.VI850	.....A-----T-----A-----TC-T-----A-----C-----T-TC-----G-----G-GA-----	613
H1G.SE.93.SE6165	.....A-----C-----A-----T-----A-----TGCA-A-----C-----A-----G-----	673
H1H.CF.90.056	.....A-AGA--C...A-----T-----A-----TGCT-A-----G-----C-----G-C-----GA-----	617
H1J.SE.93.SE7887	.....G--A-G-----T-----T-----TC-G-A-----CC-----C-----C-----G-----GA-----	593
H1K.CM.96.MP535	.....GA--A...GG-----T-----A-----TC-G-----C-----C-----G-G-A-----G-----	469
H1O1.AE.TH.90.CM240	.....AG---GC...A-A-----A-----TGCA-A-----C-----C-TT-----G-----G-----	835
H1O2.AG.NG.-.IBNG	.....AGCAG-----AGCAG-----A-----TGCA-A-----AC-----T-----G-----A-----	801
H1N.CM.95.YBF30	.....GC...A-TA-----T-----AGG-----T-----C-----C-----CTGCT-A-----C-----G-C-----G-----GA-----G-----	834
H1O.BE.87.ANT70	.....GCA-GG-----CGG-T-----ATCA-TGCC-----A-----C-----C-----C-G-----G-C-----	1316
H1O.CM.91.MVP5180	.....CCT-GG-----ACA-T-----AACA-TGCA-----A-----A-----C-----C-----G-----G-C-----	1291
CPZ.CD.-.ANT	.....CATCT-AGTGGC-AAG-G...AG-----C-----CA-AGT-G-TGCAGGA--AAT-GCAAGG-----C-AC-GA-----A-----C-----GTGT-----	682
CPZ.CM.-.CAM3	.....GA-GC...GCT---T...AG--C-----AG--A-----TGCA-A-A-GC-A-G-----C-----G-----C-G-----T-----T-----A-----	660
CPZ.GA.88.GAB2	.....TCA-A--CCACTG--A-A...AG-----AG-G-----TGCA-A-A-G-----G-----G-----C-----C-----G-----C-----T-----C-----	648
CPZ.GA.-.CPZGAB	.....GCTGGC...A-T-T...GG-----CC-C-A-A-TGCT-A-----G-----G-----C-----C-----G-----G-----	1331
CPZ.US.85.CPZUS	.....-T-C...ATT-GT---AGC-----AG-A-A-----GCT-----A-G-----A-----G-T-----C-----C-----CG-----	1318
CPZ.TZ.-.TAN1	GCAACAGAGACACCT-G...GG---T...AG-CTA---AG-GA-AACTG-TGCA-----AGTTGCAAGG-----C-T-T-----C-----C-----GG--A-----	902
H2A.GW.-.ALI	.....CACCT-G...G-ACAGG-G...GG--C-T...-CG-CAA--AGTAGC...-CA-CTACACC--GT-C-GC-GAGC-CC--C--G--T-----T-----	1582
H2A.DE.-.BEN	.....CACCT-G...GG-AAA-A...GG--C-----CG-GCAA--AGCGGGT...-CA-CTAT--C--GT-C-AC-GAGC-CC--C-----T-----G-G-----	1589
H2A.SN.-.ST	.....CACCT-GC...GG-AA-G-A...GG--C-T...-CG-CAA--AGTGGC...-CA-CTATACC--GT-C-AC-AGT-CC--CC--T-----C-----G-----	1034
H2B.GH.86.D205	.....CACCT-GC...G-----GG-----T...-AG-GCA--ACTGGCT...-CA-CTAC--C-C-T-C-GC-AGC-CC--C-----T-----GT-----	1550
H2B.CI.-.EHO	.....CCA-C...AGC...C...-T-GC--T...-AG-GCA--A-TAGCT...-CA-TTATCC--C-T-C-GC-AGC-CC--C-----C-----T-----	1551
H2G.CI.-.ABT96	.....CACCT-G...GGCAGAG-A...GG--C-----AG-GCA--AGTAGG...-CA-CTAT--T--T-C-AT--AGT--C-----C-----T-G-----	952
H2U.FR.96.12034	.....CACCT-G-GGCGCAGA...GG-----GG-GCA-----TAGGT...-CA-TTAT-CC-C-TAC-TC--AGT-----AC-C-----C-----G-----G-----	1068
MAC.US.-.239	.....CATCT-GC...GGCAGAG-A...GG-----AG-CAA--A-TAGGT...-TA-CTAT--C-C-T-C-AT--AGC-G-----A-----C-----T-GA-----G-----	1539
Gag	.....P S S . . . . . G R G . . . . . G N Y . . . . . P V Q Q I G . . . . . G N Y V H L P L S P R T L N A W V K L I E	p24
SMM.SL.92.SL92B	.....CACCT-G...G-A-A...GG--C-T...-AG-GCA--AGTAGGA...AATA-TTAT--G-CACTC-AC-G-C--AC-G-AC-C-----G-----C-G-----	973
SMM.US.-.H9	.....CACCT-G...GGCAGAG-A...GG-----AG-GCA--AGTAGGT...-CA-TTATACC--C-TAC-TT--AGT--A-----A-----T-----T-G-----	1021
STM.US.-.STM	.....CACCT-G...GGCAGAG-A...GG-----AG-CA--AGTAGG...-CA-TTAT--C-C-TAC-AT--AGC-A-----C-G-----GT--G-G-----	1195
SAB.SN.-.SAB1C	.....ACACCT-G...GGT-G-TCG...G-----C--T--CAGTGT--AT...AAT--TG--C-C--C-AT-G-C-A-G--AC-T-----T-----GA-T--G-----	1536
TAN.UG.-.TAN1	.....ACACCT-G...GGCAGACAT...GGT-----G-----A-AC-G-AT...AAT--TG--ACTC-TT-G-----AC-T--CC-----T--AC--T--G-----	1406
VER.KE.-.AGM155	.....ACACCTGG...GGC...TT...-CGC-CAA--C-AGGG...AATGC-TG--G--GTAC-AC-T-----C-C--CC-----C-----	1435
VER.KE.-.9063	.....CACCTGG...GGC...T...-AGCGCAA--AC-AGGA...AATGCCTG-----GTTC-AC--C--C-C-A-----G-----C-----G-----	1446
VER.DE.-.AGM3	.....CCACC...GGT-G--T...-AGC-CAA--AC-GGA...AATGC-TG-A--GT-C-T-G-----AC-C-C-----G-----C-----G-----	947
VER.KE.-.TY01	.....CACCTGGG...TG--T...-AGCGCAA--AC-AGGA...AATGCCTG-----GTAC--T-G-----GC-C-C-----G-----C-----G-----	936
COL.CM.-.CGU1	AAGGAGGAGG-G-AAGCAAAA--AAGCAGAGA--TTGG--ATG--A-----AACAGGGCCT...CAAGGGCCA--G-----AC-C-GAGC-C-G-G--GGA--C-----T--GTGT-----G-----	1080
DEN.CD.-.CD1	GGCGGAAGTTC-TTT-GAGAT-CAAATTCT...-A-T-CAG-CAGCT-----A--TATCAG-----A--T-AGTA-C--GGTG-----AA--A-T--AG--G-----	1037
GRV.ET.-.GRI_677	.....CACCTGG...GGC-AATCA...AG-----AG--AA-TC-G-AT...AATGCCTG-----C--C-TT-G-T--GC-C-G-----G-----C-----G-----G-----	1383
GSN.CM.99.CN166	GGGGCCGAGGG-AC-GGCGTA--T--T...-G-GA-TAG--TGC--A-A--GTATCA-----C-CAATAG--T-C-T-GA-T-----GCA--G--G-----	947
GSN.CM.99.CN71	GGGGCCGAGGG-AC-GGCGTA--T--T...-G-C-----C-----TAGA-TGCT-----A-GT-TCA--C-----C-CAACAGC--GTCC-T-GA-T-----GCA--G-----	953
DRL.-.-.FAO	.....TCA...AG--C-----C--TCA-GTTGT--AT...CAA--GGCA--C-----T--T-----A-----C-----G--A-T--G-----	780
RCM.GA.-.GAB1	.....GCG-CACCT-G...GGACAA--A...GG-----T...-A-TACT-TA-AT...CA--GCCT-AG--A-TC-T-----A-G--C-----C-----C-G--G-----G-----	698
RCM.NG.-.NG411	.....TCT-G...GGACAA--A...GG-----A-----AA-CA-TT-AT...CA--GCCT-A-----C-A-----T-----A-----T-----G-----	698
MND-2.-.-.5440	.....TCA...A-----C--TCA-GTA-TA-AT...CAAAC-CCA-----GA-T-----A-----C-G-----TGTA--G-----	783
MND-2.CM.98.CM16	.....TCA...A-----A--CA-ATA-TA-AT...CA-ACCCA--C-----GA-T-----GC-C--GC-G-----G-----TGCA--G-----	1226
MND-2.GA.-.M14	.....TCA...A-----T...-TCA-ATA-TA-AT...CAAAC-CCA--G-----A-GA-T-----A-----C-----C-----TGTA-T--G-----	1150
MNE.US.-.MNE027	.....CATCT-G...GGCAAA--A...GG-----AG-CAA--A-TAGGT...-TA-CTATACC--C-TAC-AT--AGC-G-----A-----C-----T-GA--G-----	1019
LST.CD.88.447	.....CTAAA--T...GGC-----T...-AC-C--CAGAG-A-AT...CA-AG-TG-----CACAC-AT-G-----AA--C-AA-T-----A-----	442
LST.CD.88.485	.....AAA--T...GGC-----T...-AC-CA-CAGAG-A-AT...CA-AG-TG-----CAC-C-AT-G-----AA--C-AA-T-----A-----	442
LST.CD.88.524	.....CA-AA--T...A-----TT...-T-GA-TAGAG-A-AT...CAAAG-TG--C--ACCC-TC-----C-----AA-TC-AA-----G--A-----	436
LST.KE.-.lho7	.....A-G-AAC-A...-ACAG-G...GG-----T...-CC-GA-AAGAG-A-AT...CA-AG-TG-----ACAC-TT--G-C-----AA--C-AA-----G--A-----G-----	1526
SUN.GA.98.L14	.....GCA...-A-CTTCT...TT...-AG-GCA-AGAG-GGGA...CAA--TTACA-T-----C-TC-C--G-----AG-GC-AA-----G-GA--G-----	1504
MND-1.GA.-.MNDGB1	.....T...A-G-----T...-G-CAAAG-G-TGCAGCA--A--GTATCAGT--ACTC-A--AGT--G-T-A--C-GA-----AC--G--G-----	897
MON.CM.99.L1	.....CACCA-CAGCGGC--ACC-AGTGG--C-----G-C--AGA-CACAGGG--AGGCT-CCA-----AG-GGAG--C-GCTCC-C-GA-----C-G--CA-T--G-----	934
MON.NG.-.NG1	.....GCC--C-----AG-CA-AAG--GCA--A-----GT-CCA-----C-CAATG-----TA-----AA-----GTC-A-----G--G-----	0
MUS.CM.01.1085	.....GCC--C-----AG-CA-AAG--GCA--A-----GT-CCA-----C-CAATG-----TA-----AA-----GTC-A-----G--G-----	918
DEB.CM.99.CM40	GAGCCTCAGGCAGCACCCAGACAACACCATCAGTT--C-----CT-AAGAC-GGGG...CAA--GT-T--G--AT-C-AT--AGC-C--G-TAG--AA-C--A--T-CT-----G-----	1040
DEB.CM.99.CM5	ACACAGCAA--A-CA--CCAA...CCAGCT--TT...-A--CC--AGAC-AGGA...CA--TAT--AT-C-TT--AGT--G-T-G-G--AA-C--A--T-C--G--G-----	1031
SYK.KE.-.KE51	GCAATGCCATCT--TGGC...G--TCAGG-----C--A-TAGA--CCA-AAAT--TG-A-T--GTA-GGG-CAACA-C--G-----AA-C-----G-G-C--GA-CA-----	978
SYK.KE.-.SYK173	AAACAAAGGCAGCA-CA...-C-CCACTAGAG--C--T...-AC-GC-CAGA--TCCT-AAAT--GTG-A-T--ACA-GAG-CC--GT--G--C-----AA-----GG-G-CT--GA-CT-----	1350





H1B.FR.83.HXB2	AGAAGGCTTTTCAGCCAGAAAGTGTATCCCATGTTTTTTCAGCATTATCAGAAGGAGCCACCCCAAGATTTTAAACACCATGCTAAACACAGTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAGAGAC	1400
Gag	E K A F S P E V I P M F S A L S E G A T P Q D L N T M L N T V G G H Q A A M O M L K E T	p24
H1A1.UG.85.U455	-C-----A-----G-----T-TG-----G-TGT-----C-G-----T-----T--	855
H1B.US.90.WEAU160	-A-----A-----	1399
H1C.ET.86.ETH2220	-A-----G-A-----A--T-----T-----C--	774
H1D.CD.84.84ZR085	-A-----A-----	917
H1F1.BE.93.VI850	-A-----T-----A-----G-----T-CAC-----T-----C--	743
H1G.SE.93.SB6165	-A-----C-----T-----A-----TCT-----T-----G-----T-----C--G--T--	803
H1H.CF.90.056	-A-----T-----A-----C-----TG-T-----T-----G-----T-----G--T--	747
H1J.SE.93.SE7887	-A-A-----A-----A-----G-----T-----A-A-----C-----T-----T--	723
H1K.CM.96.MP535	-A-----G-----T-A-----A-----C-----T-----G-----C-----T-----T--	599
H101.AE.TH.90.CM240	-A-----G-----T-A-----A-----C-----G-----T-TG-----T-T-----C-G-----A-----A--	965
H102.AG.NG.-.IBNG	-A-----G-----A-----G-----G-----T-TG-----T-----C-G-----A-----G-----T--	931
H1N.CM.95.YBF30	-A-----C-----T-----T-----A-----A-----ATG-C-G-----G-A-G-CTC-----C-----T-----T-----T-----A-----G-----A-----G-----C-G--G--AGT	964
H10.BE.87.ANT70	-A-A-C-T-A-----T-----A-C-T-----CATG-----G-----G-----A-TTT-CT-T-----A-T-----T-----TG-CA-A-A-----G-----TT-A-----G-----C--G--AGT	1446
H10.CM.91.MVP5180	-A-A-C-T-A-----T-----A-T-----T-----ATG-----G-----TGT-----CT-T-----A-C-----T-----G-----TG-CA-A-----C-----GG--TT-A-----G-----G--G--AGT	1421
CPZ.CD.-.ANT	-A-----AAA-----AT-----C-----C-----T-----T-----G-----A-----T-----T-----T-----T-----TG-----T-----AC-----G-----GG--C-----G--AGT	812
CPZ.CM.-.CAM3	-A-----AA-----T-A-----G-----A-----CATG-CC-G-----G-----T-----G-----T-----T-----T-----TG-TA-A-A-----G-----C-----G-----G-----G--G--AGT	790
CPZ.GA.88.GAB2	-A-A-C-T-----C-----T-----A-----CATG-----T-----G-----A-----A-----C-----G-----A-----G-----T-----GG-AC-----GT	778
CPZ.GA.-.CPZGAB	-A-----A-----T-----T-----G-----A-----A-----TC-----G-----TTA-----T-----G-----G-----T-----T-----G-----A-----G-----C-----G--AGT	1461
CPZ.US.85.CPZUS	-A-----C-----T-----A-----G-----C-----C-----G-----CATG-----C-----G-----G-----T-----C-----G-----C-----T-----TG-TA-A-----G-----A-----G-----T-----G--AG-	1448
CPZ.TZ.-.TAN1	-A-A-GG-----T-AT-----A-----A-----C-----G-----T-----G-----A-----TT-----C-----T-----GT-----C-----TG-T-----T-----A-----G-----G--AGT	1032
H2A.GW.-.ALI	-A-----AAG-----G-GG-----AG-G-----AGGA-----CAG-----C-----CTG-----G-----CT-----T-----A-----T-----TCAA-----T-----TTGT-----C-----AC-----AA-C-GG-----T	1712
H2A.DE.-.BEN	-A-----AAG-----G-GG-----AG-G-----AGGA-----CAG-----C-----CTG-----G-----CT-----T-----A-----T-----TCAA-----T-----TTGT-----C-----AT-----C-----T-----AA-C-G-----T	1719
H2A.SN.-.ST	-A-----AAG-----G-GG-----AG-G-----AGGA-----CAG-----C-----CTG-----G-----CT-----T-----A-----T-----TCAA-----T-----TTGT-----C-----AC-----T-----AA-C-GG-----A-T	1164
H2B.GH.86.D205	-A-----AAG-----G-GG-----AG-----AGGA-----CAG-----C-----TG-----TT-----T-----A-----TCAG-----TTGT-----A-----A-----G-----TA-T-G-----A-T	1680
H2B.CI.-.EHO	-A-----AAG-----G-GG-----AG-----AGGA-----CAG-----C-----GTG-----TT-----T-----A-----T-----TCAG-----TTGT-----A-----A-----G-----TA-T-GG-----T	1681
H2G.CI.-.ABT96	-A-----AAG-----G-GG-----G-G-----AGGA-----CAG-----G-----CTG-CTA-----T-----T-----A-----C-----TCAA-----TTGT-----A-----A-----G-----TA-T-G-----A-T	1082
H2U.FR.96.12034	-A-----AAAA-----TG-AG-----AG-G-----AGGA-----CAG-----C-----G-----CTG-----G-----CT-----T-----A-----T-----TCAA-----T-----TTGC-----A-----A-----T-----TA-C-GG-----A-T	1198
MAC.US.-.239	-A-----AAA-----TG-AG-----AG-G-----AGGA-----CAG-----C-----G-----TTG-----CT-----T-----CA-----T-----TCAG-----T-----TTGT-----A-----AC-----G-----T-----G-----TA-C-G-----T-T	1669
Gag	E K K F G A E V V P G F Q A L S E G G T P Y D I N Q M L N G G V G D H Q A A M Q I I R D I	p24
SMM.SL.92.SL92B	-A-----AAA-----TG-AG-----AG-G-----AGGA-----CAG-----C-----CTGT-----A-----CT-----T-----CA-----T-----TCAA-----T-----TTGT-----A-----A-----G-----TA-C-G-----A-T	1103
SMM.US.-.H9	-A-----AAAA-----TG-GG-----G-----AG-G-----AGGA-----CA-----GC-----CTG-----T-----CT-----Y-----A-----C-----TCAA-----T-----TTGT-----A-----A-----G-----TA-T-G-----A-T	1151
STM.US.-.STM	-A-----AAA-----TG-TG-----G-----G-----AGGA-----CAG-----C-----CTGT-----CT-----T-----CA-----C-----TCAA-----T-----TTGT-----AG-----TA-----T	1325
SAB.SN.-.SAB1C	-A-----AAG-----TG-----G-----AG-----AGC-----CC-----G-----T-----TT-----T-----A-----C-----TCAA-----C-----TG-----A-----A-----C-----G-----GG-----AC-----A-----AG-----G-----TGT	1666
TAN.UG.-.TAN1	-A-----AAGG-----G-GG-----A-----AG-----CCA-----CC-----C-----GTG-----TTAAGTT-----T-----CA-----T-----TCAG-----T-----TGTCA-----A-----A-----AC-----GG-----A-----G-----TA-----C-----G-----AGT	1536
VER.KE.-.AGM155	-A-----AAAA-----TG-GG-----A-----AG-----CCA-----CC-----C-----CTG-----T-----T-----T-----CA-----C-----TCAG-----T-----TGTCT-----A-----A-----AT-----G-----GG-----T-----AG-----G-----A-----T	1565
VER.KE.-.9063	-A-----AAA-----TG-AG-----GA-----AG-----G-----A-----CCA-----C-----G-----GTG-----CT-----T-----CA-----C-----TCAA-----T-----TGT-----A-----A-----AC-----GG-----AC-----A-----AG-----A-----T	1576
VER.DE.-.AGM3	-A-----AAAA-----G-AG-----A-----AG-----G-----CCAG-----T-----GTG-----A-----CT-----T-----CA-----C-----TCAA-----T-----TGTCC-----A-----AC-----GG-----GC-----A-----AG-----A-----T	1077
VER.KE.-.TY01	-A-----AAAA-----TG-AG-----A-----AG-----CA-----CC-----CTG-----A-----CT-----T-----CA-----T-----TCAG-----T-----TGTG-----C-----A-----AT-----GG-----AT-----A-----AG-----G-----T	1066
COL.CM.-.CGU1	..GGA-GAA-TGCT--TTCCT--GCT--C-T--T-AT-TACG-G-A-TAG-TT-T--A-G--TTG--C--T-TTT--ATACT--G-TT-CCT--GG-C--G--G--TGA	1208
DEN.CD.-.CD1	-C-GAAAA--TTCG--G-AG-G--AC-A--G-TCC-CG-C--GCTG-TA--TT-T--A--TCAA-C--TG--A-T-A-A--C--G--G--TGT	1167
GRV.ET.-.GRI_677	-A--AAGG-GGG-AG--AG-C--G--CCA--C-C--G--TGTCT-T-CT-T--G--TCAG-C--C--TG-T-A-A--A-AC--G-GG--AT-A--TC-T--G--AGT	1513
GSN.CM.99.CN166	---AAAG--TGCT---GACAG-TG-T---CAG-T---CTGT-TA-CT-T--A-G--CAA---C--TG-TA-A--A-ATT---GG--AG-A--G--TA-T---TGT	1077
GSN.CM.99.CN71	-A--AAG--TGCT---GACAG-CG-TC---CAG-TC-A---CTGT-TA--TT-T--A-G--CAA---C--G-T--T--A-ACT---G--AG-A--G--TA-T---TGT	1083
DRL.-.-.FAO	-A--AAG-----TG-----G-----G-----T-----C-----TG-----CC-----G-----CTG-----T-----AGCT-----T-----CA-----C-----CAG-----T-----TG-TA-----T-----AT-----C-----GT-----C-----C-----AG-----C-----T	910
RCM.GA.-.GAB1	-A-----AAAA-----TCAG-----AGCG-----C-----G-----C-----CTG-----TA-----CT-----T-----A-----TCAA-----TG-----CA-----A-----A-----C-----G-----GT-----GC-----G-----AG-----G-----AGT	828
RCM.NG.-.NG411	-A-----AAA-----TG-AG-----AG-----G-----T-----G-----C-----G-----CTG-----TA-----CT-----T-----G-----TCAG-----T-----TG-----CA-----A-----A-----G-----GT-----T-----A-----AG-----G-----AGT	828
MND-2.-.-.5440	---AAAG-----T-----A-----AG-----G-----T-----ATT-----C-----G-----TG-----T-----GT-----T-----CC-----C-----TGG-----C-----TG-----CA-----T-----A-----AT-----G-----G-----TC-----C-----AG-----G-----TGTT	913
MND-2.CM.98.CM16	---AAG-----T-----A-----AG-----G-----T-----CAT-----C-----G-----T-----GTG-----CT-----CT-----T-----C-----C-----TGG-----C-----TG-----TA-----T-----A-----AC-----G-----G-----GC-----T-----AG-----G-----TGTT	1356
MND-2.GA.-.M14	-A-----AAA-----T-----A-----AG-----T-----CAT-----CC-----G-----TG-----CT-----T-----T-----CC-----C-----TGGT-----T-----TG-----TA-----T-----A-----A-----G-----G-----TC-----C-----AG-----G-----TGTT	1280
MNE.US.-.MNE027	---AAA-----TG-AG-----AG-----G-----AGGA-----CAG-----C-----G-----CTG-----T-----CT-----T-----CA-----T-----TCAG-----T-----TTGT-----A-----A-----G-----T-----G-----TA-----C-----G-----A-----T	1149
LST.CD.88.447	-T-GA-GA-GG-AG-----ACAG-GG-A-----T-CC-CA--AA--TTTG--G-T--C--TGT--T--TG-CA-A--A-AC-----GG--A-----G--AA-T---TCA	572
LST.CD.88.485	-T-GA-GA-GG-AA-----ACAG-GG-A-----T-CC-CA--AA--TTTG--G-T--C--TGT--T--TG-CA-A--A-AC-----TG--A-----G--AA-T---TCA	572
LST.CD.88.524	-T-G-GA-GG-AG-----ACAG-GG-T-----TC-GA--AA--TTG--G--C--TGT--T--TG-TA-A--A-AC-----GG--A-----G--AA---TCA	566
LST.KE.-.lho7	-C-GA-G--GG-AA-----ACAG-G-G-A-----C--A--AA--TTTG--TG-T--TGT--T--TG--A-T--A-AT--G-GC-----AA---TCA	1656
SUN.GA.98.L14	---GA-GG-AA-----ACAG-GG-A-----G-----T-----ATGAC-----A-----TT-----G-----CA-----C-----TGT-----G-----TG-----CA-----A-----AG-----G-----T-----G-----A-----TCA	1634
MND-1.GA.-.MNDGB1	-A--AAG-GG-AA-G-G-C-C-TC-A-C--G--GA-----A-T-AGT-T--G--T-T--G--TG--A-A-AT--G-G--A-----G-----C-----TGT	1027
MON.CM.99.L1	---AAA-----TG-----G-----GG-----C-----C-----CCAG-----TC-----CATG-----TT-----CT-----T-----CA-----C-----TCAG-----A-----G-----CA-----A-----T-----AT-----G-----GG-----AC-----C-----TA-----T-----G-----TGT	1064
MON.NG.-.NG1	.....	0
MUS.CM.01.1085	-A-----AAA-----TGCTG-----TG-GG-A-----CAG-----GG-----G-----G-----A-----TA-----T-----T-----CA-----C-----TCAAC-----G-----TG-----CA-----A-----A-----A-----G-----G-----T-----A-----TA-----TGT	1048
DEB.CM.99.CM40	-A-----AAA-----TTCT-----GA-----AG-----T-----TC-----T-----CA-----TC-----GG-----G-----CTGT-----CT-----T-----CA-----T-----TGGTC-----A-----T-----TG-----A-----T-----A-----AT-----T-----G-----GG-----T-----G-----CA-----T-----TGT	1170
DEB.CM.99.CM5	-A-----AAA-----TTC-----GA-----AG-----G-----C-----A-----CA-----T-----GG-----GTGT-----T-----CT-----T-----A-----T-----TGGC-----TG-----CA-----A-----A-----AT-----TA-----GG-----A-----TA-----TGT	1161
SYK.KE.-.KE51	GC-AAAA--TGATG-CTC-C--G-C--AC-A--CCA-ATT--A-----TTTT-TA--TT-T--CC-C--TGA-----T--TG-CA-T--A-AT-----G--A-----G--A-C-----TGT	1108
SYK.KE.-.SYK173	CC-AAAA--TGATG--TCCA-AG-G--AT-A--CAG--TC--A-----TTT-TT--TT-T--C-C--TGG-----T--TG-T--A--A-AT-----G-G-----GG-AA-T---C-T	1480









PLV  
complete genomes



Accession	Sequence	Position
H1B.FR.83.HXB2	TAAGACAAGGACCAAAGGAACCCCTTTAGAGACTATGTAGACCGGTCTTATAAACTCTAAGACCCGAGCAAGCTTCACAGAGGTAATAAAATTTGGATGACAGAAACCTTGTGGTCCAAAATGCCGAACCC	1772
Gag	I R Q G P K E P F R D Y V D R F Y K T L R A E Q A S Q E V K N W M T E T L L V Q N A N P	p24
H1A1.UG.85.U455	-----G-----A-----C-G-T-----TA-A--T-----C-----T-----A-----T-----C-----C-----T--	1227
H1B.US.90.WEAU160	-----A-----A-----T-----T-----C-----T-----A-----A-----T-----C-----T-----G-----T--	1771
H1C.ET.86.ETH2220	-----A-----A-----T-----T-----CT-----T-A-----A-----A-----T-----C-----C-----T-----G-----T--	1146
H1D.CD.84.84ZR085	-C-----A-----C-----C-----A-----T-----A-----A-----T-----C-----C-----A-----A-----T--	1289
H1F1.BE.93.VI850	--A-----G-----A-----A-----T-----GTC-----T-----C-----GGG-----C-----A-----T-----T--	1115
H1G.SE.93.SB6165	-----G-----A-----TA-----T-----TG-T-G-----T-----T-----GGC-----C-----C-----A-----T--	1175
H1H.CF.90.056	--A-----G-----A-----A-----T-T-----T-----T-----CA-----T-----G-----G-----A-----T-----T--	1119
H1J.SE.93.SE7887	-----G-----A-----T-----A-----T-----G-----T-----A-----A-----T-----T-----C-----C-----A-----T--	1095
H1K.CM.96.MP535	-----G-----A-----TA-----T-----C-----T-----A-----CA-----A-----G-----C-----C-----C-----A-----T--	971
H1O1.AE.TH.90.CM240	-----G-----A-----C-----TA-----T-----C-----G-----A-----A-----C-----C-----C-----A-----C-----T--	1337
H1O2.AG.NG.-.IBNG	-----G-----A-----T-----TA-----T-----T-G-----T-----A-----A-----G-----C-----C-----C-----A-----T--	1303
H1N.CM.95.YBF30	-C-A-----A-----C-----A-----C-----A-----G-----AA-----A-----G-----AC-C-----A-----A-----T--	1336
H1O.BE.87.ANT70	-T-AG-G-----A-----A-----C-----AT-----T-----A-----T-----A-----A-----TC-C-----T-----T-----C-----T--	1821
H1O.CM.91.MVP5180	-T-G-G-----A-----A-----C-----T-----T-----T-----C-----AT-----T-----A-----T-----A-----A-----G-----C-----T-----T-----G-----T-----A--	1796
CPZ.CD.-.ANT	--A-----A-----G-----C-----T-----G-----TA-A-----AA-----T-----A-----A-----G-----CCT-----G-----GCC-----AA-----C-----T-----T--	1193
CPZ.CM.-.CAM3	-C-AG-G-----T-----A-----G-----T-----C-----G-----T-----TA-A-----C-----C-----G-----A-----A-----AA-----A-----A-----T-----C-----C-----A-----A-----G-----T-----T--	1162
CPZ.GA.88.GAB2	--A-----G-----G-----G-----T-----C-----G-----TA-A-----C-----GT-----T-----A-----A-----G-----A-----T-----G-----C-----G-----C-----CC-----A-----G-----C-----T--	1150
CPZ.GA.-.CPZGAB	-C-----G-----C-----A-----G-----T-----G-----A-----C-----T-----G-----T-----A-----G-----A-----G-----T-----C-----TC-----C-----A-----G-----G-----A-----T--	1833
CPZ.US.85.CPZUS	-C-AG-G-----T-----A-----G-----A-----C-----AT-----A-----A-----T-----CA-----A-----CC-----G-----C-----T-----A-----C-----T-----T-----T--	1820
CPZ.TZ.-.TAN1	--A-----G-----A-----A-----C-----T-----A-----A-----C-----AA-----C-----A-----A-----A-----T-----ACCA-----C-----T-----A-----A-----A-----T--	1413
H2A.GW.-.ALI	--A-----G-----A-----CCA-AG-----TA-A-----C-----GCT-----G-----A-----A-----A-----AGAT-----CA-----CA-----G-----TC-----AC-----C-----A-----G-----C-----T--	2084
H2A.DE.-.BEN	--A-----G-----A-----G-----CCA-AG-----A-----C-----GCT-----G-----G-----A-----A-----A-----AGAC-----CA-----CA-----CC-----GC-----C-----AA-----A-----G-----C-----T--	2091
H2A.SN.-.ST	--A-----G-----G-----A-----CCAGAGT-----TA-A-----C-----G-----GCT-----G-----G-----A-----A-----A-----AGAT-----CA-----CA-----CC-----AC-----C-----A-----G-----G-----C-----T--	1536
H2B.GH.86.D205	--AG-----G-----G-----CCA-AG-----TA-A-----C-----GCT-----C-----G-----A-----A-----A-----AGAC-----CA-----CA-----G-----C-----AC-----C-----A-----T-----G-----T-----T--	2052
H2B.CI.-.EHO	--AG-----G-----A-----CCA-AG-----TA-A-----C-----GCT-----C-----G-----A-----A-----A-----AGAC-----CA-----CA-----C-----AC-----C-----A-----T-----G-----T-----T--	2053
H2G.CI.-.ABT96	--A-----G-----A-----CAGTCT-----TA-A-----C-----G-----G-----R-----A-----A-----A-----GAT-----CA-----CA-----C-----G-----AC-----AC-----A-----T-----C-----T--	1454
H2U.FR.96.12034	--AG-----G-----G-----A-----CA-AGT-----TA-A-----C-----G-----T-----G-----A-----A-----A-----AGAT-----CA-----CA-----G-----TC-----G-----GC-----C-----A-----T-----G-----T-----T--	1570
MAC.US.-.239	--A-----G-----A-----G-----A-----CAGAG-----A-----C-----G-----T-----A-----A-----GA-----AGAT-----GCA-----CA-----G-----TC-----AC-----C-----A-----T-----T-----T--	2038
Gag	V K Q G P K E P F Q S Y V D R F Y K S L R A E Q T D G G V K N W M T Q T L L I Q N A N P	p24
SMM.SL.92.SL92B	-T-A-----T-----G-----CA-AGT-----TA-A-----C-----G-----T-----G-----A-----GA-----AGAT-----C-----CA-----G-----G-----C-----AC-----C-----A-----T-----G-----C-----C-----T--	1481
SMM.US.-.H9	-G-A-----G-----A-----GT-----CA-AG-----TA-A-----C-----G-----G-----A-----A-----A-----AGAC-----C-----CAS-----G-----TC-----AC-----C-----A-----T-----G-----C-----C-----T--	1523
STM.US.-.STM	--AG-----G-----A-----A-----G-----CCA-AGT-----A-----C-----G-----G-----T-----G-----A-----A-----AGAT-----CAT-----CA-----G-----TCGG-----AC-----C-----AA-----T-----T-----T-----T--	1700
SAB.SN.-.SAB1C	-T-A-----G-----T-----A-----T-----C-----AG-----T-----TA-A-----C-----GG-----G-----A-----GA-----AGAC-----CT-----CT-----G-----C-----C-----GT-----TC-----A-----G-----C-----C-----T--	2041
TAN.UG.-.TAN1	-C-----GG-----C-----G-----T-----AG-----C-----A-----A-----T-----C-----G-----C-----T-----A-----A-----A-----A-----AGAT-----CA-----CA-----G-----AC-----C-----A-----T-----G-----C-----C-----T--	1911
VER.KE.-.AGM155	-C-----G-----C-----A-----A-----C-----A-----T-----A-----A-----G-----AA-----A-----A-----A-----GGA-----A-----C-----C-----A-----T-----T-----C-----CA-----T-----G-----C-----C-----T--	1940
VER.KE.-.9063	-C-----G-----T-----A-----T-----C-----AG-----T-----TA-----C-----G-----AA-----C-----A-----A-----G-----C-----GGA-----A-----C-----A-----T-----A-----AC-----CA-----A-----T-----T-----T--	1951
VER.DE.-.AGM3	-----G-----G-----C-----A-----G-----A-----C-----A-----C-----TA-----C-----G-----AA-----T-----G-----C-----GGA-----A-----C-----G-----T-----A-----AC-----CA-----T-----G-----T-----T-----T--	1452
VER.KE.-.TY01	-T-----G-----G-----T-----A-----G-----C-----AG-----T-----G-----A-----A-----T-----C-----GG-----AA-----T-----A-----A-----C-----GG-----A-----G-----C-----A-----T-----A-----AC-----CA-----T-----T-----T-----T--	1441
COL.CM.-.CGU1	-T-----G-----AGT-----A-----GGAT-----A-----AG-----A-----A-----T-----TC-----G-----C-----G-----A-----A-----CT-----AG-----GGA-----AA-----C-----GGCC-----G-----TA-----T-----AT-----C-----AA-----A-----GC-----C-----T--	1562
DEN.CD.-.CD1	--AG-----G-----C-----A-----T-----AG-----T-----TA-----A-----C-----G-----G-----A-----A-----G-----AGAT-----AT-----CA-----G-----GCT-----CACTT-----AC-----A-----A-----A-----T-----T--	1545
GRV.ET.-.GRI_677	-TC-----G-----T-----A-----CCAG-----A-----A-----G-----C-----G-----A-----A-----A-----AC-----T-----T-----C-----T-----T-----C-----TA-----G-----C-----T-----T-----T--	1888
GSN.CM.99.CN166	-C-----A-----C-----T-----A-----G-----AG-----A-----A-----A-----T-----GT-----GCT-----G-----G-----A-----AGAT-----CC-----G-----C-----T-----GC-----C-----T-----T-----T-----T--	1452
GSN.CM.99.CN71	-C-----A-----G-----C-----T-----A-----AG-----A-----A-----A-----T-----TGCT-----G-----T-----A-----G-----AGAT-----CA-----G-----C-----TTGC-----C-----T-----T-----T-----T--	1458
DRL.-.-.FAO	-G-AG-----T-----T-----A-----A-----C-----AG-----TA-----A-----T-----G-----C-----C-----G-----A-----A-----CA-----AGAT-----GCA-----CA-----C-----AC-----C-----A-----A-----A-----T-----T--	1282
RCM.GA.-.GAB1	--AG-----G-----A-----GT-----A-----C-----AG-----T-----G-----TA-----A-----T-----G-----CT-----G-----A-----A-----G-----AGAC-----C-----CA-----C-----T-----AC-----C-----A-----A-----T-----T--	1200
RCM.NG.-.NG411	-C-AG-----G-----A-----G-----T-----AG-----TA-----A-----C-----GG-----CT-----G-----G-----A-----A-----G-----AGAC-----CA-----CA-----G-----G-----C-----T-----AC-----C-----A-----T-----T-----T--	1200
MND-2.-.-.5440	-C-----A-----G-----A-----G-----T-----AG-----G-----TA-----A-----C-----G-----GC-----G-----G-----GCA-----AGAT-----CAC-----C-----C-----GC-----G-----A-----AC-----CA-----G-----C-----C-----T--	1285
MND-2.CM.98.CM16	--AG-----G-----G-----C-----AG-----G-----A-----T-----C-----G-----GC-----G-----G-----GGA-----GAT-----CAC-----A-----C-----CC-----C-----T-----C-----CA-----G-----T-----T--	1728
MND-2.GA.-.M14	--A-----G-----A-----G-----T-----C-----AG-----G-----A-----A-----T-----G-----GC-----G-----G-----G-----A-----AGAT-----CAC-----A-----C-----CC-----C-----A-----C-----CA-----C-----C-----T--	1652
MNE.US.-.MNE027	-----G-----A-----G-----A-----CAGAG-----A-----A-----C-----GCT-----A-----A-----A-----AGAT-----CA-----CA-----G-----TC-----AC-----C-----A-----T-----T-----T--	1518
LST.CD.88.447	-C-AG-----G-----G-----A-----A-----C-----A-----A-----C-----A-----T-----G-----AT-----A-----AGG-----GC-----T-----A-----G-----A-----AG-----G-----AGA-----C-----A-----A-----A-----T--	944
LST.CD.88.485	-C-----A-----G-----G-----A-----A-----C-----A-----A-----C-----TA-----T-----G-----AT-----A-----AGG-----GC-----T-----A-----G-----A-----A-----G-----AGA-----C-----A-----A-----A-----T--	944
LST.CD.88.524	--A-----G-----T-----A-----A-----C-----AG-----CT-----TA-----T-----GG-----CT-----A-----AGG-----G-----T-----A-----G-----G-----G-----A-----G-----AAA-----C-----A-----G-----G-----A-----T--	938
LST.KE.-.lho7	-C-----A-----C-----A-----G-----T-----A-----A-----C-----TA-----T-----G-----A-----G-----A-----AGG-----GC-----T-----T-----A-----GG-----G-----AG-----G-----AAA-----C-----T-----G-----G-----A-----T--	2028
SUN.GA.98.L14	-T-----G-----A-----G-----T-----C-----AC-----TA-----T-----TC-----GG-----T-----A-----AT-----C-----GC-----G-----A-----G-----G-----A-----G-----AAA-----C-----A-----G-----A-----T--	2015
MND-1.GA.-.MNDGB1	-----G-----T-----A-----TA-----A-----C-----TGTA-----A-----G-----A-----A-----G-----G-----A-----G-----TG-----CAGC-----GCAT-----C-----TA-----AG-----A-----T--	1402
MON.CM.99.L1	-C-----A-----G-----T-----A-----A-----C-----A-----T-----A-----A-----C-----TG-----C-----A-----A-----A-----AGAT-----CA-----C-----GC-----G-----GC-----C-----C-----T-----T-----T--	1439
MON.NG.-.NG1	.....	5
MUS.CM.01.1085	-C-AG-----G-----A-----A-----AG-----T-----G-----A-----A-----T-----TG-----G-----A-----GT-----AGAC-----A-----C-----C-----CTC-----C-----A-----A-----G-----T-----T--	1423
DEB.CM.99.CM40	-T-----A-----G-----C-----A-----C-----A-----T-----T-----A-----G-----CT-----T-----A-----AGAT-----CA-----CACC-----C-----CA-----A-----T-----A-----T--	1551
DEB.CM.99.CM5	-T-AG-----G-----T-----T-----C-----A-----TA-----A-----C-----CT-----G-----G-----T-----A-----AGAT-----CA-----C-----CACC-----C-----AA-----A-----T-----G-----T-----T--	1542
SYK.KE.-.KE51	-C-----G-----T-----A-----A-----A-----AG-----T-----T-----TA-----A-----C-----TTG-----T-----A-----A-----GGAC-----C-----CT-----C-----C-----GT-----A-----AC-----AA-----T-----G-----A-----T--	1483
SYK.KE.-.SYK173	--A-----G-----C-----A-----ATT-----C-----A-----A-----A-----CC-----TTGCT-----A-----G-----GGAC-----CAT-----C-----GGG-----C-----A-----TC-----AC-----A-----T-----C-----T--	1855

see Kaye, J Virol 72(7):5877-5885(1998) Gag p2 role in RNA encapsidation

Gag p24 \ / Gag p2

	AGATTGTAAAGACTATTTTAAAGCATTGGGACCAGCGGCTCACTAGAGAAGATGATGACAGCATGTCCAGGGAGTAGGAGGACCCGGCCATAAGGCAAGAGTTTGGCTGAAGCAATGAGCCAAGTAACA	1902
Gag	D C K T I L K A L G P A A T L E E M M T A C Q G V G G P G H K A R V L A E A M S Q V T	p24
H1B.FR.83.HXB2	---C-----T-C-----G-----A-G-----G-----T-----C-----G-----A-----G-----G-----T-----CA-	1357
H1B.US.90.WRAU160	---C-----A-----G-----A-----G-----G-----C-----G-----G-----A-----A-----G-----T-----CA-	1901
H1C.ET.86.ETH2220	-----A-C-----G-----A-G-----G-----T-T-----G-----G-----T-C---C-A-----G-----G-----AC	1276
H1D.CD.84.84ZR085	-----C-----C-----G-----A-----G-----T-----G-----G-G-G-----A-----A-----G-----C-----C-----	1419
H1F1.BE.93.VI850	-----C-----A-----G-----T-----G-----G-----TA-----A-----C-----G-----T-----C-----AT	1245
H1G.SE.93.SE6165	-----C-C-----G-----A-----A-----GA-----T-----G-----A-----A-----A-----G-----G-----CCT--	1305
H1H.CF.90.056	-----C-----A-----G-----A-----A-----G-----T-----A-----G-----TA-----T-----A-----G-----G-----	1249
H1J.SE.93.SE7887	-----C-----C-----A-----T-----GA-----G-----G-----T-----T-----A-----G-----G-----G-----C	1225
H1K.CM.96.MP535	-----C-----G-----G-----T-----T-----G-----G-----G-----TA-----A-----A-----G-----G-----G-----	1101
H101.AE.TH.90.CM240	-----C-----T-C-----G-----A-----A-----GA-----T-----C-----G-----TA-----A-----G-----C-----C-----CA-	1467
H102.AG.NG.-.IBNG	-----C-----T-C-----G-----A-----A-----G-----T-----T-----G-----T-----A-----G-----G-----T-----CA-	1433
H1N.CM.95.YBF30	-----ACAGC-CC-----G-----A-----G-----GA-----CT-----G-----G-----C-----C-----G-----G-----A-----CA-----GC-----A-----A-----G-----T-----TCA-----G-----GCAG	1466
H10.BE.87.ANT70	-----C-----ACAGC-----G-----T-----A-----G-----GA-----CT-----G-----GT-----C-----C-----AACT-----C-----GC-----A-----A-----GCTAC-----CCGAG	1951
H10.CM.91.MVP5180	-----ACAA-----C-----G-----A-----AA-----TT-----GT-----C-----A-----G-----AACT-----C-----A-----A-----AC-----A-----GCTTCT-----CCGAG	1926
CPZ.CD.-.ANT	-----C-----ACAC-----CC-----G-----G-----T-----A-----GA-----CT-----CT-----T-----A-----T-----A-----C-----A-----C-----G-----A-----T-----GCTTCT-----CT-----AT	1323
CPZ.CM.-.CAM3	-----C-----A-----AC-----C-----G-----CC-----GA-----CT-----G-----C-----C-----G-----A-----C-----T-----AC-----T-----AC-----C-----CA-	1292
CPZ.GA.88.GAB2	-----CAA-----AC-----T-----GG-----C-----A-----GA-----C-----C-----G-----C-----C-----A-----C-----G-----CA-----T-----C-----A-----A-----T-----A-----CA-	1280
CPZ.GA.-.CPZGAB	-----CAA-----C-----T-----G-----C-----C-----G-----G-----T-----A-----A-----ATCT-----G-----G-----C-----A-----TCAATG-----TCAG	1963
CPZ.US.85.CPZUS	-----A-----C-----G-----C-----T-----GA-----C-----C-----G-----G-----A-----G-----CT-----G-----C-----A-----C-----T-----A-----G-----A-	1950
CPZ.TZ.-.TAN1	-----CA-----C-----C-----G-----A-----GA-----A-----T-----C-----C-----C-----A-----G-----TCT-----GC-----AAGA-----TCTG-----CT-----A-----GCA-----TGGCCTC-----G-----	1540
H2A.GW.-.ALI	-----C-----C-----ATTAG-----GC-----G-----G-----A-----GATGAATC-----CT-----G-----C-----A-----C-----C-----G-----A-----G-----A-----C-----T-----AA-----A-----CT-----A-----AGG-----G-----CC-----TG	2214
H2A.DE.-.BEN	-----C-----C-----TTAG-----AC-----G-----C-----GATGAATC-----C-----C-----G-----C-----A-----C-----C-----G-----C-----A-----G-----A-----C-----GC-----AA-----CC-----A-----AAG-----G-----CT-----TG	2221
H2A.SN.-.ST	-----C-----TTAG-----AC-----G-----C-----A-----GAT-----AAATC-----CT-----C-----A-----C-----C-----G-----T-----A-----G-----A-----C-----T-----AA-----A-----CT-----A-----AGG-----G-----CC-----TG	1666
H2B.GH.86.D205	-----C-----TTAG-----GC-----T-----G-----GC-----ATGAATC-----C-----CT-----G-----C-----A-----G-----C-----A-----GA-----C-----A-----G-----G-----GC-----AA-----C-----CT-----A-----AAG-----G-----CCCT--	2182
H2B.CI.-.EHO	-----C-----ATTAG-----GC-----T-----G-----GC-----ATGAATC-----C-----T-----G-----C-----A-----C-----C-----GA-----C-----A-----G-----G-----GC-----AA-----G-----TT-----A-----AAG-----CCCTTG	2183
H2G.CI.-.ABT96	-----ATTAG-----AC-----G-----GG-----ATGAATC-----TT-----G-----T-----C-----A-----A-----GR-----ATGA-----A-----G-----A-----A-----C-----GC-----A-----T-----A-----AAG-----CCCTT--	1584
H2U.FR.96.12034	-----C-----C-----ACTCG-----G-----GT-----TATGAATC-----CT-----G-----C-----A-----C-----A-----A-----A-----T-----GC-----GA-----A-----CT-----A-----AAG-----G-----CCCTT--	1700
MAC.US.-.239	-----C-----CTAG-----GC-----G-----GGC-----TGTGAATC-----C-----C-----G-----T-----A-----G-----G-----G-----A-----G-----T-----T-----AA-----A-----CC-----AAG-----G-----CCCTC	2168
Gag	D C K L V L K G L G V N P T L E E M L T A C Q G V G G P G Q K A R L M A E A L K E A L	p24
SMM.SL.92.SL92B	T-----C-----ACTAG-----A-----GG-----CATGAATC-----C-----C-----G-----C-----A-----C-----A-----G-----C-----A-----A-----G-----CC-----CC-----CA-----A-----G-----C-----AAG-----C-----CCCTC	1611
SMM.US.-.H9	G-----C-----ATTGG-----GC-----C-----G-----GTC-----YAT-----AAATC-----C-----TT-----C-----A-----T-----C-----R-----G-----A-----A-----A-----AS-----T-----C-----AA-----A-----G-----T-----AAG-----T-----C-----CTG	1653
STM.US.-.STM	-----C-----C-----ATTGG-----A-----G-----GTC-----CATGAATC-----C-----C-----G-----T-----A-----A-----G-----T-----G-----A-----A-----G-----A-----T-----C-----GA-----A-----CT-----AAG-----G-----CTTC	1830
SAB.SN.-.SAB1C	-----AG-----GC-----C-----GG-----A-----CATGAATC-----A-----C-----G-----C-----T-----T-----C-----A-----CA-----T-----G-----ACAA-----C-----TC-----AA-----G-----T-----CAGCG-----CCTTT	2171
TAN.UG.-.TAN1	-----C-----TTGG-----G-----G-----C-----T-----AT-----CATC-----A-----T-----G-----C-----C-----T-----A-----A-----A-----AGC-----CA-----TA-----CAA-----A-----GCA-	2041
VER.KE.-.AGM155	-----C-----AGTA-----G-----G-----GCC-----A-----GATGCACC-----C-----T-----T-----C-----C-----A-----G-----G-----C-----AA-----TT-----C-----A-----C-----A-----CA-----A-----ATG-----CAGC-----CC-----GCAG	2070
VER.KE.-.9063	-----C-----AGTC-----CC-----G-----G-----A-----ATGCACC-----CT-----C-----T-----C-----A-----G-----G-----C-----AA-----T-----C-----A-----A-----A-----ATG-----CAAA-----CA-----GCAG	2081
VER.DE.-.AGM3	-----C-----AGTC-----CC-----G-----GCC-----ATGCATC-----C-----T-----T-----A-----T-----C-----A-----G-----A-----AA-----TT-----C-----A-----A-----GA-----A-----ATG-----CAAA-----TA-----GCA-	1582
VER.KE.-.TYO1	-----GTC-----CC-----G-----G-----GCC-----A-----ATGCACC-----C-----C-----T-----T-----A-----G-----T-----G-----C-----AA-----T-----C-----A-----A-----AA-----A-----ATG-----CAGACCA-----GCA-	1571
COL.CM.-.CGU1	T-----C-----C-----A-----GA-----C-----G-----G-----C-----CAGAAAG-----C-----AT-----TT-----G-----C-----C-----AG-----C-----G-----A-----A-----C-----AG-----GC-----A-----A-----G-----T-----CAG-----GT-----CCAG	1689
DEN.CD.-.CD1	-----C-----CAG-----C-----CG-----A-----TAACC-----C-----T-----G-----G-----T-----CA-----C-----C-----G-----A-----A-----A-----CC-----TCCCAC-----A-----G-----T-----AGCTTC-----C-----TTT	1675
GRV.ET.-.GRI_677	G-----ATTG-----C-----G-----G-----ATGAATC-----A-----CT-----G-----G-----C-----A-----T-----T-----C-----G-----ACAA-----T-----AGC-----AA-----TA-----ATG-----TA-----T-----G-----CAG	2018
GSN.CM.99.CN166	-----A-----C-----CTG-----C-----G-----CA-----CC-----GG-----C-----ACT-----G-----C-----T-----C-----A-----G-----G-----AACT-----C-----AT-----T-----GC-----C-----A-----A-----C-----GCAAC-----CCCTC	1579
GSN.CM.99.CN71	-----A-----C-----CTG-----C-----G-----G-----CA-----CC-----GG-----C-----GCT-----G-----C-----T-----C-----A-----G-----G-----AACT-----AT-----T-----C-----A-----G-----C-----GCAAC-----C-----CTC	1585
DRL.-.FAO	-----C-----AGTG-----C-----C-----GCC-----A-----GATGAACC-----C-----GT-----G-----C-----TT-----T-----C-----T-----A-----A-----A-----G-----CA-----A-----T-----AG-----CAGCAG	1412
RCM.GA.-.GAB1	-----C-----A-----TGG-----AC-----C-----G-----GTC-----ATGAACC-----T-----TT-----C-----A-----C-----C-----G-----CAG-----A-----TC-----G-----AC-----A-----A-----G-----C-----CAAATGA-----GCA-	1330
RCM.NG.-.NG411	-----C-----AG-----AC-----T-----G-----GTC-----ATGAACC-----C-----G-----G-----G-----C-----A-----C-----A-----G-----C-----T-----ACAG-----C-----T-----G-----GC-----A-----CAGATG-----GCA-	1330
MND-2.-.5440	-----AG-----CCC-----C-----GG-----A-----CATGAATC-----C-----T-----G-----T-----CTG-----C-----C-----A-----G-----TAAAT-----C-----CA-----A-----GA-----A-----G-----C-----CAG-----C-----CAG	1415
MND-2.CM.98.CM16	-----AT-----C-----CC-----T-----GGC-----A-----CATGAATC-----CT-----TT-----G-----CC-----CT-----C-----C-----A-----G-----C-----TAAAT-----A-----CA-----A-----GA-----A-----G-----T-----AGG-----CCGAG	1858
MND-2.GA.-.M14	-----GC-----C-----G-----G-----CATGAACC-----CT-----TT-----T-----ACT-----A-----A-----G-----G-----AAAAT-----CA-----A-----GA-----A-----G-----C-----CAG-----GA-----GCAG	1782
MNE.US.-.MNE027	-----C-----CTGG-----GC-----G-----GGC-----TATGAATC-----C-----C-----G-----T-----C-----A-----A-----A-----A-----T-----AA-----A-----CC-----AAG-----G-----CCCTT	1648
LST.CD.88.447	-----G-----TGG-----GA-----C-----A-----GA-----G-----TT-----G-----G-----GG-----C-----A-----A-----C-----A-----G-----GA-----A-----GCAAC-----CC-----TG	1074
LST.CD.88.485	-----G-----TGG-----AA-----C-----A-----GA-----G-----GTT-----C-----AG-----C-----A-----C-----A-----G-----AGA-----A-----G-----GCAAC-----CC-----TG	1074
LST.CD.88.524	T-----C-----C-----A-----TGG-----AA-----G-----C-----A-----GA-----GT-----C-----TT-----A-----A-----AAG-----A-----A-----C-----A-----A-----G-----GCAAC-----CC-----TG	1068
LST.KE.-.lho7	T-----C-----C-----G-----TTAG-----A-----C-----A-----GA-----G-----AT-----T-----G-----A-----T-----T-----CT-----A-----G-----A-----A-----A-----GCTTC-----CC-----T	2158
SUN.GA.98.L14	-----ATTGG-----AA-----G-----C-----GA-----A-----TC-----AT-----T-----G-----G-----C-----C-----A-----G-----T-----G-----G-----GA-----A-----G-----AGA-----CC-----A-----GC-----TCT-----CTTTT	2145
MND-1.GA.-.MNDGB1	-----A-----C-----CAG-----G-----G-----AGC-----A-----GAA-----GA-----A-----TT-----G-----T-----GA-----T-----G-----ACAA-----A-----C-----T-----AA-----A-----ATG-----AAC-----GGT--	1532
MON.CM.99.L1	-----C-----ACTC-----CC-----G-----CA-----CC-----GGT-----GCT-----G-----G-----C-----C-----ACAG-----G-----T-----AATG-----C-----AT-----T-----GA-----C-----A-----T-----GCAGGG-----C-----T	1566
MON.NG.-.NG1	G-----C-----C-----CTC-----CC-----C-----A-----CC-----GG-----C-----GCT-----C-----CCA-----A-----G-----C-----CA-----AGC-----GT-----A-----A-----A-----G-----T-----GC-----AC-----C-----TT	132
MUS.CM.01.1085	-----G-----ACTC-----C-----G-----GAG-----A-----CC-----GG-----C-----T-----G-----G-----T-----CAG-----C-----G-----ACT-----AT-----C-----GC-----A-----A-----G-----TC-----GC-----TCG-----CCTTC	1550
DEB.CM.99.CM40	T-----C-----G-----TC-----CC-----G-----G-----A-----T-----AGAATC-----C-----CT-----G-----G-----C-----ACAT-----C-----A-----G-----G-----C-----ACAG-----A-----C-----C-----A-----G-----G-----TC-----CCTT--	1681
DEB.CM.99.CM5	-----GA-----TA-----C-----GG-----A-----AGAACC-----C-----TT-----T-----CAT-----C-----T-----CAA-----G-----A-----CC-----TC-----GA-----G-----G-----T-----GC-----TCG-----CCTT--	1672
SYK.KE.-.KE51	T-----G-----C-----GA-----GC-----C-----G-----A-----TCAA-----C-----A-----CT-----G-----C-----TCA-----T-----C-----A-----G-----ACAAT-----C-----A-----A-----C-----A-----A-----A-----GTAATGAC-----CA-	1610
SYK.KE.-.SYK173	-----A-----C-----GACAG-----CC-----C-----G-----A-----TCAA-----C-----CT-----G-----G-----C-----CA-----T-----A-----C-----TTA-----C-----A-----C-----CA-----A-----G-----G-----GTAATG-----CCGAG	1982

PLV Complete Genomes





PLV  
complete genomes

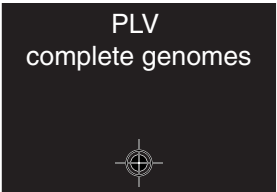


see Kaye, J Virol 72(7):5877-5885(1998) Gag p2 role in RNA encapsidation  
/ Gag-Pol fusion TF protein start

	Gag p2 / Gag p7 NC	
H1B.FR.83.HXB2	AATTCA.....GCTACC.....ATAATGATCAGAGAGGGCAATTTT...AGGAACCAA.....AGAAAG.....A	1954
Gag	N S . . . . . A T . . . . . I M M Q R G N F . . . . . R N Q . . . . . R K . . . . .	p7
H1A1.UG.85.U455	C-GA-----AGC-----GG-CG-----GA-----	1403
H1B.US.90.WEAU160	---A-----G-----G-C-----A-----	1953
H1C.ET.86.ETH2220	---A-----A-C-----A-A-----A-GG-CT-----A-GA-----G	1325
H1D.CD.84.84ZR085	-G-G-C-----T-G-TGCC-----A-A-A-----A-GG-AC-----A-----	1474
H1F1.BE.93.VI850	TCAG-C-----A-A-T-----A-GG-----GA-----G	1291
H1G.SE.93.SE6165	GG-G-----AG-AGCC-----A-----A-GG-C-----GA-----	1360
H1H.CF.90.056	---A-----AA--AGCC-----A-----C-----A-GG-----A-----T	1304
H1J.SE.93.SE7887	---A-C-----AAC-----A-----T-C-----G-----T-----A-GA-----	1274
H1K.CM.96.MP535	---C-----T-----G-----A-----C-----A-GG-----A-----	1150
H1O1.AE.TH.90.CM240	C-G-----A-----C-----A-GG-----G-----A-----	1513
H1O2.AG.NG.-.IBNG	C-GG-C-----AAC-----G-----G-----GG-----G-----GA-----	1479
H1N.CM.95.YBF30	C-GC-----A-A--TAGT-----G-CT-TGCA-A--G--A--C-----AAGG-AT-----G--A-----C	1521
H1O.BE.87.ANT70	C-AGAT-----TTG-AAGGAGGATACACAGCAG--T-C-----A--GC-AAA-----CCA-TTAGG-----A-GGA-----	2018
H1O.CM.91.MVP5180	C-AGAT-----TTA-AAGGAGGATACACAGCAG--T-C-----A--GC-GAA-----CCA--TAG-----A-GG-----C	1993
CPZ.CD.-.ANT	---G--CAGGGAA-C-----GCAG-CT-TC-----GGA-----AT-GAGG-GGA-----A--GA-----C	1384
CPZ.CM.-.CAM3	---CC-T-----A-A-----GGTG-GT-CC-----A--G--A--AA-----G--A--C-----CT-G-----	1344
CPZ.GA.88.GAB2	C-GA-----AGCTCA--T-C-----A--CAGGGGGA-----A-CT-CT-----CCT-G-----	1332
CPZ.GA.-.CPZGAB	---CA-GGGAGA--A-----GATG-TT-CT-C-----AG--AC-AGG-----GC-GG--C-----A--GA-----	2021
CPZ.US.85.CPZUS	---CC-C-----T-A-----AGCG--T-TC-A--A--A--CGC-----G--G-C-----G-G-GA-----	2002
CPZ.TZ.-.TAN1	-CAG--GGGGGA-TA-----AT--C-----G--A--AAGA-----CCACC-TT-----A-----GGTC-----	1598
H2A.GW.-.ALI	-CAC-----C-TATC-----CCATTTGCGGCAGCC--ACA-----G-----	2257
H2A.DE.-.BEN	GGAC-----AGCC-TATC-----CCATTTGC-GCAGCC--ACA-----G-----G-----	2264
H2A.SN.-.ST	GCAC-----CC-ATC-----CCATTTGC-GCAGCC--ACAG-----G-----G-----	1709
H2B.GH.86.D205	-CAC-T-----AC-ATA-----CCGTTTGCTGCCG-T--ACA-AAAGCAGGG-----A-----	2237
H2B.CI.-.EHO	-CAC-T-----T-C--AAT-----CCGTTTGCCGCCGCT--ACC-A-AGCAGGG-----AG-G-----	2235
H2G.CI.-.ABT96	---CC-T-----A-AG--CTA-----CCATTTGC-GCTGCT--ACA-AAA-CAGGA-----G-A-----A-G-----AGC-	1642
H2U.FR.96.12034	C-AC-T-----ATGC-TATA-----CCCTTTGC-GCAG-C--ACAGA-AGGCGGCCCA-----G-G-----C	1755
MAC.US.-.239	GCAC-----TGC-AATC-----CCTTTTGC-GCAGCC--ACAGA-GGGACCA-----G-----C	2220
Gag	G P . . . . . V P I . . . . . P F A A A Q Q R G P . . . . . R K . . . . .	p8
SMM.SL.92.SL92B	-CGGGTCTTTA-TGG-AGCACAGTTTAGGGGAGC-GC--AAGGCCA--G-----AT--CCT-----	1672
SMM.US.-.H9	-CGCA-----GCC-ACTC-----CCATTTGCRGCAG-C--ACAGAAAGGACAA-----R-----	1705
STM.US.-.STM	C-AC-----ACC-ACTG-----CCCTTCGC-GCAGCC--ACAGCAGGGA-----C--G-----	1879
SAB.SN.-.SAB1C	C-GCA-----CAA--AGTT-----GGGAAC--CT-TG----CA--GGCAAGA...CCC-GGGTCCCTTAGGGGAAGAGGTAGACCTCTG...-ACCCA-----	2259
TAN.UG.-.TAN1	GGAGTT-----AC--G--AG--CACCGGA...G-AGGTAGGGGAGGAGGC-----GGC-----C	2096
VER.KE.-.AGM155	-G-CAG-----AC--G-A--CAG--AGG-GGA...--GGAAG-CCA-----CCC-----C	2119
VER.KE.-.9063	-C-CA-----AC--G--CA--GGCAGCA...G-AGGAGT-AGGAGACAG-----CCC-----C	2136
VER.DE.-.AGM3	-GCCAG-----AC--G--ACAG--GG-CAG...--AGGAAG-CCA-----CCC-----C	1631
VER.KE.-.TYO1	---CA-----AC--G--CAG--AGG-CCA...-AA-GA-----CCC-----C	1617
COL.CM.-.CGU1	C-AGAG-----AGAACC-AC----TG-AGT-AAG-CAGCA-----	1725
DEN.CD.-.CD1	C-GAAG...CAA-GGTTATGCCATGGTTCAAGG-A-GG-GCC--ACCCCCAGTAGAA-----GA-----AGGG	1736
GRV.ET.-.GRI_677	...AT--G-C-AGTG--ACC-CAG...AA--GGGCCCC-----C--GG-----C	2061
GSN.CM.99.CN166	--AGGC...A--GTTCA...TAC-AC--G--G--GGG-CTCCAAGAGATCAGGAGGA...AATCC--T-----	1637
GSN.CM.99.CN71	--AGGC...-G-TCA...TAC-AT--G--A...-GAGGCCCCAGGAGG...CAAG---A-----	1640
DRL.-.-.FAO	GCAG-T...ATA-TGATGCAGCAGAAT...CCCCCAGAG-TCCACCGGGGAGCC--GGGG-CCCCCTCA...--A-----	1482
RCM.GA.-.GAB1	-G-AAT-----C--GCT--CA-TCAGAAAC...--GGG-CTCCAAGAGATCAGGAGGA...AATCC--T-----	1396
RCM.NG.-.NG411	-G-AAT-----C--GCC-ACAG--AGGACC-AGA-AAGGA--CCCCTAARACTAGGGGAGGA...CCT-G-----T	1397
MND-2.-.-.5440	GCAG-T...-TA-TGATGCAGAACTCGGGGGG-CC-CC--G-G-TCC-CCGAGACAACCCCT...-GA-----	1479
MND-2.CM.98.CM16	TCAG-T...-TA-TG...ATGCAG-ATTC-GGAGG-CC-CCGGGGGCCCC--GA--CGTCCT...-GA-----	1922
MND-2.GA.-.M14	TCAGA...-TC-TGATGCAGAACTCGGGGGG-CCACC--G-G-TCTCCGAGACAACCCCT...-GA-----	1846
MNE.US.-.MNE027	GCAC--...-GGC-ACTC...CCTTTTGC-GCAGCC--ACAGAAGGACCA-----C	1700
LST.CD.88.447	C-AAGT...CAA-TGAGA...CAG-AT--G-A-A...-TG-CACCA...CCA-GGA-TGCACAGGGAAGGTTTGTAAAGAACAGGAGGAGGAGGC...CCT-----	1168
LST.CD.88.485	C-AAGT...CAA-TGAGA...CAG-AT--G-A-A...-TG-CACCA...CCA-GGA-TGCACAGGGAAGGTTTGTAAAGAACAGGAGGAGGAGGC...CCT-----	1168
LST.CD.88.524	C-GGGG...CAG-TGAAA...CATC-C--G--CA--CACCA...CCT-GAA-TGCACAGGAGATTTGTAAAGAACAGGAGGAGGAGGC...CCA---GA-----	1162
LST.KE.-.lho7	C-GCA...CAAGTAAGA...CAA-AC--G--T--C-CCG...CTT-GAA-TGCACAGGAGGTTTGTAAAGAACAGGAGGAGGAGGC...CCC--G--A-----	2252
SUN.GA.98.L14	-GGCAG...CAAGTAGGA...AGACAAGCT--G--CA-AA-TTACCA...CCC-GAA-CAGCCAGGAGGTTTGTAAAGAACAGGAGGAGGAGGA...CCT--G--	2245
MND-1.GA.-.MNDGB1	GGACA...T-A...CAA-ATT-TG--G--CAG-GAGGGCCTCAA-GAGG-CCAGTTAGACAACTACTGGA...-G-----	1602
MON.CM.99.L1	GCC...AAC-TGCC...ATG-AT--G-A--C--GAGGACCA...CCCCA--G-AAGGGA...C	1621
MON.NG.-.NG1	---AGC...AAC-TGCC...ATG-AC--G-A-A...GGCAGAGGAGGTRGTAGCCAC-CCGACAGGCRWGC	199
MUS.CM.01.1085	--ACAG...--AGG-ACA...CTG-AC--G--A...GGAGCCAAGGGAGCCAGGGGAAGGGCAACGGCAGCCAGGGAACCGAGGA-	1632
DEB.CM.99.CM40	--AGAG...GCAGGTCCTTAGGC--G-A-ACA-A-G-GAGGA...CCCCCTGGTTCT---GA-----	1742
DEB.CM.99.CM5	---AGA...GCAGGTCCTTAGGC--G--CA-A-AGGAGGA...CAAGGAGGAGGTTCT---GA-----	1736
SYK.KE.-.KE51	C-CAGC...TTGGGC---A-A...-GCCAAGA...CAAGGAAGCAATCCC--G-GA...GGGC	1665
SYK.KE.-.SYK173	C-G---TA...-AT--G-A-AG-GCCATCAAAA...G--TCA-----	2025



H1B.FR.83.HXB2	TTGTTAAGTGTTCATTTGTTGGCAAAGAAGGGCACACAGCCAGAAATTGC...AGGGCCCTAGGAAAAAG	2022
Gag	I V K C F N C G K E G H T A R N C . R A P R K K . . . . .	p7
H1A1.UG.85.U455	GAA-----C-----A--CT---A---T.....	1471
H1B.US.90.WEAU160	C-A-----A-----T-----	2021
H1C.ET.86.ETH2220	CAA---A-----C-----G-----CT-----A-----	1393
H1D.CD.84.84ZR085	-----C-----T---A-----T.....	1542
H1F1.BE.93.VI850	-----A---T-----A---T-----A-----	1359
H1G.SE.93.SB6165	C-A-C-A-----G---A-TCT-----	1428
H1H.CF.90.056	-----A-C---C-----G-A---T-----	1372
H1J.SE.93.SE7887	-----C---C-----C---A---T---A-A---C-----A-----	1342
H1K.CM.96.MP535	-----C---C-----T-G---A---CT-----A---C-----	1218
H101.AE.TH.90.CM240	GAA---C---C-----T-G---A---CT-----A---C-----	1581
H102.AG.NG.-.IBNG	CAA-A---A---C---T-----A---CT-----A-----G-----	1547
H1N.CM.95.YBF30	CCA---A---A---C---T-----G-C---TTTG---A---C---T---A---A---G---GGA	1589
H1O.BE.87.ANT70	C-A-A-A-----C---A---G---A---T---A---A---T---A---A---A-----	2086
H1O.CM.91.MVP5180	CCA-A-A-A---C---A---G---A---T---A---A---C---T---A---A---A---G-----	2061
CPZ.CD.-.ANT	C-C-C-A-----T-C---C---T---G---C---T---T---A-----A-----A-A-GG-A	1452
CPZ.CM.-.CAM3	AAA---A-C---T---C-----G---A---TTT-----C-----A---GG---	1412
CPZ.GA.88.GAB2	GA-C---C---C-----A-G---T---T---A-----T---AA---A---G---	1400
CPZ.GA.-.CPZGAB	AAA-A-A-C---T-----G---T---CT---T-----T---AA-A-A-A-G---A	2089
CPZ.US.85.CPZUS	AAA-C-A---T-----T---G---A---TCT-----T---A---A---G---A	2070
CPZ.TZ.-.TAN1	AGC-GC---T-C---G---G---T---C---T---A-----T---T---A---A---G---A	1666
H2A.GW.-.ALI	CAA-----C-GG-----A-G-----T-G-A---C-A-----C-A-A-C-A---C---	2325
H2A.DE.-.BEN	CAA---G-A---GG-C---A-G---G---A---T-G-A---C-G-----C-A-A---A-G-C---	2332
H2A.SN.-.ST	CAA-----C-GG-----C-A-G-----T-G-A---C-A-----C-A-A---A-G-C-A	1777
H2B.GH.86.D205	CA-G-CA-C-GG-C---C---G-G-A---GC-A-----A-G-C---	2305
H2B.CI.-.EHO	CA-G-CA-C-GG-C---C---G-CG-A---T-----GC-G-----A-GGC---	2303
H2G.CI.-.ABT96	CAA-A-A---GG-----G-G-A---T---TT---C-G-----A---C-A-G-C---	1710
H2U.FR.96.12034	CA-C-GA---GG-C---G---A---T---A---GC-G-T---AA---A-G-C---	1823
MAC.US.-.239	CAA-----GG-----G---G---A---T-T-A---GC-A-----A---A-A-G-C---	2288
Gag	P I K C W N C G K E G H S A R Q C . R A P R R Q . . . . .	p8
SMM.SL.92.SL92B	-CA-C-G---T---C---A---GAC---A---T-T-G---C-G-----A---T---C---A-GG---	1740
SMM.US.-.H9	-AA-W---RG---R-A-G---A---T-T-A---C-A-T---A-----A-G-C-R	1773
STM.US.-.STM	CA-A-A-A-C-GG-----G---C-----T-A-A-C-G-T---AA-G---G-A-G-C-A	1947
SAB.SN.-.SAB1C	A-A---A---A---A---GCCT---T---CTG---T---TT---T---A---G-C-A	2327
TAN.UG.-.TAN1	CACCCCGC-C---AC---ATT---T-T-CAA-A-G---T---CCAAGAG-AG-ACC---C	2164
VER.KE.-.AGM155	CGCCA-----A---C---A---TTT---C---TGAG---GC-G-T---CCT-AG---A---TA	2187
VER.KE.-.9063	CAC-A-A---A---C---A---TTC---A---TGCAA---GC-A---CA-AA-A-A-G-TA	2204
VER.DE.-.AGM3	CA-A---A---A---A---TTT---C---T-TGCAA---C-A---CCT-AA-A-A-G-T	1699
VER.KE.-.TY01	CAC-A-GA---AT-----A---TTT---C---T-TGCAA---C-A-T---CC-AA-A---CA	1685
COL.CM.-.CGU1	. . . . .-A-C-----CAGGG-AT-----TT-----TG---TCCC-AAAGA-A-TAGG-GGAGCCGGTTCGAGGAAGAGGCCGAGGACGGGGAGGCTTTAGAGGAGCTCCAGAGGCCAGT	1850
DEN.CD.-.CD1	GAC---GA-----C---ATT---A---CT-----AGG---T---AAAG---C-A---G-GT---ACTCAACCAGG	1815
GRV.ET.-.GRI_677	CGC-A-A-A-C---T-----TTT---A---T-TGCAA---GG-A---A---A---A---AC-G-TC	2129
GSN.CM.99.CN166	CCCCA-A---AT-----AC---TTC---A---CT---A---G---G---T---CCTAAA---C-AAG---GA	1705
GSN.CM.99.CN71	CCCCA-A---A---A---A---GTTT---A---T-TC---A---G-C-T---CCCCAA---C-A-G---GA	1708
DRL.-.-.FAO	ACCC-GA---CCA-C-C---GC-GTTT---A---T---CCTG---C-A-T---CAAAG---A-----	1550
RCM.GA.-.GAB1	. . T-A-GA---A---T---GCC---A---TTT-T---T---T---AA-----A-GG---	1462
RCM.NG.-.NG411	-CC-C---A-C-AT-----AC---T---T---T---T---AAA-AAG---A-G-A	1465
MND-2.-.-.5440	ACCC-GA---CCC-----C-A---GTTT---A---TGT-CTG---GG-----A---A---A---G-GA	1547
MND-2.CM.98.CM16	ACA-C-GA---CCCT---C---A---TTT---T---TGGCTG---GG-C-T---TAAG---C-A-----	1990
MND-2.GA.-.M14	ACCC-GA---CC---C---A---GGT-T---A---TGTCTG---G-C-T---ACT---C-A---G---	1914
MNE.US.-.MNE027	CAA-----GG-----G---G---A---T-T-A---GC-A-----A---A---A---G-C---	1768
LST.CD.88.447	C-T-A-CT---C---T-----A---GCC---T---T---A---TG---T---ACAG---AC-GG---	1236
LST.CD.88.485	C-T-A-CT---C---T-----A---GCC---T---T---A---TG---T---ACAG---AC-GG---	1236
LST.CD.88.524	C-C-C-CT-----G---GCC---A---TT---A---A---TG-----CAG---AC-GG-A	1230
LST.KE.-.lho7	CCC-C-CT---C-----A---GCC---T---A---TG-----ACAG---C---AC-GG-A	2320
SUN.GA.98.L14	CCA-G-CT---T-----T---CC---A---TCT---A-ATC-G---A-A-G---A-A---GGA	2321
MND-1.GA.-.MNDG1	AACC-TC-C---C---AAT-----TGT---A---GTTC-T---A-----A-GG-A	1670
MON.CM.99.L1	AACCC---A-C---A---TTT-----TG---G-A---AA---CCAA-GG-G-	1689
MON.NG.-.NG1	AAA-C-G---AT-----G---TTT---A---TGTG---A-AG---C---T---CT---T---C---A---G-CA	267
MUS.CM.01.1085	A-CCC-GA---C-A---C---GC---T-T-C---GT-----GG-----CCCCAA---C---C-A	1700
DEB.CM.99.CM40	AGA---GA---C-----AC---ATT---A---TCTCAA---AGG-C---AAAGA---A---G-CC	1810
DEB.CM.99.CM5	AGC---GA---C---T-----AC-GATT---TGTCAA---GG---T---AAAAG---A---GTT	1804
SYK.KE.-.KE51	CAACA-GA---C---T-----GC-GTT---A---TCTTCAA-A-G-C---CCTAGG---A-AA---GCTT	1733
SYK.KE.-.SYK173	-GA---A-C-A-----GC-ATT-----TGCAA-A-G-C---AAAAG-ACCT---GCT	2093







PLV  
complete genomes

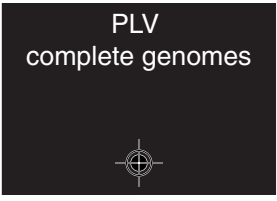


stem-loop after TTTTTT slip potentiates slippage  
Gag-Pol -1 ribosomal slip site  
Gag p7 end \\/ Gag p1 start

H1B.FR.83.HXB2	.GGCTGTTGGAAATGTGGAAGGAAGGACACCAAAATGAAAGATTGT.....ACT.....GAGAGACAGGCTAATTT.TTTAGGGAAGAT...CTGGCCTTC.....	2108
Gag	_G_C_W_K_C_G_K_E_G_H_Q_M_K_D_C...T...E_R_Q_A_N_F_L_G_K_I...W_P_S...	p1
Pol	.....F_R_E_D...L_A_F.....	Pol TF
H1A1.UG.85.U455	.....C-G-A.....C-C.....A...T.....	1557
H1B.US.90.WEAU160	.....C.....T.....	2107
H1C.ET.86.ETH2220	.....C.....C.....GAC...T.....	1479
H1D.CD.84.84ZR085	.....C.....A.....T.....	1628
H1F1.BE.93.VI850	.....GA.....C-C.....A.....T.....	1445
H1G.SE.93.SE6165	.....T.....C-C.....G.....A.....T.....	1514
H1H.CF.90.056	.....GA.....T-G.....C-C.....A.....T.....	1458
H1J.SE.93.SE7887	.....C.....C.....T.....	1428
H1K.CM.96.MP535	.....G.....T-G.....C-C.....A.....T.....	1304
H101.AE.TH.90.CM240	..-T.....C-G.....T.....C-C.....A.....T.....	1667
H102.AG.NG..IBNG	.....G-A.....C.....A.....T.....	1633
H1N.CM.95.YBF30	.....G-GC-A.....T.....AAAAATGAA.....GA.....G.....T-C.....	1681
H1O.BE.87.ANT70	..-T-C.....C.....T.....C.....GAAAT.....GA-A.....A.....C-ATA.....C.....	2175
H1O.CM.91.MVP5180	..-T-C.....C.....T.....C.....AAAAAT.....GA.....A.....TA.....C.....	2150
CPZ.CD..ANT	.....C-G.....GC-T.....A-C.....C-AGCAACA.....AATACAG-A-A-TA.....ACC...GACC-CA-GTGTG...GGG	1556
CPZ.CM..CAM3	..-A-C.....G.....GC-A-G-G-T.....A-C-C.....C-AGCTGGA.....A.....AGG...T--T-C-GTGGAG...CGG	1513
CPZ.GA.88.GAB2	..-A.....G.....A.....T-G.....G-C-C.....CAAAAT.....GAA-GA.....TG.....GG-GG...T.....	1492
CPZ.GA..CPZGAB	.....CGG.....GC-A-G.....C-C.....A.....A.....TG.....AGG...T.....CCGCAG...CGG	2184
CPZ.US.85.CPZUS	..-G-C.....G.....GC.....T.....C-C.....AGCAGGA.....A-C.....ACA...T--T-CC-ATCGTGGAGCGG	2174
CPZ.TZ..TAN1	..-T-C.....GG.....C-A-G.....T.....G-C-C.....CACCAGAAAC.....AACAGC-CTGG-TA.....ACG...CA-CC-CTTGTG...GGG	1770
H2A.GW..ALI	.....C.....C-CC.....T-TT.....GC-A-C-C.....C-A.....A.....GG.....CTTGGCCCA--GG...	2411
H2A.DE..BEN	.....C.....G.....C-CC.....ATC.....GC-A-C-C.....C-G.....A.....AGG.....TT-GGCCAC--GG...	2418
H2A.SN..ST	.....C.....C.....ATC.....GC-A-A-C.....C-A.....A.....GGG.....TT-GGCCCA--GG...	1863
H2B.GH.86.D205	..-A-C.....AAC.....ATC.....TC-A-A-C.....C-A.....A.....GGG.....TTAGGACC--GG...	2391
H2B.CI..EHO	..-A-C.....C.....ATC.....TC-A-A-C.....C-A.....A.....GGG.....TTGGCCCA--GG...	2389
H2G.CI..ABT96	..-G-C.....G.....GC.....ATT.....GCCA-A.....C-A.....A.....AGG.....TTTGGACCT--GG...	1796
H2U.FR.96.12034	..-A-C.....A.....CC.....A-C.....GCTAGC-C.....C-A.....A.....CCA-TGGG.....CTTGGCCCT--GA...	1909
MAC.US..239	..-A-C.....AATG-AC-TGTT.....GCCA-A-C.....C-A.....C.....GGG.....CCTTGGTCCA--GG...	2374
Gag	G_C_W_K_C_G_K_M_D_H_V_M_G_K_C...P...D_R_Q_A_G_F_L_G_L_G_P_W_G...	p1
Pol	.....F_R_P_W_S_M_G.....	Pol TF
SMM.SL.92.SL92B	..-A-C.....GG-A.....GTATTCA-GCCA-C-C.....C-A.....A-CCAGA.....GGG.....CCTTGGCCCA--GG...	1826
SMM.US..H9	.....C.....G.....A-C-C-TGTT.....GCTA-A-C.....C-C.....A.....GGG.....CCTTGGCCCA--GG...	1859
STM.US..STM	..-T.....C-AC-G-C-T-G.....GCCA-A-C.....C-A.....A.....TGGG.....CTTGGCCCA--GG...	2033
SAB.SN..SAB1C	.....C.....C-GCCC-AC-T-G.....CAA.....A.....T.....TTTGGCCC--GG...	2410
TAN.UG..TAN1	.AAA--CCTA.....G--CC.....TT-GCA.....C.....CGCAGT.....G.....C.....GA--GCCAACCTG...	2250
VER.KE..AGM155	.AAA--CTT.....CC--G--TT-GCA.....C-C.....GG.....G.....TG.....T-TGGCCGG--AT...	2270
VER.KE..9063	.AAA--CCTT.....G--CT.....TGGCA.....C-C.....GG.....G.....TG.....T-TGGACGG--A...	2287
VER.DE..AGM3	.A-A-C-T.....G--ACC--G--TTT-GCA.....C.....GA.....G.....TA.....T-TGGCCGG--AT...	1782
VER.KE..TY01	.AAA--CTA-G.....ATTG.....T-GCA.....C-C.....GG.....G.....TG.....T-TGGACGG--AT...	1768
COL.CM..CGU1	GA-G--TT-CC--CAATC.....ATGCAA-G.....C-C.....A-C-AG--A.....	1922
DEB.CD..CD1	A-CG-C-TT.....CT--T-T--GCA-GA.....AGA-G-AAC.....ACTAATGATC-G--TA.....TGGTCTTG--AGGAA...	1920
GRV.ET..GRI_677	.AAA--C-TT--G--C--AATT-C-TATGGCA.....C-C.....AGAAT.....G.....T-TGGCCAT--GG...	2215
GSN.CM.99.CN166	.AAG-C-TT.....C-GA-C--G--TTT-GC--CCA--AGA...-A.....GGGACACCAATC-AGG-CT-T.....ACCTTGGCCGGG--G...	1803
GSN.CM.99.CN71	.AAG--TT.....C--C--TT-GCT-GGC-A-CAAA...-A.....GGAACACCA-CC-AGG-CT--ACCTTGGCCGGG--G...	1806
DRL..FAO	..-A-C-TT-GG.....GCTTT-AT-TATGT-A-G-A--C.....C-AAAA.....TG.....T-C.....	1624
RCM.GA..GAB1	..-A-C.....TCCCC-AC-T-TCT--C-C.....AAAG.....AATA.....GAC-CCC--GG-CA...	1548
RCM.NG..NG411	..-G-C.....G.....T-T-T--G--C.....C-A.....A-G--T.....TGC...TCCCTGGAA...	1548
MND-2..5440	.....TT-G--G-TCC--T-T--GGA-C-C.....C-AAAG.....AT-TG.....T-C.....	1621
MND-2.CM.98.CM16	..-A-C-TC-G--G-TCT--AT--G-A-C.....C-AAAG.....AT-TG.....T-CCCC--GGCAG...	2076
MND-2.GA..M14	..-A-C-TT-G--GG-CAC--TATG--G-A-C-C.....C-AAAA.....AT-TG.....TGC.....	1988
MNE.US..MNE027	.....C.....C-AATG-C-TGTT--GCCA-A-C.....C-A.....C.....AGG.....CCTTGGCCCA--GG...	1854
LST.CD.88.447	..-A-C.....C--T-GCA-G-A--TAGGT-TGCTC-A--C-A.....AAG.....CCTA-G-A--TA.....CT-TGGACC--AG...	1328
LST.CD.88.485	..-A-C.....C--T-GCA--A--TAGGT-TGCTC-A--C-A.....AAG.....CCTA-G-A--TC.....CT-TGGACC--AG...	1328
LST.CD.88.524	.....C.....TGCTA--AC--AG--GCCG-G-C.....C-A.....AAA.....CCAA-G-GA--TG.....T-TGTTCC--AG...	1322
LST.KE..lho7	..-T.....C--C-GTA--AG--TAG-T-TGC-C-G-C.....C-A.....AAA.....CCAA-AG-A--TG.....AT-TGTTCC--AG...	2412
SUN.GA.98.L14	CTC--C.....ATG--TA--CAAGCTC-A-C.....C-ACAA.....GA-A--TG.....T-TGGCCC--AA...	2411
MND-1.GA..MNDGB1	..-G-C--T--GCAATG-AT-T-G-AAGCTC-A-C.....C-AAGCCAGCT.....CAGC--CAGAG--T.....T-TGGCCCT--GG...	1768
MON.CM.99.L1	.AAA--AC--T--CC-ACCG--TCGG--C-C.....C-GCAGCCC...CCCAACAA...A-C-A-GG-TG.....TCCTTT-G-...	1787
MON.NG..NG1	..-T-C-TCCG-C-C--G--TRT-TCC--A-C-C.....C-AAACGGAGGACAAACCAACCCAGA-ACAG-TG.....TCCTTTG-C-...	374
MUS.CM.01.1085	.AAG--TC.....G-GA.....TTGGCA-G-C-A--AGGTCTGATAATACC-A-GC--TG.....G-TCCCTT-G-C-T...	1795
DEB.CM.99.CM40	.AAA--TT.....GC--C--ATTGCA--A.....GGACAGACTCCGAG--A.....T-CTCC--G...CTCCAGGAGACC	1911
DEB.CM.99.CM5	.AAA--C-TC.....T--G-C--AT-GCA-GA-C.....T.....TGACAA-TACC-AG-A.....TGCTCC--G...CTCCCGGAGACC	1905
SYK.KE..KE51	.AAG-C-TT-C--GGAACG--AT-GCA-GGC-G--AGACAAC-AAGAAAAGGACAAGGCAACCTCTCTCC--C.....AGGGTGGG--AA...	1846
SYK.KE..SYK173	.AAA-C-TT-T--AAC--T-GCA-GG-CC--AGGCAAC-CAAGAGAAATCAAGGCCTGCA-TTGC--AGGATGGG--GT...CTC	2209



	Gag p1 end \ / Gag p6 start	
H1B.FR.83.HXB2	CTACAAGGAAG...GCCAGGGAATTTCTTCA...GAGCAGACC...AGAGCCAAC...AGCCCCACC...AGAAGAGAGCTTCAGG	2181
Pol	L Q G K . A R E F S S . . . . . E Q T . . . . . R A N . . . . . S P T . . . . . R R E L O	Pol TF
Gag p1	Y K G R . P G N F L Q . . . . . S R P . . . . . E P T . . . . . A P P . . . . . E E S F R	p6
H1A1.UG.85.U455	-A-----G-----C-----C-----T-----A-----	1630
H1B.US.90.WEAU160	-C-A-----C-----C-----T-----A-----	2180
H1C.ET.86.ETH2220	-A-----C-----C-----C-----C-----GAGTCT-AGACCA	1552
H1D.CD.84.84ZR085	-C-----G-----C-----C-----C-----AGCT-CGG-T	1701
H1F1.BE.93.VI850	-A-----G-----C-A-----C-----C-----G-----	1518
H1G.SE.93.SB6165	-AG-----G-----C-----A-----GA-----A-----T-C--A--C-G--	1587
H1H.CF.90.056	-AG--A-G-----A-----C-----G-----A-----C-----G-----	1531
H1J.SE.93.SE7887	-AG--A-G-----C-----C-----C-----C-----C-----G-----	1501
H1K.CM.96.MP535	-C-----G-----C-----C-----C-----C-----G-----	1377
H101.AE.TH.90.CM240	-A-----G-----C-----C-----C-----C-----A-A-GGG--	1740
H102.AG.NG.-.IBNG	-AG-----G-----A-----C-----G-A-----T-----TG--	1706
H1N.CM.95.YBF30	--T--A-G--.A--A-C-C-CC--CA-C-A--AG-AA-GAGCCAC--CT--T-ATG--	1760
H1O.BE.87.ANT70	GGGGGGCAGC--C--A-G-G--ACC-G--C-C--T--T--G--	2232
H1O.CM.91.MVP5180	GGGGGGCAGC--C--A-G-G--GTCC--T--T--G--	2207
CPZ.CD.-.ANT	G-G-G...A-----C-G-G--AGGAGGAAGTAGT.G-----T-----CAT---T--A...	1625
CPZ.CM.-.CAM3	GGGATCAAAG--A-----C-C-G--G--GAA--G-A-C--T--G--CCT--T-ATG--	1586
CPZ.GA.88.GAB2	-CGG-GC-G--C-----C-CC--A-G-AGGT--G-A--T--TCC---C---G--	1565
CPZ.GA.-.CPZGAB	GAG--C-----C-C-G-G--A--A--A--G--GAT--T-ATG--	2248
CPZ.US.85.CPZUS	GGGATCAAAG--A-----C-C--GAA--CAAG--C--T--T--GAT--G--	2247
CPZ.TZ.-.TAN1	G-G-G...A-----C-G-G--A-CC-AGA.GA-AGGG-A-----	1815
H2A.GW.-.ALI	AA-G--CCTC...CAACTTCCCCG-AACC-G...AGTTCG-A...G-GATT--ACCAAC...C-CCCCAGCAGA-	2481
H2A.DE.-.BEN	AA-G--CCTC...CAACTTCCCCG-GACC--...AGC-CCT-A...G-G-TG-TACCAAC...C-CCTCCGGCAGAT	2488
H2A.SN.-.ST	AA-G--CCCC...CAATTTCCTC-G-GGCC--...A-T-CCG-A...G-G-TG-ACCAAC...C-CCCCGA-AGAC	1933
H2B.GH.86.D205	AA-G--CCTC...CAACTTCCCCA-GACC--...AGTGCCT-A...G-GAGTG-ACCATC...T-A-C--TGAA-CCA	2464
H2B.CI.-.EHO	AA-G--CCTC...CAACTTCCCCG-C-AGGC...ACC-CAGGG...GATAGTG-ATC...T-G-C--TGAA-CCA	2459
H2G.CI.-.ABT96	AA-G--CCTC...CAACTTCCCCA-GACC--...AGT-CCT-A...G-GT-TG-TCCATC...C-CCACCGA-GGAC	1866
H2U.FR.96.12034	AA-A--CCTC...CAACTTCCCCA-GGCC--...AGT-CCG-A...G-GA-TG-CCCAAC...CCCCTCCAGCAGAA	1979
MAC.US.-.239	AA-G--CCCC...CAATTTCCTC-GGC--...AGTGCAT-A...G-G-TG-TGCCAAC...T-T-C--	2437
Pol	K E A P . Q F P H G S . . . . . S G S . . . . . G A D A N . . . . . C S P	Pol TF
Gag p1	K K P R . N F P M A R . . . . . V H Q . . . . . G L M P T . . . . . A P P	p6
SMM.SL.92.SL92B	AA-G--CCTC...CAACTTCCCCA-G-AGAC...A-CATC-TT...GACA--GT-AGCACCCAGATCC...--AG-C-G...	1898
SMM.US.-.H9	AA-G--CCCC...CAATTTCCTC-GGCC--...TGCCT-A...G-G-TG-ACCAAC...T-T-Y-	1922
STM.US.-.STM	AA-G--CCCCA...CAACTTCCCCA-GGCC--...A-TACCT-A...G-G-TG-GCCAAC...T---T-	2096
SAB.SN.-.SAB1C	-AGAGG-AA-CC...CAGGAATTTCCC--GAC...CTC--TCAG...CCAA--G-TCCCCAATGGA...ACGGGATTA-AGCAG...GCCG--GAAAATT--	2501
TAN.UG.-.TAN1	GGGA-CAAAGCC...GCAATTT--AGAG...AG-G-G...C-TT-G-GGC...TCCCCGAT...GCCT-G-ACGGATTCCC...	2328
VER.KE.-.AGM155	GGGG-CAAA-CC...AAG-AATTT-CCCGAGC...C-CTCTTGG...G-C-GA-C-AAG...T-G-C--TCC-CC--A-AA-TCT	2346
VER.KE.-.9063	GGGG-CAAA-CC...AG-AACTT-C-CGC-GC...C-CTCATGG...T-G-A-C-GAG...C-G-C--TCC-CC--CACC...	2359
VER.DE.-.AGM3	GGGAGC-AA-CC...CAG-AATTT-CCCGCGC...T-CTCTTGG...G-T-GAGC-AAC...T-T-G-C--TCC-CC--TCC...	1854
VER.KE.-.TY01	GGGGCAAA-CC...AG-AATTT-CCCGCGC...T-CTCTTGG...--C-GA-C-GAG...T-G-T--TCC-CC--GG-CC	1844
COL.CM.-.CGU1	...GCA-CCAG	1931
DEN.CD.-.CD1	GA-TCCTCCC--GAACTT...TGTACAGTACTA...CCCTCAGATGA-TCAGGATTT-	1972
GRV.ET.-.GRI_677	AGGAGC-AA-CC...AAG-AATTT...T-TG-A-TACAGAGGAGA...CAC--TTG-TC-GGAA	2270
GSN.CM.99.CN166	-CG-CGA-CCCCAG-AACTTTCCAG-AAC-TTCTGTGCTGCCAACAGC-CCGG...G--GAAGAACCC	1873
GSN.CM.99.CN71	...ACGA-CCCCAG-AACTTTCCAG-AAC-TTCTGTGCTGCCATCAGC-CCGG...G--GTGGAAACCC	1873
DRL.-.-.FAO	TCC-TG--C--CGG-AA-CCC-GGAA-T--C...TCAGGCTT...GCTGACCCCATC--CTCCTCCA--GCCA	1694
RCM.GA.-.GAB1	GGGG--CCG--...AACTTTCC--GAC-TC...CTTG-CT--CTCTG-TC-AGGGATGGAGAGCAATTA...C--CCGTGCAAGA-	1630
RCM.NG.-.NG411	GC-GGG-AA-CCGAG-AACCTTCCC--GAC-TC...CTTG--TACTG-TC-AGGGAGAGAGGAAGGATGGAATC--A...G-A-GAT	1633
MND-2.-.-.5440	-CC-TG--C--TGGAAGCCC-GGAAT-C-C...TGC--TG--ATTGACCCCATC--C-CCCCGA-GCCA	1691
MND-2.CM.98.CM16	TGG--ACCC--GAACTTCCC-GCGA-G-CATT...TGC--TG--GACCCCAACG-CCCCTCCA-GCCA	2134
MND-2.GA.-.M14	TCC-TG--C--TGGAAGCCC-GGAAT-C...TGC--TG--ACTGACCCCATC--C-CCCCGA-GCCA	2058
MNE.US.-.MNE027	AA-G--CCCC...CAATTTCCTC-GGCC--...A-TGCAT-A...G-G-TG-ACCAAC...T-T-C--	1917
LST.CD.88.447	GAGTGGTCTCC...AGG-AACTT-CC-T-GAT...G-GG-GAA...T-CAGG-GTAGTACC...	1382
LST.CD.88.485	GA-TGGTCTCC...AGG-AACTT-CC-T-GAT...G-GG-GAA...T-CAGG-GTAGTACC...	1382
LST.CD.88.524	GA-TGGACCTTC...AGGCAACTT-CCAA-CAT...G-AG-GG...T-CAGG-GTAGTACC...	1376
LST.KE.-.lho7	-AG--ACCTCC...AGGCAATT-CC--A-T...G-AG-G...T-CAGGG-GAATTC...	2466
SUN.GA.98.L14	-AGGGG-CCGCC...AGGGAACTT-CCAG-CAT...CCTTC...T-T-G--TC-AGA-	2463
MND-1.GA.-.MNDGB1	TCC-TCCAA-CC...GGGAATT-CC...T-TG-A-TACAGAGGAGA...CAC--TTG-TC-GGAA	1792
MON.CM.99.L1	TCC---AAGG-GCCAAGGAACTTCCC--CAC...CTCTGTGCAG--	1832
MON.NG.-.NG1	-GG---A-GG-GCCAAGGAACTT-CCAG-AAC...CTC-TTG-A...GCCAA--G-CCCAGC...	431
MUS.CM.01.1085	-C-G-AA-G-GCC-AGGAATTTCCC-AGAC...CTCTCT-TTG--	1840
DEB.CM.99.CM40	-AGG--CTTCTAGAAGG-ATTGGGAAGTACAT...CCCGGAGAG...	1953
DEB.CM.99.CM5	-AGG--TTTTTGAAGGGATTGGGAATACAT...CCCAGAGAG...	1947
SYK.KE.-.KE51	-AGG--ACC...C-CAGCAAATTTCTGT--TG--GGAGACCAAC--CC--C-AGCGGAGGATTGGCCATGGCAGCAA-AGA-ATGGGAAATAT-CTTACC-CAACA-	1958
SYK.KE.-.SYK173	-AGG-G-CCCCAGCCAATTTCCCAG-GAGGAGCGAGCCATCTGCTCC-CCT-T...G--G--C-T-GAAGAT	2279





PLV  
complete genomes



H1B.FR.83.HXB2	TCTGGGGTAGAGACAAC.....	AACTCCCC.....	TCAGA	2212	
Pol TF	V W G R D N.....	N S P.....	S E Pol TF		
Gag p6	S G V E T R.....	T P P.....	Q.....		
H1A1.UG.85.U455	ATG---A-A---T.....	.G-CT---T.....	-GC--	1658	
H1B.US.90.WEAU160	-TCA---A-----	-----T-----	-----	2211	
H1C.ET.86.ETH2220	GAGCCAAC-CCC-C-ACCAGA.....	.GAGCTTCAGGTTTCGAGGAAG-AA.....	A-CTT	1604	
H1D.CD.84.84ZR085	-TG--AGGAGATA-C.....	.CCCT-C-AGAA.....	A--G	1734	
H1F1.BE.93.VI850	-TCA-A-AG---T---	.CC-CT-T---	GA--C	1549	
H1G.SE.93.SE6165	-TC--A-AG---T-G-	.....	C-CCT	1609	
H1H.CF.90.056	-TC--A-AG---TG--	.....	C--T-	1559	
H1J.SE.93.SE7887	CTC--A-AG---T---	.....	C--T-	1526	
H1K.CM.96.MP535	-TC---AG---T---	.CC-CT-T---	GAG-C	1408	
H101.AE.TH.90.CM240	ATG---A-----T---	.GGGGGAAGAGATAAACCT-CTTA--	GA--C	1786	
H102.AG.NG.-.IBNG	ATG---A-----T---	.ACCTC-CT-T---	A--C	1737	
H1N.CM.95.YBF30	-T-CA--AG-----	.GAAGAGCACACAGGGGAAGGAG-TG-AGGA.....	GA-CC	1806	
H1O.BE.87.ANT70	.....	.GATGGAGGAGGAAGTGAAGGG-CAGGAGAA.....	---G	2267	
H1O.CM.91.MVP5180	.....	.AATGGAGGAGGCAGTGAAGGA-CAAGAGAA.....	-----	2242	
CPZ.CD.-.ANT	..-CA--AG---CA.....	.CAAGAGGACTCAGAAGGGTCTCAA.....	GGG-G	1666	
CPZ.CM.-.CAM3	-A-CA--A--CA---GG.....	.AAAGAGACAGGAGGAGGAGGGGAGAAGGA.....	AA--G	1638	
CPZ.GA.88.GAB2	-A--CA-AGC-----CA.....	.GACAGGAAAGAGT CAGAAGGGGGAGGAGT---	AACAG	1620	
CPZ.GA.-.CPZGAB	-ACCA--AG--AGA.....	.....GAAGAGCCAGGAGAAGAAAGAGGGAGA.....	GAGC-	2294	
CPZ.US.85.CPZUS	-A-CAA-A-----GT.....	.....GACACA.....	GG-A	2290	
CPZ.TZ.-.TAN1	..G-CTCAG--.....	.GCAGGAGACAGCACAGAC-C-AGT.....	GGTG-CA-	1856	
H2A.GW.-.ALI	C-A-CA-CG--CCTGCTGGAGCAGTATATGCAGCAGGGGAGAAAGCAAAGAGAGCAGAGAGAGAGGCCATACAAAGAGGTGACGGAGGACTTACTGCAC-T-GA.....	.....	G---G	2590	
H2A.DE.-.BEN	C-A-CA-CG--ACTGTTGGAGAGATATATGCAGCAAGGGAGAAAGCAGAGGGGAGAGAGCCATACAAAGAGGTGACGGAGGACTTCTGCAC-T-GA.....	.....	G---G	2597	
H2A.SN.-.ST	C-A-TA-AG--CCT-CTAGAGAAGTACATGCAGCAAGGGGAAAGGAGAGAGAGAGAGAGGGCCATACAAAGAGGTGACAGAGGACTTCTGCAC-T-GA.....	.....	GA-AC	2042	
H2B.GH.86.D205	G-A-A--GCAT---C-TCGGGGGGCGACACCATCTGCGCCCTTCAGATCCAGCAGTGGAGATGCTGAAAAGTTACATGCAGATGGGAGACA-CAGAGAGA.....	.....	GAGCC	2573	
H2B.CI.-.EHO	G-ATTC-GCAT---C-TCAGGGTGGGATTCATCTGCACCCCTTCAGATCCAGCAGAGGAGATGCTAAAAGAACTACATGCAACTAGGGAAGGAGAGAGGA.....	.....	GA-C-	2568	
H2G.CI.-.ABT96	H2G.CI.-.ABT96	.....	1924		
H2U.FR.96.12034	C-A-CA--G--TCTGCTGACCCCGACAGCCCTCCAGCGGACCCAGCAGTGA.....	.TCTGCTAAAGAGCTACATGCAACAGGGCAA.....	G-AA-AAAA.....	GGAGA-C-	2079
MAC.US.-.239	.....	.AGAGGACCCAGCTGTGGATCTGTAAAAGAACTACATGCAVGTGGGCAAGCAGCAGAGAGA.....	AAAGCAGAGAGAAAGC-	2514	
Pol TF	.....	G S A K E L H A V G Q A G E R.....	K A E R K Q Pol TF		
Gag p6	.....	E D P A V D L L K N Y M Q L G K Q Q R E.....	K Q R E S.....	p6	
SMM.SL.92.SL92B	.....	.GATAGTCAAGGAGTATCTAGAGAAGCACAG-GGGAAAAGAC.....	AAG--	1945	
SMM.US.-.H9	.....	.AGAGGATCCAGCTGTGGATCTCTGAAGAATTACATGAAGTTGGGAGAGGCAGAGGGA.....	GA-C-	1987	
STM.US.-.STM	.....	.AGAAGACCCAGCTGCGGATCTGCTGAGAAGTTACATGCAGCTGGGCAAGAAGCAGAGAGA.....	GAGC-	2161	
SAB.SN.-.SAB1C	-A--CA-ACAGAC-TC-GACCAGAGGACCAGGGCCAGACGATCCAGCAACAGCCCTGTAAAGCAGTATGCTGTTCAGGGGAAACGGCAGAAACAGCAGTGG-A.....	.....	AA-CC	2610	
TAN.UG.-.TAN1	..CACA-GTTC-C--G-TGCGGGAGC.....	.ATACGACCCAGCCAGGAACTGCTGGAGCAGTATGCCAA.....	AAAGG.....	AG-TC	2402
VER.KE.-.AGM155	A-ACC.....	.TTACGACCCAGCAAGAGCTCCTGCAGCAGTATGCAGAGAAGGGAAACAATGAGAAATCAGAA-AG.....	AA-CC	2425	
VER.KE.-.9063	.....	.ATACGACCCAGCAAGAGCTCCTGCAGCAGTATGCAGAGAAGGGAAACAAGTTGAGGGA-CAAGGGAA.....	GAGA-	2433	
VER.DE.-.AGM3	.....	.ATACGACCCAGCAAGAGCTCCTGCAGCAATATGCAGCAAGGGGAAGCAGTTGAGGGA-CAAGGAA.....	AA-AC	1928	
VER.KE.-.TYO1	A-CCC.....	.ATACGACCCAGCAAGAGCTCCTGCAGCAATATGCAGAGAAGGGAAACAAGTTGAGGGA-CAAGGAA.....	GA-TC	1923	
COL.CM.-.CGU1	.....	.GGGGTGGAGCTACAGACAGCCATCTTT.....	CCCCCGAAGATGAGC-AAGA--T.....	G-CAC	1987
DEN.CD.-.CD1	C-A-T--AGCCAGTG-GGGTCCAGCAGCAACCAGTACCCACAAAATTTCAAGAGGGGGAAACAGACAATCTGTGAGGAAAGCAGAGGAGC-ACAAG-AGTC.....	.....	C--TC	2084	
GRV.ET.-.GRI_677	C-AAACA-CCCCC--TGGAACAGC.....	.TTACGATCCAGCAAGAGACTCCTCCAGCAGTATGCAGAGAAGGGACAGCCGCTGAGAGAGGAGAG.....	AG-AC	2367	
GSN.CM.99.CN166	CAGCA-C-CA-TCTCT-.....	.GACACCCAGTAGGGGA.....	ACCCCA.....	GAGC-	1916
GSN.CM.99.CN71	CAGCA-C-CA-CCTGT-.....	.TCCGCCGTGGGGGA.....	GCSCCA.....	GA-C-	1916
DRL.-.-.FAO	GGATACC-CC--GAGGATCCGGCCGAGAGGATGCTCCAAAAGTACATGGAGCAGGGAGCTCAGCAGAAGAGACAACA.....	.....	GAGACAACA-CA.....	1783	
RCM.GA.-.GAB1	ATGCTAAAGA-.....	.TTATCTGAGGAGGGCAGGGGAACAAAAGAGACAACAG-GG-AGGA.....	AG---	1691	
RCM.NG.-.NG411	C-A-CA-A-----TGCTCAGGAA.....	.ATACCTAGCATTGGGGAGACAGCAACAACAGGAGCAG-GGAAGA.....	GA-C-	1706	
MND-2.-.-.5440	GGGTT--A---CC-GG-AGAAAAGATGCTGTTAGACTATATGAAGAAAAGGGCAACAACAGAGGGCAGCAGCAGGAGC.....	.....	C---	1773	
MND-2.CM.98.CM16	GGGAT--A---CC--G-AGAGAGAATGCTGTTAGATTACATGAAGAAAAGGGCAACAGCAGAAAAGCGGAAAGCAA.....	.....	A---G	2213	
MND-2.GA.-.M14	GGGAT--AG--TC--G-AGAGAAGATGTTATTAGACTATATGAGAAAAGGGCAACAACAGAGGGCAGCAAGGAGAGC.....	.....	CA-C-	2140	
MNE.US.-.MNE027	.....	.AGAGGACCCAGCTGTGGATCTGCTAAAAAATACATGCAGTTGGGCAACAGCAGAGAGA.....	AAGC-	1982	
LST.CD.88.447	.....	.ATCAGTCCCCCAATGGAGAGGAGCCCAACAAGCAGAGAGAGCCTTAGA.....	.AACATACAGGAACCTCGGACAACAGCTGAGGAG.....	G--AC	1471
LST.CD.88.485	.....	.ATCAGTCCCCCAATGGAAAAGGAATCAACAAAAGCAGAGAGAGCCTTAGA.....	.AACATACAGGAACCTCGGACAACA-CTGAGGAA.....	G--AC	1471
LST.CD.88.524	.....	.TTCAGCGCCACCCATGGAGAAAGTGCCTCCAACAAGAGCAGAGAGGGCAATAGAGACATACAGGAACCTAGGACAACA-CTGAGGAA.....	G--AC	1468	
LST.KE.-.lh07	.....	.ATCAGTCCCCCAATGGAGAGTGCCTCCAACAAGCAGAGAGAGCCTTAGA.....	.GACTTACAGGACTTTAGGACAACAGCTGAAAAG.....	G--C	2555
SUN.GA.98.L14	GA-CT-AC-TT-GGG-A.....	.CAGAATGACACCACCCAGTCAAAGGCAGAGAGAGCTCTGGAGACCTACAGACTGTTAGGACAGGGC-T-AG.....	AGCAC	2557	
MND-1.GA.-.MNDGB1	.....	.GG-ACAAGAGTGACTCCAACAGCTCCACCATTAGAGGAGAAACCTCTGCAGAAAACCTTGAGCACTTATCAGAAAATTAGGG-GAGGG-T.....	CAG-C	1887	
MON.CM.99.L1	.....	.CAGCGACCCAGCTCTGCCAGCAGGACTCGGGAGCCAGTCAATCTCAACCAAGAGATTGG-T---	GA---	1903	
MON.NG.-.NG1	.....	.GGAACCAACCCARGGGCTACAAAGTGAACCTCTCACTCAGTGGGACCCACAGAG--C-AG-	G-C--	502	
MUS.CM.01.1085	.....	.AACTGCTCCGCAGAACAGGAGGGGAAATGGTCAATCTCTC-AG-AGT.....	GGGTC	1896	
DEB.CM.99.CM40	.....	.CACAGCAGATCAGGTATATCTCGACTACCAGGGTACCAGTACAGGGGAGAGAAACAACCCCA.....	GAG-G	2021	
DEB.CM.99.CM5	.....	.CACAGCAGATCAGGTTTATTCAACAACCCAGGTACTGTGCAAGGGGAAAGGAATGCTCCCA.....	GG--G	2015	
SYK.KE.-.KE51	AAGCACA---CA-TGG-CACCCCGCTACA.....	.GCCACAGTCTCTCCAAGACAGTCAAGGACCTCTTGGTCCCGGGAAACCTTC.....	AG---	2046	
SYK.KE.-.SYK173	GGGCCATGGCTC---TGGTCAGCACAGATGAGCCAACAAGCAAGCGAAGGCACAGAACTCTCCCTCGAAGAAACCCCAACAGGGGAAGTGCTCT-G--.....	.....	GA--G	2388	



	Pol TF \ / Pol Protease p10	Gag p6 end \	
	A G A D R Q G T V S F N F P Q V T L W Q R P L V T I K I G G Q L K E A L L D T G A D D		protease
Gag p6	K Q E P I D K E L V S P L T S L R S L F G N D P L S S Q \$		p6
H1B.FR.83.HXB2	AGCAGGAGCCGATAGACAAGGAAGTGTATCCTTTAACTTCCCTCAGGTACTCTTTGGCAACGACCCCTCGTCAACAATAAGATAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGAT		2342
H1B.US.90.WEAU160	-----A-----G-----A-AC-----GT-----AA-----T-----G-----A-----A-----A-----G-----G-----TA-----		1788
H1C.ET.86.ETH2220	CT-C-A--A-GAGCTG--A--CAG-G-AG-C-----AA-----A--T--T-----A-----A--G-----G-----C-----C-----		1734
H1D.CD.84.84ZR085	-A--A--GAC-A-----G-----AA-----G--T-----C-----A-----A-----G-----A-----A-----G-----A-----		1864
H1F1.BE.93.VI850	--G--C--AAAGAC--GGG--ACTGTACCC--C--G-----AA-----T--A-----A-----A-----GA-----G-----		1679
H1G.SE.93.SE6165	CC-C-A--A-GAGATG--AA-G-G-TA-ATC-C-----AA-----G-----A-----G-----A-----G-----TA-----C-----C-----		1739
H1H.CF.90.056	-----A-C-GA-GG-CA-GGAACC--C--G-----A-----G-----T--G-----A-----A-----A-----GT-----G-----		1689
H1J.SE.93.SE7887	-A-----AG-----C-----AA-----G-----T-----GA-----G-----G-----G-----		1656
H1K.CM.96.MP535	--G--ACCAAAGACA--GG--AC--GAGCCC-----AA-----A--T-----AG-----A-----GT-----GA-----		1538
H101.AE.TH.90.CM240	--G--C--AAAGACA--GG--AC--T-CTCC-----GT-----AA-----T-----G-----A-----A-----A-----G-----G-----A-----		1916
H102.AG.NG.-.IBNG	--G--ACC--AG--GACA--GGG--CT--TATCC-----AA-----T--A--T-----G-----GA-----A-----A-----G-----TA-----C-----C-----		1867
H1N.CM.95.YBF30	--G--AG--A--AGAGA--TCTCTGTACCCA-----A-----G-----A-----A-----AAAG--G--GA-----T-----T-----		1936
H10.BE.87.ANT70	-A--A--AGG--GGCCCG--C--G--C-----G--TG--C-----AA--C-----G--CA--AA--A--T--GC--GAG--T-----C--C--TGT--T--GC--G-----		2397
H10.CM.91.MVP5180	GT--A--GG--GATC--GG--A--G--C--A--TG--C-----AA--C-----G--CA--AA--A--T--GC--GAG--T-----T--T--TGT--G--T--C--G-----		2372
CPZ.CD.-.ANT	--G-----ACC--C--CGTA-----G--AA--C-----A-----AA--GA--GGA--G--TCTC--CA--A--AA--TGTC--G-----T-----C-----		1781
CPZ.CM.-.CAM3	-A--C-----CCTGTATCC--G-----C-----AA-----G-----C--GG--G--A-----A-----G-----TA-----GC-----T-----		1753
CPZ.GA.88.GAB2	-AA-----GAGCCC--TC--ATC-----AA-----G-----C--G--GAG--CA--A-----TGT--C-----C-----G-----T-----		1735
CPZ.GA.-.CPZGAB	G-----TCT--TA-----ACC--C-----AA-----G-----A--C--G--AG--AA-----TGT--T--GC-----T-----		2403
CPZ.US.85.CPZUS	--G-----CCCT--AGC-----AA-----G-----C--G--AG--CA-----GA--GTA-----T-----C-----T-----		2402
CPZ.TZ.-.TAN1	CTGCCCA--AC--G--G--T--ACGA--A--AGGCGGGTTC-----A--A-----G-----AA--GA--G--G--AG--CCA--A--G--CTGTC-----T-----T-----		1986
H2A.GW.-.ALI	GAG--ACA--ACACA--AG--TGACAG--GGAC--GCTGCA-----AT--T-----A--A--AG--A-----GCCT--C--T--A--C--G--C--GT-----T--T--C-----C-----G--T--C--C		2720
H2A.DE.-.BEN	GAG--ACA--TCACAGAG--GACAG--GGAC--GCTGCA-----AT--T-----A--A--AG--A-----GCCT--C--C--A--AT--G--CGGTA--TCT--C-----C-----G--T--C--C		2727
H2A.SN.-.ST	-AG--ACA--ATGCAGAG--ACGACAG--GGAC--GCTGCA-----AT--T-----A--A--AG--A-----GC--C--TG--T--A--C--G--C--GTA-----T--T--GC-----C-----G--T--C--C		2172
H2B.GH.86.D205	GAG--AGA--CTACA--GG--TGACAG--GGA--GCTGCA-----AT--T-----AG--A--AG--A-----A--GC--TGT--C--A--T--GTC--GTA--TAT--C-----C-----TT--C--C		2703
H2B.CI.-.EHO	GAG--AGA--CTACA--GG--TGACAG--GGA--GCTGCA-----AT--T-----AG--A--AG--A-----A--GC--CT--T--A--T--TC--GTA--TAT--C-----C-----T--C--C		2698
H2G.CI.-.ABT96	GAA--CA--A--CTACA--GG--TGACAG--RGAH--GCTGCA-----C--T-----AG--A--AG--A-----TGC--T--C--AAA--T--G--C--GTA--TA-----C-----C-----G--T--C--C		2054
H2U.FR.96.12034	GAG--AGA--ATACA--GG--TGAC--G--GGA--GCTGCA-----AT--T-----A-----AA--A-----TGC--T--C-----C--GTAA--TA--C--GA-----C--G--T-----		2209
MAC.US.-.239	GAG--A--TTACA--GG--TGACAG--GGA--GCTGCA-----AT--T-----AGGA--AG--A-----TGCTC--T--T--AA--A--G--CTGTA--TAT--C--G-----G--T-----		2644
Pol TF	R E A L Q G G D R G F G A P Q F S L W R R P V V T A H I E G Q P V E V L L D T G A D D		protease
Gag	R E K P Y K E V G E D L L H L N S L F G R D Q \$		p6
SMM.SL.92.SL92B	G-AGCAG--TTACA--GG--TGAC--G--GGA--GCTGCA-----AT--T-----C--AG--A--AG--A-----G--GC--CTA--A--T--G--CCCTA--G--T--C--G--C-----G--C--C		2075
SMM.US.-.H9	GAG--AGA--TTACA--GG--T--AC--G--GGA--GCTGCA-----AT--T-----AG--A--AG--A-----TGCCCT--C--T--AA--AA--G--CC--TA--TAT-----G--T-----		2117
STM.US.-.STM	G-A--ACA--CTACA--GG--TGACAG--GGA--GCTGCA-----AT--T-----AG--A--AG--A-----GCCCT--T--T--A--T--G--CTGTA--TA-----G--C-----G--T--C--C		2291
SAB.SN.-.SAB1C	-CTC--CCA--AACAGAG--CCCT--CGA--G--G--G--AC--GC-----T-----AG--A--AA--GAAA--TG--GT--C--A--A--GAA--GTCACT-----C-----C-----		2740
TAN.UG.-.TAN1	-ATT--AGAAA--CAGA--GG--A--GGAAGTGGAGGACTA-----T--G--T--C-----AGGA--AA--GAGT--AG--CT--C--CA--C--G--C--GTAC--T--G-----T-----		2532
VER.KE.-.AGM155	CC--CA--C--AAC--ATCCAG--TTGGAACG--GGGA--ATT--TGAA--TC...C-----AG--A--AA--AAGG--CTGT--TA-----A--AACTGCCGTTA--G--AT-----G-----C-----		2552
VER.KE.-.9063	CA--CACCA--A--ATCCAG--CTGGGC--G--AGGA--ATT--TGAA--TC...C-----AG--A--AA--AAAA--G--C--TT--T--AA--AGTC--CTGTAA--TC--G-----G-----C-----		2560
VER.DE.-.AGM3	CA--CA--CAGT--ATCC--G--TTGACAG--GGGA--ATT--TGAA--TC...C-----AG--A--AA--AAAA--G--TT--C--AA--GTC--CC--TCAG--AT-----G--G-----C-----		2055
VER.KE.-.TY01	CA--C--CAAT--ATCCGG--TTGGACCG--GGGA--ATT--TGAA--TC...C-----AG--A--AA--AAAG--G--GT--T--AA--GTC--CC--TTA--G--A--GC-----C-----G-----C-----		2050
COL.CM.-.CGU1	C...CAG--GG--G--GGGA--AG--TC--ATC--C--T--G--A--T-----C--AG--AA--T--T--G--AG--AA--A--AAGTGCC-----G--G-----		2111
DEN.CD.-.CD1	--G--CGA--GAGACA--C--AGGGA--AG--T--ATC--GT--G--A--T-----AG--AA--A--GGCAG--CTG--G--TCA--A--AGGTA--G--TC--TA-----G-----C-----		2214
GRV.ET.-.GRI_677	--ACAAG--AAACAGA--GG--A--GAAGTGGAGGATGT-----T--G--C--C-----AGGA--AA--GAAACG--G--C--TC--AA--AACG--C--GT--C--CT--G-----C-----		2497
GSN.CM.99.CN166	--GGAC--AA--GAG--G--GCAGGAACAC--TC--ATC--C-----GTCA--T-----GG--AA--AA--GAAGGA--G--CTCC--G-----A--G--GGT--TCCATGT-----C-----G--G--C--C		2046
GSN.CM.99.CN71	--G--C--AA--GC--A--AGGAACAC--T--ATC--C-----GTCA--T-----GG--AA--AA--GAGAGA--G--CTCC--G-----A--G--GGT--TCCATGT-----C-----G--G--C--C		2046
DRL.-.FAO	...C--AA--GG--G--CCCT--CGA--G--GG--AC--AC-----T--T-----C--TCA--AACTACA--T--G--G--G--AA-----AAGGTA--G--T--C-----C-----		1909
RCM.GA.-.GAB1	GCA--A--AGAGAG--GAGCAT--T--A--G--AG--C-----AT--G-----A--G--AACTACA--T--GCTC--A--AA--GAA--GT--G--TC-----C-----G--G--C--C		1821
RCM.NG.-.NG411	--G--A--AGT--GGGAGAGCCT--CGA--G--G--G--GT--C-----AT--T-----A--G--TG--AACATGGG--G--A--A-----AAGGT--G--CT-----G-----		1836
MND-2.-.5440	-AG--AAA--AG--GTCCT--CGA--GCAG--AC--AC-----T--T-----A--CA--AACTACAGT--G--G--A--A-----AAGGTA--G--A--TC-----T-----C-----		1903
MND-2.CM.98.CM16	-AA--A--AGA--GG--GTCCAT--CGA--GCAG--AC--AC-----T--T-----A--CA--AACTACAGT--G--G--A--A-----AA--GT--G--CT--C--G--C-----T-----		2343
MND-2.GA.-.M14	--G--A--GAC--AG--GTCCCT--CGA--GCAG--AC--AC-----T--T-----A--CA--AACTACAGT--G--G--A--A-----AAGGTA--C-----C--G--C--T-----		2270
MNE.US.-.MNE027	-AAG--A--TTACA--GG--TGCCAG--GGA--GCTGCA-----AT--T-----AG--A--AG--A-----TGCTC--T--T--A--A--G--CTGCA--TA-----G-----G--T-----		2112
LST.CD.88.447	...ATCCC--C--G--ATGTG--G--CGAGCC--TGCCTGAGTTTC--TCT-----CCAG--AA--A--AA--AC--TCT--G--CAAT-----G--C--GT--TCT-----TC--T--C--T--G-----		1598
LST.CD.88.485	...GTCC--C--G--ATGTG--G--CGAGCC--TGCCTGAGTTTC--TCT-----CCAG--AA--A--AA--AC--TCT--G--CAATA-----G--C--GT--TCT-----TC--T--C--TA--G-----		1598
LST.CD.88.524	-A--AGTTC--C--G--TGTTG--G--CGAGCCCTGCAATTTC--TCT-----CCAG--AA--A--AG--A--G--CT--G--CAAT-----GA--GTAAC--C-----G--C--T-----		1595
LST.KE.-.lh07	--C--AGTCCC--C--G--ATGTG--G--CGAGCCCTGCAATTTC--TCT-----CCAG--AA--A--AA--GT-----TT--G--CAAT-----C--GT--TC--A--G--G--C--T--G--T-----		2682
SUN.GA.98.L14	---A--AG--AG--A--G--GGGAA--G--AGGAGC--A--G--T--G--ATATG-----CCAG--A-----T--A--AG--AGG--CTTC--TAAT-----CTGT--ACT--T-----G--C--G-----		2687
MND-1.GA.-.MNDGB1	--A--AT--AA--GAG--G--AG--GA--G--GGA--TCA-----G--TA--T-----CAAG--A-----AA--A--AAGAGG--CTCAG--G--AT--TGTCACT--TAAG-----C-----T-----		2017
MON.CM.99.L1	C--GG--ACC--A--AG--G--C--AGG--G--G--T--ATC--C--T--AGCA--T-----GG--AA--AA--AAGAT--TG--GCT-----A-----AAGGT--TCCATG-----G-----T--C--C--C		2033
MON.NG.-.NG1	CAGC...C--AG--G--C--AGGAA--G--T--ATC--C--T--AGCA--C-----GG--AA--AA--AG--A--G--G--A--C--CT-----GTCCGT--TCCATGT-----G--C--G--R--G--C--C		629
MUS.CM.01.1085	CCGG--GG--A--C--GG--CAGGAA--AG--TC--ATC-----GTCA--C-----CAGG--AG--AGAC--G--G-----AA-----GG--GGT--TC--ATG--GC--C--C--C-----		2026
DEB.CM.99.CM40	--G--AACAA--GGAGAGG--A--GTACCCGGA-----TGGGG--G--AA--T-----AG--A--AA--GAC--G--C--C--AA--AAG--AA--GTCC--G--TG--A-----G-----		2151
DEB.CM.99.CM5	-CACAA--C--GGAGAGG--A--GTACCCGGA-----TGGGG--G--AA--T-----AG--A--AA--GAC--G--G--T--TG--G--AA--AA--AGGTC--G--TG--A-----C-----T-----		2145
SYK.KE.-.KB51	--A--A--AGA--GA--GG--A--GGACC--T--ATC--C-----TCAA--T-----AG--AA--GA--GGA-----TG--AA--ACCAAGGTAAC--ATGT--GC--G--C-----G-----		2176
SYK.KE.-.SYK173	--AGCAGTGG--AG--GG--ACCAA--AG--TC--ATC--C--T-----C-----T-----AG--AA--GA--AGA--G--G--TG--AA--TG--C--GGT--C--ATGT--G-----G-----		2518

PLV  
complete genomes

Accession	Sequence	Accession	Sequence	Accession	Sequence	Accession	Sequence
H1B.FR.83.HXB2	ACAGTATTAGAAGAAATGAGTTTG.....CCAGGAAGATGGAAACCAAAAATGATAGGGGGAATGGAGGTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGTGGACATA	2460	protease				
Pol p10	T V L E E M S L . . . . . P G R W K P K M I G G I G G F I K V R Q Y D Q I L I E I C G H						
H1A1.UG.85.U455	-----C-----C-A-A-----A-----A-----T-----T-----A-A	1906					
H1B.US.90.WEAU160	-----C-----A-----G-A-----G-----C-----T-----A-A	2459					
H1C.ET.86.ETH2220	-----C-----A-A-----A-----A-----T-----A-----A-----T-----A-A	1852					
H1D.CD.84.84ZR085	-----C-A-A-----AG-----C-----A-----A-----T-----A-A	1982					
H1F1.BE.93.VI850	-----C-A-A-----A-----A-----A-----A-----A-----C-----T-----C	1797					
H1G.SE.93.SB6165	-----A-A-A-----A-----A-----A-----A-----A-----AG--CT--TA--A-A	1857					
H1H.CF.90.056	-----G-A-A-----G-A-----A-----A-----G--AG--GC-----A-A	1807					
H1J.SE.93.SE7887	-----C-AGAC-----C-A-----A-----A-----A-----A--CG--CG--TGAG--A-A	1774					
H1K.CM.96.MP535	-----A-A-----A-----A-----A-----A-----A-----AG--T-----A	1656					
H101.AE.TH.90.CM240	-----T-A-A-----A-----A-----G--AG--A-----T-----A-A	2034					
H102.AG.NG.-.IBNG	-----C-A-A-----A-----A-----A-----A-----A-----T-----T-----A-A	1985					
H1N.CM.95.YBF30	-----A-----GC-ACAA--A-----GAG--A-----A-----A-----G-----A-T--ACAG--C-ACAG--AGA-	2054					
H1O.BE.87.ANT70	-----C-A-CA-C-ACAA-----GA-----A-----T-A-----A-----A--G-A-----A-TG-GACAG--AGAA--AGGG	2515					
H1O.CM.91.MVP5180	-----A-TA-C-ACAA--A-----GA-----C-----T-A-----C-----A-----A--G-----A-CA-TG-GACAG--G-ACAA--A-GG	2490					
CPZ.CD.-.ANT	-----G-----G-G--TCA-----A-----AC-----G-----CA-T-T-----G-----TTCC--CA-----A-CA-AG--C--C--TG-A-ACAGG-	1899					
CPZ.CM.-.CAM3	-----A-----TA-TG-ACAAC-A-----A-----A-----G-----T-A-----A-----CA-C--GCA-----TGAG--AGA-	1871					
CPZ.GA.88.GAB2	-----A-----TA-T-ACAAC-A-----GA--G-A-----G-----T-A-----A-----A-----A-TG--GAA-----AGAG--A-A	1853					
CPZ.GA.-.CPZGAB	-----A-----GAG--ACAA--A-----A-----CTT-----A-----A-----C-----A-----A-T-----A-TG--A-----G-AGAA--GAGA-	2521					
CPZ.US.85.CPZUS	-----C-----CA-C-ACAAA-T-----GA--G-C-----G-----G-----A-----A-----A-----CG-CAAT-----AGAG--AGA-	2520					
CPZ.TZ.-.TAN1	-GT-T-TTGTA-C-C-AA--A-----AAG--CAG--C-----CC--A-----A-----A-----G-ACC--T-TG--CT-A-T-T-CAG-C-----TG-CAATA-AG	2104					
H2A.GW.-.ALI	T-A-A-G--C-G--AGAG--A-----GGGA-C-AC-AT-CC-----AG-----A-G-A-C-A-----TACC-A-G-A-----AG-TG--GAA--A--G-ACTAAATA-A-	2838					
H2A.DE.-.BEN	T-A-A-G--C-G--AGAA--A-----GGG-AC-AT-AC-CT-----AG-----A-G-A-----A-----CACC-A-G-A-----CA-AA-TG--GAA--A--G-ACTAAATA-A-	2845					
H2A.SN.-.ST	T-A-A-G--C-GCG-AGAG--A-----GGGA-C-AT-AT-GT-----G-AG-----A-G-A-C-A-----TACC-A-G-A-----AA-TG--GAA--AG-G-A-TAAATA-A-	2290					
H2B.GH.86.D205	T-A-A-G--C-GG--AGAA--A-----GGTA-C-AT-AC-CC-----AG--A-G-A-----G-C-A-----TACC-A-G-A-----CA-AG-TG--GAA--G-AGTG--A-A	2821					
H2B.CI.-.EHO	T-A-A-G--C-GG--AGAA--A-----GGCA-C-AT-AC-CC-----AG--T-G-A-----A-----A-----TACC-ATG-A-----CA-AA-TG--GAA--G-AGTA--A-A	2816					
H2G.CI.-.ABT96	T-TA-G--C-G--AGAA--A-----GG-A-T-AT-AT-GT-----AG--A-----A-----T-----TACC-A-G-A-----CA-AA-TG--GAA--TA-G-GGTA--A-A	2172					
H2U.FR.96.12034	T-A-A-G--C-GG--T-AA--A-----GA-CC-TAT-AT-CT-----CTG-----A-----A-----G-C-T-----TACC-A-G-A-----T-C-CA-TG--AA--A-G-A-TG--CA-A	2327					
MAC.US.-.239	T-TA-TG-AC-G--AGAG--A-----GGTCC-CAT-AT-CC-----AG--A-----A-----T-----TACT-A-G-A-----CA-AA-TG--GAA--G-T-TA--CA-A-	2762					
Pol p10	S I V T G I E L . . . . . G P H Y T P K I V G G I G G F I N G K E Y K N V E I E V L G K	p10					
SMM.SL.92.SL92B	T-CA-G--C-GG--AGAG--A-----GGTCC-G-AC-C-----G--TG-----G-----A-----C-A-----TACC-A-G-A-----AGAACAG--AAA-----G-T-TA--A-GG	2193					
SMM.US.-.H9	T-A-TG--C-GG--AGAA-----GGTCC-AT-AT-CC-----G-----AG--A-----A-----C-C-T-----CACC-A-G-A-----A-AG-TG--AAA--A--G-A-TA--CA-AG	2235					
STM.US.-.STM	T-A-A-G-G-CG-GG--AGAGC-A-----GGCT-CA-AC-CC-----GG-AG--T-----A-----A-----TACT-AGG-A-T-A-GA-TG--AA-----GG-A-TA--CA-G-	2409					
SAB.SN.-.SAB1C	T-----A-C-G-GT--AGAA--A-----GG--AC-AT-----G-----A-----T-----A-----A-----A-G--T-----T-----A-GCA--CC-CA-TCAGGAGG--A-----AGAA-ATA-A-	2858					
TAN.UG.-.TAN1	--CA-A-A-G--GAAGA-A-CCATTTA . . . . . GA--C-C--CGG--T--G-AG--A--T--A-----GGGA--TC-----A-----CA-T--TT-AG-GA-T--AGA--A-AG	2656					
VER.KE.-.AGM155	--TA-A-A-G--T-CAGA--ACAATTA . . . . . AGG--TC-----G-----AG-----A-----G-----AGGGT-A-C-----A-G-A-----A-TG--GAAG--C--T--GGAA-ACA-G-	2676					
VER.KE.-.9063	--CA-CA-A-----CAGACA-ACAATA . . . . . C-----CCT--G-----G-A-----A-----AGGAC-T-----T-----A-----G-----AG-A-TG--GAAG--A-----AGAA-ATA-A-	2684					
VER.DE.-.AGM3	--CA-TA-A-----GCAGA--ACAATTA . . . . . T-----C-----A-----C-----GGGAC-----T-----A-----G-----AG-G-T-GGGAAG--AG-T-GGAA-ACA-A-	2179					
VER.KE.-.TY01	--CA-A-TA-----ATGA--ACAATTA . . . . . T-----TCC--G-----C-----T-----C-----A-----AGGCC-T-----T-----A-----G-A-----A-CG-C-GGGAAG--A-----AGAA-ATA-A-	2174					
COL.CM.-.CGU1	-----T-TA-----GAAGA-G-TCAATTA . . . . . A-----CGCT--C-G--GT-CAA--TCA-----A-----G-----AGCA--T-----G-----AGAG-T-A-A-TC--A-G--GG-A--A-AG	2235					
DEN.CD.-.CD1	--CA-TA--C-----CAAGA-A-AAAATTTGGGAACAG----- . . . . . TC-----CAG-----A-----A-----T-----TCAGGCA--CCCAGG--GAA--TA-GC-TATG-ACA-AG	2341					
GRV.ET.-.GRI_677	--TA-A-TC-----A-GAC--CACTTTCCCCCA-ATAA-CC-----CGTT-C--GG-AG-----A-----T-----A-----AGGG--TC-T--C-A-G-A-----C-GGG-G--AAG--C--T--GGAG-ATA-A-	2627					
GSN.CM.99.CN166	--CA-CG-CC-----CTCAGCAA-TGAATTA . . . . . GATAC-CC-----C-----G-----A-----G-----A-----AC-----GG--C--A-A--TG--GAAG--ATCT-TAA--AGA-A-	2170					
GSN.CM.99.CN71	--CA-TG-CC-----TCAGCAA-TGAATTA . . . . . GACAC-CC-----C-----C-----G-----G-----A-----G-----A-----AC-----AGG-AC--AGA--TG--GAAG--GATCT-TAA--A-A-A-	2170					
DRL.-.-.FAO	-----A-TC--G-T--GAA--A . . . . . A-----A-----G-----C-G-T-----A-----AAC--T--T--G-----A-----TT-A-TTGTAAG--TA-G--AGCA--GA-A-	2027					
RCM.GA.-.GAB1	-----A-T--G--AGAA--A . . . . . GG-AATGAT--CC-----A-----A-----A-----G-----A-----T-----T-----A-----A-----AA-CTGTGAA--T-----AGC--A-A-	1939					
RCM.NG.-.NG411	-----TA-C-A-TT-AGA--A . . . . . GG-A-AT-----G-----T-----A-----A-----G-----A-----A-----T-----T-----A-----A-----T-----A-----C-----TTGTAAA--ACT--AGCA--GA-G-	1954					
MND-2.-.-.5440	-----A-TA--TT-AGAG--A . . . . . A-----AT-----G-----CG--T-----A-----AGCA--G-----TT-A-TTGTAAG--AGCA--A-A-	2021					
MND-2.CM.98.CM16	-----A-CA--TC-AGA--A . . . . . AA--T-AT-----C-G-T-T-A-----A-----CA-----T-----A-----T-----TTCA-TTGTAAG--AC--AGCA--CA-A-	2461					
MND-2.GA.-.M14	-----A-CA--TG-AGA--A . . . . . GA-----AT-----CC--T-----A-----AGCA--T-GG--G-A-----TTCA-TTGTAAG--AC-G-AGCA--A-A-	2388					
MNE.US.-.MNE027	T-TA-TG--C-G--AGAG--A . . . . . GGTCC-CAT-AT-CC-----AG-----A-----A-----T-----TACT-A-G-A-----CA-AA-TG--GAA--G-T-TA--CA-A-	2230					
LST.CD.88.447	--TA-C-TTC-----ATTGAG-TAGAATA . . . . . GAG--CCT-AT-C--T-G-G-AG-----A-----G-----ACAA--T-G-----A-----G-A-----CA-GG-TG-CT-T-----AGCA--A-A-	1722					
LST.CD.88.485	--CA-C-TTC-----ATTGAG-TAAAATA . . . . . GAG--CCT-AT-C--T-G-G-AG-----A-----G-----ACAA--T-G-----A-----G-A-----CA-GG-TG-CT-T-----AGCA--GA-G-	1722					
LST.CD.88.524	--A-C-TTCT-----ATTGAG-TAGAATA . . . . . GA--CCT-AT-CT--G-G-AG-----T-----A-----G-----ACAA--T-----A-----G-A-----TCAGGG-TG-TT--C-----AGCA--GA-G-	1719					
LST.KE.-.lho7	--A-T-TTCT-----ATTGAG-TAGAATA . . . . . GA--CCC-AT-CT--G-G-AG-----C-----A-----G-----GCAA--T-GG-----A-----G-A-----AGAG-TG-CT-T-----AGCA--A-A-	2806					
SUN.GA.98.L14	--A-G-GTCT--GATTC-G--TCCATA . . . . . GA--TCAC-----G--C-GGG-AG-G--T--T--A-----G-----ACAA--T-G-----A-----G-A-----CAG-TG-CT-TG-----TA-G-ATA-A-	2811					
MND-1.GA.-.MNDGB1	--CA-C-TA-T--GA-A-A-AAAATTA . . . . . AA--AT--C-G-----T-----A-----T-----AACT-A-G-----A-----A-----A-----TG--TATG-----AA--A--GA-GG	2141					
MON.CM.99.L1	--A-TA-C-----TCA-CAG-AAAATTA . . . . . GGG-AG-AT--C--T--G-AG-----A-----G-----AA-G--AG--T-A-----A-----CA-AGCA--AAG-GAT-T-TGAA-ACAGAG	2157					
MON.NG.-.NG1	--CA-CG-CC-----RCAAGACA-AAAATTA . . . . . GG--A-AT--C--C-----G-----A-----C-----A-----AC-G--T-AG--A-TCA-AGG-G--ACGG--TCCT--CAG--TA-A-	753					
MUS.CM.01.1085	--CA-A-CC-G--TGCCA-A-ACACCTA . . . . . GA--CC-----T-----C-----TCAG-----A-----G-----AAAC--ATC--C-A-G-A-----AGGGGAG--AGG-TTGTT--AA--A-A-	2150					
DEB.CM.99.CM40	--CA-TA-TTC-----AAGA-A-AGATTTA . . . . . GAG-C-CCT--CC-----CAG-----T-----A-----C-----A-----T-----C-----A-----TGT--CCCAGG--GAA--T-C--GGC-AAATA-GG	2275					
DEB.CM.99.CM5	--A-TA-TTC--G-AAGA--AGATTTA . . . . . GAT-C-CCT--CC--G-----CAG--G-----T-----A-----G-----A-----T-----A-----TGC--CCCAGG--GAG--TA--GGCA-ACA-GG	2269					
SYK.KE.-.KE51	--CA-A-A-AG--CAAGACA-CCAGTTA . . . . . ATCA-CCT--CC-----AG-----T-----A-----AAAC--A-GG-----A-----A-----C-----CAGAT-TC--AGAAA--CAG	2300					
SYK.KE.-.SYK173	--TA-A-TAG--GAAGA-A-TCAATTA . . . . . ACCA-CC-----TC-----AG-----A-----T-----A-----AAAC--T-C--T--G--A-----AGAA-T--TAGAT-CACTG-TGT-AA--CCT	2642					



		Pol p10 protease end \ / Pol p51 and p66 RT and RT-RNaseH start	
H1B.FR.83.HXB2	AAGCTATA.....GGTACAGTATTAGTAGGACCTACACCTGTCAACATAAATGGAAGAAATCTGTTGACTCAGATTGGTTGCACTTTAAATTTCCCATTTAGCCCTATTGAGACTGTA..	2573	p51
Pol1	K A I . . . . . G T V L V G P T P V N I I G R N L L T Q I G C T L N F P I S P I E T V .		
H1A1.UG.85.U455	-GA-----G-----G-----A-----T-----A-----T-----A-----	2019	
H1B.US.90.WEAU160	-----G-----G-----A-----T-----A-----T-----A-----	2572	
H1C.ET.86.ETH2220	-----C-----C-----CA-----C-----AC-----A-----C-----A-----T-----C-----A-----	1965	
H1D.CD.84.84ZR085	-----C-----C-----T-----C-----C-----A-----T-----A-----	2095	
H1F1.BE.93.VI850	-----G-----G-----A-----T-----A-----T-----A-----	1910	
H1G.SE.93.SE6165	-G-GA-----A-----G-----A-----A-----A-----C-----A-----T-----A-----	1970	
H1H.CF.90.056	-----G-----G-----A-----T-----G-----A-----A-----A-----C-----A-----T-----A-----	1920	
H1J.SE.93.SE7887	-----G-----G-----A-----G-----CA-----C-----T-----A-----T-----A-----	1887	
H1K.CM.96.MP535	-----C-----C-----C-----T-----A-----T-----A-----	1769	
H101.AE.TH.90.CM240	-----C-----A-----T-----C-----A-----T-----A-----	2147	
H102.AG.NG.-.IBNG	-G-C-----C-----A-----C-----A-----T-----A-----G..	2098	
H1N.CM.95.YBF30	---AG-T-----A-----T-----T-----T-----A-----T-----A-----C-----C-----T-----A-----A-----T-----A-----	2167	
H1O.BE.87.ANT70	-G-TACAG-----A-----G-----G-----T-----T-----T-----T-----A-----A-----AGGAT-A-----T-----AC-----C-----C-----T-----A-----C-----A-----CCC-A-----G..	2628	
H10.CM.91.MVP5180	---TACAG-----A-----G-----G-----T-----T-----TC-----G-----CA-----A-----AGGAT-A-----T-----AC-----C-----T-----A-----T-----C-----A-----CCC-A-----G..	2603	
CPZ.CD.-.ANT	C-TAC-----CA-----C-GT-G-----A-AC-A-----A-----T-----A-----T-----G-----T-----ATG-----TTT-A-----A-----T-----A-----T-----A-----TAAAG-----A-----A-----G..	2012	
CPZ.CM.-.CAM3	---A-A-C-----C-----G-----G-----A-----A-----G-----T-----C-----A-----A-----A-----A-----C-----T-----A-----T-----C-----A-----A-----A-----G..	1984	
CPZ.GA.88.GAB2	G-TAGA-----A-C-----C-----G-----A-----T-----A-----T-----T-----A-----G-----A-----C-----AT-A-----A-----T-----C-----A-----ACTC-----G..	1966	
CPZ.GA.-.CPZGAB	---TAG-----G-----A-----A-----T-----C-----A-----T-----A-----AT-G-----T-----C-----GTG-----C-----A-----TT-----A-----A-----C..	2634	
CPZ.US.85.CPZUS	---ACAG-----T-----T-----G-----A-----A-----A-----T-----T-----C-----G-----A-----T-----A-----A-----A-----T-----T-----C-----T-----A-----A-----	2633	
CPZ.TZ.-.TAN1	---TC-G-----CC-----T-----CC-----G-----GAA-----C-----CA-----T-----T-----A-----T-----A-----T-----A-----AG-----AT-----A-----A-----T-----C-----T-----A-----A-----GTA-----	2217	
H2A.GW.-.ALI	G-TA-A-----CC-----CA-----A-----GAC-----TGAC-----C-----AA-----T-----TT-----C-----CA-----T-----AGCCT-A-----CATGT-A-----CC-----A-----AG-----GC-----AAG-----A-----C-----AA-----	2951	
H2A.DE.-.BEN	G-TA-G-----CC-----CA-----A-----GAC-----GA-----C-----AA-----CT-----C-----A-----TC-----AGCCT-A-----CATGT-A-----A-----AG-----GC-----AAG-----A-----C-----AA-----	2958	
H2A.SN.-.ST	G-TA-G-----CC-----CA-----A-----GAC-----TGA-----C-----AA-----TT-----C-----CA-----TC-----AGCCT-A-----CATGT-A-----C-----A-----AG-----CGC-----AAG-----A-----AC-----AA-----	2403	
H2B.GH.86.D205	G-TA-GG-----CA-----TA-----A-----GAC-----GA-----C-----AA-----TT-----C-----A-----T-----A-----ACCT-G-----CATG-----C-----AG-----GGCAAAGG-----A-----AC-----A-----	2934	
H2B.CI.-.EHO	G-TA-G-----CA-----A-----GAC-----GGAC-----C-----AA-----TT-----C-----A-----T-----A-----AGCT-A-----CATG-----C-----C-----AG-----AGCAAAGG-----A-----AC-----A-----	2929	
H2G.CI.-.ABT96	G-ATACAG-----C-----T-----TA-----GAC-----GGA-----C-----A-----A-----T-----CT-----G-----CA-----T-----AGT-A-----C-----A-----CATGT-----C-----A-----GCAAAGG-----A-----C-----A-----	2285	
H2U.FR.96.12034	C-AT-G-----CA-----A-----GAC-----GAC-----C-----AA-----T-----TT-----T-----G-----T-----AG-----AC-----G-----CATGT-----C-----C-----A-----AG-----AAAAGAA-----CACTA-----C..	2440	
MAC.US.-.239	GGAT-A-----G-----A-----CA-----GAC-----GGAC-----C-----GA-----T-----TT-----T-----T-----C-----A-----AGCTC-G-----GATGT-----C-----AGCTAAAG-----A-----C-----	2875	
Pol1	R I K . . . . . G T I M G G D T P I N I F G R N L L T G L G M S L N N F P I A K V E P V .		
SMM.SL.92.SL92B	T-ATA-G-----AA-----C-----A-----GAC-----GA-----C-----GA-----A-----T-----TT-----C-----A-----TC-----AG-----AGCAT-----A-----GGT-----T-----AC-----AG-----GGCAAAGG-----A-----ATA-----AC..	2306	
SMM.US.-.H9	T-AT-AG-----A-----A-----TA-----GAC-----GA-----C-----AA-----T-----TT-----C-----C-----A-----AGCT-----G-----CATGT-----C-----C-----AGCTAAGG-----G-----RC-----A-----	2348	
---ATA-A-----	A-----TA-----TA-----GACT-----GAC-----C-----CA-----T-----TT-----C-----G-----C-----A-----AGCTT-G-----GATGT-A-----AGCTAAGG-----A-----AC-----A-----	2522	
SAB.SN.-.SAB1C	CTTG-A-----C-----CA-----T-----G-----GAA-----A-----T-----T-----T-----A-----G-----C-----AG-----A-----AT-----A-----AGTT-----C-----GA-----ACAA-----AGAG-----A-----AC-----A-----	2971	
TAN.UG.-.TAN1	---A---G-----AT-----C-----A-----T-----A-----AA-----A-----T-----TG-----A-----A-----TC-----AG-----A-----CTT-----A-----AGCA-----GA-----TAA-----GGGACAGCTAAA-----GAGA-----A-----TACCC..	2769	
VER.KE.-.AGM155	T-TTA-G-----A-----CC-----CA-----G-----A-----T-----CA-----T-----C-----A-----CT-----T-----AG-----C-----GCA-----AGC-----AA-----GTGA-----GGGCAATTGT-----GCAGACA-----TACC..	2789	
VER.KE.-.9063	TTTTG-G-----C-----TC-----TT-----G-----A-----C-----A-----T-----T-----T-----AC-----AG-----A-----CAGCA-----AGCT-----GAC-----GTGA-----GGGACAATTGT-----AGAAC-----A-----T-----CCC..	2797	
VER.DE.-.AGM3	TTTTG-G-----G-----CA-----GA-----AGC-----T-----CA-----A-----A-----AG-----A-----GCA-----AGC-----AA-----GTAA-----GGGTCAACTGT-----AGAAC-----A-----T-----CCC..	2292	
VER.KE.-.TY01	TTTTG-G-----A-----A-----GT-----G-----A-----T-----CA-----T-----T-----A-----T-----T-----C-----G-----C-----GCA-----TGCC-----CGG-----GTAA-----GGGACAATTAT-----AGAAA-----A-----T-----CCT..	2287	
COL.CM.-.CGU1	---TAC-G-----AGATA-----C-----A-----AC-----C-----CA-----A-----T-----G-----A-----T-----A-----A-----AAG-----AGGGAAA-----TAGTAG-----GC-----CAATTATCAGATAGG-----CCCAGTCACC..	2348	
DEN.CD.-.CD1	T---GG-----AC-----CA-----A-----CT-----A-----CTATC-----T-----G-----A-----T-----A-----AGTA-----AC-----CA-----GG-----AG-----ACAG-----AA-----GCG-----C-----CA..	2454	
GRV.ET.-.GRI_677	TCATC-CC-----CT-----A-----TC-----A-----AG-----AA-----T-----T-----A-----A-----T-----AG-----GCA-----CATG-----AA-----GT-----A-----GGGAG-----CTAT-----G-----C-----T-----AG..	2740	
GSN.CM.99.CN166	G-ATC-A-----C-----T-----C-----G-----T-----C-----A-----A-----T-----T-----A-----T-----G-----C-----AG-----A-----T-----A-----GGTA-----CA-----GGTGCAGCAGAAAG-----A-----AC-----CA..	2283	
GSN.CM.99.CN71	G-ATC-A-----CC-----T-----C-----T-----G-----C-----A-----G-----T-----T-----G-----T-----G-----A-----GCA-----G-----GCA-----G-----GCA-----G-----GT-----CA-----GGTGCAGCAGAAAC-----A-----C-----C..	2283	
DRL.-.-.FAO	TTA-CAT-----C-----T-----G-----G-----C-----C-----A-----T-----T-----A-----T-----G-----AC-----CTG-----A-----AT-----A-----AGCT-----C-----T-----TG-----T-----AAG-----A-----AC-----A-----	2140	
RCM.GA.-.GAB1	GGA-CAT-----C-----CAT-----GC-----G-----A-----A-----G-----T-----T-----A-----G-----G-----T-----A-----AGA-----AT-----A-----AGC-----AC-----A-----A-----AA-----A-----A-----A-----	2052	
RCM.NG.-.NG411	CTA-ACAT-----C-----CAT-----G-----A-----A-----A-----T-----T-----A-----G-----G-----GCA-----AT-----A-----AGC-----T-----C-----A-----A-----TAAAGCA-----A-----AA..	2067	
MND-2.-.-.5440	TTA-CAT-----C-----G-----TC-----G-----G-----C-----C-----A-----T-----T-----A-----T-----G-----AC-----T-----AGA-----T-----A-----A-----T-----T-----T-----TAAG-----A-----A-----A-----	2134	
MND-2.CM.98.CM16	CTA-ACAC-----C-----T-----CC-----G-----C-----C-----A-----A-----T-----TG-----A-----T-----G-----C-----A-----AGA-----AT-----A-----T-----AC-----AG-----TAAAG-----A-----A-----A-----	2574	
MND-2.GA.-.M14	GTA-CAT-----C-----T-----GC-----G-----C-----C-----A-----T-----T-----A-----T-----G-----TC-----T-----AGA-----T-----A-----A-----T-----C-----G-----T-----TAAG-----A-----A-----	2501	
MNE.US.-.MNE027	GGAT-A-----G-----A-----CA-----GAC-----GGAC-----C-----GA-----T-----TT-----T-----T-----C-----A-----AGCTC-----G-----AATGT-----C-----C-----AGCTAAGG-----A-----C-----	2343	
LST.CD.88.447	C-A-AG-----G-----T-----G-----A-----AG-----TG-----A-----A-----T-----A-----A-----TG-----G-----AGCA-----GA-----GTAC-----AT-----ACAATTAAG-----GACA-----A-----TACCC..	1835	
LST.CD.88.485	C-A-AG-----G-----C-----T-----A-----AG-----TG-----A-----A-----T-----A-----A-----AGTG-----G-----AGCA-----AA-----GTAC-----AG-----ACAATTAAG-----GACA-----A-----TACCT..	1835	
LST.CD.88.524	CCA-AG-----G-----T-----GC-----A-----C-----A-----AG-----T-----G-----A-----CA-----T-----A-----AC-----CTC-----A-----GGCT-----GA-----GTG-----AG-----ACAATTAAG-----GAGA-----A-----TACCT..	1832	
LST.KE.-.lho7	CTA-AG-----G-----C-----T-----C-----A-----AG-----T-----TG-----A-----G-----G-----A-----T-----A-----G-----GCA-----G-----GCG-----AA-----G-----T-----AG-----TCAGTTAAG-----GA-----A-----T-----CCT..	2919	
SUN.GA.98.L14	C-----C-----A-----G-----G-----G-----G-----A-----A-----AG-----T-----C-----A-----G-----A-----TC-----TT-----AGCCT-----A-----GG-----A-----GA-----GT-----AG-----C-----CTTAG-----GAAA-----TCCC..	2924	
MND-1.GA.-.MNDGB1	G-A-AT-T-----G-----GA-----T-----AA-----AG-----T-----A-----G-----CA-----AA-----GAAA-----AT-----A-----AG-----A-----AA-----TA-----GG-----ACAATTGT-----GA-----A-----A-----TACC..	2254	
MON.CM.99.L1	-G-T-TGG-----CA-----T-----T-----A-----T-----A-----A-----T-----TT-----A-----G-----G-----C-----AG-----A-----G-----GGTA-----CC-----C-----CA-----GG-----TGC-----G-----GGGAGACCTCCAACCCCC	2272	
MON.NG.-.NG1	C-TG-G-----CW-----T-----GC-----G-----R-----C-----AA-----T-----T-----TT-----G-----G-----CG-----CC-----T-----AAG-----G-----GATA-----CC-----G-----CA-----GY-----CAGG-----AATCGCCC-----AT-----CCC..	866	
MUS.CM.01.1085	CTAT-GG-----CA-----G-----A-----A-----A-----T-----G-----G-----CTGT-----A-----GCA-----AT-----CATT-----C-----CA-----GATACAGG-----AAAAT-----A-----AC-----AA..	2263	
DEB.CM.99.CM40	T---AG-----TCAGATC-----GA-----C-----AA-----T-----T-----TC-----A-----TAT-----AG-----AA-----A-----G-----GGTAT-----A-----AG-----ATTAGA-----TA-----ACCC-----A-----	2388	
DEB.CM.99.CM5	T---G-T-----TCAGAT-G-----TGA-----C-----AA-----T-----T-----G-----G-----TAT-----AG-----AA-----A-----G-----GGTAT-----CA-----AG-----ATTGGACT-----ACAC-----A-----	2382	
SYK.KE.-.KE51	-T-GG-G-ATAAAGGAAGTAGAG-----GTC-TC-----A-----A-----T-----TT-----A-----CA-----A-----GCA-----AT-----G-----AGCA-----AA-----GTAA-----GATAGCACAAAAT-----A-----AC-----AA-----G..	2428	
SYK.KE.-.SYK173	C-GA-A-AGAAAGCAAGTAGAG-----G-TC-T-----G-----T-----G-----A-----A-----T-----T-----A-----G-----A-----AC-----T-----CA-----AT-----G-----AGT-----AA-----GTAA-----GGTTCAA-----CAG-----AT-----A-----C-----A-----C..	2770	



PLV  
complete genomes

H1B.FR.83.HXB2	.....CCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAGTTAAACAATGGCCATTGACAGAGAAGAAAAATAAAGCATTAGTAGAAATTTGTACAGAGATGGAAAAGGAAGGAAAATTTCAAAA	2696
Po1	...P V K L K P G M D G P K V K Q W P L T E E K I K A L V E I C T E M E K E G K I S K	p51
H1A1.UG.85.U455	.....A--A-----G-----G-----AC-----AT-----A-----	2142
H1B.US.90.WEAU160	.....G-----G-----A-----A-----A-----	2695
H1C.ET.86.ETH2220	.....C-----G-----AC-C-----GA--A-----GC-----A-----G-----	2088
H1D.CD.84.84ZR085	.....G-----G-----AC-----T-----A-----A-----G-----	2218
H1F1.BE.93.VI850	.....G-----G-----G-----AC-----A--CT--A-----A-----A-----	2033
H1G.SE.93.SE6165	.....GG-----G-----AC-----A--A-----G-----A-----A-----	2093
H1H.CF.90.056	.....G-----ACG-----A-----A-----A-----C--G-----	2043
H1J.SE.93.SE7887	.....A-----AC-C-----G--A-----G-----A-----A-----G-----	2010
H1K.CM.96.MP535	.....G-----A-----A-----AC-----A-----A-----A-----C-----	1892
H101.AE.TH.90.CM240	.....C-----A-----G-----G-----AC-----A-----A-----G-----A-----C-----	2270
H102.AG.NG.-.IBNG	.....G-----G-----A-----A-----AC--C-----A-----A-----	2221
H1N.CM.95.YBF30	.....A-----G--A-----T-----AC-----G--G-----AG-----A-----A-----T--G-----	2290
H1O.BE.87.ANT70	.....C-----A-----A-----CC-AT-TA-----G--C--GAC-C--A--CAG-A-----C--A-----A-----G-----	2751
H1O.CM.91.MVP5180	.....GC--A-----A-----A-----CC-AT-TAG--G-----G-----C--ACT-C--A--CA--A-----C-----A-----C--G-----	2726
CPZ.CD.-.ANT	.....-C--T-----AGA-----A--G--A-----GC-CT--A--G-----G--G--CC--AA-----GATA--T--A--GCA--AAT--G-----T--G-----	2135
CPZ.CM.-.CAM3	.....T--A-----C-----T-----AC-----CAG--A-----A-----A-----G-----G-----	2107
CPZ.GA.88.GAB2	.....CA--A-----GG--A--G-----T--A-----G--G-----AC-----GG--A-----A-----A-----G-----	2089
CPZ.GA.-.CPZGAB	.....C-----G--A-----T-----G--A--G-----C--AT--C-----T-----C--AC-----CA--A-----A-----A--G-----	2757
CPZ.US.85.CPZUS	.....G-----A-----T--G--A--G-----T--G--A--G-----T--G--AC-----CCA--A-----A-----A--G--A--G-----	2756
CPZ.TZ.-.TAN1	.....AA--C-----AGA-----G-----G--G--G-----CC-CT-CA-G--G-----TG--G-----AC-----A-----AGACAT-----A-----TGC-----	2340
H2A.GW.-.ALI	.....GAG--G-----AA--C--G-----T--A--G-----C--A--A-----G--G--C--AA-----C--GA--A--A--C-----GA-----CC--T--AGAGG--G	3074
H2A.DE.-.BEN	.....AA--C--G-----G--AA--A-----GGC--G-----CC--A--A-----G-----G--C--AA--G--C--GA--A--A-----A--G--CC--GC--AGA--G--G	3081
H2A.SN.-.ST	.....AA--A-----TGC--G-----A-----A-----C--G--G-----C--A--A-----G--G--C--AA--G--C--GA--A--A-----GA--G--CC--GC--AGAGG--G	2526
H2B.GH.86.D205	.....AA--TG--G-----A--T--AA-----G-----GA--C--G-----TC--AT--CAGG-----G--CT--CC--CAA-----C--GA--A--A-----G--AC--T--AGA--G--G	3057
H2B.CI.-.EHO	.....AA--CC--G-----T--A--AA-----G-----A--C--G-----CC--AT--CA-----G-----CT--CC--CAA-----C--GA--A--A-----A--G--AC--GT--AGA--G--G	3052
H2G.CI.-.ABT96	.....AA-----GGC--A--T--C-----G-----A--C--GG-----CC--AT--TA--G-----C-----C--CAGG-----GACA--A-----C-----C--T--GGA--G--G	2408
H2U.FR.96.12034	.....ATT--GGGG--C-----A--C--A-----A-----T--A--G--G-----T--AT--TA--G-----T--G-----AA--G--A-----GAGA--A-----G--AC--T--GGA--G--G	2563
MAC.US.-.239	.....AA--CGCC-----A-----A-----T--G--G--G-----AT--A-----G--GTT-----AG-----C--GA--A-----T--TC--GT--GGAGG--G	2998
Po1	...K V G 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 E E	p51
SMM.SL.92.SL92B	.....AA--G--G--GAG-----G--G--T--A-----C--AT--TA--G-----G--C-----AC--G--G--C--CGAGA--A-----G--AC--T--GGA--G--G	2429
SMM.US.-.H9	.....AA-----C--C--A-----AA-----A-----T--A--G-----G--GC--AT--GA-----G--TT-----AG-----C--GA--A--A-----G--A--T--CC--GT--AGAGGR--	2471
STM.US.-.STM	.....AAG--G--C-----A-----AA-----A-----A--A-----AT--A-----G--G--G--C--AA--G--C--GA--A--A-----G-----T--C--GC--AGA--G--G	2645
SAB.SN.-.SAB1C	.....AA--CC--C-----CAA-----G-----GGA--A--GG-----T--T--A-----G--G--C--AAG--CC-----GA--CT--A-----AC-----AC--TT--GGA--G--G	3094
TAN.UG.-.TAN1	.ATAACCAAG--C--T--G--A--T--G--TCA--A--G--T--A-----CC-AT-----CC-GAC-----AAC-CAG-CTC-A-----A-----A-----T--G	2898
VER.KE.-.AGM155	.ATCACC--G--CGC-----GA-----GGCCAGA--A--CG--T--G--G-----C--CT--TA-----T--CC--GCA-----C--A--ACAT--A--GG--A-----A--T--AAGC--GG	2918
VER.KE.-.9063	.ATCACC-----C--C--G--AGAT-----GCAAGG--A--C--TA--A-----TC--AT--A--G--G-----G--C--AC--G--A--GA--C--T--A-----A-----CAGC--G	2926
VER.DE.-.AGM3	.ATTACC--T--G-----AGA-----GGCTAGA--A--TTTCT--A-----CC-CT--CA-----C--CAG-----A--GACC--AT--A--G--A-----A-----AGC--G	2421
VER.KE.-.TY01	.GTCACA--T--C-----G--GA-----GGCTCGG--A--CTGT--A--G-----TC--CT--TA-----G--G--TG-----T--CAG-----A--T--CC--AT--A--GC-----A-----CAGT--G	2416
COL.CM.-.CGU1	.....AAG--G--CG-----AGAG--C-----G-----A--G--G-----C--A--GCA-----G--GT--CA--A--G--A--GA--AGAT-----GCA--G--A-----AGA--G	2471
DEN.CD.-.CD1	.GTCACC-----A-----T--T-----T--G--A--A-----T--A--TAC--G--G--T-----T--CA--C--G--AGTG--C-----G--A--GGT--GGA--G	2583
GRV.ET.-.GRI_677	.GAAACAAA--C--C-----AGA--G--AA-----A--T-----T--G-----C--AT--AG-----TG-----T--AC-----A--C--A--C--A-----G-----G--A--T--A--T--GG	2869
GSN.CM.99.CN166	.....GA--GC--CC-----G-----AA--G--A-----GA--C-----G-----CC--T--TAGG-----G--G--G--C--ACCC-----AACACA-----CTT--GATT--TC--T--GGA--G	2406
GSN.CM.99.CN71	.....GA--GC--CC-----AT-----AA--A--A-----GA-----C-----T--TA--G-----G--G--G--C--ACCC-----AACACA--A--CTC--TT--CC--T--GGA--G	2406
DRL.-.-.FAO	.....AA-----C--C--A-----A-----A-----GA--A-----CC-CT-CA-G--G-----TG-----C--GAA-----A--GAGC--A-----G-----TC-----AAGT--G	2263
RCM.GA.-.GAB1	.....AAG--G-----T--T-----CAA-----A--CG--A-----T--A-----G-----T--AC-----C--AT--CA-----G-----A-----A-----T--G	2175
RCM.NG.-.NG411	.....AAG--G-----CAA-----G--T--G-----A-----T-----GTA-----G-----TG-----T--AC-----G-----AT--CC-----G-----A-----T--G	2190
MND-2.-.-.5440	.....AA--C--C-----A-----CA-----A--G--G--A--A-----G-----C--T--TA-----G--CCT--C--CA-----A--C--ATC--A-----A-----G--C--T--G	2257
MND-2.CM.98.CM16	.....AAG--C--C-----A-----CT--C--A--T-----A--A-----T--TA--G-----G--TTT--C--CA-----A--GTC--A-----G-----CC--G--C--T--G	2697
MND-2.GA.-.M14	.....AA-----C--C-----A-----A-----G-----C--T--TA-----G--TTT--C--CAGAC--A--GAT--A-----C-----C--T--G	2624
MNE.US.-.MNE027	.....AA--C--CC-----AA-----A-----T--G--GG--G-----AT--A-----G-----GTT-----AG-----C--GA--A-----T--TC--GT--GGAGG--G	2466
LST.CD.88.447	.ATCACAAAG--C-----G-----CTGT-----T--C--G--A-----G-----T--T--TA-----G--G--GCC--CA--C-----GATAGAT--A-----A-----A-----	1964
LST.CD.88.485	.ATCACAAAG--C--G--G-----CTGT-----T--C--G--A-----G-----T--T--A-----G--G--GCC--CA--C-----GATAGAT--A-----A-----A-----	1964
LST.CD.88.524	.ATAACAAAG--GC--T--G-----GTGT-----G-----G--A-----G-----C--AT--CA-----G--G--CA--C-----GATAGAT--A-----A-----A--CCC--	1961
LST.KE.-.1ho7	.ATTACAAAG--TCT-----A--T--GTGT-----A-----G--A--G--G-----TC--AT--A-----G-----G--GCC--CA--C--CT-----GATAG--T--A-----A-----A--TCC--	3048
SUN.GA.98.L14	.ATTACAAA--T--G-----AGA--GGCC--C-----A-----G-----CC--T--AGG--G-----G--GCC--CAGA-----GATAGAT-----GCA-----A--G--AGC--GG	3053
MND-1.GA.-.MNDGB1	.ATAACAAA--T--A--GTA-----G--A-----A--C--G--A-----T--A--GTA-----G-----GTT--GTC--TCAGA--A-----GATAGAT--A--GG-----A-----AGT--GG	2383
MON.CM.99.L1	ATCTTTAGAG--C--CT-----AG-----AA--G--A-----GT--G-----G-----AT--CAG-----G--G--TC--TAC--C-----AC--CA-----CTT--ATT--AC--T--GGA--CC--	2402
MON.NG.-.NG1	.TCCTTGA-----G--CCC--T--A-----AA-----G-----T--G-----CC-CT--TA-----G--G--TC--CACCACC-----RCCCGG-----AT--G--ATT--CC--T--GGA--G--	995
MUS.CM.01.1085	.....AA--G--CCC-----AGA--GCAA--A--G-----TG--C--G-----T-----CC--G-----G--C--G--AG--C--AACCCAGC-----CT--A--T--CC-----AGAGG--	2386
DEB.CM.99.CM40	.....AA--T-----AGAG--GGCA-----A-----GA-----C--TT--AGG-----G--G--C--AA-----T--C--TA--AGCCC-----AT--G--C-----AC--G--AGTCCCG	2505
DEB.CM.99.CM5	.....AAGT--G-----GAG--GCT-----GGA-----G-----CC--TT--AG-----G-----C--AA-----G--CATA--GGCCT-----T--G--A--C-----AC--G--AGTCCCG	2499
SYK.KE.-.KE51	.....AA--CC--C-----G--A--GCTA--T-----T--A-----G-----TC-----AG-----C--TG--G--C--GCG--C--AGT--T-----T--CTTG--AA--G--C--T--GGA--G	2551
SYK.KE.-.SYK173	.....AA-----GTC--G--A-----AC--AA--ATTG--CCGGC--A-----AT--G--T--G-----T--G--G--CC--CAAG--C--AGTAGA--C-----CTC-----C--AC--T--GGA--G	2893





H1B.FR.83.HXB2	ATTGGGCCTGAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAGACAGTACTAAATGGAGAAAATAGTAGATTTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAG	2826
Pol1	I G P E N P Y N T P V F A I K K K K D S T K W R K L V D F R E L N K R T Q D F W E V Q L	p51
H1A1.UG.85.U455	-----T-----G-----C-----G-----C---A-----G-----	2272
H1B.US.90.WEAU160	-----T-----	2825
H1C.ET.86.ETH2220	-----C---T-----A-G-G-----G-----G---C---A-----T-----	2218
H1D.CD.84.84ZR085	-----A-----G-----	2348
H1F1.BE.93.VI850	-----G-----G-----A-----T-----G-----	2163
H1G.SE.93.SB6165	-----C---A-----G-----G-----G-----G---C---A-----G---C-----	2223
H1H.CF.90.056	-A-----G-----GC---A-----A-G-G-T-----G-----C---A-----G-----	2173
H1J.SE.93.SE7887	G-----T---C-----G-----T-----C---A-----G-----	2140
H1K.CM.96.MP535	-----G-----T---C-----G-----T-----C---A-----G-----	2022
H1O1.AE.TH.90.CM240	-----T-----G---C---C-----G-----G---C---A-----G---T-----	2400
H1O2.AG.NG.-.IBNG	-----A-----G---T-----	2351
H1N.CM.95.YBF30	-----A-----G---T---C---A-T-----T---A-G---T---C-----G---T-A---A-G-C---T---T-----G---C---G---	2420
H1O.BE.87.ANT70	-----A-----T---T---A-TA-C---T---A-G---TG-----G-----T---G---T-A---A-G---G---A-GC---	2881
H1O.CM.91.MVP5180	-----A-----T---T---A-TA-T---T---A-G---T---C---G-----G---C---T-A---A---A---T-----G---G---	2856
CPZ.CD.-.ANT	-----A---A-T---C-----A---A-T---A---A-G---C-T-A---GC---T-----T-A---A---A---T---T---GA-A---	2265
CPZ.CM.-.CAM3	-----A---C-----T---A---A-T---T---A-G---T---A---C---G---T---G---T-A---A---A---G---T---C---G---	2237
CPZ.GA.88.GAB2	-----A---T-----T---CA---A---G---G-----A---G---T---T---G---A---A---T---T---A---G---	2219
CPZ.GA.-.CPZGAB	-----A---A-----T---TA-T---A---C---A-G-----C-----C---T---A---A---T---T---G---G---	2887
CPZ.US.85.CPZUS	G-A---A-----T---A---A-T---T---C---A-G---T---C-----C---T---T-A---A---G---A-----G---	2886
CPZ.TZ.-.TAN1	G---A---A---C---T---C---A---A-T-----T---G---T---CCT-----G---T---A---G---A---T---T---GT-A---GC---	2470
H2A.GW.-.ALI	GCACCT--AACT--C---T---T---C---CAC-----A-----G---G---AA-AC-----TGC-A-----T---T-A---C---GT-----T---ACA---GA---G---	3204
H2A.DE.-.BEN	GCACCT--AACT--T---T---C---CAC---A---T-----G---AG-AC-----G---TGC-GA-----T---A---A---GTG-----T---ACA---A---GC---	3211
H2A.SN.-.ST	GCACCT--AACT--T---T---C---CAC---A---C-----G---AA-AC-----TGC-A-----T---A---C---GT-----T---ACA---A---C---G---	2656
H2B.GH.86.D205	GCACCC--ACT-----C---A---CACC-----G---T-AG-AC-----TGC-A---C---T---GT-A---C---GT---C---G---ACA---CA---C---G---	3186
H2B.CI.-.EHO	GCGCCT--ACT-----T---G---CACC---C-----A---G---AA-AC-----G---TGC-A-----A---C---GT---C---A---TACA---G---C---GC---G---	3182
H2G.CI.-.ABT96	GCACCC--CACC---G---T---C---A---CACC---A---T---C---G---G---AA-AC-----TG---A---C---C---G---GT---C---G---TACT---C---G---G---	2538
H2U.FR.96.12034	GC-CCC--CACT-----C---C---CAC-----G---G---AA-AC-----TG---A---C---T---GT-A---C---GAGT-----T---TA---T---C---G---	2693
MAC.US.-.239	GC-CCC--GACC-----C---C---CAC-----T-----G---T-AG-AC-----TGC-GA-----T---G---A---G---GTC-----G---TAC-----C---	3128
Pol1	A P P T N P Y N T P G F A I K K K K D K N K W R M L I D F R E L N R V T Q D F T E V Q L	p51
SMM.SL.92.SL92B	GCCCTT--AC-----T---T---A---GACC---T---T---GA-T-AGGAC---G---TGC-CA-----T---A---C---A-TG---A---A---ACA---A---GC---G---	2559
SMM.US.-.H9	GC-CCT--AACT-----T---C---C---CAC---T---R-GR---G-A-A-----G---TGC-A-----T---T-A---C---GTC-----T---TACA---A---G---	2601
STM.US.-.STM	GC-CCT--AAC-----T---C---CACT---T-----G---AA-AC---G---G---TGC-GA-----A---AGTG---A---T---TACA---A---GC---	2775
SAB.SN.-.SAB1C	-----A---A-----T---C---C---C---A---G---G---T-AA---C---G---T---C---CA---G---C---GC-GT-A---A---T---TCA---A---C---G---	3224
TAN.UG.-.TAN1	G---AGGA---G---T-----C---G---TGT---CCG-----AA---AC---G---G---TG---A-----T---CT---C---AGC---A---G---TT---G---A---G---	3028
VER.KE.-.AGM155	G-A---GGA---C---G-----A-----CTGT-----G---AAT-AC---G---TGC-G-----G---C---C---AGCT---A-----TT---C-----	3048
VER.KE.-.9063	-----A---AGGA---G---T---T---C---A---C---TGT---C---GA---G---G---AA---C---G---TG-----C---T---A---AGT---A-----TT---G---G---	3056
VER.DE.-.AGM3	-----A---AGGA---G---G---C---G---TG-----AGT-AC-----TG-----T---G---A---C---AGC---A---T---T---TC---A---G---	2551
VER.KE.-.TY01	G-A---AGGA-----G---C---A---TG-----G---G---AAT-CC-G---G---TGC-----C---T---GT-A---G---C---T---TT---G---	2546
COL.CM.-.CGU1	GCA-AATTA-GG-----T---A---A---TG---T---G---T-AA-A-G-----G---A---G---A---C-----A-----AT-----C-----	2601
DEN.CD.-.CD1	GCGCCA--CACC--C---T-----A---T---G---TA--TCCA--G-----AA--AC---AT--G-TG---A-----GA--T-A---G---GCC--C---G-----A-----	2713
GRV.ET.-.GRI_677	-----A---AGGA-----T---T---A---G-----G---T---AA---AC-----TGC-T-----G---A---C---AGCT-----T---TT---G---GC---G---	2999
GSN.CM.99.CN166	-----A---A---A-----CT-A---C-----A---G---G---AAT-AC-----TGC-A-----T---A---T-A---G---GCT---G-----GGCT---G---C---G---G---	2536
GSN.CM.99.CN71	-----A---A---A-----CT-G---C-----A---G---AAT-AC-----TG---A-----GA---T-G---G---GCT---G-----GGCT---G---C---G---G---	2536
DRL.-.-.FAO	-----A---T---A-----C---A---G---G---TG---T---G---TG---C-----G---C---T-A-----GC---A---G-----TT---G---G---G---G---	2393
RCM.GA.-.GAB1	-----A---A---A---C---C---C---A---A---CTGT---T---A---G---TCC---A-----G---T-----T---A-----A---G---T---TT---G---G---GC---G---	2305
RCM.NG.-.NG411	-----A---A---A---T-----T-----A---G---TGT---C---A---G---TTCC---A-----GC-----T---G---A---A---T---TT---G---A---GC---G---	2320
MND-2.-.-.5440	G-A-T-T-A-----C---C-----CTGT-----A---G---TG-A-C-GC-----T---C---G---C---AGTC---G---T---TT---G---G---G---G---	2387
MND-2.CM.98.CM16	-----A---T---A-----T---C---A---G---G---TG---C---A---G---TG-A-C-GC-----G-----T---C---GT-A---C---AGT---G-----TT---G---G---G---G---	2827
MND-2.GA.-.M14	-----A---T---A-----T---A---G---G---TG---A---G---TG-A-C-GC-----T---GC---T-A---GAGTG-----G---T---TT---G---G---GC---	2754
MNE.US.-.MNE027	GC-CCC--GACC-----C---C---CAC---T-----G---AG-AC-----TGC-GA-----T---G---A---G---GTC-----G---TACA---C-----	2596
LST.CD.88.447	GCA-ATG-A-GG-C---T---C---A---A-----G---AA-A-G-G-----A-T---T---C---A---GCTTT---A-----TCAT---A-----	2094
LST.CD.88.485	GCA-ATG-A-G---C---T---C---A---A---C-----G---G---AA-A-G-G-----C-GA-T---T---C---A---GCTTT---A-----C---TCAT---A-----	2094
LST.CD.88.524	G---AT---A-G-----T---C---C---A---C-----T-----G---AA-A-G-----G-GA-----A-T-A---G---TT---A-----TCAT---GC-A---G---	2091
LST.KE.-.lho7	G-G-AT---A-GG-C-----A-----T---C---G---AA-A-G-----G---A-T-----A-G---C---G-ATT-----TCAT---GC-A-----	3178
SUN.GA.98.L14	GCA-AC--A-GG-C---T---T---C---A---A---T-----G---AA-A-G-----G---A---T---T---A---GGT-TG-----TCAT---GT-A---G---	3183
MND-1.GA.-.MNDGB1	G-A-AT--A-G---AAT-----A---TA-C-----G---G---T-AA-A-G-----A---C---T-----T-A---C---TT---A---G---T---TCAT---T---A---G---	2513
MON.CM.99.L1	-CC-A---CA-T---C-----T---CT-A---CTGT-----G---AG-GC-----TGC-CA-----C---GC---A---G-----TT---G---G---G---G---	2532
MON.NG.-.NG1	GCA-A---AACC-----T---T---TTGC---C---G-----G---G---T-AA---C-----G---TGC-CA-----G---T-A---C---AGC---A---T---TT---G---A---	1125
MUS.CM.01.1085	-----A---A---A---G---T---T---T-----TGC-----A---T-G-----G---AAT---G-----TGC-A-----GC---T-A---G---GCT---C---AC---GCA-----C---	2516
DEB.CM.99.CM40	GCA-C---CACC---C---T---CT-----G---TA-----G---G---AA-G---G---C---TG-C-A---C---A---CCT-A---C---GCC---A---G---GG-G-----G---G---GAC---	2635
DEB.CM.99.CM5	-CA-A---CACT---C---T---GCT---T---G---TT-----G---G---AG-A-----TGC-TA---CC-G---CC-----GCT---C---G---TG-----G---AC---	2629
SYK.KE.-.KE51	GCCTCC--AC-----T---T---C-----T---TA---C---A---G---G---AATTA-----G---TGCCTA-G---T---GA---C---G---AGC---A---G---T---A---G---	2681
SYK.KE.-.SYK173	GC-A-T---ACC-----T---T---C-----TA---T-G-----G---AA-AG-----G---TGC-CA-T---T---A---T-A---G---GCC---A---T---T---TC---G---A---G---	3023





PLV  
complete genomes



	Asp110 catalytic site	
H1B.FR.83.HXB2	GAATACCACATCCCGCAGGGTTAAAAAGAAAAATCAGTAACA.....GTACTGGATGTGGGTGATGCATATTTTCAGTTCCCTTAGATGAAGACTTCAGGAAGTATACTGCATT	2938
Po1	G I P H P A G L K K K K S V T . . . . . V L D V G D A Y F S V P L D E D F R K Y T A F	p51
H1A1.UG.85.U455	---C-G--A-A-G-TC-----A-----G-C-----T-----AG---T-A-----G--	2384
H1B.US.90.WEAU160	-----TT-----A-----A-----C-----	2937
H1C.ET.86.ETH2220	-G-----C-A-----G-----G-----C-----T-----GT-----A-A-----	2330
H1D.CD.84.84ZR085	-G-----T-----A-----G-----G-----A-----C-----A-----TG-----T-----C-----	2460
H1F1.BE.93.VI850	-----T-----G-----G-----A-----G-----A-G-T---A-----C-----	2275
H1G.SE.93.SE6165	-----T-----T-----A-----G-----G-----T-----A-----	2335
H1H.CF.90.056	-----C-A-----T-----G-----C-T-----A-----A-----	2285
H1J.SE.93.SE7887	-----A-----C-----G-----C-T-T-----A-----A-----	2252
H1K.CM.96.MP535	-----A-----A-----G-----C-T-----A-----	2134
H101.AE.TH.90.CM240	-----G-----A-----T-----A-----A-----T-----AG---T-A-----	2512
H102.AG.NG.-.IBNG	-----T-----G-----A-----G-----A-----T-----A-----T-A-----	2463
H1N.CM.95.YBF30	---T-----A-----A-----GC-----G-----TT-----A-A-----T-----TG-----G-CA---T-T-A-----A-T--	2532
H1O.BE.87.ANT70	-T-C-----G-GG-T-G-GC-A-GC-----T-T-----CT-A---A-A-----T-----C-TG-----CCC---T-T-A-A-----T--	2993
H1O.CM.91.MVP5180	-T-T-----A-GG-T-----GC-A-GGC-----T-T-----CT-A---A-A-----T-----C-TGC-T---CC---T-T-A-A-C---C--	2968
CPZ.CD.-.ANT	---T-T---A-C-A---GC-A-G---G-----T-----A-----C---C-CA-A-----C-G---T-A-A---A-T--	2377
CPZ.CM.-.CAM3	-C-----C-G---T-----G-A---G-----G-A-----A-----T-----CTGC-A-----A-T-T---A---A--	2349
CPZ.GA.88.GAB2	-C-T---C-T---T---G-A-G---G-----G-C-----A-----T-----CTGC-----A---T-A---A---C--	2331
CPZ.GA.-.CPZGAB	-C-----T-C-A-----G-----G-----T-A---A-A-----C-C-C-TTG---C-G---A---T---A---A--	2999
CPZ.US.85.CPZUS	-C-----C-A---A---G---G-----C-A---A-G---C-----TTG---AC-G---A-G-A---T-A-A-C---A--	2998
CPZ.TZ.-.TAN1	---C---G---G---A-G---TATG-G-----A-G---A-G---C-C---CA---C-G---CC---A---A---T--	2582
H2A.GW.-.ALI	-G-T---A---A---GCC---G-AG-A-C-T-----G---A---A-G---T-C---CA-A-AC-GC---GAG---T-AC-----	3316
H2A.DE.-.BEN	---T---C-G---AC-GCC-A-G---AGGA-CT-T-----A-T-A---A-G---C---CA-A-AC-C---T-T---C-----	3323
H2A.SN.-.ST	---T---C-A---AC-GCC---G---CG-A-T-T-----C-A---A-G---T-C---CA-A-AC-C---G-T-T-AC-----	2768
H2B.GH.86.D205	-T-T-C-C-G---TG,.GC-G-A---GGAG-A---A---A---C-C-CAGTA-C-AC---CC-A---AC---A---A--	3296
H2B.CI.-.EHO	-T-T-T-C-A---AC-GGC-TCA-G---AG-A---A---A---C-C-CAGT-C-AC---CC---AC-A---A---A--	3294
H2G.CI.-.ABT96	-C-----T-C-T---T---GGC-G-A---GGAG-A---A---A-G---C-C-CAGCA---TC---T---T---C---C-A-T--	2650
H2U.FR.96.12034	-T---C-C-A---AC-GCTG---GAGGA-----T-A---A-G---C-C-CAGCA-C-GC-GTG-----C-A---T---T--	2805
MAC.US.-.239	---C-T---AC-GC---AG-A-T-----A-A---A---C-CA-A-TC-----A-T---C---C-----	3240
Po1	G I P H P A G L G K R K R I T . . . . . V L D I G D A Y F S I P L D E E F R Q Y T A F	p51
SMM.SL.92.SL92B	-T-C-----G---CC---GG-A-TGG-AGGA-----G-A---A-A-G---C-----A---C---CCCG-A---T-AC-A---C-A--	2671
SMM.US.-.H9	-G-----C-TR---AC-GC---G-G-AGGA-C-----GT-----A-----C-CA-A-T-----A---C-A---C--	2713
STM.US.-.STM	-T-----C-T---AC-GC---GG-G-AG-A-C---CA-A---A-G---C-----G-G---T---C-----T--	2887
SAB.SN.-.SAB1C	-G---C-C-A---C---C-GC-A-GGG-GCAGA-T-----GT---CA-A-A---C---AGCTGC-T-----CC-----TCA---A---A--	3336
TAN.UG.-.TAN1	-T-T-C---A---A-G---GA-G---CAGA-C-----A---A-A-A---C---A-AGTA-A-A-G---CC---A---CCC---A-T--	3140
VER.KE.-.AGM155	-T---C---A-----G-A-TG-GCA-A---C-----A-TA-A---G---A-AGCA-A-AC-G---CCT-G---T-A-A-C-A-T--	3160
VER.KE.-.9063	---T-C-G---GC-G-A-TG---CAGA-----T-A---A-A-G---T---A---CA-A-A---CC---G-T-T-A-A-C-A--	3168
VER.DE.-.AGM3	-C-----T---AT---CG---TG-CGA-A---T-A---CA-A-G---C---A---A-A-A---CCC---G-T-A---C-T--	2663
VER.KE.-.TY01	-G---C-C-A---A---G---TG-G-CAGA-----TT-A---A-A-C-C---A---CA-A-A-G---CC-A-T-T---A---T--	2658
COL.CM.-.CGU1	-G---C-A-G---TC-T-T-G-A-G-CCCATA-----T-A---A-A-A---C---C---A---G---CCG-T-ATCAA---C-A-T--	2713
DEN.CD.-.CD1	-T-----T---A---C-C---GG-CACA-C-T-----GT-A-CT-AAA-----A-AGT-G-T-G---CC---T-TGCAC-A-C-A--	2825
GRV.ET.-.GRI_677	---T-T-C-A-G---CC-TC-G-A-G---GCA-A-C-----A-A---CA-A-G---C---A---A-A-A---TGCA-G-A---A-A--	3111
GSN.CM.99.CN166	---T-G-C-A---AA-TCCGC-CCCTGC-GG-C---GAGCAAAAGGAGCATGTAACATA---A-A---AAA---T---A-AGTA-A-A-GT---G-T---C-C-T--	2666
GSN.CM.99.CN71	-G-T-G-C-A---A.....C---GA-CAAAAAGAGCAGTAACTA---A-A---A-AAA---C-T---ACAGT-A-A-GT---A---A---C--	2648
DRL.-.-.FAO	-G-----C-G-G---C-----GCATGCG-CAGA-C-T-----A-----C---C---TGC-T-C-G---CCG---T-A---C-----	2505
RCM.GA.-.GAB1	-----A-G---A---GC-ATGTG-GAG-A---G-----T-A---CA-C-G-----TG-TTC-GT---GCCT-T---A---A--	2417
RCM.NG.-.NG411	---C-G---A-G---A---C-ATGC-GGCA-A-T-----CT-A---A-T-G---G-----TGC---C-T---GCCA-T-A---C---T--	2432
MND-2.-.-.5440	---CA-----A-G---CC---GC-ATGTG-GCA-A-T-----T---CA-A-A---C---TG---G---TG---T-A---G---G--	2499
MND-2.CM.98.CM16	---C---C---T-G---TC-C---C-ATGTG-GCA-A-T-----CA-A-G---C---C---TG---T-G---G---T-A---C---C--	2939
MND-2.GA.-.M14	---C-----G-G---C---C-ATGTG-C-A-A-C-----A-T---A-A-A---C---TG---T-G---GAG---T-A---C-----	2866
MNE.US.-.MNE027	-----C-T---AC-GC---A-GG-GAGGA-C-----A-----C---C-CA-A-TC-----A-T---C---C---T--	2708
LST.CD.88.447	-C-T---C-T---AA---G-ATGC-GCAG-----GG-A---A-A-A---C---AGCA---C-G---CC---A---AC---C-A--	2206
LST.CD.88.485	-C-T---T---AA---G-ATGC-GCA-A-----GG-A---A-A-A---C---AGTA---C-G---CC---AT-AC---C-A--	2206
LST.CD.88.524	---C---A---AA---G-ATGC---CA-A-T-----G-A---A-A-A---C---CAGCA-A-T---CCT-T-A---AC-A---G--	2203
LST.KE.-.lho7	-T-T---C-A---TA-C-G-ATGC---CAGA-T-----G-A---CA-A-A---C---CAGTA-C-T---CC-A-T-A---A---A--	3290
SUN.GA.98.L14	-G-T-T---A---A---G---TGT---CAGA-T-----G-A---G---C---C---A-A-A-G---CC---T-AT---A---A--	3295
MND-1.GA.-.MNDGB1	-T---T-C-A---AA-----TGT---AG-A-----C-A---A-A-G---C---AGTA-A-TC-G---CC---T-AT-ACCC---C--	2625
MON.CM.99.L1	---C---C-A-G---AC-TG-G---TGG-CCAT-G-----A-G-A-CA-AAG-C---C-TC-ACAGCA---GTGG---CCG---A-A-C-G-G--	2644
MON.NG.-.NG1	---T---C-G-T---AC-CC-G---C-TGGC-AAATA---T-----A-G---CA-AAA---C-T-ACAGCA-C-TC-G-----A-T-T---A--	1237
MUS.CM.01.1085	---C-----A---A---T-G-C-TG-GCAC-T-----A-G---A-CAAG-C-C-T-ACAGT-C---C---CC---T---C-A-C--	2628
DEB.CM.99.CM40	-T-----C-AT-T-CAC-CC-C-A-TGG-GAGA-C-T-----CT-A---C-T-CA---C---A---A-AG-AC---CCC-ACA-TGCA-A-C-A--	2747
DEB.CM.99.CM5	-G-C-G-C---T-G---GCC-C-A-TGG-GAGA-C-T-----C-C-A---CT---CA---C---A---G---G---CC-C-G-TGCACGT---A--	2741
SYK.KE.-.KE51	---T---A-G---C-C-GT-CA---G---A-G-----T-A-CT-AAA---C-C---ACA---A-TC-----G---AC---A---A--	2790
SYK.KE.-.SYK173	-----C-C-A-G---AC---CG-C-----C-----A-TA-A---T-AAAA-----AC---C---C-G-CA---G-T---CCA-C-G--	3132



H1B.FR.83.HXB2	TACCATACCTAGTATAAACAAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTTCCACAGGGATGAAAGGATCACCAGCAATATTTCAAAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAA	3068
Pol1	T I P S I N N E T P G I R Y Q Y N V L P Q G W K G S P A I F Q S S M T K I L E P F R K	p51
H1A1.UG.85.U455	C-----AG-C-G-----G-----T-----G-----C-----TC-	2514
H1B.US.90.WEAU160	-----A-----	3067
H1C.ET.86.ETH2220	C-----C-----A-----A-T-----C-C-----C-----G-----C-CC-----C-----GGCC	2460
H1D.CD.84.84ZR085	-----A-----T-----	2590
H1F1.BE.93.VI850	C-----G-C-----T-A-----G-C-----A-----T-----C-----TG	2405
H1G.SE.93.SB6165	C-T-----T-----G-----T-----G-----G-----C-----GC-	2465
H1H.CF.90.056	C-----T-----G-----T-----G-----C-----C-----G--	2415
H1J.SE.93.SE7887	C-T-----C-----C-----A-----A-----C-----T-GT-----A-A-----G--	2382
H1K.CM.96.MP535	C-T-----T-----A-----G-----A-----A-----CA-----C-----T-	2264
H1O1.AE.TH.90.CM240	C-----A-C-----G-----G-----G-----C-----T-	2642
H1O2.AG.NG.-.IBNG	-T-----G--T-----C-----C-----G-----G-----T-GGCA-----C-----C-	2593
H1N.CM.95.YBF30	-----T-T-----C-----T-----G-----A-C-----G-----T-T-G-----CA-----TC-----A-A-C-----G-G	2662
H1O.BE.87.ANT70	C-T-T-----G-G-T-----C-A-A-----C-----C-G-A-----T-----T-----TCA-----TC-----T-A-C-G-G-	3123
H1O.CM.91.MVP5180	C-T-T-----G-G-----C-----AG-A-----C-----C-G-A-G-----T-----C-----T-G-TCA-----G-TC-----T-A-----	3098
CPZ.CD.-.ANT	C-A-T-A-CG-G-----C-----A-A-----TGT-T-----T-A-----A-C-----T-T-T-GCA-----C-G-----CA-A-----GG-T	2507
CPZ.CM.-.CAM3	-T-----C--G--T-----A-----A-----A-----T-----A-----A-----G-----T-----T-----C-----T-----TC-AC-A-C-----C-G	2479
CPZ.GA.88.GAB2	C-----C-----G--T-----A-T-----T-A-G-C-----T-----T-A-----A-----G-----C-----C-----GG-T-----C-----C-----C-G	2461
CPZ.GA.-.CPZGAB	-A-T-C-----T-----AG-----A-T-----TT-G-----A-----G-T-T-T-C-----C-T-----TC-----A-C-C-----G--	3129
CPZ.US.85.CPZUS	-C-----G-----C-----A-----C-----T-----T-A-----A-----G-----T-----T-----G-C-T-----TC-----T-----C-----	3128
CPZ.TZ.-.TAN1	-----C--C-T--A-C-----AA-----T-----T-C-----T-A-T-A-T-----G-----T-----T-T-G-C-T-----C-----T-----	2712
H2A.GW.-.ALI	--TC--ATCAG-----CAGA-----A-AA-----ATA-T-A-CT-A-G-----G-----T-T-----CAC-CA-----G-C-G-----A-C-----G	3446
H2A.DE.-.BEN	--C--AGCAG-----AT-GA-----A-AA-----ATA-T-A-CT-G-----A-----G-----T-----T-----TAC-CA-----GGC-G-----A-----C-----	3453
H2A.SN.-.ST	--TC--ATCA-----CTGA-----A-AA-----CATA-T-A-CTCA-----G-----G-----T-----T-----GTAC-CA-----GGC-GG-----A-A-C-----	2898
H2B.GH.86.D205	C--T-G--ATCAG-----CAGAG-----A-AG-----CATT-T-A-T-A-----A-G-----G-G-C-AGT-----T-GT-----TACTC-----G--GG-A-----T-----G	3426
H2B.CI.-.EHO	--TT-G--AGCAG-----T-----CAGA-----A-AG-----TT-----A-C-----A-----G-----C-----T-----GTAC-C-----GG-AC-----C-----C-----	3424
H2G.CI.-.ABT96	C--TT--ATCAG-----CAGA-----A-AA-----CATC-----G-C-A-G-A-----G-C-----T-----C-----CA-CA-----GG-GG-AC-----C-C-----C-----	2780
H2U.FR.96.12034	--TT--ATCC-T-T-----CAGAG-----A-AA-----CATC-----T-G-T-A-C-----G-----A-----T-----T-----GTA-CA-----GG-CC-AC-G-----A-----G	2935
MAC.US.-.239	--TT--ATCAG-----T-----CAGAG-----A-AAC-----CATT-T-G-T-G-T-----G-G-----C-----C-----TAC-CT-----G-C-TG-GC-----A-C-C-----G-G	3370
Pol1	T L P S V N N G E P G K R Y I Y K V L P Q G W K G S P A I F Q Y T M R H V L E P F R K	p51
SMM.SL.92.SL92B	C-T-----ATCAG-G-----CC-AGA-----A-AG-G-----ATT-----G-C-G-----G-----G-----T-T-----GGCA-C-----GGC-GG-A-----A-C-C-----	2801
SMM.US.-.H9	--T--ATCGG-----CAGA-----A-AA-----CATC-T-G-AT-A-T-A-G-----G-----T-----T-----T-----RCA-CTR-----TG-A-----A-----	2843
STM.US.-.STM	C--T-----ATCAG-----CAGA-----A-AA-----CATC-----G-T-G-----G-----G-----T-----T-----TA-CA-----G--T-A-----A-C-G-G-	3017
SAB.SN.-.SAB1C	C-----C--ATCAG-C-T--AG-GAG-----C-A-----T-----G-----C-----C-----CA--T-T-G-CA-CAGCC-AC-----TC-GC--GAA-----GC-	3466
TAN.UG.-.TAN1	-----C--ACAAG-----T-----TAGGG-T-A-A-----G-----A-T-----TGT-A-T-A-----CA-C-----T-----GCAC-CAGCACAGCG-G--C--A-AGAAA-AG-----	3270
VER.KE.-.AGM155	C-----C-----CGG-----GG-----C-A-----A-TT-----TGC-A-G-----C-----G-----C-GA-----T-----AC-CAGCAT-----TC-----AGAAA-A-AG---	3290
VER.KE.-.9063	--A-C-----TCAG-----T-----C-AGG-----TG-----G-----A-T-----TGC-----A-C-----G-----T-----A-T-----T-----G-A-----CAGCAG-----T-----AAAA-A-A-G-G	3298
VER.DE.-.AGM3	--T--ATCAG-----T-----C-AGGG-----T-C-----T-----CTGT-----A-----G-----C-----A-T-----T-----G-AC-CAGCAG-TTCC-----TC-----AGAAA-A-A--G	2793
VER.KE.-.TY01	--T-T-----CAG-G-T-----C--GG--C-----G-----A-T-----CTGT-C-G-A-G-----T-----TA-----C-----A-----CAGCAG--TCC--T-G--GAGA-A-A--G-	2788
COL.CM.-.CGU1	--A-C--ATCAG-G-----TT-G-----C-----GT-----A-----A-G-----T-----G-----T-----A-----C-----G-A-CAG-AG-----G-----C-----A-G-C-G-G-	2843
DEN.CD.-.CD1	--AG-G--GTCC--T-----AGG-----ACAA-G--CATT-T--A-CT-A-----A-----GT-----G--TA-C-T--T--GGCC-CAGCAGAT-----GC-TAGC-AA-----G-G	2955
GRV.ET.-.GRI_677	-----C--CTCAG-----T-----ACAGGG-----A-G-----T-----TGT-G-T-----G-----T-----TA-----T-----G-A-----CGGCAG-----C--T-----GAGA-C-A--GG	3241
GSN.CM.99.CN166	-T-AG-----TCAG-----T-----C-G-C-----CTGAA-G-----T-----A-CT-A-G-----T-----CG-G--A-C-C-T--AC-TTG-AG-TTCCT-GC--CGA-AGA-----G-T	2796
GSN.CM.99.CN71	-T-AG-----ATCAG-----C-----C-----CTGAG-G-----T-----A--TT-A-T-----T-----G-CG-T--A-C-T--T--GGCC-CTG-AG-TTCCT-G--T-T-AGA-----G-C	2778
DRL.-.-FAO	C-----T--ATCGG-G--T--C--GGC-----A-C-G-----T-----C-C-----C-----T-----C-----T-----G-----C-----CAC-CT-----G--TG-GC-G-A-C-C-G-G	2635
RCM.GA.-.GAB1	--A-----AGCAG-----T-----C-AGG-----AG-G-G-----A-T-----G-----T-----G-----T-----C-C-C-T--GGCATCAGCC-AT--G-----C--A-----GG--	2547
RCM.NG.-.NG411	C-G-----TAGC--T-T--C-AGG-----T-C-----A-----T-----T-G-C-----G-----T-----T-----T-----CA-CAGCT-AC-----T-----A-----G--	2562
MND-2.-.-5440	C-----T--ATCGG-G--T--C--GGC-----A-C-G-----C-----T-----C-A-----G-----C-----C-----T-----GGCA-CAGCAGAT-----TC--A--G--AGG--	2629
MND-2.CM.98.CM16	C-----T--ATCGG-G--T--C--GGC-----A-C-G-----C-----T-----C-A-----G-----C-----C-----T-----GGCA-CAGCAGAT-----C-----C--C--G-G	3069
MND-2.GA.-.M14	--T--ATCGG-----T-----C-AGGC-----A-C-G-----C-----T-----T-----G-----C-----T-----G-----G-----GCA-CGGCAGAC-----TC--A-AA-A--C-A-G-G	2996
MNE.US.-.MNE027	--TT--ATCAG-----T-----CAGA-----A-AAC-----CATT-T-G-C-G-T-----G-----G-----G-----C-----C-----CAC-CT-----G--TG-GC-G-A-C-C-G-G	2838
LST.CD.88.447	C-AG-----ATC--T-T--CC--GA-----GA-----T-----A--A-C-A-----G-----GAGT--TGC--T-T--GC--CAG-AG--GG-C-C--TCAGAA-----	2336
LST.CD.88.485	C-AG-----ATC--T-T--CC--GA-----GA-----T-----A--A-C-A-----G-----GAGT--TGC--T-T--GC--CAG-AG--GG-C-C--TCAGAA-----	2336
LST.CD.88.524	C-T-----GTCAC-----C--GA-----A-AA-----T-----T-----G-----A-----G-----GAGT--TTGT--T-T--GAA-CAG-T-GTGG-C--C-T-CAGAG--C-G--	2333
LST.KE.-.lho7	--T-----TC-C-C-T--C-AGAG-----A-AA-----G-----A-G-----G-----GAGC--TTGT-----T--GG-A-CAG-AG--GG-C-TC-CTCAGAG-----G--	3420
SUN.GA.98.L14	C-A-----GTCAG-----T-----C-G-----A-AA-----GT--T-----C-A-----G-----AGT--TTGC--C-T--GG-A-CGG-AG-G-GCC-----AAA-----G-	3425
MND-1.GA.-.MNDGB1	--GG--ATCAG-T-T--C-AG-----A-AA-----CAT--T-----T-----T-A-G-----G-----AGT--TGT--C-T--G-G-CAG-AG--TC-C-GC-G--GTA-----G	2755
MON.CM.99.L1	CT-A-----CTC-T-----CTGA--T-CT-AA-G-----T-----G--C--C--G-----C-----G--C--A-TG-G--T--CA--CAGC-G--G-GC-ACC-C--GAGA-C--G--	2774
MON.NG.-.NG1	-T-A-----TCAG-----T-----CAGA-----CCG-G-----T-----GA--T-G--C-----R-----C-----A-----T-----GTACTCAGCAT--G-GC-GC-RA-AGRAA-C--G-G	1367
MUS.CM.01.1085	-T-GC-G--CAGG-----CTT-GG-----CA-GA-----A-T--A-----A-----C-----G-----T--TA-C-C-----GTTC-CAGCAG-C--C--AGAGA--AG--	2758
DEB.CM.99.CM40	C-A-G--ATC-G-C-T--ATACA-----AGAA-----G--T--GA--A--A--T--A--G-----GC-CT-----C--T--T--GCCTCAG-AGGG-G-CAG--C--ATC-----GG--	2877
DEB.CM.99.CM5	C-A-----TCGG-G-----ACACA-----AGAG-----G--T--GA--A--A--T--A--C-----GT-C-----C-----GGCA-CAG--GGC-GGCAGC-GA-A-T-----AGG--	2871
SYK.KE.-.KE51	--AG-G--C-----TCTT-T-----AGAG-----C-----T--G--T-G-----A-----G-----G--CA-T-C-T--GGCA-CAG-A-ATC-G--AC-CC-A--A-A--GG--	2920
SYK.KE.-.SYK173	--AG-G--A-C-----T-----CCT-C-G--AGAA-----A-TT-CA--T-A-----G-----T-----CA-T-C-----C-CA--A-ATC-G--AC--C--A-----	3262





PLV  
complete genomes



	Asp185	Asp186 catalytic site	
H1B.FR.83.HXB2	CAAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAAATAGGCCAGCATAGAACAAAATAGAGGAGCTGAGACAACATCTGTTGAGGTGGGGACTTACCACAC		3198
Pol	Q N P D I V I Y Q Y M D D L Y V G S D L E I G Q H R T K I E E L R Q H L L R W G L T T		p51
H1A1.UG.85.U455	-----C-----G-----T-----A-----G-----AT-A--GCT--A--C--T-C-TT--C-		2644
H1B.US.90.WEAU160	-----T-----A-----G-----T-----A-----G-----A--AT-A--G--A--A--T-----		3197
H1C.ET.86.ETH2220	-CC-C--A-----T-----C-----A-----G-CCC--A--T-A--G--A--A--T-----		2590
H1D.CD.84.84ZR085	-----AG-----T-----A-----G-----A--AT-A--G--A--A--T-----		2720
H1F1.BE.93.VI850	A--C-----C-----G-----T-----A--G-----AC--A--T--T-----		2535
H1G.SE.93.SE6165	A-T-----A-G-G--C-----A-----G-----T-A--G--A--AA--GT-----		2595
H1H.CF.90.056	-----T-A-G--T-C-----A-----G-----T-A--GCT--T--AA--T-----		2545
H1J.SE.93.SE7887	AG--C--A-----C-G-----C-----G-----AA--A--G--A--A--T-----		2512
H1K.CM.96.MP535	A-----A-G--A-C-----A-----G-----A-C-----A--A--G--A--AA--T-----		2394
H101.AE.TH.90.CM240	A-----A-G-----A-----C-----T-----C-----A--GCT--A--C--T--T-----		2772
H102.AG.NG.-.IBNG	A-----G--G--C-----A-----T-----G-----T--GG--AC--AA--T-----		2723
H1N.CM.95.YBF30	A-C-----G--A-C--T-C-G-----CC-C--G-----C--CA--A--GAGGC-G--A--C--T--G-T--T--A--CT--G--C-		2792
H1O.BE.87.ANT70	G-C-C--AT--AA--T-G--G-----CC-A-----A-T--CCCC-GACAG-A--A--GG--T--ATT--T--G--C--A-ATCA-----T-C-T--C-		3253
H1O.CM.91.MVP5180	AGC-C--AG--AA--T--G--A--A--C--A-----A-T--CC-T-G-CAG-A-----AG-GGG-C--ATT--T--GG--T-A-ATCA-----T--T--C-		3228
CPZ.CD.-.ANT	A-GT-----CAG--AA--T-----C-----TA-G-----TACTGCA-----GA--TG--AA--T--T--ACA-GTC-----AGAG--T-		2637
CPZ.CM.-.CAM3	-GC-----T-A-A--T-C-G--T-----C-T--G-----T-G--C--A-AG-A--GA--GG--CTA--C--G--AC-T-CT--GT--T--C-		2609
CPZ.GA.88.GAB2	A--C--AT-G--A--T-C-----C--A-----A-T--TCTT-G-AAG-A--C--G-A--G--AC-A--C--G--T-AC-C--A--TT--C--C-		2591
CPZ.GA.-.CPZGAB	A-G-----T--TAC--T-C-G-----CC-A--G-G-----TC-T--T-AT--A--A--GG-G--A--A--A--T--C-T-AA--GT-C--A--C-		3259
CPZ.US.85.CPZUS	-C-C-----TG-TA--A-----T-----C-----G--A-TC--A-CT--AAA--GGA--GG--ACT--C--T--T--C-T-CT--T-C--T--C-		3258
CPZ.TZ.-.TAN1	G-C-C--TG-G-AC--T-C-----T-----C-T--CA--T--TC-TA-TGA--A-G-A--G-A-CTG--A--A--G--T--A-CA-----T-AGAG--C-		2842
H2A.GW.-.ALI	GC--C-AG--TG-CA--C--AT-----A-C-TAA--CTAG--AGGAC-GATTTAG-A--GAC-AGGTGG-CCT-C--T-A-AGG--TC--AAATG-CCTA--T--T--C-		3576
H2A.DE.-.BEN	GC--C--TG-CA--C--ATC--G-----A-C-TAA--CTAG--AGGAC-GGTTTAG--GAC-A-GTGG-CCT-C--A-A-G--T--AAATG-CCTA--GT--T--T--C-		3583
H2A.SN.-.ST	GC--C-G--T--CA--C--AT--G-----A-C-TGA--CCAGC--AGGAC-GATTTAG-A--GAC-G-GTGG-TCT-C--A-AGG--T--AAATG-CCT--T--T--C-		3028
H2B.GH.86.D205	GCC--CAGC--TG-CA--AAT--G-----CA-CCT-A--C-AG--AG-AGTGATCT-G--CGAC-GGGT-G-GTCCC-A--A-A-G-GTTAT-AAATGACAT--T-CT-T--C-		3556
H2B.CI.-.EHO	GCC--CAAT--TG-CAC--AATC--G-----CA-TCTC--G-C-AG--AGGAGCATCT-G--GAC-GGGT-G-GTCTC-A--A-A-G-G-TAT-AAAT-ACAT--T-CT-T--T-		3554
H2G.CI.-.ABT96	GCC--GAG--TR--A--YATC--T-----CA-TCT--G-CCAG--AG-ACTGATTTAG-A--GATCA-GT-G-TTT-C--A-AGG--TCT-AAAT-AAACR--GT-CT-T--C-		2910
H2U.FR.96.12034	GC--G--GA--TGCT--C--AAT--G--T-----CA-TCT-A--CCAG--TAG-A-TGATTTAG--GATCAGTT-G-TTCAC--C-AGG-G-TA--AAT-A-ACA--T--G--C-		3065
MAC.US.-.239	GC-----TG-GACCT-AGTC--G--T-----CA-C-TAA--CTAG--AGGAC-GACTT-G-A--GAC-GGGT-G-TTTAC--TCA-AGG--TCT--AAT--CATA--GT--T--T--C-		3500
Pol	_G N P D V G L V Q Y M D D I L I A S D R T D L E H D R V V L Q S K E L L N S I G F S T		p51
SMM.SL.92.SL92B	GCT--C--TG-GC--C--GTC--G--T-----ACTAA-T--CAG-A--AG--GCC--ACAG--GAC-A--TGG--ACAC--T-A--G-CATG--CAAC-ATCT--GT-C-GT--C-		2931
SMM.US.-.H9	GCT-----TG-GACCC-GATC-----CA-C-TAA--CTAG--TAG-AC-GATTTAG-A--GAC-GGGT-G-TTTAC-AT-A-AGG--T--AAC--TATA--T--T--C-		2973
STM.US.-.STM	GCC--C--TG--AC-C-GATC-----A-C-TGA--CCAG--AG-AC-GATCTAG--GAC-GGGT-G-TTTAC--T--AGG--TC--AAAT-AT-TA--T--T--C-		3147
SAB.SN.-.SAB1C	A-G--C--TG--A--A-----A--TAA--CCAG--TAGGCC--AG-CAG-A--CT-GT--TGG--C--C--T-A--G-CT--AGAA-CT--GT--AA--C-		3596
TAN.UG.-.TAN1	A-GGT--A--TT--AA--AAT--G-----CA--GGA-----A-T--GCC-GAGAAAG-A--GATGA--GG--A--G-A--A-G-T--C-ATTAA-A--CT-CGAG--G-		3400
VER.KE.-.AGM155	G--TTAAA-C-GC-GAGC--TGTC--G-----CC-C-GG-----AC-AGA-----GGTCCAA--GATCAGCT-G--C-AAACA--T--GA-TAGAT--CAAGAA--T-AGAA--		3420
VER.KE.-.9063	G--TTAAA-CCAT-GACA--AGTG--G-----GG--G--C-AGA--TGA-TACACC--GAC-AGTT-G--AC--T--GTCA-G--AAAA-CA--T-AGAA--C-		3428
VER.DE.-.AGM3	G-GTTAAA-CC-C-ACC-TGTG-----CC-A-GG-----C-GGA--TGA-TACAC--GATCCGTTGG--AC-A--A--ATGA-AT-AAGTGCC--T-AGAA--		2923
VER.KE.-.TYO1	A-CTTG--CAC--ACC--TGTA-----A-GG--T--C-AGA-A-TGA-CACACC--GAC-A-TT-G--AC--T-A--AC-A-AT-ACAAGCC--CT-AGAA--C-		2918
COL.CM.-.CGU1	AG-GGA-A-TTG...CAA--AGC--G-----A--G-----A-TC-GCCTT--A-G-----GGGG-G--C--C--G-TGCA--C--TAC-AT--GT-AGAG--		2970
DEN.CD.-.CD1	GCC--T--TG-GA--T-AATC--G--T-----C-C-TG--C--A--TAG-TCCC--AA-----C--G-A--TG-G-TA--A--G--CTTCT--GCAGTA--CT--GAG--C-		3085
GRV.ET.-.GRV_677	-C-C--T-GGT--AA--TGTC-----C-----GGT-G-CG--A--CAT--TGAGACTAGA--ATCA-C-TGG--CATAG-A--A-GATG--C-AGAAAA--T--AGAA--C-		3371
GSN.CM.99.CN166	--GG-G--TG-G--AT-AATA--G-----ACTGA--A--TAGGA--C-G--AG-A--C--GCA-GT-G-GC-CA-AA-A--A-TTTA--AACATCC-ATAACA-CCAA--C-		2926
GSN.CM.99.CN71	--GG-G--TG-G--A--AATA--G-----C--TGA--A--TAGGA--T-G-CAG--GCA-GT-G-GC-CA-AA-A--A-TTTGT--AC-TCC-ATAACA-CCAG--C-		2908
DRL.-.FAO	A-----TG-GA--A-----G-----C--T-----AG--TAG-CC--A-CATTTA--GAGCAG-TG--TA-A--AT-A--AC--T-AAA-TTC--G--CGAG--C-		2765
RCM.GA.-.GAB1	G-G-----TG-CA-C--T-C--G-----C-C-T-----C--A--TAG-AC--AGTT-G-A--GAC-A--TG--CA-AC-AT-A--G-T--AC--TTC--CT--GAG--C-		2677
RCM.NG.-.NG411	G-G-----TG--C-----G--T-----T-----G--A--TAGGAC-GC-CATG-A--TCAG-TG--CA-AAG--C--G-G-----C--CC--TT--GAG--T-		2692
MND-2.-.5440	A-GC-----AG-CT-A-----G--T-----C--T--G--GAG--TC--AGTGCCTCTG-A--C--T-A--TG--TA--AT-A--GG-----T--CA-TTC--G--CGAG--C-		2759
MND-2.CM.98.CM16	AG-C-----TG--G-----T-----C-C-T--T--GAG--TAG--TTGCCCCAG-A--C--C-G--TG--TC-A--T-A--G--C--C--C--TTT--G--CGAG--C-		3199
MND-2.GA.-.M14	G--T-C--AG--T-A--T-----C--T--G--AG--TC-GAGTGCCTGAG--C-AC-A--TGG--A-TAG-T--GG--T--AAG-TTC--G--CGAG--C-		3126
MNE.US.-.MNE027	GC-----TG-GACCC-AGTC--G--T-----CA-C-TA--CTAG--AGGAC-GACCT-G-A--GAC-GGGT-G-TTTAC--T-A-AGG--TCT-AAAT--CATA--GT--T--C-		2968
LST.CD.88.447	TT--C-T--A--CT-G-----A--TCA-----A--TAGG--C-AGAAAGG--GATCAGGC-G--A-A--AT--G-T-TCT-AG--CA--AATT-A-AA--		2466
LST.CD.88.485	TT--C-T--A--CC-G-----ACTCA--G--A--TAGG--C-AGAAAGG--GATCAGGC-G--A-A--AT-A--G-G-TCT-AA--CA--AATT-AGAA--		2466
LST.CD.88.524	TT-----C-CC--C--G-----TCA--T--A--AGG--G-A--AAGGA--TCAGGC-G--A-A--AT-A--G--T-T-AA-A-CT--AACT-AGAA--T-		2463
LST.KE.-.lh07	TT-----GA-C--T--C-----T-----A--TCA-----A--TAG--G--AAAGGA--GATCAGGC-G--A-A--A--C--G--T--TA--CA--AACT-AGAG--		3550
SUN.GA.98.L14	--GT-----AG-TCAAT-G--C-----ACTCA-----T--AT--G-A-AA-A--GA-GA--T-T-GA-AC-AT--G--G-TA--CA--GAA--AATT-AGAA--		3555
MND-1.GA.-.MNDGB1	A-CC-----ACAG--CAGT-A-----T-----G--A--ATAC-TC--AAG--GAG-A-GCT--TA--AT-A--GGCTTTAT-AA--CA--AACT-AGAA--		2885
MON.CM.99.L1	A--T-C-----GG--AC-T-AATA--G-----CC-ACTGA-----AA-TCACCTC-GCCAG-A--C--G-A--T-G--TCA-A--GCGACC--AC-A--AAA--GA--CAA--		2904
MON.NG.-.NG1	A--T-----GG-CACCC--ATC--G-----C-ACTCA-----T--A--CATTCCC--CCTG-A--C--G-A-G-T-G--ACA-CCA-T--ACG-TR--A-CAGAAAA--GA--CAA--		1497
MUS.CM.01.1085	GC--C--TG-CAC-C--ATC-----CC-CCTGA-----CA--A--G--T--TCAAGC--C--G-A-GT-G--CAC--A-T--A-CATG--C--A-A-CTT--GA--CAA--C-		2888
DEB.CM.99.CM40	A-G-----GC--A-A--AGTC--G--T-----ACTA--G--A--T--AGG-A--AT--A--CCT--AG--GG--CAC-AT-A--GTTT--CC-AGA-A--T--G-GG--		3007
DEB.CM.99.CM5	A--T-C-----T--A--AGTC--T-----CC-A-TG-----AA--T--T--AG-TCTCA--TT-GA-C-GG--A-AC--T--GA--TT--CC-AGAAAA--GT-A-GG--		3001
SYK.KE.-.KE51	A-G--T--T--A--A--AGTA--G-----CC-T-TG-----G--A--AG-TC-GA-AA-G-A--G-TCA--TTG--AGT--AA-TGT-A--AG--A--AGCAGTA--T-CT-T-TT-		3050
SYK.KE.-.SYK173	A-GT--T--TC-CAC-C-AATA--G-----C--A-TGA-T--CA--TAG-TC-GAGAA-GCA--CAGGAG-T-G--C-AC-AA-AGTGAC-GCA--A--A-AGT--T--AGGTC-		3392



H1B.FR.83.HXB2	CAGACAAAAACATCAGAAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAAGACAGC.....TGGACTGTCAATGACATACA	3322
Pol1	P D K K H Q K E P P F L W M G Y E L H P D K W T V Q P I V L P E K D S . . . . . W T V N D I Q	p51
H1A1.UG.85.U455	-----G-----T-----G-----T-----C-----T-----CA-----T-----	2768
H1B.US.90.WEAU160	-----A-----C-----T-----G-----T-----C-----AA-----A-----T-----	3321
H1C.ET.86.ETH2220	-----G-----T-----G-----T-----C-----CA-----G-----T-----T-----	2714
H1D.CD.84.84ZR085	-----T-----T-----G-----T-----ACA-----G-----A-----T-----	2844
H1F1.BE.93.VI850	-----GC-----G-----CAAT-----A-----C-----G-----A-----T-----	2659
H1G.SE.93.SB6165	---T---G---A---G---C---G---A---G---A---G---AAA-----C---G---A-----T---C---	2719
H1H.CF.90.056	-----C-----T-----A-----G-----T-----C-----A-----G-----AAA-----T-----	2669
H1J.SE.93.SE7887	---T---G---T---A---G---T---C---A---CAA-----AGA-----T-----	2636
H1K.CM.96.MP535	-----G-----C-----T-----A-----G-----T-----C-----CAA-----C-----G-----T-----	2518
H101.AE.TH.90.CM240	-----G-----G-----A-----C-----G-----C-----AA-----T-----	2896
H102.AG.NG.-.IBNG	-----G-----T-----A-----G-----A-----G-----AA-----A-----T-----	2847
H1N.CM.95.YBF30	-T-----G-----G-----C-----G-----C-----A-----A-----C-----C-----A-----AA-----T-----A-----G-----TGTA-----T-----	2916
H1O.BE.87.ANT70	-T-----G-----A-----G-----C-----T-----A-----G-----G-----A-----C-----T-----C-----CCAAT-----TA-----C-----G-----TGTT-----A-----A-----T-----	3377
H1O.CM.91.MVP5180	-T-----T-----G-----G-----C-----TT-----A-----A-----G-----C-----A-----C-----G-----C-----CCAAT-----T-----C-----AGTG-----A-----A-----T-----	3352
CPZ.CD.-.ANT	-T-----G-----T-----AA-----A-----GT-----A-----A-----C-----T-----AAA-----AA-----A-----GCC-----TGAT-----A-----T-----C-----	2761
CPZ.CM.-.CAM3	-T-----T-----G-----G-----G-----A-----T-----TT-----G-----C-----G-----A-----CA-----T-----A-----C-----G-----G-----C-----T-----T-----	2733
CPZ.GA.88.GAB2	-T-----G-----G-----C-----A-----G-----T-----T-----C-----G-----G-----T-----A-----C-----G-----A-----CA-----TC-----G-----G-----AT-----A-----T-----	2715
CPZ.GA.-.CPZGAB	-T-----A-----G-----G-----A-----C-----TT-----A-----A-----G-----C-----A-----C-----C-----T-----CAAT-----A-----G-----T-----	3383
CPZ.US.85.CPZUS	-T-----T-----G-----G-----G-----A-----T-----A-----C-----C-----CA-----T-----A-----C-----A-----TT-----A-----T-----T-----	3382
CPZ.TZ.-.TAN1	-T-----G-----GT-----G-----A-----A-----G-----C-----G-----A-----AA-----T-----AAA-----CACAT-----A-----CC-----GCAG-----A-----G-----C-----T-----C-----	2966
H2A.GW.-.ALI	---TG---G---GTTC---A---C---A---AAA-----C---G---GTGG---AAC-----AGC---G---AAAA---CAAT---CC---G---AGTA-----A-----C---	3700
H2A.DE.-.BEN	---TG---G---GTTC---A---C---T---AA-----C---G---ATGG---AAC-----AGC---G---AAAC---CAA---CC---G---TA-----A-----C---	3707
H2A.SN.-.ST	---TG---G---GTTC---A---C---A---AA-----C---GTGG---AAC-----AGC---G---AAGA---CAAT---CC---G---AGTA-----A-----C---	3152
H2B.GH.86.D205	---AG---GTTC---A---C---G---AAA-----G---TGG---AA---A---G---A---C---G---AAAA---CAA-----AGTT-----A---G---CA---T---	3680
H2B.CI.-.EHO	---AG---GTTC---A---C---A---AAA-----G---TGG---AA---G---A---C---G---AAAA---CA---A-----GGTT-----A---A-----T---	3678
H2G.CI.-.ABT96	-T---TG---GTTC---A---C---AT---AA-----A---T---RTGG---AA---A---G---A---C---AAAG---TAAC---M-----G---CT-----A---A-----T---T---	3034
H2U.FR.96.12034	---G---TT-----C---A---CA---AAA-----C---ATGG---A---A---G---A---T---G---AAA---T---A---C---G---TGT-----A---A-----T---	3189
MAC.US.-.239	---AG---G---TTC---A---T---C---T---AA-----G---C---T---GTGG---AACAA-----AGT---G---AAAG---A---T---C---G---G---C-----A---G---T---	3624
Pol1	P E E K F Q K D P P F O W M G Y E L W P T K W K L Q K I E L P Q R E T . . . . . W T V N D I Q	p51
SMM.SL.92.SL92B	---AG---C---GTT---A---GA---T---A---A---AG-----A---TTGT---AT---C---CA---G---G---A---T---AAA---AA---C---T---G---GC---A-----A---A-----	3055
SMM.US.-.H9	---AG---G---GTTC---A---T---C---AG-----A---T---GTGG---CAC-----A---C---G---AAAA---A---T---GC---G---A---C-----A---A-----A---T---	3097
STM.US.-.STM	-T---AG---G---TT-----G---T---T---AA-----A---GT---ATGG---AACAA---G---A---C---G---AAAA---AAT---A---TC---G---G---GTT-----A---G---T---C---	3271
SAB.SN.-.SAB1C	-T---A---G---GTT---A---G---T---A---A---C-----G---T---GT---AA---G---CA---C---G---GAG---CACC---A---GG---AGAA-----A---G---T---T---	3720
TAN.UG.-.TAN1	---T---G---GGTG---AG-----AG-----ATGAA-----A---CA---GT---GT---C---TC---C---C---GTC---CA---CAAA---A---A---A---GC-----ACA-----A---G---T---T---	3524
VER.KE.-.AGM155	---G---GGTG---A---G---C---TGAG-----A---A---T---ATGG---C-----AGT---AAG---AAT---AGAGA---G---ACAA-----A---G---TC---T---	3544
VER.KE.-.9063	---GGTG---A---GA---G---A---ATGAA-----A---CA---T---ATGG---AC---A---CA---C---A---TAG---AACT---AGA---G---ACAA-----T---T---C---	3552
VER.DE.-.AGM3	---G---GTA-----A---A---T---ATGAG-----A---CA---T---GTGG---AC---C---G---CAGA---AGCAGC-----AAT---AGA---C---AGAA-----A---T---	3047
VER.KE.-.TY01	---A---G---GATG---A---A---T---ATGAG-----A---CA---TTGG---C---C---GA---C---AGCAGA---CAA---GAG-----TGAA-----C---	3042
COL.CM.-.CGU1	---A---G---T---C---GC---T---A---AAG-----C---T---A---AGC---A---T---AAAGG---CAAAA---T---C---G---T---AG-----A---A-----T---	3094
DEN.CD.-.CD1	-T---TGCC---TT---AG---AC---A---CA---A---AA-----G---T---A---C---CA---AC---C---CA---A---AAA---AC-----TTGCC---GGATGGGCAG---G---G---CA---T---	3215
GRV.ET.-.GRI_677	---G---GTC---A---G---G---GGGAA-----G---A---T---G---GA-----CA---TA---CAAA---AAT---A---CCCCTT---AG---A---GAA-----A---A---CA---A---	3498
GSN.CM.99.CN166	-C---GGCC---C---ACC---TTA---GG---AAG---C---C---G---GT---G---A---GGG---GGC---C---GG---G---AA---C---CC---G---T---C-----A---G---CA---A---C---	3050
GSN.CM.99.CN71	-C---GGCC---C---AC---TTA---G---AGA---C---C---G---GT---G---C---A---GGG---GGC---C---A---AG---G---A---C---TC-----T-----A---A-----C---	3032
DRL.-.-.FAO	-G---TT-----TT---G---CA---T---T---ATGAA-----G---CC---G---G---AA---G---G---G---AGAA---TCGCT---T---T---A---TT-----T---T---	2889
RCM.GA.-.GAB1	---TT---G---TA-----ATT---G-----G---G---G---C---AA---A---GT-----GAG---CAA---T---A---G---AGAA-----A---T---C---T---T---	2801
RCM.NG.-.NG411	-T---G---TT---G---CA---G---C---TGAA-----C---TG-----A---C-----AGAG---CAA---T---A---T-----AGAA-----A---A-----T---	2816
MND-2.-.-.5440	---GTTT---A---G---C---TGAA-----A---TG---G---C---AA---G-----T---AAA---CA---A-----A-----A---A-----T---T---	2883
MND-2.CM.98.CM16	---G---GTT---A---G---A---C---TGAG-----A---CAT---G---C---A---G-----G---AAAG---CA---T---G---AGAA-----A---A-----C---	3323
MND-2.GA.-.M14	---GTTT---A-----T---TGAG-----G---TG---G---C---AC---G-----C---AAAG---CA-----G---CA-----A---A-----T---T---	3250
MNE.US.-.MNE027	---AG---G---GTTT---A-----T---C---T---AA-----G-----T---ATGG---AACAA-----AGT---G---AAAG---A---T---C-----G---C-----T---	3092
LST.CD.88.447	---G---G---TT---AGC-----C---AT---A-----G---T---T---A---GG---GA---A---G---AAA---CTAAA-----C---T---CTTCAGAAAGACA---A---T---CA---	2596
LST.CD.88.485	---A---G---TT---AGC---G---C---AT---A-----G---T---T---A---GG---GA---A---G---AAAG---TAAA-----C---T---CTTCAGAAAGACA---A---A---CA---	2596
LST.CD.88.524	---G---G---TT---AG---CA---A---C---A---A-----G---TC---G---C---A---G---GA---A---G---AGAG---AAA-----G---T---CAAGGAAACT---A---A---A---T---	2593
LST.KE.-.lho7	---A---G---GTTT---AGC---G---A---C---AT---A-----T---G-----GG---GA---A---G---AAAG---TAAAT---A-----G---TG---TCTAACAAAAACT---A---A---C---A---	3680
SUN.GA.98.L14	---G---T---T---GG---G---A---C---ATAAG-----G---CAT---T---A---C---A---C---G---GAGA---G---AAG---AAAT---A---CCCTCT---AGAAGAG...CCA---A---G---	3682
MND-1.GA.-.MNDGB1	-T---A---G---T-----T---C---T---A-----G---GT---A---C---A-----G---AGA---G---AAAGG---TCAA---A-----TT---CAGAAACAG...CCA---A---A-----	3012
MON.CM.99.L1	-CCCAG---G---TT---G---T---CTAC---GG---G---AA---T---A---G-----G---A---C---ATCCC---T---A---A---TGCCAAA---A---A---CC---G---ACAA-----G---C---	3028
MON.NG.-.NG1	-ATAG---G---T---G---CA---GTAC---G---AG---T-----G-----C---CACCC---T-----C---AAAG---AA-----C---CAG-----A-----GG---C---	1621
MUS.CM.01.1085	-TC---AG---TT---G---C---CTAC---G---A---T---A---G-----G---A---C---CCGGGGC---A---A-----CG---CAT---C---C---AGTT-----A---T---TG---G---	3012
DEB.CM.99.CM40	---CCTG---G---GT-----G---C---A---C---T---AC-----G-----GT---A---C---A---A-----CGGT---GAT---AG---TACC---A---TG---AGAA-----A---GC---A---A---T---	3131
DEB.CM.99.CM5	-CCCGG---G---T---C---G---G---T---C---A---A-----G-----GT---A---AA---G-----CACT---GTT---A---A---A-----G---G---GGAG-----A---C---CA---A---C---	3125
SYK.KE.-.KE51	---CCAG---TGG---G---CA---TTC---C---G---AA---T---A---G---CACG---A---C-----G---GTC---AAAGG---TCAA---A-----TTACAGAATCA...CCA---A---T---GT---	3177
SYK.KE.-.SYK173	-A---AG---TGG---G---CC---TA---A---G---AG---T---A-----CACT---T-----A---C---G---CAGT---G---AAA---AA---C---TA---T---T---TGA---GAA...ATC---A---G---C---GT---	3519



PLV  
complete genomes

Table with 3 columns: Accession ID, Nucleotide sequence, and Position. The table lists various PLV complete genomes such as H1B.FR.83.HXB2, H1A1.UG.85.U455, H1B.US.90.WEAU160, etc., up to SYK.KE.-.SYK173. The sequences are aligned and the positions are listed on the right side of each row.



H1B.FR.83.HXB2	CTAGAACTGGCAGAAAACAGAGAGATTCTAAAAGAACCAGTACATGGAGTGTATTATGACCCATCAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAG	3579
Pol1	<u>L E L A E N R E I L K E P V H G V Y Y D P S K D L I A E I Q K Q G Q G Q W T Y Q I Y Q</u>	p51
H1A1.UG.85.U455	T-----T-----G-----G-----C-T-G-----A-----C-G-----A-A-----A-----	3025
H1B.US.90.WEAU160	-----G-----G-A-----G-----GC-----	3578
H1C.ET.86.ETH2220	T-----T-----G-----G-A-----G-----A-T-----T-----A-----A-T-A-----T-----T-----C-----	2971
H1D.CD.84.84ZR085	T-----T-----G-----G-----A-G-----T-----T-----A-----A-----T-----	3101
H1F1.BE.93.VI850	T-----T-----A-----T-G-----G-----G-A-----A-----A-----AG-C-G-----GA	2916
H1G.SE.93.SB6165	A-G-GT-----G-----G-A-----T-----C-----A-----G-----A-----A-----T-----A-----	2976
H1H.CF.90.056	T-G-----T-----G-----G-----A-----A-----T-----G-----A-----A-----A-----G-----	2926
H1J.SE.93.SE7887	T-G-----T-----A-----G-----G-----G-----G-A-----T-----G-----A-----G-----A-----TG-A-----G-----	2893
H1K.CM.96.MP535	T-----T-----A-----G-----G-----G-----G-----A-----A-----T-----A-----A-----A-----	2775
H101.AE.TH.90.CM240	T-----GT-----G-----G-----ACC-----T-----G-----A-----G-----G-----A-----A-----G-----A-----	3153
H102.AG.NG.-.IBNG	T-----T-----G-----G-A-----T-----G-----A-----A-----A-----G-----T-----A-----A-----	3104
H1N.CM.95.YBF30	T-----A-----G-----AT-----CC-G-----C-----GG-----A-----G-----T-----A-----A-----T-----G-----G-----G-----	3173
H1O.BE.87.ANT70	-----G-----T-----A-----AG-----A-----GGT-----C-----G-----G-----A-----C-----C-----A-----TGAT-----G-----TC-----TGG-TTA-----T-----T-----A-----AGGG-----AG-----T-----C-----G-----A-----G-----	3634
H1O.CM.91.MVP5180	-----T-----A-----A-----A-----AG-----G-----G-----A-----CC-----G-----TGAC-----GTGG-----TTAGT-----T-----T-----AG-----G-----T-----C-----GG-----A-----G-----	3609
CPZ.CD.-.ANT	T-----T-----A-----A-----T-----A-----C-----A-----C-----GC-----GAA-----A-----G-----G-----TAC-----C-----C-----G-----TGGTCT-----CCAC-----A-----ACC-----A-----A-----TC-----A-----C-----A-----A-----	3018
CPZ.CM.-.CAM3	A-----G-----T-----A-----A-----G-----R-----T-----RG-----T-----G-----G-----C-----N-----A-----AA-----TA-----T-----G-----T-----TC-----G-----	2990
CPZ.GA.88.GAB2	T-----GT-----A-----A-----T-----G-----A-----T-----G-----CACC-----C-----GA-----AC-----G-----T-----A-----AATA-----T-----G-----T-----G-----T-----G-----	2972
CPZ.GA.-.CPZGAB	T-----T-----A-----A-----T-----G-----AG-----GCAC-----G-----A-----C-----T-----GAC-----GC-----T-----A-----CA-----CT-----G-----T-----G-----A-----T-----G-----	3640
CPZ.US.85.CPZUS	A-----G-----GT-----A-----A-----A-----G-----G-----G-----C-----G-----A-----A-----A-----AAGGA-----T-----G-----C-----T-----	3639
CPZ.TZ.-.TAN1	T-----G-----GT-----A-----A-----T-----AGC-----AAG-----T-----A-----CA-----C-----T-----TAA-----TT-----CCTC-----GCA-----C-----G-----A-----G-----	3223
H2A.GW.-.ALI	GC-----G-----A-----AG-----G-----ATT-----CT-----GTC-----GGA-----CA-----G-----G-----GCAC-----CC-----AGA-----GA-----G-----G-----GA-----AC-----G-----C-----A-----AG-----TCAAG-----CAAT-----A-----AC-----C-----G-----	3957
H2A.DE.-.BEN	GC-----G-----A-----A-----ATT-----CT-----G-----GCC-----GGA-----CA-----G-----A-----TAT-----CC-----AGA-----GA-----A-----GAG-----AC-----C-----A-----AAGCCAAGG-----CAT-----CA-----AC-----C-----G-----	3964
H2A.SN.-.ST	GCG-----A-----A-----A-----ATC-----CT-----GCC-----GGA-----CA-----G-----A-----TGC-----CC-----AGAGGA-----G-----GC-----GA-----AC-----G-----C-----A-----AG-----TCAAG-----CAAT-----G-----A-----G-----AC-----C-----G-----	3409
H2B.GH.86.D205	GC-----G-----ACAG-----G-----T-----A-----ATC-----CT-----G-----C-----GGA-----CA-----G-----A-----TCC-----C-----CA-----GGA-----AGGGT-----CCGC-----GA-----AC-----G-----AA-----CCTAGC-----AAT-----G-----CA-----C-----G-----	3937
H2B.CI.-.EHO	GC-----T-----CCAG-----A-----ATC-----C-----G-----C-----GAGCAGG-----A-----TCC-----CA-----AGA-----GGGGT-----CCT-----GA-----AC-----G-----AA-----TCTAGC-----AAT-----G-----CA-----G-----C-----G-----	3935
H2G.CI.-.ABT96	GC-----T-----CA-----AGATT-----C-----R-----C-----GGARCAGG-----A-----C-----CC-----CA-----AGA-----GR-----G-----GCCT-----RGA-----CACTG-----AA-----TTTAG-----CAAT-----G-----CA-----G-----C-----	3291
H2U.FR.96.12034	GC-----T-----T-----G-----T-----AGATT-----T-----G-----GTC-----GGA-----CA-----G-----G-----G-----CT-----CAGGGA-----GA-----G-----A-----CT-----GA-----AC-----G-----TT-----AA-----TCAAG-----TAAT-----C-----CA-----G-----CC-----G-----	3446
MAC.US.-.239	GC-----TAT-----AG-----T-----A-----ATA-----C-----GTC-----GGA-----CA-----G-----A-----TGT-----CC-----AGA-----GGC-----GCCA-----GA-----CACGG-----ATA-----AGTCA-----G-----CAAT-----G-----T-----T-----A-----C-----C-----	3881
Pol1	<u>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 K S Q D N O W S Y K I H Q</u>	p51
SMM.SL.92.SL92B	GC-----T-----T-----G-----T-----ATT-----TC-----GAGCAG-----A-----AGA-----CAGAGA-----GATG-----CCGC-----GA-----ACGG-----T-----TA-----A-----TCAAG-----CAAT-----G-----GT-----A-----G-----C-----	3312
SMM.US.-.H9	GC-----TAT-----A-----AGATA-----CC-----GTC-----CAG-----G-----G-----TGT-----C-----GGAGGG-----GCCAC-----GAG-----CC-----G-----ATA-----AGTCA-----G-----TAAT-----T-----A-----C-----C-----	3354
STM.US.-.STM	GC-----TAT-----T-----A-----ATT-----C-----GTC-----GGAGCAGG-----A-----GCAA-----C-----CCGAGA-----GAC-----GCCTC-----GAG-----AC-----G-----GT-----AG-----TCA-----G-----CAAT-----T-----C-----A-----G-----AC-----C-----	3528
SAB.SN.-.SAB1C	T-----T-----A-----A-----T-----C-----CT-----G-----C-----AAGCAG-----G-----CAA-----TG-----TCTTCCT-----GG-----TA-----GG-----T-----TT-----A-----T-----AG-----T-----GG-----G-----A-----C-----GC-----	3977
TAN.UG.-.TAN1	G-----GT-----CAG-----C-----TCA-----CT-----GAGCA-----G-----G-----CT-----C-----AGT-----TCT-----T-----G-----G-----CACT-----TACC-----T-----ATG-----T-----G-----GG-----T-----CA-----A-----G-----	3781
VER.KE.-.AGM155	GC-----GTAC-----A-----AG-----C-----AC-----GAGCA-----G-----A-----TACT-----CA-----GA-----CC-----C-----T-----GG-----C-----G-----ATTA-----AG-----T-----G-----T-----C-----T-----CA-----G-----G-----	3801
VER.KE.-.9063	GC-----GTAT-----G-----A-----A-----T-----G-----AC-----GA-----CA-----G-----A-----ACT-----CA-----GA-----GG-----GCCTC-----T-----G-----C-----G-----C-----ATTA-----AAG-----T-----G-----GT-----GT-----A-----A-----G-----	3809
VER.DE.-.AGM3	GC-----TAT-----A-----A-----CT-----ACGGA-----CAGG-----A-----GACC-----C-----CA-----A-----GG-----G-----CC-----A-----C-----G-----C-----G-----A-----A-----TA-----AAG-----T-----GT-----C-----T-----CA-----G-----	3304
VER.KE.-.TY01	GCT-----TAT-----T-----G-----T-----AC-----GA-----CAGG-----A-----ACC-----CA-----A-----GG-----T-----CCTA-----T-----GG-----C-----G-----ATT-----AAGG-----A-----G-----GT-----C-----T-----CA-----A-----	3299
COL.CM.-.CGU1	GAG-----G-----ACA-----T-----A-----C-----AA-----C-----GA-----T-----G-----C-----ATA-----CATG-----G-----A-----TTGG-----T-----C-----TTTC-----G-----TTA-----A-----A-----GGC-----GC-----G-----C-----GC-----A-----	3351
DEN.CD.-.CD1	GC-----G-----TAA-----GCACA-----A-----AT-----AGTA-----TTCA-----AGT-----AGG-----GAG-----CCT-----G-----CTCT-----ATTA-----AG-----AAT-----G-----GG-----AC-----CA-----A-----G-----	3472
GRV.ET.-.GRI_677	GC-----TATAAGA-----C-----TCA-----G-----AG-----GC-----G-----A-----CA-----G-----A-----ACA-----C-----TCTC-----A-----AC-----G-----T-----A-----A-----A-----AG-----G-----C-----T-----CACC-----	3755
GSN.CM.99.CN166	A-----T-----A-----T-----G-----CT-----C-----T-----TTC-----G-----G-----T-----GC-----G-----GA-----GCCGC-----CG-----G-----TG-----TACCTCC-----TA-----AG-----CAG-----GGG-----CATGT-----CAGC-----	3307
GSN.CM.99.CN71	A-----GT-----A-----C-----T-----CTC-----G-----G-----CAGC-----C-----C-----G-----GA-----GCCAC-----TG-----G-----T-----TACCTCC-----TA-----AG-----CAG-----GG-----CATGT-----CAG-----G-----	3289
DRL.-.FAO	-----TAT-----A-----T-----A-----CTA-----AA-----G-----A-----A-----G-----G-----A-----C-----G-----GA-----CC-----C-----A-----CA-----G-----A-----A-----T-----TA-----T-----A-----T-----G-----AG-----A-----	3146
RCM.GA.-.GAB1	A-----TAT-----A-----AGAT-----AAGT-----G-----G-----TGA-----AAG-----CC-----G-----G-----A-----C-----T-----TTA-----AAGGT-----A-----G-----T-----C-----AG-----A-----G-----	3058
RCM.NG.-.NG411	T-----TAT-----A-----A-----CTG-----C-----G-----G-----GAA-----A-----G-----T-----A-----GA-----CCA-----TA-----G-----T-----ATTA-----AAGG-----G-----GC-----C-----G-----AG-----G-----G-----	3073
MND-2.-.5440	-----TAT-----A-----AGCTA-----AG-----GC-----G-----GGA-----C-----CC-----G-----GA-----CCT-----G-----A-----GG-----G-----A-----TTAACA-----A-----G-----GT-----C-----AG-----A-----G-----	3140
MND-2.CM.98.CM16	T-----TAT-----A-----A-----CTC-----AG-----GC-----G-----GGA-----G-----A-----C-----C-----A-----GAG-----GCCT-----G-----A-----GG-----T-----A-----ATT-----ACA-----G-----GT-----C-----AG-----G-----G-----	3580
MND-2.GA.-.M14	T-----TAT-----A-----T-----AGTTA-----AG-----GC-----G-----GGA-----T-----C-----G-----GG-----CCTC-----G-----G-----CA-----GG-----A-----T-----ACA-----G-----A-----GC-----G-----AG-----A-----G-----	3507
MNE.US.-.MNE027	GC-----TAT-----AG-----T-----A-----ATA-----C-----GTC-----GGA-----CA-----G-----A-----TGT-----CC-----AGA-----GGC-----GCCA-----GAG-----CACGG-----ATA-----A-----TCA-----G-----CAAT-----G-----T-----A-----G-----C-----C-----	3349
LST.CD.88.447	GA-----TAT-----A-----C-----CA-----C-----A-----AA-----C-----G-----G-----CT-----T-----TA-----T-----A-----GG-----T-----AG-----G-----C-----AGTAA-----A-----G-----AATTATT-----C-----TC-----GTGGAGA-----	2853
LST.CD.88.485	GA-----TAT-----A-----C-----CA-----C-----A-----AA-----C-----G-----G-----CT-----T-----TA-----T-----A-----GG-----T-----AG-----G-----C-----A-----TAAAA-----G-----GGTCATT-----C-----TC-----GTGGAGG-----	2853
LST.CD.88.524	GA-----TAT-----AGACC-----T-----AGACA-----AA-----A-----G-----G-----CT-----C-----T-----AGT-----AC-----TG-----T-----AGGG-----AC-----A-----CAA-----A-----GGTTATA-----T-----TC-----TGGAGG-----	2850
LST.KE.-.lho7	GA-----TAT-----GGC-----G-----A-----G-----A-----T-----AAGA-----G-----G-----CC-----GA-----G-----AC-----T-----T-----TCAGGG-----C-----AA-----CAAAA-----G-----GAT-----ATT-----T-----TC-----GTGGAGA-----	3937
SUN.GA.98.L14	GA-----TACCAGAGT-----TCAG-----G-----AT-----C-----G-----GT-----TCA-----CAGT-----C-----A-----GAT-----A-----TTGCAG-----G-----G-----A-----GTA-----A-----AAT-----TA-----C-----T-----TGG-----TG-----G-----	3939
MND-1.GA.-.MNDGB1	-----G-----TAT-----A-----C-----T-----A-----CT-----GGCT-----GA-----CA-----G-----A-----GTCT-----T-----TAAT-----G-----A-----TAT-----T-----AG-----T-----T-----AACACAGG-----AG-----TATAT-----T-----TGGA-----G-----	3269
MON.CM.99.L1	GCT-----G-----A-----G-----T-----G-----C-----G-----GGA-----TCCACC-----G-----GA-----C-----CAACG-----GCCA-----AG-----TG-----GT-----ACCTCA-----TA-----CAG-----AA-----GGC-----CAG-----T-----C-----TC-----	3285
MON.NG.-.NG1	GC-----T-----A-----T-----G-----A-----GTC-----AGTCGAC-----G-----GA-----C-----CCAGCTC-----G-----AG-----T-----G-----TACCTCC-----T-----YAG-----A-----G-----GGG-----CRGGTA-----TC-----G-----	1878
MUS.CM.01.1085	GC-----T-----T-----TCAG-----C-----TC-----C-----G-----G-----CA-----C-----GA-----C-----GCCA-----G-----G-----TGCTCA-----TA-----AG-----CTCA-----G-----GG-----A-----CT-----CACC-----GA-----	3269
DEB.CM.99.CM40	GC-----TAG-----G-----AG-----A-----C-----G-----GC-----GA-----GTCA-----T-----C-----A-----AGAGCAGG-----CCA-----T-----TCC-----ATTA-----A-----A-----GG-----GTG-----CAGA-----C-----	3388
DEB.CM.99.CM5	GC-----AGG-----G-----T-----A-----A-----C-----CC-----C-----GGA-----CTCA-----G-----A-----C-----CA-----AGA-----GATG-----GCC-----C-----T-----G-----TTCC-----AATA-----AAG-----A-----G-----GG-----GT-----AGA-----A-----	3382
SYK.KE.-.KE51	GC-----A-----T-----T-----A-----C-----T-----GA-----CA-----ACA-----GTCT-----C-----G-----GA-----CCA-----GA-----C-----C-----TAGC-----TTA-----A-----CAA-----G-----GGG-----CAT-----A-----A-----	3434
SYK.KE.-.SYK173	GC-----T-----A-----A-----G-----T-----C-----A-----A-----G-----G-----GAGCA-----GTCT-----TAA-----C-----CT-----GA-----C-----C-----CACT-----A-----TA-----TCCAG-----GG-----CATG-----A-----G-----	3776



PLV  
complete genomes

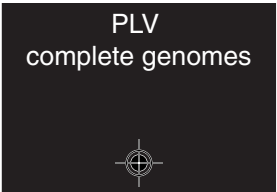


H1B.FR.83.HXB2	AGCCA...TTTAAAAATCTGAAAACAGGAAAATATGCAAGAATGAGGGGTGCCACACTAATGATGTAAAACAATTAACAGAGGCGAGTGCAAAAATAACCACAGAAAGCATAGTAATATGGGGAAAGAC	3706
Pol1	E P F K N L K T G K Y A R M R G A H T N D V K Q L T E A V Q K I T T E S I V I W G K T	p51
H1A1.UG.85.U455	-----AA--TC--T-----A-TG-----G-GT-----	3152
H1B.US.90.WEAU160	-----G-----G-----G-----	3705
H1C.ET.86.ETH2220	-A-----C-----G-G-T--A-GAG-AC-----G-G-----C-T-----G-G-CTG-----	3098
H1D.CD.84.84ZR085	-A-----G-----G-----G-----	3228
H1F1.BE.93.VI850	-C-----A-----G-----A-G-----TCG-----A-----A-----G-TTT-----G-----G	3043
H1G.SE.93.SE6165	-----AC-----G-----G-----A-G-G-TC-----G-----A-T-----G-----G	3103
H1H.CF.90.056	-----G-----A-----AAC-----A-----G-----A-----G-----T-T-----G-----A-T	3053
H1J.SE.93.SE7887	-----C-----G-----G-----A-G-----A-----G-----A-TG-----G-TTG-GC-----A-----	3020
H1K.CM.96.MP535	CA-----G-G-----ATC-----G-----A-----G-----G-----G-----A-----	2902
H101.AE.TH.90.CM240	-----A-----C-----GAG-TC--T-----G-----A-TG-----G-----	3280
H102.AG.NG.-.IBNG	-----T-A-----A-A--TC-----G-----T-----G-G-T-TG-----	3231
H1N.CM.95.YBF30	-TT...CA-----T-A-----G-----A-----ATC-----T-----A-----G-----GTT-A-TG-AAGG-----G-G-A-----T-----T-----	3300
H1O.BE.87.ANT70	-AGA...CA--G-C-C-C-----G-----A-T-GCAA-A-CCT-----A-----A-----G-----G-----A-T-A-C-G-GG-GT-TCA-----TCT--A-T-C-----ATT	3761
H1O.CM.91.MVP5180	-TGA...CA--G-C-T-----T-GCAA-A-CCT-----A-----A-----G-----GG--A-T--C-G-GG-GT-TCA-----GCT--T-----G--ATT	3736
CPZ.CD.-.ANT	-TGA...GGA--CTGT-A--G-----T-GCCT-CA-AA-T-----G--TG-----G-T-GA-T-----GGATT-----T-----T-C-----G--GT	3145
CPZ.CM.-.CAM3	-A-AG...CA--G-C-A-----C-----T-GCAA-ATCA-A-T-----CA-G-----G-----T-A-----G-T-T-----T-----C-----A-----	3117
CPZ.GA.88.GAB2	-AA...GGA-GG--A-----C-----CAA-A-CA-T-CT-----CA-G-----C-G-T-A-TG-----C-----G-TT-C-T-----T--A-CC-T-----CTGT	3099
CPZ.GA.-.CPZGAB	-A-----CA--G--T-----G-----G-----C-----CAA--TCA-A-----A-----CA-C-GG--C-GG-G-A-----A-----TG-T-T-----T-----A-----	3767
CPZ.US.85.CPZUS	-A-GG...CA--G--A--G-----C-----CAA-ATCA-A-T-----A-C-G-----GC-GTT-A-TG-----G-T-T-G-----T-C-T-----GT	3766
CPZ.TZ.-.TAN1	-AGA...GGG--T-A-----A-TCCCA--A-----C-----GA-G-----G-----GACTGA-A-G-----GG-AT-G-----A-----T-----T-TTGT	3350
H2A.GW.-.ALI	GA.....GAA--T--A--GTG-A--G-----A-----AAAA-A--T-C-C-GG--C-G-TTG--G-C--T--T-----GGA-A--GCAC-G-C-T-----CGA-T	4081
H2A.DE.-.BEN	-A.....GAG--TC-A--GT-----G-----AG-A-AAAA-A--T-C--GG--C-G-TT-C-G-C--T--T--G-----GGA-A--GGCAC--C-T-----CG-T	4088
H2A.SN.-.ST	GA.....GGA--T--A--GT-----G-----AGG-A-AAAA-A--C-C-GA-C-G-TCC--G-C-A-T--T-----GGA-A--GCAC--C-T-----CGA-T	3533
H2B.GH.86.D205	GA.....AA--GTC-A--GT-----G-----AGG-T-AAAAA-A--C-C-GG--G-T-C-GG-C-T-T--T--G-----GG-A--GC-C--C--G--G--T	4061
H2B.CI.-.EHO	GA.....GA--TC-A--GT-----G-----AGG-T-AAAAA-A--T-C--GA--G-T--GG-TC-T-T--C-----GGA-AG--GCAT-G-C-C-----G--T	4059
H2G.CI.-.ABT96	GA.....AAC-G-TC-A--GT--T-----AG-A-AAAA-A--T-A--GA--G-TT--GG-TC-T-TG-C--R--GGA-AG--GCAC-----ATT	3415
H2U.FR.96.12034	GT.....GAC-GG-CAT-A--GT-----T--C-AGG-A-A-AA-A-T--T-A--GA--TCGC-TGC--G-C-T-TG-A-G-G--GGA-A--GC-----A-T	3570
MAC.US.-.239	-A.....GAC--TA--GT-----T--AG-A-A-AA-A-A-T-C--GA--G-G--T--G-C-T-T-A-A-G-----GGA-AG--GCA-----G-C-----C-GT	4005
Pol1	E D K I L K V G K F A K I K N T H T N G V R L L A H V I Q K I G K E A I V I W G Q V	p51
SMM.SL.92.SL92B	GT....GAC-GG-T--C--GGT--G-G-T--A--A-AAAA-A-A-T-A--GAA--GGTT-C-GG-TA-T-T--T--G--GGA-A-----TC--G-----	3436
SMM.US.-.H9	-----GAC--TA--A--GT--C--T--AG--T-AAAA-A-A-T-A--GA--C-GCT--G--C-T--T--G--GGA-A--GCA--A--T--C--GT	3478
STM.US.-.STM	-A.....GAC--TAT--GT--C-G-T--A--A-AAAAA-A--T-A--GA--T-G-TT--GG--C-T-TGA-T-G-G--GGG-A--GCA--G--C--C--T	3652
SAB.SN.-.SAB1C	CAGA...AA--TA--GGTT--C-----A--C-AAACA-A--C-----AT--G-ATG--G--GCCT--A-G-G--GGA-AG-----C-C-----C--T	4104
TAN.UG.-.TAN1	-----GGA--GTA--A-----A--AG-GC-AAAA-A-T--T-A--GT-T-G-GT--GG--GAATG--C-----GGG-GTG--AG--GG--CACCT-T-----ATT	3905
VER.KE.-.AGM155	-A.....GGA--TCT-A--GGT--G-G-TC--C-A-CA--AA-C-A-T-----C--GT-GCGTGT-C-G--GA-T--A--G-----GGG-A--GGC-C-----T--C-ATT	3925
VER.KE.-.9063	-A.....GGA--TCT--GT-----G-----G--CA--A-AACA--T-A--AC--TCGCAC--GG-C-GCCT--A-----CTGT-A--GGCAC-G--C-T-----G-ACT	3933
VER.DE.-.AGM3	-----GGAC--GTAT-A--GT--T--CA--AGCA--AAAAA-A--T-C--GT--CCGTGT--GG--GATT--A-----C-TTGT-A--TCTT--T-----G--TT	3428
VER.KE.-.TY01	-A.....GGAC--GTCT--GT-----G-----CA-C-AGCAA-A-AACA--T-A--AC--TCGCAC--G-T-HTTT--G-G--TTG--A--GCTC--T--T-----G-TATT	3423
COL.CM.-.CGU1	-A.....CA-GG--CG--A-G--G-----C--AAC-C-GGC-AAAA-CAG--TT--AG-A-GC-GG-G--G-TTCA-TTA-A--G--GGT-GG--GC-----C-CC-----GT	3475
DEN.CD.-.CD1	-A.....AAAC-TCTA--C-----C-----CCAT-AGGAC-AA--AAG--TTA--CCCATATC-----G-CA-A-TGA--GC-----GGA-G--GC-T-----C--C-GACT	3596
GRV.ET.-.GRI_677	-A.....GGGGC-GTAT-A--GGTG--G-----C-AGCAA--A-AAA-T--T-----C--GGACTC--G--C-CCTT--C-----CTGT-AG--GCAC-TACC--T-----GACT	3879
GSN.CM.99.CN166	-AGGC.....G-TGT-A--T--CC--TC--AG-CC-AAA--A--TA--CTCCTATC-G--TG--T--TT--ACC--G--GGT-GG--G--TCAG--T-CC-----GT	3431
GSN.CM.99.CN71	-CAGC.....G-TG--A-G-T--C-----T--A--CC--AAC--A--T-TA--CTC-TATC-G--GG--T--TA--ACC--G--GGT-GG--TCAG--T-CT-----AGT	3413
DRL.-.FAO	-TGAT...AAC--GCCA--A-----T-AGG-A-A-AA-AGT-----C-----A-G-G-ATG--G--GACT--A-----G--G-A-A--GGCAC-G--C-----G-TT	3273
RCM.GA.-.GAB1	--T...GGA--CCA-----C-----T-AGCA--AAACA--A--C-----AA-T-G-ATGC-GG--G-TT--A-----TG-A-A--GC-----T--T-----G-CT	3185
RCM.NG.-.NG411	GAGAC...CA--CC--A--G--G-----G-----A-CAA-A-ACA-A--T-A--AA--T-G-ATGC-GG--GCTT--A--G-----TG--A--GCT-----T-----CT	3200
MND-2.-.5440	-AGAC...AAC--CC--C--GGT--G-----C--G-CA-A-AA-----A-----AT-G-GGGT-C-TG--G-TTG--A-----GG--G-T-AG--GT-TC-----C-----CT	3267
MND-2.CM.98.CM16	--GAT...AA--GCCG--G--G--G-----C--CA-A-AA--T--T-A--GC-T-G-GTGC-TG--GACT--A--G-G--TG-A-A--GCTT--G-C-----ACT	3707
MND-2.GA.-.M14	-AGAG...AA--GCC--T--GGT--G--G-----T--G-CA-AACA-----A-----GT--GGGTGC-TG--GATT--A-----G--A--GT-----G--T-----G-ATT	3634
MNE.US.-.MNE027	-A.....GAC--TA--GT-----T--AG--A-AAAA-A-A-T--C--GA--T-G--T--G--C-T-T-A-A--G-----GGA-AG--GCA-----G--C-----C--GT	3473
LST.CD.88.447	GA.....AA-TCT-T-----G-G--GG--CA--CAA-AATCA--T--A--CCCC--C-A--GG-----A-----GGA-A--G--T-----C--T-----TC-TGT	2977
LST.CD.88.485	GA.....AA-TCT-T-----G-G--GG--CA--GCAA-AA-CA--T--A--CCCC--C-A-G--GG-----A-----GGA-A--G--T-----C--T-----TC-TGT	2977
LST.CD.88.524	GA.....AA-GCC-T--A-G-G--G--G--CA--CAA-A-A--T--TG--CCCC--TC-G-----G--A--T-A--G-----GGA-A-----C--T-----TTTTGT	2974
LST.KE.-.lho7	GA.....AA--C-TCT-A-G-G-T--G-GG--CA--A-A-CA-A--A--A--CCCC--C-GA--GT--A--TA-T-G-G--GGA-A-----C-C-----CTTTGT	4061
SUN.GA.98.L14	GA.....AAAC--GTAT-A-GGGT--G--G--CA--AG-GAG--CA--A--TGAG--CC-TGTC--G--G--CT--CC-A--G--GGA-G-----T-----TTTTGT	4063
MND-1.GA.-.MNDGB1	GA.....AA--GGT-T-A-G-G--G-----GG-A-CA--AAACA--A--T-G-----CC-C-TGA--GG--GT--TAC--G--GG--GGA-G-----T-----C-----TTTTGT	3393
MON.CM.99.L1	-A.....AAATC-GTAT--G-----T--C-A--T--AA--A--T--C--ACCTATC-G--G--G--T--GC-AGCC--T-GGA-A--GGC-----C-C-----CGACT	3409
MON.NG.-.NG1	-A.....RA--GGGT--A--G--C-----C--TC--AG--T-AAAA-----T-C--CACATACC--G--GG--T--T--AACCC--C--GGG-AG--GGC-T-G--C-----CT	2002
MUS.CM.01.1085	GC.....AAGGGGTAT-A-GG-----T--C-A-G-A--TCA-TA-----AGCTATC-G--GG--C--A--GCC-----GGA-G-----C--C--T-----G--C	3393
DEB.CM.99.CM40	CT.....AAGGG--TAT-A--G-----TCT-T--GAT-AA--A--T--TA-----T-CC-T--G-TA--TA--T--T--GGA-----TT-T-----GA-T	3512
DEB.CM.99.CM5	CA.....AAGGG-TTG--A-----G--TCT--AGGAC--A--AA-T-TTA-----CT-TC-T--C--G-TA--A--T-T--GG-----G--T-----CTAT-----AGT	3506
SYK.KE.-.KE51	G---CAAGAA--CC--TC-----ACA-G--A-GCATAT-CGA--TCAC-----TACC-GGC--GG-TC--CTGA--A-T-----GGA-T-C--GC-C--TGG--T-----A-T	3564
SYK.KE.-.SYK173	--A-AAAGGCC-CCC--A-T-----GACA--A--CCTT-CA--A--T-G-----CTATC--TCTC-TG-TC-ATTGT-AA-T-----GGG-TCC--TCAC--TGGTAT-----GT	3906





H1B.FR.83.HXB2	T...CCTAAATTAAACTGCCCATACAAAAGGAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTGTAGTGGGAGTTGTGTAATACCCCTCCCTTAGTGAATATATGGTACCAG	3833
Pol1	_ P K F K L P I Q K E T W E T W W T E Y W Q A T W I P E W E F V N T P P L V K L W Y Q	p51
H1A1.UG.85.U455	C...-----G--A-----G-----TG-----G-T-----A-----C-----C--A-----	3279
H1B.US.90.WEAU160	-----C-----A-----A-----	3832
H1C.ET.86.ETH2220	-----G-T-A-----C-G-A-----G-----C-----A-----C-----A-----	3225
H1D.CD.84.84ZR085	-----G-A-----T-----C-----T-----A-----C-----T-----A-----	3355
H1F1.BE.93.VI850	ATCT-----G-----A-----TG--A--G-----T-----T-----A-----C-----C--TC-----C-----	3173
H1G.SE.93.SB6165	-----C-----A-----G--A-----T-----T-----G-----C-----C-----TC--A--G-----T--G-----	3230
H1H.CF.90.056	-----G--A--T-----A-----G--C-----A-----A-----C-----ATC--A--G-----T--	3180
H1J.SE.93.SE7887	-----G--A-----GA-----G-----C--C-----G-----C-----C--TC--A--G-----A	3147
H1K.CM.96.MP535	-----G--A-----A-----G-----C-----A-----C-----T-----C--A--G-----A	3029
H1O1.AE.TH.90.CM240	C...-----G--A-----GA-----TG-----G--T-----A-----TC--A-----A	3407
H1O2.AG.NG.-.IBNG	C...-----G--A-----GA--G-----TG-----G--T-----A-----TC--A-----A	3358
H1N.CM.95.YBF30	-----G-T-A--AG-----GTG-----GG-----C--TC-----A--T-----A-----C--C--T-----C--T--A-----T-----	3427
H1O.BE.87.ANT70	G...-----G-----AG--CACT--GA-----T-----G--G--C-----A--A-----A-----C--GC--A--C--A--GA--C-----AG--	3888
H1O.CM.91.MVP5180	A...-----C--GG-----AG--TACT--GA-----T-----T-----G--A-----G-----A-----A-----C--GC--A--C--A--GA--C-----	3863
CPZ.CD.-.ANT	A...-A-----C--T--A-----ACT--G-----T-----TG--C-----T-----C-----G-----A--A-----A--C--A--A--T--G--C--C-----TA--T	3272
CPZ.CM.-.CAM3	A...-A-----C--G--T--A--TG--T-----GGTC-----C-----T-----C--A--C--N--C--G-----C--T--A-----A	3244
CPZ.GA.88.GAB2	-----GT--A--AG--G-----G-----GT-----G--C-----T-----A-----C--A-----A--C--C--GA-----A-----C--T--A--G-----T--A	3226
CPZ.GA.-.CPZGAB	A...-----GGT--A--AG-----G--A-----GT-----G-----G-----A--C-----G--A-----C-----A--A--A--C-----C-----AGT	3894
CPZ.US.85.CPZUS	A...-A-----T-----TAG--G--G-----GTC-----T-----A-----G-----A--T-----A-----C-----C--A-----C--T-----TA--	3893
CPZ.TZ.-.TAN1	G...-----TT--T--A--TG--TCC--A--G-----AGCCAG-----T--T--C-----TT-----G--A-----A-----C-----A--AC--A--C--GGC-----A--T	3477
H2A.GW.-.ALI	A...-A--G-----C--C--A--AG--G--GA-----C-----CAG-----GAT--C--C-----TG--A-----C--A--C-----C-----ATC-----A--AC--G--C--GGC--GCA--TTA--C	4208
H2A.DE.-.BEN	A...-A-----C--C-----AG--GG--G--GA--G--C-----GCAG-----GATA--C--C-----TG--A-----C--A-----C-----ATC-----A--AC--G--C--GG--ACA--TTA--C	4215
H2A.SN.-.ST	A...-A-----C--C--A--AG--G--GA--T--C-----CAG-----GATA--C--C-----TG--A-----C--A--C-----C--CA--ATC-----G--AC--G--C--G-----GTA--TTA--C	3660
H2B.GH.86.D205	A...-AGTG--CC--T-----AG--G--GA--G-----CCAG-----T--C-----TA-----C--A-----C-----CTCG-----A--A-----A--A--G--C--GCC--A--C	4188
H2B.CI.-.EHO	A...-A--TG--CC--T--A--AG--G-----GA--G-----TCAG-----T--C-----TA-----C--A--A-----T-----CTCA-----A--A-----A--A--GG--GCC--TA--C	4186
H2G.CI.-.ABT96	A...-CTTC--C--T-----AG--G--G--C--C-----CA-----C-----C-----TA--T-----C--A--T-----C-----GTC--A--G--ATC--A--GG--GTC--A--C	3542
H2U.FR.96.12034	-----A--TG--CC--T-----AG--G--GA--GGTC-----CA-----T--C-----G--T-----G--C--TA--C-----T-----ATCC--A--A-----C--G-----G--C--GGCA--TA--C	3697
MAC.US.-.239	C...-A-----CC--CT--A--AG--TG--G-----TGT-----CAG-----C-----G--TA-----A--G--A-----T-----A--CTCA--A--A--GC-----A--G-----GTC--T--A--T	4132
Pol1	_ P K F H L P V E K D V W E Q W W T D Y W Q V T W I P E W D F I S T P P L V R L V F N	p51
SMM.SL.92.SL92B	A...-ATT--C--C--T--AG--G-----GGTT-----TCA-----T-----T-----T--T-----A-----C--CA--CTCC--A--A--G--A--C--GG--GT--TTA--T	3563
SMM.US.-.H9	G...-A--G-----C--TT-----AG--G--RA-----TT-----CA-----T-----TA-----A--A-----C-----GTCA--A-----C--G--C--GTT--T--A--C	3605
STM.US.-.STM	C...-A-----CC--CT-----AG--G--GA--GGTC-----CA-----G-----T-----G--TA--A-----A--A--A-----T-----ATCC--A-----T-----A--G-----GTC--TTA--T	3779
SAB.SN.-.SAB1C	A...-----TTA--GG-----AG--G--GA--GCTC-----CA-----T--T--C-----G--T-----C--G--A-----AA--G--C--G-----T-----AA--GA--T--G--G-----A--A	4231
TAN.UG.-.TAN1	A...-C--TCA--GG--G--A--AG--G-----GA-----GT-----CA-----GT--C-----T-----A-----C--A--A--ATA--A--CA--CTCC--A--G-----A--C-----TA--C	4032
VER.KE.-.AGM155	A...-C--CT-----G-----C--AG--GG--G--G-----CA-----G-----C-----T--GT-----A--C--A-----C--G--GTT--G-----A--CT--G-----TACA	4052
VER.KE.-.9063	A...-A--TT--G-----C--AG--G-----GT-----CA-----G-----C-----TG--G-----A-----A-----G--G--GTA--A--TC-----G-----ACT	4060
VER.DE.-.AGM3	G...-AGTCC--G-----C--A--G--G--GT-----CA-----G--T--T--C-----G--TA--GT-----A--C-----A-----C-----A-----TACC	3555
VER.KE.-.TY01	A...-AGTTC--AG-----C--G--G-----GA--GGT-----GCA-----G--G--T--C-----G--TA--G-----C--A-----T-----C--GC-----A--T--GC--C-----ACA	3550
COL.CM.-.CGU1	G...-C--A--G--G--C--AGC--A--G--GA--GGAT-----CAG-----GT--C-----AGTG--T-----C--A--ATA-----CTCAT--T--TATG-----A-----GTA--GGA--C	3602
DEN.CD.-.CD1	-----CT--C--G-----C--AG--AGC-----T--GAT-----TG-----T-----A-----TT--AT-----A-----C-----AA--G--ATCA--T--A-----C--TC--A--G--AAT--TT--TA--T	3723
GRV.ET.-.GRI_677	-----ACG--G--AC-----C--AG--G--C-----T--TG-----CAG-----C-----TAT-----A--A--A-----GC--A--A--T--C-----A-----C--G-----TTCC	4006
GSN.CM.99.CN166	ACCAGAC-----GGA--C--AG--GTC-----CA-----TG-----TGA--T-----G--A--A-----A--CACCATA--AGCAA--AC--C--A--A--A--C--C-----G--CA-----ACC	3561
GSN.CM.99.CN71	ACCTGAC-----C--GGA--C--AG--GTC-----GCA-----T--G-----TGA--C-----G--A--A-----CACCATA--AGCAA--AC--C--A--A--AT--GT-----G--CAG-----ACC	3543
DRL.-.-.FAO	G...-C--TT--CT--CT--A--AG--G-----GTT-----CA-----CAG-----C-----G--TA--A-----C-----T-----A--A--ATCA--A--C--ATC--CA--C--G-----TA--C	3400
RCM.GA.-.GAB1	G...-A--C--C--G-----A--G--G--GA--GGT-----CTA-----C--T--CC--A--C-----G--TT-----A--A--T-----A-----GC--A--A--TC--TA--T--G--C--CG-----A--C	3312
RCM.NG.-.NG411	-----A-----G--G-----A--G--GA--GT-----CAG-----T--G--T--C-----G--TTT-----A-----A-----G-----T--A--AC--A--A--GGC-----A--T	3327
MND-2.-.-.5440	A...-G-----CT--T--C--GT--G--G--G-----GGT-----CCA-----CAT--C-----TA--A-----C--A-----C--CA--TCA--A--A--A-----A--A--GG-----A--C	3394
MND-2.CM.98.CM16	C...-----G--T--T-----A--G--GA--GTC-----CCAG-----C-----A-----A-----C--A-----A-----ATCC--T--A--A--C--TA--CGGG--G-----TA--C	3834
MND-2.GA.-.M14	G...-A-----CT--T--C--C--GG--G--G-----GT-----TCAG-----CAT--T-----G--A--A-----C--A--A-----A--CA--CTCA--A--A--AC--A--A--GG--G-----TA--C	3761
MNE.US.-.MNE027	C...-A-----CC--CT--A--AG--CG--G-----TGT-----CAG-----C-----G--TA-----A--A--A-----T-----A--CTCA--G--G--AC-----A--G-----GTC--T--A--T	3600
LST.CD.88.447	G...-C--GA--C--GG--A--GG--AC--GA--GTG-----TCA-----GT--TC-----G--T-----A-----A--T-----A-----CTCA-----A--AG-----AAC--GGA-----TAGC	3104
LST.CD.88.485	G...-----GA--CC--GG--A--AG--AC--GA--GTG-----TCA-----GT--TC-----G--T-----A-----A--T-----A-----CTCA-----A--AG-----AAC--GGA-----TAGC	3104
LST.CD.88.524	A...-A--GA--AC--G--A--TG--CACT--GA-----AT-----T--A--C--T-----TA--T-----A--A--C--TA--A--CA--CTCC--A-----AA-----AAC--GGA-----TAGC	3101
LST.KE.-.lho7	G...-A--A--C--G--C--AG--GACT--GA-----GTC-----GC--CC--C-----G--TT--A-----A--C--TA--A--CA--TC-----G--AA-----AAC--GAG-----ATT	4188
SUN.GA.98.L14	G...-A--GA--AC--GG--A-----GA--G--T-----AGTCAG-----G-----C--C-----GTGT--A-----C-----TA--A-----C--GC--T-----AAG--G--A--C--GAG-----TACT	4190
MND-1.GA.-.MNDGB1	A...-A-----A--GC--GA--A-----CTAC--G--G--T-----GAT-----CAT-----GTGT--A-----A--A--AGTA--A-----A--C--GC--A-----ATG-----AA--GGGA-----TAGC	3520
MON.CM.99.L1	-----C--TT--CGC--A--AG--GGTG--A--GCA-----CG-----G-----TA-----AT--AT--G--A--A--CATA-----GCG--GT--T-----A--A--C--GC--C--G--CAG-----TACT	3536
MON.NG.-.NG1	A...-A--T-----G--T--A--AG--A--C--A-----CAG-----C-----G--C--CA-----T-----T-----A--A--ATA-----CCA--AC--CT--T--A--A--C--GC--T--G--CA-----ACC	2129
MUS.CM.01.1085	CCCTGAC--GG--C--G--A--C--AG--GATC--A-----CAG-----CCAG-----T--CAC--C-----G--T--T--G-----A-----ACCATA--AGCA--GC-----A--C--A--C--GT--A-----T--ACA	3523
DEB.CM.99.CM40	A...-AC--G--C--GG--C--AG--GGT-----A--GA-----C--AT-----CATA--TC--C-----G-----C--A--C-----GCAA--AC-----A--ATC--G--C--GC--C-----TG--	3639
DEB.CM.99.CM5	-----C-----GGT-----AG--GT-----GA-----T--AC-----CATA--CC--C-----G-----T-----C-----C-----GCG--GC-----A--AAC-----C--GCC-----TG--	3633
SYK.KE.-.KE51	C...-G-----CC--C--C--AG--A--G-----GA-----A-----TC--C-----A--A-----T--G--A--AGTAA--A--G--GC--C-----CATA-----T--G--GG--AT--A--T	3691
SYK.KE.-.SYK173	A...-A--C--C--T--C--AG--A--G-----GA-----A-----C-----C-----A--G--A--A--AGTCA--A--CA--A--G-----G--A--G--T--GG--GG--AC--TA--C	4033



PLV complete genomes

Strain	Sequence	Position
H1B.FR.83.HXB2	TTAGAGAAAGAACCCATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAGTTGTCCACCTAACTGACA	3963
Pol1	L E K E P I V G A E T F Y V D G A A N R E T K L G K A G Y V T N R G R Q K V V T L T D	p15
H1A1.UG.85.U455	-----C-----G-A-----T-----A-----GC-----G-----G-C-----TT-----G-----	3409
H1B.US.90.WEAU160	-----T-----A-----G-----C-----G-----T-----A-----	3962
H1C.ET.86.ETH2220	C-----C-----T-----T-----A-----A-----A-----G-----G-----G-----A-----TT-----A-----	3355
H1D.CD.84.84ZR085	-----G-----A-----T-----T-----AA-----G-----C-----G-----G-----C-----T-----T-----	3485
H1F1.BE.93.VI850	---A-C---G---C---C---T---A---T---T---A---C---AA---G---G---A---A---A---T---G---G---	3303
H1G.SE.93.SE6165	---A-C---G---CC---T---A---T---A---A---G---G---C---A---AG---A---A---T---G---A---	3360
H1H.CF.90.056	---A-C---G---C---T---A---A---T---T---A---A---G---G---C---G---AG---T---T---G---A---	3310
H1J.SE.93.SE7887	---A-G---A---G---T---T---T---A---AC---G---G---G---C---A---G---A---T---T---A---	3277
H1K.CM.96.MP535	---A-C---G---G---T---C---C---T---A---A---AG---G---G---C---G---G---T---A---A---	3159
H101.AE.TH.90.CM240	---A---C---G---G---T---GT---GC---G---G---C---G---G---A---TT---G---G---	3537
H102.AG.NG.-.IBNG	---A---C---G---T---T---A---GA---G---C---G---C---G---TT---G---G---	3488
H1N.CM.95.YBF30	---A-C---G---A---CAGT---G---T---T---A---A---A---G---T---T---G---AG---G---G---G---T---TA---TG---A---	3557
H1O.BE.87.ANT70	C---A-GT---T---TA---G---G---AT---A---T---A---C---G---AG---ACA---G---A---G---GA---AA---AAT---GA---G---	4018
H1O.CM.91.MVP5180	---A-C---T---T---G---T---T---A---T---A---C---G---G---AG---ACA---A---G---CA---AA---AGT---GAA---G---	3993
CPZ.CD.-.ANT	C---TT---GC---C---T---TCC---AG---T---T---T---T---AA---CT---C---G---G---G---C---G---AG---C---AGT---GG---AAAACA---CAA---G---	3402
CPZ.CM.-.CAM3	---G---A---C---CC---G---G---G---A---A---GAA---C---C---C---G---C---G---A---A---A---T---GAAA---T---	3374
CPZ.GA.88.GAB2	---G---A---T---CA---AT---T---C---T---A---A---AG---T---AT---A---G---CA---TT---T---GGAA---G---	3356
CPZ.GA.-.CPZGAB	---C---T---CC---ACCA---T---T---AT---A---A---T---A---A---ACT---G---A---AG---C---A---G---A---CA---T---G---T---GAAA---T---	4024
CPZ.US.85.CPZUS	---C---GG---AG---G---A---T---G---A---CC---AC---A---C---GAA---G---G---G---T---A---A---TT---GAAA---T---	4023
CPZ.TZ.-.TAN1	C---GTT---TCT---C---CCC---A---T---T---A---A---C---G---AAG---G---A---A---C---C---T---C---GGTCAAAGGA---T---GAGA---	3607
H2A.GW.-.ALI	C---TA---T---T---C---C---G---CAC---C---C---CTG---T---C---AT---A---GA---A---A---AG---G---G---C---G---GAAGGTA---GAAC---A---	4338
H2A.DE.-.BEN	C---TAGG---T---T---CC---C---G---CAC---AT---TGC---T---AC---T---A---GAG---A---A---AG---A---G---AAAAGTAT---GAAC---A---	4345
H2A.SN.-.ST	C---G---T---T---C---C---CAC---AT---CTGC---T---A---C---AT---A---GA---CA---A---AG---G---C---G---GAGGCTAT---GAGC---A---	3790
H2B.GH.86.D205	C---TC---C---C---A---GAG---A---CAC---T---CTGC---T---AACCT---A---GGA---C---G---G---A---G---T---G---AAAAGTGT---GAAC---G---	4318
H2B.CI.-.EHO	C---G---TC---C---C---A---T---T---A---CAC---AT---CTG---AA---CCT---A---GAR---G---G---C---G---AG---C---G---AGG---T---AAAC---AT---GAAC---A---	4316
H2G.CI.-.ABT96	C---TC---TC---AGCAGA---G---AT---AC---C---T---CTGC---T---AAACCT---C---GGAR---C---C---C---AG---C---G---AAARGT---T---GAGC---G---	3672
H2U.FR.96.12034	C---TA---TC---AT---C---A---G---A---AT---AC---CT---TTG---CTT---A---CAG---G---C---G---AG---G---A---G---C---G---AAAAC---TT---GAAC---G---	3824
MAC.US.-.239	C---T---G---C---T---AG---A---AT---AT---AC---AT---TG---T---AAC---T---A---GA---G---G---A---C---AG---G---C---A---G---C---AAAAGTGT---GAAC---G---	4262
Pol1	L V K D P I E G E T Y Y G D G S G 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	p15
SMM.SL.92.SL92B	---TA---T---A---A---AAG---AG---GTA---AT---CA---T---CT---CTGC---T---AA---C---G---GA---G---C---AG---C---G---A---G---G---CC---GC---AT---GAAC---GG	3693
SMM.US.-.H9	C---TA---G---T---CAG---A---T---A---AT---CTG---T---C---T---A---G---GA---C---G---AG---C---G---G---GC---GACAAAAC---GGAAC---G---	3735
STM.US.-.STM	---TA---T---G---A---A---A---T---G---CT---TG---C---AT---GA---G---G---C---CA---A---G---A---A---C---AAAAG---TT---GAAC---G---	3909
SAB.SN.-.SAB1C	---TA---T---CCC---A---C---GT---AT---T---AA---TT---C---GA---CT---A---AG---GGAT---G---G---AG---AT---GAGA---T---	4361
TAN.UG.-.TAN1	---G---TA---T---CC---AA---A---GGT---AT---TGC---T---AA---TT---C---G---C---TCC---A---C---C---GA---TC---C---CA---AGAAT---GAG---T---	4162
VER.KE.-.AGM155	C---GACT---G---CCCG---AG---TGT---A---A---CTG---T---AC---T---G---GAG---C---CA---A---CC---ACA---C---A---G---ACAACAG---GAAA---	4182
VER.KE.-.9063	C---GACC---G---T---CC---AG---TGTA---AT---CTG---T---AA---CT---A---GA---G---G---CA---AC---ATAT---A---G---G---A---AG---AAT---GAGC---G---	4190
VER.DE.-.AGM3	C---GACA---G---CC---AAG---A---TGT---A---G---A---TTG---T---AA---TT---A---GGGA---A---C---AC---ATAT---G---A---G---G---AA---AAT---GAAA---T---	3685
VER.KE.-.TY01	---ACA---CCCAAG---AG---CGTT---A---A---T---TGC---AA---TT---A---GA---G---G---CA---CT---AC---ATAC---A---G---G---A---AA---AT---GAAA---	3677
COL.CM.-.CGU1	---G---TTTT---G---TC---GAAG---AG---CC---T---AT---GGAC---A---G---TGCT---A---ATTAGG---GCG---G---GGA---A---TCA---A---GA---G---G---A---AT---TCA---AG	3732
DEN.CD.-.CD1	---G---TATC---G---ACAA---GCC---AT---AC---A---A---T---AC---CA---G---GG---C---CC---AT---C---AG---G---A---A---T---GGAA---A---	3853
GRV.ET.-.GRI_677	---TA---A---CAA---A---TGTT---AT---G---ATC---AA---T---C---T---G---C---GT---AG---G---A---AGT---G---A---AGGGAAT---GAAA---	4136
GSN.CM.99.CM166	C---TCTCG---C---A---AT---TT---T---AT---A---A---T---ACCT---C---C---T---T---G---C---C---A---T---GG---CCA---AGAAT---GAA---A---	3691
GSN.CM.99.CN71	C---TCTCG---C---A---AAT---TC---AT---C---A---T---ACCT---T---C---G---C---C---A---T---G---CCA---AGAA---GAA---GA	3673
DRL.-.FAO	---GCT---GT---C---A---CC---G---A---GTT---AT---G---A---A---TT---G---GA---T---C---CTCC---G---A---G---AA---AG---A---GAA---A---	3530
RCM.GA.-.GAB1	C---TT---G---T---CC---AG---GT---A---G---AA---T---G---A---G---G---A---A---G---A---G---AAAAGAAT---GAA---A---	3442
RCM.NG.-.NG411	C---TA---G---T---T---CC---A---GT---AT---A---T---A---T---G---G---G---A---AG---G---A---G---AAAAGAAT---GAAA---T---	3457
MND-2.-.5440	C---GCTA---C---A---CC---G---A---TGTA---AT---A---AA---CT---C---GGA---C---G---C---TA---CTC---T---A---AGC---G---A---AG---T---GGAA---A---	3524
MND-2.CM.98.CM16	C---GTT---G---TG---GCC---A---TGTA---AT---G---A---AA---CT---C---GA---G---G---A---AGC---AAT---A---TCC---G---AA---AG---GAGA---T---	3964
MND-2.GA.-.M14	C---GTTA---A---CC---G---AG---TGT---AT---G---A---T---AA---CT---GGA---G---G---CTA---AGC---G---A---AGT---AA---AG---T---GAAA---	3891
MNE.US.-.MNE027	C---T---G---C---T---AG---A---AT---AT---T---TG---T---AAC---T---A---GA---G---G---A---C---AG---G---C---A---G---C---GAGAGT---T---GAAC---G---	3730
LST.CD.88.447	---AGC---T---A---G---T---T---AT---A---T---AG---A---AA---TAGGG---AC---G---T---C---AC---ATC---AGG---G---AAAGAAT---A---	3234
LST.CD.88.485	---AGC---A---G---T---T---AT---A---T---AG---A---AA---TAGGG---AC---G---T---C---AC---ATC---AGG---G---AAAGAAT---A---T---	3234
LST.CD.88.524	---G---AGC---A---TA---T---A---G---A---T---AG---A---A---TAGG---AC---G---A---A---C---ATC---A---G---G---A---AAAAGAAT---CAA---T---	3231
LST.KE.-.lho7	---GG---GC---A---G---T---C---A---A---A---AG---A---A---TAGGA---AC---G---A---A---C---ATC---G---A---G---G---AAAGGAGT---	4318
SUN.GA.98.L14	C---CC---C---AG---GCC---C---A---G---A---G---A---A---CTAGAA---GAG---G---A---A---CC---GG---TCC---G---AAG---AAT---GGAAA---	4317
MND-1.GA.-.MNDGB1	---GTCCCC---TC---AG---G---T---A---AT---T---A---C---G---A---G---A---AG---T---TTT---GG---A---AAGAATATCTTAA---T---	3650
MON.CM.99.L1	C---G---TAC---C---A---G---TCC---A---A---C---G---C---C---T---ACAT---A---GGA---G---G---T---A---C---CTTT---GGC---AG---G---CC---AT---GAAC---G---	3666
MON.NG.-.NG1	---T---C---R---T---T---A---CAC---TCCC---A---G---G---C---CAT---A---GA---G---G---G---C---C---CCA---AG---G---AA---TC---AT---GAGC---G---	2259
MUS.CM.01.1085	---G---TACC---AT---AA---A---TCCC---A---AT---G---A---A---T---TAT---C---G---C---C---A---C---C---G---A---G---G---AA---T---G---GAAAGT---	3653
DEB.CM.99.CM40	C---G---T---TC---CC---AT---C---A---T---C---G---T---C---AG---GT---G---A---GT---AAAATG---T---GAGA---T---	3769
DEB.CM.99.CM5	C---T---TATC---G---A---CC---AT---T---C---AT---A---T---AT---A---A---G---C---CA---AT---AG---AT---G---A---GGG---AAA---GT---GGAAA---T---	3763
SYK.KE.-.KE51	---G---TTTC---G---AG---A---A---T---C---C---T---A---T---GC---GA---C---T---CA---A---AG---G---T---A---G---C---T---AAAG---GGT---GGAAA---	3818
SYK.KE.-.SYK173	---G---TACC---CC---A---T---C---A---T---A---C---GC---GAC---T---AG---A---G---GAT---G---C---G---G---ACAGTAT---GGAGC---A---	4163





PLV  
complete genomes



H1B.FR.83.HXB2	AGAGTTAGTCAATCAAATAATAGAGCAGTCTAATAAAAAAGGAAAAGGTCTATCTGGCATGGGTACCAGCACACAAGGAATGGAGGAAATGAACAAGTAGATAAAATTAGTCAGTCTGGA...ATCAGG	4220
Pol1	E L V N Q I I E Q L I K K E K V Y L A W V P A H K G I G G N E Q V D K L V S A G . I R	p15
H1A1.UG.85.U455	---AA-----A-C---G-----A---C---T-----G-----G-----G-----T-----	3666
H1B.US.90.WEAU160	-----G-----A-----A-----C-A-----T-A-----	4219
H1C.ET.86.ETH2220	---A-----A-----GC-----G-----C---T-----T-----A---AG-----	3612
H1D.CD.84.84ZR085	---A-----G-----A-----C-----T-----C-----AG-----	3742
H1F1.BE.93.VI850	---AA-----A-----C-----G-----C---T-----T-----G-----G-----	3560
H1G.SE.93.SB6165	---A-----A-----C-----C-----C---T-----G-----G-----C-----AG-----	3617
H1H.CF.90.056	---T-----T-----G-A-----G-----C---T-----G-----G-----C-----T-----T-----G---A	3567
H1J.SE.93.SE7887	---A-----T-----G-A-----G-----C---T-G-----G-----G-----T-----	3534
H1K.CM.96.MP535	---T-----T-----A-----G-A-----C---AT-----	3416
H1O1.AE.TH.90.CM240	---AG-----C-----G-C-----G-A-----C---T-----G-G-----T-A-----	3794
H1O2.AG.NG.-.IBNG	-----A-----C-----C-A-----C---T-----G-----AA-----	3745
H1N.CM.95.YBF30	---A-G-G-GC-----AG-C-C-----GA-T-----CT-T-----T-----T-----G-G-----C-----T-CT-A-----T-A	3814
H1O.BE.87.ANT70	CCCTA---TC-G-----AG-AC-C-----C---G---TA-----T-T-C-T-----C-A-----A-A-----A-CAAG-AT-----T-A	4275
H1O.CM.91.MVP5180	CCCTA---TC-G-G-----G-AC-C-----CGA-G---TA-----T-T-T-----C-A-----A-A-----A-CAAA-AC-----T-A	4250
CPZ.CD.-.ANT	---CCC-----GG-AG-----CC-A-A-----GA-C-A-T-C---CT-C-----T-----C-A-----G-----C-----CAA-G-----A	3659
CPZ.CM.-.CAM3	---A-G-T-----T-----AG-----G-----AAGT-C---TT-----G-----T-----T-----G-A-----A	3632
CPZ.GA.88.GAB2	---AA-----G-----AG-C-----G-A-G-----T-A-----A-----G-----G-----AG-C-----A	3613
CPZ.GA.-.CPZGAB	---A-----G-----AG-----T-A-----AA-----C-CT-C-----T-----T-----G-----A-----A-----	4281
CPZ.US.85.CPZUS	---A-G-----G-----AG-A-----G-A-----AAGT-A-CT-----T-----T-----A-----C-----T-A-G-----A	4280
CPZ.TZ.-.TAN1	CCCCA---AG-G-----T-TC-AA-----C-C-----AC-GCAA-T-----A-----C-T-----T-A-G-T-----G-----C-----G-T-----AAAAAT-----T-A	3864
H2A.GW.-.ALI	TAGAA---A-----C-----AG-AA-G-----GCCA---G-T-G---C---C-----C-A-G-----C-GG-----CC-T-----A---CAG-C-----A	4595
H2A.DE.-.BEN	TAGAA---G-C-G-C-----AG-AA-G-----G-----GCA---G-T-----C---C---T-----C-A-----CC-GG-----CC-T-----A---CAA-C-----A	4602
H2A.SN.-.ST	TA-AA---A-----C-----AG-AA-G-----GCAA---G-T-----C---C---T-----C-A-----C-GG-G-----C-C-----A---CAG-C-----A	4047
H2B.GH.86.D205	---CCAA---AGCAA---T-AG-AA-G-C-----A-GGCA-A---G-A-G-----T-----g---C-G-T-T---C-GG-----CC-CC-----A---CAA-----A	4575
H2B.CI.-.EHO	---CC-A---A-GAG-----T-AG-AA-G-C-----AA-A---G-A-G-----T-----G-C-G-T-T---C-GG-----CC-CC-----G---CCAA-----T-A	4573
H2G.CI.-.ABT96	CCC---A---A-----T-AG-AA-R-T-----A---RCTC-A---G-A-G-----T-----T-----A-----T-C-G-G-----CCRTC---A---CCAG-----T-A	3929
H2U.FR.96.12034	CCCCAC---A-G-A---C-----AG-AC-T---GG-A---GCCA-G---G-T-GC-----T-T-T-G-T-----T-C-C-G-----C-T-G-G---CAG-----T---	4081
MAC.US.-.239	CAG-C---T-----AG-AA-G-T-----TC-G-AA-T---G-A-----T-----T-----A-----CC-G-A---CC-CC-----T---CAA-G-----T-A	4519
Pol1	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	p15
SMM.SL.92.SL92B	---TA---GGCA---T-T-A---AC-G-T-C-G-A---GCCA-A---CA-A-G-----T---T-T---C-A---G---CA-TG-G---CCGT---T---CAA-----A-A	3950
SMM.US.-.H9	TAGA---A-C-G-----A-AA-G-T-----GCAA-T---G-A-----T---T---RR-A---A---C-G-----C-TC-----T---CAA-A-----T-A	3992
STM.US.-.STM	CA-----A-----G-----AG-AA-G-T-----GCAA-T---G-A-----T-----T-----A-----C-GG-G-----C-T-G-A---CAG-----T-A	4166
SAB.SN.-.SAB1C	TA-CA---AC-A---T-----G-AC-----G-----GGCA-G---A-A-----T-C-T-----G-A-T-C---GG-A-T-----C---T---CAA-----A-A	4618
TAN.UG.-.TAN1	C-T-C---A-----T---AG-AA-G---T---T---A---TGCTA---T-AA-C-----T---G-A-G-----C-G-A---GC---A---CCAG-G-----T---	4419
VER.KE.-.AGM155	CCCCAC---AG-A-----CA---A-GG-C-G-A---GCCA-----CA-----T-T-T-----T-A-G-C-----G-A---C-----A---CAAG-----G-T---A	4439
VER.KE.-.9063	CCCCA---AG-A-G-C---CATTAA-GG---G---AAGGGCAA-A---C---ACA-----G---T-----T-A-G-----G-A---C-----A---CAA-----A-A	4447
VER.DE.-.AGM3	---CCC---A-AG-A---T---CA-TAA-GG-C---AC-TC---A-A---CT---CA-----G-----G-A---C-----G-GA-----C-----A---CAA-G-----G---	3942
VER.KE.-.TY01	---CCA---AG-G---T---CCTTAA-G---C---C---AA-A---T---CAG-----T-----A-----G-GA-----G---AAA-C-----T-A	3934
COL.CM.-.CGU1	---TCTAGTAGAGGGG-G---T---TC---GCAGGT---GGCT---A---GCCA-TC---G-A---C---TGT-----T-----G-A-----TGA---C---AAG---A-----G-A---A	3983
DEN.CD.-.CD1	TC---A---GG-A-G-C---C---AGCTA-GCA-GGC-----G-AA-A---CA-CA-C---T-----C-----A-----C---AC-G-A---C---GC---T---AAA-G-----A	4110
GRV.ET.-.GRI_677	---CCC---GG-A-G---C---AGCCC---G---AGGC---CT-ACA-----T-T-G-G-A---C---AC-G-A---C---G---A---CAA-----A---A-A	4393
GSN.CM.99.CN166	T-C-C---GG-A---C---C---AG-AC-GC-----GCCA---CT-AAGT---T-G-----G---A-----CAC-G-G-G---C-----CAAAG-----C-C	3948
GSN.CM.99.CN71	TCCC---GG-A---C---C---AG-AC-GT---GGG---A---GGCC---CT---AGT---T-----G-A-----CAC-G-G---C-----CAGA-G-----C-C	3930
DRL.-.-.FAO	TCCTA---G-----T---A---AA-G---C-GT---A---GCA-T---CT-AAAAT-----T-----T-----A-----T---GG-G-AG-C---C---T---AGA-----A-A	3787
RCM.GA.-.GAB1	---G-----G-CAG---T---A-T---A-G---GGG---A---GCA-G---C---TT-C---G-C---T-G-C-C---GG-----T---AAA-----A	3699
RCM.NG.-.NG411	CCCCC---AC-A-G-T---A---AA-G---GGG---G-A-G---C---CT-C---G---T-----G---T-----C---G---G---G---CAG-----T---A	3714
MND-2.-.-.5440	CCCCA---T-GGG-----A-TCA-G---GGC-----G-A-T---CT-----C---C---G-T-A---T-----GA-----A---AAG-----A-A	3781
MND-2.CM.98.CM16	CCCCA---A-GAG---T---A-TCA-G---CC-C---GCCA-A---C---A---T-----C-A---G---G---G---C---G---CAG-----G-A	4221
MND-2.GA.-.M14	CCCTA---T-GAG-----CA-C---AA-G---GCC-A---GCT---A---CT-A-----T-T-----C-A---C-----A---CAA-----A	4148
MNE.US.-.MNE027	CAG-C---T-C-----AG-AA-G-T-----AC-G-AA-T---G-A-----G-----T-A-----CC-G-A---CC-CC---T---CAA-G-----T-A	3987
LST.CD.88.447	TCCTG---A-A-G-T-----G-A-GT-A---G-A---CC-A-G---A-G-----T-----T-A-G---C-G---CC-CC-----AAA-----T-A	3491
LST.CD.88.485	CCCTG---G-G-G-T-----G-A-GT-AG-G---A---CC-A-G---A-G-----T-----T-A-G---C-G---CC-CC-----AAA-----T-A	3491
LST.CD.88.524	TCCTA---G-GA-C---A---G---C-GCAG-G-A---TGCA-T---C---A-G-----T-T-----T-A-G---CC-GG-GA---C---T-T-AAA-----T	3488
LST.KE.-.lho7	CCCCA---A-AG-C-C---A-A-GC-AGC-G-A---TC-A-T---A-G-----G-T-T-T-----A---T---C-G-G---CC-CC---A---AAA-T-----A-A	4575
SUN.GA.98.L14	---C---A---GG-G---CG---TCA---GGC-----CC---A---C---A-GT-----G-----A-----C-G-A---C---C---T---CCAG-----A	4574
MND-1.GA.-.MNDGB1	CCCCA---A-GG-----ATTAA-G-A-GG---A---AA-T---T-AAGT---C-----G-A---G---G---A---GC---A---T---A-----A	3907
MON.CM.99.L1	TCCCA---GG-GGCC---C---C---G-AC-CT-----CC---A-TCT---T-----T-----A-----G-----CGCC---GG-CCAG-----T---	3923
MON.NG.-.NG1	CCCCA---AG-GA---C---CC---C---CC-GG---AA-GC---A---T---T-----C-G-C-----C-A-G-----G---R---C-----A---YCAA-G-----A	2516
MUS.CM.01.1085	---CCA---AG-G-----C-G-C---C-GTC---A---GCCA-A---CT-AT-----G---C---T---G-T-A-----GG---G---C---A-A-AG-----	3910
DEB.CM.99.CM40	TA---A---AG-AA-GG---C---GCC-GCA-GGG---A---C-AA-T---A-----C---G---A-----C-G-----AAA-----A	4026
DEB.CM.99.CM5	T---A---AG-AA-G-C---C---AGCC---CA-G-----C---A---TAAAT-----C-----T-A---G---T-G---C---G---A---AGA-----A	4020
SYK.KE.-.KE51	CCCAG---G-A-G-T---GC-AA---GCT---GG-G-A---A---CA-AA-C---T-----C---T-----A-----C-G-----C---A---AAG-----T	4075
SYK.KE.-.SYK173	CCCCA---GG-G-G---C---AG-GCC-----GCCA-A---G-A-----C---C---T-----A-----GCT-T---C-----AAA-----A	4420



Pol p66 RT & Pol p15 RNase H end \ Pol p31 Integrase start		
H1B.FR.83.HXB2	.....AAAGTACTATTTTGTAGTGAATAGAT...AAGGCCCAAGATGAACATGAGAAA...TATCACAGTAATTGGAGAGCAATGGCTAGTGATTTTAACTGACCCTGTAGTAGCAAAAGAAATAG	4338
	... K V L F L D G I D . . . K A Q D E H E K . . . Y H S N W R A M A S D F N L P P V V A K E I	integrase
H1A1.UG.85.U455	.....G-G-G-----G-----T-----G-C-----A-----T-C-C-----G-----G-G-----	3784
H1B.US.90.WEAU160	.....G-----G-----G-----T-----A-----T-----C-----G-----G-----G-----G-----	4337
H1C.ET.86.ETH2220	.....G-G-C-----G-----T-----A-G-----A-----C-----A-----A-----TA-C-----C-----C-C-----	3730
H1D.CD.84.84ZR085	.....G-----G-----T-----G-----C-----C-----G-----G-----G-----G-----	3860
H1F1.BE.93.VI850	.....A---G---G---G---A---A---A---A---AC---G---T---A---T---A---G---G---G---G---G---G---	3678
H1G.SE.93.SE6165	.....C-----C-----A---A---G---G---AC---C-----G---C-----TT-----A-----G-----	3735
H1H.CF.90.056	.....G---C---G---A---T---A---A---GG---T---AC---G---T---A---A---G-----G-----	3685
H1J.SE.93.SE7887	.....G-G-C-----G-----A---T---A---T---A---T---C-----T---C-----T-----A-----G-----	3652
H1K.CM.96.MP535	.....G---C---T---A---A---A---T---AC---C-----T-----A-----C-----C-----	3534
H101.AE.TH.90.CM240	.....G-G-----G-----T-----A-----A---G---A---C-----A-----TT-----A-----G-----	3912
H102.AG.NG.-.IBNG	.....G-----C-----A---A---G---A---G---C-----AG-----T-----A-----C-----G-----	3863
H1N.CM.95.YBF30	.....A---T---CC-----T---A---A---A---T---C---G-----C-----A-----C-----TA-C---CA---G-----	3932
H1O.BE.87.ANT70	.....G---C---G---CC---A---C---A---A---T---A---T-----T-----A---C---A---GGA---A---G---G---C---G---CA	4393
H1O.CM.91.MVP5180	.....G---C---G---CC---G---A---C---A---A---T---A---T-----T-----T---A-----A---AA---C---G---CA	4368
CPZ.CD.-.ANT	.....C---C---C---G---A---C-----A---T---G---C---T-----T---C-----T---T---A---AGA---A---AC---T---T---C---A---T---G---T-----TA	3777
CPZ.CM.-.CAM3	.....T---C-----T-----T---CA---A---G---A---G---A---T---AC-----T-----A-----C---TA---A---A---G---C---C-----	3752
CPZ.GA.88.GAB2	.....G-G-----T-----A---A---C-----G---A---T---C---C-----C-----GA---A---CC---A---C-----G-----C-----	3731
CPZ.GA.-.CPZGAB	.....G---G---CC-----C---G---A---A---A---GG---T-----A---T-----TT---A---CA-----C-----	4399
CPZ.US.85.CPZUS	.....T---T---CC-----C-----A---A---G---A---C---C-----T-----C---T-----C-----A---G---C---T---G---G---	4398
CPZ.TZ.-.TAN1	.....A---T---CC---G---TA---G---A---A---G---A---C---T-----A---TT---A---GA---A---A---T---C---A---T---G---T-----TA	3982
H2A.GW.-.ALI	.....C-----T---G---C---AAA---G---CC---T---G---A---A---A---G---T---C---GT---A---A---C---AT---CCA---A---A---GGAT---A---CAA---C---G---G---G---G---G---	4713
H2A.DE.-.BEN	.....C-----T---CC---G---AAAG---G---CCA---T---G---A---A---T---C---T---ATA---A---C---AA---CCA---A---A---GGAA---A---C---T---C-----G---C---G---	4720
H2A.SN.-.ST	.....C-----T---CC---GAA---A---CCA---T---G---G---A---T---C---GTA---A---A---C---AT---CCA---A---A---GGA-----CAAAAT---G---G---C---G---	4165
H2B.GH.86.D205	.....C---G---CT---G---CC---AAA---A---CCA---G---A---G---A---T---G---C---GTA---A---A---C---TCCA---A---A---CGGA---T---AAT---G---G---C---G---	4693
H2B.CI.-.EHO	.....C---A---C---C---AAA---A---CCA---T---A---A---A---T---A---GTA---A---A---C---TCCA---A---A---GGGA---T---AAT---G---G---C---G---	4691
H2G.CI.-.ABT96	.....C-----C---GR---AAA---A---CCT---T---A---A---A---R---T---T---GTA---A---AGT---AA---CA---A---G---GG---A---T---T---AAT---G---G---C---G---	4047
H2U.FR.96.12034	.....C-----G---C---AAA---A---CCA---T---G---A---A---C---T---C---GTT---A---AGC---ACAGCA---A---A---GGTA---A---AAT---C---C---C---G---	4199
MAC.US.-.239	.....C-----T---C---C---G---AAAG---G---CCA---A---A---A---T---T---C---T---GTA---A---A---T---TATTC---A---GGAT---A---CAGAA---G---C---G---C---G---	4637
Pol1	... 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	p31
SMM.SL.92.SL92B	.....C---G---C---G---CC---AA---T---G---CCA---A---A---A---C---C-----T---C---GTA---AG---AGT---A---CAG---A---A---A---C---AAT---T---C---G---	4068
SMM.US.-.H9	.....C---C---C---C---AAA---A---CCA---G---G---A---G---A---C---T---C---GTA---A---AA---T---TATTC---A---RGTT---A---TAGGC-----C-----	4110
STM.US.-.STM	.....C---G---C---CC---AAA---A---CCA---T---A---G---A---A---T---C---GT---A---AGC---A---TCTTCA---G---GG---A---A---AGGC-----GC-----	4284
SAB.SN.-.SAB1C	.....C-----CC---G---CA---T---A---G---A---A---A---G---T---C---TGCC-----GAGT---CAACAA---A---GG---T---A---TG---A-----	4736
TAN.UG.-.TAN1	.....C-----G---C---CA---G---A---G---A---G---ATC---T-----T---C-----CA---TTT---AAGAGA---CA---CGGTA---A---G---AC---G-----	4537
VER.KE.-.AGM155	.....G---A---T---G---CA---T---GCA---G---A---G---A---A---A---T---GG---C-----AATC---A---AGACACA---GGAT---AAA-----T-----T-----	4557
VER.KE.-.9063	.....G---T---A---GAA---A---G---A---T---A---C---G---T-----A---AATT---A---GGACACA---GGAT---A---AGA-----T---C-----G-----	4565
VER.DE.-.AGM3	.....A---TT-----AAA---A---G---A---G---G---A---GG---C---T---A---GAAC---A---AGACAC---A---GGG---A---AAA---T---G-----	4060
VER.KE.-.TY01	.....G---TT---C---AAA---A---G---A---T---AA---G---A---G---T---A-----A---AACC---A---AGA---ACA---A---GGG---T---AAA-----G-----	4052
COL.CM.-.CGU1	.....C---G---A---G---GGA---AA---A---GCA---AG---A---C---C---A---G---T---T---GT---CAGTATC---AAAAGAAC---A---GGGT---A---CA-----TT	4101
DEN.CD.-.CD1	.....C---A---C---C---ACA---TCCA...C---A---T---A---T-----T-----C---T---C-----A---AG---CAGAGA---A---G---A---GG---A---A---A---G---T---G---CA	4228
GRV.ET.-.GRI_677	.....C---GA---C---C---A---A---G---A---A---C---CA---G---C---T---AC-----AGT---TACAG---A---GGAT---A---TAA---A-----G-----	4511
GSN.CM.99.CN166	.....C---G---T---CC---G---AAAC---G---CCT---AGT---A---C---A---G---T-----AGTACC---AAGAGA---C---A---AC---AA---C---CG---AC---T---G---T-----	4066
GSN.CM.99.CN71	.....C-----C---CA---G---GAAC---G---CCT---AGT---G---C---A---T-----AGTATT---AAGGGA---C---A---AC---AA---C---GG---AC---T---G---T-----	4048
DRL.-.-.FAO	.....C---GT---CC---AAC---G---A---A---A---G---A---TTT---C---T---A-----T---CT---A---ACAA---A---GGTT---A---GGGA---G-----	3905
RCM.GA.-.GAB1	.....C---G---G---CC---G---A---A---A---A---A---A---G---T---A-----T---AGAA---C---AA---T---AAA---G-----	3817
RCM.NG.-.NG411	.....C---G---G---CC---C---G---GA---A---A---A---G---A---A---T---C-----AGAA---A---C---AA---A---T---AAA-----	3832
MND-2.-.-.5440	.....C---G-----G-----A---A---G---A---C-----A---GT---T---CAG---C-----A---C---CA---G-----G-----	3899
MND-2.CM.98.CM16	.....C-----G-----A---C-----A---A---A---G---T-----A---G---T---A---ACAA---C---CTGTA---A---AACA---G-----G-----	4339
MND-2.GA.-.M14	.....C-----CC---G---A---A---T---G---G---C-----C---T---A---T---AAG---CAG---A---G---A---A---C---A---G-----	4266
MNE.US.-.MNE027	.....C---C---C---C---G---AAAG---G---CCA---A---A---T-----C---T-----GTA---A---A---T---TATTC---A---GGAT---A---CAGAC---G---C---C---G---	4105
LST.CD.88.447	CAAAAAC-----G---C---GAAAG---G---CCC---A---G---A---GT---AGC---T---T---A---GCT---A---ATT---A---AAGAAA---A---C---A---A---C---A---G---T---GC-----	3615
LST.CD.88.485	CAAAATAC-----G---C---AAA---G---CCA---A---G---A---G---AGC---T---T---A---GCT---A---ATC---A---AAGAAA---A---C---A---C---A---G---T---GC-----	3615
LST.CD.88.524	ACAAAAC---G---CA---G---G---GAAG---A---CCA---TGT---A---G---GA---G---T---A---GCA---A---ACT---A---AAGAAA---G---A---GGGA---A---C---CA---G---T---GC-----	3612
LST.KE.-.lho7	CAGAAAAC---CA---G---C---AAAG---A---CCT---AGT---A---AGC---T---C---T---AC---GCA---AG---ATC---A---AAGAAA---A-----A---C---CA---G---C---C---T---	4699
SUN.GA.98.L14	AAGAGAC---G---A---G---A---AAAG---A---CCA---TGT---A---G---G---G---TC---T---AC---GCAGCCT---TT---CAGGAGATG---G---A---A---C---TA---G-----GC---G---	4698
MND-1.GA.-.MNDGB1	.....CT---CC---C---AAAT---A---CCA---A---G---A-----T---C---GAAGC---CA---T---AAGAGAGA---A---CC---T---A---G---C-----C---C---G---T---	4025
MON.CM.99.L1	.....C---C---G---A---G---AAATT---G---A---CCC---AAGG---A---T---A---G---T-----A---TACT---AAGAGA---ACC---ACC---A---T---TA---C---GT---G---C-----	4041
MON.NG.-.NG1	.....C---G---C---C---CA---G---GAAC---T---A---CCR---G---G---T---A---C-----C---TACT---CAGGA---AC---A---GG---A---T---A---AC---CC---C-----G-----	2634
MUS.CM.01.1085	.....C-----C---CA---G---ACAG---A---CCT---AA---G---A---T---A-----AGTACC---CAGGGA---A---A---ACC---AA---C---G---T---GT---C-----	4028
DEB.CM.99.CM40	.....GGA---C---C---G---AA---GG---G---G---A---T---G---C---A---C---TG---G---A---CA---T---AAGGGA---A---ACC---A---A---TA---AC---CA---T---GC---TA	4144
DEB.CM.99.CM5	.....G---A---C---C---G---AA---G---A---G---T---G---T-----TG-----A---CA---T---AGAGA---A---C---GT---A---TA---CT---A---C---GC---G---TA	4138
SYK.KE.-.KE51	.....A---G---C---G---AAC---CCCA...A---A---A---G---A-----C---GTAGA---TAC---AAGGCAA---A---C---T---A---CAGACAG---C---CT---	4193
SYK.KE.-.SYK173	.....A---C-----CC---AA---TCCC...C---A---A---G---A---T---A---GG---C---TCA---AT---GAGTATC---TAGGCAG---G---C---T---A---C---GACAG-----C---TA	4538



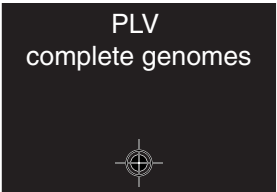
PLV  
complete genomes



	Asp64 catalytic site	
H1B.FR.83.HXB2	TAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAGGAAAAGTTATCCTGGTAGCAGTTTCATGTAGC	4468
Po1	V A S C D K C Q L K G E A M H G Q V D C S P G I W Q L D C T H L E G K V I L V A V H V A	p31
H1A1.UG.85.U455	-----A-----A-----G-----G-----T-----C-----A-----A-----C-----	3914
H1B.US.90.WEAU160	-----T-----G-----A-----A-T-----G-----T-----C-----G-----A-----C-----	4467
H1C.ET.86.ETH2220	---TT-----G-----A-----A-T-----G-----T-----C-----G-----A-----C-----	3860
H1D.CD.84.84ZR085	---T-----G-----A-----C-----G-----T-----C-----C-----G-----A-----A-----C-----	3990
H1F1.BE.93.VI850	---T-----G-----G-----C-----G-----T-----C-----C-----G-----A-----A-----C-----	3808
H1G.SE.93.SE6165	-G-----G-----G-----C-----G-----T-----C-----C-----G-----A-----A-----C-----	3865
H1H.CF.90.056	---T-----G-----G-----C-----G-----T-----C-----C-----G-----A-----A-----C-----	3815
H1J.SE.93.SE7887	---T-----G-----G-----C-----G-----T-----C-----C-----G-----A-----A-----C-----	3782
H1K.CM.96.MP535	---T-----C-----G-----A-----G-----T-----C-----C-----G-----A-----A-----C-----	3664
H101.AE.TH.90.CM240	-A--A-----A-----G-----T-----G-----T-----C-----C-----C-----A-----C-----C--C--G--	4042
H102.AG.NG.-.IBNG	-G-----C-----A-G-----G-----G-----G-----T-----C-----G-----A-----T-A-----C-----	3993
H1N.CM.95.YBF30	---C-----C-----G-----G-----CA-T-----G-G-----T-----C-----G-----A-----C-----T--G--G--C-----	4062
H1O.BE.87.ANT70	-T--T-----CC-----TA-----G-----A-T-----T-G-----C-----A-G-----A-----C-----A-G-----C--A--C--AA--A--T--T--C-----	4523
H1O.CM.91.MVP5180	-T--T--T-----CC-----C--TA-----G-----A-C-----T-----AC--C-----AG-----A-G-----C-----A-C--AA--A--T--T--C-----	4498
CPZ.CD.-.ANT	---ACAG-----CG--G-----G-----C-----G-----G-----C--C-----T-----A--A--A-----G-----T-----	3907
CPZ.CM.-.CAM3	---C-T-----C--A-----G-----G-----C-----C-----C-----T-----C--C-----CC-----G-----A-----G-----A-----	3882
CPZ.GA.88.GAB2	---CA-----T-----A--A-----G-----G--T-----C-----G--C-----G--G--C--C--C-----G--A--A-----G-----A-----G-----	3861
CPZ.GA.-.CPZGAB	-G-----CAT-----G--C--G-----G-----G-----C-----G--T-----G-----G-----C--C-----C-----G--A--A-----G-----C-----	4529
CPZ.US.85.CPZUS	-G-----C-----C-----G-----G-----A-----G-----G-----T--C-----T--C-----G-----G--A--C--TA--A--G-----A-----G-----	4528
CPZ.TZ.-.TAN1	-T--TCAG--CCA-----C--TA-----G--T--A-----G--G--AC-----A--C-----A--C-----C--CC-----G--C--A--A-----G-----	4112
H2A.GW.-.ALI	--AA--CA--CCC-----AG--G-----G--T--A-----G--GA--TGCAGAA--T--C--CT-----A--G--C--C-----C-----C--A--AA-----G-----T-----	4843
H2A.DE.-.BEN	--AA--TCA--CCC-----C--A--AG-----G-----A--A-----G-----A--TGCAGAAAT--CG--T-----A--G--C--AC-----C-----A--C--TA--A-----A-----T-----	4850
H2A.SN.-.ST	--AA--CA--CACCC-----AG--G--G--T--A-----G-----A--TGCAGAAAT--C--CT-----A--G--C--C-----C-----A--C--TA--A-----A-----T-----	4295
H2B.GH.86.D205	--AATTC-----C--A--A-----G-----T--T-----G-----A--TGCAGAC--T--G--C-----GA--G--C-----C-----A-----AA--A--G-----C-----	4823
H2B.CI.-.EHO	--AATTC-----C--A--A-----G-----T--T-----G-----A--T--CAGAA--T--G--C-----A--G--C-----C-----G-----G-----AA--A--G-----	4821
H2G.CI.-.ABT96	--AATTC--ACC--C--GT-----C--A--AG-----T-----G-----A--TGCAGAAATG--T--CT-----A--G--C--C--C-----G-----A--A-----G-----	4177
H2U.FR.96.12034	--AAT--CT--C--GTT--C-----AG-----G-----A--TACAGAAGT--G--C-----A--G--C-----C--C-----G-----G--A--C--G-----A-----G-----	4329
MAC.US.-.239	---A--C-----T--AG-----G--T--A-----G--G--C--A--T--CAGA--T--G--CT-----A--G-----C-----C-----G-----A--A--A--T-----A-----	4767
Po1	V D T C D K C H Q K G E A I H G Q A N S D L G G W Q M D C T H L E G K I I I V A V H V A	p31
SMM.SL.92.SL92B	---AATGCT--A--C-----C--A--AG--G-----G--A-----T--GAC--A--TGCAGAGGT--C--C-----GA--G--C--C--C-----G-----G-----A--A-----A-----G-----	4198
SMM.US.-.H9	---A--CA-----C--C--G-----Y--A-----G-----A--TGCGRRA--T--G--CT-----A--G--C-----CC-----G--C-----A--A--TA--A--T-----A-----G--G-----	4240
STM.US.-.STM	-G--AT--CA-----C--G--C--AG-----T--A-----G-----A--TGCAGAGTT--G--CT-----GA--G--C-----C-----C-----GA--A--TA--A-----A-----	4414
SAB.SN.-.SAB1C	---GGCA--CC-----C--A-----G--GT--TG--A-----G-----TGC-----G--G-----A--G-----CC-----A--C-----A--A--G--T-----T-----	4866
TAN.UG.-.TAN1	---TGCT--CCCA-----AA--TCG-----C--G--A-----G-----TGCC--GT--GG--C-----A--G--C-----G-----A--C--TA--A-----C-----G-----	4667
VER.KE.-.AGM155	---A--TG--CCCA-----AG-----G--C--A--A-----G-----TGC--TCA-----G--G-----GA--G--C--C-----A-----A--AG--GA--A-----G--C-----	4687
VER.KE.-.9063	---A--TG--CCA--TGCCA--AA-----G--C--A--A-----G-----GCCTCC-----GG-----A--G--C--C--G-----G-----A--AG--A--A-----G-----	4695
VER.DE.-.AGM3	---TG--CCA-----A-----G--C--AG--C-----G-----TGCTCG-----GG-----A--G--C-----C-----C--G--A--A--A-----C-----	4190
VER.KE.-.TY01	-G-----TG--CCA-----AA-----G-----C--AG-----G-----G--TGCTCA--T-----C-----GA--G-----T--C-----AA-----AG--A--A--T--G--C-----	4182
COL.CM.-.CGU1	GG-AA--A--TCAG--G-----AAC-----C-----TG--T-----C--GC-----T--A-----CTAT--CT-----T-----T-----GAG-----G-----T--A-----A--C--CTG	4231
DEN.CD.-.CD1	-TAATGT-----CCA-----AC--C--T--G-----C--A--CAGA-----G--G--TGCTC--T-----C-----GA--G-----C--C--C-----GAAT-----A--AA--A-----AA--C-----	4358
GRV.ET.-.GRI_677	---GGCA--CCC-----C--AA--G-----C--T--A-----G-----GCCTCCATT--A--CT-----GA--G--C--C--C-----AA--A-----A-----	4641
GSN.CM.99.CN166	-GAA--AA--TCC--G-----G--C--T--G-----C--A--AA-----G--G--A--TGCAGAAAT--C--G-----GA--G--C--C--G-----G-----G--C--T--A--G--G--C--C-----	4196
GSN.CM.99.CN71	-AA--AG--CTCC-----A--C--T--G-----C--A--AA-----G-----A--TGCAGAAAT--C-----A--G--C--C-----C--G-----C-----T--A--G--C--C-----	4178
DRL.-.FAO	-G--ACAG--CC-----AA--C--T-----GC--A--A-----C-----TGCA--C-----CT-----A--G--C--C-----GA--A--A--A--G-----C--C-----	4035
RCM.GA.-.GAB1	---ACAG--CCA-----AG-----G-----A--C-----G-----G--TGCA-----G--CT-----A--G--C--C-----G-----A--A--A-----C-----G-----	3947
RCM.NG.-.NG411	---CAA--CCCC--G-----G-----AG--A-----G-----TGC-----T--CT-----A--G--C-----C-----C-----A--C--A--A-----C-----	3962
MND-2.-.5440	-T--ACAA--CCA-----AC--G-----C--TG--T-----G--G--A--TGCAGA-----G--C-----GA--G--C--C--C-----A--C--A--A-----C-----	4029
MND-2.CM.98.CM16	---CAG--CCA--G-----AAC-----GC--A--A-----C--G-----TACATCC-----CC-----A--G--C--C-----A--G-----A--A--A--C-----C-----	4469
MND-2.GA.-.M14	-T--ACAG--CCCA--G-----G-----C--T--C-----G--G--TGCGCC-----C-----A--G--C--C--C-----G-----G-----A--A-----C-----G-----	4396
MNE.US.-.MNE027	---A--CA-----T--AG-----T--A-----G--G--A--T--CAGA--T--G--CT-----A--G-----C-----C-----G-----A--A--A--A--T-----A-----	4235
LST.CD.88.447	-GAAT--T--A-----AC--G-----AACA-----TAC--TCA-----T--T-----A-----C--C--T--CA--G--G--C--G--A--A--TAAT-----G-----	3745
LST.CD.88.485	-GAA--T--A-----AC--G-----AACA-----TAC--TCA-----T--T-----A-----C--C--T--CA--G--G--C--G--A--A--TAAT-----A--C-----G-----	3745
LST.CD.88.524	-AATGAG--C--G-----AG--G-----AACA--T-----TGTATCAGTG--T--T-----G-----T--CC-----GC--A--C--A--AAT-----C-----A--G-----	3742
LST.KE.-.lho7	-CAATGA-----CA--C-----AAAG-----AACA-----G--TGCTCAGTG--T--T-----G-----C-----T--C-----C-----G--C--TA--AAAT-----C-----G-----	4829
SUN.GA.98.L14	-CAATGAA-----CAC-----A--AG-----G-----AACA-----TGCTCAGTG--C--C-----GA-----C--C--CA--G--A-----C--AA--A-----G--G-----G-----	4828
MND-1.GA.-.MNDGB1	-GCAA--T--CAG--G--CTGT--ATC--T-----GC-----AA--G-----GAC--TG--TCA--TT--G--C-----GA-----C-----C--G--AATC--A-----TA--A-----G-----	4155
MON.CM.99.L1	-CAA--CAT--C--C-----AAACC--T-----C-----AATCA-----G--A--GCAGAAGT--G-----GA--G--C--C--C-----G-----GCA--A--TT-----A-----	4171
MON.NG.-.NG1	-CAA--CAT--CCCC--G-----CAC--C-----C-----AA-----G-----Y--CAGA--AT--G-----A--G-----C--T-----AT-----GTA--C-----A-----C-----	2764
MUS.CM.01.1085	--AA--CAT--CCA-----C-----AC--C--T-----C--A--AAACA--G-----CA--TGCAAGAT--GG-----A--G-----C--C--CC--G--G-----A--C--T--A-----C-----G-----	4158
DEB.CM.99.CM40	-A--ACAA--CCCC-----AC--G-----G--GC-----CA-----T--G--G--TGCTCAATG--C--C-----GA--G--C--C--C-----C--G-----AAC--A--A--AA--A--G-----C-----G-----	4274
DEB.CM.99.CM5	-T--GCAG--CCA--G--C--A-----G-----GC--T-----T-----TGCTCCATG--GG-----A-----C-----C--CC-----AAC--C--A--A-----G-----	4268
SYK.KE.-.KE51	--AA--TG--CCC-----C--AA--C--G-----C--A--AA-----G-----ACAGA--T--T--C-----A--G-----C--C-----A--TTGT-----AA--CAC-----	4323
SYK.KE.-.SYK173	-CAACAA--CCA-----AAAT--G-----C--A--AA-----G-----TGTAGA--AT--TAT--AT-----GA--G--C-----T--GAG-----C-----A--TGT-----AA--CAC-----	4668



		Asp116 catalytic site	
H1B.FR.83.HXB2	CAGTGGATATATAGAAGCAGAAGTTATTCCAGCAGAAACAGGGCAGGAAACAGCATATTTCTTTTAAAATTAGCAGGAAGATGGCCAGTAAAACAATACATACTGACAAATGGCAGCAATTCACCGGT		4598
Po1	S G Y I E A E V I P A E T G Q E T A Y F L L K L A G R W P V K T I H T D N G S N F T G		p31
H1A1.UG.85.U455	-----C-C-----C-----A-----G-----C--A-AC-----GT-----C-A--A-C		4044
H1B.US.90.WEAU160	-----C-----G-----G-----A-----C--A-C-----C-----A-----A--		4597
H1C.ET.86.ETH2220	-----C-C-----G-----G-----A-A-----C--AC-----G-----C-GGGT-----A-T-----T-C-----A--		3990
H1D.CD.84.84ZR085	-----C-----C-----G-----C-----C-----GT-G-G-----A-----A--		4120
H1F1.BE.93.VI850	T--AG-C-C-----A-----C--CA-A-----G-----T-----A-----A--		3938
H1G.SE.93.SE6165	-----C-----A-----C--A-A-----G-----C-GT-----A--T-----A--		3995
H1H.CF.90.056	-----C-----C-----AA-----C--C--G--G-----C--A-C-----GT-----A-----GA--		3945
H1J.SE.93.SE7887	-----C-----C-----A-----G-----T-----A-A-----C-----GT-----A-----C-----A--		3912
H1K.CM.96.MP535	T--C--C-----C-----G--A-A-----C--C--A-AC-----C-----GT-----A-----C-----A-C		3794
H101.AE.TH.90.CM240	-----C-----GC-----C-----G-----C-----GC-----C-----GT-----C-A-----C-T-----A-C		4172
H102.AG.NG.-.IBNG	-----C-----C-----A-----G-----C--A-A-----G-----GT-----C-A-----A--		4123
H1N.CM.95.YBF30	-----C--CT-----T-----G-----A-----A-----G-----T-----TA-----GT-----C--T-----ATC-----TA--		4192
H1O.BE.87.ANT70	A-----G-TC-----G-A-----A-A-----T--C--C--G-----C-G--T-C-----T--T--GT-----A-----C--GCCT-----T-AA--		4653
H1O.CM.91.MVP5180	A-----AC-T-----G-G-A-----A-----T--C-----C--G-----T--C-----GT-----A-----ACCT-----T-AA--		4628
CPZ.CD.-.ANT	T-----C-TC-----T-----A--GG--AT--G-----A--AAGT-----C--C--G-----CA-C-----C-----AGCT-----AA--		4037
CPZ.CM.-.CAM3	-----G--CC-----G-----G--C-----T-----G--T--A-----T-----A--C-----T-----GT--C--C-----T-----GCCT--C--T-----A--		4012
CPZ.GA.88.GAB2	A-----CT-----A-----T--G--T--A-----T-----A--C-----G-----G--AG-----C--A--T-----A--CA-----A--		3991
CPZ.GA.-.CPZGAB	-----C-----C-----T--G-----A--A-----T-----C--G-----T-----T--T--C--A--T-----GCCA-----T--AA--		4659
CPZ.US.85.CPZUS	-----CC-----G--C-----T-----A--A--G-----C-----A--C-----T-----GTG--T-----T-----ATCT--C--T--AA--		4658
CPZ.TZ.-.TAN1	T-----T-TC-----C--A-----A-----AAGAC-----C--T--C--CA-CC-----G-----T-----G--A-----A--T-----ACCA-----T--TA--		4242
H2A.GW.-.ALI	-----TC-----G--C--C--GCAG--T-----AAGAC-----CTC--C--A-----C--G--TA--T-----A--C--CACT--G--C--A--T--G-----TGT--C-----TCA		4973
H2A.DE.-.BEN	A-----TC-----C--C--CAG--T-----AAG--C--G-----CTC--C--A-----C--G--CA--T--G-----A--CGCACT--G--C--A-----CC--C-----TTCA		4980
H2A.SN.-.ST	A-----T-----C--C--CAG--T-----AAG--C--G-----CTC--C--AC-----C--G--CA--T--G-----A--C--CATT--G--C--A-----TGC--C-----TTCA		4425
H2B.GH.86.D205	-----G--T-----G--A--CCA--G-----AAGAC--G-----TCTC--C--AC-----G--G--CA--C-----TA--C--C--CACC--C--A-----C--TGC--C-----TCA		4953
H2B.CI.-.EHO	-----TC-----A--A--CCA--G-----AAGAC--G-----TCTC--C--G-----GC--G--CA--C-----TA--C--C--CACC--G--C--A-----C--TGC--C-----TTCA		4951
H2G.CI.-.ABT96	-----C--T--G-----A--C--CA-----AAG--C--G-----TA--C--G-----G-----TA--C--C--CACC-----A-----TGC--C-----TCG		4307
H2U.FR.96.12034	-----C--T-----G--C--GCAG--T-----TAG--C-----CCTG--C--AC-----GC-----G-----TA--C--CTCATT--G--C--A--T-----TGC--C-----TTCA		4459
MAC.US.-.239	T-----TC-----G--A--CA-----G-----AAGAC--G-----CTA-----G-----G-----TA--T--C--CATC-----C--A--T-----TGCT--C--TG--TTCG		4897
Po1	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 I T H L H T D N G A N F A S		p31
SMM.SL.92.SL92B	T-----C--T-----G--A--A--TAGG-----AAGAC--G-----TG--C--C-----A-----TC-----TA-----GCATC--G-----A-----C--TGC-----TCC		4328
SMM.US.-.H9	T-----TC-----G-----A--C--GCRG-----AAGAC-----CTG-----GC-----CA--C-----CA--C--C--CATC--G-----T-----TGC-----ATCA		4370
STM.US.-.STM	T-----G--T-----G-----A--C--CA-----AGAC--G-----TG-----G--G--G-----A--C-----T--C--C--CACC--G--C--G--T-----GCT--C-----TCG		4544
SAB.SN.-.SAB1C	T-----T-----A--C--CAG-----AA--A--CT-----C--C--C--GC--G--G-----CA--C-----CA--C--C--CAGC--G-----A--T-----T--C--C--T--TA--C		4996
TAN.UG.-.TAN1	-----C--TCT-----C--AG--TAGG-----AA--A-----A--A--C--G-----A--CATT--A--T-----TA--T--C--A--T-----GCCA-----TGTTA--		4797
VER.KE.-.AGM155	-----G--T-----G-----C--TAGG-----AA--A--G-----A--G--CT--G-----A--AT-----CA--C--CTCACC--C-----A--T-----ACCA-----TTC--		4817
VER.KE.-.9063	-----TC-----C--A--TAGG-----AAGA--G-----CA--A--C--C-----A--TTGA--T-----A--GTCCAGT-----C--A--T-----ACCA-----T--TTCC		4825
VER.DE.-.AGM3	-----TC-----A--TAG-----A--A-----A--A--T--A-----GA--CT--A--T-----CA-----C--CA--C--G-----A-----ACC-----T--GTC--		4320
VER.KE.-.TY01	-----TC-----C--A--TAGG-----AA--A-----G--A--G-----A--A--CTGA--T-----TA--C--CAGT-----C--A-----GCCT--C--T--TCC		4312
COL.CM.-.CGU1	T--C--CTG--TCTGCTGG-----ACCA--AT--AAAGAG-----G--A-----GTAGGGCC--CA-----G--TA--TCAG-----GAG--T--G--CA--G--T--C--A--T-----GCCA-----TGTA--		4361
DEN.CD.-.CD1	-----G--CC-----GA--A--A--T--G-----G--G-----CA--A-----CTA-----T--AC--GC--GA-----T--CT-----A--C--CC--A--T-----C--A-----ACCA-----T--T--A--		4488
GRV.ET.-.GRI_677	-----TC-----G--G--C--AG-----T--A-----G-----C--C-----GC--G--C--GTT--C-----G-----CATC-----C-----T-----CCA--C--T--TC--		4771
GSN.CM.99.CN166	A-----G--G--CTGG--TAGGA--A--A--CCAG-----CAGAC-----T-----TGAAG--C--G-----C--C-----T-----CCCACC-----A--T-----CC-----GT--TCC		4326
GSN.CM.99.CN71	-----G--G--TTGG--TAGGA--A--CCA-----CAGAC-----T-----TAAAG--A--G-----C--C-----C--C-----CTCACC-----A--T-----CC-----T--TC--		4308
DRL.-.-.FAO	A-----C--C-----A--A-----AG--A-----T-----GC-----T--C-----TC--CACT--G--C--A--T-----ACCA-----A--		4165
RCM.GA.-.GAB1	A-----G-----A--A-----AA--A--G-----C--C--G--G-----C-----CC--AA-----G--AGC-----C--A--T-----AGCA-----AA--		4077
RCM.NG.-.NG411	-----C-----A--A-----G-----AA--A--G-----TC-----C--G-----C-----C-----G-----CATT-----C--A-----A--CA--C--T--AA--		4092
MND-2.-.-.5440	-----T-----G--A--A-----G-----AA--A-----GC--C--C--G-----GC-----T--G-----CATC-----C-----TCCT--C--TGT--A--		4159
MND-2.CM.98.CM16	A--A--G-----C-----A--A--CA-----AA--A-----C--C--C-----T-----CATC-----C--A--T-----CCT-----A--		4599
MND-2.GA.-.M14	-----C-----G--A-----AA--A-----C--C--GC-----C-----G-----CATT-----TCCA--C--TGT--A--		4526
MNE.US.-.MNE027	T-----TC-----A-----CA--G-----AAGAC--G-----CTA-----G-----G--A--C-----TA--T--CGCATC-----C--A--T-----TGC--C--T--TTCG		4365
LST.CD.88.447	-TC-----T--G--T-----A-----AT-----A--AAC-----C--C--C--G-----C--GTGCA--C-----C--GCA--G--C-----A-----ACCA--C--TGT--A--C		3875
LST.CD.88.485	-T-----T--G--T-----A-----AT-----A--AAC-----A--C--C--G-----C--GTGCA--C-----C--GCA--G--C-----A-----ACCA--C--TGT--A--C		3875
LST.CD.88.524	ATCA--C--TC--G--TT-----G-----AT-----A--AACT-----A-----T--G-----G--TGTA--C-----T-----CA-----T--C--A--T-----TCCA-----TGTGA--		3872
LST.KE.-.lho7	-TCA--C--TC--G--TT-----A-----AT--G-----AA--AAC--GT--C-----C--G-----G-----TGCA--T-----T-----CA-----TCCA-----TGTA--		4959
SUN.GA.98.L14	TTCA--G-----G--T-----AC--A--CAAT--GCA--A--AC-----CAC--GG--G-----GC--CTGT--CC--TG-----CAG-----A--T-----GCCA-----T--A--		4958
MND-1.GA.-.MNDGB1	TTCA--C--TC--GA--G-----AA-----T--AA--AA--G-----G--G-----G-----C--CA-----TA--T--GT--A--C-----C--A--T-----GCCT--C--T--TA--		4285
MON.CM.99.L1	-----T--G--TGG--CA--GA--CC--A-----AGG-----AA--ATGT-----GGATAGCC--G--GG--GC--G--C--CT--TG-----G--C--CTCAG-----A--T-----GCCA-----A--C		4301
MON.NG.-.NG1	T-----C--CG--TTGG--CA--A--C-----CAG-----A-----GCAT--GCC--G--GG--C--T-----CG--C-----A--C--TCATG-----C-----ACCA-----A--		2894
MUS.CM.01.1085	T-----G--C--GTGG-----A--GA--AT--G-----AG-----AG--C--G-----T--TAGGC--AC--GG--G-----TC--CT-----T--CG--C--CA-----C-----A--T-----GGCT--C-----T--A--		4288
DEB.CM.99.CM40	T-----G--CA--GA--CC--A-----AA--A--G-----TCTG--C--G-----C--G-----G--C-----CTCAG-----C--A-----ACC--C--T--T--A--		4404
DEB.CM.99.CM5	T-----T--G--GA--TA--G--CT--AT-----T--AA-----C--TA--CT--A-----C--G--G--G--C-----CA--T-----CAG-----C--A-----ACCA-----TAA--		4398
SYK.KE.-.KE51	-----G-----C-----A--TA--A--AT--GAA--AG-----A-----G--CTA--C-----GC--A-----C-----A--T-----CAT-----A-----TCCA--C-----T--A--		4453
SYK.KE.-.SYK173	T-----A--CA--GA--CT--AAA--AGG-----G--G--T-----TG-----CA--GC--A-----CA--T-----A-----CAG-----C-----T-----ACCA--C-----TGT--A--		4798



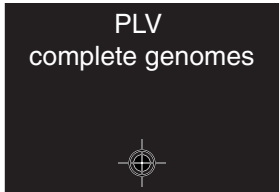
PLV  
complete genomes

Table with 3 columns: Accession ID, Nucleotide sequence, and Position. Includes a reference sequence for Glu152 catalytic site. Accession IDs include H1B.FR.83.HXB2, H1A1.UG.85.U455, H1B.US.90.WEAU160, H1C.ET.86.ETH2220, H1D.CD.84.84ZR085, H1F1.BE.93.VI850, H1G.SE.93.SE6165, H1H.CF.90.056, H1J.SE.93.SE7887, H1K.CM.96.MP535, H101.AE.TH.90.CM240, H102.AG.NG.-.IBNG, H1N.CM.95.YBF30, H1O.BE.87.ANT70, H10.CM.91.MVP5180, CPZ.CD.-.ANT, CPZ.CM.-.CAM3, CPZ.GA.88.GAB2, CPZ.GA.-.CPZGAB, CPZ.US.85.CPZUS, CPZ.TZ.-.TAN1, H2A.GW.-.ALI, H2A.DE.-.BEN, H2A.SN.-.ST, H2B.GH.86.D205, H2B.CI.-.EHO, H2G.CI.-.ABT96, H2U.FR.96.12034, MAC.US.-.239, SMM.SL.92.SL92B, SMM.US.-.H9, STM.US.-.STM, SAB.SN.-.SAB1C, TAN.UG.-.TAN1, VER.KE.-.AGM155, VER.KE.-.9063, VER.DE.-.AGM3, VER.KE.-.TY01, COL.CM.-.CGU1, DEN.CD.-.CD1, GRV.ET.-.GRI\_677, GSN.CM.99.CN166, GSN.CM.99.CN71, DRL.-.-.FAO, RCM.GA.-.GAB1, RCM.NG.-.NG411, MND-2.-.-.5440, MND-2.CM.98.CM16, MND-2.GA.-.M14, MNE.US.-.MNE027, LST.CD.88.447, LST.CD.88.485, LST.CD.88.524, LST.KE.-.lho7, SUN.GA.98.L14, MND-1.GA.-.MNDGB1, MON.CM.99.L1, MON.NG.-.NG1, MUS.CM.01.1085, DEB.CM.99.CM40, DEB.CM.99.CM5, SYK.KE.-.KE51, SYK.KE.-.SYK173.





Accession	Sequence	Accession	Sequence	Accession	Sequence	Accession	Sequence
H1B.FR.83.HXB2	ATCAGGCTGAACATCTTAAAGACAGCAGTACAAATGGCAGTATTCATCCACAATTTTAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAAGA..ATAGTAGACATAATAGCAACAGACATACA	D O A E H L K T A V Q M A V F I H N F K R K G G I G G Y S A G E R . . I V D I I A T D I Q	4855	Pol1	-----A-----	-----A-----	p31
H1A1.UG.85.U455	-G--A-----C-----T-----A-----	-----A-----	4301	H1B.US.90.WEAU160	-----A-----	-----A-----	4854
H1C.ET.86.ETH2220	-A--A---G--C-----T-----G-----	-----A--T-----T-----	4247	H1D.CD.84.84ZR085	-----A-----	-----A-----	4377
H1F1.BE.93.VI850	-----A-----T-----	-----A-----	4195	H1G.SE.93.SE6165	-----A-----T-----T-----	-----A-----	4252
H1H.CF.90.056	-C--A--A---C-----T-----A-----	-----A-----	4202	H1J.SE.93.SE7887	-A--A---C-----A-----	-----A-----	4169
H1K.CM.96.MP535	-----C--A-----A-----	-----A-----	4051	H1L.CM.96.MP535	-----C--A-----A-----	-----A-----	4429
H101.AE.TH.90.CM240	-G--A-----C-----A-----T-----	-----A-----	4429	H102.AG.NG.-.IBNG	-----A-----T-----T-----	-----A-----	4380
H102.AG.NG.-.IBNG	-----A-----T-----T-----	-----A-----	4380	H1N.CM.95.YBF30	-A--A---A-----G--G--T-----C-----	-----A-----	4449
H1N.CM.95.YBF30	-A--A---A-----G--G--T-----C-----	-----A-----	4449	H1O.BE.87.ANT70	-C--A--A---CT-A-GA-----TG-T-----C-----A--G--	-----A-----T-----T--C-A-----	4910
H1O.BE.87.ANT70	-C--A--A---CT-A-GA-----TG-T-----C-----A--G--	-----A-----T-----T--C-A-----	4910	H1O.CM.91.MVP5180	-C--A--A---G--T-A-A-----C--TG-T-----C-----G--	-----C--A-----C-----T--C-A-----	4885
H1O.CM.91.MVP5180	-C--A--A---G--T-A-A-----C--TG-T-----C-----G--	-----C--A-----C-----T--C-A-----	4885	CPZ.CD.-.ANT	-C--A--A---AT-A-----GT-----GCA--T-----CAC-T--AC-G--	-----T-----C-----	4294
CPZ.CD.-.ANT	-C--A--A---AT-A-----GT-----GCA--T-----CAC-T--AC-G--	-----T-----C-----	4294	CPZ.CM.-.CAM3	-A--A--A---CT-G-G-----C--G-----G--T-----C-----	-----TA-----C-----	4269
CPZ.CM.-.CAM3	-A--A--A---CT-G-G-----C--G-----G--T-----C-----	-----TA-----C-----	4269	CPZ.GA.88.GAB2	-C--A--A---CT-A-----T-----T-----C-----A-----	-----GC-----T-----	4248
CPZ.GA.88.GAB2	-C--A--A---CT-A-----T-----T-----C-----A-----	-----GC-----T-----	4248	CPZ.GA.-.CPZGAB	-A--A---T-A--A-----G--T-----C-----	-----A-----	4916
CPZ.GA.-.CPZGAB	-A--A---T-A--A-----G--T-----C-----	-----A-----	4916	CPZ.US.85.CPZUS	-A--A--G--T-A-----T-----G-----T-----	-----C-----A-----T-----AC-----	4915
CPZ.US.85.CPZUS	-A--A--G--T-A-----T-----G-----T-----	-----C-----A-----T-----AC-----	4915	CPZ.TZ.-.TAN1	-C--A--A--G--A--G-G-----ATC-----G-AT-----A--C-----	-----C--T-----C--A-T--A-T-----	4499
CPZ.TZ.-.TAN1	-C--A--A--G--A--G-G-----ATC-----G-AT-----A--C-----	-----C--T-----C--A-T--A-T-----	4499	H2A.GW.-.ALI	-G----AA-TACAG-GG-A--AT---T-----TCATTG-ATG-----GG-A--A--A--ATATG-CCC--CA-----	-----C--CA-CA-T--G--CT-C-----ACA-G--	5230
H2A.GW.-.ALI	-G----AA-TACAG-GG-A--AT---T-----TCATTG-ATG-----GG-A--A--A--ATATG-CCC--CA-----	-----C--CA-CA-T--G--CT-C-----ACA-G--	5230	H2A.DE.-.BEN	-A----AA-TACAA-AG-A--AT---TG-----TCATTG-ATG-----GG-A--A--A--ATATG-CCC--CA-----	-----C--A-CA--G--TA-C-----ACA-G--	5237
H2A.DE.-.BEN	-A----AA-TACAA-AG-A--AT---TG-----TCATTG-ATG-----GG-A--A--A--ATATG-CCC--CA-----	-----C--A-CA--G--TA-C-----ACA-G--	5237	H2A.SN.-.ST	-G----AA-CACAG-AG-A--AT---T-----TCATTG-ATG-----GG-A--A--A--ATATG-CCC--CA-----	-----C--A-CA-T--GG-CA-TG-----ACAGG--	4682
H2A.SN.-.ST	-G----AA-CACAG-AG-A--AT---T-----TCATTG-ATG-----GG-A--A--A--ATATG-CCC--CA-----	-----C--A-CA-T--GG-CA-TG-----ACAGG--	4682	H2B.GH.86.D205	-C--A--A--T-TCAA-AG-----TT---T-----ACTCA-TG-ATG-----GG-A--A--A--ATATG-CCC-T-CA-----	-----C--TA--G--A-C-----GCA-G--	5210
H2B.GH.86.D205	-C--A--A--T-TCAA-AG-----TT---T-----ACTCA-TG-ATG-----GG-A--A--A--ATATG-CCC-T-CA-----	-----C--TA--G--A-C-----GCA-G--	5210	H2B.CI.-.EHO	-----A-T-TCAA-AG-----TT-GTT-----ACTCA-TG-ATG-----GG-A--A--A--ATATG-CCC-T-CA-----	-----C--A--G--A-T-----ACA-G--	5208
H2B.CI.-.EHO	-----A-T-TCAA-AG-----TT-GTT-----ACTCA-TG-ATG-----GG-A--A--A--ATATG-CCC-T-CA-----	-----C--A--G--A-T-----ACA-G--	5208	H2G.CI.-.ABT96	-----A-AA-TTCAA-AG-A--CAT---GTT-----CATTG-ATG-----GG-A--A--A--ATATG-CCC-T-CA-----	-----A-TA--G--TA-T-----ACA-G--	4564
H2G.CI.-.ABT96	-----A-AA-TTCAA-AG-A--CAT---GTT-----CATTG-ATG-----GG-A--A--A--ATATG-CCC-T-CA-----	-----A-TA--G--TA-T-----ACA-G--	4564	H2U.FR.96.I2034	-A----AA-TACAA-GG-----CAT---C-T-----CTCA-TG-ATG-----GG-A--A--A--ATATG-C-C-T-CA-----	-----C--A-TA--G--T-C-----ACA-G--	4716
H2U.FR.96.I2034	-A----AA-TACAA-GG-----CAT---C-T-----CTCA-TG-ATG-----GG-A--A--A--ATATG-C-C-T-CA-----	-----C--A-TA--G--T-C-----ACA-G--	4716	MAC.US.-.239	-A--A--AA-TTCAG-AG-A--CAT---TT-----TCATTG-ATG-----GG-A--A--A--ATATG-C-C--CA-----	-----T--A-TA--G--CA-T-----ACA-G--	5154
MAC.US.-.239	-A--A--AA-TTCAG-AG-A--CAT---TT-----TCATTG-ATG-----GG-A--A--A--ATATG-C-C--CA-----	-----T--A-TA--G--CA-T-----ACA-G--	5154	Pol1	E O A N S V E T I V L M A V H C M N F K R R G G I G D M T P A E R . . L I N M I T T E Q E	-----L I N M I T T E Q E	p31
Pol1	E O A N S V E T I V L M A V H C M N F K R R G G I G D M T P A E R . . L I N M I T T E Q E	-----L I N M I T T E Q E	p31	SMM.SL.92.SL92B	-A--A--A--GTCAAG-AG---CCT---T-----C-CATTG-ATG-----GG-A--A--A--ATATG-C-C-T-CA-----	-----A-T--G--CA-C-----AC-G--	4585
SMM.SL.92.SL92B	-A--A--A--GTCAAG-AG---CCT---T-----C-CATTG-ATG-----GG-A--A--A--ATATG-C-C-T-CA-----	-----A-T--G--CA-C-----AC-G--	4585	SMM.US.-.H9	-A--A--AA-TTCAA-AG-A--TAT---T-----A-TCAATTG-ATG-----GG-A--A--A--ATATG-C-C-CA-----	-----T--CA-T--G--TA-C-----ACA-G--	4627
SMM.US.-.H9	-A--A--AA-TTCAA-AG-A--TAT---T-----A-TCAATTG-ATG-----GG-A--A--A--ATATG-C-C-CA-----	-----T--CA-T--G--TA-C-----ACA-G--	4627	STM.US.-.STM	----A--AA-CACAG-AG-A--T-T---TC-----TCA-TG-ATG-----G-A--GG-A--AC-A--ATATG-CCC--CA-----	-----T--TA-T--G--CA-C-----ACA-G--	4801
STM.US.-.STM	----A--AA-CACAG-AG-A--T-T---TC-----TCA-TG-ATG-----G-A--GG-A--AC-A--ATATG-CCC--CA-----	-----T--TA-T--G--CA-C-----ACA-G--	4801	SAB.SN.-.SAB1C	-AG-T--A--AGAT-GG-A-----CAT---GCA--T-----TC-----CA-----C--A-CA-T-----CAT---AC--G--	-----C--A-CA-T-----CAT---AC--G--	5253
SAB.SN.-.SAB1C	-AG-T--A--AGAT-GG-A-----CAT---GCA--T-----TC-----CA-----C--A-CA-T-----CAT---AC--G--	-----C--A-CA-T-----CAT---AC--G--	5253	TAN.UG.-.TAN1	-GG-CTG-C-TT-T-GG-A--G---C-TC-----CTGCCA---T-----G--A--A--A--TATG-CCC--CA--G--G--	-----C--A--A--G--A--C---C-AC--G--	5054
TAN.UG.-.TAN1	-GG-CTG-C-TT-T-GG-A--G---C-TC-----CTGCCA---T-----G--A--A--A--TATG-CCC--CA--G--G--	-----C--A--A--G--A--C---C-AC--G--	5054	VER.KE.-.AGM155	--G-CTG-C-T--ACAG-A-----TT-----CTGCCA---T-----G--A--A--A--CTA-CA--T-CA--G--	-----C--A--A-T--G--A---C-AT--G--	5074
VER.KE.-.AGM155	--G-CTG-C-T--ACAG-A-----TT-----CTGCCA---T-----G--A--A--A--CTA-CA--T-CA--G--	-----C--A--A-T--G--A---C-AT--G--	5074	VER.KE.-.9063	--G-TTG-C-GT--ACAG-A-----TT-----TGCCA-----G--A--A--A--TA-CAC--CA--G--	-----C--A-TA-T--G--A---C-AT--G--	5082
VER.KE.-.9063	--G-TTG-C-GT--ACAG-A-----TT-----TGCCA-----G--A--A--A--TA-CAC--CA--G--	-----C--A-TA-T--G--A---C-AT--G--	5082	VER.DE.-.AGM3	--G-CTG-C-T-CACAG-A-----TT-----TGCCA-----G--A--A--A--TA-CAC-G-CA--G--	-----T--A-CA-T--G--TA-T---C-AT--G--	4577
VER.DE.-.AGM3	--G-CTG-C-T-CACAG-A-----TT-----TGCCA-----G--A--A--A--TA-CAC-G-CA--G--	-----T--A-CA-T--G--TA-T---C-AT--G--	4577	VER.KE.-.TY01	--G-TTGCC--T--ACAG--G-----TG-----TTGCA-AC-T-----G--A--A--A--AC-G-C-T--CA--G--	-----C--A-TA-T-----A---C-AT--G--	4569
VER.KE.-.TY01	--G-TTGCC--T--ACAG--G-----TG-----TTGCA-AC-T-----G--A--A--A--AC-G-C-T--CA--G--	-----C--A-TA-T-----A---C-AT--G--	4569	COL.CM.-.CGU1	----A--A--ACAT-AG-A-GCAA---GC-----C-ATGCG-T-----G--A--AT-A--A-G--C--T-----CAAG--A-AGAGC--AT-GA--	-----T--G--	4618
COL.CM.-.CGU1	----A--A--ACAT-AG-A-GCAA---GC-----C-ATGCG-T-----G--A--AT-A--A-G--C--T-----CAAG--A-AGAGC--AT-GA--	-----T--G--	4618	DEN.CD.-.CD1	-AG-T--A--AGAT-AG-A-----C--GC-----ACCC-----T-TT-----GG-A--AC-A--ATATG-C--CA-----TACA-TA-----GC--AAT--C-IT--G--	-----T--G--	4745
DEN.CD.-.CD1	-AG-T--A--AGAT-AG-A-----C--GC-----ACCC-----T-TT-----GG-A--AC-A--ATATG-C--CA-----TACA-TA-----GC--AAT--C-IT--G--	-----T--G--	4745	GRV.ET.-.GRI_677	--G-TTG--GAGAT-GG-----G-----TACGCAT-----G--A--A--A--TAT-TC-AGT-CA-----T-G--TA-T--GC--A--C-AC--G--	-----T--G--TA-T--GC--A--C-AC--G--	5028
GRV.ET.-.GRI_677	--G-TTG--GAGAT-GG-----G-----TACGCAT-----G--A--A--A--TAT-TC-AGT-CA-----T-G--TA-T--GC--A--C-AC--G--	-----T--G--TA-T--GC--A--C-AC--G--	5028	GSN.CM.99.CN166	-AG--TCAC-T-CT-AG-A-----GC-CA--TGC-A--T-T-----G--A--A--A--ATATGT-CC-TACA-----	-----A-T--G--TCAC--AT--G--	4583
GSN.CM.99.CN166	-AG--TCAC-T-CT-AG-A-----GC-CA--TGC-A--T-T-----G--A--A--A--ATATGT-CC-TACA-----	-----A-T--G--TCAC--AT--G--	4583	GSN.CM.99.CN71	-G--TCAC-T-CT-AG-A-----GC-CAA--TGC-A--T-T-----G--A--A--A--ATATGT-CC-TACA-G-G--	-----T--GA--G--CCAC--AT--G--	4565
GSN.CM.99.CN71	-G--TCAC-T-CT-AG-A-----GC-CAA--TGC-A--T-T-----G--A--A--A--ATATGT-CC-TACA-G-G--	-----T--GA--G--CCAC--AT--G--	4565	DRL.-.-.FAO	----A--A--AG-----T-----C-G--T-----CA-----	-----T-----CT-T---C-CTT	4422
DRL.-.-.FAO	----A--A--AG-----T-----C-G--T-----CA-----	-----T-----CT-T---C-CTT	4422	RCM.GA.-.GAB1	-A--A--A--A-AT-AG-A-----C-----T--GG-T-----A-----	-----A--T-----T--GC	4334
RCM.GA.-.GAB1	-A--A--A--A-AT-AG-A-----C-----T--GG-T-----A-----	-----A--T-----T--GC	4334	RCM.NG.-.NG411	----A--A--GT-AG-----T-----GC--T-----A-----G-----	-----A-----T-----T-----	4349
RCM.NG.-.NG411	----A--A--GT-AG-----T-----GC--T-----A-----G-----	-----A-----T-----T-----	4349	MND-2.-.-.5440	----A-----AGG--AG-A-----G-----GC-A--T-----C-T--A-----A-----A-C-----C-T	-----A-C-----C-T	4416
MND-2.-.-.5440	----A-----AGG--AG-A-----G-----GC-A--T-----C-T--A-----A-----A-C-----C-T	-----A-C-----C-T	4416	MND-2.CM.98.CM16	-C--A--A--GA-A--AG-A-----T-----A--T-----C--A-----T-----C-C-T	-----C-C-T	4856
MND-2.CM.98.CM16	-C--A--A--GA-A--AG-A-----T-----A--T-----C--A-----T-----C-C-T	-----C-C-T	4856	MND-2.GA.-.M14	----A--A--GAGA--GG-----C-----A-----A-----A-----G-T-----A-----T--T	-----T--T	4783
MND-2.GA.-.M14	----A--A--GAGA--GG-----C-----A-----A-----A-----G-T-----A-----T--T	-----T--T	4783	MNE.US.-.MNE027	-A--A--AA-TTCAA-GG-A--CAT---TT-----TCATTG-ATG-----GG-A--A--A--ATATG-C-C--CA-----	-----G--T--C-TA--G--CA-T-----ACA-G--	4622
MNE.US.-.MNE027	-A--A--AA-TTCAA-GG-A--CAT---TT-----TCATTG-ATG-----GG-A--A--A--ATATG-C-C--CA-----	-----G--T--C-TA--G--CA-T-----ACA-G--	4622	LST.CD.88.447	-AG-T--C--AT-A-----T--C-T-----AC-CA--T--T-----CA--GG-A--AC-A--CT-TCAC--CA-----	-----G--T-TA-TA-T--G--TAATG--T--G--	4132
LST.CD.88.447	-AG-T--C--AT-A-----T--C-T-----AC-CA--T--T-----CA--GG-A--AC-A--CT-TCAC--CA-----	-----G--T-TA-TA-T--G--TAATG--T--G--	4132	LST.CD.88.485	-AG-T--C--AT-A-----T--C-T-----CA--T--T-----CA--GG-A--AC-A--CTTTCAC--CA-----	-----G--T-TA-TA-T--G--TAATG--T--G--	4132
LST.CD.88.485	-AG-T--C--AT-A-----T--C-T-----CA--T--T-----CA--GG-A--AC-A--CTTTCAC--CA-----	-----G--T-TA-TA-T--G--TAATG--T--G--	4132	LST.CD.88.524	-AG-T--AC-G--G--A-----GTT-----CCA--T--T-----TATCAC--CA--G-----TACA-CA-T--G--CAATG--TT--G--	-----T--G--	4129
LST.CD.88.524	-AG-T--AC-G--G--A-----GTT-----CCA--T--T-----TATCAC--CA--G-----TACA-CA-T--G--CAATG--TT--G--	-----T--G--	4129	LST.KE.-.lho7	-AG-T--AC-G--A--A-----T-----CCA--T--T-----CA--GG-A--AC-A--CTT-CAC--CA--G--	-----T--TA-TA-T--G--TAATG--AC--G--	5216
LST.KE.-.lho7	-AG-T--AC-G--A--A-----T-----CCA--T--T-----CA--GG-A--AC-A--CTT-CAC--CA--G--	-----T--TA-TA-T--G--TAATG--AC--G--	5216	SUN.GA.98.L14	-AG-T--AC-GG-AT-G--A--T-----T-----TC-GCA--T--T-----CA--GG-A--AC-A--AATG-CA--CA--G--G--	-----T--TA-CA--G--TAATG--TC--G--	5215
SUN.GA.98.L14	-AG-T--AC-GG-AT-G--A--T-----T-----TC-GCA--T--T-----CA--GG-A--AC-A--AATG-CA--CA--G--G--	-----T--TA-CA--G--TAATG--TC--G--	5215	MND-1.GA.-.MNDGB1	-AG--TTGCA--G-AT-A--A-----GCC-----CAC--T--T-----CA--GG-A--AC-A--ATG-CA-----A--G--	-----A-T--G--CAAT--AT--G--	4542
MND-1.GA.-.MNDGB1	-AG--TTGCA--G-AT-A--A-----GCC-----CAC--T--T-----CA--GG-A--AC-A--ATG-CA-----A--G--	-----A-T--G--CAAT--AT--G--	4542	MON.CM.99.L1	--G--T--T-C--CCCA-----T--GC-CAA--C-G-----TTA-----G--A--A--A--ATATGGCCC-T-TA--C--	-----T--TA--A--G--CAC--AT--G--	4558
MON.CM.99.L1	--G--T--T-C--CCCA-----T--GC-CAA--C-G-----TTA-----G--A--A--A--ATATGGCCC-T-TA--C--	-----T--TA--A--G--CAC--AT--G--	4558	MON.NG.-.NG1	--G-A-T-C-GTTC--CTCT-----C-GGC-CAAA-GC-C-A--T-TA-----RG--A--A--A--ATAT--C-ATA-G-G--	-----T--CA--G--TCAT--TT--G--	3151
MON.NG.-.NG1	--G-A-T-C-GTTC--CTCT-----C-GGC-CAAA-GC-C-A--T-TA-----RG--A--A--A--ATAT--C-ATA-G-G--	-----T--CA--G--TCAT--TT--G--	3151	MUS.CM.01.1085	-AG--TCACTT--T-GG-A-----GC-CAA--C-----TTTA-----AG--G--A--AG-G--ACCTATCAAAT-CT--G--G--	-----A-CA--G--CAC--AC--G--	4545
MUS.CM.01.1085	-AG--TCACTT--T-GG-A-----GC-CAA--C-----TTTA-----AG--G--A--AG-G--ACCTATCAAAT-CT--G--G--	-----A-CA--G--CAC--AC--G--	4545	DEB.CM.99.CM40	--G-T--A--GAGAT-GG-A-----GC-----TACC-A--AATT-----GG--A--A--A--ATATGTC-C-GATA-----	-----C-----A--G--A-----A--G--	4661
DEB.CM.99.CM40	--G-T--A--GAGAT-GG-A-----GC-----TACC-A--AATT-----GG--A--A--A--ATATGTC-C-GATA-----	-----C-----A--G--A-----A--G--	4661	DEB.CM.99.CM5	-AG-T--A--AGA--AG-A-----GC-----CACC-----TATT-----GG--A--A--A--ATATGGCA--ATA-G--	-----C-----A--G--A-T-----T--GG--	4655
DEB.CM.99.CM5	-AG-T--A--AGA--AG-A-----GC-----CACC-----TATT-----GG--A--A--A--ATATGGCA--ATA-G--	-----C-----A--G--A-T-----T--GG--	4655	SYK.KE.-.KE51	-AG-A-TA--AGAT-AG-----G--GGC-----CACG-----TTTA-----G--A--A--A--ACT-CAC--CA--T--G--	-----TATA-TA--GC--TAC--AC--	4710
SYK.KE.-.KE51	-AG-A-TA--AGAT-AG-----G--GGC-----CACG-----TTTA-----G--A--A--A--ACT-CAC--CA--T--G--	-----TATA-TA--GC--TAC--AC--	4710	SYK.KE.-.SYK173	-CG-C-TCAC-----T-GC-A-----GGC-----ACC-----TTA-----G--A--A--A--ATATCAC--A--G--	-----TACA-TA-T--GC--TAT--AC--	5055
SYK.KE.-.SYK173	-CG-C-TCAC-----T-GC-A-----GGC-----ACC-----TTA-----G--A--A--A--ATATCAC--A--G--	-----TACA-TA-T--GC--TAT--AC--	5055				





PLV  
complete genomes



H1B.FR.83.HXB2	AACTAAGAATTACAAAAACAAATT....ACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAAATCCACTTTGGAAAGGACCAGCAAGGCTCCTCTGGAAAGGTGAAGGGGCAGTAGTA	4979
Pol1	T K E L Q K Q I . . . . . T K I Q N F R V Y Y R D S R N P L W K G P A K L L W K G E G A V V	p31
H1A1.UG.85.U455	-----T-----G-C-CA-----A-A-----	4425
H1B.US.90.WEAU160	G---C-C-----C-----G-----T-----	4978
H1C.ET.86.ETH2220	G-----C-C-----C-----TT-----A-----G-C-TA-----C-A-A-----	4371
H1D.CD.84.84ZR085	-----G-----C-----C-----G-----A-----T-----	4501
H1F1.BE.93.VI850	-----G-----G-----C-----C-----G-----	4319
H1G.SE.93.SE6165	-----C-----C-----G-----G-----C-A-A-----	4376
H1H.CF.90.056	-----T-----C-----A-----G-C-A-----A-----	4326
H1J.SE.93.SE7887	-----G-----G-----A-----A-----A-A-C-----	4293
H1K.CM.96.MP535	-----TT-----C-----A-----G-A-A-----A-----	4175
H101.AE.TH.90.CM240	-----G-----C-----A-----A-----A-----	4553
H102.AG.NG.-.IBNG	-----C-----G-----T-----G-----CA-----A-A-----	4504
H1N.CM.95.YBF30	G-A-C-A-T-----C-----TT-----N-----C-A-----A-----G-----CA-----C-A-T-G-----A-----G-----	4573
H1O.BE.87.ANT70	-----A-C-----C-----TT-----N-----C-A-----C-----A-----G-C-TA-C-----G-----G-----G-----A-----G-----C-----	5034
H1O.CM.91.MVP5180	-----A-C-----C-----TT-----CA-C-----C-----A-----T-----G-C-TA-----G-----C-A-----G-----G-----C-----	5009
CPZ.CD.-.ANT	G-A-CTC-----T-----TT-----C-A-----C-----T-----G-C-TG-G-----C-A-T-G-----C-----	4418
CPZ.CM.-.CAM3	-----A-CTA-TC-----TT-----G-----C-----C-----T-----C-GA-----T-----G-----	4393
CPZ.GA.88.GAB2	-----AC-----C-----CT-----A-----A-----G-C-CA-----CTCT-----G-----G-----T-----	4372
CPZ.GA.-.CPZGAB	-----A-GC-----C-----TT-----G-----A-----A-----G-C-A-----G-----CCT-A-----G-----G-----	5040
CPZ.US.85.CPZUS	-----AG-CTT-----TT-----G-----C-----C-----A-----T-----G-----C-A-----T-----G-----	5039
CPZ.TZ.-.TAN1	G-A---C-----TT-----G-----T-----GC-----G-----A-----G-----C-----A-----G-----	4623
H2A.GW.-.ALI	-----TAC--TTCC-C-----C-A--A-----TTG--T-A-----C-C--C-----AG--G--AG-G-----T-GGG--A-G-----G--C--A--CA--	5354
H2A.DE.-.BEN	-----TAC--TTCC-C-----G-A--A-----T--TT-A-----C-A--C-----A--AG--G--AG-G-----T-GTG-A--A--G-----G--A--A--CA--	5361
H2A.SN.-.ST	-----TAC--TTCC-C-----GC-A--A-----T--T-A-----C-----T--A--AG--G--AG-G-----T-GGG-A--A--G-----G--G--C--A--CA--	4806
H2B.GH.86.D205	-----TAC--TTCC-C-----GC-A--A-----TT--T-----C-A--C-----A--AG--G--A--C-----G--T-GTG-A--AT-G-----G--A--CA--C-----	5334
H2B.CI.-.EHO	-----TAC--TTCC-C-----C-A--A-----TT--T-C-----C-----C-----A--AG--G--A--C-----G--T-GTG-T--AT-G-----G--A--CA--C-----	5332
H2G.CI.-.ABT96	-----TAC--TTTC-A-----C-A--A-----T--T-A-----C-----A--AG--G--C--AG-C-----T-GTG--TT-G-----G--A--T--CA--C-----	4688
H2U.FR.96.12034	-----TAC--TTCC-C-----TC-A--A-----T--T-A-----C-----A--AG--G--C-A--T-G-----C--G-CT-A--G-----G--A--A--G--CC-T-----	4840
MAC.US.-.239	G-TAC--TTTCA-----TC-A--AC-----T--T-A-----C-----A--AG--G--A--G-----G--C-GTG--AT-G-----G--A--CA--C-----	5278
Pol1	I Q F Q O S K N . . . . . S K F K N F R V Y Y R E G R D Q L W K G P G E L L W K G E G A V I	p31
SMM.SL.92.SL92B	-----CC--T-C--A--TTC-----A-----T--T-----A-----A--AG-A--G-C-A--G-----G-----G-----A-----T--T-----	4709
SMM.US.-.H9	-----TAC--TTCCA-----TC-A--A-----T--T-A-----C-----A--AG--G--C--AG-G-----C-GTG--AT-G-----G--A--CA--C-----	4751
STM.US.-.STM	-----TAC--TTTCA-----TC-A--A-----T--T-CA-----C-----AG--G--A--A--G-----C-GTG--AT-G-----G--A--CA--C-----	4925
SAB.SN.-.SAB1C	-----C--AC-C-T-----C--A-----T-----C-----AG--G--C--TG-G-----T--G--T-----A-----A-----T-----	5377
TAN.UG.-.TAN1	-----TAC--C-CA-----C--CAA-----CA-----TC-----AA--C--C-----A--AG-A--G-G--G-----C-CT-----G--A--A--G-----	5178
VER.KE.-.AGM155	-----TC--CACTC-----CCA-----C-----CA-----TTG-----A-A--C--C-----A--AG--G--G--G-----T--TCGC--GA-----A-----C--G-----T-----	5198
VER.KE.-.9063	-----CTAC--AC-----CCA-----CA-----TT-----A-A--C--C-----A--AG--G--G--G-----C--AT-AA-----G-----T-----G--C-----	5206
VER.DE.-.AGM3	-----TTAC--C-CC-----CCA-----CA-----TT-----A-A--C--C-----A--AG--G--G--TG-C-----G-C--T-AA-T-----G-----T-----G--C-----	4701
VER.KE.-.TY01	-----TAC--C-T-----CCA-----CA-----TT-----A-A--C--C-----A--AG--G--G--C--TG-G-----G--C--AT-AA-----G-----A-----G--C-----	4693
COL.CM.-.CGU1	T--AC--A-CC-----C-A--T-ACAAAATCA-----T--A-----AA--CC--C--GG-AA--GCA--...CAGG-GAG--C-----G-G-A--GG-G-----A--G-----C-----	4742
DEN.CD.-.CD1	-----TAC--C-TACCA-----C--CCA-----T--TCG-----AA--C--C-----A--GG-AGCT-----G-AA--C-----TG--G--T-----A-----T--C-----	4863
GRV.ET.-.GRI_677	-----CTA--TACTC-----C-----C-----CA-----TTG-----AA--C--C-----A--AG-T--G-----G--G-----C-GGA-----A-----A-----C--G--G-----	5152
GSN.CM.99.CN166	-----AC--C-C--A--C-C-----A-----T-C--T-----A-----C--AG-AGCT--C--T-----C-----GTA-----A-----A-----	4707
GSN.CM.99.CN71	-----AC--C-C--A--C-C-----G-----T-C--T-----A-----C--AG-AGCT--T-A--C-----GTA--T-----G-----A-----G-----	4689
DRL.-.-.FAO	-----A-GCA--C-----C--A-T-----CTC-----A--GG-A--G--AGT-G-----G-A--TG-----G--A--G--G--C-----	4546
RCM.GA.-.GAB1	-----C--TA-----T-----T-----A--AG-A--GG--A--G--G--T-----T--GA-----A-----A-----C-----	4458
RCM.NG.-.NG411	-----A-CCA-----C-----T-----T-----A--AG-A--GG--AG-AG-----G-----T--A--A--G-----A-----T--C--G-----	4473
MND-2.-.-.5440	-----A-CCA-----G-C--A-T-----T-----A-----A--AG-A--GG-C-A--AG-----G-----A--T-----A-----C--T--G-----	4540
MND-2.CM.98.CM16	-----A-CTA-----C-CA-T-----CA-----A--AG-A--GG-C-A--AG-----G--A--A--T-----A-----A--G--G-----	4980
MND-2.GA.-.M14	-----CCA-----C--A-T-----T-----A--AG-A--GG-G-AG--C-----G--A--A--T-----A-----A--G-----	4907
MNE.US.-.MNE027	-----TAC--TTCCA-----TC-A--AC-----T-----T--A-----C-----A--AG--G--A--G-----G-----C-GTG--AT-G-----G--A--CA--C-----	4746
LST.CD.88.447	-----AC--T-T-----T--A-----T-----TT--A--AA-----C-----AC-AG-AC--G-----C-AG--G-----T--C--T--G-----A-----T--T-----G-----	4256
LST.CD.88.485	-----AC--T-T-----T--A-----T-----TT--A--AA-----C-----AC-AG-AC--G-----T-AG--G-----T--C-AA-T--G-----A-----T--T-----G-----	4256
LST.CD.88.524	-----C--T-CA-----TT--A-----T-----TT--A--AA-----C-----AC-AG-A--G-----AG--G-----T--TC-A--T--G-----A-----T--T-----G-----	4253
LST.KE.-.lho7	-----AC--T-TC-----TT--A-----T-----TT--A--AA-----C-----AC-AG-A--G-----T-AG-----GC-A--T--G-----A-----T--T-----G-----	5340
SUN.GA.98.L14	-----AC--T-C-----G--AT--A-----T-----TT--A--AA--C--C-----A--AG-A--G--T-AG-----T--T--A--T--G-----G-----A-----C-----	5339
MND-1.GA.-.MNDGB1	-----TA-C--T-TCA-----T-----T-----A--TT-----AA-----C--T--A--AG-A--G--AG--G-----T-GT-TC--TT-G-----A-----A-----T-----	4666
MON.CM.99.L1	-----CTAC--ACC-C-A-C--C-----T-----T--TC--A-----T-----ACAG-TGCCG-C--T-AG-----GCGC--A-----T--C-----C-----	4682
MON.NG.-.NG1	-----CTAC--C-C-C-A-C--C-----T-----T--TCG--A-----T-----AACAG--GCTG-----TTCG-----G-----C--T--T-----G--G-----A--TT--G--G-----	3275
MUS.CM.01.1085	-----CTAC--C-C-----C-----AA-----TAC--T-----T-----ACAG-TTCTG-----GTC-----G-----C--CGTCT-A--G-----G-----A-----C-----	4669
DEB.CM.99.CM40	-----A--ACT-----C--A-----T-----C--TT-GGC--CAA-----C-----A--AG-TGCTG-----AC-----G-----CC--G--G-----A-----T-----G--G-----	4785
DEB.CM.99.CM5	-----AC--ACC-----C--A-----T-----C--TT-GG--ACAA-----A-----A--AG-AGCTG--C--AC-----G-----C--T-CCT-A--G-----A-----A--G--G--G-----	4779
SYK.KE.-.KE51	-----CTAC--...AATAC-C-CAC-----CA-----T--TCG-----AAA--A-----T--AC-TG-T...-C-T-TGA-----C--G-----TTCT--A--G-----G--A-----A-----G-----	4825
SYK.KE.-.SYK173	-----TTAC--C--AATAC--C-TC-----C-----T--TCG-----A-----C--C--G...G-G--GAACGAG-----G-----TCGT--A--T-----T-----G-----	5173





PLV  
complete genomes



H1B.FR.83.HXB2	.....AACATGGAAAAGTTTAGTAAAACACCA	5123
Vif	.....T W K S L V K H H	Vif
H1A1.UG.85.U455	.....C-----T--	4569
H1B.US.90.WEAU160	.....	5122
H1C.ET.86.ETH2220	.....T-----G----	4515
H1D.CD.84.84ZR085	.....C-----T--	4645
H1F1.BE.93.VI850	.....C-----T-T--	4463
H1G.SE.93.SB6165	.....C-C---G-----T--	4520
H1H.CF.90.056	.....C-----C-----GT--	4470
H1J.SE.93.SE7887	.....C-----GT--	4437
H1K.CM.96.MP535	.....T-T-----T--	4319
H101.AE.TH.90.CM240	.....C-----T--	4697
H102.AG.NG.-.IBNG	.....T-----T--	4648
H1N.CM.95.YBF30	.....C-----T--	4717
H1O.BE.87.ANT70	.....	5181
H1O.CM.91.MVP5180	.....	5156
CPZ.CD.-.ANT	.....C---G-----	4562
CPZ.CM.-.CAM3	.....CC-----TGT--	4537
CPZ.GA.88.GAB2	.....T-----GTTT--	4516
CPZ.GA.-.CPZGAB	.....T-----T-T--	5184
CPZ.US.85.CPZUS	.....C-G-----T-T--	5183
CPZ.TZ.-.TAN1	.....CC---T---T--	4767
H2A.GW.-.ALI	.....	5510
H2A.DE.-.BEN	.....	5517
H2A.SN.-.ST	.....	4962
H2B.GH.86.D205	.....	5490
H2B.CI.-.EHO	.....	5488
H2G.CI.-.ABT96	.....	4844
H2U.FR.96.12034	.....	4995
Pol p31 integrase end \		
MAC.US.-.239	AGGTGGCATAG...CCTCATAAAATATCTGAAATATAA.....	5434
Pol	E V A \$	
Vif	R W H S . L I K Y L K Y K	Vif
SMM.SL.92.SL92B	CGCTGGCATAG...CTGCATTAATTCATAAGTATAA.....	4868
SMM.US.-.H9	AAGTGGCATAG...CCTCATCAAACACCTGAAATATAA.....	4907
STM.US.-.STM	AGGTGGCATAG...CCTCATAAAACATTAAAGTATAA.....	5081
SAB.SN.-.SAB1C	AGGTGGACTAG.....CC--GTGA-AT-	5518
TAN.UG.-.TAN1	.....TGGAGAGCATATAG-C-AA-GGAAGGGCATAGT	5319
VER.KE.-.AGM155	AAATGGCAGGG...GATAGTTAGATATTGGATGAATAA.....	5357
VER.KE.-.9063	AAATGGCAAGG...AATAGTCAAATATTGGATGACTAA.....	5365
VER.DE.-.AGM3	AAATGGCAAGG...GATAGTGAGGTACTGGATGAGGAC.....	4860
VER.KE.-.TY01	AAATGGCAGGG...AATAGTCAGATATTGGATGAGACA.....	4852
COL.CM.-.CGU1	AGGGAATGGATTAA.....	4886
DEN.CD.-.CD1	GTGAGACTGACATATAA.....GATTGGGGAAACCAAATT-GA-CATTGG-TGAGCACTTTGA-ATC	5034
GRV.ET.-.GRI_677	AGATGGAGGGGATAGTGACTTACAAGATCAGGAATAA.....	5314
GSN.CM.99.CN166	GTAGACTGGTTGCTTAG.....AGCTACTAAACACCACATATGGTCCAGGCAAAACCCATTGTGTATATGTCCATCACTACCAATTGCAACACCA-...AGATT	4929
GSN.CM.99.CN71	GTAGAATGGCTGATTAG.....AGCAACTAAACATCATATATGGTCCAGGAAAGTCCATTGTGTATATCCACCACCTATCAACT-CAACATCA-...AGGTT	4911
DRL.-.-.FAO	.....-GG----CTCCC-GTGC---T-T--	4690
RCM.GA.-.GAB1	CA.....C-CTC---G-----T-T--	4608
RCM.NG.-.NG411	.....GCAG---C-TTCC-----GT-T--	4620
MND-2.-.-.5440	.....ACAGGACGTCAAGTTGAG-GAT---TTCAA-----T-T--	4687
MND-2.CM.98.CM16	.....ACATGCAGTAGA--G-----TTCAC-G-----T--	5124
MND-2.GA.-.M14	.....-A-----TTCA-----C---T--	5051
MNE.US.-.MNE027	AGGTGGCATAG...CCTCATAAAATATCTGAAATATAA.....	4902
LST.CD.88.447	.....G-GCCTAGT-GAGAAGT-CCT-G-TAT	4391
LST.CD.88.485	.....G-ACCT--T-GAGAA-T-CCT-G-TAT	4391
LST.CD.88.524	.....G-GC-TAGT-GAGAA-T-TCT-G-TAT	4388
LST.KE.-.lho7	.....-T-C-TA-T-GAAAAGT-TTT-G--AT	5475
SUN.GA.98.L14	.....AAGAGAATGGGAAGATAGAAGCAGTATAAGATAGTTAGAAATAGTATGGCTTATAGATAG--TTGCTGT-GAGAA-T-TTT-G-TTT	5525
MND-1.GA.-.MNDGB1	.....TCAGAGAATAGAAAAGTGGCAC---TT-GTAAG-AG-C-GATGGC	4819
MON.CM.99.L1	TGGCTAATTAG.....ATGCACTAAATGGCACATACTCACAGGAAGTGCACCCTTTGTCTATGTACATCACTACCA-CTGCACAATCAG...AGATT	4898
MON.NG.-.NG1	ATGGAATGGCTGATTAG.....AGCAACCAAATATAACATATATACAGGAAAGTGCACCCATTGTGTATGTACATCACTATCA-CTT-C-CACCCAG...-G-TT	3491
MUS.CM.01.1085	ATAGATTGGATAATTAG.....GGCAACAAAATACTACATCCTTTCCAGGAGAGCAGCCCTTTACATATGTGCATCACTATCAATTGTACCATCAG...AGGTT	4885
DEB.CM.99.CM40	.....-TT-GAT-GGT-G-GCAGTGCTATAA-	4926
DEB.CM.99.CM5	.....-TT-GAT--GT-G-GTAGTGCCATAA-	4920
SYK.KE.-.KE51	AA.....-TGCAGCAT-T-TTAA-	4966
SYK.KE.-.SYK173	AG.....-TTGCAGCAT-T-ATAA-	5314



H1B.FR.83.HXB2	TATGTGTGTTTCAGGGAAA...GCTAGGGGATGG...TTTTATAGACATCACTATGAAAGCCCTCATCCA...AGAATAAGTTCAGAAGTACACATCCCCTAGGG...GATGCTAGA...TTGGTA	5235
Vif	M Y V S G K . . A R G W . . F Y R H H Y E S P H P . . R I S S E V H I P L G . . . . D A R . . L V	Vif
H1A1.UG.85.U455	C-----C-CAA-----CAA-T-----TAGA-T-----G-----A-----A-----	4681
H1B.US.90.WEAU160	-----A-----A-----A-----C-----C-----A-----T-----A-----G-----A-----	5234
H1C.ET.86.ETH2220	-----C-----A-----G-----AT-----G-----T-----C-----AGA-----AGG-----T-----G-----AA-----	4627
H1D.CD.84.84ZR085	-----C-----A-----A-----AA-----C-----C-----C-----A-----T-----A-----C-----	4757
H1F1.BE.93.VI850	-----AA-----C-----AA-----C-----T-----C-----AGG-----G-----A-----A-----T-----A-----A-----	4575
H1G.SE.93.SE6165	-----C-----AA-----A-----C-----C-----C-----AGG-----G-----A-----C-----A-----	4632
H1H.CF.90.056	-----C-----A-----A-----A-----C-----G-----G-----C-----AA-----A-----A-----G-----T-----A-----G-----A-----C-----	4582
H1J.SE.93.SE7887	-----A-----AA-----AAA-----C-----G-----G-----C-----AA-----A-----A-----G-----T-----A-----G-----A-----	4549
H1K.CM.96.MP535	-----C-----A-----AA-----ACC-----A-----A-----AGG-----A-----A-----AGAG-----A-----	4431
H101.AE.TH.90.CM240	-----A-----C-----AA-----AAAAG-----T-----AG-----AGG-----C-----T-----A-----G-----A-----	4809
H102.AG.NG.-.IBNG	-----C-----AA-----A-----T-----T-----TAGG-----A-----G-----T-----G-----A-----	4760
H1N.CM.95.YBF30	-----G-----AAA-----G-----A-----AA-----A-----T-----T-----CA-----A-----C-----A-----G-----T-----G-----A-----G-----	4829
H10.BE.87.ANT70	-----CAGG-----TA-----A-----G-----A-----CGA-----AAC-----GG-----C-----T-----TC-----AGAA-----G-----G-----AGT-----T-----T-----G-----TG-----CCAT-----G-----A-----	5290
H1O.CM.91.MVP5180	-----A-----CA-----G-----TAA-----G-----CGC-----AAC-----CG-----G-----T-----TC-----AGGA-----A-----G-----C-----G-----CG-----GT-----T-----G-----CA-----A-----GAT-----A-----A-----G-----	5265
CPZ.CD.-.ANT	C-----A-----GG-----AGA-----TAAAGT-----CT-----AACC-----AAG-----T-----ATGA-----C-----AG-----A-----G-----AGA-----T-----A-----CCAACATTG-----AAA-----A-----	4680
CPZ.CM.-.CAM3	-----AAAA-----AA-----A-----A-----AC-----G-----C-----T-----TT-----T-----A-----G-----A-----A-----G-----A-----A-----G-----	4649
CPZ.GA.88.GAB2	-----CAAG-----AAA-----G-----C-----A-----G-----C-----T-----T-----CA-----T-----A-----G-----C-----A-----CA-----A-----A-----A-----A-----G-----	4628
CPZ.GA.-.CPZGAB	-----T-----AGG-----AAA-----G-----A-----A-----T-----C-----T-----TCA-----A-----A-----A-----G-----TGC-----T-----A-----TT-----CA-----AGAT-----T-----T-----A-----AA-----	5299
CPZ.US.85.CPZUS	C-----AGA-----AA-----AA-----AA-----C-----C-----G-----T-----G-----A-----A-----C-----G-----A-----TA-----A-----A-----A-----A-----A-----G-----	5295
CPZ.TZ.-.TAN1	-----C-----T-----ACAA-----CAAATGC-----TG-----AA-----AT-----AAG-----T-----CTGA-----ACA-----A-----G-----CGAGGG-----A-----A-----T-----ACA-----GAAAGAT-----A-----A-----G-----	4882
H2A.GW.-.ALI	-----ACA-----GATCTAGAAAAGGT-----TGC-----A-----GT-----CCC-----C-----TA-----G-----TGG-----ATGGGCATGG-----T-----G-----CTT-----CAGCAGG-----ATAT-----T-----CAA-----GAAGA-----TCATC-----A-----AG-----	5619
H2A.DE.-.BEN	-----ACA-----GACCTAGA-----AGGT-----CGC-----A-----GT-----CCC-----C-----A-----G-----TGG-----ATGGGCATGG-----T-----G-----CTT-----CAGCAGG-----ATAT-----CAA-----GAAAA-----TCATC-----A-----A-----G-----	5626
H2A.SN.-.ST	-----ACAGG-----GATCTAGA-----AAGGT-----TGC-----A-----GT-----CCC-----C-----TA-----G-----TGG-----ATGGGCATGG-----T-----G-----CTT-----CAGCAGG-----ATAT-----T-----AAA-----GA-----AA-----TCATC-----A-----AG-----	5071
H2B.GH.86.D205	-----ACAGG-----GAGTTGCAAAAGGTCTCT-----A-----GT-----CCC-----C-----A-----G-----T-----G-----ATGGGC-----TGG-----T-----G-----CTT-----CAGTAG-----A-----ATAT-----T-----C-----AAC-----A-----A-----GAGC-----TGCC-----A-----A-----	5599
H2B.CI.-.EHO	-----GACA-----GACTTGCAAAAGGTCTCT-----A-----GT-----CCT-----C-----TA-----G-----T-----G-----ATGGGC-----TGG-----T-----G-----CTT-----CAGTAG-----A-----ATAT-----T-----C-----GAAA-----A-----A-----GAGC-----CATC-----A-----A-----	5597
H2G.CI.-.ABT96	-----ACA-----GACCTACA-----AAGGC-----GTT-----A-----GT-----CCT-----C-----TA-----G-----T-----G-----ATGGGC-----TGG-----T-----G-----CTT-----CAGCAGG-----ATAT-----T-----ACA-----A-----A-----AAGC-----CAT-----A-----A-----	4953
H2U.FR.96.12034	-----CACA-----GAGCTACA-----AAG-----GT-----TGT-----A-----G-----T-----G-----ATGGGC-----TGG-----T-----G-----CTT-----CAGCAGG-----ATAT-----T-----AAA-----A-----A-----GATC-----CAT-----AG-----	5104
MAC.US.-.239	-----ACT-----GATCTACAAAAGTTTGC-----A-----GTGCCC-----TTTA-----G-----TCG-----ATGGGCATGG-----T-----G-----CCT-----CAGCAG-----AT-----T-----CA-----A-----A-----GA-----CCAT-----A-----A-----	5543
Vif	T K D L Q K V C Y V P H F K V G W G W . . W G C S R V I F P L Q . . . . E G S H L E	Vif
SMM.SL.92.SL92B	-----ACA-----G-----GAGCTAGA-----AA-----GCCTGC-----A-----GTGCCC-----A-----A-----T-----G-----ATGGGCATGG-----T-----AC-----C-----GCC-----TAGG-----GATTT-----CT-----G-----A-----G-----GA-----TCAC-----A-----AG-----	4977
SMM.US.-.H9	-----CACT-----GACCTACA-----AAGGCTTGT-----A-----GTACCC-----TA-----A-----TTG-----ATGGGCATGG-----T-----G-----CTT-----CAGCAG-----AT-----T-----GT-----A-----AGGCTCAT-----A-----	5016
STM.US.-.STM	-----CACT-----GAGCTA-----CAAGGCATGC-----A-----GTGCCC-----C-----TA-----G-----TTG-----GTGGGCATGG-----T-----G-----CTT-----CAGTAG-----GATTT-----T-----CT-----GCAA-----GA-----AAGC-----CACCA-----A-----A-----	5190
SAB.SN.-.SAB1C	CCACATGCA-----GTGTCC-----CAGTG-----TTCAT-----CGG-----A-----ACCCCT-----ACTA-----AAT-----ATGGA-----TGG-----TACTCCTA-----CA-----TGGGT-----A-----C-----GAA-----GAGCTCTAA-----AAA-----	5639
TAN.UG.-.TAN1	G-----AA-----CCACATGA-----CCGGTTACA-----AT-----ACA-----A-----CT-----TG-----CT-----C-----ATGTG-----ATGGGC-----TGG-----TAC-----C-----T-----AGTAG-----T-----C-----T-----CT-----A-----GA-----AAG-----AAAA-----A-----	5440
VER.KE.-.AGM155	-----AGGAA-----CT-----AA-----GAA-----ACA-----A-----TG-----T-----TC-----AATC-----CTTGGGCATGG-----TAC-----CT-----TGAGCAG-----TATGTA-----A-----C-----CCCA-----GAAG-----G-----GAAA-----CCAT-----	5460
VER.KE.-.9063	-----AGGAACCTTAA-----CAA-----ACATG-----TG-----CT-----C-----AATT-----ATTGGGC-----TGG-----TAT-----CT-----TG-----GCAG-----TACAT-----A-----A-----TTA-----AAG-----GAAA-----ACAG-----	5468
VER.DE.-.AGM3	-----TAGAAAATTA-----AC-----AAA-----A-----CGA-----TG-----CT-----C-----AATT-----CATGGGCATGG-----TAC-----C-----TGAGTAG-----TATG-----G-----A-----C-----ACA-----G-----AGTA-----ACAT-----	4963
VER.KE.-.TY01	-----AGAAA-----TT-----AA-----AAC-----A-----TG-----T-----C-----AATT-----CATGGGC-----TGG-----TAC-----CC-----TGAGTAG-----TATGTA-----A-----A-----A-----ACA-----G-----GGAAA-----ATGT-----	4955
COL.CM.-.CGU1	-----GAAGCAGTTTC-----GAATGCCAATAT-----ATTAT-----AGGA-----A-----ACAG-----G-----CACCCGAATGG-----TGGACAGA-----GGT-----AAAT-----GT-----TT-----T-----AGGG-----TCAGCTATTA-----GTTA-----AGCA-----	4998
DEN.CD.-.CD1	-----AATATACAA-----ACA-----TGAGTTA-----AA-----GACAGTGA-----ATACC-----TTTAGAATC-----CATGG-----ATGG-----TAT-----CT-----TGA-----A-----A-----G-----G-----GAGCTTACA-----A-----	5152
GRV.ET.-.GRI_677	-----ACAAAT-----CCT-----GAA-----ACAGCAT-----TGGC-----G-----TGCAATGG-----GTTTT-----T-----G-----CCTACGCC-----GT-----CATT-----CT-----TCA-----A-----A-----A-----GAT-----ACA-----	5411
GSN.CM.99.CN166	-----CAC-----AAACAGATTTCG-----TTA-----A-----T-----ATACTAGGAAGGTAG-----GG-----AG-----AGTA-----GCAA-----AT-----AT-----GAG-----C-----ATCCT-----TGGG-----CAA-----CAGT-----C-----GACC-----GCTAGCT-----	5047
GSN.CM.99.CN71	-----CAC-----AAACAA-----ATTTCG-----TTA-----ATTA-----ATACTAGGCAGGTAG-----G-----GG-----AGTA-----GC-----A-----AT-----ATT-----GAG-----CCATCCT-----TGGG-----CAA-----TAGT-----C-----GACC-----GCCAGTC-----	5029
DRL.-.FAO	-----A-----C-----AAGGG-----AA-----GTATTTAGAAA-----CAT-----A-----GCACCT-----CTT-----C-----ATGCTCA-----GG-----TGG-----T-----G-----CTCA-----AGTC-----GTGAC-----CT-----CAAA-----GACA-----AAG-----A-----A-----A-----	4808
RCM.GA.-.GAB1	-----CA-----CAAAGG-----AA-----GCA-----A-----AA-----AG-----GAG-----A-----GTGCCT-----CTTTA-----A-----T-----CCATGGGGATGG-----T-----GTGCA-----T-----A-----TT-----A-----GAA-----GAT-----A-----AG-----AAA-----	4729
RCM.NG.-.NG411	-----CA-----AAGGG-----AA-----GGAA-----AA-----AT-----GAG-----A-----GTACCC-----TTTA-----A-----T-----CCTTGGGGTGG-----T-----GTCTC-----C-----CT-----G-----A-----AATA-----CA-----AG-----A-----AAAG-----	4741
MND-1.-.5440	-----A-----AAAGG-----A-----CATTTAGAAAAT-----CAA-----A-----CCCC-----TTTC-----ATGCTCA-----GGATGG-----T-----G-----C-----CAC-----TC-----AAGA-----G-----CT-----TAA-----GAT-----G-----T-----A-----AG-----A-----T-----	4802
MND-2.CM.98.CM16	C-----AA-----AAAGG-----AA-----CATTGACAAA-----GAT-----AA-----GGC-----TTTC-----ATGCTCA-----GGATGG-----T-----G-----C-----CAC-----C-----AA-----T-----T-----T-----C-----AA-----GAA-----AACAG-----A-----C-----	5239
MND-2.GA.-.M14	C-----AA-----AAGGA-----AA-----CACTTAGATAG-----GAAC-----A-----TCCAC-----CTTTC-----ATGCTCA-----GGATGG-----T-----G-----CGCAC-----TC-----AAGATT-----CT-----CAAA-----GAT-----GCT-----A-----AG-----A-----T-----	5166
MNE.US.-.MNE027	-----ACT-----GATCTACAAAAGTTTGC-----A-----GTGCCC-----TA-----G-----TCG-----ATGGGCATGG-----T-----G-----CCT-----CAGCAG-----AT-----T-----CAAA-----A-----A-----AA-----CCAG-----T-----	5008
LST.CD.88.447	C-----G-----AGA-----CACATAAAG-----GACTCAA-----AT-----AT-----GTAAGC-----TACATGGAACAGGG-----T-----G-----TGGG-----ATTT-----TAT-----CTTA-----AATA-----TA-----GTG-----A-----TT-----A-----ACA-----GA-----CCTTA-----	4509
LST.CD.88.485	-----G-----AGA-----CACATAAAG-----ACTCAGTT-----AT-----GTAAGC-----TACATGGAACAGGG-----T-----G-----TGGG-----ATTT-----TAT-----CTTA-----AATA-----A-----GTA-----A-----TT-----A-----ATA-----GA-----CCTTA-----	4509
LST.CD.88.524	C-----AAAGGAA-----CATTTCAG-----AGTAAA-----AT-----AT-----GTTAG-----TACATGG-----ACAGGA-----T-----G-----ATGGG-----GTTT-----TAT-----CTTACAATA-----T-----TA-----T-----AT-----ACA-----GA-----CTTTA-----	4506
LST.KE.-.lho7	AG-----CAGGT-----AAACACAG-----GGCAAAAGAAAAT-----GTAGGC-----TACATGGAACAGG-----T-----G-----TTGGG-----ATTC-----TAT-----CCTA-----CATA-----T-----GTG-----A-----T-----A-----TCA-----GA-----CTTTA-----	5593
SUN.GA.98.L14	A-----GAAGAA-----GCATA-----G-----ACAAAAGAT-----AT-----GTA-----C-----ATGTATGGAACAGG-----ACTG-----ATGGG-----GTGG-----TAC-----C-----TACAATA-----A-----TATT-----T-----G-----ACT-----T-----GG-----C-----GTA-----	5643
MND-1.GA.-.MNDGB1	ATG-----GCCAC-----G-----AAT-----TGAG-----AAG-----AT-----T-----TGCC-----G-----CCT-----TTATG-----CTTATA-----G-----ATGG-----TAT-----CTT-----CAGTA-----GTG-----TAT-----A-----AAT-----AGG-----CATA-----TTA-----A-----	4940
MON.CM.99.L1	CTCAC-----GAACAA-----ATT-----CTCC-----CTTA-----ACCTAGGAA-----CAC-----CA-----G-----AGGAG-----GTC-----T-----GG-----AACATAT-----TTGGAG-----TAA-----ATTTATTGGGA-----TT-----AC-----AAC-----TG-----GACCTAGTGCAC-----C-----	5019
MON.NG.-.NG1	-----TC-----C-----ATGCAAGATT-----GCTTC-----TTA-----AGAT-----GGTA-----CACACA-----R-----GRRGG-----AC-----T-----GG-----AACATAC-----TRGAG-----TCA-----GATCC-----TTGGGA-----AYTACCAAT-----TA-----GGCCCGC-----CAC-----	3612
MUS.CM.01.1085	CTCCC-----GAGCAA-----ATC-----G-----ATAC-----ATT-----AC-----CAAAATACCCGGGA-----TCAGGTGCA-----T-----GAAAG-----AC-----TAC-----T-----GA-----CTGA-----TCC-----GTGGGATGTGACCAAT-----TG-----GACC-----GCAAGTTGC-----	5006
DEB.CM.99.CM40	A-----GAG-----AAAA-----ACAAA-----CCAAA-----AATTA-----ATACCTGTACAA-----GG-----T-----C-----T-----TTAAGGA-----CT-----TATTAT-----C-----CAAA-----A-----A-----GTAT-----CA-----CACA-----AAA-----A-----GGG-----GTA-----	5047
DEB.CM.99.CM5	A-----GAGGAAAG-----ATAA-----CCAAA-----AATTA-----GCAATAAAG-----GG-----T-----C-----T-----TTAAGGAGG-----TTCTAC-----CTCAGA-----A-----A-----GT-----T-----CG-----CACA-----GAA-----A-----GA-----A-----GTA-----	5041
SYK.KE.-.KE51	GG-----AC-----CAAG-----ATAAA-----GCAAAACAATTA-----AGAAAGCCACA-----A-----T-----C-----G-----TTA-----ATGGGAGTGGT-----G-----CTC-----AG-----C-----TGGG-----A-----A-----G-----A-----C-----A-----G-----GT-----TTA-----	5084
SYK.KE.-.SYK173	A-----CAC-----AAG-----ATAAA-----GCAAA-----AATTT-----AGAAAGCCACA-----A-----C-----TAGAATGG-----A-----TGGT-----CCTA-----GCC-----TGGACA-----A-----G-----A-----GA-----CTATA-----	5432



PLV  
complete genomes



H1B.FR.83.HXB2	ATAACAACATATTGGGGTCTGCATACAGGAGAAAGAGACTGGGCAT.....TTGGGTTCAGGGAGTCTCCATAG	5302
Vif	<u>I T T Y W G L H T G E R D W H . . . . . L G Q G V S I</u>	Vif
H1A1.UG.85.U455	G---G-----A-----C-----T-G-----	4748
H1B.US.90.WEAU160	-----A-----T---A-----T-----A-----	5301
H1C.ET.86.ETH2220	---A-----T---A-----T-----T-----	4694
H1D.CD.84.84ZR085	G-----A-----C-----C-----	4824
H1F1.BE.93.VI850	-----G---CC-----A-----C-----	4642
H1G.SE.93.SE6165	G-----G-----A-----A-----A-----T-G-----	4699
H1H.CF.90.056	---C---C---A-----A-----A-----A-C-----	4649
H1J.SE.93.SE7887	-----G-----G-----T---C-----	4616
H1K.CM.96.MP535	G-----T-----G-----G-----	4498
H101.AE.TH.90.CM240	---G-----A-----AG-----C-----T-----	4876
H102.AG.NG.-.IBNG	G---G-----C-----C-----T-G-----	4827
H1N.CM.95.YBF30	-C-GTC-T---G-AAACA---CAGTCT---C-A-A-T---A-----	4896
H1O.BE.87.ANT70	G---C-----AT-ATGC---G-----TGAA---A-T-G---TAGT---	5357
H1O.CM.91.MVP5180	G-C-C---AT-AATGC---G-----GGAA---C---A-T-G---TAGT---	5332
CPZ.CD.-.ANT	G-G---GTG-TC---AT-A--ATGT---C-CCCA---C---A-A-T-T-GAG---	4747
CPZ.CM.-.CAM3	G-----C-----AT-AATGC---C-GCAT-C---C---G-T---G-----	4716
CPZ.GA.88.GAB2	G-CTGT---C---AA-T-AATGC---CA-CCA---C-A-C-T---G-----	4695
CPZ.GA.-.CPZGAB	G-G---C---CCT-AAGCC---G-C---A-C-T---TC-----	5366
CPZ.US.85.CPZUS	---T-T---G---GT-A-A---CA-C---C---A-G-T---A-A---	5362
CPZ.TZ.-.TAN1	G-TTT-CAT-----AGCCTGT-----CCA-----C-A---T---CA-AGGAT---	4949
H2A.GW.-.ALI	---CAGG-----AAC...TA---CC---A-GA---TCTCCTTTATGCA...G-A...AGA---A	5683
H2A.DE.-.BEN	---CAGG-----AAC...TA---CC---A-GA---TCTCCTTTATGCA...G-A...AGAT--A	5690
H2A.SN.-.ST	---CAGG---C---AAC...TA---CC---A-GA---TCTCCTTTATGCA...G-A...AGAC--A	5135
H2B.GH.86.D205	G-CCA-GG---AAC...TA---CC---G-GA---TCTTGAGTTCCTATGCT...G-A...AGAC--A	5663
H2B.CI.-.EHO	G-CCA-GG---C---AAC...TG---CC---G-GA---TCTTGAGTTCCTATGCT...G-A...AGAC--A	5661
H2G.CI.-.ABT96	---CA-GGG---AAC...TTG---CCCT---AG-GA---TTAAGCCAGTATGCA...G-A...AGAT--A	5017
H2U.FR.96.12034	G-GCA-G-C---AA...TTG---CC---G-AG-GA---TATCCTTTATGCA...G-A...AGGT--A	5168
MAC.US.-.239	G-CA-GGG---CA...TTG---CC---A-GG---TCAGTACTTATGCA...G---AGG---A	5607
Vif	<u>V Q G Y W H . . . . . L T P E K G W L S T Y A . . . . . V . . . . . R I</u>	Vif
SMM.SL.92.SL92B	G-CA-GTG---C---AA...TG---CCC---A-GA---TTATCAAGCTATGCA...G-T...AGG--TA	5041
SMM.US.-.H9	G-CA-GG---AA...TTG---CC---GRA-GG---TCAGTACTTATGCA...G---AGA--A	5080
STM.US.-.STM	G-TCA-GGG---AA...TTA---CC---A-GA---TCAGTGAATATGCT...G---AGG--A	5254
SAB.SN.-.SAB1C	G-T-T-AT---CA...TTA---CCC---G-A-GA---TTAGAAACCTATGCA...ACA...GGA---	5703
TAN.UG.-.TAN1	G-GATTGT---C---CAC...TA---GCC---CAG-GA---TCTCTACCTATGCA...G-A-CAATA...AGT-TTGAGA	5513
VER.KE.-.AGM155	G-GGAT-TC---CA...TTAG-TCC-A---CA-GA---TCTCAACTTATGCA...G-A-AAATA...CAA-ATG-TA	5533
VER.KE.-.9063	---GACTGC---CA...TTA---CC---G---TCTCTACTTACGCC...A-A-AAATA...CAA-A-T-G-	5541
VER.DE.-.AGM3	G-GATCT---CA...TG---CC---AG-GA---TATCAACATATGCT...GA---GATA...CAG-ATC--A	5036
VER.KE.-.TY01	G-GACCT---CA...TTA---CC---GCA-GA---TATCCACATATGCA...G-A--ATA...CAA-ATG--A	5028
COL.CM.-.CGU1	---TACTTGGGCACT...TTA---CCC---T---T-TCATAGAGGACCACCAGTC...TCTT	5053
DEN.CD.-.CD1	G---AGCTC---CA...TTG---CCC---A-GA-AT-TGTCACAAGTAGCAATCAGCATAGCTTATGTA...CACC	5228
GRV.ET.-.GRI_677	GA-GTG-ATAT--ATCACAAC-TC---CCC---GA---TCTCAAGTCATGGA...G-A-GTTA...TC--ATTACC	5487
GSN.CM.99.CM166	TCCCAGT-CACC-ATTAC-AA--AG--TCAT--T--A--GTATACAACAGAAGCAA...C-A--AGCA...AGA...	5120
GSN.CM.99.CM71	TCCCG-T-CACC-ACTAC-AA--GG--TCAC--T--G--GT-TACAACAGAAGTAAG...C-A--AGCA...AGG...	5102
DRL.-.-.FAO	G-----GTCCTG---AAC...TC---CCC---C-A--GG---TTAAGTAAATATGCCATA...A-A---	4872
RCM.GA.-.GAB1	---C-C---AA-T---T---AG-GA---TA...GGGACATATGGG...G-A-A---CT	4793
RCM.NG.-.NG411	G-T-----A---C...AT---AG-GA---TTA...GGAACCTATGGA...G-A-AAATA-C-TA-ATAGATA	4817
MND-2.-.-.5440	---C-TC---CGTC-T-T-GAACTTG---CCC---A--GG---TGTCTCAATACGCCTTT...A-AG---	4869
MND-2.CM.98.CM16	---TC---CATCCT-T-GAACTTA---CCC---A-GA---TTATCAACTTATGCCATG...A-A---	5306
MND-2.GA.-.M14	---T-TT---TGCCCTAT-GAACTTG---CCC---A-GA---TTATCTCAATATGCAATC...A-AGC---	5233
MNE.US.-.MNE027	GA-GT-CA-GGG-ATT-GAATTTA---CC---GG---TCAGTACTTATGCA...G---AGA---A	5075
LST.CD.88.447	G-G-G...-GC-AT-CA-ACTTGG---C-G---GA---ATCTCTCAGTGGGCA...G-A---A---	4573
LST.CD.88.485	G---G...-GC-AT-CA-ACTTGG---C-G---GA---ATCTCTCAGTGGGCA...G-A---A---	4573
LST.CD.88.524	G---G...AT-AT-C-AT-TGG---C-GG---GA---AT-TCACAGTGGGCA...G-A---A---	4570
LST.KE.-.lho7	G-C-G...ATA-AT-CA-ATTTGG---C-G---G-GA---ATCTCACAATGGGCA...G-T---	5657
SUN.GA.98.L14	G---GG...ATC-AT--A-AC-TC---CC---C-A--GG-T-AT-AACCAGTGGGGA...-GT...AGT--G-	5707
MND-1.GA.-.MNDGB1	G-T-G...AGC-ATT-G-ATTTGCA-AT---GGT-GA---CTTAAGTACTTATGCAGTA...AGCA-A-AAGCA-T--TAGACC	5022
MON.CM.99.L1	TC-C---GTG-ACAACAAA--GG-CTACAC--T--A--GT-TATTGGATGCAAAAG...GTGGAACGGTCCCTTTAGATCCCCTGTGG...AATCAA-CCC...AGAGATA	5128
MON.NG.-.NG1	TCTC--T--GTCAAAT-GAA---AG--TACATC-T--A--GTATATTGGAGGAGCCAGGGTGCCACAGCAAGAGGGAGTAGTAGATCGAGTGTGGTTTCTTCTAGTCCCA-AAGACCG-A-CGG...	3739
MUS.CM.01.1085	TCT-A--GCACA--T-GAA--AT-CTTTATCCT--G--GTATATCTAAGGAGACAT...CCACAG...ACAG--A-CAGGATTACA	5091
DEB.CM.99.CM40	G--GT-GA--TACTAT-GTGCTTG---CCC---G-A-GA---TACCCACTATGGCAGTA...G-A---	5114
DEB.CM.99.CM5	G--GTGGA-ATA-TCT-GAGC-TGG-TCC---A--GG---ATGCCCTCTATGGCGGTA...GGA--CT	5108
SYK.KE.-.KE51	---T-T-A--TC-ATCA-AATTTG---CC---G-A--G-CT-TTAAGAACAGAGGGAGTA...GGAT--A	5151
SYK.KE.-.SYK173	TGG-TC--G-TC-ATCACAATTTGG---CCC---GA---TTACATATGCAGGGAATT...GGA	5496



	Thr96 phosphorylated	Cys114 needed for APOBEC 3G degradation	
H1B.FR.83.HXB2	AATGGAGGAAAAAGAGATATAGC..	ACACAAGTAGACCTGAAC TAGCAGACCAACTAATTCATCTGTATTACTTTGACTGTTTTCAGACTCTGCT..	ATAAGAAAGGCCTTATTAGGACACATAGT
Vif	E W R K K R Y S . . . . . T Q V D P E E L A D Q L I H L Y Y F D C F S D S A . . . . . I R K A L L G H I V		
H1A1.UG.85.U455	-----CTG-A- - - - -	-----T- - - - -	-----C- - - - -
H1B.US.90.WEAU160	-----C- - - - -	-----G- - - - -	-----T- - - - -
H1C.ET.86.ETH2220	-----ATTG-GA-C- - - -	-----GC-G- - - -	-----A-C- - - -
H1D.CD.84.84ZR085	-----G- - - - -	-----GC-G- - - -	-----A- - - - -
H1F1.BE.93.VI850	-----C-GGG-AG- - - -	-----G- - - -	-----G-C- - - -
H1G.SE.93.SE6165	-G-----C-G-GA- - - -	-----A- - - -	-----G-C- - - -
H1H.CF.90.056	-----ACTG- - - - -	-----G- - - -	-----GC-G- - - -
H1J.SE.93.SE7887	-----C-G-GA- - - -	-----G- - - -	-----GC-G- - - -
H1K.CM.96.MP535	-----CTG- - - - -	-----A- - - -	-----C- - - -
H101.AE.TH.90.CM240	-----C-G-GA- - - -	-----A- - - -	-----T- - - -
H102.AG.NG. . .IBNG	-----A-C-G- - - -	-----A- - - -	-----CT- - - -
H1N.CM.95.YBF30	-----ACT- - - - -	-----G- - - -	-----A-G- - - -
H1O.BE.87.ANT70	-----C-TT- - - -	-----G- - - -	-----AC- - - -
H1O.CM.91.MVP5180	-----CAAT-C- - - -	-----G- - - -	-----AA- - - -
CPZ.CD. . .ANT	-----ATGTGGA-AG- - -	-----T- - - -	-----A- - - -
CPZ.CM. . .CAM3	-----ATGTGG-TC- - -	-----T- - - -	-----A- - - -
CPZ.GA.88.GAB2	-G-----C-G-GA-C- - - -	-----A- - - -	-----T-GA- - - -
CPZ.GA. . .CPZGAB	-----CT-GGATCT- - -	-----G- - - -	-----G-T-CTTAC- - -
CPZ.US.85.CPZUS	-----AC-GGA-CT- - -	-----C- - - -	-----AG- - - -
CPZ.TZ. . .TAN1	-----AC-GGA- - - -	-----C- - - -	-----T- - - -
H2A.GW. . .ALI	CT---TAT-C-G-A-AG-TCT-G. . . . .	G-T-TAC- - - -	A-CTGT- - - -
H2A.DE. . .BEN	CC---TAT-C-G-A-AG-TCT-G. . . . .	G-T-TAC- - - -	A-CTGT- - - -
H2A.SN. . .ST	CT---TAT-C-G-A-A-TCT-G. . . . .	G-T-TAC- - - -	A-CTGT- - - -
H2B.GH.86.D205	C---TATG-G-G-AC-T-TAT. . . . .	G-T-ACA- - - -	TG-G- - - -
H2B.CI. . .EHO	C---TATG-G-G- - - -	G-T- - - -	ACT- - - -
H2G.CI. . .ABT96	CC---TAC-C-GA-A-T-TAT. . . . .	G-T-GACA- - -	A-GAC- - - -
H2U.FR.96.12034	CT---TAC-C-GAG-T-T-T-G. . . . .	G-T-ACA- - - -	TTGC- - - -
MAC.US. . .239	CC---TACTC- - - -	A-C-T-T-G. . . . .	G-T-ACA- - - -
Vif	T W Y S K N F W . . . . . G D V T P N Y A D I L L H S T Y F P C F T G G E . . . . . V R R A I R G E Q L		
SMM.SL.92.SL92B	CT---TACTC-G-A-A-T-T-G. . . . .	G-T-AC- - - -	A-TG- - - -
SMM.US. . .H9	CC---TACTCC-GA-AT-TCT-G. . . . .	G-T-C-ACA- - -	TTAT- - - -
STM.US. . .STM	CT---TAT-CC-GA-AT-TCT-G. . . . .	T-G-T-ACA- - -	CTGT- - - -
SAB.SN. . .SAB1C	G-AC- - - -	GG-GA- - - -	GGTTT- - - -
TAN.UG. . .TAN1	-T---CAA- - - -	CG- - - -	CA- - - -
VER.KE. . .AGM155	GCCTAGTA- - - -	TG- - - -	T-A- - - -
VER.KE. . .9063	G-GCAGTAGG- - - -	AT-TC- - - -	A- - - -
VER.DE. . .AGM3	GCAAT- - - -	G-TCCTT- - - -	G- - - -
VER.KE. . .TY01	GCAATTTAG- - - -	TCT- - - -	A- - - -
COL.CM. . .CGU1	-TGAATTAC-G-TAG-TA-GT-GGAA- - - -	G-TT-AG-TGGC- - - -	A-TAGGGTAC-G-CA-CATAAGGGGGA-G- - - -
DEN.CD. . .CD1	TCA-C-CAGGG. . . . .	T-G. . . . .	GCT. . . . .
GRV.ET. . .GRI_677	-TCAA-A-GG- . . . . .	AG. . . . .	G- - - -
GSN.CM.99.CN166	-GG-AGAC-T- - - -	TTT-G- - - -	C- - - -
GSN.CM.99.CN71	-ACAGAC-T- - - -	CAT-G- - - -	C- - - -
DRL. . .FAO	--CAT-TTC-GG--AG-T-T-T. . . . .	CIT-A- - - -	AAC-AC- - - -
RCM.GA. . .GAB1	T---TA- - - -	GGAGAT- - - -	GTA. . . . .
RCM.NG. . .NG411	-CA-AT-TG-TCCCT-T-CTT. . . . .	G-TA-T- - - -	AAC-G- - - -
MND-2. . .5440	--ATCAA- - - -	GG-A-AT- - - -	TT- - - -
MND-2.CM.98.CM16	--AT-A-TTC-CA-AC-CTA. . . . .	T-CA- - - -	CC-AC- - - -
MND-2.GA. . .M14	--AC-AC- - - -	GG-GAC- - - -	TAT. . . . .
MNE.US. . .MNE027	CC---TACTC- - - -	G-AC-T-T-G. . . . .	G-T-ACA- - - -
LST.CD.88.447	-----TTT-TGG-AC- - - -	CAA. . . . .	G-A- - - -
LST.CD.88.485	-----TTT-TGG-AC- - - -	CAA. . . . .	G-A- - - -
LST.CD.88.524	-----TCT-TGGAC-AC- - - -	CAG. . . . .	G- - - -
LST.KE. . .lho7	-----GTCT-T-GAGAC- - - -	CAA. . . . .	G-A- - - -
SUN.GA.98.L14	-----TTT-T- - - -	C- - - -	T- - - -
MND-1.GA. . .MNDGB1	C-CCCTTGG-G-AGAG-GGT-T. . . . .	G-GA-ACT- - - -	A-GG- - - -
MON.CM.99.L1	G-GA-GTC-G-CAT-G- - - -	GCT. . . . .	C-TC-GACG-G-C-G- - - -
MON.NG. . .NG1	--ATTGAT-G-TCT-G- - - -	GCT. . . . .	C-CT-C-ACA-A-TG- - - -
MUS.CM.01.1085	TC---TATTCT. . . . .	T-CT-GAG- - - -	CAC- - - -
DEB.CM.99.CM40	C- - - -	G-CA-A-ACAT-GAAG- - - -	G-C-ACA- - - -
DEB.CM.99.CM5	CT- - - -	G-GC- - - -	C-GGGT-GCAG- - - -
SYK.KE. . .KE51	TT---GCCC-TCC-TCTGGAT-GAGT- - - -	G- - - -	ACA-AC-AC- - - -
SYK.KE. . .SYK173	TTCAATACC- - - -	TG-ACC-GT-GAAC- - - -	TG-CT-ACA-A-C-G- - - -

PLV  
complete genomes

SLQYLAL needed for APOBEC 3G degradation

	Cys133 needed for APOBEC 3G degradation		SLQYLAL needed for APOBEC 3G degradation	
H1B.FR.83.HXB2	TAGCCCTAGGTGTAATATCAAGCAGGACATAAC		AAGGTAGGATCTCTACAATACTTGGCACTAGCAGCATTAATA	ACACCAAAAAAGATA...AAG 5520
Vif	S P R C E Y Q A G H N		K V G S L Q Y L A L A A L I	T P K K I K
H1A1.UG.85.U455	---A---		---T---	---G---
H1B.US.90.WEAU160	---TT---		---A---	---G---
H1C.ET.86.ETH2220	---C---		---T---	---G---
H1D.CD.84.84ZR085	---T---		---CT---	---G---
H1F1.BE.93.VI850	---A-C---		---C---	---T---
H1G.SE.93.SE6165	---T---		---T---	---G---
H1H.CF.90.056	---A-C---		---A---	---G---
H1J.SE.93.SE7887	---C-C---		---T---	---G---
H1K.CM.96.MP535	---T---		---CT---	---G---
H1O1.AE.TH.90.CM240	---A-G---		---C---	---T---
H1O2.AG.NG.-.IBNG	---A---		---G---	---T---
H1N.CM.95.YBF30	ATTA---		---C---	---G---
H1O.BE.87.ANT70	ACTGA-C-A---		---GA---	---CT---
H1O.CM.91.MVP5180	GCTGA-C-A---		---GA---	---T---
CPZ.CD.-.ANT	ATTGA-ATAC---		---GA---	---G---
CPZ.CM.-.CAM3	ATAT---		---G---	---AT---
CPZ.GA.88.GAB2	G-CAAAGGT---		---C---	---TC---
CPZ.GA.-.CPZGAB	AGCG---		---C---	---T---
CPZ.US.85.CPZUS	A-TT-C-AA---		---C---	---T---
CPZ.TZ.-.TAN1	ATTGA-ATTC---		---GA---	---G---
H2A.GW.-.ALI	ATTGT-CT-C---		---CC---	---G---
H2A.DE.-.BEN	ATTGT-CT-C---		---CC---	---A---
H2A.SN.-.ST	ATTAT-CT-C---		---CC---	---G---
H2B.GH.86.D205	ATTGT-CTAC---		---CC---	---G---
H2B.CI.-.EHO	ATTGT-CCAC---		---CC---	---G---
H2G.CI.-.ABT96	ATTGT-CT-C---		---CC---	---G---
H2U.FR.96.12034	ATTGT-CT-C---		---CC---	---G---
MAC.US.-.239	GCTGT-T-C---		---CC---	---G---
Vif	L S C C R F P R A H K Y		Q V P S L Q Y L A L K V V S	
SMM.SL.92.SL92B	GTTAT-CTAC---		---GT---	---CCA---
SMM.US.-.H9	GCTGT-T-C---		---CC---	---AG---
STM.US.-.STM	ACTGT-T-C---		---CC---	---AG---
SAB.SN.-.SAB1C	-CTTTGG...		---G---	---AT---
TAN.UG.-.TAN1	-TTGTGG...		---C---	---G---
VER.KE.-.AGM155	C-T-TTC...		---GG---	---T---
VER.KE.-.9063	CGT-TTC...		---GG---	---C---
VER.DE.-.AGM3	C-C-TTC...		---GG---	---C---
VER.KE.-.TYO1	-GT-TTC...		---GG---	---C---
COL.CM.-.CGU1	G-AT---		---GG---	---G---
DEN.CD.-.CD1	CCTGAGC-CT---		---CCA---	---AG---
GRV.ET.-.GRI_677	---CGTGG...		---CC---	---CA---
GSN.CM.99.CN166	CCTAGGA-AT---		---C---	---A---
GSN.CM.99.CN71	CTTAGGATAT---		---C---	---A---
DRL.-.-.FAO	-CAAG---		---G---	---T---
RCM.GA.-.GAB1	-CTAGTG...		---GA---	---G---
RCM.NG.-.NG411	GCTG-T...		---G---	---G---
MND-2.-.5440	GGTAG---		---GA---	---G---
MND-2.CM.98.CM16	A-CAGTG...		---G---	---G---
MND-2.GA.-.M14	AGTAG-G...		---C---	---G---
MNE.US.-.MNE027	GCTGT-T-C---		---CC---	---AG---
LST.CD.88.447	A-TGG-ATTC---		---TGT---	---TCTC---
LST.CD.88.485	A-TGG-A-TC---		---TGT---	---TCTC---
LST.CD.88.524	-CTG-AC---		---TGT---	---TCTC---
LST.KE.-.lh07	A-CTATGT-T---		---TGT---	---TCTC---
SUN.GA.98.L14	ACTA-A-GA---		---CT---	---TCTC---
MND-1.GA.-.MNDGB1	GTTAAAAGTT---		---CAAGGT---	---TCTC---
MON.CM.99.L1	CTTGGGACA---		---GA---	---C---
MON.NG.-.NG1	CCTAGGG-AC---		---A---	---CT---
MUS.CM.01.1085	-CTAGGGGAT---		---C---	---CT---
DEB.CM.99.CM40	ACTG-AACA---		---GGACC---	---CCT---
DEB.CM.99.CM5	ACTA-AACA---		---GGACC---	---CCC---
SYK.KE.-.KE51	G-CTT-GCAC---		---C---	---CTT---
SYK.KE.-.SYK173	G-CAT-ACAC---		---C---	---CCC---





H1B.FR.83.HXB2	CCACCTTTGCCTAGTGTACGAAACTGACAGAGGATAG	5558
Vif	P P L P S V T K L T E D R	5557
H1A1.UG.85.U455	-----A---T-A-----	5004
H1B.US.90.WEAU160	-----A-----	5557
H1C.ET.86.ETH2220	-----GT---T-AGT-----A	4950
H1D.CD.84.84ZR085	-----C---A---G---A---A-----	5080
H1F1.BE.93.VI850	-----GC---CCA---AGT---C-----	4898
H1G.SE.93.SB6165	-----C---G---T-AG---A-----	4955
H1H.CF.90.056	-----GA---G---AGT-----	4905
H1J.SE.93.SE7887	-----CA---G---AGT-----	4872
H1K.CM.96.MP535	-----G---A---AGT-----	4754
H101.AE.TH.90.CM240	-----C---A---G---A---A-----	5129
H102.AG.NG.-.IBNG	-----G---GT-AG---A---C-----	5083
H1N.CM.95.YBF30	-----C-A-----G---T---G---A-----	5152
H1O.BE.87.ANT70	-----T---CC-A-C---CCA---T-A---A-----	5613
H1O.CM.91.MVP5180	-----T---CC-A-C---CCA---G-T-A---A-----	5588
CPZ.CD.-.ANT	AG--AG--C---G-T-T-TCT-A---A-----	5015
CPZ.CM.-.CAM3	-----C---C---T---GT-A---A-----	4972
CPZ.GA.88.GAB2	-----CC-A-----C---T-TCT-AG---A-----	4951
CPZ.GA.-.CPZGAB	-----A---CG-C---T-A-----	5622
CPZ.US.85.CPZUS	-----C-A-----G---T---T---A-----	5618
CPZ.TZ.-.TAN1	AGG--C---ATC---A---A---C-----	5220
H2A.GW.-.ALI	.....ATGGCA.....AACCCAGAGAAAACAGTACCACCAGGA.....AACAGTGGCGAAGAGACTATCGGAGAGGCCCTTCGAGTGGCTAG	5966
H2A.DE.-.BEN	.....ATGACA.....GACCCAGAGAAAAGGGTACCACCAGGA.....AACAGTGGAGAAGAGACCATTGGAGAGGCCCTTCGAGTGGCTAG	5973
H2A.SN.-.ST	.....ATGGCA.....GGCCCCAGAGAGACAATACCACCAGGA.....AACAGTGGCGAAGAAAACAGTACCAGAGGCCCTTCGAGTGGCTAG	5418
H2B.GH.86.D205	.....ATGGAT.....CCAGGGAGAGAGTACCACCAGGA.....AACAGCGACGAAGAAAACAGTAGGAGAAGCATTTCGCTTGGCTAG	5949
H2B.CI.-.EHO	.....ATGGAT.....CCAGGGAGAGAGTACCACCAGGA.....AACAGCGGCGAAGAAAACAGTAGGAGAAGCATTTCGAAATGGCTAG	5947
H2G.CI.-.ABT96	.....ATGGAT.....CCAAGGGAGAGAGTACCACCAGGA.....AACAGCGACGAGGAGACAATCGGAGAGGCCATTTCGAAATGGCTAG	5300
H2U.FR.96.12034	.....ATGGGA.....GATCCAAGAGAGAGAAATCCACCAGGA.....AATTGAGAAGAAGAACTGTAGGAGAGGCCCTTCAGTGGCTAG	5448
MAC.US.-.239	/ Vpx start	
Vpx	.....ATGTCA.....GATCCCAGGGAGAGAAATCCACCTGGA.....AACAGTGGAGAAGAGACAATAGGAGAGGCCCTTCGAATGGCTAA	5887
Vif	M S D P R E R I P P G N S G E E T I G E A F E W L	Vpx
	D V R S Q G E N P T W K Q W R R D N R R G L R M A K	Vif
SMM.SL.92.SL92B	.....ATGACA.....GACCCAAGGGAAAAGAAATCCACCAGGA.....AACAGTGGAGAAGAAAACAATAGGAGAGGCCCTTCGAATGGTTAC	5324
SMM.US.-.H9	.....ATGTCT.....SRTCCCAGGGAGAGRATCCACCTGGA.....AACAGTGGAGAAGAAAACRRTAGGCGAGGCCYTTCCGACTGGCTAC	5360
STM.US.-.STM	.....ATGTCA.....GATCCCAGAGAGAGAAATCCGCCCGGA.....AACAGTGGAGAAGAGGCCAATAGAGGAAGCATTTCGAGTGGCTAC	5534
SAB.SN.-.SAB1C	.....	5910
TAN.UG.-.TAN1	.....	5738
VER.KE.-.AGM155	.....	5740
VER.KE.-.9063	.....	5745
VER.DE.-.AGM3	.....	5243
VER.KE.-.TY01	.....	5235
COL.CM.-.CGU1	.....	5280
DEN.CD.-.CD1	.....	5450
GRV.ET.-.GRI_677	.....	5682
GSN.CM.99.CN166	.....	5336
GSN.CM.99.CN71	.....	5318
DRL.-.-.FAO	.....ATGGCAGAA...AGA...CAG...TCAGTGGAGAGAGCTCCA...GCGGAGCCAATGGGAGCAGGAGAGGTAGAGTTAGAAGAAATGGCTAC	5164
RCM.GA.-.GAB1	.....ATGGCAGAA...AGAGCCCCGAGGTCCCACTGGG...GCTGGAGAAGCA...GAGTTCCAGCCTTGGCTAC	5076
RCM.NG.-.NG411	.....ATGGCAGAG...GGCAGAGAAAAGAGTGGCCAGAGGCCCCCACTGGGGCTGGAGATGTA...GAGTTTGCCCCCTGGCTTC	5106
MND-2.-.-.5440	.....CTCAGAG-TA--A--CATGGCAGAG...AGG...GCACCAGAG...GCG...CCAGAAGGAGCAGGAGAGGTAGGACTGGAGCAATGGCTGG	5146
MND-2.CM.98.CM16	.....ATGGCAGAAGGG...GCCCCAGAG...ATC...CCAGAAGGAGCAGGAGAGGTGGATCTAAATACCTGGTTGG	5583
MND-2.GA.-.M14	.....-TCA-GTTTT-CAGAAATGGCAGAA...AGA...GCACCAGAG...GCA...CCACAGGGGGCAGGAGAGTGGACTAGAACAGTGGCTGG	5510
MNE.US.-.MNE027	.....ATGTCA.....GATCCCAGAGAGAGAAATCCACCTGGA.....AACAGTGGAGAAGAGACAATAGGAGAAGCCCTTCGAATGGCTAA	5355
LST.CD.88.447	.....	4777
LST.CD.88.485	.....	4777
LST.CD.88.524	.....	4774
LST.KE.-.lho7	.....	5858
SUN.GA.98.L14	.....	5908
MND-1.GA.-.MNDGB1	.....	5226
MON.CM.99.L1	T-CG--CCATTGGACC-CC--G	5389
MON.NG.-.NG1	.....	3962
MUS.CM.01.1085	--CAGCGCC-TGGCAC-CCA--GTGGTC-CA-CC-CC	5352
DEB.CM.99.CM40	.....	5327
DEB.CM.99.CM5	.....	5321
SYK.KE.-.KE51	.....	5358
SYK.KE.-.SYK173	.....	5700





PLV  
complete genomes



H1B.FR.83.HXB2	.....	5558
H1A1.UG.85.U455	.....	5004
H1B.US.90.WEAU160	.....	5557
H1C.ET.86.ETH2220	.....	4950
H1D.CD.84.84ZR085	.....	5080
H1F1.BE.93.VI850	.....	4898
H1G.SE.93.SE6165	.....	4955
H1H.CF.90.056	.....	4905
H1J.SE.93.SE7887	.....	4872
H1K.CM.96.MP535	.....	4754
H101.AE.TH.90.CM240	.....	5129
H102.AG.NG.-.IBNG	.....	5083
H1N.CM.95.YBF30	.....	5152
H1O.BE.87.ANT70	.....	5613
H1O.CM.91.MVP5180	.....	5588
CPZ.CD.-.ANT	.....	5015
CPZ.CM.-.CAM3	.....	4972
CPZ.GA.88.GAB2	.....	4951
CPZ.GA.-.CPZGAB	.....	5622
CPZ.US.85.CPZUS	.....	5618
CPZ.TZ.-.TAN1	.....	5220
H2A.GW.-.ALI	ACAGGACAGTAGAGGCCCTTAAACAGAGAGGCGGTGAATCACCTG . . . CCCCAGGAGCTCATTTCCAGGTGTGGCAAAGGTCCTGGAGATACTGGCATGAT . . . GAACAAGGAATGTCACAAAGTTACAC	6090
H2A.DE.-.BEN	AGAGGACCATAGAAGCCTTAAACAGGAGGAGCAGTGAACCATCTG . . . CCCCAGGAGCTCATTTCCAGGTGTGGCAAAGGTCCTGGAGATACTGGCATGAT . . . GAACAAGGGATGTCAGCAAGCTACAC	6097
H2A.SN.-.ST	ACAGGACGGTAGAAGCCATAAACAGAGAGGCGAGTGAACCATCTG . . . CCCCAGGAGCTTAATTTCCAGGTGTGGCAAAGGTCCTGGAGATACTGGCATGAT . . . GAACAAGGAATGTCAAATAGTTACAC	5542
H2B.GH.86.D205	AAAGAACAATAACAGAGCTCAACAGGAGTACCGGTCAACCATTTG . . . CCCCAGGAACTTATTTCCAGGTCTGGCAGAAGTCTTTGGGCATATTGGCGTGAG . . . GAACAGGGCATGTCAATTAGTTATAC	6073
H2B.CI.-.BHO	AGACAACATTAGAACATCTCAACAGAGTAGCATCAATCACTTTG . . . CCCCAGGAACTTATTTCCAGGTCTGGCAGAAGTCTTTGGGCATATTGGCGTGAG . . . GAACAGGGCATGTCAATTAGTTATAC	6071
H2G.CI.-.ABT96	AAAGAACAGTMACAGAGATCAACAGGATAGCAGTCAATCATTTA . . . CCCCAGGAGCTTAATTTCCAGGTGTGGCAAAGGTCCTGGGAATATTGGCATGAT . . . GAACAAGGGATGTCAGCTAGTTACAC	5424
H2U.FR.96.12034	AAAGAACCTGTGAGAGATATCAACAGGGCAGCAGTGAACATCTT . . . CCCCAGGAACTTATTTCCAGGTCTGGCAAAGGTCCTGGCAATACTGGCGTGAT . . . GAACAAGGGCATGTGAGAAAGTTATAC	5572
MAC.US.-.239	ACAGAACAGTAGAGGAGATAAACAGAGAGGCGGTAAACACCTA . . . CCAAGGGAGCTAATTTCCAGGTTGGCAAAGGTCCTGGGAATACTGGCATGAT . . . GAACAAGGGATGTCACCAAGCTATGT	6011
Vpx	N R T V E E I N R E A V N H L . . P R E L I F O V W Q R S W E Y W H D . . E Q G M S P S Y V	Vpx
Vif	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 S	Vif
SMM.SL.92.SL92B	ACAATACAGTGGAGGCCCTGAATCAACAGCAGTACAACACCTA . . . CCAAGAGAGTTAATTTCCAGGTCTGGAGAAGGTGTTGGGAATACTGGGTAGAT . . . GAGCAAGGATACTACCAAGCTATGC	5448
SMM.US.-.H9	ACAGAACASTAGAAGAAATAAACAGGGCAGCAGTGAATCATTTG . . . CCRAGGGAGCTAATTTCCAGGTCTGGCAGAAGGTCCTGGGAATACTGGCATGAT . . . GAAATGGGGATGTCAGTCAGCTACAC	5484
STM.US.-.STM	ACAGAACGGTAGAGGACATAAACCCAGAGAGGCGAGTAAACCATCTA . . . CCGAGGGAACTGATTTTCCAGGTCTGGCAGAAGGTCCTGGGAATACTGGCATGAT . . . GAGCAAGGGATGCCCGGAGCTATGT	5658
SAB.SN.-.SAB1C	.....	5910
TAN.UG.-.TAN1	.....	5738
VER.KE.-.AGM155	.....	5740
VER.KE.-.9063	.....	5745
VER.DE.-.AGM3	.....	5243
VER.KE.-.TY01	.....	5235
COL.CM.-.CGU1	.....	5280
DEN.CD.-.CD1	.....	5450
GRV.ET.-.GRI_677	.....	5682
GSN.CM.99.CN166	.....	5336
GSN.CM.99.CN71	.....	5318
DRL.-.-.FAO	AGAGGAGTCTCTTAAGAATCAACAGGAGGCTCGATTACACTTCCACCCA . . . GAGTTCCTTTCCCGT . . . CTTTGGAAACACCTGCATGGAGCACTACCATGATGCTCTTCAGTTATCTTTTACTTTACAG	5288
RCM.GA.-.GAB1	GGGACATGCTAGAAAAGGTCAACTTGGAGCCAGGTTACACTTTCATCCA . . . GAGTTCATTTCCCGT . . . CTGTGGCGCACTTGTGTGGAACACTGGCATGACGTACATCAGAGAAGTTTAGAATATGC	5200
RCM.NG.-.NG411	ACAGAATGCTAACAGAAGTCAACTTAGAAGCCAGGTGCACTTTCATCCA . . . GAGTTCATTTCCCGT . . . CTGTGGCGTACTTGTGTGGAACACTGGCATGATAGGCTTGGAGAAGCCTTGAATATGC	5230
MND-2.-.-.5440	AAACGTCACCTGGAGAGAATCAACCCGGAGGCCCGGTTACACTTCCACCCA . . . GAGTTCCTTTCCCGT . . . CTCTGGAACACATGTGTAGAACACTGGCATGATAGACATCAGAGATCTCTTGATTATGC	5270
MND-2.CM.98.CM16	AAAGATCTCTGGAGAAAATCAATCAGGAGGCAAGATTACACTTCCACCCA . . . GAGTTCCTTTCCCGT . . . CTCTGGAACGATGCATAGAACACTGGCATGACAGACATCAGAGAAGCTTAAGCTATGC	5707
MND-2.GA.-.M14	AAAGATCATTTGGAACAAGTGAACAGAGAGGCTCAATTACACTTCCACCCA . . . GAGTTCCTTTCCCGT . . . CTTTGGAAACACCTGTGTAGAACACTGGCATGATAGACTTAGAAGAACCCTTAATTTATGC	5634
MNE.US.-.MNE027	ACAGAACAGTAGAGGAGATAAACAGAGAGGCGAGTAAACACCTA . . . CCAAGGGAGCTGATTTCCAG . . . GTTTGGCAAAGGTCCTGGGAATACTGGCATGATGAACAAGGGATGTCACCAAGCTATGT	5479
LST.CD.88.447	.....	4777
LST.CD.88.485	.....	4777
LST.CD.88.524	.....	4774
LST.KE.-.lho7	.....	5858
SUN.GA.98.L14	.....	5908
MND-1.GA.-.MNDGB1	.....	5226
MON.CM.99.L1	.....	5389
MON.NG.-.NG1	.....	3962
MUS.CM.01.1085	.....	5352
DEB.CM.99.CM40	.....	5327
DEB.CM.99.CM5	.....	5321
SYK.KE.-.KE51	.....	5358
SYK.KE.-.SYK173	.....	5700



H1B.FR.83.HXB2	.....	5558
H1A1.UG.85.U455	.....	5004
H1B.US.90.WEAU160	.....	5557
H1C.ET.86.ETH2220	.....	4950
H1D.CD.84.84ZR085	.....	5080
H1F1.BE.93.VI850	.....	4898
H1G.SE.93.SE6165	.....	4955
H1H.CF.90.056	.....	4905
H1J.SE.93.SE7887	.....	4872
H1K.CM.96.MP535	.....	4754
H101.AE.TH.90.CM240	.....	5129
H102.AG.NG.-.IBNG	.....	5083
H1N.CM.95.YBF30	.....	5152
H1O.BE.87.ANT70	.....	5613
H10.CM.91.MVP5180	.....	5588
CPZ.CD.-.ANT	.....	5015
CPZ.CM.-.CAM3	.....	4972
CPZ.GA.88.GAB2	.....	4951
CPZ.GA.-.CPZGAB	.....	5622
CPZ.US.85.CPZUS	.....	5618
CPZ.TZ.-.TAN1	.....	5220
H2A.GW.-.ALI	AAAAATAGATACTTGTGCTTAATGCAGAAAGCTATGTACACACATTTTCATGAAAGGGTGCACC.....TGCTGGGGGGAGGGCATGGGCCAGGAGGATGG...AGATCAGGACCTCCCCCT	6205
H2A.DE.-.BEN	AAAAGTATAGATATTTGTGCCTAATGCAAAAAGCTATATTTACACATTTCAAGAGAGGGTGCACCT.....TGCTGGGGGGAGGACATGGGCCGGGAAGGATTGGAAGACCAAGGACCTCCCCCT	6215
H2A.SN.-.ST	AAAAGTATAGATATTTGTGCCTAATGCAGAAAGCTATGTTTCATACATTTCAAGAGAGGGTGCACCT.....TGCTGGGGGGAGGACATGGGCCGGGAGGATGG...AGATCAGGACCTCCCCCT	5657
H2B.GH.86.D205	CAAAATATAGATACTTGTGCTAATGCAGAAAGCAATGTTTGTGCACATATCAAAAGGGCTGTAGG.....TGCCCTGCAGGAGGGCCATGGGCCAGGGGGATGG...AGATCAGGACCTCCTCCT	6188
H2B.CI.-.EHO	CAAAATATAGATATTTGTGCTAATGCAGAAAGCAATGTTTATACACTTTGCAAAAGGGTGTGGA.....TGCTGCAGGAGGGCCATGGGCCAGGGGGATGG...AGATCAGGACCTCCTCCT	6186
H2G.CI.-.ABT96	TAAGTATAGATATTTGTGCTGATGCAAAAAGGCTATGTTTGTACATTTGAAGAAAGGATGCACA.....TGCTTGCAAGAGGGCCATGGGCCCTGGAGGATGG...AGACAAGGACCTCCTCCT	5539
H2U.FR.96.12034	TAAGTATAGATATCTCTGCTGATACAGAAAGCTTTATTTATGCATTGCAAAAAGGGATGTCAA.....TGCTTGAGGGGAAGGCCATGGACCAGGAGGATGG...AGATCAGGGCCTCCTCCT	5687
MAC.US.-.239	AAAATACAGATACTTGTGTTAATACAAAAAGGCTTTATTTATGCATTGCAAAAAGGGTGTAGA.....TGCTTAGGGGAAGGACATGGGGCAGGGGGATGG...AGACCAGGACCTCCTCCT	6126
Vpx	<u>K Y R Y L C L I Q K A L F M H C K K G C R</u> ..... <u>C L G E G H G A G G W</u> ..... <u>R P G P P P</u>	Vpx
SMM.SL.92.SL92B	AAAATATAGATATGTACAGTTAATGCAAAAAGCAATGTTTCAGCACTTCAGAAAGGGATGCACA.....TGCAAGGAGAGAGGGCATTTCAGGGGGGGCTGG...AGAAGTGGACCACCACCT	5563
SMM.US.-.H9	TAAATRCAGATACTTGTGCTTGATACAGAAAGCTCTGTTTATGCATTGCAAAAAGGGTGTAGG.....TGCTTAGGAGAGAGCATGGGGCAGGGGGATGG...AGACCAGGGCCTCCTCCT	5599
STM.US.-.STM	TAAATATAGATACTTGTGCTTGATACAGAAAGCTTTGTTTATGCATAGCAAGAGGGATGCAGA.....TGCTTGAGGGGAAGGACATGGGGCAGGAGGATGG...AGACCAGGGCCTCCTCCT	5773
SAB.SN.-.SAB1C	.....	5910
TAN.UG.-.TAN1	.....	5738
VER.KE.-.AGM155	.....	5740
VER.KE.-.9063	.....	5745
VER.DE.-.AGM3	.....	5243
VER.KE.-.TY01	.....	5235
COL.CM.-.CGU1	.....	5280
DEN.CD.-.CD1	.....	5450
GRV.ET.-.GRI_677	.....	5682
GSN.CM.99.CN166	.....	5336
GSN.CM.99.CN71	CAAGTATAGATACTTGTGTTTACAGAAAGCCATGTTTCATGCACCTTCAGCAAGGATGCTCATGTCTGCAGGGAAGGCAT...CCACCT...CCCCTCAGACCAGCAGGAGATAGACTTCTCCTCCT	5318
DRL.-.-.FAO	CAAGTATAGATACTTGTGTTTACAGAAAGCCATGTTTCATGCACCTTCAGCAAGGATGCTCATGTCTGCAGGGAAGGCAT...CCACCT...CCCCTCAGACCAGCAGGAGATAGACTTCTCCTCCT	5412
RCM.GA.-.GAB1	TGCTTATAGGTACTTACTGCTGATGCAAAAAGCTTTGTTTATCCACTGTCAGACAGGGTGTAGTCAG.....AGA.....CATGGA...CCCAATCCTAGGGCAGTAGGAGAAAGGATA	5306
RCM.NG.-.NG411	AGGCTATAGATATCTGCTTCTGATGCAAAAAGCTCTGTTTATTCATTGCAATCAGGGTGTCTCAG.....AGA.....CATGGACAGGGACAAGCAAGGGAAGCAGGAGAAAGAATC	5339
MND-2.-.-.5440	CAAGTATAGATACTGCTGTTGATGCATAAGGCCATGTATATCAGATGCAACAGGGATGCCCATGTAGAATGGGCGCCCAAGG...GGA.....	5358
MND-2.CM.98.CM16	AAAAGTATAGATACTGCTGTTGATGAATAAAGCAATGTTTACTCAGATGCAACAAGAATGTCCTTGTAGAAGTGGGCATCCGAGA...GGA.....	5795
MND-2.GA.-.M14	CAAAATATAGTATCTTCTGCTGATGCAGAAAGCCATGTTTGTTCATATGCAACAAGGATGCCCATGTAGAAGTGGACATTCAGA...GGA.....	5722
MNE.US.-.MNE027	AAAATACAGATACTTGTGTTAATACAAAAAGGCTTTATTTATGCATTGCAAAAAGGGTGTAGA.....TGCTTAGGGGAAGGACATGGGGCAGGGGGATGG...AGACCAGGACCTCCTCCT	5594
LST.CD.88.447	.....	4777
LST.CD.88.485	.....	4777
LST.CD.88.524	.....	4774
LST.KE.-.lho7	.....	5858
SUN.GA.98.L14	.....	5908
MND-1.GA.-.MNDGB1	.....	5226
MON.CM.99.L1	.....	5389
MON.NG.-.NG1	.....	3962
MUS.CM.01.1085	.....	5352
DEB.CM.99.CM40	.....	5327
DEB.CM.99.CM5	.....	5321
SYK.KE.-.KE51	.....	5358
SYK.KE.-.SYK173	.....	5700

PLV  
complete genomes

Accession	Sequence	Accession
H1B.FR.83.HXB2	..... / Vpr start ..... ATGGAACAAGCC..... CCAGAAGACCAAGGGCCAC	5589
Vif	..... W N K P ..... Q K T K G H	Vif
Vpr	..... M E Q A ..... P E D Q G P	Vpr
H1A1.UG.85.U455	.....	5035
H1B.US.90.WEAU160	.....	5588
H1C.ET.86.ETH2220	.....	4981
H1D.CD.84.84ZR085	.....	5111
H1F1.BE.93.VI850	.....	4929
H1G.SE.93.SE6165	.....	4986
H1H.CF.90.056	.....	4936
H1J.SE.93.SE7887	.....	4903
H1K.CM.96.MP535	.....	4785
H1O1.AE.TH.90.CM240	.....	5157
H1O2.AG.NG.-.IBNG	..... AG .....	5114
H1N.CM.95.YBF30	..... G--A .....	5183
H1O.BE.87.ANT70	..... A .....	5644
H1O.CM.91.MVP5180	..... A .....	5619
CPZ.CD.-.ANT	.....	5046
CPZ.CM.-.CAM3	..... G--A .....	5003
CPZ.GA.88.GAB2	.....	4982
CPZ.GA.-.CPZGAB	..... A .....	5653
CPZ.US.85.CPZUS	..... TG .....	5649
CPZ.TZ.-.TAN1	..... A .....	5251
H2A.GW.-.ALI	CCTCCCCCTCCAGGTCTAG-CT-..... ATGACTGAAGCACCACAGAGTTTCCC--G-CC-GAATG--A--C-	6274
H2A.DE.-.BEN	CCTCCCCCTCCAGGTCTAG-CT-..... ATGACTGAAGCACCACAGAGTTTCCC-----TGGGACC----	6284
H2A.SN.-.ST	CCTCCCCCTCCAGGTCTAG-CT-..... ATGACTGAAGCACCACAGAGTTCTCCC-G--G--TAGGACC----	5726
H2B.GH.86.D205	CCTCCTCCCCCAGGCTGGCCT-..... ATGGCAGAAGCAGCCCCAGAGATCCCT----GA--G-GAAC----	6257
H2B.CI.-.EHO	CCTCCTCCCCCAGGCTTAGCCT-..... ATGGCAGAAGCAGTCCCAGAGATTCCT--G--TA--AAC----	6255
H2G.CI.-.ABT96	CCTCCTCCTCCAGGTCTAGCAT-..... G..... ATGGCAGAA..... GAAATCCCC-----TG-G-CC--C-	5600
H2U.FR.96.12034	CCTCCCCCACCAGGATTG--CT-..... ATGGCA..... GAAATACCT--G--GGG-CC----	5744
MAC.US.-.239	CCTCCCCCTCCAGGACTAGCAT-..... ATGGAAAGAAAGACCT-----A-TG-----A----	6184
Vpx end in HIV-2 and SIV-SMM	.....	
SMM.SL.92.SL92B	..... P P P P G L A \$ ..... M E E R P P E N E G P	Vpr
SMM.US.-.H9	CCTCCCCCACCAGGTTTGGCCT-..... ATGGAA...CATGCT-----G--TG-GACCAAC-	5617
STM.US.-.STM	CCTCCCCCTCCAGGACTAGCAT-..... A.ATGACAGAAAGACCT-----TG--CC----	5658
SAB.SN.-.SAB1C	..... CCTC--GAGGTGGCTCCACCAGTAGGAGGGATCCTCCAAGGATCCCCCAAGAATCCAGAGAAGAGATACCAGGATGG.....-T---ACATGG-ATTTG-	6016
TAN.UG.-.TAN1	..... -C-G--GAAGA..... GATTCCAGGGAAGAAAGACCCGGATGGTTAGAAATTTGG--T-TCA-C....	5801
VER.KE.-.AGM155	..... -CCTC--GAAGA..... GATCCAGAGAGGAGAGACCAGGAGGA.....-T---ATTTGG-AT-TCA	5801
VER.KE.-.9063	..... -C-TC--GGAGA..... GGTCCAGAGAAAACAGGCCAGGAGAA.....GT--GAT-TGG-AT-T-A	5806
VER.DE.-.AGM3	..... -CCTC--G-AGA..... GATCCAGAGAAAGCAAGACCCGGGGAA.....-T---ATATGG-ATTTGA	5304
VER.KE.-.TYO1	..... -CCTC--GGAGA..... GATCCAAGAGAGGCAAGACCAGGAGAA.....GT--ATTTGG-AT-T-A	5296
COL.CM.-.CGU1	..... -AG..... GTCCGAAAAGGACTTCACA.....	5306
DEN.CD.-.CD1	..... -AGGTTA..... CCACCATCACACCACCAGCA.....TTAGTTTCAAG-A	5496
GRV.ET.-.GRI_677	..... -C-TC--GAAGA..... GATCCAAGAGAACCTTTACCAGGATGG.....-TG--AT-TGG-AT-T-G	5743
GSN.CM.99.CN166	..... -AG..... ATCCCACCAGTCATCCTATGCCT.....TGG-TCA-C-G-A	5382
GSN.CM.99.CN71	..... -AG..... ACCCCACCAAGTCACCCCTTACCT.....TGGATCA-C-G-A	5364
DRL.-.-.FAO	CCTCCTCCTCCATGATTGG-CT--T--.T..... TGTGTCTTACAGATGGAAAGGCCA--G--G--TG---A--G-	5483
RCM.GA.-.GAB1	ACCATCCTACCGGGATGTAATG..... ATGGCCCTCTCTTTACAGAGATGGAGCTGCC--T-G-TG---T-T-G-	5380
RCM.NG.-.NG411	CAGATTCTCCGGGAATGTAAT..... CAAGCACTCTCTCATTGAGAGATGGAGATGCTC--G--G--TG---AA---	5415
MND-2.-.-.5440	CCTCCTCCTCCAGGATGGCTT-..... TTCTCTGTTTTACAGA..TGGAACAGCCA--T-G--TG-G-CT--C-	5429
MND-2.CM.98.CM16	CCTCCTCCTCCAGGATGG-CT--..... CCTCTGTGTTACAGA..TGGAACAGCCA--C--G--G-G-CC--T-	5864
MND-2.GA.-.M14	CCTCCTCCTCCAGGATGGCTT-..... CTTCTGTCTTGCAGA..TGGAACAGCCG--G--TG-G-CT--	5791
MNE.US.-.MNE027	CCTCCCCCTCCAGGACTAGCAT-..... ATGGAAAGAAAGACCT-----TG---C----	5652
LST.CD.88.447	..... -TCCAG-A-AAGG..... CAGCAAAGGAGTCCAGAAAACCTCCA.....-G--TTT-----	4835
LST.CD.88.485	..... -TCCAG--AAGG..... CAGCAAAGGAGTCCAGAAAACCTCCA.....-G--TTT--A-----	4835
LST.CD.88.524	..... -TCCAG-CRAAGA..... CAGCAGCGCAGCCAGAAAGGCCCCCA.....-G--TTT--T--T-	4832
LST.KE.-.lho7	..... -TCCAGGCAGAGG..... CAGCAAAGAAAGTCCAGAGAGACCACCA.....-G--TTT--A-----	5916
SUN.GA.98.L14	..... -CCAGC..... AGGAGAGAGCCAGTAGAACAGCCACCA.....-G-G-----	5960
MND-1.GA.-.MNDGB1	..... -TT..... GGCCAGAAGAGAGATGAGCAAGTATCA.....-----T-----A--T-	5272
MON.CM.99.L1	..... -TT.....	5395
MON.NG.-.NG1	..... -C-C..... CACCCAGGC.....CATCCT--C-TTGGACCTCC-GCG	4001
MUS.CM.01.1085	..... -CGC.....	5358
DEB.CM.99.CM40	..... -AGG..... TACCCACCTCACACCCA--C-TTT-ACCTCAAG-A	5373
DEB.CM.99.CM5	..... -AGG..... TATCCACCTCACACCCA--C-TTT-AC-TCAAG-A	5367
SYK.KE.-.KB51	..... -C--G--TTTC..... TTCAGAGTTTACCAGCGGCAGCCAGGGAAACCACTATTTTACATTCCA.....AGAA-T-TCAATT	5434
SYK.KE.-.SYK173	..... -CGG-G--TTTC..... TTCATCCTCCAGCATGTGCAAGGAACACCATGGTTCTTACATTCCA.....AGAA-T-TAGAGT	5776



Vif end in HIV-1 \		5714
H1B.FR.83.HXB2	AGAGGGAGCCACCAATGAATGGACTAGAGCTTTTAGAGGAGCTTAAAGTGAAGCTGTAGACATTTT...CCTAGGATTGGCTCCATGGCTTAGGGCAACATATCTATGAACTTATGGGGAT..	Vif
Vif	R G S H T M N G H \$	Vpr
Vpr	Q R E P H N E W T L E L L E E L K N E A V R H F . P R I W L H G L G Q H I Y E T Y G D .	Vpr
H1A1.UG.85.U455	-----GT-GC-----A-G-T-----T-G-----A-----C-----A-----...CAG-----A-----A-----A-C-C-----..	5160
H1B.US.90.WEAU160	-----T-----G-----A-----G-----G-----G-----A-----A-----A-----A-----..	5173
H1C.ET.86.ETH2220	-----A-----T-----G-----C-----C-----G-----C-----C-----...ACCA-----AA-----A-----T-----C-----..	5106
H1D.CD.84.84ZR085	-----T-----T-----T-----A-----G-----C-----G-----C-----...G-----T-----A-----T-----C-----..	5236
H1F1.BE.93.VI850	-----T-----G-----AA-----G-----A-----A-----G-----...CCA-----T-----C-----C-----A-C-C-----C..	5054
H1G.SE.93.SE6165	-----A-----A-----CT-----G-----T-----A-----G-----A-----A-----A-----...C-----T-----A-----G-----A-C-----A-C-----..	5111
H1H.CF.90.056	-----A-----T-----C-----G-----A-----A-----G-----G-----G-----...AG-----CA-----CAA-----A-----G-----A-C-C-----A-----..	5061
H1J.SE.93.SE7887	-----A-----T-----C-----G-----A-----A-----G-----G-----G-----...CCA-----T-----A-----T-----A-----GT-----AGC-----A-----..	5028
H1K.CM.96.MP535	-----A-----A-----CT-----T-----A-----G-----GA-----...CCA-----AATC-----G-----A-----G-----AC-----..	4910
H101.AE.TH.90.CM240	-----A-----A-----CT-----T-----A-----G-----A-----A-----...CCC-----A-----A-----A-----GT-----C-----A-C-A-----..	5282
H102.AG.NG.1.IBNG	-----GTC-----A-----G-----A-----AC-----C-----...CCG-----A-----A-----A-----A-----..	5239
H1N.CM.95.YBF30	-----A-----CT-----T-----G-----AT-----A-----A-----AT-----A-----GC-----...A-----A-----G-----A-----C-----A-C-A-----A-C-----..	5308
H10.BE.87.ANT70	CT-AA-----CT-----G-----C-----A-----A-----AGCA-----A-----A-----C-----...CC-----A-----C-----G-----A-----T-----C-----T-----G-----..	5769
H10.CM.91.MVP5180	CT-----TT-----C-----G-----A-----G-----AG-----A-----A-----C-----...CC-----T-----A-----A-----C-----GT-----GT-----C-----T-----G-----A-----C-----..	5744
CPZ.CD.-.ANT	-----A-----A-----ATG-----G-----TTGT-----AAC-----A-----AA-----A-----G-----A-----...ACA-----ACAT-----A-----ACA-----C-----T-----AA-----TTGGG-----A-----C-----A-----C-----A-----..	5171
CPZ.CM.-.CAM3	CA-----A-----T-----G-----T-----A-----GC-----G-----A-----A-----AC-----C-----...C-----AGCA-----A-----A-----A-----GT-----T-----C-----A-----A-----..	5128
CPZ.GA.88.GAB2	CAG-----A-----TT-----G-----G-----AA-----A-----G-----CT-----A-----A-----C-----G-----...ACCA-----A-----C-----CAG-----A-----TTA-----CA-----T-----A-----A-----..	5107
CPZ.GA.-.CPZGAB	CA-----A-----TT-----C-----G-----G-----T-----ACC-----G-----A-----AT-----A-----A-----A-----...ACCA-----T-----G-----CAA-----A-----TTC-----T-----C-----A-----A-----..	5778
CPZ.US.85.CPZUS	CC-----A-----A-----T-----G-----G-----T-----TG-----C-----A-----A-----AG-----A-----G-----T-----GC-----...ACCC-----A-----A-----A-----GT-----C-----CAGT-----A-----C-----..	5774
CPZ.TZ.-.TAN1	-----A-----A-----T-----C-----G-----CTGT-----CA-----CC-----A-----AA-----A-----AC-----A-----G-----A-----...A-----ACCAATAT-----A-----G-----GG-----AA-----TTGGG-----TCACC-----T-----A-----C-----..	5376
H2A.GW.-.ALI	-CCA-----GAG-----CAGG-----G-----GT-----A-----AG-----CC-----GAGA-----AA-----A-----AG-----A-----...T-----A-----G-----GAC-----C-----C-----G-----...AAT-----CTC-----T-----CGGCT-----ACT-----GAC-----A-----C-----..	6399
H2A.DE.-.BEN	G-----CTT-----GGG-----G-----GT-----A-----AAC-----C-----GAG-----AA-----A-----AG-----A-----...CT-----A-----GAT-----CC-----CT-----G-----...AAT-----CTC-----T-----CT-----CT-----C-----A-----T-----GAC-----A-----C-----..	6409
H2A.SN.-.ST	C-----GGGG-----G-----GT-----A-----AACCC-----GAGA-----A-----A-----AT-----A-----...T-----A-----AG-----C-----GAT-----C-----CT-----G-----...AAT-----ACTC-----T-----CA-----CT-----CT-----GAC-----A-----C-----..	5851
H2B.GH.86.D205	-A-----A-----A-----GTGGG-----A-----G-----T-----GGG-----A-----CC-----G-----AA-----A-----C-----A-----...CT-----A-----AG-----GAT-----C-----CT-----G-----...AAC-----CGC-----T-----TA-----CTT-----CAGT-----GGC-----A-----..	6382
H2B.CI.-.EHO	-A-----A-----A-----GTGGG-----AC-----G-----GT-----G-----G-----CG-----CC-----G-----AA-----A-----AC-----A-----...CT-----A-----AG-----GAC-----CC-----CT-----G-----...T-----AAC-----CAC-----T-----CA-----TTT-----CA-----T-----GGC-----AA-----..	6380
H2G.CI.-.ABT96	-----A-----A-----T-----GGG-----G-----GT-----G-----AG-----C-----G-----A-----AA-----C-----AG-----A-----...CT-----A-----A-----GAT-----C-----CT-----G-----...AAC-----CGC-----T-----TA-----CT-----T-----GGC-----A-----C-----..	5725
H2U.FR.96.12034	-A-----A-----TGGGG-----C-----G-----GTGAG-----AA-----AC-----G-----A-----A-----AG-----A-----...A-----G-----GAT-----C-----CT-----G-----...AAC-----CGC-----T-----TA-----CT-----A-----T-----GGC-----A-----..	5869
MAC.US.-.239	-A-----A-----TGGG-----GT-----G-----G-----C-----G-----A-----A-----G-----AG-----A-----...T-----A-----A-----GAT-----C-----CT-----G-----...AAC-----CAC-----T-----TA-----T-----A-----T-----GAC-----A-----C-----..	6309
Vpr	Q R E P W D E W V V E V L E E L K E E A L K H F D P R L . L T A L G N H I Y N R H G D . M E T	Vpr Tat ex 1
SMM.SL.92.SL92B	CA-----A-----A-----TGGG-----G-----T-----AG-----TG-----AC-----G-----A-----AT-----G-----AG-----A-----G-----AC-----G-----A-----...GAC-----C-----CT-----G-----...TAC-----CT-----AA-----TT-----CG-----A-----T-----C-----A-----A-----..	5742
SMM.US.-.H9	-----A-----GTGG-----G-----GTRG-----AG-----C-----G-----AA-----KA-----AG-----A-----CC-----G-----AC-----K-----...GAT-----C-----CT-----G-----...AATY-----CGC-----T-----TA-----CT-----T-----GAC-----A-----C-----..	5783
STM.US.-.STM	-----A-----A-----TGGG-----G-----GT-----G-----G-----AG-----C-----G-----AA-----A-----AC-----A-----...CC-----GC-----G-----GAT-----C-----CT-----G-----...AAGC-----CGC-----T-----TA-----CT-----T-----GAC-----A-----C-----..	5957
SAB.SN.-.SAB1C	CC-----A-----GTT-----G-----C-----CTC-----G-----CA-----G-----C-----C-----C-----TTCA-----G-----CCAGT-----C-----C-----C-----...-A-----A-----CTC-----TTTCC-----TC-----TT-----TGGA-----TGTG-----GAACCA-----CCAT-----..	6141
TAN.UG.-.TAN1	-----GTGGG-----C-----ATT-----AG-----CA-----GG-----CA-----A-----C-----TC-----G-----ACAAC-----C-----...GGG-----AGAAGCTC-----TTCCAAG-----T-----A-----C-----...T-----G-----C-----GGAAG-----A-----AAG-----..	5923
VER.KE.-.AGM155	GC-----GTGGG-----C-----CTCAG-----TA-----GG-----A-----AA-----C-----C-----CAAAGT-----...GGCC-----AGAAGCTC-----AT-----CCAAG-----T-----A-----T-----...TAT-----G-----C-----GGAGG-----A-----GAG-----..	5925
VER.KE.-.9063	-C-----A-----A-----TGGG-----C-----CT-----AGG-----CA-----G-----A-----T-----C-----TC-----G-----CAAAT-----...GGC-----AGAGCTC-----TTCCAAG-----TT-----A-----T-----...TAT-----G-----C-----GGAGG-----A-----A-----GAG-----..	5930
VER.DE.-.AGM3	GC-----GTGGG-----C-----CT-----AG-----CA-----GC-----A-----TA-----C-----TC-----G-----CAAG-----TG-----...GGG-----CGAGCTC-----GTTCCAAG-----T-----A-----C-----...TAT-----G-----C-----GGAGG-----G-----A-----AAG-----..	5428
VER.KE.-.TY01	GC-----GTGGG-----CA-----CT-----AG-----TA-----G-----C-----A-----T-----C-----CG-----G-----GAGGCT-----...GGG-----AGAAGCTC-----TTCCAAG-----T-----A-----C-----...TA-----G-----C-----GGAGG-----A-----A-----AAG-----..	5420
COL.CM.-.CGU1	-----TTG-----AG-----C-----AGG-----A-----CA-----A-----CCC-----G-----G-----CAG-----A-----G-----C-----CCCTTT-----CTC-----ACAT-----CT-----CAGT-----GATA-----GT-----...GGT-----G-----ACT-----GAAG-----...	5402
DBN.CD.-.CD1	T-----GCA-----AA-----CACTC-----AAGGCA-----CT-----AG-----AGCAG-----TG-----TA-----C-----CAG-----G-----G-----AAGA-----A-----C-----C-----...AGC-----AAGA-----A-----A-----AC-----AA-----T-----G-----T-----...CA-----GCTGGT-----A-----TACCA-----CGCT-----..	5620
GRV.ET.-.GRI_677	-C-----TGGG-----C-----CT-----A-----CA-----GC-----CAG-----T-----A-----CG-----A-----CAG-----A-----G-----C-----...GGA-----T-----ACAT-----AATCC-----AG-----T-----A-----T-----...TA-----G-----T-----GAGG-----G-----AAGGAG-----..	5867
GSN.CM.99.CN166	G-----GT-----CCCA-----AC-----TG-----CTATCG-----T-----A-----A-----CGC-----C-----GTG-----AA-----A-----CG-----G-----A-----G-----A-----AA-----A-----C-----C-----...T-----A-----GACGA-----G-----AA-----TT-----T-----...GATGTGACGGAGCTCCCT-----CCGA-----..	5506
GSN.CM.99.CN71	G-----GT-----CCCA-----AC-----TG-----C-----GTTG-----T-----A-----AGC-----GCA-----GTG-----AA-----A-----CG-----A-----G-----A-----AA-----A-----C-----C-----...T-----A-----GAAGA-----G-----AA-----TT-----T-----...GATGTGACGGAGCTCCCT-----CCGA-----..	5488
DRL.-.-.FAO	CA-----A-----A-----CT-----TG-----G-----CT-----G-----TAC-----...TA-----AA-----AC-----AG-----G-----T-----G-----AG-----GATAGGC-----CT-----A-----...T-----G-----AATG-----TC-----TGGG-----...GA-----A-----..	5608
RCM.GA.-.GAB1	-AC-----A-----A-----TT-----TG-----...TT-----A-----G-----TACCC-----ATA-----T-----GC-----AG-----A-----AAG-----A-----...ACATA-----GCTT-----G-----...AACG-----CAAA-----AG-----TT-----G-----A-----GCAAC-----A-----A-----..	5505
RCM.NG.-.NG411	-----A-----A-----CT-----TG-----...CT-----A-----G-----TACCC-----TATA-----A-----AC-----AG-----A-----AAG-----A-----...ACA-----GAAT-----G-----...TAC-----CAAGA-----AA-----TT-----A-----GAC-----A-----..	5540
MND-2.-.-.5440	-----A-----GA-----T-----T-----A-----CTGG-----GTAC-----G-----CA-----AA-----AC-----AG-----G-----T-----A-----AG-----CGATAGGC-----CT-----G-----...-A-----CAG-----CTC-----TGGG-----G-----GCAAC-----A-----A-----C-----..	5554
MND-2.CM.98.CM16	-A-----A-----A-----T-----...CT-----A-----T-----CAC-----CA-----AG-----C-----AG-----A-----T-----A-----AG-----GATAGAC-----CT-----A-----...T-----G-----CAG-----AAGCTGG-----CA-----C-----..	5989
MND-2.GA.-.M14	-A-----A-----A-----T-----T-----G-----CTGA-----TACA-----CA-----AA-----C-----AG-----A-----T-----G-----AG-----GATAGGC-----CT-----A-----...T-----G-----CAG-----AAGCTGG-----T-----A-----C-----..	5916
MNE.US.-.MNE027	-A-----A-----TGGG-----GT-----G-----G-----G-----C-----G-----A-----G-----AG-----A-----...T-----A-----A-----GAT-----C-----CT-----G-----...AAC-----CGC-----T-----TA-----TT-----T-----GAC-----A-----C-----..	5777
LST.CD.88.447	CA-----A-----TGGGG-----C-----CT-----GC-----CACCA-----G-----AA-----A-----AG-----G-----G-----CAGAC-----C-----...-A-----GAACAA-----TTTG-----C-----CA-----TCTC-----G-----ATA-----G-----TC-----...G-----..	4945
LST.CD.88.485	CA-----A-----TGGGG-----C-----TT-----GC-----CACCA-----G-----AA-----A-----G-----G-----G-----CAGAC-----C-----...-A-----GAAGCA-----TTTG-----C-----CA-----TCTCGG-----ATA-----G-----TC-----...G-----..	4945
LST.CD.88.524	CA-----A-----TGGGG-----C-----CT-----GC-----CACAA-----G-----AA-----C-----G-----A-----G-----AAGAC-----C-----...-A-----GAAGCA-----TTTG-----C-----CA-----TCTCGG-----ATA-----G-----TC-----...G-----..	4942
LST.KE.-.lho7	CA-----A-----TGGGG-----C-----TT-----GC-----CACCA-----G-----AA-----A-----AG-----G-----G-----AAGAC-----C-----...-A-----GAAGTAA-----TTTG-----C-----CA-----TCTCG-----ATAT-----G-----T-----...G-----..	6026
SUN.GA.98.L14	C-----A-----TTTG-----C-----CTGGC-----TACCA-----G-----A-----AA-----C-----G-----G-----C-----G-----AA-----G-----C-----...AGCAGCAATAG-----TGCC-----...CA-----TCTTG-----ATA-----G-----CG-----...G-----..	6070
MND-1.GA.-.MNDGB1	CC-----A-----T-----C-----G-----CT-----GC-----TAC-----A-----G-----AA-----A-----G-----A-----AAGA-----AG-----C-----C-----...CTC-----ATC-----AA-----CAG-----TCAG-----...TAT-----G-----TGCAA-----C-----...G-----..	5388
MON.CM.99.L1	...CCCAT-----G-----A-----GAC-----GGCCT-----AG-----CCGCCA-----TTG-----...A-----CG-----A-----CC-----A-----A-----C-----C-----...T-----A-----GAAG-----G-----...ATTTGGG-----ACAGGTAAACGGAGCTCCCT-----C-----GA-----..	5514
MON.NG.-.NG1	T-----GTTCCAT-----G-----GCG-----CGCTCT-----AG-----CCGCCA-----TG-----C-----A-----CG-----G-----G-----CC-----TYG-----C-----C-----...T-----A-----A-----GAAG-----G-----...ATCTGGG-----CAGGTCAGTGA-----TGCCCT-----CCGA-----..	4125
MUS.CM.01.1085	...CA-----G-----A-----AGCCA-----GTG-----AG-----A-----CG-----A-----G-----A-----AA-----A-----C-----...T-----C-----AGAAG-----G-----A-----...ATTTGGA-----GAGGTAAACGGAGCTGCCA-----CCGA-----..	5462
DEB.CM.99.CM40	CA-----TACCCAT-----GACA-----GGCT-----G-----TTG-----AGC-----GCCA-----GC-----A-----T-----A-----CG-----G-----CC-----A-----A-----C-----...A-----C-----AGAAG-----ATT-----AG-----TT-----A-----T-----TGTTGA-----CTTACCC-----CCC-----ACC-----..	5500
DEB.CM.99.CM5	CAGTACCAAT-----GACA-----GGCT-----GCAT-----A-----C-----AGCCA-----GC-----T-----A-----CG-----G-----T-----A-----A-----C-----C-----...A-----A-----AGAAG-----T-----ATGG-----AG-----CT-----A-----T-----TGTTGA-----CCT-----CC-----CCC-----ACC-----..	5494
SYK.KE.-.KE51	G-----GAC-----G-----GGAGTAA-----A-----TG-----CACT-----CCAAAGC-----AC-----G-----AGCCC-----A-----ATGC-----AG-----...A-----AGC-----GAAGAAA-----ATT-----AG-----TT-----A-----T-----ATG-----TA-----GGTGA-----CA-----G-----CC-----..	5561
SYK.KE.-.SYK173	T-----CCCCA-----A-----GTGATCAAT-----GTCACT-----TAAAGC-----ACT-----G-----AGT-----CA-----ATCC-----AG-----...A-----CCTCAAGAAA-----AT-----...GG-----TTTGA-----CAA-----CCTT-----ACG-----A-----A-----CAGG-----..	5897



PLV  
complete genomes



H1B.FR.83.HXB2 .....ACTTGGGCAGGAGTG...GAAGCCATAATAAGAATTCTGCAACAACACTGCTGTTTATCCAT...TTTCAGAAATGGGTGTCGACATAGCAGA...ATAGGC GTT..... 5808  
Vpr ..... T W A G V . E A I I R I L Q Q L L F I H . F R I G C R H S R . I G V . . . . . 5808  
Vpr ..... T W A G V . E A I I R I L Q Q L L F I H . . . . . 5808  
normal Vpr continues  
\* F Q N W V S T \$  
frameshift (extra T)  
causes premature end in HXB2

H1A1.UG.85.U455 .....-AG----T.....-T.....-T-A-----T-----C-A-----A----- 5253  
H1B.US.90.WEAU160 .....-A-----C-----C-----AA-----T-----A-C-AG-----A----- 5806  
H1C.ET.86.ETH2220 .....-T-G---A.....-TT-----C-----AA-----A-----A-C-AG-----A----- 5199  
H1D.CD.84.84ZR085 .....-C-----T-----T-----A-----T-----A-----A-----A----- 5329  
H1F1.BE.93.VI850 .....-AG----T.....-T-----G-AT-----A-----C-T-----A----- 5147  
H1G.SE.93.SE6165 .....-A-----T-----A-----A-----A-----C-A-----A----- 5204  
H1H.CF.90.056 .....-T-----T-----TT-----CG-----A-----T-----C-A-----AA----- 5154  
H1J.SE.93.SE7887 .....A-----A-----C-----A-----AT-----T-----C-----C-AT-----A----- 5121  
H1K.CM.96.MP535 .....-A-----T-----T-----AT-----A-----T-----C-AC-----A----- 5003  
H101.AE.TH.90.CM240 .....-A-----G-T-----T-----GT-----A-----G-T-----A-----A----- 5375  
H102.AG.NG.-.IBNG .....-A-----G-T-----A-----T-----T-----AT-----G-T-----A-----A----- 5332  
H1N.CM.95.YBF30 .....-C-----AG-G-A.....-G-A-T-C-G-A-----T-A-----AT-G-----C-C-AG-C-----GA-C..... 5401  
H1O.BE.87.ANT70 .....-T-----T-----ATG-A-T-----CT-A-----T-----C-----AT-----A-C-A-----T-----AA----- 5862  
H1O.CM.91.MVP5180 .....-A-----T-----ATG-A-T-----CT-A-----A-----C-----AT-----A-C-A-----T-----AA----- 5837  
CPZ.CD.-.ANT .....T-A--AG-----C-G-TTT-G---TTC-C-T--A--GCT---CA---AT--CAC-A--GCT---G...--ACA..... 5264  
CPZ.CM.-.CAM3 .....A-----T-A-----A-C-T-----A-----G-T-----T-GC-----C-----A-----GA-C..... 5221  
CPZ.GA.88.GAB2 .....A-----A-----G-----C-----A-----T-----G-G-----T-G-----C-----A-----G-----GA..... 5200  
CPZ.GA.-.CPZGAB .....A-----T-----A-----G-----C-----CT-A-----CT-----T-----C-----C-AG-----A-----A-C..... 5871  
CPZ.US.85.CPZUS .....-AG-----A-----A-----T-----G-CT-A-----TT-----AT-G-----C-----AAT-----G-----GA--TCCTTG..... 5873  
CPZ.TZ.-.TAN1 .....T-C--AG-----C-G-AAT--C-AG-CT--GAG-GCTT--C--C...ATC-CCA--T--ATC-C--...AGA--ATCA..... 5469  
H2A.GW.-.ALI .....CCTT-A-A--CC...AGG-AGC-C-T-ATGCC-A---G-GCC-C-C-G-G-C...GCA-A--G-C-G-TCA---G-T--CAA..... 6492  
H2A.DE.-.BEN .....CCTT-A--C-CC...AG-AGC-C-T-A-CC-A---G-GCC-C-C-G-G-C...AAC-GCTCA---A--T--TAA..... 6502  
H2A.SN.-.ST .....CCTT-A--C-CC...AG-GGC-C-T-G-C-A---G-GCC-CC-CT-G-C...GCA-A--CG-C-GCTCA--G...-T--TCAG..... 5944  
H2B.GH.86.D205 .....CCTT-----CA...-G-AGC-C-T-A-C-C---G-GCC-C-CC-C...GCC-T---A--CTCA--G...-T--ACAA..... 6475  
H2B.CI.-.EHO .....CCTT-A--CA...-G-AGC-C-T-A-C-C---G-GGCCA-C-CC-C...GGG-T--C-A--CTCA--G...-T--ACAA..... 6473  
H2G.CI.-.ABT96 .....CCTC-AG--CA...-G-AAC-C-T-A-C---GAG-GCT-C-CC-G-C...TC--GGA-AC--A--TGCTC--G...-T--A-GG..... 5818  
H2U.FR.96.I2034 .....CCTT-A--CA...-G-AGC-C-T-A---C---AG-GCAA-C-CC-C...GGG-A---G-CTCA--G...-C--AGC..... 5962  
MAC.US.-.239 .....CCTT-AG--CA...-G-AAC-C-T---C-C---G-GCC-C-C-G---GGC--A--CATC--CTC---C--CAA..... 6402  
Vpr ..... T L E G A . . . . . G E L I R I L Q R A L F M H . . . . . F R G G C I H S R . . . . . I G O . . . . .  
Tat ..... P L R E Q . . . . . E N S L E S S N E R S S G I . . . . . S E A D A S G P E . . . . . S A F . . . . .  
SMM.SL.92.SL92B .....AATA-A--G-CA...-GG-AA--T-T-AG--AT-A---A-GGCAT--CT-A-C...T-GCA--A--CACC--TCA--G...-T--ACAG..... 5835  
SMM.US.-.H9 .....CCTT-A--CA...-G-AGC-C-T---C-C---GGG-GCA-C-C...GG--A--CA--CTC---C--CAG..... 5876  
STM.US.-.STM .....CCTT-A--CA...-G-AGC-C-T-AG-C-C---G-G-GCC-C-C-G-C...GGC--T--C--C--CTC--G...-T--CAA..... 6050  
SAB.SN.-.SAB1C .....GATCATGGGCAAACC-GGCT--A--CTG...T-TAAATATTGT-----A-GGCTT-A--G-G...A-G-A-G-A-A--T-TA-G...CC-AAGAC..... 6243  
TAN.UG.-.TAN1 .....AATGGGGCACC-A-TGATA-A-A-G-CT...T-TAGATATTAT--GT-GG-A--GA-GGCTT-A--G-G...TC-TG--C--CA-GAGA-A-CT...CCTTT-AG..... 6029  
VER.KE.-.AGM155 .....CAGGGAAGACC-TAGC-A-A-G-CA...T-TAAGTATTATC-CT-AG-T--GA-GCT-C--G-G...C-GTG--A---CAGG--ACA...CCCTTT-AG..... 6031  
VER.KE.-.9063 .....CGGGGAGCCCA-TGAC--AGA-G-CC...T-TAGATACTAT--GT-GG-T--GA-GCA-C--G-T...C-TTG--T---TAGA--ACAG...CCCTT-AA..... 6036  
VER.DE.-.AGM3 .....AATCGCACTCCC-TGCTA-A-A-G-CT...T-TAAATATTAT-A-T-GG--A--GCT-C--G-G...C-GTG--A--C--CAGA--ACA...CCCTTT-AA..... 5534  
VER.KE.-.TY01 .....CATGGTACTCCC-TGAT--A-A-G-CC...T-CAATATTAT--GC-AG-A--A-GGCT-C--G-G...TC-TG----CA-GAGA--GCAG...CCCTTT-AA..... 5526  
COL.CM.-.CGU1 .....CACTGG...GGGAGACCGTGG...A-G-CT-GT...CTG--TTATG-G----CA--G-GTGTGCC--C-ACT-G-C..... 5471  
DEN.CD.-.CD1 .....CCCACATTGGACAGAAGGACAAGTCAT----CC-CCGCAT--ATT-C-----CA--GA-GAAA-A-GG-AG...-AT--G-A-A--TTC----GGA-GCAGA-A-A-- 5741  
GRV.ET.-.GRI\_677 .....CATAATACC...CCATGGAAT-AGAT-GC...T-CAAGTACTAT----G-T--A-GTCTA--G-A-----TG--T--A--AGG-AG...CCTTTTC..... 5973  
GSN.CM.99.CN166 .....TCCAGATTGGACAGTGGATCAGG-AGCAAT--C-TGTGCAAT--ATTAC--C-G...G---AC--T-T-C---GA-C...T-A-GA--A--CTTT--C--GGTCAAT...TTATCA..... 5621  
GSN.CM.99.CN71 .....TCCGGATTGGACAGTGGATCAGG-AGCAAT--C-TGTGCAAT--ATTATG-C-G-GAG--GAC-T-A-C---GA-C...T--TA-----TTC--C--ATACAAT-G-ATT--C..... 5609  
DRL.-.-.FAO .....T-A-TA-A-----C-CCAGC-T--T-C-CTT--AG-GCA-T-T-CT-G...T--TA--A--C-AGG-A-T---C-T-A..... 5698  
RCM.GA.-.GAB1 .....GCATA-AG-----C-G--A--G--TC-GC-G-T--AG-GCCT--C-T---T---A--A--GC-GGG--T-G...-T--AACA..... 5598  
RCM.NG.-.NG411 .....G--TA-AG-----A--A---T-G---CCT-G--TA-TA-GCA-T-CT-G-C...GCA--A--GA-GGG---G...-T--A-GA..... 5633  
MND-2.-.-.5440 .....--TA-A--T--C...C--CAGC-----C--T---AG-GCA-----CT-A-C...T--TA--A--A-GGA--TC-C...-C--ACAA..... 5647  
MND-2.CM.98.CM16 .....-A-TA-A--G-----CG-CAGC-T--CTC-----T--A--GCT---C--G-C...-AT--TA--A--CA-GG-A--C-T...-C--ACAA..... 6082  
MND-2.GA.-.M14 .....C-T--AG-G-----C-GAAGC-T--C-CCC-CT-A--GAG-GCT--A-C-G...T--GA--A--CA-GG-A--C--T...-T--ACAA..... 6009  
MNE.US.-.MNE027 .....CCTT-AG--CA...-G-AAC-C-T-A---C---GGGC--C-C-G-C...GGC--C--CAAC--CTCT---C---CAA..... 5870  
LST.CD.88.447 .....AGCAGTACCGGGTCAGAGGA-AA--CCTGC...ATGAAAT-C-T-CC--AG-AA-TAG-GCA-C---T---GCCT...ACC---CCAGCTATGC...GC--CAAGA..... 5050  
LST.CD.88.485 .....AGCAGTACCGGGTCAGAGGA--A--CCTGC...ATGAGAT-C-T-CC--AG-AA-TAG-GCA-C---T---GCCT...ACC---CCAGCTATGC...GC--ACAAGA..... 5050  
LST.CD.88.524 .....AGCAGTACCGGGTCAGAGGA--AG-CTTGC...ATGAGAT-C-C-CTT-AG-TA-TAG-GCT-C-A-T-A...C-GCCT...CAA---CCAG-TATGC...GC--C-AGA..... 5047  
LST.KE.-.lh07 .....GACAGTACAGGGTCAGAGGA--AG-CCTGC...ATGAAAT-C-C-CCC-AG-CA-TAG-GCTT---C-T-C...C-CC...ACT---CCAGCTATGC...GC--C-AGA..... 6131  
SUN.GA.98.L14 .....GACTCTACAGGGTCAGAAAA--A--GTGT...ATGAGAGC--C-CTT-G--CA-TAG-GCCT---C-T---C-GCC...AT--CTTC-TA--TCAGA--CA-AAG-GGATCA..... 6184  
MND-1.GA.-.MNDGB1 .....ACAGGGAGTGAGAA-AG-CTGT...GAAAT-T-T-CCT-AA--A-TAG-GCCA-T-GGG---C...C-AGCTCAA----GATGGA-C-TTC...-GG-AAAGA..... 5490  
MON.CM.99.L1 .....CCCGCCTGGAATGCAGACCAGG-A-----C-TGTGCCATT-ATTAT-CT--TGGG-C---ACCA-C--T-A--GG-C...-A--GGAA--C--TAC---ATATGCAGAGCAAA-AAGAAGA..... 5641  
MON.NG.-.NG1 .....CCCAGGTGGAACCGCAGACCAGG-A-----C-TGYGCTATT-ATTAC-CC--GCGGG---AC-A-RA-A---GG-C...A-G--GGA-----TAT--C--G...CATGCC-A--G..... 4243  
MUS.CM.01.1085 .....TCCTGATTGGACAGTGGATCAGG-GGCAAT--C-TGTGCCAT--A-TATG-T--G-GAG---AC--TG-C---GG---GAC--A--CTAT----GTATAAC-G--TAA-GAGAAGA..... 5589  
DEB.CM.99.CM40 .....AGATTGGACGGGAGAAACAAGCATGGGCT--TCG...ATT-A-TAC-----A--AG-T--A--A--T-A-GGT-G-C...C-G--GAA-CT--CTTT--C--GGAG...G--AG-C-ACCAGA..... 5621  
DEB.CM.99.CM5 .....AGATTGGACGGGAGAAACAAGCATGGGCT--GTC--A...ATT-A-TAT---A--AG-A--AG-A---T-GGC-T---C-G--GAA-CA--CTTC---AGAG...GG-AA-C-ACTAGG..... 5615  
SYK.KE.-.KE51 .....AGAAGAAACACCCACCATGGCTTGGATTAG-T-CA---CT--AT--GCAT--GCA--CA-CTTTA---T--GAA---TGCTGCA--A---C--G--G-C...AGTA--CG..... 5676  
SYK.KE.-.SYK173 .....GACTGATCCCCGACAATGGCTTGGGAAAG-ACCA---TTG-A---GG---GCA--CA-T-TTA---C-GAA---TGCTGCA--A---CT--AC-G-CC...-G-TAT-CC..... 6012

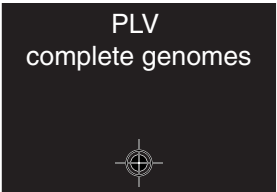
ex1



Accession	Sequence	Position	Label
H1B.FR.83.HXB2	.....ACTCGA.....CAGAGGAGAGCAAGAAATGGAGCCAGTAGATCCTAG.....	5850	Vpr
Vpr	.....T R.....Q R R A R N G A S R S \$		
Tat	.....M E P V D P R.....		Tat ex1
H1A1.UG.85.U455	.....-T-C.....GG--A---GC-G-----A.....	5295	
H1B.US.90.WEAU160	.....-C.....-A-----A.....	5848	
H1C.ET.86.ETH2220	.....TTA-----A-----A.....	5241	
H1D.CD.84.84ZR085	.....-C.....-A-----T-----A.....	5371	
H1F1.BE.93.VI850	.....GT--C.....-A---T-----A.....	5189	
H1G.SE.93.SE6165	.....-A-C.....-G--A--G-T--GG----C-G-----A.....	5246	
H1H.CF.90.056	.....-A-C.....-A---T-----C-----A.....	5196	
H1J.SE.93.SE7887	.....-TA-C.....-A-A---GG-G-----A.....	5163	
H1K.CM.96.MP535	.....-TA-C.....-A---GG-----T-----A.....	5045	
H101.AE.TH.90.CM240	.....-TG-C.....GG--A---GC-G-----A-TG-----A.....	5417	
H102.AG.NG.-.IBNG	.....-T-A.....GG--A---GC-G-----TG-----A.....	5374	
H1N.CM.95.YBF30	.....-CT.....-A---AGG-AT-----A.....	5440	
H10.BE.87.ANT70	.....-AC-C.....AGAGGA--AG--AG-----T-----GA.....	5907	
H10.CM.91.MVP5180	.....CTC-C.....TCTAACACAAGAGGAAGAG-A--AG--AT-----T-----GA.....	5891	
CPZ.CD.-.ANT	.....GGA-CC.....CGT--AC-C-GA.....	5282	
CPZ.CM.-.CAM3	.....-TC-CT.....GG-----GAG--AT-----GA.....	5260	
CPZ.GA.88.GAB2	.....GAA-----AG--AG--G-GA.....	5225	
CPZ.GA.-.CPZGAB	.....CTC-C.....-A-A---AG-TCC-----T---A-----GA.....	5913	
CPZ.US.85.CPZUS	.....GCAAGGAGA---CT.....-AG-----AGG-AT-----T-----A.....	5921	
CPZ.TZ.-.TAN1	.....	5469	
H2A.GW.-.ALI	.....AC-AGGGGAAGA...AATCCT-TCTCAGCTAT-CCG-CCCTAGA-AC-TGCAA--A...	6546	
H2A.DE.-.BEN	.....AC-AGGAGAAGA...ACTCCTTGCCAGCT--CCG-CCCCTAGAGCC-TGCA--A...	6556	
H2A.SN.-.ST	.....-CCAGGGGACGA...AATCCTTTATCAGCTAT-CC--CCCCTAGAG-C-TGCCA--A...	5998	
H2B.GH.86.D205	.....TC-GGGGGAGGA...AATCCT-TCTCA-CTAT-CCGCCCTTAAG-C-TGCCA--A...	6529	
H2B.CI.-.EHO	.....-CGGGGGAGGA...AATCCT-TCTCAGCTAT-CCGCCCTTAAG-C-TGCAG--A...	6527	
H2G.CI.-.ABT96	.....AACGCAGGCGGA...AATCCT-TATCA-CTAT-CCGCCCTTAGAG-CGTG-T--A...	5872	
H2U.FR.96.12034	.....A-CAGGGGAGGA...AATCCT-TC-CA-CT-T-CCGCCCTTTAG-CGTGCT--A...	6016	
MAC.US.-.239	.....-CTGGGGGAGGA...AATCCT-TCTCAGCTAT-CCGCCCTTAGA--C-TGCTA--A...	6456	
Vpr	.....P G G G N P L S A I P P S R S M L \$		Vpr
Tat	.....L G E E I L S Q L Y R P L E A C Y N		Tat ex1
SMM.SL.92.SL92B	.....-C-AGGGGACCA...AATCCC-T-G-ATCTAT-CCGTCC-CTAGAGACGTG-TA--A...	5889	
SMM.US.-.H9	.....TCTGGAGGARGA...AATCCT-TCTCA-CTAT-CCGCCCK-TAGAG-CGTG-TA--A...	5930	
STM.US.-.STM	.....-C-GGAGGAGGA...AATCCTTT-GCA-CTAT-CCGCCACTAGAG-AGTG-TG--A...	6104	
SAB.SN.-.SAB1C	.....CACCCCTG-GTATGGACCAGGAGCAGGAGGC-C-CCCC--GTCTGGGA----TTCAG--GGAGCTGCATCGGCCCTCCAGGCTGTGA...	6333	
TAN.UG.-.TAN1	.....C-CTAC.....G-AGAA--GAG--ATGGA-TC-GAG-GG--CGGG-TGGC...	6074	
VER.KE.-.AGM155	.....C-ATAC.....G--GA--GAG--ATGGACA--GGG-AG--AGACCAGGACGTGTCCACCAGGACTTGATTAA...	6100	
VER.KE.-.9063	.....C--TAC.....G--GAA--AG-GATGGACA--GGG-AG--AGAGC-AACCCTGTCTCCACCAGGACTTGATTAG...	6105	
VER.DE.-.AGM3	.....C-ATAC.....G-AGAA--GAGGATGGACA--GGG-AG--CGAGCAGGGCGCTACCACCAGGACTTGATTGA...	5603	
VER.KE.-.TY01	.....C-ATAC.....G--GA--GAG-GATGGACA--GGG-AG--CAGAGCAAATCGTGTCCACCAGGACTTGAGTGA...	5595	
COL.CM.-.CGU1	.....GTT.....CAT...GCGGGGCA-TGTC-CTTTC---CTACGAAAACCCCTTAG...	5516	
DEN.CD.-.CD1	.....AGACACTATCCAAACATAAGACCCTTA-GCCCTC-CAG-GA-G-GCC-TGA...ACG-GG-CAGCATAGACCCCTTCGCTGGCAATAAA...	5827	
GRV.ET.-.GRI_677	.....C--TAC.....G-AGA--GAG--ATGGACA--GAG-AG--...GCCCCACCCCTCCTCCAGGACTTGCATAG...	6039	
GSN.CM.99.CN166	.....GGG-GATACCCAGCAATTAGGCCCTCAAG-G-C-CC--CCC-CCA-ACAGT-ACTCTGTT-CTCATGCCATCCTGAGCAGCCTAGGAGACCCTCAAGATATCGCATGGATGAATAG...	5738	
GSN.CM.99.CN71	.....AGA-GGTACCAGTAATCAGGCCCTTAAG-G-T-CC--CCC-CCA-ATAGT-ACTCTGTTCTCATGTGTCTGAGCAGCCAGAGACCCTCAAGATACCGCATGGATGAATAG...	5726	
DRL.-.-.FAO	.....-C-GCAGGACGCTGGAACCCATTTCC-A-CCTTTCC--GG-CGAA-CAACA--TG--A...	5755	
RCM.GA.-.GAB1	.....TCC.....AGAGGAAGTA-TCCTCT-CG-TCC-T-CCGCAA-CG--AA-ATC...	5646	
RCM.NG.-.NG411	.....G-GA--GGAGGAA-TCCTCTTAG-TCC-T-CTTCA-TCG--AA-ATC...	5681	
MND-2.-.-.5440	.....GA-GGAGGAAGATATAACCCA-TTC-ATCCTTTCCG-GCCCGAA--ACCC-TG--A...	5704	
MND-2.CM.98.CM16	.....GA-GGAGGGGATGTTACCCTTAC-ATCCTTTCCG-GATCG-A--ACCCCTG-GA...	6139	
MND-2.GA.-.M14	.....GA-GGAGGCAAATATAACCCC-TT-ATCCTTTCC--GGCCGAA--ACCC-TG--A...	6066	
MNE.US.-.MNE027	.....TCTGGAGGAGGA...AATCCT-TCTCA-CTAT-CCGCCCTTAGA--C-TGCTA--A...	5924	
LST.CD.88.447	GCAGGAACAG--ACCCGACAGAAGCAGCAGCAAGACCC-CTAA----G-CAT-CAA--AG-CCGT--CTAA...	5122	
LST.CD.88.485	GCAGGAACAG--ACCCGACAGAAGCAGCAGCAAGACCC-CTA--G--AT-CAA--AG-CCGT--CTAA...	5122	
LST.CD.88.524	GCAGGAACAG-A-ACCCAAACAGAAGCAGCAGCAAGACCC-CT-G---G--AT-CAA--AG-CCAT--CTGA...	5119	
LST.KE.-.lho7	GCAGGAACAG-C-ATCCAAACAGAAGCAGCAGCAAGACCC-AT-A-G--G-CAT-CAAC-AG-CAGTG-CTAA...	6203	
SUN.GA.98.L14	GGAAATGTC-AC-C-AGGCCACCAACAGCAAGACCCAG-GCA-G-GG-CATTAGA--AG-CCTAC-AGACTAA...	6259	
MND-1.GA.-.MNDGB1	.....-GA-C.....-ACT-CCCC-CTC-GGATTCAGGCCA--GGAG-TAGATTATAA...	5541	
MON.CM.99.L1	TACCTGTCTT-A-GCCAATGAGAGGACTGTCT-CCG-CCC-A-C--TTC--TTC-ACAGCCGATCCAGACAATCCAGAAGACCTTCCAGATATAGAATGGATGAATG...	5751	
MON.NG.-.NG1	...RTTCGC-GGTACCCTYRGTGCGACCACTAAGAG-C-CT--CCC-GGACCCA----TCTGT-CCTCATGCAGACCCAGAGAGCCCTTTCACCCGCTCCAGATACCACCACGATGAATAG...T...	4364	
MUS.CM.01.1085	TATCCAACC-TAA-GCCTTTGAGAGGCACACAA-CCCCTCCTT-T-ATCAACCC-A-A-GCTGA-CCTGTACCTCCTTTGAGTCCTTCTAGGTACAGGATGGATGAATAA...	5700	
DEB.CM.99.CM40	GGGTATCCT-ACATTAGGCCATTGACAGGAAGAA-T--AGA--TT--GG-C---AATAA...	5681	
DEB.CM.99.CM5	AGATACCCC-A-ATTAGACCCCTAACAGGAAGAA-C--GAG-TC--GG-----AATAA...	5675	
SYK.KE.-.KE51	.....-GA-AC.....AG-G-ATACC--CATCC-TCTTAG...	5706	
SYK.KE.-.SYK173	.....-GA-AC.....AGAG--TACC-CCATCC-TCTTAG...	6042	

PLV Complete Genomes

471





PLV  
complete genomes



H1B.FR.83.HXB2 .....ACTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCCTAAACTGCTGT...ACCAATGTCTATTGTAAAGGTGGCTTTCATTGCCAAGTTTGTTCATAACAAAAGC 5955  
Tat ex1 .....L E P W K H P G S Q P K T A C . . T N C Y C K K C C F H C Q V C F I T K A Tat ex1  
H1A1.UG.85.U455 .....C-----A-C-G-----C-----G--C-T-----GT-----GG-----C--C--TC-G-A-----G 5400  
H1B.US.90.WEAU160 .....G-----G-----GA-----G-----G-----G-----G-----G-----G-----G 5953  
H1C.ET.86.ETH2220 .....C--A--C-----G-----ATC-A-T-----A--A--A-----T-T-----C--TC-G-----G 5345  
H1D.CD.84.84ZR085 .....CA--T--C-----GG-----A-----T-----A-----A-----G-----G-----G-----G 5476  
H1F1.BE.93.VI850 .....CT--T--T-----C-----C-----C-----A--T-----CGA-----TGG--C--T-C--G--G--G 5294  
H1G.SE.93.SB6165 .....C-----T--G-G-----C--C-----A--A--T-----GT-----GG-----C--TC-G-AC--G 5351  
H1H.CF.90.056 .....C-----C-----C-----A-----T-----C-----A-----A--G--C--TT--AG--G--G 5301  
H1J.SE.93.SE7887 .....CAG-----T-----T-----T-----T-----A-----T-----G--C--TC--CAG--G--G 5268  
H1K.CM.96.MP535 .....CA--C-G--C--A-----C-----A--C-G-----CG-----A-----A--A--C--TT--AG--G--G 5150  
H101.AE.TH.90.CM240 .....C-----T-----G-----C-----G--G--T--C-----A-----GG-----C--A--C--TC--G--A-----G 5522  
H102.AG.NG.-.IBNG .....C-----C--C-G-----C-----G--G--T--C-----T-----C-----T--G--C--TC--G--AC--G--G 5478  
H1N.CM.95.YBF30 .....T-----T-----C--A-----A-----C-----AT-----GA-----A--C--TT--TA--C--C--AG--G--G 5545  
H1O.BE.87.ANT70 .....GG-GCCC--T--C-C-C-T-----CC-G-TCC-----AT-----C-----GA--C-----A-----T-T-----G--G--G--G 6012  
H1O.CM.91.MVP5180 .....GA-GCCC--T--C-T-C-T--G--CA--CC--CC-----AT-----C-----GA--C-----A-----T-T-----C--AG--G--G 5996  
CPZ.CD.-.ANT .....-AC-CCT--A--TTA--C--TCCTGC-ACA--AGC--GC-A--C...-AT--C--T--C-----TGC-----T-AC-----CTC-C--C--C--A-----G 5387  
CPZ.CM.-.CAM3 .....T-----CTA-----G-----AT--C-----T-----C-----AT-----TGAG-A-----G 5365  
CPZ.GA.88.GAB2 .....ATGGATCCTATAGATCCTAGT--T-----T-----A-----A-----C...-ATTCA-----C--C--G--C--A-----TATGC--C--CC-GC--G 5350  
CPZ.GA.-.CPZGAB .....C--G-----A--C-----C-----A--G--A--T-----AT-----C-----GCT-----T--A--C--TAT--TA--C--C--A-----G 6018  
CPZ.US.85.CPZUS .....CA-----A-----G-----C-----G-----C-----AT-----C-----C-----TGTC-----T--CG--AG--G--G 6026  
CPZ.TZ.-.TAN1 .....TGAATCCATAGATCCTCAGG--CA--A--G--A-----GC-GCA--G--AC-----A--C--T--C-----A--C-----C--T--CG--A-----G 5593  
H2A.GW.-.ALI .....CA-----GG-A--C--AC-----GC-G--TT--AT--G--G 6603  
H2A.DE.-.BEN .....A-----GCGA--CA--T--C-----GC-G-----TCG-A--G--G 6613  
H2A.SN.-.ST .....A--T--C-----AC-----GA-G-----TT--AC--G--G 6055  
H2B.GH.86.D205 .....T-CA-----C--G--A--C--AC-----GC-----TC--T--A--G--G 6586  
H2B.CI.-.EHO .....T-CA-----C-----C--AC-----GC-----TC--T--A--G--G 6584  
H2G.CI.-.ABT96 .....T--A--T--C-----A--C-----A-----T--GT--A--C--TC--T--AC--G--G 5929  
H2U.FR.96.12034 .....T--A-----A-----A-----GC-A-----TC--T--A--G--G 6073  
MAC.US.-.239 .....CA-----AC-----GT-----TC--T--A--G 6513  
Tat ex1 .....T C Y C K K C C Y H C Q F C F L K K G Tat ex1  
SMM.SL.92.SL92B .....GC--T--T-----GC-----TC--T--GG--G--G 5946  
SMM.US.-.H9 .....G--Y--C-----G--A-----A-----T--GCA-----TC--Y--A--G--G 5987  
STM.US.-.STM .....G-----T--C-----A-----A--C-----GC-G-----TG--T-----G 6161  
SAB.SN.-.SAB1C .....T-AT--C-----TC--C-----GTT--C-----TATTC--C--C--CAT--A--G--G 6393  
TAN.UG.-.TAN1 .....AG-CTCCTC--GGACTTGCA-AG---TT---AC--A---ACT--A---A---T---C---G---A---C---A---C---T---TC--GCAG--G--G 6179  
VER.KE.-.AGM155 .....A--A--TAC--GA--A--CCTTGAGA--A---ACA--AT--A---T---C---A---C---T---A---T---C---C---T---CGG--G 6190  
VER.KE.-.9063 .....A--A--TAT--AGA---GCT---AA---AGA--AT--G---T---C---A---C---T---GC-----TC--GCA--G 6195  
VER.DE.-.AGM3 .....A--A--CTC--AAGCA--CCTG-AGCGG---ACA--A--G---T---G---T---C---T---A---C---T---GC-----TC--CA--G--G 6593  
VER.KE.-.TY01 .....AGACTATCAGA---CTGCAGA---AAA--AT--A--T--T---C---A---T---AC--C--T--GC-----C--TC--GCA--G 5685  
COL.CM.-.CGU1 .....AG--G---AATGA--A---GG---GGCC-----C--T--GC--A--C--TC--TT-----G 5586  
DEN.CD.-.CD1 .....CC---C---A---CA-----C---C---GCT--AC-----T--G--C--TC--CA-----G 5895  
GRV.ET.-.GRI\_677 .....ATT-CAAC--A---ACT--AT--A-----C---A---C---T--A---C---TG--GC-----C---GCAG--G--G 6117  
GSN.CM.99.CM166 .....AGACCCGGAAGAGTGTCTCT-GCCCAA---C---G---GCGGC---CGCC--CC--G--C...T-TGC-----C--C--G---GC--C---C--T--GT--G--C--TC--GG--G--G 5861  
GSN.CM.99.CM71 .....AGACCCGGCAGAGTGTCTTT-GCCTAAG---C--A--G---GCGGC---A--C--CC--G--C...T-TGC-----C---G---GCT-----T--GT--G---TC--GG--G--G 5849  
DRL.-.-.FAO .....T-CA-----GG--C---T--C---CGCT--AC--C--T--GC-----C--TC--TCA-----G 5812  
RCM.GA.-.GAB1 .....A-----A--C--T-----A-----T--A--C-----C--C--C---C--GCAG--G 5709  
RCM.NG.-.NG411 .....T-----AT--GC--T-----C-----GA--GTG--A--C--T--GC-----C--TC--TCAG--G---G 5744  
MND-2.-.-.5440 .....T--G---TC-----GTT-----A-----T--GC--G--C---T--GCAG--G---G 5761  
MND-2.CM.98.CM16 .....CA--T--C-----GA-----A--C-----GC--A--C---T--GCAG--G---G 6196  
MND-2.GA.-.M14 .....A--T-----C--G--A-----A--C-----GTG--C---T--GCAG--G---G 6123  
MNE.US.-.MNE027 .....CG-----C-----AC-----GT-----TC--T--A-----G 5981  
LST.CD.88.447 .....ACTG--G--C---AAC--A--G---GC-----TATTC---C--TCAGCAG--G--G 5200  
LST.CD.88.485 .....ACTG--G--TC---GAC--A--G---GC--C---G-----TATTC---C--TCAGCAG--G--G 5200  
LST.CD.88.524 .....T--ACTCCAAT--C---CCAG--AT--G---C--C---GTT-----C--TAT--C---CAGCAG--G--G 5197  
LST.KE.-.lho7 .....AG-CCTGCAA--G---CAG--AT--CC--T--GC-----A-----TATTC---C--TCAGCAG--G 6281  
SUN.GA.98.L14 .....CTTGGAG--C---GAT--A--A---GG---GG--GA--C-----C--T--C---C--TC--GCAG--G 6334  
MND-1.GA.-.MNDGB1 .....AACT-CTAG-AGAGTATTA-----TTGCAA-----GAG-AT--A---GG--C--G--A--C-----TATGC---C--TCA--AG--G--G 5643  
MON.CM.99.L1 .....ATGGAACCTGTGACCCAGACT---CCCAAAGAAC--A--C---CCC--GCAACC--ACGC--AC-----T--C-----T--G--GTT-----C--T--GA---C---C--C-----G 5876  
MON.NG.-.NG1 .....GCAASGATAGTGAACCCAGC--T---CCC---GA--C-----CTC--GCGACA--GCTC--CC--G---A---C---T--C--G--TT---GC--C--G--CY--T--G---C---C--CCA-----G 4489  
MUS.CM.01.1085 .....TGGATCCATCAGTAGAGGAGT-GCCAAAGGAAC--A--GA--G--TGCCGA--GGCC--GC-----ATC-----TC-----GCT-----GT--C--C---TCA--AG-----G 5824  
DEB.CM.99.CM40 .....CCTTTCACCTTTGCTTCAGA--G---AG-ATAG--T--CATT--A--GAAAGAG---T--AC--CCAC...C--GC---TATG--C---GCC---CATT--A-----GC--G---T---CA---G--G 5804  
DEB.CM.99.CM5 .....CCTCTCTTTTACATCATAGA--G--AGAAATAG--T--CATTCAAGGAAACAG---T--AC--CCA...C--T---TATG--C---GCC---ACTA--G-----GC--G--C--T--GCA-----G 5798  
SYK.KE.-.KE51 .....AAG--A--CATT--TCTC--A--GG---C--G---C-----G--GA-----C-----A-----TT--TAGA--C--TC--TCA-----G 5803  
SYK.KE.-.SYK173 .....AAG--A--CATT--T--AG--AAGGA--ACC---AC--A---AT--A-----T--C-----A-----GC-----C--TCA--G--G 6139





	/Rev exon 1 start	Tat, Rev exon 1 \ intron	
H1B.FR.83.HXB2	CTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGA.....CAGCGACGAAAGAGCTCATCAGAAC.....AGTCAGACTCATCAAGTTCTCTATCAAAGCAGTAAGTA.....		6052
Tat ex1	L G I S Y G R K K R R . . . Q R R R A H Q N . . . S Q T H Q A S L S K Q		Tat ex1
Rev ex1	M A G R S G D . . . S D E R E L I R T . . . V R L I K L L Y Q S		Rev ex1
H1A1.UG.85.U455	-----T-----A-----A-----CC-----G-C--C--A.....GGC--A-GA-----A-CT-A-C-----A.....		5498
H1B.US.90.WEAU160	-----T-----C-----C-----A-----AAC-----A-----A-----A-----A-----A.....		6050
H1C.ET.86.ETH2220	-----T-----A-----C-----A-----AGC--A-GA-----AA-CT-A-----CCAAATAATAGATGTAATG..		5461
H1D.CD.84.84ZR085	-----C-----C-----C-----AGC-----A-C-A-C-----A-----A-----A-----A.....		5573
H1F1.BE.93.VI850	-----AC--A--C--A-----AGC--AGTG-----AA-----C-----GTGTTA.....		5398
H1G.SE.93.SB6165	-----AG-----C-----G-A-C-----AGC--A-AGGC-----A-C-G-C-----G-----GAAAAA.....		5455
H1H.CF.90.056	-----A-T-----C-----AC--A--C-GCA-----AGTTTG--AGA-----AA-----A-----A.....		5398
H1J.SE.93.SE7887	-----T-----TC--C--CT.....GGC--A-----A-CT-A-C-----AATCAA.....		5371
H1K.CM.96.MP535	-----T--A-----A-----CC-----C-A--C-T-C-----AAC--G--AC-----A-C--CG-----G-----GTGCTA.....		5254
H101.AE.TH.90.CM240	-----AG-----C-----G-A-C-----AGC--A-GA-----AA-C-A-C-----A-----A.....		5620
H102.AG.NG.-.IBNG	-----G-----G-A-C-----AGCC--GA-----AA-C-G-C-----G-----GTAATA.....		5582
H1N.CM.95.YBF30	-----A-----T-----A-----C-----AGC--A--G-----A-CT-A-C-G-----AAA.....		5645
H1O.BE.87.ANT70	T--G--A-----AG-----GA-C-GCTGCT--A..AGC.....CA--CAGA--A--A-C-G-C-----G-----6103		6103
H1O.CM.91.MVP5180	T--G--A-----C-----A-----AGA-C-GC-GCT--GCAAGC.....TA--CAGA-A--A--C-G-C-G-----6090		6090
CPZ.CD.-.ANT	---G-G--T-----GAA-AGTCTGA..AG-AACA--C--A-AAC-GCTG-A..AGC--G-A-A-A-----A-C-G-----T-----5490		5490
CPZ.CM.-.CAM3	---A-----A--G.....GGA--C--CTGTGCT-GCTGCA..AGC--AC-CA-A--G-----CAGC-----5465		5465
CPZ.GA.88.GAB2	---T--T-----GGA--A--C--A-GGC--TT--G..GGC--A-A-GCA-----A-C--CG-C-----G-----5447		5447
CPZ.GA.-.CPZGAB	---A-----A-CCACA..AGA--C--CG--C--GCAGG.....A-A-A-A-----A--A-C-----6118		6118
CPZ.US.85.CPZUS	---A--T-----G-GCGA..-GCA-----C-A--GC-A--G.....GC-GA-A-----A-C--CGC-----GCCCTTTTGG..		6134
CPZ.TZ.-.TAN1	A-----A-----GAA-AG-----GCAA-TCTGCT-TA...-C-GT.....AC--A-A-----A-C-G-CG-C-----CCCATG.....ATA		5693
H2A.GW.-.ALI	AC-C--G--A-GG-----A-C-A--GGCA--C-----AGAA--TCTCC--AAGA-A.....T-A--G-----TCGT-----GC--G-CA--G--5693		6693
H2A.DE.-.BEN	GC-C--A--A--A--AGC--A--GGCA--AC-----AGAA--GACTCC--AGGA-A.....C-A--G-----C-TCCG-----GC--C-G-CA--G--5703		6703
H2A.SN.-.ST	GC-C--G--A--GG-----AAC--A--GGCA--A-----AGAA--ACTCCGGAAGA-A.....C-A--G-----TCGT-----GC--G-CA--G--5645		6145
H2B.GH.86.D205	TC-T--G--A-GT-----A-C-CTC--GAA-----TCTGC--AA-----A--G-----AC-GC-C-----GC--G-CA--G--5670		6670
H2B.CI.-.EHO	-C-G--G--A-GT-----AAC-CTC--GAAA-----TCTTC--AA-----G--A--G-----AC-AC-T-----GC--C--CG--G--5668		6668
H2G.CI.-.ABT96	AC-C--AG-A-GT-----AGCAACT-GAA-A-----AGAACT-CGAA--AGA--A--G.....G-A--TC-TT-----GC-----CG--5616		6016
H2U.FR.96.12034	-C-T--G--A-GT-----AGCAACC--GAA-A--G.....AGT--C--AAG--AG-T--A--G.....G--T--TC--T--G--C-----CA--5617		6157
MAC.US.-.239	-G-G--A-GT-----AGCAATCAGCAA.....AGAA--ACTCCGAAAA-G.....GC-A--G--A--AC-T--GC-----CA--5603		6603
Tat ex1	L G I G Y E Q S R T P S R K R . . . R R T P K K . . . A K A N T S S A S N K		Tat ex1
Rev ex1	M S N H E R D . . . E E L R K R . . . L R L I H L L H Q T		Rev ex1
SMM.SL.92.SL92B	-C-T-----TA-T-----CA-AACCC-GAA-A-----GTTAAGAAG-AG--A-G.....G-A--C-TT-GA-AC-----C-----6030		6030
SMM.US.-.H9	---G--A--A--G-----AGCA-C-ACGAA--A-----ACT--CGAAG-AGA--A--G.....G--A--AC-TT-----GC--G-CA-----6071		6071
STM.US.-.STM	TC-T--A--AA-----AGC-ATCA-GAA-A-----GTTAAGAA--AG--A--G.....T--C--AT-----GC-----CA-----6245		6245
SAB.SN.-.SAB1C	-C-----TCG-----TATGTCCT-G-CA-----AGAGCTTCT-AGAAG.....ATTT-G--A--TCAAGT-TCTCT-C-CA-C--GTA.....		6477
TAN.UG.-.TAN1	-----A-----CATGTCTCT-G-ATC-----AGA-CTAAG-AGAA-----TCA-A-----GAA-CT-G-----C-----6266		6266
VER.KE.-.AGM155	AC---T--TA-----CATGCCTTT-G-ACC-----AGAAAGAAG-TT--TCG.....G--G--GCAT-C--G--C-GC-----6274		6274
VER.KE.-.9063	-C-----TG-TA-----CATGCCCT-G-ACC-----AGAAAGAA--T-T--G.....C--A--GCCTA-----A-CAGG-C-----6279		6279
VER.DE.-.AGM3	A-----TG-TA-----CATGCCCT-G-ATC-----AGAAAGAAG-TT--CG.....CT-G--GCTT-C--GA-CA-----A--G-----5777		5777
VER.KE.-.TY01	-C-----TG-TA-----CATGCCCT-G-ACC-----AGAAAGAAG-TTCG--TCG.....CT-A--TTG--C--T--CAGC-C-----5769		5769
COL.CM.-.CGU1	GC-G--A--CG-----AAT-AG.....A-TGCT-C-CAA.....CA-----6537		5637
DBN.CD.-.CD1	AC---T--TGTT-----GCTCCAA-A-AG.....AGA-C-A--GCTCG--AAGC.....G--AC-C--G-----AAG--G-----TTG...ATGGATCTCCTCTCA		6006
GRV.ET.-.GRI_677	-----AG-GCGT-----CATGTCTCT-G-AAA.....AGAAA-AC--CA-----A.....GA-A-----ACC--A--CGGC-----6201		6201
GSN.CM.99.CN166	AC---TC-T--CTCCATG-C-GATCA--C-AG..AGGA-A--C--A-----ACAG-ACCTGATCCT--CGT-T-GTC-A-TCAAGACTCTG--T--GTCC-GTA.....		5967
GSN.CM.99.CN71	AC---TC-T--CT-CATG-C-GATCC--C-A.....CGGA--A--TC-GA-----TCAG-ACCTGCTCCT--CGT-C-GAC-A-TCAAGAC-CTAT-C-AATCC-GTA.....		5955
DRL.-.-.FAO	-----G--ACGT-----AATGTC-GC-G--CC-----A--A--ACCGCA--CGTGGTT..TCAAGATATCTGAGAC--GGT-ACA-G-CTGTGGCAAG--		5909
RCM.GA.-.GAB1	AC-----AAAT-----CTTCC-G-GCAA-AC-----AGAA--TCT-AG-AAG-A.....AATAAGGC-G-T-AATT--CT-TAC-----T-----5800		5800
RCM.NG.-.NG411	T-----AAA-----CTTCT-G-GAA-A-----A--G-C--C--AGGAA-T--GAAGAAG-A-----T-A--C-TT-GA-AC--TG-----5844		5844
MND-2.-.-.5440	T-----GG-ACAT-----CATGTCT-C-GA-A--G.....-GA-CT--C--GG-T--T.....GAGAAAGTA-CA-GCA--AGTCAAG-AGCT-TGGGAAG--		5860
MND-2.CM.98.CM16	-----AG-ACAT-----CATGCCT-C-GAAAC--G.....AGATCTA--CAGAGA-T--T.....GAAAAA-TATCAGAAG--AGT--AG-G-CTGTGGGAAG--		6296
MND-2.GA.-.M14	T-----G--ACAT-----CATGTCT-C-GAAC--G.....AGAACT--G-AGAGA-T--T.....GAAAAA-TATCAGAG--AGT-AAG-AGCTGTGGGAAG--		6222
MNE.US.-.MNE027	---G--G--A-GT-----AGCA-TCACGA-A.....AGAA--ACTCCGGAAGA-G.....GT-A--G--A--AC-T-----GC-----CA-----6071		6071
LST.CD.88.447	A-----G--ACGA---TATGTCC-C-G-A-----A-AA--GTCGCCCAGGTATTT.....GAAG--A-GCAAGA--CTGTGG--CGG----A.....5287		5287
LST.CD.88.485	A-----G--ACGA---TATGTCC-C-G-A-----AGAA--GTT--CCCAGGTATTT.....GAAG--A-GCAAGA--CTGTGG--CGG----G.....5287		5287
LST.CD.88.524	G-----G--ACGG---TATGTCC-C-G-ACC-----AG-A--GTT--CTACATA--TT.....GAAAC-G-GCAGGA--CTGTGG--CAG----AA.....5284		5284
LST.KE.-.lho7	G-----GC-GCA---TATGTCC-C-G-AAAC--G.....AGA--GTT--CCT-GATA--TT.....GAGAC--GCAGGA--CTGTGG--CAG----G.....6368		6368
SUN.GA.98.L14	G-----G--CA---TATGTCT-C-G-G-AC-----AT-A-TC-ATATC-----AGG-A-GCAAGAGGCTGT-CGAAG-----6418		6418
MND-1.GA.-.MNDGB1	T-----A--AAGG---CCATGTCT-C-G-AAAC--T.....GTA-C-G--CTAA-A-GA-----TA-CTGG-AG-GGT-AAGAAGCTAT-CGAAG-----5736		5736
MON.CM.99.L1	-C-C--G--T--A-----ACG-AA-----GAA--GCG-CGT--C-GTC..CCTGGCCT-AGC-GCAG-A-GAA-C--GC-AG-----G-----5976		5976
MON.NG.-.NG1	AC-T--A--A-----GGCCGAAA--G.....-GC--CG--TCC-AGA--CT--CGATTGGCGA--G--GCT-----G-A-AA--CGGG-GTCA-----G.....4592		4592
MUS.CM.01.1085	A-----T--T-----CGTTTCA-A-A.....-GATCG--GCT-A-TCG.--CCTGCCGAGCCATT-C-GATCCTGTA-AAA--A-GT--GTAGC.....5920		5920
DEB.CM.99.CM40	TC---G--T-----CCACGC-G-A--AGAGGAAGTCA--A--GA-TACACG-CAGCTGCTGAA--A-TCAG-C-G-TCAAGA-T-TGT-CCAG--C-GTA.....5913		5913
DEB.CM.99.CM5	TC---G--T-----TCACGC-G-A--AGAGGAGA-GCG--TC--CTT-GC-GC--TGATAAGGA--TCAAG-GTTTATAC-AGAG.....5898		5898
SYK.KE.-.KE51	-----G--AA-A---T-C--G--GAAA-----A--ACCG-AGAACAAGCTGCTGAGACTT-GCTG-G--AGCT-AA-AACT.....G-----5894		5894
SYK.KE.-.SYK173	A--G--A---A-T-----C--ACCA-G-AA-C-----GCAGCT--C--TAT-----T--G--GG-T-A-AGCGC-C-T-CAGG-----6221		6221



PLV  
complete genomes



	/ Vpu start (ACG start codon in HXB2)	
H1B.FR.83.HXB2	...GTACATGTAACGCAACCT...ATACCAATAGTAGCAATAGTAGCATTAGTAGTAGCAATAATAATAGCA...ATAGTTGTGTGGTCCATAGTAATCATAGAA...TAT...AGGAAAATAT	6161
Vpu	T Q P I P I V A I V A L V V A I I I A I V V W S I V I I E Y R K I	Vpu
H1A1.UG.85.U455	TTA-CTT-----T-AC-----T-GGA---CTGG-----AC-GGC-GA-----C---CT-----A-----A-T-----GGT-----AAA....AA---T-GC	5612
H1B.US.90.WEAU160	.....T-----T-----T-G-A-----T-----C-----GG-----C-----	6159
H1C.ET.86.ETH2220	..GT-GAT-TACTA---AAGTAG--TAT--G-A---T-----CA-----C-----C-----A-----C-TAT-----T-G-	5573
H1D.CD.84.84ZR085	.....T-----T-----T-A-----T-----T-----GC-----A-----T-----C-T-G---A	5682
H1F1.BE.93.VI850	ATA--T-----T-TC-TA...T-GTT-GC-A--GG---C-----A-----C-----C-----A-T-----TAT-----A---C-GG	5510
H1G.SE.93.SE6165	TTA--T-----T---GT-A...T---GT---T-----G---A-----T-C---GC---C-----A-----T-T-----G---A	5567
H1H.CF.90.056	...T-TCATA--T-T-TATA...T--GG-T---G-AT-GG--CGCT-G---AC-TTT--C---GCC...G-A---A---TAT-----AA---T-GG	5504
H1J.SE.93.SE7887	GTA-C-T-----T-AT-----T-G-A-----C-----TA-----AC-TTT---TT-CC-T...G-A---A-T-----GTAT-----A---C-G-	5483
H1K.CM.96.MP535	GTA-----T-GTGT-C...T-G---TTC-AT-G---C-TT-G---C-CT---T-GC---A---A---A---TAT-----T-GG	5363
H101.AE.TH.90.CM240	TAA--T-----T-AC-----T-GGA---TAGT-----G-C-GA-----GC---CT---A---AA-T-----GCT-----GT-----A-----C	5732
H102.AG.NG.-.IBNG	TTA-C-T-----G-T-----T-A-----ACT-----G-C-----T-C---GC---C-----AG-----TAT-----A	5694
H1N.CM.95.YBF30	CCT--T-----T-----G-TGTC-T-G-G-T-CA---G---G-C---G-AT-C---TAT-G---A---A---A---A	5745
H1O.BE.87.ANT70	.....ACGC--T---T-A...GGGACC-GC---A-ATTA-TAGT-CTTTGC-GT-T---AAT...G-A-T-A---GGGT-TA-TC-T-G-A---TTAG--CA-A	6209
H10.CM.91.MVP5180	.....ACGC-G-T---T-AA...GAGAACC-GC---CT--A---T---AGT-CTTTGTGC-T---AAT...G-C-A-A---TGT-TAAC-T-G-ATT...TT-G-GC	6193
CPZ.CD.-.ANT	.....-TCT--T-ACTAA...TTT...GAGTAT-CTTTTC-TGCCT-TAGT...AC-A---ATT-CTGT--ACCTAT-CTCTATA-GCTTTAT-AA-T-TATA	5590
CPZ.CM.-.CAM3	ACTTC---CT-T-T-A...TGGGA-CA-A---GCT--A-T-GTA---G-A---A---T---T---T---T---T---A-CC---GGG---CTT-T-A---G...GG...A---GG-A	5577
CPZ.GA.88.GAB2	..TCCC-TCT-T-T-TTT-CATGTGGGT-GCTA---G-C-A-T-G-A---GTACTCTTT--G-T---AAT...G---GGT-T---GGA-CTCTGTCTATA-AAGGTGG-AA-G-CACA	5572
CPZ.GA.-.CPZGAB	.....T-T-ACT-TG...T-GTCGGT---TTC-CA--TT-G---GCT---TGC-TGGAAC...TTG-A-A---GGATACA-T-T-A-TGG...GGA---AGAAGGTAT---G-C	6225
CPZ.US.85.CPZUS	.....CT-T-TTAA...TGGTTTGA-A---GTT--A---G-A---A-GGT-TT--T...G-GA-AA-C---GGAC---TGCT-G-TTG...GGAGACAA-TA---A	6244
CPZ.TZ.-.TAN1	AAAA--GTA--GGGAAGTGTGTC-CTAATG-CA---GC--TC-TTGTA--T--C-GATTT---GGG-G...GGCT-GC-AATAGGT---GT--A-G-AG...G-G...TT-GA-A	5805
H2A.GW.-.ALI	.....	6693
H2A.DE.-.BEN	.....	6703
H2A.SN.-.ST	.....	6145
H2B.GH.86.D205	.....	6670
H2B.CI.-.EHO	.....	6668
H2G.CI.-.ABT96	.....	6016
H2U.FR.96.12034	.....	6157
MAC.US.-.239	.....	6603
SMM.SL.92.SL92B	.....	6030
SMM.US.-.H9	.....	6071
STM.US.-.STM	.....	6245
SAB.SN.-.SAB1C	.....	6477
TAN.UG.-.TAN1	.....	6266
VER.KE.-.AGM155	.....	6274
VER.KE.-.9063	.....	6279
VER.DE.-.AGM3	.....	5777
VER.KE.-.TY01	.....	5769
COL.CM.-.CGU1	.....	5637
DEN.CD.-.CD1	TCAC-GAT-T---GCTTTT-AAATA--CATT--TTGCTT-G-TTGTATA-T--TA-TATT-C-G-G---TTCCTGACCCA..TA--ATGC-TAAG---AGTT-GGACCAATAGAATACGTTTGTCT-GG	6134
GRV.ET.-.GRI_677	.....	6201
GSN.CM.99.CN166	.....GTA-GGCA-GCATCCT-C---G-TTGGTGGTGGGGTGTCTATTATA-CCT-CA-A-ATCTTTGT--GCGC-TCCTGCTCTC--CTTAGCTT--G-T-AG	6071
GSN.CM.99.CN71	.....GTA-GGCA-GAGTGCT-C---TC-CTGGTGGTGGGGGCTGTCT-T-ATA-CCT--A-A-AT-T-TGCT--GCA--CTTTGCTCTC--CTTGGCGT--G-C-AG-	6059
DRL.-.-.FAO	.....	5909
RCM.GA.-.GAB1	.....TGG-G-GCCCTGG-CTAGTA-TGCTT-TAG---CTC-GCT-A-AAG...G-TT--A-AGATCTTT---T-CA---ATTAAGGAA-ATAGAAT-A--T--G..	5899
RCM.NG.-.NG411	.....TGG---GTCCTGGTCT---C-TGCTT-TAG-TGCT--GCTTATAAG...G-GA-A-GAGATATA--TAG-GAAG-TTTTATTTGTAGAGTTTAT--A---TAGA	5945
MND-2.-.-.5440	.....	5860
MND-2.CM.98.CM16	.....	6296
MND-2.GA.-.M14	.....	6222
MNE.US.-.MNE027	.....	6071
LST.CD.88.447	.....	5287
LST.CD.88.485	.....	5287
LST.CD.88.524	.....	5284
LST.KE.-.lho7	.....	6368
SUN.GA.98.L14	.....	6418
MND-1.GA.-.MNDGB1	.....	5736
MON.CM.99.L1	.....GAATTATTGGTGGTC-C-G-T-GC-A--ACTTAC-GCC-G---CCTCATAGCCC-CCCCGTAG-AGCTTGGGCTTGGTGGCGATAT---AAAATC-CT--G-GG-	6082
MON.NG.-.NG1	.....GAAT-CCTATTGG---C---CA--TCT--TTGGTGTG-T--T-TGCTTT...GC-C-TTGACCC-ATACCTCTC...TATCAGTCTT-T-AGG	4686
MUS.CM.01.1085	.....GAACTATTGGTATT---C---TG-CA-T--A--GGC--CTACTTTGTG--T-C-A-A-TTG--T-T--GT-AGCTT-CCAA...AGGT--TGT-AGC	6017
DEB.CM.99.CM40	.....	5913
DEB.CM.99.CM5	.....	5898
SYK.KE.-.KE51	.....	5894
SYK.KE.-.SYK173	.....	6221



Accession	Sequence	Annotation	Position
H1B.FR.83.HXB2	TA...AGACAAAGAAAAATAGACAGGTTAATTGATAGACTAATAGAA.....AGAGCAGAAAGACAGTGGCA.....ATGAGAGTGAAGGAGAAATATCAGCACTTGTGG.....	/ Env gp160 start signal peptide	6257
Vpu	L...R Q R K I D R L I D R L I E...R A E D S G...N E S E G E I S A L V...		Vpu
Env		M R V K E K Y Q H L W...	sig pept
H1A1.UG.85.U455	---AG---A-----T-AA-C--A---G-----T--G--T-C-GAG-A-T-ATCCT.....		5708
H1B.US.90.WEAU160	-----A-----A---G---C-----G--TCAGGA--A-T-ATCA.....		6255
H1C.ET.86.ETH2220	-----G-----T-----A-A---ACT-GG-----A-----T--G--T-C-GAG-A-T-GTCAA.....		5669
H1D.CD.84.84ZR085	A-----G-----T-----A---G---A-----A-----G--T-A-GAG-A--ATCAA.....		5778
H1F1.BE.93.VI850	-----G-----A-T-AA---TA-A-A--A---G-----G--G--TGC-GAG-A-T-G-CA.....		5606
H1G.SE.93.SE6165	G-----A---G-A-G-----GA-A---C-----A---G-----C--G--T-C-GAG-A-T-G-AA.....		5663
H1H.CF.90.056	-----G---AG-----C-----A---GG-----ACGAG-GTG.....T---C-C-GAG-A-T-ATCCA.....		5600
H1J.SE.93.SE7887	-----G-----T-A-----A---G-----T-ACGAG-GTG.....T---C-C-GA--A-T-G-CA.....		5579
H1K.CM.96.MP535	-----A---A-G---A-T-----A---A---G-----T-AGAGTGAGGGG-ATGCA-A.....		5450
H1O1.AE.TH.90.CM240	-----G-----G--A-G--A---G-----A-----C-C-GAT-A-T-G-CCA.....		5828
H1O2.AG.NG.-.IBNG	GG...A---G-AG-----C-----A---G-----T-----T--G--T-C-GA--A-T-ATCCA.....		5790
H1N.CM.95.YBF30	A...TTG--GGA-----A-ACACA--GAC-G--A---G-----A-----T--G--TGC-GATGTTG-T-AT.....		5841
H1O.BE.87.ANT70	AG...GA---GAC-G--AG--G--AGA---ACT-GAGAGGT--AG-AGAATAAGAGAAATTAGGGATGATAGT--CT-TGAAA---		6326
H1O.CM.91.MVP5180	A-AGA-A---GAT-G--GG--GCA-GA--ACT-GA-AG-T--AGGAGAATAAGGAAATCAGGGATGACAGT--CT-TGAAA-T.....		6313
CPZ.CD.-.ANT	AGCAGCAG---T-G-T-ATA-G--AAATCAAAGA-TCA--GA--TGTTAAGTAGA.....CTTAGTATAGA-A-TGCAATAGAAGA---		5701
CPZ.CM.-.CAM3	A-GAAGA-G---AC-G-GC-C-A-AAA-T-GGAC--TT-A--GA-AG-ATAATATCA.....CAG-----T---A-----A-----T---GATCA-AGG-AGC-T.....		5685
CPZ.GA.88.GAB2	A-GAAGAG-----A-C-----TTGA-T--A--ATCACAGAA.....T---A-----A-----T--G--GACAAG-AGACGC--CAA.....		5671
CPZ.GA.-.CPZGAB	ATAGGCTTG-G-C-G-----G-----T-ACCT-ATATTGAGA.....GA-AG--C--AGA-A-TGGAA.....T-A-AGT-ATGG--A-GAA-AA-AGAGA.....		6320
CPZ.US.85.CPZUS	A-GAA-ATACTCA-C-----G-TT-A--ATCTTTTGAAGA.....-TCAG-AT-AGAGAA-AGATAGTGGA-----T-A-AGTGATGG--A-GAA-AA-AGACT.....		6342
CPZ.TZ.-.TAN1	GGGAA--G---CATC--G--TATTAGA--GGCTAGCTAG--G-TT-AGCATAGAC.....-T-G--T--AGAA-----T-A---T-TAATT.....		5892
H2A.GW.-.ALI		T-AT-TCTAG-AGAA-TCA-CT-.....	6717
H2A.DE.-.BEN		T---CCTGG-AG-A-TCA-CT-.....	6727
H2A.SN.-.ST		AA-T...TGTGG-AG-A-TCAACTA.....	6169
H2B.GH.86.D205		T-CAT-T-T-AGCAG-CGCCT-.....	6694
H2B.CI.-.EHO		T-CAC-TGT-A-TA-T-ACCTA.....	6692
H2G.CI.-.ABT96		T-CAT-TCT-GGCA-TCA-CT-.....	6040
H2U.FR.96.12034		T-CATGTCT-GGCA-TCA-CT-.....	6181
MAC.US.-.239		T--GATGTCT-GG-A-TCA-CT-.....	6627
Env		M G C L G N Q L.....	sig pept
SMM.SL.92.SL92B		T--C-TGTCC-GGA-TTCACCT-.....	6054
SMM.US.-.H9		T--GATGTCT-GG-A-TCA-CT-.....	6095
STM.US.-.STM		T--CCTGCC-GGAA-TCAACT-.....	6269
SAB.SN.-.SAB1C		GTAT--GCT--TTACAG-ACT-.....	6501
TAN.UG.-.TAN1		T--GACC--TAAG-GGAAAAG--.....	6290
VER.KE.-.AGM155		T-ACA--G-TCTTAGGAA-T-TT-.....	6298
VER.KE.-.9063		T-A-ACCTCT-TTAGGTA-AGTT-.....	6303
VER.DE.-.AGM3		T-A--CTGACATTA-TGA-AG--.....	5801
VER.KE.-.TYO1		T-AG-T-TACAATAATAACC-TA.....	5793
COL.CM.-.CGU1		T-CTT-G----TCTTTA-AAAC.....	5661
DEN.CD.-.CD1	--TTTGC-GTTGTT-TT-CCTGG-TAGC-GCCATAG-TG--GGGT-TCTTGCTTATAGAGCCATAAA--TTATAG--AGAGCTT-GGTAT--A--C-A-G-TT-TGGAG-ATAG-TAGTG-AT....		6260
GRV.ET.-.GRI_677		T--G--G--TG-TTATAAAAATA.....	6225
GSN.CM.99.CN166	GGGTA-A-GG--A-CC--A-CCA-CACA-G-A-CAGT-A-T-G-CT-----T--A-----T-AG--GACAGT-G-A-CTAT.....		6149
GSN.CM.99.CN71	GGATA-AGGG--AGCC--AGATACCA...G-A-CAGTTA-T-GGCT-----GTT-A-----T-AT--GGA-AGT-GGA--T-T-.....		6134
DRL.-.-.FAO		---T---A-GAT-ATTAGACATATA..	5933
RCM.GA.-.GAB1		T---T--G.....	5908
RCM.NG.-.NG411	AGGTA-AT-TTTAG-T.....	T--CT--G.....	5970
MND-2.-.-.5440		T-AACAGT-AG-AT-TTAAGA-AT.....	5884
MND-2.CM.98.CM16		TTAACAGT-AG-AT-TTAG--AC.....	6320
MND-2.GA.-.M14		TAATAGT-AG-AT-TTAGACAT.....	6246
MNE.US.-.MNE027		T--GATGTCT-GG-A-TCA-CT-.....	6095
LST.CD.88.447		T-A-CTGCC-GGA-TAA-ACTT.....	5311
LST.CD.88.485		T---CTGTCT-A--T-C-ACTT.....	5311
LST.CD.88.524		T--CATGTCCAGGTA--A-A-T-.....	5308
LST.KE.-.lho7		AG-T...GC--G--CAGGA--AG-A.....	6392
SUN.GA.98.L14		AA-T...CGG-GC-CTG-AA-A-T-.....	6442
MND-1.GA.-.MNDGB1		T-A--TGTCAGGT-TT--TCTT.....	5760
MON.CM.99.L1	-TAAA--AT-GACC--GA-ATACA-AGGC-GAT-CAGA--CAT--GAGAAGGAGAC.....T-ACTC-GGGGT-ACACA-AATCTGAA.....		6168
MON.NG.-.NG1	AGCAC-AG-GGGTGTCCCGCTT-GAACAG...-CTTT-A--GGCTT.....GT-CAG.....Y-TATC--GAGYAT-ATTCA-G-TATGAG.....		4766
MUS.CM.01.1085	CC...AGA-GGT-G-GG-CAGTGTCA-CCGCCTCTTGA-GAG-G-GACGGTGACAGCGGCATCTTGA--AT-CC--GGA--A.....		6100
DEB.CM.99.CM40		GCAA-TCTGC--GTAA-TGCTAAAAGGGAA.....	5942
DEB.CM.99.CM5		GT--A-CTGC---GTAT-TAGTATACAAAT.....	5929
SYK.KE.-.KE51		CACA-T-A-ACTT-T-A-TAT-A-AA-TGTAGT.....	5928
SYK.KE.-.SYK173		AAA-T--CAGCT-T-AGAACT-AC.....	6245

PLV complete genomes

Table with columns for sample ID, sequence, and coordinates. Includes headers for 'signal peptide', 'Vpu end', and 'Env gp120 start'. Lists various samples like H1B.FR.83.HXB2, H1A1.UG.85.U455, etc.



PLV complete genomes

Table with 3 columns: Accession ID (e.g., H1B.FR.83.HXB2), Sequence (e.g., TGGTAAATGTGACAGAAAATTTTAAACATGTGG...), and Length (e.g., 6580). The table lists numerous PLV complete genomes from various sources and dates.



V1 loop

H1B.FR.83.HXB2	TGTAATAAATTAACCCACTCTGTGTTAGTTTAAAGTGCACT.....GATTTGAAGAATGATACTAATACCAATAGTAGTACGGGAGA.....	6662
Env	V K L T P L C V S L K C T . . . . . D L K N D T N T N S S S G R . . . . .	gp120
H1A1.UG.85.U455	-----G-----T-----C-----CG-----G-T-----CA-AACATCACCATTAACA-CACC--T--CA-C-C-----	6101
H1B.US.90.WEAU160	-----C-----T-----AATGTGAATGTGACTA-----G-----T--AG-G.	6672
H1C.ET.86.ETH2220	-----G-G-----C-----T-A-GCTATCAAACAATAC-AA-GTT-CAA--A--G--	6065
H1D.CD.84.84ZR085	-----G-----G-----C-----T-----GATGCATCGAGAAACAGCACTG-T--A-C-GC-CA	6177
H1F1.BE.93.VI850	-----G-----T-----G-C-----T-----AATGCCACCAATAACAG-CA-G-A--GCCAGGGGCA	6002
H1G.SE.93.SE6165	-----GC-----T-----CC-----C-----T-----GATGTAAACCAACAA-GCAAC--A-GAA--A--GC--TG--A--C-CT-AA-CCAACAATAGCACTGTGGATAACCCA	6104
H1H.CF.90.056	-----C-----C-----C-----T-----AATGTGAGAAACAAATACC-CT--C-GCAC--GC-G-	5996
H1J.SE.93.SE7887	-----GA-----T--A-----C-----T-----T-G-AACATCACTAGTAACAG-AAT-CA-C-AG--AC-G--GT	5978
H1K.CM.96.MP535	-----G-G-----T-----C-----C-----T-----GATTACAAAGGAACCA--AGC-CC--A--G-A-C--G--C	5861
H101.AE.TH.90.CM240	-----G-----T--T-----C-----T-----CAATGCTAATTTGACCA--GGC-GT-GCA-A--C--GT-TC--AC	6233
H102.AG.NG.-.IBNG	-----G-----T-----C-----C-----G-T-----TCA-AATTTCAACAACAGCT-CAGC--C-G-AGC-ACCTA--T-G-	6189
H1N.CM.95.YBF30	-----T-----T--A-----A-C-A-GCTT--T-ACGATAGCTATGGGGAG--AAG--C--ACA-A-	6240
H1O.BE.87.ANT70	-----TC--A-G--TTTC--G-----ACAAA--GG--T--AAACATAGCTGGAACAAACAATG-A.	6692
H1O.CM.91.MVP5180	-----AG--A-G--TTTCT-A-----ACAAA--G--C--TGATGATCTGCAACAAATA-AACAGGC	6682
CPZ.CD.-.ANT	-----C-----A--TA-G--A-A-AAA-G--T--T--AGGATACAATGGAAACACC-ACA-CACCA.....GT-CA-CA-CA--TACA	6088
CPZ.CM.-.CAM3	-----T--GC-G--G--A-A-----C-CCA-GG-A--T-GAAAGGTTACTTTCAATAGCACTTCT--AGG-A--A--TTC--C--TG-CTACT	6093
CPZ.GA.88.GAB2	-----G-A--TT-A-----GCAA--C--T-----TCTATAATGCTACT-GAAACG-A-C-A-G--C-GGA	6064
CPZ.GA.-.CPZGAB	-----G-G-----TT-A--A--C--GC-----G-AAGGCTAACTTTAGC.....CAGGCA-AA-ACCTAAACA	6717
CPZ.US.85.CPZUS	-A-----G-----T-G-----C-A-G-CT--TCT-AAACCC	6706
CPZ.TZ.-.TAN1	-----A--T--T--CA-A-AGA-G-CA--T..ACAATGACTAATACCACAAT--A-CC	6264
H2A.GW.-.ALI	-----C--G-----A--TT-A--A-AGCAA-G-----T-GCAA..ATAAGCACA--GAGC-CA-CCACAT-CCCG-G-CCAG-G--C-CACTC	7086
H2A.DE.-.BEN	-----C--GC--A--T--A-----GGCAA-G--T--T-GCAGG..GTTCAAGGA--ACC-C-CCCCG-A-CCC-GG-CCTCG--TC-ACA-CCTCGAGACCACCCACATCC	7117
H2A.SN.-.ST	C--C--C-----A--CT-A-----AGCAA-GCCT--T-ACAGC..CAACTGCAA-AAAC-CA-CCTCC--ACCA...CA-CC-CC-CAACAGC	6538
H2B.GH.86.D205	-----C--C-----G-----GCAA-G--C--T-G-AAAACCGAAACAAACCCAGG--TGCCAG--G--C-----CC-C--AGCCTACT-CCACCTCT	7075
H2B.CI.-.EHO	-----T-----T-----G--C-CAA-G-----T-A-AAAGCGTGGAGCTCA-C-AGC--AG-GACC--CGT--TCCGCGTCCCT-A-ATCTTCTACTCAG	7076
H2G.CI.-.AHT96	-----T-----A--T--G--CA--GCCA-G-GA--A--AAACTGAGACAACA-A-G-GGTCTAACAGGG--CC--C-GTA-CCCCT-CA-C-GTCCCACAACCTAAAATGGTGACAGCT	6442
H2U.FR.96.12034	-----G-G--A--C--T--A-AGCAA-G-GA--T-G-AAACTGAAACAAATAGA-G-GGATTAACAGGA-CAG-AGCA-CA-CA-CTA-CCC-GCAATGCAAGAATCAAGT	6574
MAC.US.-.239	-----T-----T--A--CA--C-A-G-GA--A-AAAAGTGAGACAGATAGA-G-GGATTTGACA-AATCA.....TA-CA-CAACAGC-TCAACAACATCAACGACAGCA	7017
Env	V K L S P L C I T M R C N K S E T D R W G L T K T S . . . . . I T T T A S T T S T T A . . . . .	gp120
SMM.SL.92.SL92B	-----G-----T-----T-A--A--GC-A-G--A--T-A-AAAATGAGACAGACAGA-G-GG-TTGACA-GAGCAG-T-C--C--C--TCACC-ACTACTACTAGCCCTTAACCTGCTGCTAGCCCA	6462
SMM.US.-.H9	-----C--C-----A--A--CA-A-G-G--T-A-AAAAGTGAGMCAAGMAGA-G-GGTTTAAACAGAGGAGSAC-AGCACA-CA-CAACACA-ACATCAACAACACCACCATCAACCAATAATAGCA	6503
STM.US.-.STM	-----G--T-----A--A--GCAA-G-G--T-A-AAGAAATGAACACAGACA-A-G-GGATTTGACAGGG--A--AGTA-CA-CAGTGAACC-ACAGCAGCAGCAGCAGCAACA	6665
SAB.SN.-.SAB1C	-----T--GC-CT-A--A--G--CA-A-AAA-G--C-----TA-CGCTTAGAAGGAGGA-CAGCA-CA-CAACAT-ACCAT-A-CATCA-CAGCAC-CCCCGAGGTAGTGAGT	6900
TAN.UG.-.TAN1	C-----C-G--A--TA-G--CA-A-AGA-G--T--T--AAATTAACAAGTACC-CCCCT-CATCAAG--CCC--ATCA--C--C-CTAC.....GAT	6677
VER.KE.-.AGM155	-----GC--T-A--T--A--A--C--AAA-G--T--TGTAGAGTTAAAAGGCTCC-CAACTCT-CCCCG-A-CCT-T-C--CGGAG-AACC-A-.....CTA	6682
VER.KE.-.9063	C-----T--T--A--A--A--AAA-G--GC--TGTAGAGTTGAACTCAAGTAGAAGA--CCCAAC-C-A-C-T-AGC-TCA-CA-CAAAAAGCT...CCAAAAACAGGT.....GAT	6696
VER.DE.-.AGM3	-----T--A--TT-G--CA--AAA-GTCC--TGTAGAATGAACTCCTCT--GCCT-CC-CCAC-C--A--GT-CC-CGGCTCAAAC-CCAATATCACAGCCTCAACACCAACT.....TTG	6206
VER.KE.-.TY01	-----G--G--T--A--A--CA-C-AGA-G--T--TGTAGAGTTAAAATCCACCAAGAGAA-G-GCGACA--ACC--A-CG-CGCCG-AATCT-CCGGC.....CTA	6174
COL.CM.-.CGU1	-----CT--G--G--AG--TA-----C-CAA-G-TT-----CAAAGTAAAT.....GACTGA-AATGGCACTAGTACAGTAGCCCAACAACA...	6027
DEN.CD.-.CD1	-----C-----A--TG-A-----C--AA-G--T--T--AAGAAAAGAGGATAACAGCACA--CGCC-TC--CGTGT--TC.....	6656
GRV.ET.-.GRI_677	-----T--GC-CT-----CA-A--A--AAA-GTCC--TGTAGAGCTGAATGGTACA-CCAC--CA--G-CC--C--TGCA-C--CA-CAAT--CT.....ACC	6597
GSN.CM.99.CN166	-----C-----G--T--A--G-----GG-GA-GT--T--T-ACCTCACAACACATCCACA...GCG-ACCA--A-CA--T--T...CCA--G.....	6548
GSN.CM.99.CN71	-----C--C--G--T--GA-G-----AC-AA-G--A--T-AGCTCCCAAAATACGAGCACAACA-CA-CACCGT-C-CA--A-CAGCC-CA-CA.....	6542
DRL.-.-FAO	-----T--A--C--AT--TTAT-----G-AAA-GG-A--T-G-GATCTC.....AAA-TAGTAA--G-AGCC.....AGCACAAACCAACCAACA...	6305
RCM.GA.-.GAB1	-----C--G-----G--T--A--A--CCA-G--T--T-AGCAATAAATGGTGTAGCTGGGATGGA.....ATCCCT.....	6274
RCM.NG.-.NG411	-----GG-----A--A--G--A--A--CAA-GT--T--GCATTGAATGGTAGTTGGGATGGT--C--CC--CC-AT-ACC-----	6342
MND-2.-.-5440	-----GG-----A--TTAT-----G-AGA-G--A--T--AGAAGGTCAAATGAAACAGA-C-AGCAACAG--G.....ACCACAACACCTGTACCC...	6268
MND-2.CM.98.CM16	-----G--C--G--T--TA-----C-AGA-GC--T--AAAAGGAAAAGACCTT-C-ACCCTT-TACCC--CA.....	6677
MND-2.GA.-.M14	-----G--A--TA-----C-AGA-GC--A--T--AAAAGGAGAAAACAGTAG--ACC-GC.....ACTACCCTCCCATGCCC...	6612
MNE.US.-.MNE027	-----T--A--CA--C-A-G--A--A--CAAAAGTGAACAGATA-A-G-GGATTTGACA-AATCAT-A-CA-CA-CAGCA-CACC-ACAACAAAAACAACAACA...	6491
LST.CD.88.447	-----G--AT--TA-----G-GA-G--A--TGAAGGGGGAAGGAATAGGAACAACA--AGCAACA-AA-CA-AA-CG.....TCAACAACACCCATG	5692
LST.CD.88.485	-----GA--AT--TA-----G-AA-G--A--T--AGGGGGAGGAATAGGACCAAAA-CA-CACCAT-ACCA-TAGGA-AA.....ACAACAACACCCATG	5692
LST.CD.88.524	-AC-----A--AT--TTAT-----C-CA-G--A--T--AGGGGAAGGACTAGGTA-GAA--GA-CAACA--A-CA--A-CACCA.....ACAACAACACCCATG	5695
LST.KE.-.lho7	-----A-C--AT--TAT-----G-AA-GC--T--GGGGAAGTTACAAAACACCAACA-CA-CACCA-AA.....ACTACTACACAGATG	6761
SUN.GA.98.L14	-----GA-C-AT--CTAT-----C-AAA-G-----TC-ACCGAACCCGACAACTCCC-CCTCC--CTCA--AGTC-AA-G--CT--GATTATT-GACGACCACCAAGGCTAAGACTACAACCTCAGACT	6850
MND-1.GA.-.MNDGB1	-----GA-T-A--TAT-----AGA-GC--A--TCAGGAAACAGAAAAATGTATCAGCA-CA-CA-C--AGCC--TA.....ACTACACT...	6128
MON.CM.99.L1	-----G-----A--CA-G-----GA-GCTC--TGT-GAAGTTAACACCCGTTTCGAATGCC-G-ACC--CC-G--CC.....	6574
MON.NG.-.NG1	-----G-----TA-G-----A-CA-GC--A--GA-GC--T--TCAAARGATCAATGAAAGCACAAACAGT--CRCCA--CC--CY.....	5187
MUS.CM.01.1085	-----C-----TA-G-----A-AAA-G--A--TG-GAAAGTAAACGAATTAACACCAGATCT-CG-GAGGG--	6500
DEB.CM.99.CM40	-----GC-----A--TAT-----CA-GC-----TAAAAACATCAACAAATCAACACCA-CAA-C--ATCA--A-CA-CA-CA-CAATAGC-ACAACAACAAAAACACTACAGATTTGGTCAGGG	6390
DEB.CM.99.CM5	-----G-----TAT-----GC-GA-G--C--T-G-AGTCAACAACAACAACAACAACAACAACAACA--A-CA-TA--A-CA-CA-CT-ACT-GTCAGGAGAGAAATAGACAGTACACAG.....	6353
SYK.KE.-.KE51	-----T--T--TA-G--A-C--AACGTTA--TGTAGAGGACAAGGACAGTAG-AGTGGT--CT--C-CA--A-TC--CCA-CAACA-C-ACAACAAGCAGCAACACCACCTGCATGGTGG	6340
SYK.KE.-.SYK173	-----G--GC--G-A--CA-G--A--GA-GTT--T--ATTGGACAACAGCCCT-C-ACA-GT.....CCCC--CTACA--TCCCCCTACTACACTCCAAATGAAACGTGGTGG	6644

PLV  
complete genomes

Table with columns for sample ID, sequence alignment (V1 loop, V2 loop), and genome size (gp120). Rows include samples like H1B.FR.83.HXB2, H1A1.UG.85.U455, etc.





		V2 loop	
H1B.FR.83.HXB2	Env	TAAGAGGTAAAGTGCAGAAAAGATATGCATTTTTTTATAAACCCTT...GATATAATACCAATAGATAATGATACTT...ACC...gp120	6791
H1A1.UG.85.U455		-----A---AAA-----T---T---C-----G-----A---CA---AAC-GA-AAC-----G-----AT...	6236
H1B.US.90.WEAU160		-----AA---AGAA---CT-----C-----G-----G-----C-----A-----A-----	6801
H1C.ET.86.ETH2220		-----A---AAAAG---G-C-----C-----G-----G-----C-TA---C-G-T-----T-----	6194
H1D.CD.84.84ZR085		---CG-A---AG-A-C---T---C-----G-----G-----G-----A-----GGG-AAAATGAAATCAAT...GATACTTATGGT...	6333
H1F1.BE.93.VI850		-----A---AA---T---TT---C---GC-----GG-----G-----GC---A-C-A-AGC...-GT...	6113
H1G.SE.93.SE6165		-----A---AAAA-----C---G---C-----G-----G-----G-----TA---CA-----T-----	6215
H1H.CF.90.056		-----A---CA-----T---C---C-----G-----G-----G-----T---CA---G-----T-----	6122
H1J.SE.93.SE7887		-----AA---AAG-A-C---G---C---GC---C---G---AA...G---G-----T---G-A-C-A---AA...	6104
H1K.CM.96.MP535		-----A---A---AA-A-A-G---C---C---C-----G-----G---C-G---C-TA---GGC---GGG-AATAACAGTAGT...-T...	5999
H101.AE.TH.90.CM240		---C-A---AA-----G-TCC---C---C-----G-----G-----A---T---AG---A-G-AGACTAGT...-GT...	6371
H102.AG.NG.-.IBNG		-----A---AAAA---GATGC---C-----G-----G---G---A---TA---G-AA---A-GGT...-GT...	6315
H1N.CM.95.YBF30		---C-A---AAA-A-C---TT---CT-T-C-G---GT-GAA...G---G-----CA---GCC...TATAAT-AA...	6384
H1O.BE.87.ANT70		---C-A---AC---AAA-G---A---C-G---TC-A---C---GT-TCA...-T-G---GGA-C-GA---G-GACA-GCAGCACAAAT...AAGACAAAACAGC-AA...	6827
H1O.CM.91.MVP5180		---C-C---AC---AAA-G---C-A---C-G---TC-A---C---GT-TCA...-C-G-GTAAGG-TA---G-CTCA-A-GCAGTAAAT...GGA---A...	6826
CPZ.CD.-.ANT		-T-A-A---AAA-A---ATGA-A-CA-A-----GGA...-C-T---GAAGTGTG-GG-CA-C-A-GAGACTAAC...-T...	6244
CPZ.CM.-.CAM3		-----A---A-A-A-A---T---T-CC---GT-GA...-C-G---TTG-T---C---CGAT...-CGAT...	6231
CPZ.GA.88.GAB2		-----A---AAAA---C---ATT---T-C-A---C-GG---AA...-C-TG---GGT-CTA-C---AA...	6196
CPZ.GA.-.CPZGAB		-----A---AA-A-AC-G-TC---T-C-A---GTGGAG...-G-GG---AACC---GG---G-ACAAC...-G...	6852
CPZ.US.85.CPZUS		-----AC---AAA-A-AC---TG---T-C-A---GT-GA...-C-GCT-AT---CA---AAT...-AAT...	6835
CPZ.TZ.-.TAN1		-T---A---AAAA-AC---AT---T-C-G---G-GAA...-TG---AA-GAG---GG-A-C...-ATAAT...	6435
H2A.GW.-.ALI		-GGTCAA-TGTCG-TTC---CATGACA-GG---ACAGAGAG-TAAGCCAA-ACAGTATAATGA-ACA...TGGTACTCAAAGATGTGGT...CAG---	7278
H2A.DE.-.BEN		-G-TGCAATGT-A-TTC---TATGA-G-G---AGAAC-AG-TAAGAAAAGG-GGTATAAGGACACA...TGTA-TTAGAAGATGTGGT...GCTGG...	7297
H2A.SN.-.ST		-GGTC-ACTGTCA-TTC---TATGACA-G---AGAGAGGG-TAAGAAA-AC---TATAATGA-ACA...TGTTACTCAAAGATGTAGTCTGTGAATCAAATGACACCAAGAAAAGAG...-AA...	6724
H2B.GH.86.D205		---TGCAATGTAATTTT---GTATGACG-G-C-AAGAAGAG-TGAGCTAA-ACA-TATAA-GACACC...TGTTACTCAGAAGATTTAGAGTGTAAATAATACCGAAGTATACC...-G...	7264
H2B.CI.-.EHO		-G-T---ACTGTCAATTC---ATGACA-G---AAAAAGAG-TGAGTCAA-ACA-TATAAGGACACC...TGTA-AAACAAGATTTAGTGTGTGAAAGGGGACAGGAGTAAT...GAA-G...	7259
H2G.CI.-.ABT96		-GGTT---CTGCAATTT---CATGACA-GGC-AAAAAGAG-T-AGAGAA-GGA-TACAATGAGAC...TGTA-TCACAAAGATTTGCAATGYAACAGTCCAATAAGAGTGAAT...GAG-GT...	6628
H2U.FR.96.12034		-GGTGA-TGTAGATTT---TATGACA-GC-AAAAAGAG-TAAGAAAC-AGAGTATAA-GA-ACA...TGTA-TCTCATGACATAGTTTGTAACAGAATACTAGTGAAGT...GAA-GT...	6757
MAC.US.-.239	Env	-G-T-A-CTGTAATTC---CATGACA-GG-AAAAAGAG-CAAGAAA-AGAGTACAATGA-AC...TGTTACTCTGCAGATTTGGTATGTGAAACAAGGGAAT...AACCCTGGTAATGAA-GT...gp120	7215
SMM.SL.92.SL92B		-G-T---TGTAATTT---CATGACA-G-C-AAAAAGAG-C-AGAAAAGGCGTACAATGA-ACC...TGTTACTCCAGAGATCTAGTCTGTGAGCAAGGAGGA...AATGAA-G-AGT...	6651
SMM.US.-.H9		-GGT-A---TGTAATTT---CATGACA-GG-AAAAAGAG-CAAAAAGAGAGA-TATAATGA-ACA...TGTA-TCAAGAGATTTAGTGTGTGAAACAAAATAGCAATGGAAATGAAACTGAC-GT...	6692
STM.US.-.STM		-GT---TGTAATTT---CATGACA-G-C-AAAAAGAG-TAAAAGAGAGAGTACAATGA-ACC...TGTTACTCTAGGATCTAATCTGTGAAACAAAAT...GTCCTGGCGAAGAG-G...	6854
SAB.SN.-.SAB1C		G---G-A-GTAAA-A-A---TT-C...T---ACAG-A-GGG-TGA-CAA---GG-GG-TTGTGA---G...-GA-GGGAAAAATCAAAC...GCAACA...CATACTGTG...	7077
TAN.UG.-.TAN1		GG---A-GTAAAA---T-T...AATAG-AC---GGT-TGA-CAG---GT-GG---TGTGAGA-GG-AA-C-ACACCACA...GGCAGC-GA...	6866
VER.KE.-.AGM155		-T---G-ACC-AAA-A---TT-C...T---G-AG-A-GG-TGA-GCA-A---CT-TTGT-AGCG...AG---A...TCCGATAAT...GGGACA-AA...	6883
VER.KE.-.9063		-T---A-C---AA-A-A---TT-T...T---G-AG-A-GG-TGA-GCA-G---CTATTGT-AGA-C...AGCT---AGCTCTAACTCT...AGCACA-AA...	6903
VER.DE.-.AGM3		-----A-C---AAAA---GT-T...T---G-GG-G-GG-TGA-GCA-A---C---GTGT-AGA-G...-G...-AACAAT...TCTAAC-GA...	6401
VER.KE.-.TY01		-----A-C---AA-A-A---TT-T...T---T-G-GG-G-GG-TGA-GCA-A---CTATTGC-A-A---A-G---AATAGC...ACTAGC-AA...	6375
COL.CM.-.CGU1		-T-A-A-A-GAAA-G---TT----TA-G-----CAGTAG---GA-TTAA-C-G-GC-TGG---AT...A-A-A-GACAGTAAATAGT...-T...	6198
DEN.CD.-.CD1		A-G-CTT---G-ATA---G-T-ATGTATGAAGC-C---TGTAGAATA---AGATG---ATGGA---GG---A---AACCAAAACA...-T...	6842
GRV.ET.-.GRI_677		A---A-GTAAAA---TT-T...AGCA-GACC-GGT-TGA-CAG---GT---G-CTGC-ATA...A-A-A-AGGAAGTGAAAAG...GGAAGT-AG...	6774
GSN.CM.99.CN166		-T---ACC-AAA-A-AC-GATG---AGCC-C---C-GG-G-GAA...-C---T---GGA-GAGACCC---G...AACCAG-GT...	6725
GSN.CM.99.CN71		-T---G-AC---AAA-AGAC-GATG---AGCC-C---GG---GA...-T---GAGGGC-C-GG---G...AATGAC---T...	6716
DRL.-.-.FAO		AC---TATGTAAAT-C-T-TGACA-G---A-GCAGGG-CTG-AAG---AGAG---CA---GA-A-C...TT---GGTATGATGATGATGTTTGT...AAATGTGGCAATGAAACTGAATCACCAGAAAAGCAA	6532
RCM.GA.-.GAB1		-G-TGTC-TGT-AATTTGCT-T-GCA-G-C-AAAGAGAG-TGAGAAA-C-AGTATAATGATACC...TGTA-TCTAGAGACCTTTGGTGTGAAAAGGAAAACAAATTTCTACAAATTTCTACA-AAAAG...	6508
RCM.NG.-.NG411		-C-TGAA-TGT-AATTT-CG-TGGCA-GG-AAAAAGAG-TGAAAATAATG-A-G---AATGACACA...TGTTACTCAAGAGATCTCTGGTGAAGGCAGGT...AATAATAATACTAGG...	6552
MND-2.-.-.5440		AT---TATGTAAATTC---CACGACA-G---A-G-AGGG-CTGCAAAATTTGGA---GAGGA-A-C...TT---GATATGAAGATGTAACCTTGC...ACAAAGCTTAAATAAATCTGGT-GTGCTAC	6489
MND-2.CM.98.CM16		AC---G-TCTGTAAATTC---TACTACA-G---A-G-AGGG-TTGCAAAATTTGGAG---GAG-A-A---TT---GATATGAAGATGTAACCTTGT...ACAAATGGCACAAATAACACATAT...	6884
MND-2.GA.-.M14		AT---G-TATGTAAATTT---CAC-ACA-G---A-G-AGGG-TAG-AAAATAGAG-TAA-GAGAGC...TT---GATATGAAGATGTAACCTGT...ACTAAAAGGAAAATGACACTGAA...	6804
MNE.US.-.MNE027		-G-T-A-CTGTAATTC---CATGACA-GG-AAAAAGAG-CAAGAAAAGAGAGTACAATGA-AC...TGTTACTCTGCAGATTTGGTGTGTGAAACAAGTGAAT...AGCCTGAAGATGAA-GT...	6686
LST.CD.88.447		AT---AGTGTAGAT-T-T-TTACA-G-C-A-G-AGGG-CTG-AGGACAGA---AAGCAGAG...TT---GGTATGATGAGACAACTGTGATAAAGATGGA...-T...	5908
LST.CD.88.485		AT---AATGTAAAT-T-T-T-ACA-G-C-A-G-AGGG-CTG-AAA---AGA---AAGCAGAG...TT---GGTATGATGAGACAACTGTGATAAAGATGGA...-T...	5905
LST.CD.88.524		AT---G-AATGTTCAT-T-T-TGACA-GGC---G-AGAG-TTGCAAAG---AGAG---AAGCA-A---TT---GGTATGATGAGACAACTGTGATAAAGATGGA...-T...	5911
LST.KE.-.lho7		ATGC-AATGCCA-T-T-T-T-ACA-GG---A-GCAGAG-CTGCAGGACGGA---AA-CA-AGC...TT---GATATGATGATGTAACCTGTCTCAGGAGAGAGGGAGAACAGA...-T...	6977
SUN.GA.98.L14		AT-A-TATGTAAAT-T-T-TGACA-GGC---G-AGAG-TTG-AAAACGGAGG-GGA-CA-A-C...TT---GGGATACAGAAGTACTTGTGCAATGGTAAATGAC...-T...	7039
MND-1.GA.-.MNDGB1		ATCAT-TATGTAG-T-T-CAT-ACA-G-C-A-GCAGGG-TTCGAAG---AGA---GT---C-A---TT---GAGGGGATGATGTGAAATGTGAAAATAAT...-T...	6308
MON.CM.99.L1		A-TTCA-GG-TAAAA---GC---ATGTATAGC---A-T-TGGAAGGAA---C---GAAGGAG---AGGA...-T...	6724
MON.NG.-.NG1		A-TTTA-AG-CAAAAR---C---ATGTATAGC---G-T-KGGAAGGAG-C---T---G-A-GACA-GG---T...	5337
MUS.CM.01.1085		AGTTCA-GG-TAAAA---G-C-GATGTAT-CA---G-TCTTTG---GAT---C-G---GAG-GACAGCTCA---G-G-GAGAAAT...-T...	6686
DEB.CM.99.CM40		...TTT---T-T-C-CA-GTCC---A-AGAG-TAAGAAA---A-AG-GCT---G-C---TA...TGCTAGAGGATGATATCCAGTGGGCGGATAATAAAA...	6513
DEB.CM.99.CM5		...TTT---T-TTACA-GGCCA-A-AGGG-CAAGAG---A---GCA-AGTGC---TG...TGTTAGAGGATGACATACAGTGGGAGACAGAAAAC...	6457
SYK.KE.-.KE51		...TTT---CATGACA-GGGG---A-GG-TAAGAAC---CAGTATAATGC-TTC...TT-TACAAGTGCAGATATAATGAGAGGAAAGGAAAT...	6463
SYK.KE.-.SYK173		...TTT---CTTGACA-GGGG---CA-GG-TAAAAAC-GCA-TATAG-GCTTTC...TTCTA-AAGGATGACTGTGATGAAAGAGGAAAGCAAT...	6767

PLV  
complete genomes

Table listing PLV complete genomes. Columns include: Accession ID (e.g., H1B.FR.83.HXB2), Gene name (e.g., Env), Nucleotide sequence (e.g., ..AGCTATAAGTTGACAAGTTGTAACACCTCAGTCATTACACAGGCCCTGTCCAAAGGTATCCTTTGAGCCAATTC...), and Genome size (e.g., 6896 bp120).



H1B.FR.83.HXB2  
Env  
H1A1.UG.85.U455  
H1B.US.90.WEAU160  
H1C.ET.86.ETH2220  
H1D.CD.84.84ZR085  
H1F1.BE.93.VI850  
H1G.SE.93.SB6165  
H1H.CF.90.056  
H1J.SE.93.SE7887  
H1K.CM.96.MP535  
H101.AE.TH.90.CM240  
H102.AG.NG.-.IBNG  
H1N.CM.95.YBF30  
H1O.BE.87.ANT70  
H1O.CM.91.MVP5180  
CPZ.CD.-.ANT  
CPZ.CM.-.CAM3  
CPZ.GA.88.GAB2  
CPZ.GA.-.CPZGAB  
CPZ.US.85.CPZUS  
CPZ.TZ.-.TAN1  
H2A.GW.-.ALI  
H2A.DE.-.BEN  
H2A.SN.-.ST  
H2B.GH.86.H2D05  
H2B.CI.-.EHO  
H2G.CI.-.ABT96  
H2U.FR.96.H2034  
MAC.US.-.239  
Env  
SMM.SL.92.SL92B  
SMM.US.-.H9  
STM.US.-.STM  
SAB.SN.-.SAB1C  
TAN.UG.-.TAN1  
VER.KE.-.AGM155  
VER.KE.-.9063  
VER.DE.-.AGM3  
VER.KE.-.TY01  
COL.CM.-.CGU1  
DEN.CD.-.CD1  
GRV.ET.-.GRV\_677  
GSN.CM.99.CM166  
GSN.CM.99.CN71  
DRL.-.-.FAO  
RCM.GA.-.GAB1  
RCM.NG.-.NG411  
MND-2.-.-.5440  
MND-2.CM.98.CM16  
MND-2.GA.-.M14  
MNE.US.-.MNE027  
LST.CD.88.447  
LST.CD.88.485  
LST.CD.88.524  
LST.KE.-.lho7  
SUN.GA.98.L14  
MND-1.GA.-.MNDGB1  
MON.CM.99.L1  
MON.NG.-.NG1  
MUS.CM.01.1085  
DEB.CM.99.CM40  
DEB.CM.99.CM5  
SYK.KE.-.KE51  
SYK.KE.-.SYK173

ATTCTAAAATGTAATAATAAGACGTTCAATGGAACAGGACCA.....TGTTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCACTGCTGTTAAATGGCAGTCTAG  
I L K C N N K T F N G T G P . . . C T N V S T V Q C T H G I R P V V S T Q L L L N G S L  
-----G---GG---CCTGA-----A---G-----C---GG-----C---A-----G-----  
-----G---G-----A-----AG-----C-----C-----A-----G-----C---T-----  
-----G---GAG-----A---C-----CCAT-----C-----AA-----G-----AT-----T---A-----  
-----G---CG---A---G-----C-----C---A-----T-----G---G-----  
-----G---A---GA-----G-----G-----C---AG-----AA-----G-----T---T-----C-----  
--T---G---GGG---GA-----A-----T-----A-----A-----G-----A---AC---G-----T---  
-----G---C---A---A-----TT-----T-----A-----G-----A---A-----A---C-----  
-----G---AT-----T-----C---A-----T-----C-----A-----G-----A---C-----A-----  
-----G---GG---A---G-----G-----A-----T-----C-----AA-----G-----C-----C-----  
--T---G---G-----AT-----G-----G-----C---A-----T-----C-----A-----G-----T-----  
-----G---GG---AGG-----C---A-----C-----C-----C-----A-----G-----T-----  
--A---G-----G---AGGA---AC---T---G-----AT---AGC-----G---T---T-----C-----A---A-----G---A-----C---GT---AA---CC-----A---CT---A  
--CT---T---G-----C---GC---CAGAA---T-----CA-----C---G---CA---A---CGGT---TACT-----C-----AC---AGT---G---AA---A-----G---CA---CT  
--CT---T---G-----G---C---CAGAC---T-----C---T-----CCAC---A---TTCAGTG---TACT-----C---C---A---AC---AGT---AA---AC---G-----G---CA---CT  
--AT---G---G---G---AG---G---GAT---T---CA---GT---GATG-----A-----TCAGT-----T---C---T-----A---C---A---G---GG---ATGGT---A---A-----A---C---TACC  
--CA---G---C-----TGCCA---AC---T-----A---GGAAGA-----AC---A---A-----C-----T-----C---C---A-----G---A---T---A---T---A---CC---T-----A---G---  
C---GT---G-----G-----AGAT---ATCCA---A-----AAG-----A-----A---T-----G---T---C-----A-----ACT---A---T-----GT---A---AC---C-----CT-----  
-----T-----C---G-----AGAC---TTCA---A---TAA-----T-----A---T-----G---T-----AA-----G---TA---C-----GT---A---TA---T-----A---T-----  
--C-----G---A---AGAT-----A---AG---GAG-----A-----T-----T---G---G-----T-----C---A---AA-----G---GA---T---A---G---CA---AA-----T-----  
C---GT-----GAG---C---AT-----CA---G---A-----A-----CT---C-----TCAGT---T---C-----G---TAT---TA---GA---G---C---AGC---T---A---AC-----GTCC---G---  
C---A---G---CG---G---TC---AT---ATTCA---CTTT---C---C---.AAT---CT---T---A---AGTAG---T---CTAC---C---AGGATG---GGA---A---GCA---ACT---T---GGT---TGGC---T-----C---AGG---  
C---AT---G-----G---CC---AC---ATTCA---CTTT---A---T---.AAG---C---T---A---AGTAG---T---CTTC---C---AGGATG---GGAAA---GCA---ACT---T---GGT---TGGC---T-----C---AG---  
C---G---G---C---G---CC---AT---ATTCA---CTTT---AG---C---.AAT---T---T---G---AGTAG---T---CTAC---C---AGGATG---GGAAA---GCA---ACC---C---GGT---TGGC---T-----C---AGG---  
C---A---G-----G---CC---AC---ATTCA---CTTCATG---C---.AAC---C---GT---G---AGTAG---GTCTCC---C---AGAATG---GGAAA---CAGTCC---T---ATGGT---TGGC---C-----T---CAAGG---  
T---G---G---C---G---CT---AA---ATTCA---CTTCATG---C---.AAC---GT---G---AGTAGT---TCT---TG...TACAGAATG---GGAAA---CAGACC---T---ATGGT---TGGC---C-----T---CAAGG---  
T---G---G-----G---CT---AC---ATTCA---TTT---C---T---.AAT---CT---T---A---AGTGGTCTCTAC-----AGAATG---GGAAA---CA---ACC---T---ATGGT---CGG---T-----A---CAAGG---  
C---GT---G-----G---CT---AT---A---TC---TTTATG---T---.AAT---CT---T---A---AGTAGTCTCTAC---C---AGAATG---GG---A---CA---ACC---T---GGT---TGG---T-----C---AG---T

L L R C N D T N Y S G F M P . . . K C S K V V V S S C T R M M E T Q G S T W F G F N G T R  
T-G-T-G-----G---CA---AT---ATTCA---CTTT---CT---T---.AAT---GT---G---AGTAGT---TC---TC-----AGGATG---GGAAA---CAGACT---T---ATGGT---TGGC---T-----T---C---AG---  
T-G-T-G-----G---CC---AT---ATTCA---CTTT---CT---T---.AAC---T---G---AGTAGTCTCTTC---C---AGAATG---GGAAA---CA---ACC---T---GGT---TGGT---C-----T---C---AG---  
T-G-C-G---C---G---CA---AT---CA---CTTT---C---T---.AAT---CT---T---G---AGTAGTCTCTTCG---TAGAATG---GGA---A---CA---ACT---TGGT---TGGC---T-----A---C---AG---  
T-G---G-----GCAG---CAGAT---ATTCA---CATAAG...GCA---GG---A---CAGT---TC---GCT---C---TAGGTT---C---ATA---TACT---AGC---AGG---A---AGGAA---T-----T---TAT---  
--CT-----GG---CT---AT---CA---TTT---TGTT...---G---TTCAGT---TAGT---TGATTG---G---ATA---TAC---GAGCT---AGC---T---TGGCA-----AG---  
T-G-T-----G---TG---GAT---ATGCA---TTTAAAG---.GAA---T---T---TTCAGT---G---T---C---AC---TT---G---ATA---AC---A---C---GGT---T---A---G---A---CTATT  
C---GT-----CC---G---C---TGAT---AT---TTATAAG---A...AAT---GT---TTCAGT---C---TGGACTT---G---ATA---AC---T---T---AGG---A---G---A---CTATT  
T---AT-----C---CG---T---TGAT---ATGCA---GTTTAAAG...AAC---T---T---TTCAGTG---G---T---TA---CTTG---A---ATA---AC---GA---T---GG---T---G---G---CTACT  
T-G-----G---G---AGAT---AT---GTATAA---A...AAT---CT---ATCAGT---G---T---GGCTT---G---ATA---AC---GA---AGGGT---T---C---G---A---CTATC  
C---T---G---G---A---A---GGG---G---G---.---C---AC---G---CAG---GTCC---C---AGT---ATAT---AC---CATGT---TG---AT---AG---C---C---TCAAAA---  
C---C---T---G-----G---C---TC---T---GG---G---GTAT---C---TT...---A---A---G---G---AG---G---AG---G---AC---CA---G---CCAG---CTGGT---A---CA---G---CA---T---CAG---  
T-G-----G---GAT---AT---GA---CTTT---CT...AAG---C---AG---TTCAGT---G---T---TAGATT---C---ATA---TACTA---A---T---AGGGA---AGGA-----T---AG---T  
-G-----C---G---GTT---AT---CA---GT---CA...---CTAC---GTCAG---ACC---C---C---CAA---CT---G---ATGG---C---AT---G---CCTATC  
-G-----C---G---GCC---AT---CA---GTG---CAAG...---T---C---ATCAG---AC---C---G---CAA---TT---G---ATGG---T---ATC---G---C---TATC  
T-G---GG---G---A---ACTAAAT---TAA---T---GTAAG...---A---A---A---CAG---G---G---TAGACAT---GCCTG---CAC---A---C---T---GC---TT---TGGC---C---CAAAGC  
T-A---GGTC---C---AG---A---AT---AT---CA---CTTT---ATA---C...---GTT---T---CTG---CACTTC---C---ATG---ATA---TACT---GG---CT---AGGGT---TGG---ATCAA---TA  
-A---G---CTG---GCA---AT---AT---CA---CTTT---ATA---C...---CA---T---GAC---ACGTC---C---CATG---A---ATA---AC---ATC---T---TGG---ATC---A---TA  
T-A-----G---A---ATT---AA---GAC---A---TT...---GGG---TCTG---G---G---G---GCACCG---GCCAG---CAC---A---C---T---AATGT---TGGC---C---A---AAGC  
C---CT---G---GG---CCTAG---A---GATT---AATGTCTCT---A---AA...---C---A---CAG---G---G---GCC---T---GCCAG---CACCACC---ATGTGTGGA---CT---G---G---CAAA---C  
C---G---G---G---CG---G---G---ATT---AAT---A---AA---TT...---G---C---T---TCAG---G---T---G---CCC---GCCTG---TACTA---ATGT---TGGC---T---A---CAAA---C  
T-G---T---G---G---C---CA---A---ATTCA---CTTTATG---T...AAC---T---T---G---AGTGGTCTCTTC---C---AGAATG---GGA---A---CAGACT---T---TGGT---TGGC---T---A---C---AG---  
C---T---G---C---CG---A---ATT---AATTTCTC---G---TAAT...---C---A---A---CAG---ACC---CT---C---G---CTAT---G---TTAGTTT---GAGT---G---TTCT---TGGC---T---A---CCAATC  
C---T---G---G---A---ATT---AATTTCTC---G---TAAT...---C---A---A---CAG---ACT---CT---C---G---CTAT---G---TTAGTAC---GAGT---G---TTCT---TGGC---T---A---CCAATC  
C---T---G---G---A---CTAAA---TT---TC---G---TAAT...---T---A---A---CAG---TAC---C---C---G---TAC---G---TAAGTTT---TAGT---GCTTCT---TGGC---T---A---CCAATC  
C---AT---G---G---C---G---AC---ATTA---AA---TT---A---T---A---AA...---GA---A---A---CAG---TAC---C---TGATAT---GCTTAGTAGT---AGT---G---TTCT---TGGC---T---A---CAATC  
T---AT---G---G---G---A---ATTAAATGC---ACT---AGAA...---A---A---A---CTG---TACC---CT---TA---TAT---G---CAAGTAC---TAGT---GCTTTT---TGGC---T---A---C---AG---C  
--GT---G---A---A---ATCAG---TTAAAT---AAT---A---TTG...---GC---A---A---ATCAG---G---G---T---GCAC---AGTAG---CAC---G---AGT---GCTTTT---TGGC---T---A---C---A---GC  
T---AT---G---G---G---C---CCCG---C---T---CG---CA---GAGT...---T---T---TTCG---AC---T---CT---CCAA---TA---G---T---TGGT---C---A---G---CA---T---CGGGTA  
--AT---G---G---G---G---CCA---C---CA---GCAGAACAA...---CT---R---TCAG---AC---C---G---T---A---T---A---T---TGGT---T---A---C---T---CT---CCTCT  
T-G-----C---G---GCC---AT---T---C---GCA...---AT---G---CTG---TAGG---T---C---G---C---ATTA---CT---G---TG---A---A---C---T---A---C---TATT  
C---A---G---G---G---A---AT---T---T---GTT...---C---AC---C---G---CTG---GGC---A---CTC---ACACA---CA---G---C---C---ATGGG---A---AA---T---GTC---GAT  
C---GT---G---G---G---C---GA---AC---T---T---C---CTT...---C---G---CTG---GGC---A---CTC---ACATA---CA---G---C---ATGGA---AA---C---TTCAGAT  
C---GT---G---G---G---CAGAA---TG---A---GAT---CTGAG...---CA---A---CAG---TTC---C---CCTT---C---ACA---TT---G---CTAGT---ATGGT---T---A---CCTATA  
T---GT---G---G---G---CA---AT---TG---A---GAC---ATGTC...---A---A---CTG---TC---C---G---A---T---G---ATA---T---G---CTAGC---ATGGT---C---AAC---G---CCTACA

7020  
gp120

6465  
7030  
6423  
6562  
6342  
6444  
6351  
6333  
6228  
6600  
6544  
6613  
7056  
7055  
6473  
6466  
6425  
7081  
7064  
6664

7510  
7529  
6956  
7496  
7488  
6860  
6989

7447  
gp120

6883  
6924

7086  
7306

7095

7115  
7135  
6633

6607  
6427

7071

7006  
6954  
6945

6780  
6737  
6781

6737  
7110  
7036

6918  
6134  
6131  
6137  
7203

7265  
6531  
6965  
5581  
6918  
6760  
6711  
6698  
7002

PLV Complete Genomes



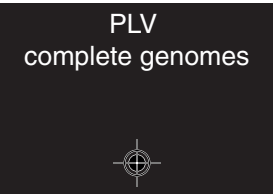
PLV complete genomes



Table with columns for sample ID (e.g., H1B.FR.83.HXB2), sequence alignment (A, C, G, T), and genome size (e.g., 7126, 9p120). The table lists 100 different PLV complete genomes.



	loop tip	V3 loop end -	
H1B.FR.83.HXB2	CAATACAAGAAAAAGAAATCCGTATCCAGAGA..GGACCAGGGAGAGCATTGTGTTACAATA.....GGAAAAATA..GGAATATGAGACAAGCACATTGTAAACATTAGTAGAGCAAAAATGG		7238
Env	N T R K R I R I Q R . . . . . G P G R A F V T I . . . . . G K I . . . . . G N M R Q A H C N I S R A K W		gp120
H1A1.UG.85.U455	T...-AT-A---GGT-T-TATA---T---ACA---CTA-GT-C-.....-T---ATA-GG---A-----TG-C-----AGGG-C---		6686
H1B.US.90.WEAU160	-----A---AAC-T-A.....-T-C-TA-C-.....-G---ATA---G---A---G-----C-----A---GT---		7245
H1C.ET.86.ETH2220	T-----G---T-AA-G-A.....-ACA-A---CTA-G-C-.....-G-C-ATA---G---A-----GA---A---		6638
H1D.CD.84.84ZR085	A--GGA---C---G-CA-C---A.....-A---CA---GC-CTA-C-.....A-GT-T-C-ACAA-G-TA-AG-----T-----T---G---T---		6777
H1F1.BE.93.VI850	-----G---G-T-A-A-T-A.....-ACA-A---CTA-G-C-.....-GC---ATA---G-C-A---A-G-----G-A---C---		6557
H1G.SE.93.SE6165	-----TG---G---AA-A-G...G-ATT---ACA-A-G---CTA-G-C-.....-TGCC---ATA---G-C-A-----TG---C---A-AG---		6665
H1H.CF.90.056	T-----G---C---T-A-A-T-A.....-G---AC---CTA-G-C-.....-TG-C-CATA---G---A-----T-----A---G-C---		6566
H1J.SE.93.SE7887	-----G---G-T-A-AC-G.....-ACA-TGC-CTACG-C-.....-G---ATA---A---G-A-----GA-AG-G-T---		6548
H1K.CM.96.MP535	-----T-A-A-G.....-A---A---CTA-G-C-.....-TG-T-ATA---G---A---G-----C-----G---A---		6443
H101.AE.TH.90.CM240	-----C---T-AAC---A.....-AC---T---CTA-G-C-.....-G-T-ATA---A---A---T---G-G---A-G-A---		6815
H102.AG.NG.-.IBNG	-----G-TG-A-A-A.....-CA---CTA-G-C-.....-TG-C-ATA---GG---A-----TG-C---A-A-G---		6759
H1N.CM.95.YBF30	T-----G---GG-CAGG-G-AG-A.....-T-CT-TGA---TA-AC-.....-A---ATA---G-C-T-----T-C---TG-CTC-A---A-CT---		6816
H1O.BE.87.ANP70	A...GAC-T-C-GAG-GA-A-A.....-T---AT-GCCTGG-ACAGC-TGGG.....ATAGGG---C-GC-GGAAAC-GCTCA-GGC---TT---C---GTA---A-GCCA-TG-T---		7274
H1O.CM.91.MVT5180	TGCAGAGT-C-GAT-ATA-CA.....-T---AT---TGGCGAG-TG-C-CTTAAAGAAGTAAC-T-C-TCACC-GATCA-GGT---TT---CATA---A---AGA-TGT---		7282
CPZ.CD.-.ANT	T-GG---GT-G---ATC-A-AA-A.....-A-TGA-T---CTA-ACG-.....-A---T-GC-ACA---G-C-CT-GA---GTTC---CAG-C-A---AGA-GCT---		6688
CPZ.CM.-.CAM3	T-NC-GC---GG-CA---A-AG-A.....-C-TGA-C---TA-AC-.....-A---T-TGTA---G---CC---G---T---C-G---C-A---GA-GTC---		6666
CPZ.GA.88.GAB2	-----G---G-GC---AAACC-T.....A-C-----C-A---CT-C-AC-CG.....-A-GCCC-ATA---CCCT-GA---GTCC-CC-C---A-G-TA-CT---		6649
CPZ.GA.-.CPZGAB	-----GGG-GAGG-G-AG-A.....-G---A-TGA-C---TA-AT-.....-A---TG-GTA---G---CC---TCT-CT-C---G-A-A-G-GA-C-C---		7296
CPZ.US.85.CPZUS	-----GTGGG-ATG-A-AAC-A.....-A-TGA-T---TA-AC-.....-CC---G---GTA---G-G-A---G-----CTCA-A-CTG-C---		7270
CPZ.TZ.-.TAN1	-----GGTCA-G-A-AG-A.....-T---T-TGA-C---TA-AT-.....-A---TG-GTA---GG-C-CC-GA---TT---TCAG-C-A-GC-A---C---		6876
H2A.GW.-.ALI	T-G---GTTGT-CC---AACAC-TAT-TC.....GTT-ATATTT...CACTCC.....CAGCC---CAATAA-GACCC-----TGG-CTGGT-C...-A-GCG-G---		7716
H2A.DE.-.BEN	T-G---GTTTT-CC---AACAC-TAT-TC.....TT-T-TTT...CACTC.....CAGCC---CAACAC-GGCCT-G-G---TGG-CCGGT-...G-GCG-G---		7735
H2A.SN.-.ST	-----G---GTTGT-CC---AACAC-AT-TC.....GTT-T-TTT...CACTCC.....CAGCC---CAATA-GACCC-G---TGG-CTGGT-C...-A-GCG-G---		7162
H2B.GH.86.D205	-----G---GGTTGT-CC---AA-A-C-GT-TC.....T-CITTTTC...CA-TCA.....CAGCCT-CAATAAG-GACCC-----TTGG-CTGGT-...-AG-G-C---		7702
H2B.CI.-.EHO	-----A-TGTTGTGCC---AA-A-C-GT-TC.....TATTCCTTC...CA-TCA.....CAGCCT-CAATAA-GACCT-A---TTGG-CTGGT-...-A-G-C---		7694
H2G.CI.-.ABT96	T-A---GT-TT-CC---TAC---TAT-TC.....G-T-TCTTC...CACTC.....CA-CC---CAAT-A-GACCC---G-TTGG-TGGT-...-G-GC---		7066
H2U.FR.96.12034	T-G---GTCTT-CC-G-AACC---TAT-TC.....T-T-TTC...CA-TC.....CA-GCT-CAATAAC---CCA-A---TTGG-TGGT-...G-T-G---T---		7195
MAC.US.-.239	T-G---GTTTT-CC-G-ACC---TAT-TCT.....TTG-TTTTC...CACTCA.....CA-CC---CAAT-AT-GGCCA-AG-G---TGG-TGGT-...G-G---		7653
Env	K T V L P V T I M S . . . . . G L V F . . . . . H S . . . . . Q P I N D R P K Q A W C W F . . . . . G G K W		gp120
SMM.SL.92.SL92B	T-G---GTCCT-CC-G-ACC---TAT-TC.....TTTG-TCTTT...CA-TCG.....CA-CC---AAT-AG-GGCCA-----TGG-CTGGT-...G-G---		7089
SMM.US.-.H9	T-G---GTCCT-CC-G-ACC---TAT-TC.....GTTG-TCTTC...CA-TCG.....CA-CCC---AAT-AG-GACCA-A---G-CTGG-CTGGT-...G-G-G---		7130
STM.US.-.STM	-----A---GTCTT-CC-G-ACC---TAT-TC.....CTTG-TCTTC...CA-TCG.....CA-CCC---TAAT-AG-GACCA-A---TGG-CTGGT-...G-G-G---		7292
SAB.SN.-.SAB1C	T-G---GT-TTGCTG-AACC---AT-GCT.....CTTG-TCTTC...CA-AGC.....CAG---TATAACACC-GGT---A-----TTGG-CTGGT-...G-GT-T---		7518
TAN.UG.-.TAN1	T-G---GTCCT-CC-G-AACC---AT-GC.....T-TC-T-TTC...CACTCA.....CA-C---TATAATAC---AT-A-----TGG-CC-TT-C...CAG-GTG-T---		7307
VER.KE.-.AGM155	T-G---GTCCTGCC-G-AACG---AAT-GC.....TT-TCTTC...CACTCA.....CAG---GTATAATAC---GAC-A---G-G---CTGG-CC-T-C...CAG-GC-T---		7327
VER.KE.-.9063	T-G---GTCCT-CC-G-AACC---AAT-GC.....TTT-TTTTC...CA-TCG.....CAG---TACAAATACG-GC-T---G-CTGG-C-TT-C...CA-GC-GG---		7347
VER.DE.-.AGM3	-----A---GTCTT-CC-G-AACA---AT-GC.....G-T-T-TTT...CA-TC.....CAG-GGTACAAATC-GGC-----TTGG---C-T-C...CAG-GC-C---		6845
VER.KE.-.TY01	-----G---GTCCT-CCGG-AACG---AAT-GCG.....-TG-TTTTC...CACTC.....CA---TACAAATACG-GC-T---G-TTGG---C-T-C...CAG-GC-T---		6819
COL.CM.-.CGU1	T-G-T-TCAC-G---ATC-AAAC-CTGCC-AT.....G-AAATTCTAC-A-AG.....CTT-TTCTTATTC---AGGC-TTT-T-G-AGG---C-GT-...-T-CC-AT-		6642
DEN.CD.-.CD1	TGTA-CCT---GG---T-AAAC---GGCAGC.....-AT---TC-TG.....AGAG-C-CT.....ATT-T-CC-AGAA-GGC---T-GAG---GCA-C-G-AGTAG---GC-C---		7295
GRV.ET.-.GRI_677	T-A---GTGTTGCC-G-AACA---AT-GC.....GTT-TATTT...CACTC.....CAG---TATAATACC-GGT-A-A-----GTGG-CC-T-C...CA-G-G-T---		7218
GSN.CM.99.CN166	T-A-C-T-G---ATT-G-AA-A.....-G-G-A-TGA---TA-T-CCA.....-TG-T-G-GGA---G---CT-GA---CT---CCGAC-GA-CCT-AT-GG---		7172
GSN.CM.99.CN71	T-A-CCT---ATT-A-AG-A.....-G---A-TGA---CTACT-CCA.....ATT-T-G-GGG---G---CT-GA-G-CT---C-AC-C-ACCATA---G---		7163
DRL.-.FAO	T-GG-GC-T---T-TCT-CA-CCTCAGCCTC.....GTTGTT-TTTATCA-GG.....TT-G---CC-GGAAA---T-A-A-AG-G-ATG---AT-...GA-G-C----		7001
RCM.GA.-.GAB1	T-GA---GTG---G-G---ATCGC-AGCA-C.....GTCTTT-TC.....TCA.....CT-G-G-GAGAAG-GACCA-A-GG---TTGG---GAT-...GA-GG-T---		6940
RCM.NG.-.NG411	-----GTG---G---TCAT-AGCA-C.....GT-TTC-TC.....TC.....CT---GGG-GCAAAA-GGCCA-A-GG---TTGG---GAT-...GAT-G-G-T---		6984
MND-2.-.5440	T-GAT-T-TC-TTTCC-CA-CA-GTCT-C.....TTGCT-TTCTATCA-GG.....TTGG-GCC-GGAAAG---C-T---A-G-GCA-G---C-GT-A...GA-GG-----		6967
MND-2.CM.98.CM16	-----GAT-TGTG-TCTC---CA-C-GTGCA-CG.....TTGATCTTTATAG-GG.....TT-G-CC-GGAA---C-T-AGA---GTATG---C-GT-G...-AG-G-C---		7340
MND-2.GA.-.M14	-----GGT-T-T-GTTTC-G---CA-GTGATC.....TT-ATTTCTATCA-GG.....TT-G-CC-GGCA-G---C-A-A---G-ATG-C-GGT-A...T-G-C---		7257
MNE.US.-.MNE027	T-G---GTTTT-CC-G-ACC---TAT-TCT.....TTG-TTTTC...CACTCA.....CA-CC---CAAT-A-GGCCA-A---G---TGG---GGT-...GA-G---T---		7124
LST.CD.88.447	T-GGT---GA-GT-TCT-C-AT-GTTCA-C.....TTCT-TTTTAC-A-GC.....TTGG-CATGGTA-C-GGC---TTG-T-G---AT-...GA-G-C-G---		6361
LST.CD.88.485	-----GGT-GA-GTCTC-C-ATC-GTTCA-C.....TTCT-TTTTAC-A-GC.....TTGG-CATGGTA-C-GGC---TTG-T-G---AT-...GA-G-C---		6358
LST.CD.88.524	T-GGT---GA-GTCTCC-CAATC-GTTCA-C.....TTCT-TTTTAC-A-GG.....TT-G-CATGGTCT-GAC-A---TTG---A-C-AT-...GAG-GTC---		6364
LST.KE.-.lho7	T-GAT---GA-GT-TCC-CAAT-G-TCT-C.....TATTTATTTAT-A-GG.....CT-G-CATGGCA-C-GAT-A---TT-----G---AT-...GA-GCC---		7430
SUN.GA.98.L14	T-GGT---CA-GT---T-CTATA-GTAGT-C.....-TGCT-TTTTAT-A-GG.....CT-G-GCC-GGTTCT-GT-A---TG---G---AT-...GT-G-C---		7489
MND-1.GA.-.MNDGB1	T-GGT---GT-GTG---T-CT-CATCAGCT-C.....TT-TTATTTATCA-GG.....TT-G---CCTGGAAAG---T-A-AGA---GCATG-C-C-T-C...-A-G-CGT---		6764
MON.CM.99.L1	-----A-T-TT-GG-ATT-A-AG-A.....-G---C-TGA---CTA-T-GCA.....CTT-T-G-GGGA---T---CTC-TA---CT-C-C-AG---A-AGA-CC---		7180
MON.NG.-.NG1	T-G---CTT---G---ATT-A-AG-T.....-TG---A-TGA---CTACT-GCA.....ATC-T-G-GGGA---C-CY---A-G---T-C---C-GGGA-A-A-AGA-C-G---		5796
MUS.CM.01.1085	-----G---T-GG-ACC-T-AA-A.....-G---C-TGA-T---CTA-T-GCAG.....CTT-T-G-GAT---CG---CAC-CAG---CT---C-A-A-A-AGA-C-G---		7136
DEB.CM.99.CM40	T-G---CTACCG-GCC---A-AC---GGCT-C.....T-TGTCCTTTTATAC-ACC.....TTT-T-CA-AGGCTG-GG---A-AGAG-----GAC-A...-AC-G---GC---		6990
DEB.CM.99.CM5	-----G-GTAT-G-GCTG-T-A---GGCC-C.....G-T-TCTTTTATACAACC.....TTC-T-C-AGGCTG-GG---A-AGAG-----GG-G...-GT-T---		6929
SYK.KE.-.KE51	TG-GT-T-T---AC---AAT-GGCAGC.....TTTTTCTTGC-G-GA.....CAAGGGAGATG---C-GGGAA-GC-GC-A-AG---TT---C-AG-C...-A-G-G---		6922
SYK.KE.-.SYK173	TG-GT-T-T---AT---T-AAT-GGCTGC.....CTATTTCTTCC-G-AA-A.....CAAGGGAAATG---CCGGGA-GG-CGCA-A-GT---CTTC---CGAG-A...-C-G---T---		7226





PLV  
complete genomes

H1B.FR.83.HXB2	AATAACACTTTAAAACAGATAGCTAGCAAATTAAGAGAAACAATTGGAAATAATAAAAACAATAATC.....	7304
Env	<u>N N T L K Q I A S K L R E O F G N N K T I I</u> .....	gp120
H1A1.UG.85.U455	----GA--A--C-----G-----GAAC-----AGA--A-G--AAT--C--A-C-----	6743
H1B.US.90.WEAU160	-----T-GAA-----AT-AAACA-TT--G--T-A--C--A-----	7311
H1C.ET.86.ETH2220	-----A--C--C--A--G--AAGGAA-----CA-A-G--C--CCCT-----	6695
H1D.CD.84.84ZR085	-----C--CG--G--G-----A-----G--A-C-TT--GAACC-A-CA-----	6834
H1F1.BE.93.VI850	-----C-GG-GT-TG--AAGGCAG-----AGTCG--T--CCCT-----	6617
H1G.SE.93.SB6165	--AG-GG-C--C-GA-TG-C--GCAG--C--G-GA--ATC--AAT--G-GC-GCGAG-AC-----	6728
H1H.CF.90.056	-----G-----C--C--G--T--CAC-----G--AT--C--GAAC-----GA-C-----	6623
H1J.SE.93.SE7887	-G-----CGTAGAG-----CA--C-----C--AAT--A-CA-----	6602
H1K.CM.96.MP535	--C-TG-----GCAGAG--AAGGAA--GC--A-----T--AAG--GGA-C--T-----	6503
H101.AE.TH.90.CM240	-----AGT-----G--A--GAA-----A--G--C--AAT--G-CA-----	6869
H102.AG.NG.-.IBNG	-----A-----C-T--G--T--CAC-----AGACGT--C--AAG--C-CC-C-----	6816
H1N.CM.95.YBF30	G-ACCA-TG-GG--TAGA-C-AGAGAGG--A--AGA--ATCC-G--G--A--C--C-----	6873
H1O.BE.87.ANT70	GGA--A-TA-----A-C--GAA-GG-ATTT-----T-G-AAAC--CAGGT-----	7331
H1O.CM.91.MVP5180	G-A--TG-CC-C--A-C--TA-GG-ATTT-A-T-TTG-AAACC-A-CAG-G-AT-----	7342
CPZ.CD.-.ANT	G-AC-AG-ACGT--CA-A-C--AGCA-GTTC-TGCG--G--T-GGAA--AGTAG-C-AC-A--C-AAT-----	6757
CPZ.CM.-.CAM3	G-CG-A-GAC-----TG-A-CT-G-CAAGCC--GC-----G-TT--AC--CCT--C-CA-G-C-A-TTTACAGTAAGT-----	6744
CPZ.GA.88.GAB2	-----T-C-----CAGA--AAACAA--GA-T-A-A-CAGCACACGTGGC-C-GGGGGGAC--ACA-----	6718
CPZ.GA.-.CPZGAB	---CGA--GG-GG--G-AG--AAA-AGGCT--GC-ACCTC--CAAATCGA-CAGC-G-C-AC--AACACTCAACAGAGCC-----	7377
CPZ.US.85.CPZUS	G-G--ACAGAG--T-C-CTTTAGAG-T-A--A-A-GG--GCAAACCTG-CA--GGT-GAGT-AATTCCAAAT-----	7345
H2A.GW.-.ALI	--GG--AG-CA-GC-GG--G-GAAGGAA--CCC-TGT-A--TCCCA-GT--AGG--C-ATGA-ACA-----	7785
H2A.DE.-.BEN	GGG-AG-CA-GC-GG--G-GAAGCAA--CCC-TGT-C--TCCCA--T-C--AGG--TC-ATGATACA-----	7804
H2A.SN.-.ST	--GG-AG-CA-G--GG--G-GAAGCTA--CCC-TGC-A--TCCCA-GT--AGG--C-ACGA-ACA-----	7231
H2B.GH.86.D205	-CAG-AG-CA-----GG--G-GAAA--G-CCA-C-T-A--TCCCA-GT--AGG-GGTGC--AAAAT--ATCACA-----	7777
H2B.CI.-.EHO	-CAG-AG-CA--C-GG--G-GAAAGAG--CCA-T-A-A-T--TCCCA--T--TCAGG----C--ATATC-----	7763
H2G.CI.-.ABT96	--GG-AG-CA-GC-GG--G--AAGGAA--CCG-GGTCA-G--CCCCA-CT--AGG--ATGA-ACA-----	7135
H2U.FR.96.12034	--AGGGG-AA-CC-GG--G-GAAGGAG--CCA-TGCCA-C--TCCCA-GT--C--GGT--C-AG-ATATT-----	7264
MAC.US.-.239	--GG-TG-AA--G--G-GAAGCAG--CCA-TGTCA--TCCCA-GT--C-GG--T-AC-ATACT-----	7722
Env	<u>K D A I K E V K Q T I V K H P R Y T T G T N N T</u> .....	gp120
SMM.SL.92.SL92B	-GAG-AG--A-GC-GG--G-TAAG-AA-CCA--GTCA--CCCCA-GT--C-GG--C-ATGATACT-----	7158
SMM.US.-.H9	--AG-GG-CA-CCGGG-AG-GAAGMAMCC--GGTCA--TCCCA-GT--GGG--T-ATRAKAYT-----	7199
STM.US.-.STM	-GAGGAG-AA-----G--G-TAAGGAA-CC--GGTCA--TCCCA-GT--CAGG--C-ATGA-ACG-----	7361
SAB.SN.-.SAB1C	-GAGGAG-C-GG--GG-AG-CAAAGAA-C-A--GTGAG--T-CCACC--A--GT-T-GTGGC-CAAAT-----	7587
TAN.UG.-.TAN1	--AGGAG-C-GG--GGG-AG-GAGA-AA-C-A--GTG--G-TTCCAAA-G-G--AT-TCG-GGG-CAAAT-----	7376
VER.KE.-.AGM155	--AGGAG--GG--GG-AG--CAAGAGG--A--GT-A--TT-CCAAA-G-ACGGT-CCA-GGC-C-AAT-----	7396
VER.KE.-.9063	-GAG-AG-C-GG--GG--G-TAAGAAAG--A--TCA--TT-CCAAA-G--GGT-TCGGGG--C-AAT-----	7416
VER.DE.-.AGM3	-GAGGAG-C-GG--G-AG--AAA-ATG--A--GT-A--TT-CCAAA-G--GAT-CCA-GG--C-AAT-----	6914
VER.KE.-.TY01	-GAGGTG-C-GGCGGG-AG--AAACAA--A--GT--GTT-CCAAA-G-C-GGT-T-A-GG--C-AAT-----	6888
COL.CM.-.CGU1	-C-GGGCAGAAT--GA-A-ACAAA-CAC---CGC-AT-G--A-CAA--G--TT---GCTTGGCTAGAGCGCATTTCCAGCAAAAAATA-----	6735
DEN.CD.-.CD1	-C--GA--AAGGC-GG-A---AA-AAGC-G-T-A--GG-G--AAC-TC--GGC-----	7355
GRV.ET.-.GRI_677	--AGGGG-A-GG--G-AG-CAGAGAAG--G--AGA--GTGAAAAATCT--CAG--GT--GC--AGAA-----	7287
GSN.CM.99.CN166	---GGG-AA---G-G-AGCTATG-AAGC-A-G-AT---T-GG-A---A-CTGGG-----	7232
GSN.CM.99.CN71	G---TAG-AA--GG--AGCCATG-AGGCCA-G-AGA-C--T-GG-A--G-TC--C-ATGAT-C-ACC-----	7232
DRL.-.-.FAO	GGA--GG-----GG-AGTC-----C-AAG-----A-TC-GT-AG---TCC-T-TGG-G--AT-CTACT.....AATAGCACATGCATAAGTCTCAAAGGCCAAAAGATCCAGGACAGAAAACGTGA	7122
RCM.GA.-.GAB1	-CGG-TG-A-GG--G-AG--AAAGAG-G-G-G-A-AC-AC-AAA--GT--CGAGGT--T-GT-A-ACA-----	7009
RCM.NG.-.NG411	--AGGTG-G-GG--G--AAAGAGGC-A-T-ACA-GAC-AAA--T---AGGT-----	7041
MND-2.-.-.5440	GGAC-GG-----C-TAGTC--T-ATTAG-GC-----A-GAT-AA--AT-GC-T-T-C-A-GAC-A-CAC.....AATATGACATGTAAAAGCAGTAACAACAAA.....AAGAACACAA	7076
MND-2.CM.98.CM16	GGA-TAG-CA-GC-TG-TC-----A-TAG-----A--AT-GACA-CTCC-TATGG-G--ATG-AACA.....AAGGGCTATTGCAAGGGGTAAAAGAGAAAG.....GAAAATAGAA	7449
MND-2.GA.-.M14	GGA-GAG---G--CGCAT--T-AC-TG-----AGA--GT-AA--C--G--T-T-T-G-GGTT-TAAT....GCAAGTGGCCCTTGCAGGGGATAAACAATAGAGGA....GTGGTAAACAAA	7375
MNE.US.-.MNE027	--GG-GG-AA-----G--G-GAAGCAG-CCA-TGTCA--TCCCA-GT--C-GG--T-AC-ATACT-----	7193
LST.CD.88.447	GGA-GA-TG--TC-TA-TT---GA-A--TGC-T-A-AGG-TTAA--C--CAGCG-TG-ATTAT-C-CATGGAAAC.....TGTATAGATAACACAAAAGCCTTGT.....G	6461
LST.CD.88.485	GGA-GG-TG--T--TA-TT---GG-AA-TGC-TCA-A-G-TTAA--C--CAGCA-TG-ATTAT-CAAT...GGACAA...TGTCTGAGGGAAAAACAGCCTTGT.....G	6458
LST.CD.88.524	GGA-GA-TG--T--CA-TT---GA--G-TG---AGA-G-TTAA--AG-C-GC--TG-ATTAT-AAGAA...GGAAAT...TGTTCACACCAGGAAAAACCTTGT.....G	6464
LST.KE.-.lho7	GGA-GA-TG--T--TA-TC---GC-AA-TGC--AG--GTTGAA--C-G-AGCC-TG-ATTAT-CAGAAGGGACAGGGACATGTGATTCTAAGAAAAACAACCTTGT.....G	7536
SUN.GA.98.L14	GGA-GAG-C--TGCCACCT--GG-AACT--G--GC-GGTGGAACCTTTAGC--C-TG-GT-ATGGCTGT...ACATTT...GATAATATAACAAGACATGCAGATTTACC...AATGGTACAGATT	7610
MND-1.GA.-.MNDGB1	GGGTTAG-AC-TTGGAGTC-----AAG--C--ATA--TT-AA--ACTCC-TC--GTG-ACCAGACCTGTAAAAATTTTACTAGCACTGGAGAGGAGAAACAAAACACCGGACAAGCAAAAAGGAGT	6894
MON.CM.99.L1	G-G-CTG-CC--C-GGCAG-TCACGAGGC-G--A-ACTG-G-GG-A--G--A--C-AT-----	7240
MON.NG.-.NG1	G-CGCTG-C--C-TGCAGCCCA--TRGCTA-C-RGR-GG--GG-A--GC--A--T-AT-----	5856
MUS.CM.01.1085	G--CCAG-A--G-G--AGCTATGCTAGCCA-G--A--T--GGCATGG--GATTGCA-----AAACGAGAACCACAGT	7212
DEB.CM.99.CM40	GCA--G-AAC--GG-A--GCGACAA--GA--TTG--AT-----GGCA--T-GG-C--AT-----	7056
DEB.CM.99.CM5	-CAGGAG-AG-T--GG-----AGG--G--AG--GAT--A-AATGTG-CA---T--C-G-A-----	6995
SYK.KE.-.KE51	GGA--GTT---TG-G--G-CCA--ATG-GAGT-T-A-GGTG...T-G--A--GTG--CTC--CA-----	6985
SYK.KE.-.SYK173	-CAG-ATT--T-G--G--CA-GAAC-GGCC-CGA--ACC...T-G--A--GT--C-AC-----	7286







PLV complete genomes

Table of PLV complete genomes with columns for sample ID, sequence (including V4 loop), and genome size. Samples include H1B.FR.83.HXB2, H1A1.UG.85.U455, H1B.US.90.WEAU160, H1C.ET.86.ETH2220, H1D.CD.84.84ZR085, H1F1.BE.93.VI850, H1G.SE.93.SE6165, H1H.CF.90.056, H1J.SE.93.SE7887, H1K.CM.96.MP535, H101.AE.TH.90.CM240, H102.AG.NG.1.IBNG, H1N.CM.95.YBF30, H1O.BE.87.ANT70, H1O.CM.91.MVP5180, CPZ.CD.1.ANT, CPZ.CM.1.CAM3, CPZ.GA.88.GAB2, CPZ.GA.1.CPZGAB, CPZ.US.85.CPZUS, CPZ.TZ.1.TAN1, H2A.GW.1.ALI, H2A.DE.1.BEN, H2A.SN.1.ST, H2B.GH.86.D205, H2B.CI.1.EHO, H2G.CI.1.ABT96, H2U.FR.96.12034, MAC.US.1.239, Env, SMM.SL.92.SL92B, SMM.US.1.H9, STM.US.1.STM, SAB.SN.1.SAB1C, TAN.UG.1.TAN1, VER.KE.1.AGM155, VER.KE.1.9063, VER.DE.1.AGM3, VER.KE.1.TY01, COL.CM.1.CGU1, DEN.CD.1.CD1, GRV.ET.1.GRI\_677, GSN.CM.99.CM166, GSN.CM.99.CM71, DRL.1.FAO, RCM.GA.1.GAB1, RCM.NG.1.NG411, MND-2.1.1.5440, MND-2.CM.98.CM16, MND-2.GA.1.M14, MNE.US.1.MNE027, LST.CD.88.447, LST.CD.88.485, LST.CD.88.524, LST.KE.1.lh07, SUN.GA.98.L14, MND-1.GA.1.MNDGB1, MON.CM.99.L1, MON.NG.1.NG1, MUS.CM.01.1085, DEB.CM.99.CM40, DEB.CM.99.CM5, SYK.KE.1.KE51, SYK.KE.1.SYK173.





V5 region

H1B.FR.83.HXB2	AAAGTAGGAAAAGCAATGTATGCCCTCCCATCAGTGGAA.....CAAATTAGATGTTTCATCAAATATTACAGGGCTGCTATTAAACAAGAGATGGTGGTAAT.....AGCAACAATG	7614
Env	<u>K V G K A M Y A P P I S G . . . . . Q I R C S S N I T G L L L T R D G G N . . . . . S N N</u>	gp120
H1A1.UG.85.U455	-G-----C-----GA-----CAA-----GT--A-G--GA-----C-----A--A-----G-----G-----C-AAC-----G-----T-CAA	7044
H1B.US.90.WEAU160	G-----C-----A-----GAA-----T-----T-----A-----G-AGT-----G-AGG-G--CC	7618
H1C.ET.86.ETH2220	GG-----CG-----TGAA-----AT--A-TG--AG-----C-----A--A-----C-G--C-T-----A-CC--A-----GAGCCAC-T-GCA	6993
H1D.CD.84.84ZR085	G-----C-----TGAA--G-----AT-----A-----G-----G-----GC-----AT--T-C-C	7135
H1F1.BE.93.VI850	GG-----CG-----A--T-----TGCA-----A-C--CC--AAC-----TT-A--G-----G-----G-G-G-A	6891
H1G.SE.93.SE6165	-GG--G--C-----GCA-----A-C--GA--AAC--GC-----A-AA-----G-----AAT-----AACAAATACA-AT-CA-G--	7032
H1H.CF.90.056	-G-----CG-----CAA-----A-C--TG--GT-----A-AA-----G-----TT--C-AG--C-----GCGTCTGCA	6933
H1J.SE.93.SE7887	-GGAC--C-----C-----TGCA-----A-C--CG--CA-----AT-A--G-----C-T-----A--A--CAGG-----GGGAATGGC--TG-G--	6921
H1K.CM.96.MP535	-G-----C-----A-----TGCA-----AGT--AC--AGC-----AA-A--G-----G-----AT--T-C-C	6774
H1O1.AE.TH.90.CM240	GG--C--C-----T-----AG-----AT--GT-----AA-A--G-----G-----GT-AAT-----T-CG-	7167
H1O2.AG.NG.-.IBNG	-C-----C-----CAA-----AT--A--GAT--C-----A--A--G-----A--G-----T-G-A	7120
H1N.CM.95.YBF30	-GG-----G--T--A--A--A--TC-G-----GTTT--AC--A-C-C-----T-A--G-TC-GA-TATAG-----GCC-----G-----G	7159
H1O.BE.87.ANT70	-GG-G-CAGTCG-G-C-C--A-----AA--T-----A-TC-A-C--ATG--C--A--T--AA--A-CC--CA--TG--AACACATGGAACAGCAGCAACAC--ATGTA-CAT	7668
H1O.CM.91.MVP5180	-G-G--AGTCGAG--C--A-----CCC-C-----A-CT-A-C--CAT--C-C--A--T--AA--A-TC--CAGTT--CAACCATGG-----ATTG--CA-	7670
CPZ.CD.-.ANT	-T--TTC-----GC--T--CTTAG--C--T-GG--A--G-----A-TG--TCC--A--T--C-GC--A--T--AA--TA--G--G...GA--G-CAAATAT--	7040
CPZ.CM.-.CAM3	-G--G--G--G--G--T-----A--A--T--A-----GTT--A--T--CA-----CG-----T--A--TA--C--TGAG--C--GGCA--C--TC-----AAC--A--G--TCA	7024
CPZ.GA.88.GAB2	-G--G--G--GT--C--T-----G--T-GA--A-----ACT--A--CC--AAT--C--CC-----A--C--C-----GA--GT--C-GAA-----G--CGGGA	6983
CPZ.GA.-.CPZGAB	-GG-----G--G--A--A--C--A-----A--C--C--CT--CAAT--T--C-----T--A--T--C-----TTC--CAC--CC--GTGACT-----AAC-----G--	7636
CPZ.US.85.CPZUS	-G-----G--C--T-----A--A--AA--T-----GTGC--A--T--AAC-----A--G--AA--A--TC--TGAC--T--AGCATAA--CGCAGTC-----AATAAC...G--T-G-A	7616
CPZ.TZ.-.TAN1	-T--GTC-----GGG--T--CCTT--C--A-GG--G-----AC--G--A--A-----CAC-----C--C-----A--TA--CA--G--GC--AAAA--AC--CAAT-----	7230
H2A.GW.-.ALI	----C--G--AATG--A--TTG--T--GGGAA-----G--GT--G--CC--CAAC-----CAG--A--CA--CT--A--TGCT--AC--TT--CACG--A--GGC-----	8070
H2A.DE.-.BEN	----G--AATG--A--TTG--T--GGGAA--G-----G--GT--GGCC--GA-----CAG--A--CA--AA--TGCT--AC--TT--CATA--A--AAAT-----C	8084
H2A.SN.-.ST	-G--G--G--AATG--A--TTG--T--GGGAA-----G--GT--A--CC--CAAC--T--CAG--G--CA--CA--AA--TGCT--AC--TT--C--A--AG--G-----	7516
H2B.GH.86.D205	-G--T--G--AAT--A--TTG--A--GAGAA--T-----G--GC--CTC-----AAT--C--C--G--G--TA--C--CA--GCC--AC--TCA--CTCA--AC-----A	8060
H2B.CI.-.EHO	-GA--T--G--AATG--TTG--A--GAGAG--T-----G--C--CTCC--AAT--C--C--G--CA--C--CA--GCC--AC--TT--CT--GATAG--AAG-----	8048
H2G.CI.-.ABT96	----AATG--CTTG--C-----G--TC--G--CG--AAT-----CAG--C--CA--C--AA--GCG--AT--T--A--AGAA--GGTCTAAC-----	7447
H2U.FR.96.12034	----G--AATG--TTG--A--GAGAA-----G--TC--C--CC--CAAT--T--CAG--A--TA--CA--A--TGC--AT--TT--C--GAT--G--GGA-----	7570
MAC.US.-.239	----C--AATG--T--TTG--A--GAGAG-----G--CC--C--CG--AAC--C--CAG--G--CA--T--CA--GC--AC--T-----T--GAT--G--GGA-----	8028
Env	<u>K V G K N V Y L P P R E G . . . . . D L T C N S T V T S L L I A N I D W I D G . . . . .</u>	gp120
SMM.SL.92.SL92B	-GG-----AATG--C--TTG--A--GGGAA-----G--CT--A--CT--AAC--C--CAG--G--CA--C--AA--GC--AT--TT--T--GAT--G--CAAC-----	7479
SMM.US.-.H9	----C--AATG--R--TTG--T--GGGAA-----G--CC--G--C--AAC--C--C--G--A--TA--C--CA--GC--GA--T-----T--GATC--GGC-----	7517
STM.US.-.STM	----G--AATG--T--TTG--A--GACAA-----G--CC--CT--AAT--C--C--G--G--CA--TA--AA--GC--AC--TT--CT--GAC--CAAT-----	7682
SAB.SN.-.SAB1C	-CCT--GTCT--AAG--CA--C-----GAGAA--C-----CT--GGAG--AAC--G--CAGCC--C-----TATG--GA--TTGA--CTA--AAC--G--AAA-----	7917
TAN.UG.-.TAN1	-CGC--CA--T--GAA--CC-----A--A--A--GGGAA-----CT--AGA--CA-----CAG--G--TCTA--A--GG--T--CTG--TA--AAC--GCAAA-----	7721
VER.KE.-.AGM155	-CTA--TC--GAA--CA-----T--A--A--GAGAA-----TT--GCAG--CA--G--C--CAG--T--A--AC--G--GAGCT--A--CTA--AA--CCAG-----	7741
VER.KE.-.9063	-CCA--TC--GAA--G-----T--A--A--GGGAA-----TC--GGA--CA-----C--CGG--C--T--TA--AC--G--GA--CT--A--TA--AAC--AAC-----	7761
VER.DE.-.AGM3	-C--C--TC--GGA--CC-----A--G--A--GAGAA--G-----CT--GCA--CA-----C--CGG--A--G--TA--TC--G--GGAGCT--A--TACAA--G--AAG-----	7259
VER.KE.-.TY01	-CCC--TC--GAA--CT-----T--G--A--GAGAA-----TT--GCA--CAGG--C--CAG--C--T--A--AC--G--TGAGCTTA--TA--AA--G--AAA-----	7230
COL.CM.-.CGU1	-CTCAT--CC--GGATCC-----GG--G--TCCAG--A--T-----TC--ACA-----AAC--GGG--G--AACAGCCAG--TA--TGCGTTC--TG--GCACCATAG--AGGA....GACAATGATGGT----	7012
DEN.CD.-.CD1	-C-----C--TG--AT--T--CTTA--A--AGCAGAA-----G-----C--CA--C--C--G--GT-----CA--TA--TG--GAC--CG-----TCCA--CCGACAGA-----	7652
GRV.ET.-.GRI_677	-CT--CTCT--AAGG--A-----T--A--A--GGGAA--T-----TT--GGAG--AAC--C--CAG--C--G--CA--ATACG--GG--T-----TA--AAC--CAAG-----T	7621
GSN.CM.99.CN166	T--T--GTCT--T--C--CTTA-----ACGGCAG--T-----T--A--AT--CA-----C--C-----T--T--G--A--GATG--C--C--ATGTA--G--AAG-----	7553
GSN.CM.99.CN71	T--T--GTCT--GGT--G--C--CTTA-----ACGGCAA--T-----T--A--AT--A-----T--C--C--T--C--C--A--A--GAGC-----CGATGTA--GGACAT-----	7559
DRL.-.-.FAO	-G-----C--GGAAG--T--ATG-----TG--ATCA--TTTAATAATG--A-----A--CA--G--G--C-----AAA--T--C--TGAG--TCC--AAAATTGG--GACAGAGGAAACAGGAGATTCC--C.	7454
RCM.GA.-.GAB1	-C-----TCG--GGAAGG--A--CA--G--A--G--GGCCA--AT-----GCCT--A--AG--CAGTG--C--GG--AT--CTAT--T--GGC--GAC--T--CTA-----TT-----	7307
RCM.NG.-.NG411	-C-----GTCT--G--AA--G--A--CA--T--A--A--GGCCA--AT-----GCCT--G-----CAACG--C--C--G--G-----TACT--T--GC--GAT--TT--CTA--ACAG-----A	7351
MND-2.-.-.5440	-----C--GAA--T--ATG-----TG--GTC-----GTTTAAATAATG--A-----A--TAATG--G--G-----AAA--T--C--TGAGGTCC--AAAAC--G--GATGAC-----	7394
MND-2.CM.98.CM16	-----GAAG--T--ATG-----TG--ATC-----TTTAAATAATG--G--A-----TAATG--G--A-----AAA--T--C--TGAGGTTC--AAAGACAG--AGAGGGA-----	7763
MND-2.GA.-.M14	-----G--GAAG--T--ATG-----TG--GTC-----TTTAAATAAT--T--A-----CA--TAATG--G--C-----A--A--T--T--TGAGGTCC--ACTA--T--G--AGGA-----AA	7690
MNE.US.-.MNE027	-----C--AATG--T--TTG-----A--GAGAG-----G--CC--C--CG--AAC--C--CAG--A--CA--T--CA--GC--AC--T-----T--GAT--G--GGA-----	7505
LST.CD.88.447	TCT-----G--GAA--T--CT-----A--CATCA--CTTTAACAACAG--A-----A--TCAT--GAG--A-----AAA--TGG--TCA--TG--AAAGTT--G--ACCAGAAGAAAACCTGG--AGGA--T	6794
LST.CD.88.485	-CT-----G--GAA--C--AT-----A--CATCA--CTTTAACAACAG--A-----A--TCAT--GAG--A-----AAA--TGG--TGA--CTG--AAAGT--GG--GCCAGAGAAAACCTGG--GGGA--T	6791
LST.CD.88.524	-CT--T--G--GAA--T--CTT-----A--CATCA--CTTTAACAATAG--A-----A--TCAT--GAG--A-----AAA--TGG--TGA--TG--AAAAT--GG--ACCACATGAAGACTTAG--AGGA--C	6797
LST.KE.-.lho7	-C--A-----GAAG--T--CTA--A--T--CTTCA--CTTTAACAATAG--C-----A--TCAT--GAG--A-----AAA--T--C--TGAG--TG--AAAAT--GG--ACCACATGAGACTTAG--AGG--CC	7869
SUN.GA.98.L14	TC-----G--C--GAAG--T--CTTG--C--A--CATCA--GTTTAAATAACAG--A-----G--AAC--GAG--G-----AAGCATGG--C--CTG--AA--G--TGG--AGACTGGAAGGTAATG--ATCA--CA	7937
MND-1.GA.-.MNDGB1	C-----G--AAG--A--TTA--A--T--CATCA--TTTAAATAAT--C--A--G--A--CAT--GGG--A-----AAA--TAC--TGA--TGC--AAAGATAG--AGTAATGAACAACAAA-----	7217
MON.CM.99.L1	T--T--GTC-----GAGT--C--CCTA-----ACGG--AA--C-----TG--GCAG--A--C--T-----C--T--CC--C--GA--T--G--G--GTTGTA--C--AAAC-----	7579
MON.NG.-.NG1	T--CA--CTCC-----GAGTR--C--CTG--A--A--GACAA--C-----TG--GCAG--A--C--T-----C--C--C--G--G--GATG--G--GTTGTA--GGCAGC-----	6204
MUS.CM.01.1085	T--T--GTC-----GT--T--C--CTTG-----ACG--AG--C-----TTG--C--GAG--T-----A-----T--CA--TGACG--CCATGTA--GAGAT-----	7547
DEB.CM.99.CM40	-CG-----GC--CAT--C--CTA-----G--AAGAA-----G--GT--G--AG--CAGC--C--C--T--T--CAT--TG--T--TGATGT-----CACTA-----GGG-----	7350
DEB.CM.99.CM5	-G-----C--CATG--A--TTG-----A--AGGAA-----G--T--G--AG--CAT--TTC--G--T--T--CA--TG--T--TGAGGT--C--CA--TA--GGGCAA-----	7271
SYK.KE.-.KE51	T--T--ACC--GCT-----TTG--A--T--CAGAA--G-----C--AG--CA--T--T-----T--CTG--AT--AC--GAT--T--GTACTA--CCAAAT-----	7273
SYK.KE.-.SYK173	T--T--A-----TTG--A--CCTG--C--T--CGGCA--G-----T--A--A--A--T--C--G--C--T--CAG--AT--GACGGAC--T--ATACTACCAGGA-----	7577



PLV complete genomes

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

Table with columns for sequence identifier (e.g., H1B.FR.83.HXB2), V5 region, sequence alignment, and genomic position (e.g., gp120, 7171, 7745, etc.).



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

gp120 / gp41

H1B.FR.83.HXB2	GCAG.....AGAGAAAAAGAGCA...GTGGGA...ATAGGAGCTTTGTCTCTGGGGTCTTGGGAGCAGCAGGAAAGCACTATGGGGCAGCCTCAATGACGCTGACGGTACAGCCAGA	7850
Env gp120	_Q_..._R_E_K_R_A_..._V_G_..._I_G_A_L_F_L_G_F_L_G_A_A_G_S_T_M_G_A_A_S_M_T_L_T_V_Q_A_R_	gp41
H1A1.UG.85.U455	-G-...-A-C-...-T-...-C-...-A-C-...-A-...-T-...-A-...-G-...-G-...-A-...	7280
H1B.US.90.WEAU160	.....ATGC-...A-...-A-...-G-...-A-...-A-...-A-...	7857
H1C.ET.86.ETH2220	-G-...-GCA-...-C-...-C-...-A-...-A-...-A-...-A-...	7226
H1D.CD.84.84ZR085	-G-A-...-A-A-...-C-...-A-...-A-...-G-...-G-...-A-...	7371
H1F1.BE.93.VI850	.....-C-...-C-...-TC-T-G-ACAGCA-GGAGCA-...-G-G-G-A-...-A-...	7127
H1G.SE.93.SE6165	-G-...-T-...-C-G-...-G-C-...-A-...-A-...-G-...-G-...-A-...-T-...	7268
H1H.CF.90.056	-G-...-CT-...-G-...-G-...-G-...-A-...-A-...-G-...-G-...-A-...-G	7169
H1J.SE.93.SE7887	-G-...-G-...-G-...-A-...-T-...-A-...-T-...-A-...-G-...-A-...-T-...	7157
H1K.CM.96.MP535	.....-A-...-T-...-G-...-T-...-A-...-A-...-G-...-G-...-A-...-A-	7010
H1O1.AE.TH.90.CM240	-G-...-G-...-G-...-A-...-A-...-A-...-A-...-A-...-G-...-G-...-A-...-A-	7403
H1O2.AG.NG.-.IBNG	-G-...-G-...-T-...-C-G-...-T-...-G-C-...-T-...-A-...-A-...-G-...-G-...-A-...-A-	7356
H1N.CM.95.YBF30	-AGT-...-GCCT-T-...-C-...-T-...-GC-...-T-...-TC-T-...-G-...-G-...-A-...-A-...-C-G	7398
H1O.BE.87.ANT70	T-...-A-...-T-G-...-ATGC-A-...-T-G-...-G-TC-AA-T-...-T-...-GG-...-CA-...-G-...-A-...-CAC	7904
H1O.CM.91.MVP5180	AA-CATTACACCCCTCAC-G-...-A-...-T-G-...-ATGC-A-...-T-G-...-G-GC-AA-T-...-T-...-GG-...-CAG-...-G-A-...-CAC	7921
CPZ.CD.-.ANT	AA-ACAACACTCC.....-C-...-GCATT-...-GCTG-C-...-T-G-TC-TC-CA-T-...-T-C-T-A-...-G-...-AG-...-C-...-A-...-G	7282
CPZ.CM.-.CAM3	ATCA.....GCCT-T-...-C-G-T-...-GC-...-T-...-A-...-A-...-A-...-G-...-GT-...-C-...-C-G	7263
CPZ.GA.88.GAB2	-GCAAAAGCACTTGAT.....-TCC-...-GCCT-T-...-C-G-G-...-GC-...-T-...-A-...-C-T-...-G-...-A-...-TG-...-T-...-A-...-TC-...	7234
CPZ.GA.-.CPZGAB	-GCAAGACAGAAAGAC.....-C-G-...-GCCT-C-...-T-...-C-G-...-AC-...-C-...-T-...-GG-...-G-A-...-A-...-C-...-C-	7887
CPZ.US.85.CPZUS	-A-A-...-T-...-AC-...-G-...-T-...-AC-...-G-...-TC-...-T-...-A-...-G-AGT-...-C-...-C-...-	7855
CPZ.TZ.-.TAN1	TA-GCCAATCATACT...-GTC-G-...-ATGTG-...-C-...-CTG-...-T-...-A-...-TC-TA-T-...-T-...-A-...-G-...-AG-...-C-...-	7476
H2A.GW.-.ALI	T...ACT....CCAAGG-GA-T-...GTGTGT-C...GTGC-...G-...-T-A-T-TC-C-GA-...TTCTG-A-...A-G-AG-TT-A-...-T-T-CT-...-T-TC-G	8307
H2A.DE.-.BEN	A...ACT....CCAGTG-GA-C-...GTGTGT-C...GTGC-...G-...-T-G-T-TC-C-GA-...TTCTG-A-...GCGG-CC-...-T-A-CC-...-T-C-G	8323
H2A.SN.-.ST	T...GCT....CCAGTG-GA-T-...GTGTAT-C...GTGC-...G-...-T-A-T-TC-CAGCA-...GCTG-A-...G-G-CT-...-T-CT-...-T-TC-G	7753
H2B.GH.86.D205	A...GTA....AAACCG-GA-T-...GTGTAA-...GTGC-...G-...-T-G-A-C-T-C-ATG-...TTCTG-A-...A-G-TT-...-T-A-CT-...-T-C-G	8299
H2B.CI.-.EHO	A...GTG....ACACCG-GA-T-...GTGTAC-T...GTGC-...G-...-TT-G-A-C-T-C-GA-G-...TTCTG-A-...G-G-CT-...-T-T-CT-...-T-C-G	8285
H2G.CI.-.ABT96	T...GTG....ACACCG-A-A-T-...GTGTAT-T...GTGC-...G-...-TT-G-A-TC-T-C-ATG-...TTCTG-A-...A-...G-GT-...-T-CT-...-T-C-G	7684
H2U.FR.96.12034	A...GTG....ACACCG-GA-T-...G-GTAT-T...GTGC-T-G-...-T-G-T-TC-C-CCA-G-...TTCTG-A-...T-G-G-GT-...-T-A-CT-...-T-C-G	7807
MAC.US.-.239	TGGTGGC....ACCTCA--A-T-...GGGICT-T...GTGC-...G-...-T-G-T-TC-C-A-A-G-...TTCTG-A-...G-G-GT-...-C-CT-...-T-C-	8268
Env gp120	_G_G_..._T_S_R_N_K_R_G_V_F_..._V_L_G_..._F_L_G_F_L_A_T_A_G_G_S_A_M_G_A_A_S_L_T_L_T_A_Q_S_R_	gp41
SMM.SL.92.SL92B	A...AGT....ACCTCA-AGA-T-...G-GTCT-T...GTGC-...G-...-T-G-T-TC-C-GA-G-...TTCTG-A-...G-GC-A-...-T-CT-...-T-C-G	7716
SMM.US.-.H9	AACTGGT....GCCTCA--A-T-...GGGTCT-T...GTGC-...G-...-T-G-T-TC-C-GA-...TTCTG-A-...G-GT-...-T-CT-...-T-C-G	7757
STM.US.-.STM	T...AGT....ACCTCA--ACT-G-...GGGTCT-C...GTGC-...G-...-T-G-T-TC-C-GA-...TTCTG-A-...G-GC-...-CT-...-T-C-G	7919
SAB.SN.-.SAB1C	C.....CCCGAA--C-...G-TGCCAT-C...GTGC-...G-...-A-T-TC-C-GA-...T-GCTG-A-...T-...GG-G-CAG-C-...-C-...-T-TCAG	8151
TAN.UG.-.TAN1	--CAACT.....C-...G-TGCCCT-C...GTGT-G-G-...-A-C-T-...G-...T-...GCTG-A-...T-...A-AG-G-CAG-T-...-C-...-T-CAG	7955
VER.KE.-.AGM155	A...GGC....CAAGAG-GC-...C-...TCCCCT-C...GTGC-...G-...-A-C-...T-...T-G-CTG-A-...A-...GG-G-CAG-C-...-C-...-T-TCAG	7978
VER.KE.-.9063	A...GGA....CAAGAC-TC-...C-...TCCCCT-T...GTGC-...G-...-A-C-...A-...T-...T-CTG-A-...A-...GG-G-CAG-C-...-C-...-T-TCAG	7998
VER.DE.-.AGM3	A...GGT....CATGAC-AC-GC-...TCCCCT-C...GTGC-...G-...-A-C-...A-...T-...T-G-CTG-A-...A-...GG-G-CAG-C-...-C-...-T-TCAG	7496
VER.KE.-.TV01	A...GGA....CATGAG-GC-...G-...TCCCCT-T...GTGC-...G-...-T-A-C-...-G-T-...T-CTG-A-...A-...GG-G-C-GC-...-C-...-T-C-G	7467
COL.CM.-.CGU1	TGTAATAATGGCATCATGGG--C-...G-...GCATCT-T-C-TTCTCCATC.....T-A-CCC-...A-T-GG-C-TGCTG-A-...T-...G-TG-G-...-A-C-...-TCA-	7257
DBN.CD.-.CD1	A-CAGAATCGGCAGAGTCTGTTTC-CCGC-AGGCGG-AA-TGTCTT--TGCT-GG-C-A-A-A--TC-CA-C-...CTGTC-...G-T-GG-...CAGGCT-AG-...-T-C-G	7907
GRV.ET.-.GRI_677	C-CCACA.....G-TGCCAT-C...GTGC-...G-...-T-A-C-...T-...T-CTG-A-...G-...GG-CA-...-A-C-...-T-TC-G	7857
GSN.CM.99.CN166	C-CC.....GCTGTG-G-...C-...TTC-...T-G-...ATA-CC-...T-GAC-...TC-CA-T-G-...T-C-...G-...GG-...CAG-...-A-...-T-C-G	7787
GSN.CM.99.CN71	T-CAGGGGTG....CATGTGA--G-G-...C...A-TTC-...T-G-...CATA-CC-...T-GAC-...TC-CA-T-G-...TGC-...G-...GG-...CAG-...-AT-...-C-G	7799
DRL.-.-.FAO	--C-GGA.....CAC-GC-G-G-...GTGCA-...TCGTGC-T-G-...A-...-CC-...A-CTTG-...TTC-G-A-...T-...-TGG-G-G-C-CA-...-T-C-...-T-CAG	7691
RCM.GA.-.GAB1	-...-A-C-G-...-G-T-...CCTTTG...G-GCTG-G-...-T-A-T-...C-CTC-...T-CTG-A-...GG-G-CAG-C-...-T-C-...-T-C-G	7537
RCM.NG.-.NG411	A.....-A-C-...-G-T-...CCTTTG...G-GCTG-G-...-T-A-T-...C-CTCG-...T-CTG-A-...GG-G-CAG-C-...-A-C-...-T-C-G	7581
MND-2.-.-.5440	--CTAAT.....ACA-GC-...G-GCG-...TCTTAC-T-...A-G-C-CC-C-A-...TTTG-...TTC-G-G-...T-...TGG-GG-G-C-...-T-C-...-T-CAG	7633
MND-2.CM.98.CM16	--CAGAA.....GCT-GC-G-G-...G-GCG-...TCTTGC-T-G-...A-GT-C-CC-C-A-...TTTG-...TTC-G-G-...T-...TGG-GG-G-C-...-T-C-...-T-TCAG	8000
MND-2.GA.-.M14	A-CAGAT.....GCA-AGC-G-G-...G-GCA-...TCTTGC-T-...A-G-C-CC-...CTTG-...TTC-G-G-...T-...TGG-GG-G-C-...-T-C-...-T-CAG	7929
MNE.US.-.MNE027	TGGTGGC....ACCCCA--A-T-...GGGTCT-T...GTGC-...G-...-T-A-T-TC-C-C-A-G-...TTCTG-A-...G-G-GT-...-C-CT-...-T-C-G	7745
LST.CD.88.447	C.....AGTAGT-AGC-G-G-...GCACC-CTGGCTC--T-...GC-T-G-AC-TC-CA-T-T-...T-CTG-A-...TT-TA-G-C-TA-A-T-C-A-...-CA-	7030
LST.CD.88.485	C.....AGTAGT-AGC-G-G-...GCACC-CTGGCTC--T-...GC-T-G-AC-TC-CA-T-T-...T-CTG-A-...TT-TA-G-C-TA-A-T-C-A-...-CA-	7027
LST.CD.88.524	C.....AGGGGA-GC-...G-...GCTCCACT-GCTC-T-...GC-T-G-AC-TC-CA-T-T-...T-CTG-A-...TT-TA-G-C-TA-A-T-C-A-...-CA-	7033
LST.KE.-.lho7	C.....AGGGGA-GC-...CC-...GGCCTCTGGCTT-...GC-T-G-AC-TC-CA-T-T-...T-CTG-A-...TT-TA-G-C-TA-A-T-C-A-...-CA-	8105
SUN.GA.98.L14	C.....AGG-AG-...G-G-...GCACCT-TTGCCC-...GCTT-G-CAC-TC-CA-T-T-G-...T-CTG-A-...TT-TA-G-C-TA-A-T-C-...-CA-	8173
MND-1.GA.-.MNDGB1	A-CTGAA....CATCAT-A--G-G-...G-GCA-...T-ATAC-T-T-...A-...-TC-GC-CTCGCTG-...TC-G-G-...T-...TG-...GG-G-A-...-T-C-...-AT-TCAG	7457
MON.CM.99.L1	-GTA.....CAC-G-...G-...ATT-...ACTT-G-...TATGGCC-...T-G-C-...C-CA-CA-G-...G-...G-...GG-G-CAG-C-...-T-TC-G	7813
MON.NG.-.NG1	-TA.....ATG-...G-C-...TATC-...ACAT-G-...ATGGCC-...T-G-C-...TC-CA-CA-G-...G-...G-...GG-G-CAG-C-...-T-C-G	6438
MUS.CM.01.1085	C-CA.....CATGTGCAT-C-G-GCCATCTCA...TTGGGC-...ACC-...-TT-A-A-TC-CA-C-...G-...G-...GG-G-CAG-C-...-T-TC-G	7784
DEB.CM.99.CM40	A-C.....GAGTCGGTA--C-C-G-G-...G-GCA-GAATTGCTT-T-T-...C-GG-A-CA-TC-CA-TA-...TGCTG-...G-...CAG-...-T-TC-C	7590
DEB.CM.99.CM5	A-C.....GAGTCGTA--C-...G-GGCG-GAATTGCTT-T-T-...C-AG-G-CT-TC-A-TA-...TGCTG-...G-...CAG-...-T-TC-C	7511
SYK.KE.-.KE51	--CT.....ACG-A-C-G-...GCACCATGGCAC-T-G-...-G-A-AC-CTCG-...G-T-CTG-A-...GAGC-TG-G-CAG-...-T-AC-C-...-T-CAG	7510
SYK.KE.-.SYK173	--CT.....AATACC--G-...G-...GCACCATGGCAC-T-G-...-G-A-AC-CTCG-...T-G-C-CTG-A-...G-GC-TG-G-CAG-...-A-AC-C-...-T-TCAG	7817





PLV  
complete genomes



Strain	Sequence	Position
H1B.FR.83.HXB2	CAATTATGTCTGGTATAGTGCAGCAGCAGAAACAATTGCTG...AGGGCTATTGAGGGCGAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAA	7977
Env	<u>Q L L S G I V Q Q Q N N L L . R A I E A Q Q H L L Q L T V W G I K Q L Q A R I L A V E</u>	gp41
H1A1.UG.85.U455	-----A-----G-----C-----	7407
H1B.US.90.WEAU160	-T-----A-----C-----CG-----	7984
H1C.ET.86.ETH2220	-----G-----A-----A-G-----AA-----A-----G-----G-----T-----A-----G-----A-A-----	7353
H1D.CD.84.84ZR085	---G-----A-----A-----A-----	7498
H1F1.BE.93.VI850	-----A-----A-----A-----	7254
H1G.SE.93.SE6165	-----C-----A-----AGG-----C-----	7395
H1H.CF.90.056	-----G-----C-----A-----A-G-----A-----AC-----G-----A-----G-----G-----T-----A-----G-----C-----	7296
H1J.SE.93.SE7887	-----G-----C-----A-----A-G-----A-----A-----A-----A-----A-----G-----T-----T-----A-----G-----G-----A-----	7284
H1K.CM.96.MP535	-----G-----C-----A-----A-G-----A-----A-----A-----A-----A-----A-----G-----G-----T-----A-----G-----A-----	7137
H1O1.AE.TH.90.CM240	-----G-----C-----A-----A-G-----A-----A-----A-----A-----A-----A-----G-----G-----T-----A-----G-----A-----	7530
H1O2.AG.NG.-.IBNG	-----C-----C-----A-----A-----G-----AA-----A-----T-----A-----A-----G-----T-----A-----G-----C-----	7483
H1N.CM.95.YBF30	AC-----A-----G-----A-----T-T-C-T-----A-----A-----A-----A-----A-----T-----A-----T-----A-----A-----A-----	7525
H1O.BE.87.ANT70	ACT-GC-AAG-----A-----G-----CC-----A-----A-----A-----AC-----C-----G-----AT-----C-----AGG-----AT-----TR-----A-----T-----GA-----A-----GA-----TC-----CC-----G-----A-----CT-----A-----	8031
H1O.CM.91.MVP5180	AGTG-C-AAG-----A-----G-----CC-----A-----A-----G-----AC-----C-----G-----A-----CT-----C-----AGGT-----AT-----T-----A-----T-----T-----GA-----A-----GA-----TC-----CC-----G-----AA-----CT-----A-----	8048
CPZ.CD.-.ANT	A-T-GY-CCA-----T-----A-----A-----GC-----C-----CAA-----C-----A-----A-----A-----G-----A-----C-----A-----G-----T-----G-----AG-----A-----A-----A-----G-----T-----A-----C-----G-----	7409
CPZ.CM.-.CAM3	-----A-----G-----A-----CC-----C-----A-----A-----A-----A-----T-----T-----T-----T-----A-----A-----A-----A-----G-----T-----A-----	7390
CPZ.GA.88.GAB2	A-CC-GC-T-G-A-----A-----A-----T-----C-----C-----A-----C-----A-----A-----T-----G-----A-----C-----C-----G-----AT-----A-----A-----A-----G-----GG-----T-----A-----	7361
CPZ.GA.-.CPZGAB	---G-----GC-----G-----A-----T-----C-----T-----AA-----A-----A-----T-----C-----A-----C-----A-----AT-----A-----AG-----A-----A-----A-----A-----C-----T-----G-----T-----A-----	8014
CPZ.US.85.CPZUS	-----A-----G-----A-----T-----C-----A-----A-----A-----A-----C-----T-----A-----AT-----T-----T-----A-----T-----A-----C-----G-----A-----T-----	7982
CPZ.TZ.-.TAN1	GG-----C-----C-----T-----A-----AC-----A-----CC-----T-----CA-----C-----A-----A-----A-----CT-----A-----G-----T-----T-----A-----G-----	7603
H2A.GW.-.ALI	ACT-----G-----C-----G-----A-----C-----AC-----GC-----T-----GAC-----TGG-----CA-----AGA-----G-----AA-----G-----G-----C-----A-----CG-----AA-----T-----G-----ACT-----A-----C-----G-----	8434
H2A.DE.-.BEN	ACT-----C-----G-----C-----G-----A-----C-----AC-----GC-----T-----GAC-----TAG-----CA-----AGA-----AG-----AA-----G-----G-----C-----A-----CG-----AA-----C-----G-----ACT-----A-----C-----G-----	8450
H2A.SN.-.ST	ACT-----G-----C-----G-----A-----C-----AC-----GC-----T-----GAC-----TGG-----CA-----AGA-----AG-----AA-----G-----G-----C-----A-----CA-----AA-----T-----G-----ACT-----A-----C-----G-----	7880
H2B.GH.86.D205	ACT-----C-----G-----C-----G-----A-----C-----AC-----GC-----T-----GAC-----TGG-----CA-----AAGA-----G-----A-----GG-----G-----C-----G-----CG-----AA-----C-----G-----ACT-----CA-----C-----G-----	8426
H2B.CI.-.EHO	ACT-----C-----G-----G-----A-----C-----AC-----GC-----G-----GAC-----TGG-----CA-----AAGA-----AG-----A-----GG-----G-----C-----G-----CG-----AA-----C-----G-----ACT-----CA-----C-----G-----	8412
H2G.CI.-.ABT96	ACT-----G-----G-----G-----A-----C-----AC-----GC-----T-----GAT-----TGG-----CA-----AGA-----G-----AA-----G-----G-----C-----G-----CG-----AA-----C-----A-----T-----G-----ACT-----CA-----C-----G-----	7811
H2U.FR.96.12034	ACT-----C-----G-----G-----A-----C-----AC-----GC-----T-----GAC-----TGG-----CA-----AAGA-----AG-----AA-----GG-----G-----C-----A-----CT-----AA-----C-----A-----T-----G-----ACT-----CA-----C-----G-----	7934
MAC.US.-.239	ACT-----G-----G-----G-----A-----C-----AC-----GC-----T-----GAC-----TGG-----CA-----AGA-----AG-----AT-----G-----G-----C-----A-----CA-----A-----C-----A-----T-----GG-----ACT-----CA-----C-----G-----	8395
Env	<u>T L L A G I V Q Q Q Q Q L L . D V V K R Q Q E L L R L T V W G G K N L Q T R V G A I E</u>	gp41
SMM.SL.92.SL92B	ACT-----G-----G-----G-----A-----A-----C-----AC-----GC-----T-----GAC-----TGG-----CA-----AGA-----AG-----AT-----G-----G-----C-----A-----CT-----AA-----T-----A-----T-----G-----ACT-----CA-----C-----G-----	7843
SMM.US.-.H9	ACT-----G-----G-----G-----A-----A-----C-----AC-----GC-----T-----GAC-----TGG-----CA-----AGA-----AG-----AT-----G-----G-----C-----A-----CT-----AA-----C-----A-----T-----G-----ACT-----CA-----C-----G-----	7884
STM.US.-.STM	ACT-----A-----G-----G-----A-----C-----AC-----GC-----T-----GAC-----TGG-----CA-----AGA-----AG-----AT-----G-----G-----C-----A-----C-----A-----C-----A-----T-----G-----ACT-----A-----C-----G-----	8046
SAB.SN.-.SAB1C	---G-----C-----CG-----G-----C-----A-----C-----T-----GC-----G-----A-----CA-----G-----AA-----C-----AA-----G-----TA-----T-----TG-----A-----AA-----C-----A-----T-----CC-----CG-----ACT-----C-----T-----G-----	8278
TAN.UG.-.TAN1	---C-----TG-----A-----G-----T-----G-----C-----GC-----AG-----CA-----G-----GA-----C-----AA-----G-----CA-----T-----TG-----G-----AA-----C-----A-----T-----CC-----CG-----ACT-----CC-----T-----	8082
VER.KE.-.AGM155	---T-----C-----TG-----G-----T-----G-----C-----GC-----G-----G-----GA-----T-----GA-----A-----G-----G-----CA-----T-----TG-----G-----AA-----C-----A-----T-----CC-----CG-----ACA-----C-----T-----G-----	8105
VER.KE.-.9063	---T-----C-----TG-----G-----A-----A-----C-----GC-----G-----G-----T-----AA-----A-----G-----G-----CA-----T-----TG-----G-----AA-----C-----A-----T-----CC-----CG-----ACA-----C-----T-----G-----	8125
VER.DE.-.AGM3	---T-----C-----TG-----G-----C-----G-----GC-----G-----G-----T-----GA-----A-----G-----G-----CA-----T-----TG-----G-----AA-----C-----A-----T-----CC-----CG-----ACA-----C-----T-----G-----	7623
VER.KE.-.TY01	---T-----GC-----TG-----G-----C-----G-----GC-----G-----G-----T-----GA-----A-----G-----G-----CA-----T-----TG-----T-----AA-----C-----A-----T-----CC-----CG-----ACA-----CC-----T-----G-----	7594
COL.CM.-.CGU1	TCCC-TAATGGAA-GGCTAGTGCCAGCAGC--CGAA-T-CTG-AACTGG-C--A--GTCTGCA--C-A-G--G--GG-T-AA-C--T-C-G-TGCCA--A-A--G	7387
DEN.CD.-.CD1	G-T--C-TG-C-G--G--A--C-AG-GC-T--A--G--G--A-GT--CAGTGCA--AC--GT-GT--A--T--AA-C--A-C--TC-CC-GACT--AA-A--G	8034
GRV.ET.-.GRI_677	---T--GC-TG--G--T--G--G--C--GC--G--G--ACA--A--GT--A--G--G--CA-T--TG-G-AA-C--A-T--CC-CG--ACA--C--G	7984
GSN.CM.99.CN166	TCT-G--G--G--A--G--G--CC--G--G--A--C--AGCT-A-A-A-G-CAT-T-A--A--A--A--AA-CC-CC-GTCAT-AC-A--G	7914
GSN.CM.99.CN71	TC--G--G--G--G--G--C--G--G--A--C--AGCT-A-A-A-G-AT-T-A--A--A--A--A--CC-CC-GTCTT-AC-A--G	7926
DRL.-.-.FAO	GCT--GC--A--G--G--G--A--GC-AC--C--CCTGG-G--ACA--AG-AT--G--G--G--CA--G--TA--G--A--A--C--CC-CC--ACA--C--T--GG	7818
RCM.GA.-.GAB1	---T--GC-TG-A-G--T--A--G--CC--GACATAG--A--CG--G--A--C--AA-G--C--A--CT-AA-T--GC-TG--ACT--A--T--G	7664
RCM.NG.-.NG411	---T--GC-TG--G--T--G--G--G--GACATAG--ACA--G--G--A--C--AA-G--C--A--CT-AA-C--GC-TG--ACT--A--T--G	7708
MND-2.-.-.5440	ACT--C--AA--G--T--G--G--G--GGT-C--CCTG--A--TCA--CTCCG-GT-A-AA--A--TA--TG-A--A--T--T--TCC-CC--ACA--CT--GG	7760
MND-2.CM.98.CM16	GCT-G--AA--G--T--G--G--G--GGT-C--CCTG--A--TCA--CTCCG-GT-A-AA--A--TA--TG-A--A--T--T--TCC-CC--ACA--CT--GG	8127
MND-2.GA.-.M14	GCT--G--AA--G--T--G--G--G--GGT-C--CCTG--A--TCA--CTCCG-GT-A-AA--A--TA--TG-A--AA--T--T--TCC-CC--ACA--CT--GG	8056
MNE.US.-.MNE027	ACT-----G-----G-----G-----A-----C-----AC-----GC-----T-----GAC-----TGG-----CA-----AGA-----AG-----AT-----G-----G-----C-----A-----CA-----A-----C-----A-----T-----G-----ACT-----CA-----C-----G-----	7872
LST.CD.88.447	GT-G--A-ACAA-G--T--G--G--G--GT-CT-G--AAA-G--AG-G--C--AAG--CA--AG-G--A-T--A--C-CC-GACA--A-T--GG	7157
LST.CD.88.485	GT-G--A-ACAA-G--T--G--G--G--GT-CT-G--AAA-G--AG-G--C--AAG--CA--AG-G--A-T--A--C-CC-GACA--A-T--GG	7154
LST.CD.88.524	GT-G--A-ACAA-G--T--G--G--G--GT-CT-G--AAA-G--AG-G--C--AAG--CA--AG-G--A-T--A--C-CC-GACA--A-T--GG	7160
LST.KE.-.lho7	GC-G-C--CAA-G--T--G--G--G--GT-CT-G--AAA-G--AG-A--AAG--CA--A--AG-G--A-T--A--C-CC-GACA--CC-T--GG	8232
SUN.GA.98.L14	GC-G-C--CAA-G--T--G--G--G--GT-CT-G--AAA-G--AG-G--AC-AAG--CA--A--AG-G--A-T--A--C-CC-GACA--CC-A--GG	8300
MND-1.GA.-.MNDGB1	TCT--GG--A--G--G--G--G--A--A--AC-G--T--A--CTC--A--CAA--GTCTG--A--C--AA--G--A--A--TT--A--A--TC-CC-GACCAG--T--G	7584
MON.CM.99.L1	TC--C--G--G--G--G--G--CC--CG-AACA--C--AGCT-AC--A--G--T--T--T--AG--A--A--A--CC-CC-GAC--A--G	7940
MON.NG.-.NG1	TCT--G--G--G--G--G--C--CG-AACT--C--G--AGCT--A--A--G--T--T--A--AT--A--CC-CC-GAC--A--G	6565
MUS.CM.01.1085	TCT--G--G--G--C--G--G--G--GCA--C--A--CG--A--A--C--A--GT--GT--T--T--T--A--A--A--TC-CC-GACA--GC--T--G	7911
DEB.CM.99.CM40	TCT--GC--AG--G--G--G--A--C--AG-GC--T--A--G--G--T--CGG--G--AAC--T--A--C--AG-G--GAA--C--A--CA--CC-CC-GACT--A--G	7717
DEB.CM.99.CM5	TCT--GC--AG--G--G--G--A--C--AG-GC--T--A--G--G--T--CGG--G--AAC--T--A--C--AG-G--GAA--C--A--CA--CC-CC-GACT--A--G	7638
SYK.KE.-.KE51	ACT--G--G--G--G--G--C--A--G--G--GA--G--A--A--C--C--AGGCT--A--G--TG--G--AA--C--TA--C--CC--C--ACT--C--A--G	7637
SYK.KE.-.SYK173	ACT--G--G--G--G--G--C--A--G--G--GA--AG--A--A--G--A--C--C--AGGCT--A--T--TG--AA--C--TA--C--CC--GC--ACA--T--A--G	7944



Accession	Sequence	Position
H1B.FR.83.HXB2	GATACCTAAAGGATCAACAGCTCCTGGGGATTGGGGTTGCTCTGGAAAACCTATTGGACCACCTGCTGTGCTTGGAAATGCTAGTTGGAGTAATAAATCT	8078
Env	R_Y_L_K_D_Q_Q_L_L_G_I_W_G_C_S_G_K_L_I_C_T_T_A_V_P_W_N_A_S_W_S_N_K_S	gp41
H1A1.UG.85.U455	-----C-----A-A-----G-----C-----A-----C-----CT-----	7508
H1B.US.90.WEAU160	-----G-----G-----A-----C-----C-----A-----G-----	8085
H1C.ET.86.ETH2220	--C-----G-----A-----C-----C-----C-----CT-----G-----	7454
H1D.CD.84.84ZR085	-----G-----A-A-----A-----A-----A-----C-----CT-----G-----	7599
H1F1.BE.93.VI850	-----A-----C-----C-----C-----AA-----C-----CT-----	7355
H1G.SE.93.SE6165	-----A-----C-----C-----C-----AA-----C-----C-T-----	7496
H1H.CF.90.056	-----G-----A-----C-----C-----AA-----C-----CT-----A-----	7397
H1J.SE.93.SE7887	-----A-----A-----C-----C-----AA-----C-----CT-----	7385
H1K.CM.96.MP535	-----A-----A-----C-----C-----AA-----C-----CT-----C-----TC	7237
H101.AE.TH.90.CM240	-----A-T-----A-AC-----C-----A-----C-----C-----CT-C-C-----G-----	7631
H102.AG.NG.-.IBNG	-----GA-----A-----C-----C-----A-----C-----CT-----C-----A-----	7584
H1N.CM.95.YBF30	-----T-G-----G-AA-----AA-TC-A-----C-----A-----ACA-A-----TAT-CA-----AG-C-----C-C-TA-CTCT	7629
H1O.BE.87.ANT70	CC-TA--C-A--G-A--AA-CC-A--C-TAAA--G-AG-C--TA--AT-A-AAAA--AGA-CA--TAGGA	8126
H1O.CM.91.MVP5180	CCCTTA--C--A--G--A--G--AAACC-A--C-TAAA--A--C--TTA--AT-A-AAAA--CA-ATCA--TCAGGA	8151
CPZ.CD.-.ANT	AG-----GA-----A-AT--A-CC-C-----TG--AC--GG-G-CC--TCA--A-G-----AA-TCC--GTA--CTTCA-GCAAACATGT...GCAAAGAACAGCAGTGA	7536
CPZ.CM.-.CAM3	-----C-GA-----G-AA--A-CC-A-----C-----A-----GC-----TTAT--A-----AA-C-----GC--T-----ACATC	7493
CPZ.GA.88.GAB2	-----T-----G--AG--A-CAC-A-----C-----ACTG-G-----TA-T-A-----A-C-----A-TCC--AC-G--CAAA	7464
CPZ.GA.-.CPZGAB	-G-----GC-----G--A-T-----CC-G-----C-----GGCTG-C--TTAT-CA-G-----CAACTC--CC-GGG...AGC	8117
CPZ.US.85.CPZUS	-----T-----G--A--A-TC-G-----C-----A-----ACA--TTAT-CA-----A-CC-----C-C-----CTCTC	8085
CPZ.TZ.-.TAN1	A--A--GA-C-----AA-CC-C-----A-TG-AAC--T-GG-G--TCA-G-AG-----A--CCTC-CC--GC-G-AG-T--ACAAAGTGAATCAGTGCAGTGCAGTGA	7733
H2A.GW.-.ALI	A--T-----C--GGC--GG--AAATTCA-----A-TG-ATTT-G--AAG-C--CA--A--A-A--GTA--AC-CC	8523
H2A.DE.-.BEN	AG-----C-----GGCA-AG--AAATTCA-----A-TG-GTTT-G--AGG-C--CA--A--A-G--GTAA--GAC-CT	8539
H2A.SN.-.ST	A--T-----C--GGC--AA--AAATTCA-----A-TG-GTTT-G--AAG-C--CA--A--A-A--GTA--GACACC	7969
H2B.GH.86.D205	A-----GGCA-AA--AAATTCA-----A-TG-TTC-G--AGG-C--TCA--GA--A--A--CCAAA-GAAACC	8515
H2B.CI.-.EHO	A--T--C--A--C--GGCA-AA--AAATTCA-----A-TG-TTC-G--AGG-C--TCA--GA--A--A--GTA--GAA-CC	8501
H2G.CI.-.ABT96	A--T-----C--GGCA-GG--AAATTCA-----A-TG-GTTT-G--AGG-C--CA--A--A--CCT-G-TGC--A-A	7912
H2U.FR.96.12034	A-----GGCAAGT--AAATGCA-----A-TG-CTTT-G--AGG-C--CAT--A--A--TAAACGACACA	8023
MAC.US.-.239	AG--T-----C--GGC--AG--AATGC-----A-TG-GTTT-G--AAG-C--CA--A--A--CCAAA-GCAA-T	8484
Env	K_Y_L_K_D_Q_G_Q_L_N_A_W_G_C_A_F_R_Q_V_C_H_T_A_T_V_P_W_P_N_A_S	gp41
SMM.SL.92.SL92B	A-----GGCA-AG--AAATTCA-----A-TG-ATTT-GG-AAG-C--CA--A--A-TA--CCAAA-GACA-T	7932
SMM.US.-.H9	AG-----GGC--CRN-AAATTCA-----A-TG-TTT-GG-AGG-C--CA--A-K-A-A--CCAAA-GACACA	7973
STM.US.-.STM	A-----GGCA-AG--AAATTCA-----A-TG-GTTT-G--AGG-C--CA--CA--A--A--CCAAA-GA--CT	8135
SAB.SN.-.SAB1C	A-----G-----GGCA-GGT--AAC-----A-TG-TTC-GG-AGG-G--TCA--GA-G-C-T-----GTA--A-AAC	8367
TAN.UG.-.TAN1	A-----G-----GACA-GG--AAATTG-----A-TG-ATTC--AGG-G--TCA--A-G--G--C-TTC-ACAAT	8171
VER.KE.-.AGM155	AG-----GGC-----GGCA-GGT-AAACGC-----G--GT-G--AAG-A--TCAT--AA-A-A--C--CGTGG-A-AAT	8194
VER.KE.-.9063	AG-----G-----GGCA-GAT--AACG-C-----G-AT-G--AGG-C--TCAT--CA-A-A--C--ATGGCAAAC	8214
VER.DE.-.AGM3	AG-----G-----C--GGC--GGT--AATGC-----G--AT-G--AAG-C--TCAT--AA-C-A--G--C-GTGG-A-AAT	7712
VER.KE.-.TY01	AG-----G-----GGCA-GA--AAACTCC-----G-AT-G--AAG-A--TCAT--CA-A--GAG--CCCTGG-CAAA	7683
COL.CM.-.CGU1	-T--TT-GG--A--GCAAAG--T-CC-GCATT--G-TG-AAAC-TG-AA--T-GA--AT--C--CAAG-CC	7478
DEN.CD.-.CD1	A--T--T--A--C--GCA--G--AAATGAA-----A-TG-CT-G--G-AA--TCA--CT-A--AGAA--GCAA-CAA-CA-TACC-GG--TG	8140
GRV.ET.-.GRI_677	AG-----G-----GGCA-GG--AAATTCA-----A-TG-GT-G--AAG-A--TCA--CA-A--A--GTA--A-AAC	8073
GSN.CM.99.CN166	A--TT-G-G-----ACCA-T--CCAAGCA-----C-TG-GAAC-GG-CA--C--TCA--CATA--A--CA-CTCA--GC--GGC--G	8015
GSN.CM.99.CN71	A-----G-G-----ACAA-T--CCAAGC-----C-TG-GAATC-G-CA--C--TCA--CATA--A--ACTCA--GC--G--C--G	8027
DRL.-.-.FAO	A--A--GGA-C--GCAA-GT--A-CC-C-----A--CTTGC--AAG-A--CAT--AAA--GTA--CCAA--GAG-CA	7907
RCM.GA.-.GAB1	A-----GCA--C--TCT--AT--AAT-CA-TT--G-TG-CT-G-G--AAG-C--CAT--A-TG--G--CATTC	7749
RCM.NG.-.NG411	A-----GC--C--TCT--AT--AAT-CA-TT--C-TG-AT-G-G--AAG-G--TCA--AA-G--GAG--TATA--AG	7793
MND-2.-.-.5440	A--G--GC--C--TCAAGG--CTC-G-A--A--T-ATTCCTC--AAG--TCAT--AG--AAAA--CCAA--A-A-T	7849
MND-2.CM.98.CM16	AC--TG--GCT--GCAA--ATCC--G--C--ATTTCGC--AGG-A--CA--ACA--A--A--CCAA--GAC-CT	8216
MND-2.GA.-.M14	AC--TG--GC--C--TCAAGA--A-CAG-A--A--ATTTCGC--AAG-G--TCAT--AA--T--G--CCAA--GAA--T	8145
MNE.US.-.MNE027	AG--T-----C--GGC--AG--AAATGC-----A-TG-ATTT-G--AAG-C--CAT--A--A--A--CCAAA-GCAA-T	7961
LST.CD.88.447	AG-----C-----ACT--G--T-CTTCC-----G-TCAAT-G--AAG-A--TCAT--AA--AGAA--CTA--A	7243
LST.CD.88.485	AG-----C-----GCT--G--T-CTTCC-----G-TCAAT-G--AA--A--TCAT--AA--GAA--TA--A	7240
LST.CD.88.524	AG--T-G--A--C--GCA--A--T-CCTC-----G-TCAGT-G--AAG-G--TCAT--AA--A--A--TA--A	7246
LST.KE.-.lho7	AG--TG-G--AC-----GC-----T-CCTC-----G-TCAAT-G--AGG--TCAT--AA--GAG--CATA--A	8318
SUN.GA.98.L14	AG--TG-GC-A--C--AGC--TT-A-CTTC-----G-TCAAT-G--AAG-A--CAT--AA--C--TA--A	8386
MND-1.GA.-.MNDGB1	AT--TA-C-----C--GCTT-G--TCTCAA-----G-T--AT-GGC--AGG-G--TCAT--T--AGAG--C-AA--CAA-C	7673
MON.CM.99.L1	AG-T-A-----C--ACAT-G--AAATGCA-----A-TG-CAAT--GC-G-G--TCA--AA-A--G--CAA-TC--GCA--AGGTCAC	8041
MON.NG.-.NG1	AG-T-A-----C--ACAT-GT-AAATTCA-----A-TG-CAATCGCGCAG-R--CA--ACAA--TA--CAA--CG--GC--AGGCCAC	6666
MUS.CM.01.1085	AG-T-A-----C--GCC--T--AATC-C-----A-TG-AAAC-GG-AG--TCA--AGA--A--A-TCA--GCA--C-CA-C	8012
DEB.CM.99.CM40	AG--T--A--C--GCAAAAT-AAATGAA-----C--G-CTTC--AG--C--TCAT--CA-G--C-----AACCTC-CT-GAG	7809
DEB.CM.99.CM5	AG-----A--GAGCAAAGT-AAATGAA-----A-TG-CTTC--AA--C--TCA--A-G--C--CAA--CA--G-A	7730
SYK.KE.-.KE51	CG--T--C-GA--C--GCCA-TA--TCA--A--G-ATTT--G-AA--TCA--A--A--AA-G--C-GCAAGCA--TG-C--C--CGT	7743
SYK.KE.-.SYK173	CG--T--C-GA--C--GGCCA--T--TCA--A--G-ATTC--G-AG--C--CAT--A--A--A-G--G-GAAAGCA--TG-A--C--C--CAAC	8053

PLV  
complete genomes

Table with columns for sample ID, sequence, and genome size. Includes samples like H1B.FR.83.HXB2, H1A1.UG.85.U455, H1B.US.90.WEAU160, etc.





PLV complete genomes

Table with columns for sample ID (e.g., H1B.FR.83.HXB2), sequence alignment (I V N R V R Q G Y S P...), and accession number (8443, 8450, etc.). The table lists various PLV complete genomes and their corresponding sequence data.





H1B.FR.83.HXB2	AGACAGAGACAGATCCATTTCGATTAGTGAACGGATCCTTGGCACTTATC.....TGGGACGATCTGCGGAGCCTGTGCCTTTCAGCTACCACCGCTTGAGAGACTTACTCTTG	8552
Env	D R D R S I R L V N G S L A L I ..... W D D L R S L C L F S Y H R L R D L L L	gp41
Rev	R Q R Q I H S I S E R I L G T Y ..... L G R S A E P V P L Q L P P L E R L T L	Rev ex2
H1A1.UG.85.U455	--G-A-----G-----C-----G-----T--A-GA-GC-----C-----A-----A-----CGC----	7982
H1B.US.90.WEAU160	-----GGAA-----G-T-----T--AA-----T-C-A-----CT-----TC-----A-----	8559
H1C.ET.86.ETH2220	--G-----A-C-----G-T-----T--A-A-T-T-----C-----T-----A-----A-----	7928
H1D.CD.84.84ZR085	-----GT-C-----G-----T-----CA-----A-----A-----A-----A-----A-----	8073
H1F1.BE.93.VI850	--G-A-----G-G-C-----C-----T--A-T--GC-----A-----A-----G-A-----CA-A--A	7829
H1G.SE.93.SE6165	--G-----G-----AG-G-----G-----T--AC-G-----C-----C-----CA-----CCA----	7970
H1H.CF.90.056	-----G-GA-----T-----AC-G-G-----C-C-----CA-----G-T-----A-----	7871
H1J.SE.93.SE7887	--G---AC-----G-----T-C-A-T--GC-----C-----A-----T-----A-----CG----	7859
H1K.CM.96.MP535	--A-A-----AG-----G-C-T--A-G--GC-----C-----A-----A-----G-AA--A--A--A	7712
H101.AE.TH.90.CM240	--G-----AG-G-----G-----T--A-----GC-----A-----A-----G-----AC----	8105
H102.AG.NG.-.IBNG	-----A-----G-G-----G-----T--A-----GC-----A-----A-----A-----A-----	8058
H1N.CM.95.YBF30	-----G-----G-G-----G-----T--CA-T--G-----G-C-C--A-----TGA---CT-----C-----C-----	8103
H1O.BE.87.ANT70	--G---GCC--G-GG--A-CC-CGCC-C-A-G-T-----C---GT-GTACACGGACCTCAGAACATAATATTGTGGACTTACCA--C-TGAG-AA-TTAGCATCAG-GA-CCAGA-GG-GA--AGC	8621
H1O.CM.91.MVP5180	-----GCC--AG-GG-CAGCC--GCCACCA--T-----CA--AGT-G-----ACAGC--CA--CAA-AAT-T-G-GG-CT-----T-----CA-----A-A-CA	8626
CPZ.CD.-.ANT	-----GAT--G-GG-GGGCC--GCA-C---G-T---C---CT-G-----TG-C--AC--A-AAT--AG-GG-T-----GAT--GC--AC--GT-----	8011
CPZ.CM.-.CAM3	--G-A-G-----C-----G-----T--A-T-----A-----C-----A-----CCTG-----T-T-----A-----A-----	7968
CPZ.GA.88.GAB2	-----G-----TG-GA-----C-G-----TT--A-T--CTG-----CA-C--A--ACGA---CT-----G-T-----T-GT---C-	7939
CPZ.GA.-.CPZGAB	-----AG-----AG-GA-----G-A-G-G---C-----C-C-A-A--G-GA-T-GG-----GA-----C-AG---GC-GC	8592
CPZ.US.85.CPZUS	-----AC-T-----A-CGA-----C-G-----T--A-G-G-----A-----T-----A-----CCTGA---CT-----A-----GAG-----A-A	8560
CPZ.TZ.-.TAN1	-----ACAT--G-GG-CG-CC-CGCCAGCA--T-----CAGTA-G-----G-C-CA--A--CCT-A--GG-TA-----GAC---TCA-A--CA--G-	8208
H2A.GW.-.ALI	-----AC-TTG-GGA--GATTG-GGCCCTGGCCGAT-GCATATT-ACATTTCTGATCCACTGTAGTCT-CCT-TTGA-CG--CTGTACA--A--G--GG--TTA-TA-CC--GAT--CC-CG...	9030
H2A.DE.-.BEN	--G-GACAGTG-T-A-GACTTG-GGCCTTGGCC-ATAAACTATG-GCAGTTCCTGATCCACTACTGACT-CCT-TTGA-CG--CTATACAA-A--G--AG--TTA-TA-CC-AGA--CC-CG...	9046
H2A.SN.-.ST	-A---C-TTG-GA--A-T-G-GGCCCTGGCCGATAGAATATA-ACATTTCTGATCCGCGAGTCTGATTC-CCT-TTGAACA-CTATACAA-A--G--GG--TTA-TA-CC--GAG--C-AG...	8476
H2B.GH.86.D205	-A--GAC-GGGAC-A--GATCC-GGCCCTGGCCAGATAGAATATA--CATTTCTACTGCGCCAGCTGAGGAACCT-TTAA-TT--CTATACAA-GG--G--AAC-TTA-TG---AGAC---C-AA...	9022
H2B.CI.-.EHO	-A---AC--GG-C-A--GATCT-GGCCCTGGCCAGAT--GAATACA-CCA-TTCCCGATTGCGCAACTGAGGGACCT-TTGA-TT--CTATACAA-GG--G--AAC-TTATTA-CA-AGAC---C-AA...	9008
H2G.CI.-.ABT96	-A---A--GTG-C-A--GATCT-GGCCCTGGCCAGATAGAATATA--CATTTCTGMTCCGCCAGCTAGGGAACT-TTGACTT--CTATACA-AA--G--AGC-TTA-TA-CC--GAT-Y-C-AG...	8419
H2U.FR.96.12034	-----C-GTG-C-A--GATCT-GGCCCTGGCCAGATAGAATATA-ACATTTCTGATCCCAACTGATTC-CCTGTTGACTT--CTATACA-AG--G--GG--TTG-TA-CC--GATT-G-AG...	8530
MAC.US.-.239	---AG-C-GTG-CAA--GCTCC-GGCCCTGGCCAGATAGAATATA--CATTTCTGATCCGCCAAGCTGATAC-CCT-TTGACTT--CTAT-CA-AA--G--AAC-TTG-TA-C-----TA-AC-AG...	8991
Env	E G G G N S S W P W O I E Y I H F L I R Q L I R L L G W L F S N C R T L L S R V Y Q	gp41
Rev	R R R W Q Q L L A L A L A D R I Y S F P D P P T D G P L D L G I Q Q L Q N L A I E S I Y P D	Rev ex2
Tat	E K G V G T A P G L G R S /	
Tat end /		
SMM.SL.92.SL92B	-A-G--GTG--A--GATTG-GGCCTTGGCCAGATAGAATATA--CATTTCTGATCCGCGAGTGTATGAC--ATATGACTT--CTATACAA-AAT--G-CTCG-TTAGCA-CC---C--AC-AG...	8439
SMM.US.-.H9	--R-----GTG-CAA--GATCT-GGCCTTRGCAGATAGAATATA--CATTTCTGATCCGCGAGTGTATGAC-CCT-TTGACTT--CTAT-CA--AGT-G--GG-TTGG-TA---GAN--GC-AA...	8480
STM.US.-.STM	-----GTG--AT--ACTCC-GGCCTTGGCA-ATAGAATATA-CATTTCTGATCCGCGAGTGTATGAC-CCT-TTGACTT--CTGTACAA-AA-----GGCTTGT-TA-CC--GAT--AC-AG...	8642
SAB.SN.-.SAB1C	GCTA-ACAG-GACAG--GGTC--ACACTTGGCAGAGAGAATTCT-G-GACACCTGTGTACAGATTGATCACTTGGCTCAG-AATTGACCA-TTGG--TCGACAAT-TT-AGC-ACCT-CAT-G...	8868
TAN.UG.-.TAN1	CAG--CAG---GA--AGTCGCCGCCCTCA-TGAAGAAT--TT-AT-GCCCAAAGAGGAGGGATCCAAGC--G--GAGAGCTT-G-G-CAG-AT--G-C-A-TTGGT-----C-AT-AGCAGT-G-	8657
VER.KE.-.AGM155	-TCT-CGC-TTAGCAG-AAGA--T-G-GGAC--GA-AA-AGGACC-GTTGGTGCAGGCAATTGACCAAT--TT-TTGACA-TC--A-C--GT-ACA--A-AG-TG--TGACCTCCTTC-CAAG...	8695
VER.KE.-.9063	CTCAGATTCTT-GCAG-AAG--CT-GA-CGA--CAAAA-A-G-GC-TTGGAGCAAACGCTTGACCAAT--CT-TGCAC-TTC--ACC--G--A-A--A-AC-TT--TGACCTCCTAA-T-AT...	8718
VER.DE.-.AGM3	CTC-GAGCCTT-GCAG-AAGA--CT-GC-CA-C-GAG-G-AAGAGC-A-TGGTGAAGCGATTGACCAAT--TG-TCGA-CA-C--A-C-G--A-A--A-AGTTG--TGACCTCCTAG-T-AT...	8216
VER.KE.-.TV01	-TTAGAGTCTT-G-AG-TVAGAC-CCAA--G-A-G-GTA--CAGT-G-C-GCCTGGCTGACGAGGCTCAAC.....ACT--G-A-A--A-AGTTG--TGACCTCCT-A-T-AG...	8172
COL.CM.-.CGU1	---A-ACAG--CAG--GATTT-GGCGATTGATGGAAGA-TGGTGG-G-----	7911
DEN.CD.-.CD1	.-G--ACAGGCTCAAGTGGT-GCA-TATCCGA--GGA--CTTGAG--ATT.....A-----CTCACAG--CTGACG-GG--T---TGAC-----A-----G--	8609
GRV.ET.-.GRI_677	GATA---TTGCAGGAAGATGCC-C-ACACCTTCG...AG-ATCAAC-A-TGGTGGCTCAACTTCAAGAGC--CAG-TTGAGAATA--GACT--GTGT-A--A-AT-TG--TGACCTCCT-A--T--...	8571
GSN.CM.99.CN166	--G--AG-T--G--G-GA--ATC-A-CA--T--C-AGT--CT-G.....C--A--A--GTGA-T-GG-T-----GAT---GCCAG---GC--G-	8486
GSN.CM.99.CN71	--G--AC-T--G--G-CA--ATC-C-CA--T--C-AGT--CT-G.....C--A--T--GTAA-T-GG-T-----GAT---GCCAC---G--G-	8498
DRL.-.-.FAO	--G-GAGTGG...AA---CA---C--AACC-TCC.....AAGCCAGGATCCTCCAAGCTTGGAGCGAA--AACAG-AG--C-TG--TGAAGGAATCGAGAGGATA-A-CT-GCTGAAGAA-G--	8402
RCM.GA.-.GAB1	T-GAGA--AG-TGG--ACAAG-ACAAC--TT-GCTGAGA-A-TA-TG.....ATT-C-A--ATCCA--C--TGAG-AGG-TT-GGACG-AGC-ATCGC-GA-TTG-AGA	8230
RCM.NG.-.NG411	T-GAGA--G-G-TGGACACAA--ATTAC--TT-GCAGAGA-G-TA-T-T.....ATAT-CTGA-A-TCC--C-ATAGAGA--CTT-G-AGGA-GC-CTATC-AA-CTG-AGC	8280
MND-2.-.5440	---A-A-C-G...AA---CAAG-C--A--G-TCC.....AGACAAGAACTTAGGAAGCCTTGGGAAGCCA--GC-AG-CA----TG-CTGAAAAGGTCGACG-TTTA-A-TT-GCTGAAGAA-C-C	8341
MND-2.CM.98.CM16	---A-A-C-G...AGT-CCAA--C--AC-G-TCC.....AAG-AGAATTTGGGAAGCCTTGGAGCAA--GC-AA-CA-AGATTG-CTGAAGACTTCAAGGGG-TA-G--T-GCTGAAGAAATC-A	8705
MND-2.GA.-.M14	---AGA-C-G...AA---CAAG-C--A--G-TCC.....AAG-AGAATTTGGCAGACCTTGGAGGCCA--GC-AA-AA--ATTG--TGAAGGAGTCAACATTTA-A-TT-GCTGAAGAAATC--	8634
MNE.US.-.MNE027	-----C-GTG-CAA--G-TCC-GGCCTTGGCCAGATAGA-TATA--CATTTCTGATCCGCCAAGCTGATAC-CCT-TTGACTT--CTAT-CA-AA--G--AAC-TTG-TA-C-----CA-AC-AG...	8468
LST.CD.88.447	---A-ACAT...-A---CAAC---AAC-GTACAAGAAA-A-TC-TC-CTCCACCCTGGAACCTGGAC--AG--G-CTTTACAGAC-CA--G--TGTGACT-T--TGAA-T-GAT-AAGGC--T	7750
LST.CD.88.485	---AGACAT...-A---CAAC---AAC-GT-CAAGAAA-A-TC-TC-ATTCCACCCTGGAACCTGGAC--AG--G-CTTTACAGAC-CA--G--TGTGACT-T--TGAA-T-GAT-AAGGCA--T	7747
LST.CD.88.524	---A-ACAT...-A---CAAC---A-C-GT-CAAGAAA-A-TC-T--AGGCCACCCTGGAACCTGGAC--A--G-CTTTACAGAC-CA--G--TGTGACT-T--TGAA-T-GCTGAAGGC-G-	7753
LST.KE.-.lho7	---G-CTAT...-A---CAAC---A-C-GT-CAAGAAA-A-TC-TC-AGGCCACTGTGGAACCTGGAC--A--G-CTTTACAGAC-CA--G--AGTGA-T--TGAA-T-GCT-AAGGAAGG-	8825
SUN.GA.98.L14	GA-AGA-AGA-TC-A---CAAC---A-C-GTCAAGAAA-A-TC-T-GCCACCCTTGGACAGTAGAC-----G-CTTT--AGAC-CT--G--AGTGA-TT-T-AA-T-GCTGAA-GCT-G-	8887
MND-1.GA.-.MNDGB1	---ACT--GA...GAGCACAAC-CC-A---TTG.....AAGGAAGAATCCTTAAACAGATCCTTGATA--G-C--ACC--T-TGC-AGGGAGT-ACCAATTT-GC-TT-GCTGAA-AATC--	8153
MON.CM.99.L1	CAG--ACTC---GCCAGC--ATC---A---TT--TA-CA-----A-----CA--A--TGT---GG-CA-----GAT-----A-----C-G-A--A	8512
MON.NG.-.NG1	--G--GA---G--R-CAGT-AT-C--GA---TT--AA-CA-----T-----A-----CA-----CTG-AGA--AATAG--GG-C--T---G-G-TA-----CCGG--	7137
MUS.CM.01.1085	-AG-----G-GA--ATC--G--T--CT-T--CT-----C-C-C-A--AGTGA-T-GG-----GAC--A-T-AG--C-GC--CC	8483
DEB.CM.99.CM40	C-G---G-C---G-GG-CAGC--GCC--GA---T--CAGT--C-----G-----C-AACAA--CTGACG-GGCT-----GAC--TC-CA--GCA-T-G-	8274
DEB.CM.99.CM5	C-T--AGAG--G-GG-CGAC--CGCC--GA---T--AGG--C-----CA-CTCAGC--TCTGACG-GGCT-----AAC-----GA--CGCC--	8195
SYK.KE.-.KE51	TT--AGCC---CGT-TC-ACG--CAGC-GA-A--TT--AG--AATCGCTGCGAGAAGCGCAGAACTT--A--AGAG-TTC-GAGAGC--AGA-CACTGATTAGAG-AC-TGT-ATTGCCTGG...	8242
SYK.KE.-.SYK173	---AG--CT--ACGT-TC-ACC--CAGC-GA-A--T--AGG-AATCTTTGGAAGCAGGACAGCAGCTT--AGGACAG-CT-C--AGC-T-AGA-CTCTGATTAGA-AGC-T-CTAT-ACCTGG...	8543



PLV  
complete genomes



H1B.FR.83.HXB2	ATTGTAACGAGGATTGTGGAACCTCTGGGA.....CGCAGG...GGGTGG.....	8594
Env	I V T R I V E L L G.....R R G W.....	gp41
Rev	D C N E D C G T S G.....T Q G V.....	Rev ex2
H1A1.UG.85.U455	-----G-----AGC-----CGCAGCAGCCTCAAGGGACTGA-ACT-----	8045
H1B.US.90.WEAU160	-----C--A-----	8601
H1C.ET.86.ETH2220	-----C-G-----CA-----CGCAGCAGTCTCAAGGGACTA-AG-----	7991
H1D.CD.84.84ZR085	-----C-G-----	8115
H1F1.BE.93.VI850	-----C-G-----CAGGGGCTG.....A-G-----	7871
H1G.SE.93.SE6165	-----G-----CA-----CGCAGCAGCCTCAAGGGACTGA-ACT-----	8033
H1H.CF.90.056	-----GT-----C-----G-----A-A-----A-----	7913
H1J.SE.93.SE7887	-----C-G-----C-----G-AC-----T-----	7901
H1K.CM.96.MP535	-----G-----CC-----AGGGGACTG.....A-GG-----T-----	7754
H101.AE.TH.90.CM240	-----C-G-----CG-----CACAGCAGTCTCAAGGGACTGA-AC-----	8168
H102.AG.NG.-.IBNG	-----C-G-----C-----CACAACTGTCTCAAGGGACTGA-ACT-----	8121
H1N.CM.95.YBF30	-----AC-G-G-----C-C-----C-----CAGAGTCTCAGCAGGGGACTG-AACTA...CT-AAT...	8166
H1O.BE.87.ANT70	TA-C-G-G-CTTGAC--TGGA-C--A--G.....CAGAAGATAAATAATGTT.....TGCAGAATT...	8678
H1O.CM.91.MVP5180	GGGA-CCG---C-GA-C--CTAC.....CTGGGACTGTGGATCCTGGGA-AA-A...ACAATT...GAAGCTGTAGACTT...	8704
CPZ.CD.-.ANT	-ACC-TTG-GCAG-CC-CC--AC--C.....TGCAGAATT.....	8047
CPZ.CM.-.CAM3	-----AC-GGG-----CAT-----AAC-----CAGAGCCTCAACAAGGGACTG-AACA...CT-A--	8031
CPZ.GA.88.GAB2	G-CC-TTG---CC-GC-----AAT-----CAACAGGTCCTGAGAGGACTGA-ACTA...CTCA--	8002
CPZ.GA.-.CPZGAB	-A---GTG---ACAAC-CA-GACGT-----CATCTAATA-TTCAC...A-CCTC...AGAC	8647
CPZ.US.85.CPZUS	-----C-G-G-----CA--AC--A-C-----CAGAACATCAACAAGGGACTG...CAA...CT-T-...A	8621
CPZ.TZ.-.TAN1	-----CC-CTG--TC-GCC--C--GCA--AA-----CAGGGATAATCAGCTTGGCA...CAGAGCCTAGTA...ATAGTGCATAGAACTA	8287
H2A.GW.-.ALI	.....ATCCTCCAACCGATCTTCCGGAGTCTTCAGAGAGCGCTGACA-CA...ATCA--	9081
H2A.DE.-.BEN	.....ACCCGCGACTGATCTCCAGAGTCTAACAGCA...ATCA--	9085
H2A.SN.-.ST	.....ACCTCCAACCTAATCTCCAGAGTCTTCGGAGAGGACTGACAGCA...-TCA-A-	8527
H2B.GH.86.D205	.....ATCCTCCACCAAACTCAACCAATCTCCAACCACTCAGGCTCCCG...	9067
H2B.CI.-.EHO	.....ACCTCCAACCGACTCTCCAACCGCTCAGGCTTCCA...	9044
H2G.CI.-.ABT96	.....ACTCTCCAACCGCTTCCAAGAACTCTCAGGACCTC-AAGCA...ATTA-A-	8470
H2U.FR.96.12034	.....AGCCTCCAGCCGTTGTTCAGAGCATC.....A-A-	8560
MAC.US.-.239	.....ATCCTCCAACCAATCTCCAGAGGCTCTTCGCGACCTA-AG--...ATTC-A-	9042
Env	I L Q P I L Q R L S A T L Q R.....I R.....	gp41
Rev	-----P P T N T P E A L C D P T E-----D S-----	Rev ex2
SMM.SL.92.SL92B	.....AACCTCCAACAACCTGTGCCAGAGACTCTCAGAGATCAGT-AGCC--...ATTA-A-	8490
SMM.US.-.H9	.....ATCCTCCAGCCAGTGTCCAGAGCCTCTCAAGGAGCGTG-AG--A...-CCC-T-	8531
STM.US.-.STM	.....ACCCTCCACCAACCTTCCAAGGATCTCCAGGATACTG-AG--A...ATTA-A-	8693
SAB.SN.-.SAB1C	.....TGCCTC-AGGAC...ATCCAA...	8886
TAN.UG.-.TAN1	T-AT-G-GACTTTACCA-ATC--C-GCA-G.....AGCCTCAG-CT...CTTCTT...	8705
VER.KE.-.AGM155	.....	8695
VER.KE.-.9063	.....	8718
VER.DE.-.AGM3	.....	8216
VER.KE.-.TYO1	.....	8172
COL.CM.-.CGU1	.....AGCTTGTTCGCTCATTGCAGCAATGGGAT-CTT...ACTGTC...	7950
DEN.CD.-.CD1	T--C-G-TT---CC.....GCCTGTA-TC--ATTT--GAC.....	8645
GRV.ET.-.GRI_677	.....	8571
GSN.CM.99.CN166	-C-C-CCGTG---C.....-G.....AAGTTCCTGT-GGAACAGCT-CA-AAAGCAATCCAGAAAGCAATC...CAGCAAG	8556
GSN.CM.99.CN71	-C-C-CCA---C.....-C.....AAGTTCCTGT-GGA-CAGCT...CAGAAGGCGTTC...CAACAAG	8556
DRL.-.-.FAO	-AA-C-GTT.....CTTGAATATGGGT--CA.....GAACTCCAAG	8439
RCM.GA.-.GAB1	-GCTGC-GCTCA--AA-CTTC-AGAGC.....CTCGGTGATTTTGTAGTAAGATACAATAT-----	8293
RCM.NG.-.NG411	-ACTC-CTCTCAG-GA-CT-C-AGAGCCTCCAGTCAATCCCTTCTCGAATCCTTCTTATTGCGGTGGATAGGAGTAAG-TACAAGAT-----	8376
MND-2.-.-.5440	CAA-C-GTA.....ATTGAATATGGGT--CAA...GAGCTCAAAG	8378
MND-2.CM.98.CM16	CAA-C-GTA.....ATTGAATATGGGT--CAA...GAGCTCAAAG	8742
MND-2.GA.-.M14	CAG-C-GTA.....ATTGAATATGGGT--CAA...GAGCTCAAAG	8671
MNE.US.-.MNE027	.....ATCCTCCAACCAATATTCAGAGATTCTCCAGCACCCTA-AG--A...-TCC-A-	8519
LST.CD.88.447	GGAA-TCTTCTCC-CAGCCTTG-GTG-CAGTTCCCTTTCGTGGCTAGGCCATTTAGTGATCCTTTTCTTCCAACATGGGCAA-ATCTT...C--AAA...	7843
LST.CD.88.485	GGAA-TCTTCTCC-CAGCCTTG-GTG-CAGTTCCCTTTCGTGGCTAGGCCATTTAGTGATCCTTTTCTTCCAACATGGGCAA-ATCTT...T--AAA...	7840
LST.CD.88.524	GGAA--CT-CTCC-CACCCTCG-GTG-CA-TTCCTCACTGGCTGGACATATAGTGATCCTTTCTTACCACCATGGGCAA-ATCTT...T--CAA...	7846
LST.KE.-.lh07	GGAA--CTCCTCC-ATCACT-G-GTG-CAGAGCCTCAGTTGGCTGTGGCATTGTGTGATCCTTTTCTTCCAACATGGGCAA--CTT...T--CA...	8918
SUN.GA.98.L14	GGAA--GTCCTAGCACAAA-CA-CTACCACCTGCTAAGTTTCTTGTGGCACCTACTTACGACTTCCTTCCATCATGGGCAA--CTT...T--CA...	8980
MND-1.GA.-.MNDGB1	-GA-C-G-A.....ATTGAATAT...---TC...	8177
MON.CM.99.L1	G-GA-CTACC---C.....CTCCAGA-AGT-TCCCA-A--...CTGCCACCCCTCCTGCACATAC	8570
MON.NG.-.NG1	-CGA-CTACC---GC.....CTCCAGA-AGT-CGCA-CCAC...CTGATCCCTCTTCTCCGCGACC	7195
MUS.CM.01.1085	-CCC-TTG-.....TGGCTACTTCAGAGCAGCTG-A-CT-ATACA-AA-GGGTCAATAAGATTGTGGTTACGCTCAGGGCAC	8562
DEB.CM.99.CM40	C-GA-T-GA--AGGCT-T--T-A--CA-GAAGGAATATATAGACTGTCAATTTGCTCTTTGG...CAAAGA...CTGA-A...CTCTTG	8356
DEB.CM.99.CM5	C-GA-T-GA--AGGCT-T--G--C--CA-GAACAATATATAGACTTTCATTTGCTCTTTGG...CAAAGA...CTGA-A...CTCTTG	8277
SYK.KE.-.KE51	.....	8242
SYK.KE.-.SYK173	.....	8543



Rev exon 2 end \

H1B.FR.83.HXB2	..GAAGCCCTCAAATATTGG.....TGGAACTCTCCTACAGTATTGGAGTCAGGAACATAAAGAATAGTGTCTGTAGCTTGTCTCAATGCCACAGCCATAGCAGTAGCTGAGGGGACAGATAGGGT	8710
Env	. E A L K Y W . . . . . W N L L Q Y W S O E L K N S A V S L L N A T A I A V A E G T D R V	gp41
Rev	<u> G S P Q I L . . . . . V E S P T V L E S G T K E S</u>	Rev ex2
H1A1.UG.85.U455	...-G-G---CT-.....-T-G-T---G---G---A-T---A---CT---TG-C---TGT---AG-----G-T---T-----	8161
H1B.US.90.WEAU160	...-T---T---T.....-T---G---C-----A-----A-----A-----A-----A-----A-----A-----	8717
H1C.ET.86.ETH2220	...-A---T---CT-.....G-A-GC-TG-G---G---T---G---A---G---A---ATC---T---A---A---T---G---A---A-----AT-	8107
H1D.CD.84.84ZR085	...-CT-.....G---A-----GA---G---G---A-----A-----G-TG---A---A-----A-----A-----	8231
H1F1.BE.93.VI850	...-T---CT-.....G-----ACG-G-----A---C---C---T---A---A---T-----A-----AA-	7987
H1G.SE.93.SE6165	...-GT---T---T---T.....-T---GTT---G---CAG---T-----A---T---TG---A---AGT---A-----A---TT-----	8149
H1H.CF.90.056	...-CTC.....T---A---C---G---A-----A---GAT---T---CA---A-----A-----G---A	8029
H1J.SE.93.SE7887	...-GAT---CT-.....GT---C---TG---TG---G---G-----A---T---T---A-----A-----A---A-----AA-	8017
H1K.CM.96.MP535	...-CT-.....CT---TG---CA-----T---A-----A-----GA---A-----AA-	7870
H1O1.AE.TH.90.CM240	...-G---CT-.....G---T---GTTA---G---C-----A---T---A---TCT---TG---T---A-----C---G---T---C-----	8284
H1O2.AG.NG.-.IBNG	...-G---CT-.....-TA---TCA---C---GT-----A---AT---T---A---A---T-----A---CT-----A---C	8237
H1N.CM.95.YBF30	...-CT-AGA-C-C-C-T.....-GGAA-A---TGCA---G---AA---A---GT---G---G-----A---C-----T---A---A---T---T---T-----A---A---A-----A	8282
H1O.BE.87.ANT70	...-T.....GCAGC-G-AAC---A---C---CTA---A---T---GC-----ACA-----AG-CA-ACCT---AG-G---CA---TT---T---CG-CA-	8779
H1O.CM.91.MVP5180	...-T.....G-AGC-G-AA-G---A---CTA---T---G---A-----ACA-A---C---TG---A---T---TT---AG-GT---T---CA---TT---T---CG-CA-	8805
CPZ.CD.-.ANT	...A---TT---G---CTG---CAACCATCTGGAGAAC-----AGCACCCCTC-----ACAAT-A-C-GA-C-GAGATCA---AGAACA-TG-CAGACTT---T---TTGG---GG---AAAA-----CA-	8169
CPZ.CM.-.CAM3	...A---CTT---AG---G---CCT-----GGAG-GA-ACC---C---G---AAG---G---C---A---C---CA---A-----AT-AG-----A---A-----C---AA-	8147
CPZ.GA.88.GAB2	...-GCTT---G---GCC---GCT-----AAAGGA-AGG---A-----CTG---A-----GA---C-----A---C---TG---A---T-----TAT-----C---AA---A---T-----CAA-	8118
CPZ.GA.-.CPZGAB	TGCT-CGGGAG-G-CTG---CCTGCTG.....G---GGAA-TA-T---A-----G---AA---A---T-----A---A---A-----A---A---T-----AA-	8771
CPZ.US.85.CPZUS	AT---CT-AGAGC-CGC---C.....-GGGG-AA-CGCT---C---GCAAG---G---AGT-----ACA-----AT-AG---A---A-----A---A---T-----AA-	8739
CPZ.TZ.-.TAN1	TCAT---TAGGAGTTAGACA-ATC.....-T-GAGTGGAGCAG-AAT-C-T-T-CTAGCTTA-GAGT-TTGC-A-TACAAGC---TA-A---G-CTTGCTAACT-TA-A-G-T---TGACAGATT-	8405
H2A.GW.-.ALI	...-CTGG---G---G---CT-AAA.....GCAGCCTA---G-----G---T---CG---TGGA---CC---AG-AGCCTGCCGGGC-----TGCA---GGATTG-GAGAGAGA-TC-TA-AA-CACCTGGAGAGACT-	9197
H2A.DE.-.BEN	...-CTGG---G---G---CT-AAA.....GC-GCC-AA-G---A---G---T---CG---TGGA---CC---AG-AGC-TTCCAAGCA---CGCG-GGA-T---GAGAGAGA-TC-T---G-GC-C-TGG-GAT---T-	9201
H2A.SN.-.ST	...-CTGG---G---G---T-AAC.....ACAGCCTA---G---A---G---G---CG---TGGA---CC---AG-AGCCTGCCGGGC-----CGCG-GG---T---G-GAGAGA-TC-TA-AA-CCTGGAGAG-CT-	8643
H2B.GH.86.D205	...GTTGCGTA---C-----G---T---CAGCTGCT---CC---AG-AGCCTGCCGGGC-----CGCG-GG---T---G-GAGAGA-TC-T---GAGT-CCGG---G-CCT-	9165
H2B.CI.-.EHO	...CCTGCATA---C---GA---G---T---CAGCTGCT---CC---AG-AGCCTGCCGGGC-----CGCG-GG---T---G-GAGAGA-TC-T---GAGC-C-G-GAGG-CCTC	9142
H2G.CI.-.ABT96	...CAT---CGGCTAGAA.....GCAGC-TATT-CAGC---G---TTC-GCTGG---CC---AG-AGCCTGCCGGGC-----CGCG-GG---T---G-GAGAGA-TC-TA-AAAGCAC-TGGAGAGCCT-	8586
H2U.FR.96.12034	...CGT---GC-CCT-GAA.....ATAGC-TA---A---G---T---G---T---T---C---AG-AGCATTCCAAGCG---CGGG---A---TG-GCGAGAGA-TC-TAGCCGCAC-GGGAGAGAAC-	8676
MAC.US.-.239	...-T---T---GGAC-GAA.....CT---CCTA-----A---G---T---GAGCT-TT---CC---TG-GGCG-TCCAGC-G-CTGG-GAT-TG-GA-AGAGA-TC-T---G-GC-C-TGG-GAGACT-	9158
Env	. E V L R T E . . . . . L T Y L Q Y M 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	gp41
Rev	<u> R S P Q D S</u>	Nef
	\ Nef start	
SMM.SL.92.SL92B	...-GCTTG---G---AGAGAA.....GCAGGCTATA-C-GT---G---T---GA---TT---CT---C---TCG---AGCCTGCCGGAAGCAGT-GAT-TG-GCAAGA---TA-T-TC-GC-C-TGG-GACTCA-	8606
SMM.US.-.H9	...-T---A---G---GT-GA.....ATA-CCTA-----G---T---G---GT-T---C---AG-AGCA---GCARR---G-TGG---ATTTYRCGAGAGA-TC-T---AAGC-C-TGGAGAGACT-	8647
STM.US.-.STM	...-T-G---G---CTAG-A.....GCAGCCTATT-----G---T---CATCTGGA-CC-AG-AGCA---GCAAGCAGCATGG-GA---TG-G-GAGAGA-TC-T---GAGC-C-GGGAGAGACT-	8809
SAB.SN.-.SAB1C	..C-GAGAAC---G-C-GCTA.....ACAGCTC---TGA-T-TGGG-----A---G---C---AGCAGCA---A-CGTCAC-CACTGTG---TGTG-T-CA---T-CGA-AAGC-CC-GT---TCA--	9002
TAN.UG.-.TAN1	..C---CTG---T-G-C-GGA.....TGCCAAT-CA-T-AGT-TGGG-----C---T---C---AG-GG-A---A-CA---G-CTT-TG-G---TTTG---G-GC---T-CGCAAGC-CC-GTCC---C-C-	8821
VER.KE.-.AGM155	...-T-G-AGCGCC.....TGGA-T-CT-A-A-T-TGGGCTT...GG---G---C---AGCCGCA---ACAAGAAGCAG-T---CCATCTG---G-GCTTT-C-CGCA-C-C-G---C-CCA-A-	8805
VER.KE.-.9063	...-T-G-GGCGCT.....TGGA-T-CA---A-T-TGGGCTT...GG---G---C---A-CGCA---ACAAGAAGCA-GTCT---TGTG---ACGCTTT-C-CAAA-C-CAGGTC-CCAAA-	8828
VER.DE.-.AGM3	...-T-GGAGCGCT.....TTCCA-T-CA-A-A-T-TGGGCTT...GG-----C---AGCCGCA---ACAAGAAGCAG-TGTC---TTTG---ACGCTTT-C-CAAA-C-C-GGCT-CCA-A-	8326
VER.KE.-.TY01	...-T-GGA-AGT.....TTTCAAT-C-G-A-T-TGGGCTC...GG-----C---A-CG-C---ACAAGAAA-A---C-AA-TCTG---AGGC-TT-C-CAAA-C-CATGTC-CCA-A-	8282
COL.CM.-.CGU1	..CTGC-T-AAC-TCCGCCA...ACAACCTCAGAC---GAGGAGGAGTGC-----ACT---TCG-T---CTA.....-G-ACCCAGGGGTC-A-GCCTTCAGAT-CTT-A---CT-CA-AG	8063
DEN.CD.-.CD1	..TTT-TA-AT-GGC-CA-CTTAATTTGCTGTG-CT-CT---A-T-TGGGCTCCAAG-ACTC-A-G---CAGGCTC.....-GGATT-GGGGA-T-GC-----CCA-CA---GAAG---C-----CGGAA-	8767
GRV.ET.-.GRI_677	...A---GGACAGCA.....GTGG-AT-C---G-AGT-TGGGCTCCAG-ACTC-A-G---CGACGAA-A...G-GC-TGCTC-A---TCTG---G-GG---T-CGAGG---A-CCTGG-GC---AC-	8681
GSN.CM.99.CN166	CGAG...-AGCTCAGGAA.....GTAGCAGCC-GAG-GA---GCC---CATTT---CT-T-GGA-CC-AG-ACTCAA-CAGCAGCAACAGGAAT-CTC-A---GCCTT-CTATCTTAC-TGGA-C....	8672
GSN.CM.99.CN71	CGAG-CAG---GGGG-AGT.....GCAGCC-GAGCAGC-GCC---ATT---CT-T-GGA-CC-AG-ACTCAA-CAGCAG-CTCAGGAAT-CTC-A---GCCTT-CT-TCTTCAATGGA-C....	8672
DRL.-.-.FAO	AA-CG-GAAGAGGCATA---C.....CGCT-CTACAAGAGG-TGCGC-GAGG-T-TGGCGC-A---GGAGGC-ACTTGG-CTATCATC	8521
RCM.GA.-.GAB1	...-AGT-C---AG-ATTCAGCAGCT-G-----A---GT-C---AG-ATTCAGCAGCT-G---TGCTG-GATGG-GCTTCA-AAT-CTTACT-CACCTGGAGAG---T	8367
RCM.NG.-.NG411	...-A---GT-C---AGGCTTCAGCAGCT-GC-TGCGG-GATGG-GAGG-CTAAT-CTTACT-CACCTGGAGAG-AC-	8450
MND-2.-.-.5440	CA-CG-GAGCAGC---TA-AT.....C-AGTATTACAAG---TGCGC-GAGGCT-TGG-GCAG---GGTACC-ACTTGG-CTATCA-G	8460
MND-2.CM.98.CM16	CA-C---GAGGA---ATA-TT.....AGTACTACAAG---ATGCGC-G-GGCT-TGG-GCAG---GGCACC-AT---GG-CTATCAAC	8824
MND-2.GA.-.M14	CA-CG-GAGCA-GGATA---C.....C-AGTATTACA---TGCGC-GAGGCT-TGG-GC-G---GGCACC-ACTTGG-CTATCA--	8753
MNE.US.-.MNE027	...-T---GGAC-GAA.....CTA-CCTA-----A---G---T---GAGCT-CT-CC-AG-AGCG-TCCAAGT-GCCTGG-GAT-TG-GA-AGAGA-TC-T---G-GC-C-TGG-GAGACT-	8635
LST.CD.88.447	...-CCTG-AG---G-CGGT---GT-G-AAA-GCT---A-GATTGCGAGCTGGAGCTG---C-AGATCAGA---GAACAGA-GA-A---T-T-AGC---T.....CATG---AGA-ACC-	7950
LST.CD.88.485	...-CCAG-AG-GG---GGAT---GT-G-AAA-GCT---A-GATTGCG-GCTGGAGCTG---C-AGATCAGA---GAACAGA-GA-A---T-T-AGC---T.....CATG---GAGAGACA-	7947
LST.CD.88.524	...ACCCT-AG---G---TCCT---GT-G-GAGTGT---A-ACTGCGAGCTGGA-C-G---C-AGAGCAGA---GAACAGAGGA-AGCT-T-A-A-A-T.....AATGG-ACA-ACA-	7953
LST.KE.-.lho7	...ACCAG-AG---G---GGAT---GT-G-AAA-GCT---A-GATTGAGCTGG---C-G-G-AATGCAGA---GAACAGGAGCAGTT-T-G-GCA-T.....AT-G-A-A-ACA-	9025
SUN.GA.98.L14	..ACCCT-AGAGGT-GGCT.....G-TTCG-A---TATTC-GGCT-CGAGCAGGA-C-G---CGCGTGAGA-C-AGCAGGG-GAGAGT-T-ATCGT-TCAGAAG-CAA-A-GTAGA-CCT-	9096
MND-1.GA.-.MNDGB1	...TCT---AG---C---G-AGC---ATC-GAAG-T-GGTT-CGCTACTCTGCTAGCT-TGC-GCA-T-AG-TGG---GGCCGGTT-CA---GG-TGGC---CA---C-AA-AGAATCTTCAAG---TG	8293
MON.CM.99.L1	GCCTGT---AACTG-GGGAAGCCTTAGAGCC-T-CTAGC-TACTGCC-G-ATG-ATGCA---G-ACATCAGA---A-C-C-C---CT---A-G-CCTTGTAGTT-CA-AATC-TCTGGCAGGAT-C	8700
MON.NG.-.NG1	TGTGCCGA-AGCTTCCGGAAGCCAGTAGCCGCTTCT-GC-TAC-T-C-G-ATG-G-TCC---GAGYTCC-ARCA---CTGC-CAGGAGC---T---ATG-CCT-GC-CGGT-CA-A-TCATCTGGACAGAC-C	7325
MUS.CM.01.1085	TGATT-AG-CTCTCAGGAAGCCTTGCAAAGACCCTCGCTTCTTAC-GCATG-GATCC---GAGCTCC-A-CA---CAGC-CAGCAG---CT---ATCAGCT-GCCA-CT-CA-CTG-AACTGGACGGAAAC	8692
DEB.CM.99.CM40	GG---CG-T-T-G-GTGCTTGGG.....GGGTA-T-----G---T-GA-A---T-C---AG---CCATCAGT-GGCAGTGGG-AGCTTTGGCAGTGGC-CCAG-G-ATTG-T---G-T---C	8474
DEB.CM.99.CM5	GG---CATT-A-G-GTGTCTGG.....GG-TATTT---A---G---T-GA---T-C---AG-C-CATCA---T-GGCAGGGC-A-AA-TCTGGGAGTGGC-CCAA-G-ATTG-TC-GGT---C	8395
SYK.KE.-.KE51	...GGAT---A-CAGT---G---GTCACA---G---C---AG-AGCA---CA-CTCGC---GGAGCA-AAGTG---GCACCAGA-G-TG-CCATTTGGC-AGCCC-	8337
SYK.KE.-.SYK173	...GGAT---A-CAGC---G---TTCA-C---G---C---A-TCGCA---A-C-TCGC-CGGA-GG-AA-T-CG-GATTGG---G---G-CTATTTGGC-AGCAA-	8638

PLV  
complete genomes

	Env end \	/ Nef start	
H1B.FR.83.HXB2	TATAGAAGTAGTACAA.....GGAGCTTGTAGAGCTATTTCGCCACATACCTAGAAGAATAAGACAGGGCTTGGAAGGATTTTGGCTATAA.G.....ATGGGTGGCAAGTGG		8811
Env	I E V V Q . . . . . G A C R A I R H I P R R I R Q G L E R I L L \$ # . . . . . M G G K W		Nef
H1A1.UG.85.U455	-----A--G-----AC-AT-G-----T-A-----C-----A-----A-----GC-----C.....		8262
H1B.US.90.WEAU160	-----A--G-----A--A-----T-----A-----A-----GC-----C.....		8818
H1C.ET.86.ETH2220	-----T--A-----A--AT--G-----T-CT--A-----C-----A-----GCAGC-----A-----A-----G-----C-AT-		8208
H1D.CD.84.84ZR085	-----CA-----AGG-----A-----G-G-TT-----C-C-----T--A-----GC-----A-----A-----G-----A-----		8332
H1F1.BE.93.VI850	-----TT-G-----A-----G-----G--T-A-----C-----A-----GCA-----GC-----G-----A-----G-----		8088
H1G.SE.93.SE6165	-----C-----A-----G-----C-T-A-----C--G-----A-----AGC-----A-----A-----A-----		8250
H1H.CF.90.056	-----T--A-----G-----A-----G-----T-----T-----A-----A-----GC-----TA-----A-----A-----		8131
H1J.SE.93.SE7887	C-----A--C-----A-----T-----TT-----T-----A-----AGC-----A-----A-----GAA-----A-----		8118
H1K.CM.96.MP535	-----A--G-----A-----T-----T-A-TT-----TC-C-----AGC-----A-----A-----A-----		7971
H1O1.AE.TH.90.CM240	-----C-----A-----G-----C-----T-----C-----T-----A-----C-----C-----AA-T-----		8385
H1O2.AG.NG.-.IBNG	-----A--G-----A-----TGG-----A-----C-----C-----C-----GC-----C-----		8338
H1N.CM.95.YBF30	-----T--C-----A--ATAG-A--G-GA--ATTA-----C-----A--C-A--AGCAC--A-----AAAG-TT-----		8383
H1O.BE.87.ANT70	A--C-C--GGA-----A--ATAG-A-C--GA-----TA-----C--A--G--T-----A--A--G--AT-G--C-----AAA-GCA-T-		8880
H1O.CM.91.MVP5180	C--CTT--GTC-----A--ATAG-ACA--GAT-C-TT-----C--A--T-----A--TGCA--A--C--AG-G--C-----GAATGCA-		8906
CPZ.CD.-.ANT	AC-TCT--CTC-C---ACTATAGTC---ATC--AA-GG-AG-----GC-C-C---A--G-----TTGCA--AAAT---AC.....TCTGCA---		8271
CPZ.CM.-.CAM3	-C-----T-C-----AT-ATAG-----GC--TG--T---A-----T-----T--A-----A-----C-----CAA--G-----		8248
CPZ.GA.88.GAB2	A-----C--ACT---A--ATAG-G---GC--ATTG--T---A-----C--G--A--T--A-----AGC--A--G--TAA-----A-----CAA--A---		8222
CPZ.GA.-.CPZGAB	A-----CTT-T---A-----TTA-ACT---AT---A-AA---C---GC-C---A--GC-----AGC--A--T---CACTGCATTAAG-----AAC--A---		8884
CPZ.US.85.CPZUS	-----T--AC-AG---A--GCTC-T-CT--G---ATA-----A-----T--G--A--C--A--G--A--GC--AT-----AA-----		8838
CPZ.TZ.-.TAN1	A--CAT--A--G-GTGGTT---TACATAGCC--G-GA--CA-AA-T--T-----T-----T-----TC-----CTAGCC--AAAT---A-----AAA--TA-TT		8509
H2A.GW.-.ALI	GTGG-GG-C---G--G---T--G--TCG--G--GAGG--A--T--GCAG--C--A--G--G--C--G---GGCA-----TTGCC--C--G--G---		9282
H2A.DE.-.BEN	ATGG-----C--CG--G---C--CATCG--G--G--GA--A--T--GCAG--T--A-----C--G---AGCA-----CTCGCC--C--G--G---		9286
H2A.SN.-.ST	CTGG-GGAC-C-GGG---CA-AT-G-G-G-GA--A--TTGCAG--C--A---G--C--G---GGCA-----TCGCC--C--G--G---		8728
H2B.GH.86.D205	ATGG--G--CTC-CAGG---A-----GGCG---A--CAT--GCG--C--C--G---C---A--AC--T---CTC-CCC-CT-G-G---		9250
H2B.CI.-.EHO	GTGG-G---CC-CAGG---A-----GGCAG--AG--CATTGCA---C--G--G--C---A--GGCT---CTCGCC--CT-G-G---		9227
H2G.CI.-.ABT96	ATGGA--ACTC-GGG---A--G-T-G-A---GG--A--T--GCA---G-----C--G--A--GC--T--GCTC--C--C--CT-G-G---		8671
H2U.FR.96.12034	ATGG--GACTC-GGG---C--G-T-G-A---TGCC---AGCT---A--G--G--C--G--A--A--T--GCTCGCG--TT-G-G--A-----		8762
MAC.US.-.239	ATGG--GACTC-TAGG---A--G-G-A---TGG--A--T--GCA--C--C--G--G--T---A--GC--T--GCTC--C--C--CT-G-G---		9243
Env	W E G L R . . . . . R G G R W I L G I P R R I R Q G L E L G L L \$		
Nef	Y G R L L . . . . . G E V E D G Y S Q S P G G L D K G L S S L S C E . . . . .		Nef
SMM.SL.92.SL92B	ATGG--GACTC-GGG---C--G-T-G-A---GG--CGCAGCA--C--A--G--G--C---C-T---CTC--GC-TAAC--G-----		8691
SMM.US.-.H9	ATGG--GACTC-GGG---A--G-T-G-A--G-GA--A--T--GCA--C--MR--C-C--C--G--A--GC--T--GCTC--C--C--CT-G-G---		8732
STM.US.-.STM	ATGG---ACTC-TGG---A--G-T-G-G---AGG--CG-GGC---C--A--GC--C--C---A--GC--T---CTC-CAC-CT-G-G---		8894
SAB.SN.-.SAB1C	ACG-C-T-CTTGCA---TCCAT-GT---G--AG-CAT-GCACAT--A-----GC-C--A--AGC--T--G--ATGG--TAAT--G-----		9087
TAN.UG.-.TAN1	ATGGA-T-CTTGCA---TCC---A--C--G--A--C--T--G--ACAT--A-----GC---A--AAC-----C--TGG--CAAC--G-----		8906
VER.KE.-.AGM155	ATGGCTT-CTTGCA---TCC---A--C--G--A--CAT--A--TCT--A-----G--GC---A--GC--T--GGAAG--CC--TAAT--G-----		8890
VER.KE.-.9063	ATGGCTT-CTTGCA---TCCA---A---GCAC--CAT--AG--GT--A-----G--C---A--GC--T--GGAA--CC--TAAT--G-----		8913
VER.DE.-.AGM3	ATGGCTT-CTTGCA---TCC---A---G--A--CAT--A--TCT--A-----G--GC---A--C--T--G--A--CC--TAAT--G-----		8411
VER.KE.-.TYO1	ATGGCTT-CTTGCA---TCC---A---GAAC--CGT--A--GT--A-----G--GC---A--C--T--GGAA--CC--TAAT--G-----		8367
COL.CM.-.CGU1	GTGGCT-CG-TCTGCGCG---C--T--GC---GC--TCGAGCG--CTG--ATAT--TGC--G--TGG--T--T--A--CGTCCCAAGG--CCTGCC---		8151
DEN.CD.-.CD1	GG--TTG-CG-CT-TTACC---C-----AGCA---AAG--GGTGGC---C--C---C--G--C---C-T---TTG--CC-CAAC--G-----		8855
GRV.ET.-.GRI_677	GGGT-CTA-T--C-G---TCC---A--C--G--AG-CAT--A--GT--A-----G--GC--G--A--C--T---AAG--CC--GGG---		8766
GSN.CM.99.CN166	.TGGAC--AG-CCGT-CTTACGCTTGTAGAC---TA--GC---AAT---TGGCT--C--CC--C--C---A--GGC--G--TAC--A--TAAC--G--A-----		8772
GSN.CM.99.CN71	.TGGAC--AGCC-TCGTTTCACTGTTGGAC---TA--GC---AAT-CTTGGCT--C--CC--C--C---GGC--G--TAC--AC-TAAC--G--A-----		8772
DRL.-.-.FAO	AGCTAG--CCC-C-G---CCCT-GC-CAG-AAG-GGCAGC---T--C--G--G--T---A---GCA--GGTCC--C--CAAC--G-----		8606
RCM.GA.-.GAB1	ATGT-CG--G-C--G---ACTT-GC-G--TGGCCAGC-ATGG-GTGC---G--C--G---C--C--G--AC--GTAAT--G-----A--G--A---		8456
RCM.NG.-.NG411	ATGC-CG--TTGCAG---ACC--GCAG--TGGCCAGC-ACGT-GTGC---G--C--G---C--G--AT--C--TAAC--G-----A--G--A---		8539
MND-2.-.-.5440	-G-TAG--G--C-GC---CCTT-G-C--G-GC--CT-GA---C---C--C--C---A--TGCT--GGCC-AC-CAAC--G-----		8545
MND-2.CM.98.CM16	AGCT-C-TGCT-CAG---CCAT-GC---G-GC---AT-A---C--A-----TC-----AGCA--GTCC-CC-CAAC--G-----		8909
MND-2.GA.-.M14	---TAG--G--C-GCG---CTAT-G-A--G--C---G-GA---C---C-G-C--C---A--AGCT--GTCC-GC-CAAC---G-----		8838
MNE.US.-.MNE027	ATGG--GACTC-GGG---A--G-T-G-A---TGG--A--T--GCA--C---G--G--C---A--AGC--T--GCTT--C--C--CT-G-G---		8720
LST.CD.88.447	C-G-CTG-G-CA-A-G---AAGAGA--G---TTC-GATT-AGAGG-AGA--TG-GT-TCCCAGT-AGACCAC-GA--CCGCT--G--G-----		8035
LST.CD.88.485	C-G-CTG-G-CA-A-G---AA-AGA--G---TTC-GATT-AGAGG-AGA--TG-GT-TCCCAGT-AGACCAC-GA--CCGCT--G--G-----		8032
LST.CD.88.524	---GGTTG-G-CA-A-CAGA---A--AGA--G---TTC-GATT-AGA-----TG-GC-TCCCAGT-AGACCAC-GA--CCACT--G--G-----		8035
LST.KE.-.lho7	CCA-TTG-GCAA-A-G---AAGAGA--G---CTC-GATT-GGAGG-AGA--TG-G--TTCCCAGT-AAGCCAC-GA--CCGCTT-G--G-----		9110
SUN.GA.98.L14	---GTCTC-GCAG-A---T--GAGGCC-A-TGG-A-A-AACA-G-GGA--T---TCCCAGT-AGACCAC-GA--CCACTT-G---		9181
MND-1.GA.-.MNDGB1	C-G-----C-A-CATT---CT--GG-G-CGTGC--AT-G-----TC---G--T---AA-TGCA---C-ACCC-CAAC---		8378
MON.CM.99.L1	GC--CTTCAT-CTGG-GGA---C--CTA--G---A--CGTGGCT--C--C--C--C---AGC---TAT--CC-TAAC-G-G-CGAGAA---		8796
MON.NG.-.NG1	GG-CCTTCGCT-GGG-GGA---C--TTA--GC---GAC-CGTGGCTG-C--C--C--C---GGC---G-TTC-C-TAAC-G-----		7413
MUS.CM.01.1085	AG--CTTCA--CTGGTCGA---A---TA--GC---AAT-CTTGGCT--C--C---C-G-C--G---GGC---G-TAC--CC-TAAC-G-----		8780
DEB.CM.99.CM40	GT--C-G-G-C-TGTC---A---AG-AG---A---T-AGAG-G--GCT---C-----AGCA--GCTC--CC-CAAC-G-----		8559
DEB.CM.99.CM5	G--C--G--C-TGTC---T--G-GAG--C---AACC--T--GC---T---GC---C---A--AGCA--GTT--T-AC-CAAC-G-----		8480
SYK.KE.-.KE51	AC-C-CCTAT-CCAG---C---T-GC-GAGAA-G-CGCAGC-T---AC-----C-TC-----GC---G-TCTACC-TAAC-G-----		8422
SYK.KE.-.SYK173	ATAC-CC-CTACCAG---C---T-GTCGAG--CG--GCAGC-T-G--GC-G--GC-TC---A-----G-TCTACC-TAAC-G-----		8723



H1B.FR.83.HXB2	TCAAAAAGTAGTGTGATTGGATGGCCTACTGTAAGGAAAGAATGAGACGA.....GCT.....GAGCCAGCAGCAGATAGGGTGGGAGCAG	8893
Nef	S K S S V I G W P T V R E R M R R ..... A ..... E P A A D R V G A	Nef
H1A1.UG.85.U455	-----AG-CAGAG-G-A-----GAG-T--A-----GAAA.....CCTG-----A-AG-A--A-----	8344
H1B.US.90.WEAU160	-----C-----GT-G-----G-A-----A-----GT-----AG-----	8900
H1C.ET.86.ETH2220	-----T-----CCAG-A-----G-A-----A-----A-----CT-----GG-A--A-----	8290
H1D.CD.84.84ZR085	-----A-AG-----G-A-----A--AA.....ACTGATCCAAGGGAAGAGAGAC-----G-----	8438
H1F1.BE.93.VI850	-----C-A-AG-----G-----G-----A-----A-C-----CCTA-----AG-----	8170
H1G.SE.93.SB6165	-----CA-AG-----CGAG-----A--AAC.....A-C-----CCTA-----AG-A--A-----	8332
H1H.CF.90.056	-----GA-GG--G--T--A-----G-----G-----A-----T-----AG--A-----	8213
H1J.SE.93.SE7887	-----A-AG-----GAG-----G-----GCTC-C-----CT-----G-A-----	8191
H1K.CM.96.MP535	-----A-AG-----G-A-----GCACGACCAGCAGCAGAGGGTGGGAA-A.....C-A-----CG-----	8080
H101.AE.TH.90.CM240	-----CA-AG-G-----CAG-C-----A--A-AG.....CAAA.....CCT-----A--AG-A--A-----	8467
H102.AG.NG.-.IBNG	-----C-CA-AG-G-----AG-T-T-A-----CAAA-C.....CCTA-----ACAG-A--A-----	8420
H1N.CM.95.YBF30	-----G-C-CC-AG-A-----AGAAA-C-A-----A-----CAAACGCAAGAACCA--A.....-TAGAGC-----TAG-A-CA-----	8480
H1O.BE.87.ANT70	AG-----G--AAT-TGAG-----G-AG-A-----A-----A-----ACTAGAACTTTCCCT-AG.....TCT--A--TGC--CC-G-A--A--CAGA	8980
H1O.CM.91.MVP5180	AGC-----C-AAT-TGCA-----T-AGAA-----A-T-----TCTCTCTGTATCCTCAA.....C-A--TGT--CC-G-A--A--T-	9003
CPZ.CD.-.ANT	--T---TC-AGTG-G---GCAAGACAA-C--TTAGG-A--A.....CA.....-AA--AATC---GATA-A--GC-CT	8347
CPZ.CM.-.CAM3	-----CC-G-A-G-----AGA--C-A-----C-A.....AGCGAACTC.....A-C-TGCTT--CCAG-A--A--C-	8330
CPZ.GA.88.GAB2	--T---C-CA-TG-A-----GCAG-C-A-G--C-T--G--G.....ACT.....C-GA-----AG-A--A--A--	8304
CPZ.GA.-.CPZGAB	--T---C-G-A-----GAG-C-AAG-----A.....AGGGAA---CCAA-----GG-A--A--A--	8966
CPZ.US.85.CPZUS	--T-G---CA-AG-A-----AGAA--C-AA-C--T-----AGA.....CAAACCTCAGACAACAG-----AG-A--A--C-T-	8926
CPZ.TZ.-.TAN1	GGT-G-.....TG-CC--GGCC-GG-AA-CC-TC--GATC-T.....CA.....A-CA-CT--AGT--GCCT--A--CAG-	8579
H2A.GW.-.ALI	.....-G-CAGCGGTATCA-CAG-G--	9305
H2A.DE.-.BEN	.....G-G-CAGC-GTATCA-CAG-G--	9309
H2A.SN.-.ST	.....G-G-CGGCGGTATCAACAG-G--	8751
H2B.GH.86.D205	.....	9250
H2B.CI.-.EHO	.....	9227
H2G.CI.-.ABT96	.....-GGCA-C-ATACAC-CAG-GTC	8694
H2U.FR.96.12034	.....-CT-AG-GCCA--GCTATAGT-A-G--	8790
MAC.US.-.239	.....G-G-CAGA-ATACAATCAG-G-C	9266
Nef	.....G Q K Y N Q G	Nef
SMM.SL.92.SL92B	.....-GAT-G-TATA-ATACAAT-T--GG-	8717
SMM.US.-.H9	.....GC-TCAGA-ATATTGT-AG-GTC	8755
STM.US.-.STM	.....GC-TCAGAGATATAAT-A--GTC	8917
SAB.SN.-.SAB1C	.....	9087
TAN.UG.-.TAN1	.....	8906
VER.KE.-.AGM155	.....	8890
VER.KE.-.9063	.....	8913
VER.DE.-.AGM3	.....	8411
VER.KE.-.TY01	.....	8367
COL.CM.-.CGU1	.....	8151
DEN.CD.-.CD1	.....GTT-C	8860
GRV.ET.-.GRI_677	.....	8766
GSN.CM.99.CN166	.....	8772
GSN.CM.99.CN71	.....	8772
DRL.-.-.FAO	.....	8606
RCM.GA.-.GAB1	.....	8456
RCM.NG.-.NG411	.....	8539
MND-2.-.-.5440	.....	8545
MND-2.CM.98.CM16	.....	8909
MND-2.GA.-.M14	.....	8838
MNE.US.-.MNE027	.....G-G-CAGA-ATACAATCAG-G-C	8743
LST.CD.88.447	.....	8035
LST.CD.88.485	.....	8032
LST.CD.88.524	.....	8035
LST.KE.-.lho7	.....	9110
SUN.GA.98.L14	.....	9181
MND-1.GA.-.MNDGB1	.....	8378
MON.CM.99.L1	.....	8796
MON.NG.-.NG1	.....	7413
MUS.CM.01.1085	.....	8780
DEB.CM.99.CM40	.....	8559
DEB.CM.99.CM5	.....	8480
SYK.KE.-.KE51	.....	8422
SYK.KE.-.SYK173	.....	8723



PLV  
complete genomes

H1B.FR.83.HXB2	CATCTCGAGACCTGGAAAAACATGGAGCAATCACA...AGTAGCAATACAGCAGCTACCAATGCTGCT.....TGTGCCTGGCTAGAAAGCACAAAGAGGAGGAGGAG....	8991
Nef	A S R D L E K H G A I T . . . S S N T A A T N A A . . . . . C A W L E A Q E E E E . . . . .	Nef
H1A1.UG.85.U455	T----A---TT-A-T---T-----G-----..-C-T---T-T---T---CAG-.....G-----G-----A-GA--C-....	8442
H1B.US.90.WEAU160	T-----T---C-----GAG-----..-AT-----..-C-T---T-A-----..-T---GAG..	8992
H1C.ET.86.ETH2220	-G-----T-A-C---T---G---C-T---..-C---C---C-C-AT-C-A-.....G-----G-----A-AGAG..	8391
H1D.CD.84.84ZR085	T-----A-----G-----G-----..-T-----..-C-----..-G-----G-----A-AGAG..	8539
H1F1.BE.93.VI850	TG---A---TT-A-C-G-GG---T---..-C-----TAG-A-..-T---C---AC-.....-TG-----G-----A-....	8268
H1G.SE.93.SB6165	T---A---TT-A-T-GG---G-----..-C-----..-C-AT---C-AT---C-....G-----G-----CTCAGAA..	8433
H1H.CF.90.056	TG---A---TT---T-G-GC-G---G-----..-T---AT-----T---T---CG-A-.....GCC-----G-----C-G---GAA..	8314
H1J.SE.93.SE7887	TG---A---TT---CT-G---G-----..-C-----..-TT---CAC-AT---C-A-.....G-----G-----ACA-....	8289
H1K.CM.96.MP535	T---A---T---A-T-----G-----..-A-----..-T-----..-ATAGAT---A-.....T---GAG-----C-....	8178
H101.AE.TH.90.CM240	T---A---T---A-T-----G-----..-A-----..-T-----..-ATAGAT---A-.....T---GAG-----C-....	8559
H102.AG.NG.-.IBNG	---A---TT-A-T-G-----G-----..-C---T-----CAA-T-CC-A-.....G-----G-----T---A-T-....	8518
H1N.CM.95.YBF30	-T---A---T---A-CT--T-GA--G-C---C...-TA-A---TAG--AC-AT---AAAG-.....ATA-T-----A-A-A---GAA..	8581
H1O.BE.87.ANT70	TC--CA-G-AT-A-C-GCTAGA---GG-AC---..-TC-C---TC-TCAA-A---A-C-.....CT--A-TC-----AGT--CC-A---A-....	9078
H1O.CM.91.MVP5180	TC--CA-G-AT-A-C-GCTAGA---GG-AC---..-TC-C---TC-TCAA-A---A-C-.....CT--A-TC-----CAGC--CA-A---T-....	9101
CPZ.CD.-.ANT	GTGGAAAC--AT---CC-GTAGA--G---C-A-C-.TCCTCT-C-T-...-G---AG-AAAA-A-.....GTAAT-ACCTATTC--A-G-TC-TACA-----A...GG	8444
CPZ.CM.-.CAM3	TC--A-A---T-A-GGC-----G-----..-C---GC-C---C-CAG-AT---CAGA-.....CTG-----T-----A-ATGC-A--AC-CA-TGAAGA	8433
CPZ.GA.88.GAB2	TG---AG--TT-A-C---GG-G-----..-A-T---TCC-AC---CC-TCAG--T---CAAA-.....CT--A-----A-ATG-AC-----T-....	8402
CPZ.GA.-.CPZGAB	TT--GAAG---A---G---C---T---..-G-C-CC---AG-T---CAAA-.....CTA-T-----G-A-ATG-CA-T-A-...GA	9063
CPZ.US.85.CPZUS	TC--A-A---T-A-C-G-----G-T-----..-C---A---C-CAG-A---C-CAAA-.....CTG--A---T---T-AGATGACTA-TC-T---AAGTGA	9029
CPZ.TZ.-.TAN1	-C-A-A---CC-G--TA-A---GTC---T...-C-A---CCCT---G---TCA-A-A-.....GTGTTAGAATACTCT--G-CC-TACT-A--AGAA..	8677
H2A.GW.-.ALI	AT-T-ATGA-TACCC--TGGAGAACC--GCA---..GAA--AG-G-A--A-TTGTAT--GCAACAG-.....AA-ATGGATGAT-T--ATTT--T-T-T--TAGCCT	9408
H2A.DE.-.BT	AG-ACATGA--AGCC--TGGAGAACC--GCA---..GAA--AC-G-A--A-ATTTGTAT--GGCAGCAA-.....AA-ATGGATGAT-T--ATTTCT--T-T-T--C...CT	9409
H2A.SN.-.ST	AT-T-ATGA-TACCC--TGGAGAACC--GCA---..GAA--AG-G-A--A-TTGTAT--GCAACAG-.....AA-ATGGATGAT-T--ATTTCT--T-T-T--C...CT	8848
H2B.GH.86.D205	.....G-G-CAA-AT...GCC--GGGG--A--GGGGAGGGCAACAA-.....GATGC---T-A-AGT--T---C--TGAA..	9320
H2B.CI.-.EHO	.....T-G-CAA-AG...-C-TTGGGAG--A--GGGGAGGGAAACAA-.....GATTC---T-AGG-T--T---CA-TGAA..	9297
H2G.CI.-.ABT96	A-T-ATGA-TACCC--TGGAGAACC--GCA---..GAA--AGCA-AGTT--ATA--GGYAGCAA-.....AATAATGAT--T-T-G-TAGT--T--TA-TGACTT	8794
H2U.FR.96.12034	AG-TCATGA-TACACCTTGGAGAACC--GCA---..G-G-AAG--A-TT-CAATAT--GCAACAA-.....AACATGGATGAT-T-ATG---A--TA-T--T...TT	8890
MAC.US.-.239	AG-A-ATGA-TACTCC-TGGAGAACC--GCTGA...GAG--AG-A-A-TT---ATA--GAAAACAA-.....AA-ATGGATGATAT--ATG-GT-A--T-T--C...TT	9366
Nef	Q Y M N G P W R N P A E . . . E R E K L A Y R K Q . . . . . N M D D I D E S D D . . . L	Nef
SMM.SL.92.SL92B	A--T-ATGA--ACCC--TGGAGAACC-TGCT---..GAA--AC-A-AGCA-CAGTAT--GGCAGCAG-.....--ATGGATGATAT--ATG-G--T-T-T---CT	8817
SMM.US.-.H9	A--TCATGA-TACCCCTTGGAGAACC--GCA---..GAA--AGCA-A-TT--G-TAT-GACAACAA-.....AACATGGATGAT-TG-ATA-T--A--T-....TT	8849
STM.US.-.STM	A--T-ATGA-TACCCCTTGG-AAA--GCGAG--..GAA--GC--AGCT--AATAT-GACAGCAA-.....AACATGGATGAT-T-ATG-G--A--T--TA-T...CT	9017
SAB.SN.-.SAB1C	.....G-AA-G-G-CTCCTCTAG--CAGA-..GAAG-TGTA-A-ATGAAACATC-A-GAAGA-.....CAGC---C-TGGT-T-ATG-G--T--A--A-....	9175
TAN.UG.-.TAN1	.....GAGG--AAG-G--C...CC-GAAGGA-G-AA-TTGG-AGC--A--AC-.....ACATGGATGACTGG-A--C-----A-A-A-....	8985
VER.KE.-.AGM155	.....-A-G-G-A---CAAC--G--T...CCAGAAGGA-G-CGTCTACAGG-C-GG-AC-.....CAATGGGATGA-TGGT--G-T--A--A--T--A-....	8975
VER.KE.-.9063	.....-A-G-GGA---CAAC--G--T...CCAGATGGA-G-CGCTGCAAG-A-GA-AC-.....GAATGGGATGA-TGGT--G-T--A--A--T--A-....	8998
VER.DE.-.AGM3	.....-A-G-G-A---TAA--G--T...CCAGAAGGA-G-AA-TTACAAG-A-GA-A-.....AAATGGGATGA-TGGT-TG-T--A--A--T--A-....	8496
VER.KE.-.TY01	.....-A--C--A---CAAC--G--G...GAAGAGG-A-GGAGGCT-CAAG-A-GA-AC-.....ACCTGGGAAGAGTGGT-GG-T--T---A-A-....	8452
COL.CM.-.CGU1	.....-T-GCC--TTGCC-GCT-TGCTA-AGACATCCAGATGTATATTACCACAACC-AGAA-C---G-ACGT--A--T-TA-A...GC	8232
DEN.CD.-.CD1	TGGAG-AT--TGA---CG--G-AACCTTTG-AGATGAGAAGAA-TC-TCCA-T-AT-TCCA--ACTGGACCTCAACATCCAAGAATACACCAATATGT-G---CCATG-T--T--A--A-....	8985
GRV.ET.-.GRI_677	.....A--TAA---G--T...CCAGA-GGCGCCGCTGCAAG-A-GA-AC-.....ACCTTTGATGAGTGG-ATG-T--T--A--A-....	8842
GSN.CM.99.CN166	.....GAGCAAAACCCAAAGAG--G-AGCCTC-----T-T...GA	8810
GSN.CM.99.CN71	.....GAA-CAC---CCAGGC--AATTGCATC--T...T-T...GA	8807
DRL.-.-FAO	.....GCTTAGA--CT-C-CATATGCAAG-AGCA--A--AGAGAAAGATGTACCCTTCCCACAGAA-AAGAA-A-CCCT--G---A-A-A--AGAA..	8703
RCM.GA.-.GAB1	.....-CATT-----GCAG-A--GCA----T-A-TA-A--AC-.....AT--A---TAT-----G-TACT--T--CACCTTG..	8525
RCM.NG.-.NG411	.....GTTT-----TAG-A--TCAGCAGT-AGTG-A---A-.....ATA-AT---TAT-----G-TACA-T--TACCTTG..	8608
MND-2.-.-5440	.....GCGGGATCTAAGGACCAGCAGGATGTAACCTCTCCCTCAGAAAAAGA--A-CC-AGTG-T-A----A---A-....	8618
MND-2.CM.98.CM16	.....CG-----A--AAC-....	8925
MND-2.GA.-.M14	.....AG---T-ATG-GC-AA-A--A-....	8863
MNE.US.-.MNE027	AG-T-ATGA-TACTCC-TGGA-AAACC--GCTGG...GAG--GG-A-A-TT---ATA--GAAAACAA-.....AA-ATGGATGATAT--ATG-G--A--T--T-GC...TT	8843
LST.CD.88.447	.....	8035
LST.CD.88.485	.....	8032
LST.CD.88.524	.....	8035
LST.KE.-.lho7	.....	9110
SUN.GA.98.L14	.....	9181
MND-1.GA.-.MNDGB1	.....AGATGAGAA..ACTTGAC-TAA-ACA-AG--TC--G-----A--A-....	8427
MON.CM.99.L1	.....ATGGCAAACAACGCC-GCAACCT--G--C-ACG-C--C--C--C--C--C-CTTGT	8848
MON.NG.-.NG1	.....G-GGCG--CCACCGC...TCAAACCGCGCCCAACCAC--T-GTTGAT--C--CG--C-----CC-A--TTTCC	7485
MUS.CM.01.1085	.....AGGACCCAGTAAAG-CCT-TGAT--A--C-T---CGT..	8820
DEB.CM.99.CM40	.....GTT-TGG-GAAG-CCTAGAAGAGAAATAAGGTGACACAA-AAGCAA-A-C---ACTT---TC--ACA-A...GC	8631
DEB.CM.99.CM5	.....G-A-C-ACT-T---A-ACCCCAACAGAGAGATAAATGCACCTCAGAGACC-AG--GTGTT-T-T---AC-T...GG	8558
SYK.KE.-.KE51	.....-CC--TCTTGA-CC---A-G--ACTT-GA-TC--GG-GGAAAAATGGAATGCGGGAGACATC--CATGATGAG-G-ATG-G-...-T...CTAGTAGG	8518
SYK.KE.-.SYK173	.....GCC--TTGACCCACA-GGCG-G-ACAGAGA-TGG-GACA-TCGGGAGGACAAAAATTCACCTG-A-ATATAG-GC-G-ATG--GA--CACA-GT...CT	8822



Accession	Sequence	Position
H1B.FR.83.HXB2	.GTGGGTTTT...CCAGTCCACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTAAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCAAAGAAGACAA	9117
Nef	V G F . . P V T P Q V P L R P M T Y K A A V D L S H F L K E K G G L E G L I H S Q R R Q	Nef
H1A1.UG.85.U455	..-A-C-C-...-T-GG-A-...-T-...-T-T-...-T-...-T-...-A-G-A-...	8568
H1B.US.90.WEAU160	..-A-C-...-T-G-...-GCAT-...-T-...-T-...-A-...-A-...	9117
H1C.ET.86.ETH2220	..-A-C-...-G-...-T-...-AT-C-...-C-...-T-...-T-...-A-G-A-...-G-G	8517
H1D.CD.84.84ZR085	..-C-...-G-...-G-...-C-T-...-C-T-...-G-A-...-G-...-T-...-A-G-A-...-G-G	8665
H1F1.BE.93.VI850	..-A-C-...-G-...-G-...-C-T-...-C-T-...-T-...-T-...-A-G-A-...-GG-	8394
H1G.SE.93.SE6165	..-A-C-...-G-A-A-...-G-...-TT-...-GT-...-T-T-...-C-...-TT-...-T-...-T-...-A-G-A-...	8559
H1H.CF.90.056	..-A-C-...-GG-...-G-...-C-T-...-G-...-T-T-...-C-...-TT-...-T-...-T-...-A-GCA-...-G	8440
H1J.SE.93.SE7887	..-A-A-...-A-...-A-...-A-...-G-...-T-...-T-...-T-...-TT-...-T-...-T-...-A-G-A-...	8415
H1K.CM.96.MP535	..-A-...-G-...-A-...-T-...-T-...-CG-...-TT-...-T-...-T-...-A-G-A-...	8304
H101.AE.TH.90.CM240	..-A-C-...-TG-G-...-C-...-T-...-G-...-T-T-...-TT-...-T-...-T-...-A-G-A-...	8685
H102.AG.NG.-.IBNG	..-A-C-...-GG-A-...-G-...-T-...-G-...-C-...-C-...-T-...-T-...-T-...-A-G-A-...	8644
H1N.CM.95.YBF30	..-A-C-...-ACGC-...-A-...-G-...-A-C-...-T-ACAG-...-T-T-...-TC-...-TT-...-T-...-G-TGG-...-AG-...-A-...-G-	8707
H10.BE.87.ANT70	..-A-...-AG-...-A-G-...-C-...-G-...-C-T-...-A-T-...-C-...-C-...-TT-...-A-...-T-...-T-...-A-G-A-...-GC-	9204
H10.CM.91.MVP5180	..-A-C-C-...-A-G-...-A-G-...-C-...-G-...-C-TT-...-A-...-CT-T-...-C-...-C-...-TT-...-A-...-T-...-T-...-T-...-A-G-...-GC-	9227
CPZ.CD.-.ANT	GACA-C-...-T-G-G-...-A-C-CA-G-...-C-...-AG-A-...-ATT-...-C-...-GTCATGN-...-TN-...-T-NN-A-TT-A-...-GC-	8571
CPZ.CM.-.CAM3	A-A-C-...-T-G-...-A-...-C-T-...-T-...-C-TT-...-A-...-T-T-...-T-A-...-A-...-T-...-AT-...-NNNNNNNN	8560
CPZ.GA.88.GAB2	..-A-G-...-A-G-...-A-...-A-...-T-...-A-...-CT-T-...-C-...-T-...-G-...-T-T-...-A-G-...-G	8528
CPZ.GA.-.CPZGAB	A-A-A-C-...-A-G-...-A-...-T-...-AACG-...-T-...-T-...-T-...-TCA-...-T-...-G-...-T-...-AGG-...-G	9190
CPZ.US.85.CPZUS	A-A-C-C-...-T-TCGC-...-A-...-A-...-G-...-C-...-ACA-...-AT-C-...-T-AG-...-TT-...-T-...-G-...-T-...-AG-...-G	9156
CPZ.TZ.-.TAN1	..-A-...-G-...-AGCA-...-CA-G-...-C-...-AG-G-...-CT-...-AA-...-GTCATGG-...-C-...-T-...-T-...-TT-...-T-C-...-A-...-GC-	8803
H2A.GW.-.ALI	A-A-GG-C...-T-T-...-AAGA-...-AA-...-A-...-A-...-ATTG-A-...-A-...-GTCA-...-T-...-AA-...-G-...-A-GT-...-T-...-AGTG-...-G-...-T	9535
H2A.DE.-.BEN	AA-A-AG-...-T-T-...-AAGA-...-ACGG-...-GA-...-C-...-T-...-ATTG-...-AA-...-A-...-GTCA-...-T-...-A-...-C-...-A-GT-...-T-...-AGTAGG-...-G-...-T	9536
H2A.SN.-.ST	A-A-GG-C...-T-...-AAGA-...-A-...-GA-...-A-...-T-G-TTG-...-AAG-...-A-...-GTCA-...-T-...-GA-...-GTA-...-T-...-AGTG-...-T-...-G-...-GT	8975
H2B.GH.86.D205	..-AGCCATATGTAAG-C-AATA-...-AC-G-...-A-...-ACT-...-A-...-CA-GTCT-...-T-...-A-...-C-...-A-GTA-...-T-...-TAGTG-...-G-...-C	9449
H2B.CI.-.EHO	..-A-GG-C...-GT-A-G-...-CGG-...-C-AC-G-...-G-...-A-T-...-ACT-...-A-...-CA-GTCT-...-T-...-A-...-A-...-TT-CT-...-TAGTG-...-G-...-T	9423
H2G.CI.-.ABT96	A-A-AG-A...-TAC-CA-A-C-GC-...-GT-...-G-...-ATTG-...-AA-...-CA-GTCA-...-T-...-A-...-A-...-TTA-...-T-...-AGTG-...-G-...-T	8921
H2U.FR.96.12034	A-A-AG-A...-G-...-A-G-...-AAGA-...-C-AC-G-...-A-...-ATTG-A-...-A-...-GTCA-...-T-...-A-...-A-...-T-...-T-...-T-...-T-...-G-...-T	9017
MAC.US.-.239	G-A-AG-A...-T-...-G-GG-...-AA-A-...-T-CC-...-A-...-G-...-ATTG-...-AA-...-CA-GTCT-...-T-...-A-...-A-...-TTA-...-T-...-AGTGC-...-T	9493
Nef	V G V . . S V R P K V P L R T M S Y K L A I D M S H F I K E K G G L E G I Y Y S G R R H	Nef
SMM.SL.92.SL92B	G-CT-AG-A...-GTT-...-TGG-...-AA-A-...-CC-...-G-...-T-A-...-ATTG-...-AA-...-CA-GTCA-...-T-...-A-...-T-...-A-...-TTA-...-T-...-AGTG-...-G-...-T	8944
SMM.US.-.H9	A-A-G-...-G-GC-...-AAGA-...-C-CG-...-ATC-...-A-...-ATTG-...-AA-...-CA-GTCT-...-T-...-A-...-T-...-A-...-TTA-...-T-...-AGTATT-...-T	8976
STM.US.-.STM	A-A-AG-A...-G-...-ACAT-...-AAGA-...-C-AC-T-...-GGA-...-TTG-...-AA-...-C-...-GTCA-...-T-...-A-...-AGC-...-A-...-TTA-...-T-...-AGTG-...-G-...-T	9144
SAB.SN.-.SAB1C	..-A-C-...-A-GG-...-TGCT-...-AC-C-...-G-...-A-...-T-...-TTG-...-AA-...-CT-...-G-...-A-...-TTA-...-T-...-AGTG-...-G-...-A-G	9301
TAN.UG.-.TAN1	..-C-...-A-G-...-A-GA-...-G-...-A-...-A-...-T-...-TTG-...-A-...-T-CTC-...-A-...-A-...-G-...-CA-...-TA-TGG-...-A-C-...-AG-...-G-G	9111
VER.KE.-.AGM155	..-A-...-A-G-...-AAGA-...-G-AC-...-A-...-A-...-ACTT-...-A-...-T-...-TCG-...-A-...-T-...-A-...-TTA-...-T-...-G-...-T-...-G-...-A-T	9101
VER.KE.-.9063	..-A-...-T-...-A-A-...-AAGA-...-A-G-...-G-A-...-ATT-...-A-...-G-...-CT-...-TCG-...-A-...-T-...-A-...-TTA-...-T-...-G-...-G-...-GA-...-T	9124
VER.DE.-.AGM3	..-A-A-...-A-G-...-AAGA-...-G-AC-...-A-...-C-...-T-...-ATT-...-G-...-CT-...-TCG-...-A-...-T-...-A-...-TTA-...-T-...-G-...-C-...-G-...-GA-...-T	8622
VER.KE.-.TYO1	..-A-...-G-G-...-AAGA-...-C-...-G-A-...-T-...-ACTT-...-A-...-G-...-T-...-TCG-...-A-...-T-...-A-...-TTA-...-T-...-G-...-T-...-G-...-A-T	8578
COL.CM.-.CGU1	A-GAT-...-A-...-G-TGT-...-GAC-G-...-AG-...-C-...-AGGC-...-CATT-...-CT-CTG-...-G-...-CA-CTCTTG-...-G-...-T-...-G-...-A-...-T-...-AGAA-...-CTGAG-...-G-T	8356
DEN.CD.-.CD1	..-A-C-...-ACGG-...-A-GAC-...-AC-GCAG-...-CC-...-C-...-T-...-A-ATATGA-...-A-...-GTCT-...-T-...-G-...-A-...-C-...-G-ATA-...-TT-...-T-...-AGT-...-C-...-G-...-T	9111
GRV.ET.-.GRI_677	..-A-C-...-T-...-GRI-...-GA-C-...-C-...-AG-...-C-...-T-...-ATT-...-A-...-G-...-CT-...-TC-...-T-...-A-...-TA-...-T-...-TG-...-G-...-G	8968
GSN.CM.99.CN166	ACCATCA-GC...-A-G-...-A-G-...-AC-T-...-GACCCC-...-C-...-A-TCATG-...-G-...-CTCT-...-G-...-A-...-G-...-CA-GT-...-T-...-GTG-GGAC-...-G-...-C	8937
GSN.CM.99.CN71	ACCATCA-GC...-G-G-...-A-...-G-...-GC-T-...-GACCCC-...-C-...-ACTTATG-...-G-...-CTCG-...-G-...-A-...-G-...-CCA-GT-...-T-...-GTG-...-GAT-...-C	8934
DRL.-.FAO	..-A-...-ATGT-...-ACGC-...-ACGC-...-C-...-A-ATCTA-...-C-...-CT-...-CT-...-G-...-A-...-TTGGTGG-...-T-...-TC-...-G-...-G	8829
RCM.GA.-.GAB1	..-A-A-...-G-A-...-A-...-AC-T-...-G-...-CT-...-AA-...-CA-GTCT-...-T-...-A-...-A-...-TTA-...-TGGAGTATC-...-T	8651
RCM.NG.-.NG411	..-A-G-...-G-G-...-A-A-...-C-CA-...-G-...-T-...-ATT-...-AA-...-ATCT-...-T-...-A-...-TTA-...-TT-...-AG-...-G-...-AGG	8734
MND-2.-.5440	..-A-C-...-TAC-...-AAC-CAG-...-AG-GCCTGA-...-GCA-...-C-...-ACTTGA-...-CA-GTC-...-A-...-TTGGTGG-...-TA-...-G-...-G	8744
MND-2.CM.98.CM16	..-C-...-ATAT-...-A-...-TG-...-AG-...-GA-CCC-...-C-...-ACTTG-...-T-...-A-GCC-...-A-...-TTGG-...-T-...-TA-...-G-...-G	9048
MND-2.GA.-.M14	..AG-...-C-...-TTAT-...-CACACAG-...-G-GC-TGA-GCA-...-C-...-AAACTGA-...-A-GTC-...-A-...-TTGGTT-...-TAG-...-G-...-G	8989
MNE.US.-.MNE027	G-A-AG-A...-G-GG-...-A-GA-...-T-...-C-...-G-...-ATTG-...-A-...-A-...-GTCT-...-T-...-A-...-A-...-TTA-...-T-...-AGTG-...-T-...-T	8970
LST.CD.88.447	.....CCCT-C-T-...-A-TCTGA-...-C-CTCA-...-T-...-A-...-T-...-TA-TGG-...-G-...-G	8123
LST.CD.88.485	.....CCCT-...-T-...-A-TCTGA-...-C-CTCA-...-T-...-A-...-T-...-TA-TGG-...-G-...-G	8120
LST.CD.88.524	.....CTCT-C-T-...-AA-CTGA-...-C-CTCA-...-T-...-A-...-T-...-TA-TGG-...-G-...-GA-T	8123
LST.KE.-.lho7	.....CCCT-C-T-...-A-TCTA-...-C-CTCA-...-T-...-A-...-T-...-TA-TGG-...-G-...-G	9198
SUN.GA.98.L14	.....GCCA-A-...-T-...-CAGCTAA-...-CTCA-...-T-...-A-...-T-...-CTGGT-...-AG-...-CC-...-G-...	9269
MND-1.GA.-.MNDGB1	..-T-A-...-T-...-GTGT-...-GC-...-A-...-TT-C-...-GTGCCAT-A-...-A-ATCTGA-...-CT-CTCT-...-T-...-A-...-G-...-A-...-TA-TAG-AGG-...-G-...	8553
MON.CM.99.L1	C-GTT-...-A-GG-...-C-GA-...-T-...-GC-TC-CGACCT-...-GG-...-CTCATGA-...-G-...-C-CTCT-...-T-...-A-...-G-...-A-...-A-...-GT-...-T-...-GTG-...-GAC-...-T	8972
MON.NG.-.NG1	T-GCT-...-G-TCGC-...-AT-...-G-AC-T-...-GACCC-...-A-GG-...-ATCATGA-...-G-...-TCT-...-T-...-A-...-TAA-...-GT-...-T-...-GTGTWAGACC-Y-...-T	7609
MUS.CM.01.1085	..ACCT-C-...-G-GCGT-...-GAG-...-G-AA-CC-TGACCCC-...-C-...-CTGATG-...-G-...-A-GTCC-...-T-...-C-...-A-...-GT-...-CGT-AGTAC-GA-...-G-T	8943
DEB.CM.99.CM40	A-A-C-...-ACGC-...-C-...-T-CA-CC-...-GACCCC-...-A-...-T-...-TT-ATGA-...-CTACTCT-...-T-...-G-...-T-...-G-ATG-...-TT-...-T-...-TGCC-...-G-...-T	8758
DEB.CM.99.CM5	GA-A-C-...-G-T-...-G-...-AAG-...-C-AA-T-...-GGA-CCC-...-ATTGATGA-...-CTACTCT-...-T-...-A-...-T-...-G-ATA-...-TT-...-T-...-TGCT-...-T-...-T	8685
SYK.KE.-.KE51	CT-T-...-T-...-TTT-...-A-...-A-...-C-...-G-...-CT-A-...-ACTT-...-CA-...-CTC-...-T-...-A-...-AG-...-A-...-C-...-G-...-CA-GT-...-TT-...-GT-...-AG-...-T	8639
SYK.KE.-.SYK173	C-A-...-C-...-T-...-GTGT-...-AAC-...-AC-C-...-A-...-T-A-...-C-...-ATTG-...-AA-...-C-CTCT-...-T-...-A-...-A-...-C-...-G-...-CA-G-A-...-T-...-GTG-...-AG-...-G-T	8949









H1B.FR.83.HXB2	AGCTAGTACCAGTTGAGCCAGATAAGATAGAGAAGAGGCCAATAAAGGAGAGAACCACCAGC.....TTGTTACACCCGTGTAGCCCTGCATGGGATGGATGAC.....CCGGAG..	9327
Nef	K L V P V E P D K I E E A N K G E N T S . . . . . L L H P V S L H G M D D . . . . . P E .	Nef
H1A1.UG.85.U455	-----T---C-G-AG---G-A-T-C-GG-----A---T....C-AC-----A-AT---AA-----AG-A---T.....GAA---..	8778
H1B.US.90.WEAU160	-----G-A---G---G-----A-T---G-----A---T.....C-AC-----A-AT---AA-----AG-A---T.....GAA---..	9327
H1C.ET.86.ETH2220	-----C---AG-G-AG-G---AAT---G-----A-T-----C-G-----C-----A---G-T.....GAA---C..	8727
H1D.CD.84.84ZR085	-----T---GGT-G---A---C-G-----AG---A-T-----C-G-----A-CT---A---A---G---G.....GA---..	8875
H1F1.BE.93.VI850	-----T---AG-G-G---A---G-----A-T-----C-A-----CA---AA---A---G-T.....GAA---C..	8604
H1G.SE.93.SB6165	-A-----A-G-T---CAG-G---G-A---G-----A---T....C-A-----CA-CT---A---A---G---G.....GAA---C..	8769
H1H.CF.90.056	-----AA-T---C-GG---G---C-----G-----A---T....C-----CA-----A---A---G-T.....GAC---G..	8650
H1J.SE.93.SE7887	-----T---AGCG-AG---G-A-T---G-----A---T....C-----CA-AT---A---A---A---G-T.....GAA---A..	8625
H1K.CM.96.MP535	-----T---CAG-AG---A---CAG-----G---A-T....C-----CA-A-A---A---A---G-T.....GAAC---C..	8514
H101.AE.TH.90.CM240	-----C-A-AGAG-AG---G---A---C-----A---T....C-----CA---A---A---A---G-T.....GAA---A..	8895
H102.AG.NG.-.IBNG	-A-----A-G-T---CAG-G---G-A---G-----A---T....C-A-----CA-CT-T-AA---A---G---G.....GAC---T..	8854
H1N.CM.95.YBF30	-A-----T-GTCAG-T---AG-AG-----T---G-----ATGC-----C-C-----CA-AT-T-AA---AGCA---T.....GATC---T..	8917
H1O.BE.87.ANT70	-A-----GTCAGA---AG-GC---AGACTAGGA-TAC-TGTG-G-GGGCT...AATC-CC-G-T-A-CAT-TGCG---AT-T-A-T.....A-AC-T..	9417
H1O.CM.91.MVP5180	-A-G-----GTCAG---AG-GC---GAGACTGGG---TAC-A-TG-AGATGCT...AGTC-TC---T-A-CTT-TAAT---AGCT-G-T.....G-AC-C..	9440
CPZ.CD.-.ANT	--T-----A---AT---CCACCTGAT.....AT---GG-A---TA....C-CC-C-----CTT-TACA---AGAT-GA-----AC-T..	8766
CPZ.CM.-.CAM3	-A-G-T---C-GACA---AG-AG-G---C-----G-----T-T-A-T....C-AC-G---CA-CT---A---A---A---T.....G-AC-T..	8770
CPZ.GA.88.GAB2	-----G---ATCA---GGCAG---A---C-G-----G-T---AT...CTCC-A---T-CC-T-ACT---AT-T-G-T.....GAA---..	8738
CPZ.GA.-.CPZGAB	-A-----G-CC-GACAGAG-GC-G---C-A---G-----T---A-T....C---G-T---CA-AT-T-A---A---T.....GA---C..	9400
CPZ.US.85.CPZUS	-T-----C-TC-CACAGA---GG-AG---GC-----C-A---TA....C-C-G---CA-T---A---A---A---T.....GAAC---T..	9366
CPZ.TZ.-.TAN1	--T-G---G-A-C---CCAG-AGAT.....G-T-AGA---ATC---C---T-A-CCT-TAGC---A-CTACC---T.....A---T..	8998
H2A.GW.-.ALI	-----A---TGTCCAC-AGA-GG---A-C-G-GACT....CA-T-----C-C-----A-ACAAACAAGCA---CAT-----A-C-T..	9739
H2A.DE.-.BEN	-----T---A-A-TCTCAC-AGAG-C---AAG-G-G-CC---A-T-----AG---A-CACAAACAAGCA-ACAT---T.....GA-C-T..	9740
H2A.SN.-.ST	-----A---TGTCCAC-AGAG-G---T-A-G-G-GACT....CA-T-----AG-G-T---A-CACAAACAAGCA-T-T-----C-T..	9179
H2B.GH.86.D205	-----G---A---GTGCCAGCAGCCCGA-AGG-GG-G-A-ACC...CATT....C-AA-G---G-CACAGA-CTCCTCATG-----ATCC-T..	9656
H2B.CI.-.EHO	-----G---A-AA-TATGATAGCAGA-CC---A-G-GG-ACC...CATT-T....C-G-G-T---A-CACAGACTCC-CATG-----CC-C..	9627
H2G.CI.-.ABT96	-A-----CAC-GTCTC-G-TGAG-C-C---AGG-CG-GACT....CATT....G---T---A-CACAACTAGCCCATG-----G---..	9125
H2U.FR.96.12034	-----A---TG---A-G-T-A-C-C---AGG-G-GA-G...CTGTAT....C-G-GGGTT-A-CACAGACTCCG-T-CGA---G---ACTG---..	9221
MAC.US.-.239	-AT---C---T-AA-TGT-TCAG-TGAG-C-C---AGG-G-G-AG....CATTAT....AA-G-T---A-CTCAAACCTCCCA-TG-----TTG...	9697
Nef	K L V P V N V S D E G Q E D E E . . . . . H Y . . . . . L M H P A Q G S O W D D . . . . . P W .	Nef
SMM.SL.92.SL92B	-----G-G---A-CGTCTCG-TGAG-C-A-CAAT-G-E-GACC...CATT-----G---T---A-CCAGACAT---CA-CATTCA-----TTG...	9148
SMM.US.-.H9	-AT---WC---G---AA-TGTCTCAG-TGA-CTC---AAG-CG-GAC...CATT-----A-G-T---A-CACAACTTC-CA-TG-----CTG...	9180
STM.US.-.STM	-A-----A---TATGTCA-TGAG-C-C---AGG-G-G-T-GACA...CATTAT....C-G-G-T---A-CACAGACA---CA-TG-----TTG...	9351
SAB.SN.-.SAB1C	-A-G---T---G-C-TGTGAG-GAG-CTA---AAT-G-G-GAAC....CA-T-----C-C-G-T---CTCAAG-GCCTATGA-----G-CTG...	9505
TAN.UG.-.TAN1	--T-----G-G-A---TGTCTCCC-GA-CCC---ATG-G-G-.....CA-T-T....C-C-----CCCAG-TG-AT---GAAAGC-----TTG...	9312
VER.KE.-.AGM155	-----G---CC-TGC-G---GA-C---AA-ATG-G-A-G...CATT-----G-----A-CACAA---AAGACCC---TGGTATA.....AATC-T..	9311
VER.KE.-.9063	-----G-A---ATGC-G-AGAG-C---A-TG-G-GA...CATT-----GCG-T---CACAAA-AA-A-A-GACCCC---TGGCATC.....AGTC-T..	9334
VER.DE.-.AGM3	-AT-G-----A-C-TGC-G-GA-C---A-TTG-G-GA...CATT-----C-G-G-T---A-CACAAAG-AGG-AAGACCC---TGGAATC.....AACC-T..	8832
VER.KE.-.TY01	-----G---G---CTT-C-G-GA-C-C-AA-ATG-G-A...CATT-----C-AG-C-T---A-C-CAGA-GGA-AAGATCCA---TGTTATC.....AGCC-T..	8788
COL.CM.-.CGU1	-AT---CTACA-A---A-CT-GGAGAGT-T-CATA---GCT-T.....C-----C-----C-----C-----C-----C-----C..	8533
DEN.CD.-.CD1	-AT-G-C-G-G-GACTAT---GG-GA-T-T-CAAT-TG-G-A-.....ATT-T....C-C-G-TGACAGATATGA-GG-CAACAA-C-----TTG...	9315
GRV.ET.-.GRI_677	-----C---G-C-TGC-G-GAG-C-CGCAA-GG-G-GA...CA-T-T....C-A-G-T---A-CACAGA-GGG-AAGATCC---TGGAATA.....GATC-T..	9178
GSN.CM.99.CN166	GT---GCGG---A-CCA---AC-G-GACTCT---C-TGGGG-T-AT-AT-T...CTG....C-ACC-----CCTAT-A-GGAC---GT-A-----A-C-T..	9144
GSN.CM.99.CN71	GT---CGG---G-CCATGAC-G-GACTCT---C-TGGGG-T-AT-AT-T...CTG....AAC-T-C-CCTAT-A-GGAC-CAA-G-----ACC-T..	9141
DRL.-.-.FAO	-----G-C-T-G---GT---G-AGATCTG-TAAATC-GC-GT-TA-C-GG....C-----G---T-CAGCCAGA-GG-CCT-T-----A-A..	9033
RCM.GA.-.GAB1	-AT-G-G---A---TGT-TC-G-TGA-C-AGA-AAG-G---AG....CAT-T....C---T---A-CAGAAACAAG-----A-----ATG...	8855
RCM.NG.-.NG411	-AT-----A---TGTGAG-G-TGA-C-ACA-AG---G-GA-G...CATT-T....C-CC-G-T---A-CAGAAACTAG-----C-----CTG...	8938
MND-2.-.-.5440	-----G-C-G-G-CAGTGTGCCCCCTG---AGG-G-G-T-CA-C-GG....C-C-A-T---TCTCAA---GGAATCCA---A---A---A---..	8948
MND-2.CM.98.CM16	-----TTGC---G-G-CAGTGC-CCAGACC-G---AA---G-GT-TA-C-A...C-C---GT-CTCTCAG---GGAATCCA---A-A-----T-CTT...	9252
MND-2.GA.-.M14	--T---A-C---AC-AGT-CCACCAGACC-G---AA-C-GG-GT-TA-C-GG....C-C-TA---T-CTCTCAG---AGGAATCCA---A-A-----T-CAT-G..	9194
MNE.US.-.MNE027	-AT---C-T-AA-TGT-TCAG-TGAG-C-C---AGGG-G-G-AG....ATTAT....AC-G-T---A-CTCAAACCTCCCA-TG-----TTG...	9174
LST.CD.88.447	-----G-A-----G-CCTGT---G-TACAC---TC---TCAG-TTACAGA...CAGT-----C-C-G---T-AAGTCAG---GGG-TCCA---G-----CTG...	8330
LST.CD.88.485	-----G-A-----G-CCTGT---G-TACAC---TC---TCAG-T-CCAG...CAAT-----C-C-G---T-AAGCCAG---GGG-TCCA---G-----CTG...	8327
LST.CD.88.524	--T-G-AG---A-CATGT---G-T-TA---TTC---TCAG-T-AGAGA...CAAT-T....C---TT-AAGCCAGT---GGG-TCC-A-A-----CTG...	8330
LST.KE.-.lho7	--T---A-----G-CTTGT---G-TATA---TC---TCAG-T-AGAGA...CAAT-T....C-T-G---TT-AAGCCAG---GGG-TCC---A-----CTG...	9405
SUN.GA.98.L14	-----C---CACCATT---G-AGCAGG---TC-T-----C-TC-TT-TCAAGCCC-AC-G---TT-CAGCCAG-AAGGG-TT-AT---A-----TTG...	9476
MND-1.GA.-.MNDGB1	-----G-A-----GAGCAG-C-TAGG---T---TATGCAGCTAACA-TC-T....C-----T-CTCCAGT-ATG-CCTCA---A-T.....A-A..	8757
MON.CM.99.L1	GCT-GCGC---G-A-CCA-CACGG---GATTCC---AGGGAG-T-AG-CTT....CTCC-AACC---T-G-CCTAT-AAGGCC-----G-----CC-T..	9179
MON.NG.-.NG1	GC---GCCT---G-A-CCATCC-GG-AGACTCG---A-GGAG-T-AY---TT....CTCC-ACC---G-CCTA---A-GG-C---GAT---A-----CC-T..	7816
MUS.CM.01.1085	GT---GAG-GA---G-CCATGACGG---GGCTCT---AGGGAG-T-ACC-AT-....CTCC-AGAC-----CCTTT-A-GGGA-AGAA---A---ACC-T..	9150
DEB.CM.99.CM40	-A-G-G---A---AAT---CAG-CCC---TT-T-AG---G-T-AGAG....ATC-----TGA---CCCA---A-GGGAT-----A.....GACCCCCA	8964
DEB.CM.99.CM5	-----G---A---AT---CAG-TCCT---CT-T-AG---G-T-AGAGA---T....ATCC-C---TGA---CCCA---AAGGACAAGCA---A.....GACCCCCA	8891
SYK.KE.-.KE51	-AT-G-CC---TAA---AAT---AGG-TACTTG-TA-ATG-GG---ACACA....CTGA---A-G---A-CAGCAGGAGTG---GCTTCA---A.....GACCCACA	8842
SYK.KE.-.SYK173	-A-----T---AA---AT---GGGATGCTTG---TATG-GG-GCATACA....CTCC-CC-C-T---A-CCTTGGA-G---GTTGAG---ATGGGAGAACCCTCA	9158



PLV  
complete genomes

Table with columns for sample ID (e.g., H1B.FR.83.HXB2), sequence alignment (with amino acid translations above), and Nef protein length (e.g., 9417). The table lists numerous samples and their corresponding Nef protein lengths.



Accession	Sequence	Reference	Position
H1B.FR.83.HXB2	.....CATCGAGCTTGCTACAA.....	NF-kappa-B-II GGGACTTTCCG.CTG..	9448
H1A1.UG.85.U455	.....AAG.ACTG.CTGA--CA--AG---GAC-.....	.....A-.....	8907
H1B.US.90.WEAU160	.....-T-T-----T-----A-.....	.....A-.....	9448
H1C.ET.86.ETH2220	.....CACAGA--G-AC-TTCCGCC.....	.....A-.....	8851
H1D.CD.84.84ZR085	.....GACTGCTGACACAGAGATT.GCTG.ACACAGAA--ATC-.A-A-.....	.....A-.....	8746
H1F1.BE.93.VI850	.....CACAGAA--TTGC--AC-AG-.....	.....AC-.....	8892
H1G.SE.93.SB6165	.....CACAGAA--T.C-.G-A-.....	.....A-.....	8766
H1H.CF.90.056	.....CAAAGAA--TTTC-.AG-G-.....	.....A-.....	8746
H1J.SE.93.SE7887	.....CAAGGAA--TTTC-.ACT.....	.....A-A-.....	8598
H1K.CM.96.MP535	.....CACAGAA--TTGC--AC-.....	.....A-.....	9014
H101.AE.TH.90.CM240	.....CAAGGGAC--TT-CTGCTGACA-.....	.....-ATACT-.....	8975
H102.AG.NG.-.IBNG	.....AAACTGCTGACCTGAA--TTGCTGACACTGT.....	.....-A-----AG-AAAG	9044
H1N.CM.95.YBF30	.....CTGACACTGC.....	.....-AGAC.TG	9555
H1O.BE.87.ANT70	.....TAACAGGGCGATATAAAAC-GCT--CAGA--AGACTTGTGGTCT.....	.....-ACTAT..	9553
H1O.CM.91.MVP5180	.....TGCCAGCG.CCTG.CGCAGTAGA-A--GCTG-----TAA.....	.....8918	8903
CPZ.CD.-.ANT	.....CGCAACTGGCGCATGCGCTCTAGAA-TG-TGAGCCTGA--TTTTACAA.....	.....AA.....	8888
CPZ.CM.-.CAM3	.....TGACTCTG.....	.....AA.....	9539
CPZ.GA.88.GAB2	.....TGTAACCGC.GCAGGCGCAATA.A-A--GCTG--TG.....	.....A-----TAA.....	9503
CPZ.GA.-.CPZGAB	.....CTCAGAACTGCTGACAG.AGGAC.TTTT.....	.....GGACTC.....	9145
CPZ.US.85.CPZUS	.....AACAGC-GAG.AC---A.....	.....-AGAA... 9964	9964
CPZ.TZ.-.TAN1	.....AAAGCAAGAGGGATAACCATTTAGTGAATAACAGGAACAACCATACTTGGTCAAGGCAGGAAGTAACTACTGAA.....	.....CTAAG-AAC-.GC--AGG-TGCA.....	9967
H2A.GW.-.ALI	.....AAAGCAAGAGGGATAACCATTTAGTGAATAACAGGAACAACCATACTTGGTCAAGGCAGGAAGTAACTACTGAA.....	.....ACAGA-AAC-.GC--AG--TGCA.....	9406
H2A.DE.-.BEN	.....AAAGCAAGAGGGATAACCATTTAGTGAATAACAGGAACAACCATACTTGGTCAAGGCAGGAAGTAACTACTGAA.....	.....ACAGA-AAC-.GC--AG--TGCA.....	9406
H2A.SN.-.ST	.....AAAGCAAGAGGGATAACCATTTAGTGAATAACAGGAACAACCATACTTGGTCAAGGCAGGAAGTAACTACTGAA.....	.....GAAGG-AACTAGCA-AC--TGCA.....	9887
H2B.GH.86.D205	.....AAAGCAAGAGGGATAACCATTTAGTGAATAACAGGAACAACCATACTTGGTCAAGGCAGGAAGTAACTACTGAA.....	.....GAAGG-AACTAGC--A--TGCA.....	9858
H2B.CI.-.EHO	.....AAAGCAAGAGGGATAACCATTTAGTGAATAACAGGAACAACCATACTTGGTCAAGGCAGGAAGTAACTACTGAA.....	.....AGAAAGA.GG.AC-AGC--AC--GCA.....	9316
H2G.CI.-.ABT96	.....AAAGCAAGAGGGATAACCATTTAGTGAATAACAGGAACAACCATACTTGGTCAAGGCAGGAAGTAACTACTGAA.....	.....CAGAAAGG-AAC-AGC--AG--GCA.....	9417
H2U.FR.96.12034	.....ACCAGCAAGAGGCTCCAGTAAAGAATTGCTGA.....	.....Nef end \	
MAC.US.-.239	.....ACCAGCAAGAGGCTCTTAACATGGCT.....	.....GACAAGAAGG-AACTCGC--AA--GCA.....	9889
Nef	.....T A R G L L N M A.....	.....D K K E T R S.....	Nef
SMM.SL.92.SL92B	.....GCTAACAAGCCAAAACCCGAGAAGAAGATGGCG.....	.....GATAAGAAGG-AAC-AGC--AG--GCA.....	9346
SMM.US.-.H9	.....ACCAGCAAGAGGCTCTTAACATGGCT.....	.....-ACA--A.....	9373
STM.US.-.STM	.....ACCAGCAAGAGGCTCTTAACATGGCT.....	.....GACAAGAAGGAAA--AGCT-AG-.AC-GC.....	9543
SAB.SN.-.SAB1C	.....ACCAGCAAGAGGCTCTTAACATGGCT.....	.....CCACAAAACCACATCTATGGAGTTGTC.ATGGTGAT--CA--AAG-ACTGCTGACTGAGA.....	9673
TAN.UG.-.TAN1	.....ACCAGCAAGAGGCTCTTAACATGGCT.....	.....CCTAACCCGAGGCTTCCGGTTAAGCTAACTAGGCCGTTGCC.TAGGAGAT--CA-AAAG-AGTGTGATCAGCA.....	9493
VER.KE.-.AGM155	.....ACCAGCAAGAGGCTCTTAACATGGCT.....	.....CCGACCACAGGCTTGCCTTTCTGTTGCC.TAGGAGAT--CA--AAG-ACTGCTGAC.....	9489
VER.KE.-.9063	.....ACCAGCAAGAGGCTCTTAACATGGCT.....	.....CTAACCCGAGGCTTGGTTAAGCCGTTGCC.AGGGAGAT--CA--TGA-ACTGCTGACAG.....	9508
VER.DE.-.AGM3	.....ACCAGCAAGAGGCTCTTAACATGGCT.....	.....CTAACCCGAGGCTTGGTTAAGCCGTTGCC.AGGGAGAT--CA--TGA-ACTGCTGAC.....	9008
VER.KE.-.TY01	.....ACCAGCAAGAGGCTCTTAACATGGCT.....	.....CCAGACCACAAGCCTGCGGTTAGAACAT.CACC.ATGGAGAT--CA--AAA-ACTGCTGAC.....	9865
COL.CM.-.CGU1	.....ACCAGCAAGAGGCTCTTAACATGGCT.....	.....G-T-CTGA...GAC-GCA.....	8634
DEN.CD.-.CD1	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....T-TCA--ACTTTGCGGTT.....	9492
GRV.ET.-.GRI_677	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....-C-A.....	9303
GSN.CM.99.CN166	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....-A.....	9282
GSN.CM.99.CN71	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....-A.....	9279
DRL.-.-.FAO	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....TGTTGCAAGAGGAAGTGGTGGCAGTTCAAGCCAACACAGAAAGTTGTCATGGAGACAACTTCAGAAAGTGTGACTGAGA.....	9249
RCM.GA.-.GAB1	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....CCTTGGTAAACCAGGCAGAAAGATCTGCTGATGCAAAAGGGACTTTCCACTGGTGCATGCGCACTGGGGAA.....	9065
RCM.NG.-.NG411	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....CCTTGGAAACAGCGCATGCTACTGAGAAGATGCTGACACAGAA.....	9122
MND-2.-.-.5440	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....AAGCCTGACTGCTGCAAGAGGAAGTGGTGGCAGTTCTAGCTAACCACAAGGGTTGCTAAGCAACTGCTAACTCTGGA.....	9168
MND-2.CM.98.CM16	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....AAGCCAGAGTGTGCAAGAGGAAGTGGTGGCAGTTCTAGCTAACCACAAGGGTTGCTAAGCAACTGCTAACTCTGGA.....	9484
MND-2.GA.-.M14	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....AAATCAGACTGCTGCAAGAGGAAGTGGTGGCAGTTCTAGCTAACCACAAGGGTTGCTAAGCAACTGCTAACTCTGGA.....	9419
MNE.US.-.MNE027	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....AAATCAGACTGCTGCAAGAGGAAGTGGTGGCAGTTCTAGCTAACCACAAGGGTTGCTAAGCAACTGCTAACTCTGGA.....	9366
LST.CD.88.447	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....ACCAGCAAGAGGCTCTTAACATGGCT.....	9366
LST.CD.88.485	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....AAGTAGCCAACCCGAGACCCGACACATCTCTTGCAGCCTGGTTGCCAGGGCAACCAGGCTAGCGCATGCGC.....	8545
LST.CD.88.524	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....AAGTAGCCAACCCGAGACCCGACACATCTCTTGCAGCCTGGTTGCCAGGGCAACCAGGCTAGCGCATGCGC.....	8548
LST.KE.-.lho7	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....AAACCCGAGAACATCTCTTTCAGCCCGGTTGCCCTAGCAACCCGGGCTAGCGCATGCGC.....	8551
SUN.GA.98.L14	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....TAGCCAACCCGAAACCCGAGCACATCTCTTGCAGCCCGGTTGCTAAGGCAAC...CGGGCTAGCGCATGCGC.....	9626
MND-1.GA.-.MNDGB1	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....CATCTCTTGCATGCGCTAGGCAACG.....	9670
MON.CM.99.L1	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....CTAACCCGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	8979
MON.NG.-.NG1	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....TTGCTTGGCAACTGCTTAGCAACTGGACTGGGCTTGGCGC.CTAGG-A.....	8979
MUS.CM.01.1085	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....G-.....	9317
DEB.CM.99.CM40	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....C--T-----AGTA-TC	9062
DEB.CM.99.CM5	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....G-----AGT--CC	9131
SYK.KE.-.KE51	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....C--T-----AGTA-TC	9062
SYK.KE.-.SYK173	.....AACCAGAAAACCACATCTACTGAGAACTTGTGAGAGGCTCA.....	.....ACTCCATGGTGACAAGCTCGGCCAGGGGA.....	8987
			9312





PLV  
complete genomes



	NF-kappa-B-I	Sp1-III	Sp1-II	Sp1-I	TATA Box					
H1B.FR.83.HXB2	GGGACTTTCCAGGAGCG	GCCT	GGG	CGGGACT	GGGG	AGTGGC.GA	CCCTCAG.ATCCT	GCATATAAGCA	9519	
H1A1.UG.85.U455	GA-GA-G	T	GT	T-G	AGT	GA	TG	C	TAG	9879
H1B.US.90.WEAU160	G	AT	A							9519
H1C.ET.86.ETH2220	G	AG-AG-G	T							8919
H1D.CD.84.84ZR085										8975
H1F1.BE.93.VI850	A-G-CG-TCC	AGAG								8817
H1G.SE.93.SB6165	G	CG	A	AGT						8963
H1H.CF.90.056	G	A	T	GA						8837
H1J.SE.93.SE7887	GA-GAG			GT						8818
H1K.CM.96.MP535										8598
H101.AE.TH.90.CM240	GA-GTG	G	AGT	A	T	A				9086
H102.AG.NG.1.IBNG	GA-GTG	GT	T-G	A	AGT					9046
H1N.CM.95.YBF30	GGGACTTTCCGCCA	A-GTG	GT	GGAGT	CT					9118
H10.BE.87.ANT70	ACTGCTGACTGCG	T-G-AG-GA		CAGGGG	C-GT	C	G	T	A	9640
H10.CM.91.MVP5180	CTGACACTGCG	C-T-GAGG-AT	AAGG	CGG	T	C	G	T	A	9636
CPZ.CD.1.ANT		GGGA	TGA	CTATGGGC	GGCA	T-G	TTTC	G-CT	G	8976
CPZ.CM.1.CAM3		TTC	A-G							8974
CPZ.GA.88.GAB2	G-GG-GG		TTC	AA-GA	G-A-T-G	T	TTTG	A	G	8949
CPZ.GA.1.CPZGAB	G	AG	GTC	AGG	GT-G	T	TTTG			9612
CPZ.US.85.CPZUS	A-GACGTT	CAAGGGGGTG	GTC	A-G	GC	ACA	C	T	TTTA	9585
CPZ.TZ.1.TAN1		TGTGGGT	GT	ACTGG	GC	GACA	G	T	TTT	9205
H2A.GW.1.ALI	G-G	A-CCAG-GAAG-AC	ATGG	A-GAG	T-TGG	AA	C			10031
H2A.DE.1.BEN	G-G	A-CCA-GAGG	A-A	T-GA	GC	T-T	G-G-AA	CG		10036
H2A.SN.1.ST	G-G-T	CCAG-GAGG	A-A	T-GA	GC	C-T	G-G-AA	CG		9475
H2B.GH.86.D205	A-G-G	A-CA-T-GAGG	AG-A	A-GA	GC	T-T	T-G-AA	CG		9955
H2B.CI.1.EHO	G-G	A-CA-T-GAG	GAC	AT	GG	A-GAG	CA	GAG-AA	CG	9926
H2G.CI.1.ABT96	T-G-G	A-CA-GAG	GTA	CTATG	GG	ACT	G-T-G	ACG		9384
H2U.FR.96.12034	GA-GAGG	A-CCA-GAG	AGC	AA	GG	A-GAG	CCAT	G-G-AA	CG	9485
MAC.US.1.239	GA-G-TAC	G-GAGGTACTGGG	GAGGAGCC	GTGC	AACG					9956
SMM.SL.92.SL92B		G-GAG-TACAA	GGGAGGAGTCT	GG	CGGGG	A-GG	AAC			9414
SMM.US.1.H9	GA-G-T	C-G-GAGGTTCT	GGGG	AGGAGCT	G-G	AAC				9439
STM.US.1.STM	G-G-AACA	G-GAGGTACTGGG	AGGAAGTGGTG	GGAACG						9610
SAB.SN.1.SAB1C			TGG	A	ACT	G	CGG	ACT	GGAGT	9728
TAN.UG.1.TAN1		T-T-T	CG	AGGG	C	TA	CT			9548
VER.KE.1.AGM155	A-C		GACATGGG	C	TAC	G	A	T	GCTTTA	9556
VER.KE.1.9063			GGC	CCAT	G-G	TAC	G-G	AG	TG	9575
VER.DE.1.AGM3	C-G	TCATG	G-G	TACGGG	A-TGG	CTTTA				9075
VER.KE.1.TY01	A		GG	ACAT	G-G	TCC	G-G	AG	TG	9032
COL.CM.1.CGU1	A-TT	GAGG		GACCT	GC	GA	G	AGGGCGTG		8706
DEB.CD.1.CD1	GTTTGC	TAGCAAC	C-GGGC	ACTCA	GG	AGGGCCTGGGAGGTCT				9577
GRV.ET.1.GRI_677		CA		GTGGTG	ATCGGA	CGG	ACA	GGCG	TACTGGGAGTG	9385
GSN.CM.99.CN166	GTCACCATGACTACGGGG	C-TT-CT-AG	CAACCGGGCGGACTC	TGG		GAGG	ACTGG	A-T	CGGGGAGT-GCTCTA	9387
GSN.CM.99.CN71	GTTACCATGACTACGGGG	C-TT-CT-AG	CAACCGGGCGGACTC	TGG		GTGG	ACTGG	A	C.GGGAGT-GCTCTA	9383
DRL.1.FAO	CATGCGCTCTGTGGCAG	T-A	TG-GT	TGGG		AGTGGTCTG				9327
RCM.GA.1.GAB1	CGTGGGAGGGGAGTGGTTCAG									9109
RCM.NG.1.NG411	CCTGCGCGTGGTGTA	ACT	A	T		GGAG	GT	TT	TCCT	9201
MND-2.1.5440	CATGCGCACTGGACT	G	A	ACT	ACG	CGGG		TGGG	AGTGGCCAG	9246
MND-2.CM.98.CM16	CATGCGCACTAGACT	G	T	CA	ACT	A				9518
MND-2.GA.1.M14	CATGCGCACTGGACT	G	T	ACT	A	TC	GG		TGGG	9497
MNE.US.1.MNE027		G	GATGTCAT	GGGGGGTACTGGG	A	GAG	T	T	G	9433
LST.CD.88.447	GCGG	G	AGG	G						8580
LST.CD.88.485	GCGG	G	AGG	CA						8581
LST.CD.88.524	GCGG	G	AGG	G						8586
LST.KE.1.lho7	GCGG	G	AGG	G						9661
SUN.GA.98.L14	G	A	G	A	GCT	TA				9705
MND-1.GA.1.MNDGB1		AA	ACAGG	AGGG	A	GC	C			9021
MON.CM.99.L1	GTCACCATGACTACGGGG	C-TT-CT-CG	CAACCGGGCGGTC	CAAAAGGA		T	T	GGT	G	9421
MON.NG.1.NG1	GTTGCCCTGACTACGGGG	C-TT-C	AG	CAACCGGGCGGACTC	TGG	GAGG	ACTGG	A	TACGGGAGT-GCT	8059
MUS.CM.01.1085	GTCACCATGACTACGGGG	C-TT-CT-AG	CAATCGGGCGGACTC	TGG		GTGG	ACTGG	C	TACTGGGAGT-GCT	9393
DEB.CM.99.CM40	TAGGAAATATCATGGCAACC	GGGGAGG	CCT	G		G	TACGG	GA	TG	9208
DEB.CM.99.CM5	AAGGAGACTCCATGGTT	AC	GGGGAGG	ACT	G		CG	TACGG	GA	9139
SYK.KE.1.KB51	AGACAGAGACCCGCC	GGGGA				GGA	TCGG	GC	GG	9050
SYK.KE.1.SYK173						GGA	CCTG	GC	GG	9354



	3'LTR U3 end \ / 3'LTR R repeat start		
	TAR element	stem	bulge loop stem
H1B.FR.83.HXB2	G..CTGC...TTTTTG.CCTGT...AC..TGGG.TCTCTC..TGGTTAGA..CCAGATCTGAGC.CTGGGAG.CTCTCTGGC...TAA		CTAG..... 9590
H1A1.UG.85.U455	-----C-----	-----T-----	-----C-----G-----G----- 9049
H1B.US.90.WEAU160	-----C-----	-----T-----	-----G----- 9590
H1C.ET.86.ETH2220	-----C--T-----	-----C-----	-----A-G-----T-----G----- 8990
H1D.CD.84.84ZR085	-----C-----	-----A-----	-----T-----G-----G----- 8975
H1F1.BE.93.VI850	-----C-----	-----T-----	-----G-----G----- 8888
H1G.SE.93.SE6165	-----C--C-----	-----T-----	-----T-----G-----G-----AG----- 9034
H1H.CF.90.056	-----C--T-----	-----T-----	-----G-----A----- 8908
H1J.SE.93.SE7887	-----C--T-----	-----T-----	-----G-----A----- 8889
H1K.CM.96.MP535	-----C-----	-----T-----	-----G-----A----- 8598
H101.AE.TH.90.CM240	-----C--C-----	-----T-----	-----G-----A----- 9156
H102.AG.NG.-.IBNG	-----C--C-----	-----T--C-----	-----G-----GGA----- 9117
H1N.CM.95.YBF30	-----C--C--T-----	-----T--C--G-----	-----TA--T-----ATAT--G----- 9182
H1O.BE.87.ANT70	-----C--T-----	-----C-----GG--TAGAG--	-----G-----C-----CT-----CTAGC----- 9711
H1O.CM.91.MVP5180	-----CC--T-----	-----C-----TAG--TAGAG--	-----G-----C-----C-----CT-----CTAGC----- 9707
CPZ.CD.-.ANT	-----GC--T-----	TA..A--C-----G..CCTAGGAG..	-----A-----TA.AT-TTT-----GGCAG----- 9053
CPZ.CM.-.CAM3	-----C--C--T-----	-----T--C-----	-----G----- 9045
CPZ.GA.88.GAB2	-----C--C--T-----	T-----T-----A--G--	-----T-----G--C----- 9019
CPZ.GA.-.CPZGAB	-----C-----	-----T-----CAC--G--	-----T-----AC-----G-----TG--A----- 9682
CPZ.US.85.CPZUS	-----C--C--T-----	-----C--C-----	-----AA-----C-----G-----TGTA----- 9655
CPZ.TZ.-.TAN1	-----GC--CT-----	A..A-----TG.CCTAATCTG.	-----T-----AG--G-----GCTGGC----- 9280
H2A.GW.-.ALI	TAC-C--T..GC--GCAT----	ATTCA-TC.G--GC.G-AGAG-CTGG--	-----C-----GT-----CC...AGC-CTA...GCAGGTAGAGCCTGGGTGTTC-- 10134
H2A.DE.-.BEN	TAC-C--...T-C--GCAT----	ATTCA-TC.G--GC.G-AGAG-CTGG--	-----C--A--GT-----CC...AGC-CTA...GCAGGTAGAGCCTGGGTGTTC-- 10139
H2A.SN.-.ST	TAC-C--...TAC-CGCAT----	ATTCA-TC.G--GC.G-AGAG-CTGG--	-----C-----GT-----CC...AGC-CTA...GCAGGTAGAGCCTGGGTGTTC-- 9578
H2B.GH.86.D205	TAC-C--...T-C-CGCAT----	ATTCA-TC.G--GC.G-AGAG-CTGG--	-----C-----GT-----CC...AGC-CTA...GCAGGTAGAGCCTGGGTGTTC-- 9978
H2B.CI.-.EHO	TAC-C--...T-G-GCAT----	ATTCA-TC.G--GC.G-AGAG-CTGG--	-----C-----GT-----CC...AGC-CTA...GCAGGTAGAGCCTGGGTGTTC-- 10029
H2G.CI.-.ABT96	TAG--...A--CGCT----	ATTCA-TC.G--GC.G-AGAG-CTGG--	-----C--C--GT-----CC...AGC-CTA...GCAGGTAGAGCCTGGGTGTTC-- 9487
H2U.FR.96.12034	TAC-C--...TC--GCAT----	ATTCA-TC.G--GC.G-AGAG-CTGG--	-----C-----GT-----CC...AGC-CTA...GCAGGTAGAGCCTGGGTGTTC-- 9588
MAC.US.-.239	TCA-----A--C-CT-----	ATTCA-TC.G--GC.G-AGAG-CTGG--	-----C-----GT-----CC...AGC-CTA...GCAGGTAGAGCCTGGGTGTTC-- 10059
SMM.SL.92.SL92B	TAC-C--...T--GCC-CT-----	ATTCA-TC.G--AC.G-AGAG-CTG-----	A--C-C-A--GT-----C...AGC-CTT...AGTAGGTAGAGCCTGGGTGTTC-- 9518
SMM.US.-.H9	CAA-----A--C-CT-----	ATTCA-TC.A--GC.G-AGAG-CTGG-----	C-----GT-----CC...AGC-CTA...GCAGGTAGAGCCTGGGTGTTC-- 9542
STM.US.-.STM	CAA-C--...T-A-A-CT-----	ATTCA-TC.A--GC.G-AGAG-CTGG-----	C-----GT-----CC...AGC-CTA...GCAGGTAGAGCCTGGGTGTTC-- 9713
SAB.SN.-.SAB1C	-----GCG-----	A..A..ACA-----GC.-T-GAG-CTG-----	-----T..T-----TAAGTCTGAACCAGCTTAGAGCCTGGGTGTTCGCTGGGGAA...GG 9836
TAN.UG.-.TAN1	-----CGC--G-T-----	A--CA-----T.ACTAGG-----	C-----T--G--T...G-GCCTA..ACCTGGTCTGGCCATCCAGGGG--A... 9646
VER.KE.-.AGM155	-----CGC--G-T-----	A--CA-----T.ACTAGG-----	C-----T..T--G--T...GCCTA..ACCTGGT.TGGCCA.CCAGGGG--A... 9652
VER.KE.-.9063	-----CGC--G-T-----	A--CA-----T.ACTAGG-----	C-----T..T--G--T...GCCTA..ACCTGGT.TGGCCA.CCAGGGG--A... 9671
VER.DE.-.AGM3	-----CGC--G-T-----	A--CA-----T.ACTAGG-----	C-----T..T--G--T...GCCTA..ACCTGGT.TGGCCA.CCAGGGG--A... 9171
VER.KE.-.TY01	-----CGC--G-T-----	A--CA-----T.ACTAGG-----	C-----T..T--G--T...GCCTA..ACCTGGT.TGGCCA.CCAGGGG--A... 9128
COL.CM.-.CGU1	AG.AGC-T..CGAGAGCTT.C-AGG.C		9728
DEN.CD.-.CD1	-----C-G-AA-AGTC..GGTTGC-A-----	CA-CGGCCTGGG--CC-----A.GGCCTTAG-AGCGCCTGGGAGCCCTTAAGTATTGCAACAAGTACC	9679
GRV.ET.-.GRI_677	-----CGC--G-T-----	A--CA-----T.ACTAGG-----	C..A-----T..T--G--T...GCCTA..ACCCGGT.TGGCCA.CCGGGG--A... 9481
GSN.CM.99.CN166	A..AC--T..C.....	AG-G.AC....TCGGAGTC.....	9414
GSN.CM.99.CN71	-----AC--T..C.....	AG-G.AC....TCGGAGTC.....	9410
DRL.-.-.FAO	-----CT--T..T.AC-GC.A-----	AAAC-----CC--GAG-CTA-----	C-----TGT-----TGAGT.C..TGCTGAAGGAGGCTACCTGCTCGAG... 9425
RCM.GA.-.GAB1	-----T..C.-GCGC.T-----	A-AACG-----CC--GAG-CTA--G-----	T.GT-----TAAGT.C..TCTAGG.AACTCCAGCTTGGACCTG...G 9208
RCM.NG.-.NG411	-----A--C-CT-----	A--CA-----CC--GAG-CTA-----	TTGT-----TAAGT.C..TCTAGG..AGACTAGCTTGGACCTG...G 9299
MND-2.-.-.5440	-----CT--T..T.GCCGC.------	AAGT-AG-----CC--AGAG-CTG-----	C-----TTGT-----TGAGTTC..TTGAAGGAGTGCCTGCTTGTAGC-C... 9347
MND-2.GA.-.M14	-----C-C--T..T.GCCGC.T-----	AT--A-----CC-	9529
MNE.US.-.MNE027	TCA-----A--C-CT-----	ATTCA-TC.G--GC.G-AGAG-CTGG-----	C-----GT-----CC...AGC-CTA...GCAGGTAGAGCCTGGGTGTTC-- 9536
LST.CD.88.447	-----T..C-----		8588
LST.CD.88.485	-----T..C-----		8589
LST.CD.88.524	-----T..C-----		8594
LST.KE.-.lho7	-----TCTCCGC--T..TCGGGAG.....	CT-GCTCT-GGGACCAG--TCGCTCTGGGT-AGCA-A--GGTCAGCTAAGCAGTCAGCCTGGGAGGCTACA.....	CTTAGCACCCA 9771
SUN.GA.98.L14	-----TCTTCGCC--T..TCGGGAG.....	GCTCA-AGGATAGG-GACGACCCTGAGA--TCTCAC--AGA--TGG.CTTGTGCTCTCTGAGTGCAGCCTGTGTAAG-ACCTCACCCGA	9821
MND-1.GA.-.MNDGB1	-----A..T.-CGC.T-----	TCGG-A-----ACTACAGAG-CTAA.....	TTGTATCTCTGAGC.-GATCCCTTAGAGCAAGACCAGAGT...CCTGA... 9112
MON.CM.99.L1	-----AC--C..C.....	AG-T.ATCCT--GGAGTC.....	9448
MON.NG.-.NG1	-----AC--T..C.....	AGMT.ATCCGC-GGAGTCA-C-CTGCTCT-CACAGC-GC-TGG-A-CCCT-GGAGCTAGGTAA-GCGGCCTGGGAGCCCTTGCTAGAGAGTGCCTGCTCTCG	8172
MUS.CM.01.1085	-----AC--T..C.....	AG-T.ATCCC--GGAGTC.....	9420
DEB.CM.99.CM40	-----A--T..C.....	AGCT-TCCA.....	9227
DEB.CM.99.CM5	-----A--T..C.....	AGCT-TCCA.....	9158
SYK.KE.-.KE51	-----A--T..C.....	AGCT-TCC.....	9068
SYK.KE.-.SYK173	-----ATGCTCTGTAGC-ATC-GC-GAGTTG-AACTA-C-----	CA.....	C-GC-----G-C--AGA.GACTGGTAGAGCGGCC..GGAGCCCTTGCCAGTAAAGGCTAG 9459





PLV  
complete genomes



	TAR element     GGAACCC...ACTGCTTAA... 		
H1B.FR.83.HXB2	.....GGAACCC...ACTGCTTAA.....	.....GCC.TC.....	9611
H1A1.UG.85.U455	.....	.....	9070
H1B.US.90.WEAU160	.....	.....	9611
H1C.ET.86.ETH2220	.....	.....	9011
H1D.CD.84.84ZR085	.....	.....	8975
H1F1.BE.93.VI850	.....	.....	8903
H1G.SE.93.SE6165	.....A.....	.....	9056
H1H.CF.90.056	.....	.....	8929
H1J.SE.93.SE7887	.....	.....	8910
H1K.CM.96.MP535	.....	.....	8598
H101.AE.TH.90.CM240	.....A.....	.....	9178
H102.AG.NG.-.IBNG	.....	.....	9138
H1N.CM.95.YBF30	.....	.....	9182
H1O.BE.87.ANT70	.....T.....G-CG.....	.....CG.....	9732
H1O.CM.91.MVP5180	.....T.....G.....	.....CG.....	9728
CPZ.CD.-.ANT	.....-AG.....	.....-T.....	9068
CPZ.CM.-.CAM3	.....A.....	.....	9066
CPZ.GA.88.GAB2	.....	.....	9039
CPZ.GA.-.CPZGAB	.....	.....	9703
CPZ.US.85.CPZUS	.....	.....	9675
CPZ.TZ.-.TAN1	.....TA-AG.....G.....	.....CG.....	9302
H2A.GW.-.ALI	.....GCTA--CT-TC.ACC-G---GGCCGGCACTGGGCA..GA.....	.....CGGCTCCACGCT.TGCTTGC...TTAA.AGAC-TC-T.....	10204
H2A.DE.-.BEN	.....GCT--CT-TC.ACC-G-A---GGCCGGTACTGGGCA..GA.....	.....CGGCTCCACGCT.TGCTTGC...TTAA.AGAC-TC-TC.....	10210
H2A.SN.-.ST	.....GCTA--CT-TC.ACC-G---GGCCGGCACTGGGCA..GA.....	.....CGGCTCCACGCT.TGCTTGC...TTAAAAGAC-TC-T.....	9649
H2B.GH.86.D205	.....	.....	9978
H2B.CI.-.EHO	.....GCTA--CT-TC.ACC-GCA---GGCCGGTGCTGGGCA..GAC.....	.....GGCTCCACGCT.TGCTTGCTTAA.AA..CCCT-T.....	10099
H2G.CI.-.ABT96	.....GCTA--CT-TC.ACC-GCA---GGCCGGTGCTGGGCA..GAGC.....	.....GGCTCCACGCT.TGCTTGCTTACAAC.GACCT-T.....	9560
H2U.FR.96.12034	.....GCTA--CT-TC.ACC-AAAG.....	.....	9607
MAC.US.-.239	.....GCTA--CT-TC.ACC-GCA---GGCCGGTGCTGGGCA..GAG.....	.....TGACTCCACGCT.TGCTTGC...TTAA.AGCC-TC-TC.....	10131
SMM.SL.92.SL92B	.....ACTAA-CT-TCAACC-GCAACAGGCCATTGCTGGGTA..GAC.....	.....GGCTCCACGCT.TGCTTGC...TTGA....-TC-.C.....	9585
SMM.US.-.H9	.....GCT--CT-TC.ACC-GCA---GGCCGGTGCTGGGCA..GAG.....	.....TRGCTCCACGCT.TGCTTGC...TTAA.AGAC-TC-TC.....	9614
STM.US.-.STM	.....GCTA--CT-TC.ACC-GCA---GGCCAGTGCTGGGTA..GAG.....	.....TGCTCCACGCT.TGCTTGC...TTAT.ATAT-TC-TC.....	9785
SAB.SN.-.SAB1C	CAAGCCTAAGA-C--T-TGGTG-GGT---GGTAGACCCTCGCC.....	.....TGGCGACTGGCCATTGCCAGTAGCAG.AGACTCC.....GCTTGCTTGCTTG..ATTG-C-.....	9933
TAN.UG.-.TAN1	.....-AGC-T-.....CT-----GAAGGCTAT.....	.....TG.....	9674
VER.KE.-.AGM155	.....--CT--.....TTG----GAA.....	.....AG-.....	9674
VER.KE.-.9063	.....--CT--.....TTG----GAA.....	.....AG-.....	9693
VER.DE.-.AGM3	.....--CT--.....TTG----GAA.....	.....AG-.....	9193
VER.KE.-.TY01	.....--CT--.....TTG----GAA.....	.....AG-.....	9150
COL.CM.-.CGU1	.....	.....	8728
DEN.CD.-.CD1	.....	.....	9679
GRV.ET.-.GRI_677	.....--CT--.....TTG----C-TA.....	.....TAG-.....	9505
DRL.-.-.FAO	.....CCCCG--TGTTTC.GCAGGGC---GGCCTTTGGGCTC.TGGGTAGCTCGGTGAGGAGGCTTCT.GGAGGTAGATCACTAGGGGCTGCTCTGG..TCTTGGCTTGTTAGACCTT-GCTACTTCTAGT	.....	9545
RCM.GA.-.GAB1	GTGTTTCGCTG-TG.T-TCTGAACAG---GCT.GGGGTGCCTCTCGCTCTTCGGGTAGACCCAGTTGAGG....CTCGGCCGGCCTCAACGGG..AGAGATC.....	.....	9304
RCM.NG.-.NG411	GTGTTTCGCTA-T.T-TCTGAA.GG---GCTGGGAGCGCCTTGGCTGGGTCTTGGTAGACCTTCAGCTGGG....TACTTGGCCAGTCCAGTGGGAGTAG.....	.....	9394
MND-2.-.-.5440	.....TGGGCG-TTCG-AGGCCCTG---GT....GGCTCTGGGTAGCTCGTCAAGTGTCTGGAAAGGCTTACGGGAGGACACTCTTGCTTGGTCTTGATAGACCTTAGCAATC--T-AGGCCCTGGAG	.....	9469
MNE.US.-.MNE027	.....GCTA--CT-TC.ACC-GCA---GGCCGGTGCTGTGCA..GAG.....	.....TGCTCCACGCT.TGCTTGC...TTAA.AGAC-TC-TC.....	9608
LST.CD.88.447	.....	.....	8588
LST.CD.88.485	.....	.....	8589
LST.CD.88.524	.....	.....	8594
LST.KE.-.lho7	CCTGGCCTGATC-G--TAGGG-GCTAAG.....	.....GGCTCCTAGCTTGTACTTTAAA..GCCTTC.....	9829
SUN.GA.98.L14	GCCTGATCAGCTTGGGGAGCTAGA.....	.....GGCTCC...TTGCTTTGCAGTACAAGCCTTC.....	9873
MND-1.GA.-.MNDGB1	.....GTGACT.GGGT.CTG-GCA..CCTCACTCGGGGCTG.....	.....ATCACCTCGAGGTAGTGGAA....CTCCTTGT.TGCTTGCTAT.....TGCTTTC.....	9190
MON.NG.-.NG1	GCTCGCTTGCT--TAGG.CTTTACCAGCCCAATAAAGTGCTTGAAGCAA.....	.....	8260
SYK.KE.-.SYK173	TGAGGTCCGG-TTTGG-C.GCCCGGT--CGGCAGCTCTCCGTT..GCTT.....	.....GCTTTGCTTACG..CTGA..TCAG-CC.A.....	9530



TAR elements of HIV-2 and SIV have extra stem-loop, see Berkhout, Nucleic Acids Res 20(1):27-31(1992)

	Poly-A signal	3'LTR R repeat end	3'LTR U5 start		
H1B.FR.83.HXB2	AATAAAG...CTTGCC	TTGAG	TG.CTTCAA.GTAGTG	..TGTGCCCGTC.TG.TT..GTG.TGACTCT...G 9668	
H1A1.UG.85.U455				..GTG--G-.....A-T..... 9128	
H1B.US.90.WEAU160				..... 9668	
H1C.ET.86.ETH2220				..... 9031	
H1D.CD.84.84ZR085				..... 8975	
H1F1.BE.93.VI850				..... 8903	
H1G.SE.93.SE6165		AA-CC		..... 9074	
H1H.CF.90.056			T	..... 8953	
H1J.SE.93.SE7887			G-CGCATGC.AAGCC	..... 8943	
H1K.CM.96.MP535				..... 8598	
H101.AE.TH.90.CM240			A	..... 9203	
H102.AG.NG.-.IBNG				..... 9196	
H1N.CM.95.YBF30				..... 9182	
H1O.BE.87.ANT70			AG	..... 9754	
H1O.CM.91.MVP5180			AGAAGC.AGT	---CT-A.---CAACC.CTGG-G...C 9785	
CPZ.CD.-.ANT				..... 9068	
CPZ.CM.-.CAM3			A-G.-C-	---TA-.CA.-C.A.-A- 9123	
CPZ.GA.88.GAB2			CTC-A-	---T-A-...C.A.-A- 9096	
CPZ.GA.-.CPZGAB			TGTAT-G.-C-	---TA-.A.-C.A.-AC- 9761	
CPZ.US.85.CPZUS			CT-A-	---TA-.C.-C.A.- 9731	
CPZ.TZ.-.TAN1			A-TG-T-	..... 9326	
H2A.GW.-.ALI		AA	AGA.A-CAGTT.AAG	---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG...C 10269	
H2A.DE.-.BEN		AG	AGA.A-CAAGTT.AGT	---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG...C 10275	
H2A.SN.-.ST		AG	AGA.A-CA	..... 9672	
H2B.GH.86.D205				..... 9978	
H2B.CI.-.EHO		A	TA	---AGA.A-CAAGAC.AGT	---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG...C 10164
H2G.CI.-.ABT96		C.T.GC-AAA	AGA.A-TAAG.C	---AGT---TGTGT..... 9599	
H2U.FR.96.12034				..... 9607	
MAC.US.-.239		AT.T	AGA.A-TAAG-T	---GT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG...C 10196	
SMM.SL.92.SL92B		G	AGA.A-CAAGTC	..... 9613	
SMM.US.-.H9		G	AGA.A-TA	..... 9638	
STM.US.-.STM		AA.T	AGA.A-TAAG-.A	---GT-G-TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG...C 9850	
SAB.SN.-.SAB1C		T	TT-A.-AAGT-AAGCA	---AGTGTGTG-CATT-AT-CCTCA-GAC-AC-CTGGT 10002	
TAN.UG.-.TAN1		TG.C	A-T-ATCT	---CAAGTGCCT-ATTG-.GCGCCCTTCC.-C...G... 9736	
VER.KE.-.AGM155		TG.CA	AGA.GCTTA-TG	---AGTCAAGTGCCT-ATTAA.---CC-CAC-C.-TGAA-G... 9741	
VER.KE.-.9063		AC.T	TG.CA	---AGA.GCTTA-TG	---AGTCAAGTGCCT-ATTAA.C-CC-CAC-C.AAG-AGG... 9762
VER.DE.-.AGM3		T	CG-TG.C.A	---AG..CAGAGA-TT---GAGAGACAT---T-GT.GCA-CCA-CA.CA-G-GA... 9261	
VER.KE.-.TY01		TG.CA	AG	..... 9170	
COL.CM.-.CGU1				..... 8728	
GRV.ET.-.GRI_677		CC.TGCT-GC	AGT.C-CTA-ATT	...G-A-TCAA---T-A-T.GCT.GC-CC.GAG-CTCTAGA 9570	
DRL.-.FAO	GGCCATAGAAGTAGGGGAGTAGGCACCGCTTGCTTATCCTAATTGCTC	C.TACC-ATT	TAGA-GATGA-TAAGTGT	ACTC-TGTC-T-TGTTT-T.A-TGAGA.CTCTGG-TA...C 9664	
RCM.GA.-.GAB1	ACCGCTTGCTT.ATAGCCTTG...AA.GCTC	C	ATGCCAG-AGTTTACTG-AAGCAAGT	---TGCCGTGTTTAC.-CTCAGCAGTTAACGA-CTG- 9403	
RCM.NG.-.NG411	GCT.CCGCTTGCTTTCTGG...TTAATTGCTC	C	TGCATGCTCTTAG	---A...CCAGAATCAAGT---TGCCCATTTTA---CTCAGT-AAAC-----G- 9495	
MND-2.-.5440	ATTGTG...GGATTGACTACCGCTT.GCTTGCT...ATTATTGCCC	C.T-AC-TGA	TTAGAAT-CAA-TAAGAGT	ACTCGTCCA-T-TG.TC-TCAAT.AGAAACTCTGG-TA...C 9581	
MNE.US.-.MNE027		AT	AGA.A-TAAG-C	---GT---TGC-CCCAT-TCT.CCTAGTC-CC 9664	
LST.CD.88.447				..... 8588	
LST.CD.88.485				..... 8589	
LST.CD.88.524				..... 8594	
LST.KE.-.lho7		AA-TTGCTCATTG-T	AAA-TACAAGCTAGT	---CTCA-AGTAGTC-T-TTC-CCTTC-CCC... 9897	
SUN.GA.98.L14			AATAAAGCCTGCCT	---TTTACT.....GCAAGCAA 9905	
MND-1.GA.-.MNDGB1		T.AACTTAGA	A	---AGA.GCA..... 9215	
SYK.KE.-.SYK173		T.G-G-TAAA	CGCAA-TA	ACGTC-GCT.C-CA-CTGT-C---TG-GT.CACC-CCCGA.A-GTG..... 9597	





PLV  
complete genomes



Strain	Sequence	Position
H1B.FR.83.HXB2	GT..AA.CTAGAGATCC.....CTCAGAC..CCTTTTAGTCAGTGTGGAAAATCTCTAGCA	9719
H1A1.UG.85.U455	-----AC-A---A-T.GTGTA-----	9178
H1B.US.90.WEAU160	-----A-----	9719
H102.AG.NG.-.IBNG	-----	9201
H10.CM.91.MVP5180	TA..G-----	9793
CPZ.CM.-.CAM3	-----A.CTA--A--A..GTA-GT-----	9170
CPZ.GA.88.GAB2	-----CTAAACATAA.G-A-----	9146
CPZ.GA.-.CPZGAB	-----TTAAA--A---.TCAAG-A---	9811
CPZ.US.85.CPZUS	-----A.CT--AA-A..GTA-GT-----	9781
H2A.GW.-.ALI	A...TCGG-GTTC-C-TGA..GTAACAAGACCCT...GGTCTGTTAGGACCCTTTCTGCTT..TGGGAA-CGG--GCA-----C-----	10353
H2A.DE.-.BEN	A...TCGG-GTTC--TGA..GTAACAAGACCCT...GGTCTGTTAGGACCCTT--GCTT..TGGGAATCCA--GCA-----C-----	10359
H2B.CI.-.EHO	A...TCGG-G.....-TC..CACTAGAAACCCT...GGTCTGTTAGGACCCTTCTGCTT..TGGGAA-CCA--GCA-----C-----	10242
MAC.US.-.239	AAC.TCGG-...CTCAATA..ATAAGAAGACCCT...GGTCTGTTAGGACCCTTTCTGCTT..TGGGAA-CCG-AGCA-----C-----	10279
STM.US..STM	AAC.TCGG..CTCAA.....GACCCT...GGTCTGTTAGGACCCTT....T TG	9892
SAB.SN.-.SAB1C	.....TACT-AG	10009
TAN.UG.-.TAN1	-----CCTC-T-A-TG..G	9752
VER.KE.-.AGM155	-----GAAGTTCCT-A-TG..G	9758
VER.KE.-.9063	-----C-GTTCCT-A-TG..G	9779
VER.DE.-.AGM3	-G...-GACCCT--GGAA..T.....-AACC-T..GGAGAAGT	9294
GRV.ET.-.GRI_677	-GTG-----CCTCTCTTACTGGTT-.T-C-G--CC---GTG---G--A---C----	9623
DRL.-..FAO	TGGAG-T-CCTCAGA-TAGGTGAGAGCTATCCTGAGCTTTAGTGAGAGTAAAGAAAA.AT-TC-AGCAG	9732
RCM.GA.-.GAB1	-GTAGGGA-CCCTCA..GATTCTTGTGGCAGAAGAGCCTTGGGCTAAGAA.AATTC-CT--CAGT	9465
RCM.NG.-.NG411	-GTTGGGA-CCCTCA..GATATTGTGGCAGAAGTACTGAAGCCTAAGAAAAAATC-CA--CAG	9557
MND-2.-..5440	CGGAG-T-CCTCAGAT-TGTGGCAGAGATTGATATCTA..GTGAGAGTAGAGAAAA-AT-TC-AGCAGT-G-GCCCGAAC-GGGACT-GA--AGGGTTCAGTTCGCACCCACCTGTAAGGGAC	9703
LST.KE.-.lho7	.....TGGTTCAGTGATCTCTCACTAGAGAGATTGGAGCCTTGTGAT-CGG---AGGGC--CCAG	9957
SUN.GA.98.L14	--GTTCC-GT.....GTTCTCTCT..TTCTCCT.TCTAAACCCCTGTTCGACTG--C-C--CT--G-A-C--TTA--GGCCGCCTAGTGATCCGGGGCTGACGGCTAGCA	10006